nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-020-01364-1

Host diversity slows bacteriophage adaptation by selecting generalists over specialists

Duhita G. Sant, Laura C. Woods, Jeremy J. Barr and Michael J. McDonald ✉

Most viruses can infect multiple hosts, yet the selective mechanisms that maintain multi-host generalists over single-host specialists remain an open question. Here we propagate populations of the newly identified bacteriophage øJB01 in coculture with many host genotypes and find that while phage can adapt to infect any of the new hosts, increasing the number of hosts slows the rate of adaptation. We quantify trade-offs in the capacity for individual phage to infect different hosts and find that phage from evolved populations with more hosts are more likely to be generalists. Sequencing of evolved phage reveals strong selection and the genetic basis of adaptation, supporting a model that shows how the addition of more potential hosts to a community can select for low-fitness generalists over high-fitness specialists. Our results show how evolution with multiple hosts alters the rate of viral adaptation and provides empirical support for an evolutionary mechanism that promotes generalists over specialists.

Phages impact microbial communities by the horizontal spread of genes[1,2] and the dispersal of nutrients[3], and are potential tools for microbiome engineering[4]. The potential for a phage to impact a newly encountered microbial community largely depends on host range, with many phage able to infect multiple hosts[5–7]. Theory suggests that trade-offs between fitness on different hosts should drive a preference for single-host specialists[8–10] and empirical studies confirm that trade-offs exist for many traits and species[11–14], including phage[15–17]. The dynamics and molecular causes of phage evolution have been studied using evolution experiments[16,18–28] and confirm that phages can rapidly adapt to exploit new bacterial genotypes[25–29]. However, these experiments have resulted in the evolution of both generalists and specialists[25–27] and the reasons behind these different evolutionary outcomes are unclear[15,30,31]. Another important factor in the evolution of generalists is genotypic diversity in the community of hosts and the presence of potential new hosts[10]. Non-susceptible 'potential' hosts provide opportunities for phage diversification that could alter the outcomes of phage evolution. The importance of the broader ecological community for host evolution has been emphasized by a recent experiment showing that increasing the number of phage genotypes speeds host adaptation[32]. However, it remains unknown how the number of host genotypes alters the speed or mechanisms of phage adaptation[33].

To separate the effects of trade-offs and the presence of potential new hosts on the evolution of generalists, we tracked the host-range expansion of many replicate populations of a newly isolated phage in communities with one, two or three bacterial genotypes. We quantified trade-offs and the relative incidence of generalists and specialists by measuring the infectivity of evolved populations and phage clones, and then used genome sequencing to identify the genetic basis of generalism and specialism. Finally, we examine how the evolution of generalists and the rate of adaptation depends on the number of unique host genotypes.

We began by isolating, from an environmental waste water sample, a novel phage that could infect the enteropathogenic *Escherichia coli* strain O127:H6 EPEC E2348/69 (ref. [34]), which we named *Escherichia* phage JB01 (øJB01) (Methods; Extended Data Fig. 1). To study how the inclusion of additional non-permissive *E. coli* strains in the environment might impact øJB01 evolution, we cocultured øJB01 with its permissive host of isolation (hereafter referred to as EPEC) along with three other non-permissive *E. coli* strains (*E. coli* MG1655, *E. coli* O103:H2 REPEC E22 and *E. coli* BL21). We propagated replicate populations of øJB01 and hosts in one of three different treatments. For the first treatment, we cultured phage on a single host (4 × one-host treatment groups, $n = 16$ for each). For the second treatment, phage populations were cocultured with two *E. coli* strains, the host of isolation (EPEC) and one of the other non-permissive hosts (3 × two-host treatment groups, $n = 16$ for each). The third treatment had øJB01 added to cultures of three strains—the host of isolation (EPEC) and two of the other three non-permissive hosts (3 × three-host treatment groups, $n = 16$ for each). It is important to note that only phage populations were able to evolve in this experiment: *E. coli* cultures were kept evolutionarily naive by bringing in fresh *E. coli* cells after each coculture cycle and the experiment continued for ten cycles (Fig. 1; Methods).

Adaptation to the novel host requires the presence of ancestral host

Phage host-range expansion depends on ecological opportunity as well as sufficient genetic variation. Given that high titres of øJB01 could not infect our non-permissive *E. coli* strains (Supplementary Table 1), we expected that only coculture combinations of host and non-permissive *E. coli* strains would result in host-range expansion[28]. To test this, we measured the capacity of phage populations from each treatment to infect strains other than the host of isolation, *E. coli* EPEC, for each day of the evolution experiment (Fig. 1). As expected, we found that phage propagated on a non-permissive strain alone were unable to evolve the capacity to infect that strain as a new host. However, phage that evolved in the presence of two or three strains, including the original host, adapted to infect these strains (Fig. 1c,d). We found that øJB01 was unable to evolve to infect *E. coli* REPEC in the two-host treatments but did evolve to infect *E. coli* REPEC in the three-host treatment.

Increased host diversity slows the rate of phage adaptation

Previous studies have shown a correlation between the number of virus genotypes and host evolution[32]. While increased host diversity

School of Biological Sciences, Monash University, Monash, Victoria, Australia. ✉e-mail: mike.mcdonald@monash.edu

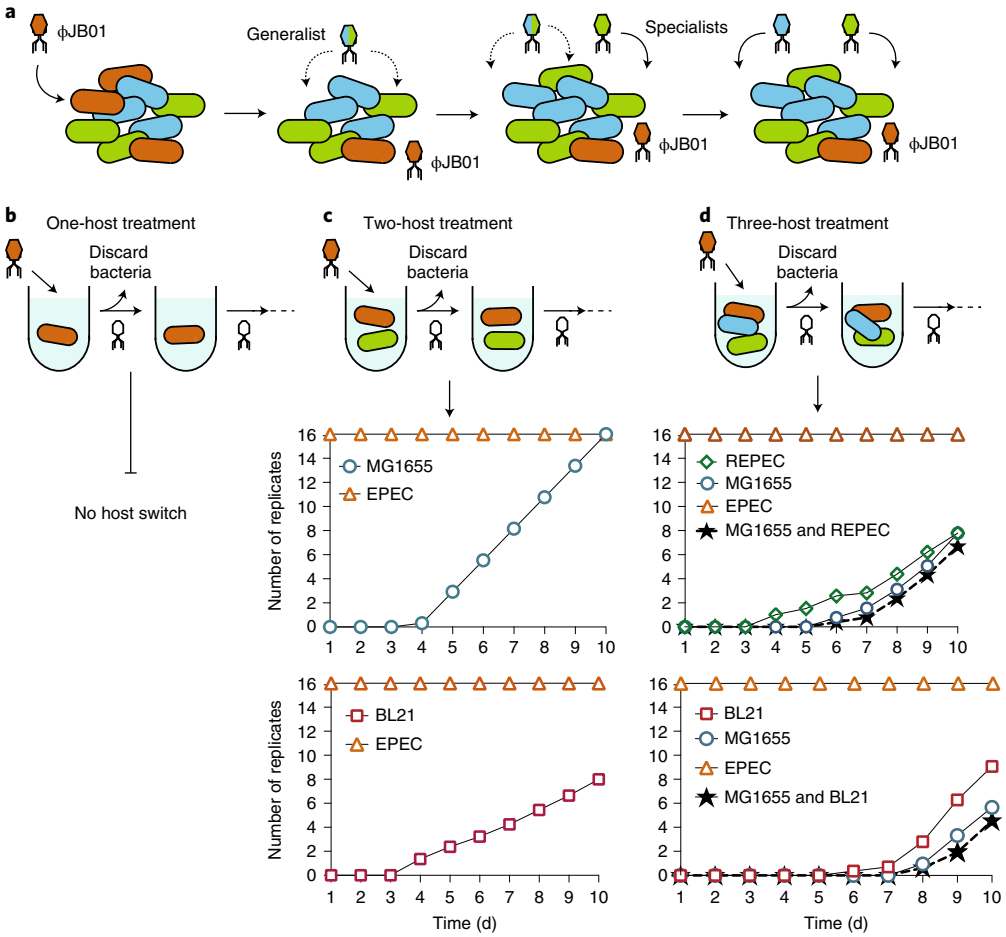


Fig. 1 | Adaptation of the novel phage øJB01 to a community of permissive and non-permissive *E. coli* hosts. a, Ancestral øJB01 were propagated on permissive host (orange) and other non-permissive *E. coli* strains (green and blue), and evolved to infect new hosts by way of generalist or specialist phenotypes. **b**, Replicate phage populations were passaged in cultures with a single host. **c**, For the two-host treatment, phage were cultured with the permissive *E. coli* EPEC host and one of the non-permissive *E. coli* strains. **d**, For the three-host treatment, phage populations were passaged with one permissive and two non-permissive strains. In all the host treatments, phage were collected from the evolved population while the bacteria were discarded every cycle to keep the bacterial populations evolutionarily 'naive'. Trajectories in the bottom four graphs show the number of replicate population that evolved to be able to infect the initially non-permissive *E. coli* strains MG1655 (blue circles), BL21 (red squares) and REPEC (green cones). In all treatments, infectivity was maintained on the original host, EPEC (orange triangles). Stars show the number of populations in the three-host treatment that evolved to infect both novel hosts.

might drive higher rates of virus adaptation, recent experiments in a bacterial community found that increasing community complexity slowed adaptation[35]. When we compared the time taken for phages to infect the novel host, we found that phage populations adapted more slowly to the three-host treatment than to the two-host treatment (Mann–Whitney *U*-test, BL21: $U=15$, $P=0.004$; MG1655: $U=0$, $P=9\times10^{-5}$). Specifically, the ancestral phage genotype øJB01 took an average of 5 d to adapt to infect MG1655 in the two-host treatment populations but required 8.5 d to adapt to MG1655 in the three-host treatment. Similarly, øJB01 required approximately 6 d to adapt to BL21 when it was present in the two-host treatment compared to 8 d in the three-host treatment (Fig. 2a). Population size can influence the rate of evolution[36] and individual *E. coli* strains will have smaller census population sizes in the three-host treatment compared to the two-host treatment. For example, in the two-host treatment, the ancestral host *E. coli* EPEC makes up 50% of the population (novel:ancestor, 1:1) but in any three-host treatment, *E. coli* EPEC makes up 33% (novel:ancestor, 2:1). The smaller census population size of *E. coli* EPEC in the three-host treatment will result in a smaller phage population and will influence the rate of adaptation.

We carried out control experiments to determine whether population size differences play an important role in the rate of adaptation to novel hosts in the three-host treatments. Phage populations were evolved with two hosts, with three different proportions of novel and ancestral host, 1:1, 2:1 and 3:1 (Fig. 2b). We observed a marginal (but non-significant) increase in the time to adapt to the novel host in the 2:1 and 3:1 treatments compared to the 1:1 treatments (Mann–Whitney *U*-test, $P>0.05$), however population size differences do not explain the magnitude of difference observed in the three-host treatment compared to the two-host treatment (Fig. 2a).

Given that phage in the three-host treatment were slower to evolve to infect a new host, we predicted that phage isolated from the three-host treatment should also have a lower fitness on the novel host than the phage isolated from the two-host treatment populations. We carried out this test of evolutionary rate by measuring the capacity of each evolved population to infect each host. These data showed that where øJB01 was able to evolve to infect a new host, phage from the two-host treatment populations had higher infectivity on that new host than phage from the three-host treatment populations (Fig. 2c–e) (Mann–Whitney *U*-test, $P<0.05$).

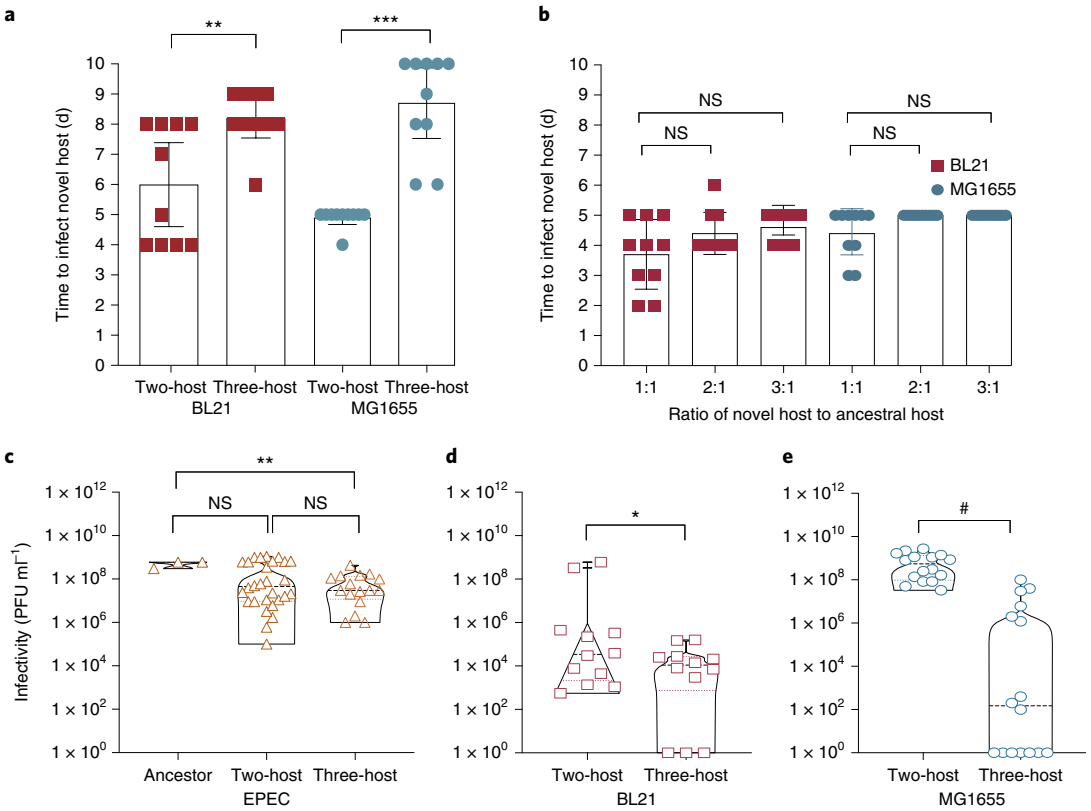


Fig. 2 | Increased host diversity slows phage adaptation. a, Phage populations adapt to a novel host more quickly in two-host than in three-host treatments. **b**, Two-host experiments with the ancestral host *E. coli* EPEC and either BL21 (red) or MG1655 (blue) as the novel host. **c**–**e**, Whole-culture lysates from single-host, two-host or three-host cultures were assayed for infectivity on *E. coli* EPEC (**c**), *E. coli* BL21 (**d**) and *E. coli* MG1655 (**e**). The two-host PFU counts from BL21 and MG1655 were standardized to account for population size differences (Methods). Each point represents an independent replicate from the evolution experiment and all infectivity measurements shown are the mean of three replicates. Mann–Whitney *U* one-tailed tests, $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$, $^{\#}P<0.0001$. NS, not significant.

This supports that the rate of adaptation to novel hosts was slowed in environments with more bacterial genotypes.

Phage clones show trade-offs for infectivity on different hosts

There is broad experimental support for trade-offs in a range of organisms[11,13,37,38], including bacteriophage[39]. Evolutionary theory predicts that the presence of novel potential hosts will apply selective pressures on a phage population that could drive two outcomes. First, the phage population could evolve a generalist phenotype with the capacity to target multiple hosts. Alternatively, the phage population could diversify into multiple subpopulations each specialized to a different host. To distinguish between these possibilities, we obtained phage lysates from the evolved two-host and three-host cultures. Each of these lysates was plated onto the ancestral host (EPEC) and onto each of the novel hosts that were present in the treatment conditions. For example, for the two-host treatment of øJB01 evolving in a coculture with *E. coli* EPEC and *E. coli* BL21, we plated the evolved population phage lysate onto two different bacterial lawns, one of EPEC and the other BL21 (Fig. 3a). Phage clones isolated from a bacterial lawn of EPEC were measured on independent cultures of EPEC and BL21 (Fig. 3b). Similarly, the infectivity of phage clones isolated from BL21 was measured on EPEC and BL21 (Fig. 3c). Using this method, we systematically measured the infectivity of individual phage clones to determine whether they were specialized to infect a single host or were capable of infecting multiple hosts.

We plotted the infectivity of individual phage clones against all the hosts that were present in their respective two-host treatment (Fig. 3g,h). These plots show that some phage clones from the two-host treatments were able to infect multiple hosts (Fig. 3g,h). However, very few clones showed high infectivity on more than one host, suggesting trade-offs between infectivity on different hosts. We quantified the overall trend by calculating the correlations in performance of phage clones on different hosts in the two-host treatment (Fig. 3g,h). This provides a measure of whether performance on the original host declines as performance on the novel host improves. We compared our measurements of phage infectivity to a hypothetical null model where an individual phage could evolve a high fitness on the novel host while maintaining the same high fitness that the ancestral phage has on the ancestral host of isolation. We used simulated data as our expectation under the null model of no trade-offs between phage fitness on two different hosts. Our comparison supports the existence of trade-offs, with phage isolated from both novel hosts (BL21 and MG1655) showing stronger negative correlations than the null expectation (Kolmogorov–Smirnov test, $P \leq 1 \times 10^{-4}$; Fig. 3i).

Host identity and host number impact the evolution of generalists

While our two-host treatment data suggest a general trend towards specialization, the individual phage isolates varied in their capacity to infect multiple hosts. We next looked at the effect of the number and identity of non-permissive hosts on the evolution of the generalist or specialist phenotypes across both the two- and three-host treatment populations (Fig. 4). We defined generalists as those phage clones that were able to infect more than one of the hosts

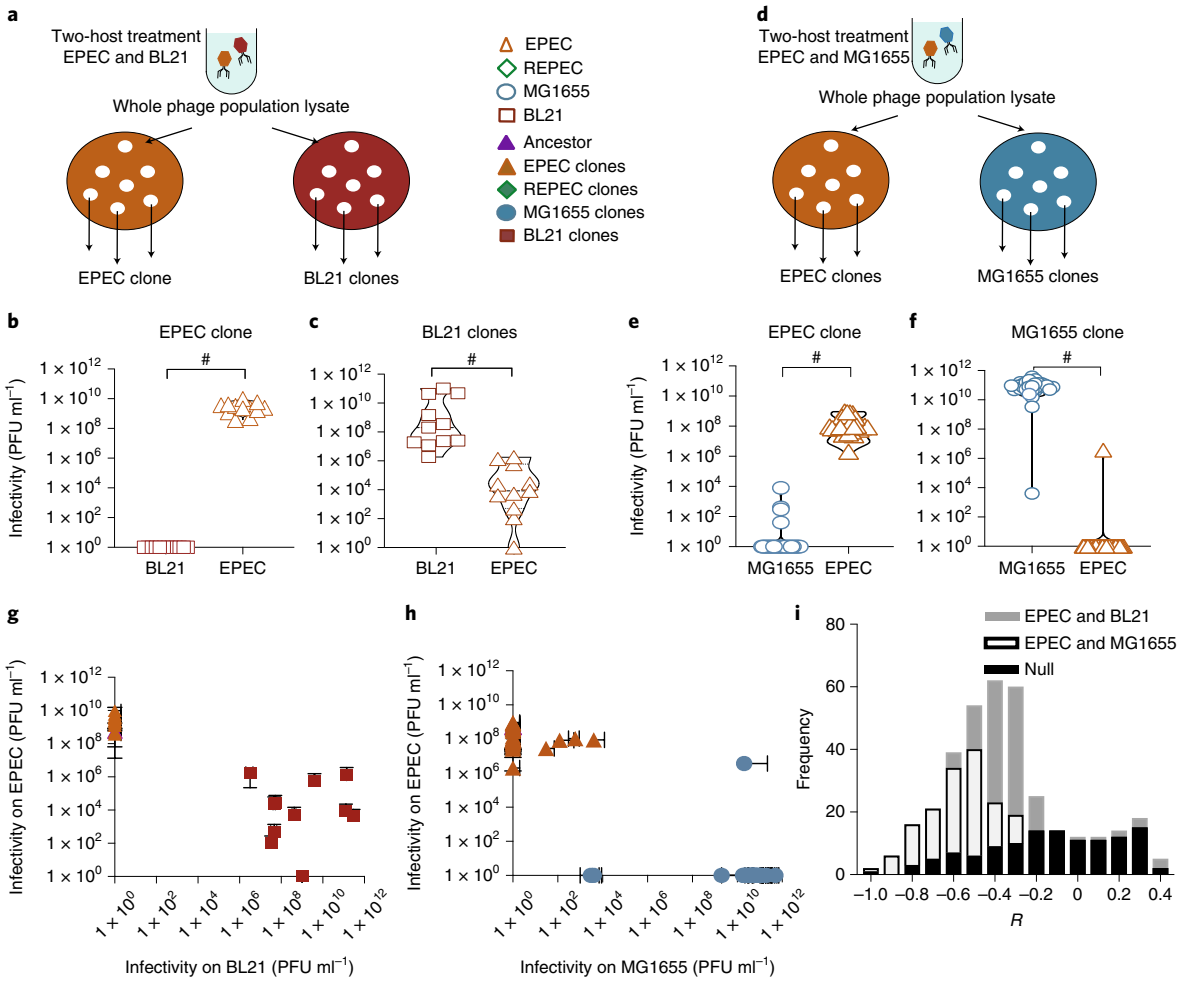


Fig. 3 | Evolution with multiple hosts is shaped by trade-offs for infectivity on each host. To isolate phage clones from two-host treatments, whole population phage lysates were plated onto individual hosts. **a**, Phage from the two-host treatment EPEC and BL21 were plated onto agar plates of *E. coli* EPEC (orange plate) and *E. coli* BL21 (red plate). **b**,**c**, Single plaques were isolated from each plate and infectivity was measured on the both EPEC (**b**) and BL21 (**c**). **d**, Phage from the two-host treatment EPEC and MG1655 were plated onto agar plates with lawns of *E. coli* EPEC (orange plate) and *E. coli* MG1655 (blue plate). **e**,**f**, Single plaques were isolated from each plate and infectivity was measured on both EPEC (**e**) and MG1655 (**f**). Phage clones isolated from EPEC could successfully infect EPEC (open orange triangles) but formed significantly less plaques on BL21 (open red squares) and MG1655 (open blue circles) (Kolmogorov–Smirnov D criterion = 1, $P \leq 0.0001$ for EPEC and BL21 clones, Kolmogorov–Smirnov D criterion = 1, $P \leq 0.0001$ for EPEC and MG1655 clones). **g**,**h**, The infectivity of phage clones on the two hosts from their respective treatments, BL21/EPEC (**g**) and MG1655/EPEC (**h**). **i**, Correlation coefficients for performance on two hosts were calculated for observed data (grey and white bars) and compared to a simulated distribution of correlation coefficients given a null model of 'no trade-offs' (black bars). The distribution of correlation coefficients between host treatments EPEC–MG1655 and EPEC–BL21 were significantly different to the null expectation of no trade-offs (null distribution versus EPEC–BL21: Kolmogorov–Smirnov D criterion = 0.5667, $P \leq 0.0001$; null distribution versus EPEC–MG1655: Kolmogorov–Smirnov D criterion = 0.6833, $P \leq 1 \times 10^{-4}$).

present in the evolution experiment treatment. First, we found that phage isolated from the novel host BL21 were more likely than phage isolated from any other host to be able to infect an alternative host (Fig. 4e) (two-proportion Z-test, $P < 1 \times 10^{-5}$). Next, we found that phages isolated from three-host treatments were better able to infect alternative hosts than phages from the two-host treatment (Fig. 4f,g) (two-proportion Z-test, $P = 2 \times 10^{-4}$ for MG1655 clones, $P = 0.03$ for EPEC clones). These results led us to hypothesize that the slower rate of adaptation in the three-host treatment compared to the two-host treatment may be because the presence of multiple hosts provides selection pressure for the maintenance of a generalist phenotype.

Phage adapt by parallel evolution in phage tail fibre protein gp17

The genetic causes of phage adaptation to novel hosts can be determined by whole genome sequencing. The phage used in this study is an uncharacterized isolate, so we started by sequencing its genome, finding that *Escherichia* øJB01 is a novel phage, 6% diverged from *Escherichia* phage N30 (NCBI: txid2340719) that falls within the broad class of T7 *Escherichia* phages (Supplementary Table 2). T7 phage attach to lipopolysaccharide molecules on the surface of their *E. coli* host, with differences in the rapidly evolving *gp17* gene able to distinguish between hosts[29]. To determine the genetic basis of phage adaptation to novel hosts in our study, we sequenced individual phage clones at the end of the evolution experiment and found that all had evolved non-synonymous substitutions in the gene encoding the tail fibre protein gp17. The protein gp17 has been shown to mediate the attachment of T7 phage to *E. coli* lipopolysaccharides[40] and mutations in the C terminus of gp17 contribute to host specificity[41] (Extended Data Fig. 3).

Given the high amount of genetic variation that evolved at the *gp17* locus, we sequenced additional clones only at the *gp17* locus,

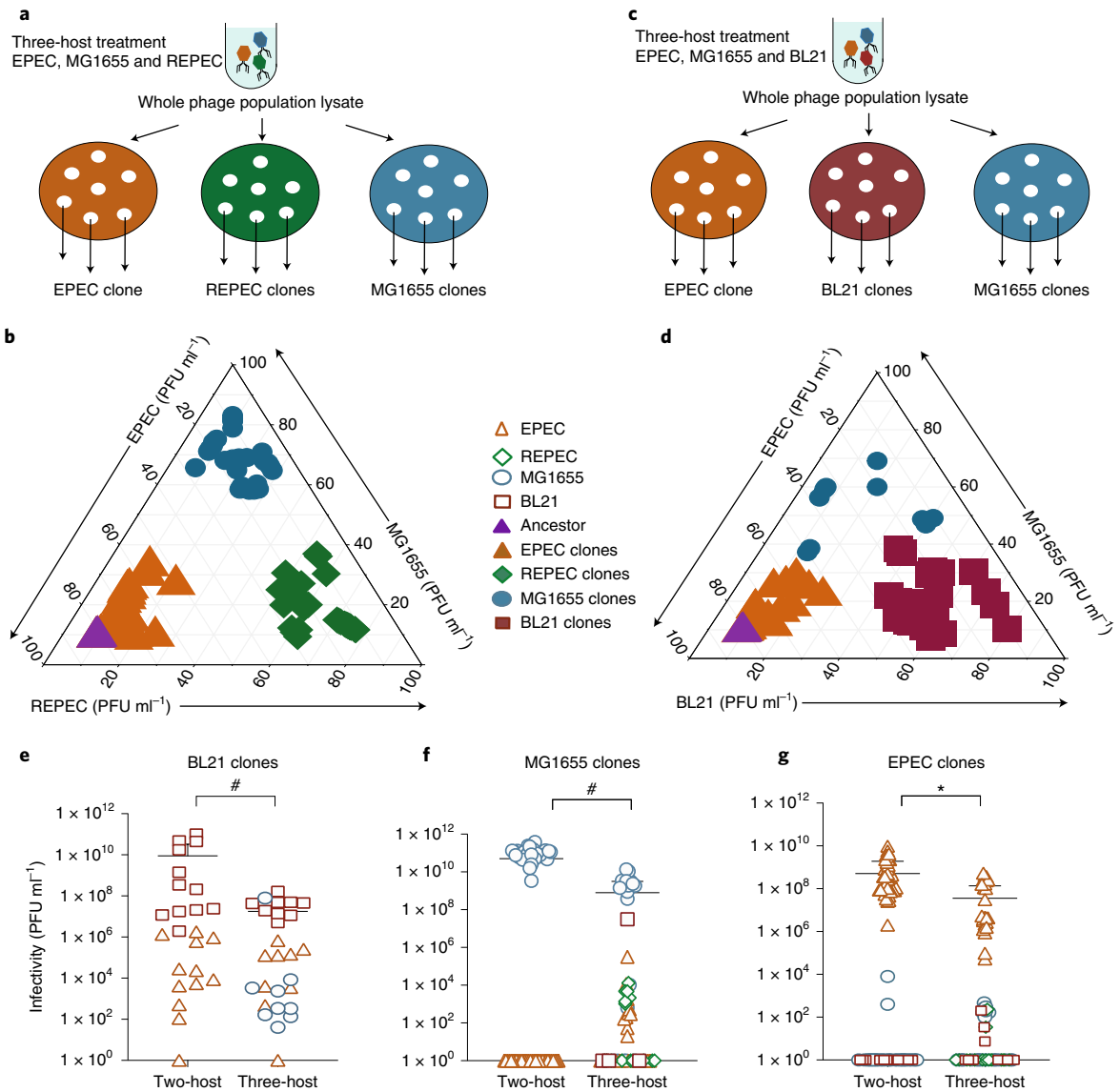


Fig. 4 | Host identity and host diversity impacts the evolution of generalists. To isolate phage clones from the three-host treatments, whole population phage lysates were plated onto individual hosts. **a**, Phage from the three-host treatment, EPEC, MG1655 and REPEC, were plated onto separate agar plates with a lawn of either *E. coli* EPEC (orange plate), *E. coli* REPEC (green plate) or *E. coli* MG1655 (blue plate). Single plaques were isolated from each plate and infectivity was measured on the three hosts involved in the treatment. **b**, Simplex plot for three-host treatment, EPEC, MG1655 and REPEC, shows infectivity of EPEC clones (orange triangles), MG1655 clones (blue circles) and REPEC clones (green cones) on all three hosts. **c**, Phage from the three-host treatment, EPEC, MG1655 and BL21, were plated onto separate agar plates with a lawn of *E. coli* EPEC (orange plate), *E. coli* BL21 (red plate) and *E. coli* MG1655 (blue plate), single plaques were isolated from each plate to measure the infectivity. **d**, Simplex plot for three-host treatment EPEC, MG1655 and BL21, shows the infectivity of clones isolated from EPEC (orange triangles), MG1655 (blue circles) and BL21 (red squares) on all three hosts. Plot axes are scaled to a range of 1–100 from a log scale; the underlying data are shown in Extended Data Fig. 2. All measurements of phage clones or phage populations shown are the mean of three replicates. **e**–**g**, Infectivity for each MG1655 (**e**), BL21 (**f**) and EPEC (**g**) clone on each host from their respective two-host and three-host treatments. Phage clones isolated from three-host treatments were better able to infect alternative hosts than phage clones from the two-host treatment (for MG1655 clones, two-proportion $Z=-3.5327$, $P=2\times10^{-4}$; for EPEC clones, two-proportion $Z=-1.8708$, $P=0.03$). BL21 phage clones were more likely to infect an alternate host irrespective of the host treatment (two-proportion $Z=6.0894$, $P<1\times10^{-5}$). $^{*}P<0.05$, $^{\#}P<1\times10^{-4}$.

for a total of 53 phage clones isolated from populations in the two-host and three-host treatments (Figs. 3 and 4) (Supplementary Table 3 and Supplementary Data). Around half of these clones had been isolated from one of the three-host treatments (EPEC, MG1655 and REPEC) (Supplementary Table 3). We used all of the non-synonymous substitutions discovered in *gp17* to look at the phylogenetic relationship amongst these 26 clones and found that they clustered by host of isolation, rather than by experimental replicate (Extended Data Fig. 4). This suggests the evolution of divergent subpopulations adapted to infect individual hosts (specialists) or multiple hosts (generalists).

The parallel evolution of mutations in the same genes in independently evolving populations is evidence for natural selection[42]. In this experiment, the evolution of mutations at the same amino acid sites in the *gp17* gene across independently evolving populations suggests that natural selection is driving the fixation of mutations at these sites. We found three non-synonymous substitutions at the *gp17* locus that were significantly enriched above expectations,

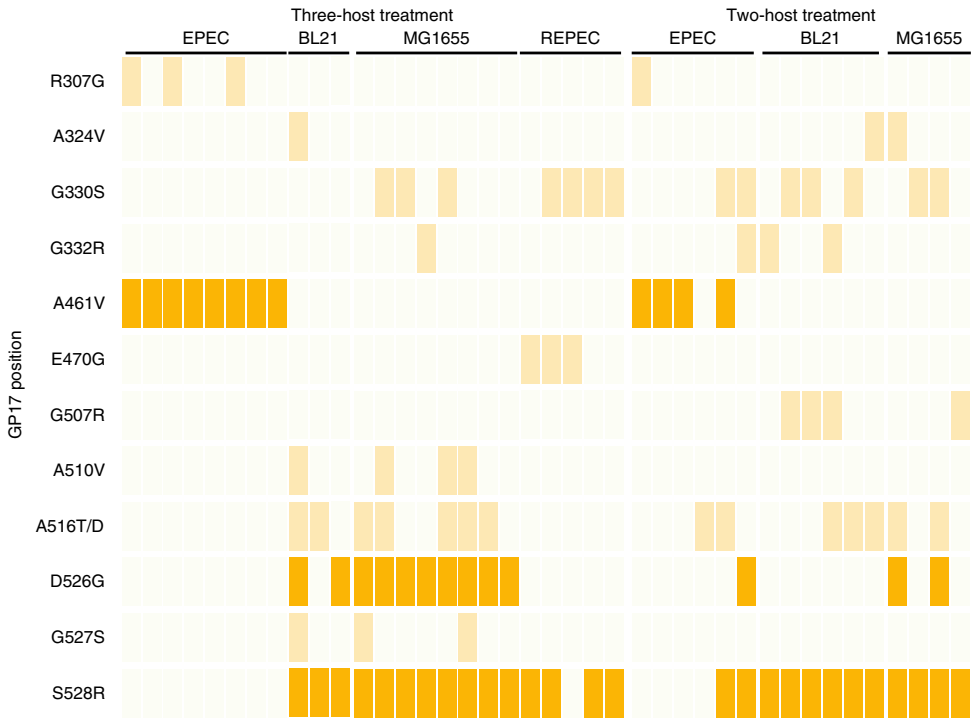


Fig. 5 | Parallel evolution of the tail fibre protein gp17. The *y* axis shows amino acid substitutions that were discovered in the *gp17* gene in at least three different phage clones out of the 53 that were sequenced. Phage clones were isolated from three-host and two-host treatments by plating on EPEC, BL21, MG1655 or REPEC *E. coli* strains, as shown in Figs. 3 and 4 (for full set of clones see Supplementary Table 3). Each cell shows the genetic state as wild type (light green), mutant (light orange) and/or a mutant determined to be significantly enriched by natural selection (bright orange) (Bonferroni-corrected hypergeometric test, $P<0.01$).

suggesting that these contribute towards øJB01 adaptation to its host (Bonferroni-corrected hypergeometric test, $P<0.01$; Fig. 5; Methods). First, we found that the S528R substitution was present in all except one of the phage that had adapted to any of the novel hosts but was not found in phage isolated from *E. coli* EPEC, the host of isolation (Fig. 5). We also found specific amino acid changes associated with isolation from two hosts, *E. coli* EPEC (A461V) and *E. coli* MG1655 (D526G) (Fig. 5). To test whether the presence or absence of these mutations could predict host specificity and range, we carried out spot tests of these sequenced clones against all the hosts used in this study (Supplementary Tables 4 and 5). While these sequencing data are consistent with the alleles at the *gp17* locus contributing to the genetic basis of generalist and specialist phenotypes, we found that it is highly likely that alleles at other sites in the genome influence host-range specificity.

Incorporating community context into phage fitness can explain the maintenance of generalists

Our data show that the rate of phage adaptation is slowed in communities with a greater number of potential novel hosts (Fig. 2). Moreover, even though we find trade-offs between infectivity on multiple hosts, there is a higher prevalence of phage clones that can infect multiple hosts in the three-host treatment (Fig. 4e–g). These data suggest that the presence of multiple hosts in a community may slow adaptation of a phage to any one of the hosts by promoting the maintenance of generalist phenotypes over specialist phenotypes. This hypothesis is supported by a population genetic model developed by Whitlock that incorporates alleles that have niche-dependent fitness effects[43]. In this model, the selection coefficient for a given allele is modified by the proportion of time spent by that allele in the environment where it is beneficial. Similarly, in a community with many host genotypes, the fitness of any given phage genotype (s) can be given by the probability of encountering a permissive host (ρ) and infectivity (i) so that fitness of a given phage genotype is: $s=\rho i$. For the purpose of this model, infectivity encapsulates such traits as the efficiency of binding, exploiting host replication machinery and burst size.

We wanted to understand the implications of host diversity on the fixation probability of generalist and specialist mutations that arise in our experiment. First, we consider that generalist mutants will, by definition, have a higher probability of encountering a permissive host than will a specialist. Here, we assume that generalists can infect all hosts, although the results of the model only depend on generalists being able to infect more hosts than can a specialist. We therefore assign probabilities to generalist and specialist alleles on the basis of the number of host genotypes (n):

$$\rho_{\mathrm{gen}} = 1, \rho_{\mathrm{sp}} = \frac{1}{n}$$

Next, we assume that generalist infectivity is always lower than specialist infectivity for any given host:

$$i_{\mathrm{gen}} < i_{\mathrm{sp}}$$

These simple assumptions about infectivity and the probability of encountering a permissive host suggest how environments with more hosts could favour generalists. In the single-host treatment, a specialist and a generalist are equally likely to encounter a permissive host. However, since the specialist can generate more offspring from every encounter with the host, it will rapidly outcompete the generalist and fix in the population. In the two-host treatment, a newly evolving specialist has a one-in-two probability of bumping into a susceptible host, while the generalist genotype can infect both of the hosts. In the three-host treatment, a newly evolved specialist has only a one-in-three probability of bumping into a susceptible

host, while the generalist genotype can still infect any host it encounters. The reduced probability of encountering a host means that when a specialist arises in a three-host treatment, it will have a lower fitness than in the two-host treatment. It is possible that the increased infectivity of the specialist could overcome the reduced probability of encountering a host and eventually outcompete the generalist. However, even when this is true, the specialist will experience an increased time to fixation and the population-wide rate of adaptation to the novel hosts will be slowed.

Implicit in our assumption $i_{gen} < i_{sp}$, is that there is a trade-off between the capacity to infect multiple hosts and infectivity on any one host. Previous studies support that trade-offs do exist between many traits across a wide range of species[11–14], including phage[15–17]. Our data also support this assumption (Fig. 3i), although the strength of the trade-off between generalists and specialists is highly variable (Fig. 4 and Extended Data Fig. 2). To parameterize our model, we used our measures of specialist (i_{sp}) and generalist (i_{gen}) infectivity from the three-host treatments (EPEC, MG1655, REPEC; Fig. 4a,b) and (EPEC, MG1655, BL21; Fig. 4c,d). We estimated the strength of trade-offs by using our plaque assays as a proxy for a phage clones' fitness. Specialist fitness (i_{sp}) was taken as the maximum mean infectivity (plaque-forming unit, PFU) of a phage clone measured on its own strain of isolation. For example, the i_{sp} estimate for EPEC is the maximum mean PFU count obtained from any of the clones isolated from *E. coli* EPEC, when measured on *E. coli* EPEC. Generalist fitness (i_{gen}) was taken as the maximum mean infectivity that we measured for a phage isolated from a different host. For example, the i_{gen} for EPEC, is the highest infectivity measured on *E. coli* EPEC for a clone that was a generalist, in other words, a clone that was isolated from one of the other hosts in the treatment. We calculated the fold advantage of the specialist over the generalist (i_{sp}/i_{gen}), finding ratios ranging from just over one to 1×10^7 (Supplementary Table 6). We used our model to determine ρ and i values for generalist and specialist mutations in communities containing up to 20 potential hosts. These calculations show that, where the specialist has <20-fold advantage over the generalist, communities with as few as two new potential hosts can favour the evolution of generalists over specialists (Fig. 6).

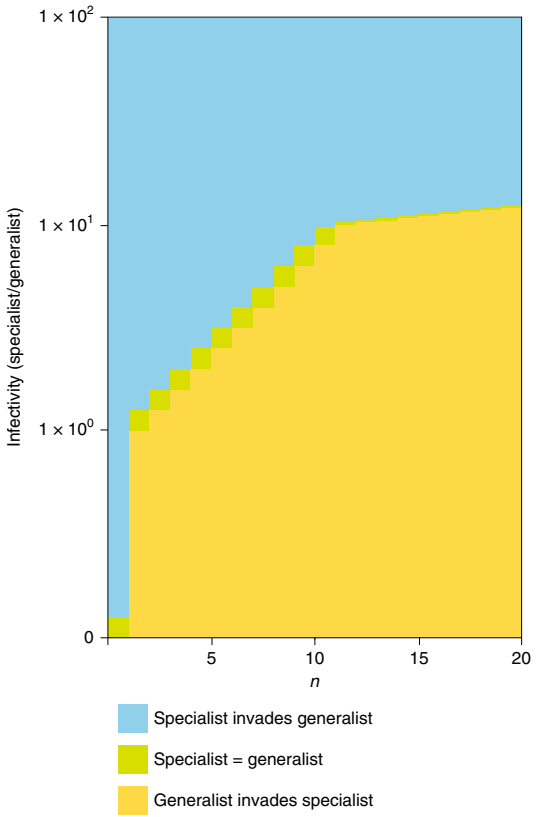


Fig. 6 | Community diversity and trade-offs predict the relative fitness of generalists and specialists. The outcome of competition between specialist and generalist genotypes of fitness i_{sp} and i_{gen}. The *y* axis shows advantage of the specialist over the generalist (i_{sp}/i_{gen}) for hypothetical competing pairs of specialists and generalists. The specialist advantage here ranges from 0 (generalist and specialist have equivalent fitness) to 1×10^2 (specialists are two orders of magnitude fitter than the generalist). The *x* axis indicates the number of potential hosts in the community. The relative probability (ρ) of a specialist encountering a permissive host declines with an increasing number of hosts ($\rho = 1/n$) but does not change for the generalist ($\rho = n$). Yellow-shaded areas show parameters where generalists are expected to outcompete specialists, blue-shaded areas show where specialists outcompete generalists and green-shaded areas show where generalists and specialists have equivalent fitness.

Conclusions

Our results show that a newly identified phage quickly evolved to infect non-permissive *E. coli* genotypes and that the rate of adaptation depended on the number of distinct hosts (Fig. 2). Similar to previous studies[15–17], we found evidence for trade-offs in the capacity for phage to infect multiple hosts (Fig. 3i). However, the capacity of individual phage clones to infect multiple hosts was variable and depended on the number and identity of host genotypes in the evolution treatment where the clones had evolved (Fig. 4e–g). Previous host-range expansion experiments with a two-host experimental system showed that the evolution of generalists or specialists depended on the ratios of the two hosts and the strength of trade-offs between the fitness on each host[15,27]. Our results support the importance of trade-offs for the evolution of generalists and specialists but emphasize a second factor, the number of potential hosts in the community.

An elementary assumption of epidemiological models is that increasing the proportion of non-susceptible (non-permissive) hosts reduces the chances of a specialist pathogen encountering a susceptible host[10]. Studies of multi-host pathogens incorporate non-susceptible hosts by multiplying fitness by the probability of encountering a susceptible host, similar to a population genetics approach that was developed to show how species with greater niche breadth have a slower rate of evolutionary response[43]. In this study, multiplying a measure of fitness (infectivity) with the probability of encountering a host (ρ) shows how generalists can have a high fitness relative to specialists, even while slowing the rate of adaptation to any one host. Altogether, these results suggest the ecological conditions, and an evolutionary mechanism, that can promote the evolution and maintenance of multi-host generalists, even when fitness trade-offs predict an advantage for single-host specialists. However, it should be noted that the advantage of generalists depends strongly on the strength of trade-offs, which may change over longer evolutionary time-scales when the *E. coli* hosts are able to co-evolve. Moreover, the evolutionary history of the newly isolated phage used in this study could affect the trends observed in this experiment.

Methods

Strains and culture conditions. The T7-like novel bacteriophage øJB01 was isolated from untreated waste water, kindly supplied by Christelle Schang and David McCarthy from Monash University, Department of Civil Engineering. The øJB01 phage was isolated from *E. coli* EPEC E2348/69 serotype O127:H6 (referred to as EPEC). Phage øJB01 was unable to infect *E. coli* K12-MG1655 (referred as MG1655), *E. coli* strain B/BL21-DE3 (referred as BL21) and *E. coli* REPEC E22 serotype O103:H2 (referred as REPEC); these were used as 'novel' potential host strains in the evolution treatments. Phage lysates were stored at 4 °C and filter sterilized before use. All the bacterial strains and phage lysates were grown in

LB medium (10 g l^{-1} Bacto Trytopne, 5 g l^{-1} yeast extract, 10 g l^{-1} NaCl) at 37 °C. Bacterial cultures were revived from glycerol stocks stored at −80 °C by streaking on LB agar plates and a single colony picked for overnight culture in 3 ml of LB. Before starting the evolution experiment, hosts were grown overnight to an optical density (OD_{600}) corresponding to a density of 1×10^8 cells ml^{-1}.

One-step phage kill curve and bacterial growth dynamics. Overnight cultures of bacterial host *E. coli* EPEC and lysates of øJB01 were mixed at a multiplicity of infection (MOI) of 0.1 (1×10^7 PFU ml^{-1} to 1×10^8 colony forming units, CFU ml^{-1}) in LB. The suspension was incubated at 37 °C with aeration for 5 min. It was then centrifuged at 4,000*g* for 10 min to remove unadsorbed phages. The supernatant was removed and the pellet was resuspended in 10 ml LB and incubated at 37 °C with aeration. The sample was withdrawn after every 5 min up to 20 min and then after every 10 min up to 90 min. At each time point, sample from the suspension was transferred into chloroform-saturated PBS, vortexed and then centrifuged at 3,500*g* for 3 min. The supernatant was diluted in PBS and plated for quantification of free phage particles[44]. The overall experiment was carried out in triplicate. Latent period was measured from the first time point when phage was added to bacterial population to the point where newly formed phage particles starts to appear. The burst size was calculated by dividing the average PFU ml^{-1} of the latent period by the average PFU ml^{-1} at plateau phase of the growth curve[44,45] (Extended Data Fig. 1a)

The growth of *E. coli* EPEC in the presence and absence of phage øJB01 was assayed by growth curves measured in a microplate reader. *E. coli* EPEC cultures were grown overnight in 3 ml Falcon tubes (Nunc) and diluted 100-fold into 96-well plates. Growth was tracked by measuring OD_{600} every 10 min. The measurements including phage were carried out at MOIs of 0.1 and 1 (Extended Data Fig. 1b).

Evolution experiment. The basic evolution transfer protocol, consistent to all treatments was as follows. Phage populations were evolved by serial transfer into fresh media containing mixed populations of bacteria at an MOI = 1×10^{-1}. After 24 h of growth, the phage populations were collected by the chloroform method, whereby bacteria are lysed, and only phage lysate recovered. These phage lysates were used to found the next-day cultures on 'naive' bacteria. We referred to these bacteria cocultures as naive since they were freshly cultured from stocks each day and had not previously been cocultured with phage. The different evolution treatments involved mixing cultures of multiple *E. coli* strains before inoculation with phage lysate. These mixtures were prepared by adding equal proportions of ancestral and novel host before each transfer with a total density of 1×10^8 cells ml^{-1}. The evolution experiment involved three treatments. First was the single-host treatment wherein phage was propagated on either the host of isolation (EPEC) or one of the non-permissive strains. Second, was the two-host treatment in which phage populations were evolved on two *E. coli* strains, one of which was always EPEC and the second was one of the non-permissive *E. coli* strains mixed in a 1:1 proportion (except for population size control experiments where the ratio was varied as described in Fig. 2). The three-host treatment had the host of isolation cocultured with two of the other non-permissive hosts, for a total of three *E. coli* strains. Again, each strain was mixed in 1:1:1 proportion (except for the control shown in Fig. 2). Overall, there were four single-host treatments (EPEC, MG1655, REPEC and BL21), each replicated 16 times, three two-host treatments (EPEC–MG1655, EPEC–REPEC and EPEC–BL21) each with 16 replicates and two three-host treatments (EPEC–MG1655–REPEC, EPEC–MG1655–BL21), each with 16 replicates. In total, there were 160 replicate populations with all evolution treatments carried out in 96-well plates and incubated overnight at 37 °C under shaking conditions in a microplate shaker. The overall process was repeated every day for 10 days.

Spot assays. After an evolution experiment, spot assays were performed to determine the changes in the capacity of the evolved phage populations to infect novel hosts for all three treatments each day for 10 days[46]. There was one spot per replicate per day. A total 5 µl of evolved phage lysate were plated onto a lawn of host cells to check the lytic activity after incubation. The spotted plates were left to dry and incubated at 37 °C for overnight. Presence of zone of clearance on a lawn of host cells was counted as positive. For all treatments, evolved phages from single-host treatment were spot inoculated on the corresponding host and the phage evolving on two- and three-host treatment was spot inoculated on each host present in the treatment. This was done for all 16 replicate populations of øJB01.

Plaque assays. Plaque assays were performed[47] to quantitatively determine the capacity of evolved phage populations to infect each of the hosts present in their treatment. Each individually evolving replicate population, from the two- and three-host treatments was measured in triplicate. All plaque assays to measure the infectivity of phage clones were also carried out in triplicate. The plates were incubated at 37 °C overnight and the number of plaques were counted after the incubation. Plaque assays from whole-culture lysates (Fig. 2c–e) were standardized so that PFU counts from two-host cultures could be compared to PFU counts from three-host cultures. To do this, we exploited whole-culture lysate PFU counts on *E. coli* EPEC. Since all treatments contained EPEC, we expected that differences between EPEC PFU counts in two- and three-host treatments would capture differences due to population size. The mean EPEC PFU counts from two-host cultures was divided by mean PFU counts from three-host cultures. Since the proportion of the total population that EPEC makes up in a two-host treatment is 0.5 and in the three-host treatment 0.33, we expected to find 1.5-fold less PFUs in the three-host treatment. Our calculations showed that three-host cultures had 3.62-fold lower EPEC PFUs than the two-host cultures. We followed this more conservative estimate, dividing all two-host PFU values by 3.62 and using these standardized values for statistical comparisons.

Isolation of phage clones. Evolved populations of phage from the two- and three-host treatments were initially plated onto the original host and each novel host present in the treatment (Fig. 3a,d and Fig. 4a,c). Individual plaques were isolated from each host and amplified in cultures with that same host of isolation. The infectivity of each single clone/plaque isolate was then measured on each host from the evolution treatment. For instance, in the case of evolution treatment with the two *E. coli* strains EPEC and BL21, the lysate was plated on a bacterial lawn of EPEC and another bacterial lawn of BL21. A plaque that was obtained from a bacterial lawn of EPEC, was then measured for infectivity on EPEC and BL21. In turn, phage clones isolated on BL21 were assayed for infectivity on EPEC and BL21. In this way, we could isolate individual plaques from all the replicates that could efficiently infect novel hosts in the case of two-host (EPEC–MG1655 and EPEC–BL21) and three-host (EPEC–MG1655–REPEC and EPEC–MG1655–BL21) treatments. Each clone was measured with three separate replicate assays.

Sequencing and genetic analysis. To identify the novel phage used in the study, and to compare the genome-wide mutations which had occurred in the evolved phage clones, we performed next-generation sequencing with the ancestral phage and ten evolved phage clones. Phage DNA samples were prepared for sequencing using DNeasy blood and tissue kit by Qiagen as in ref. [48], with some modifications. The concentration and quality of the DNA was assessed with the Qubit double-stranded DNA quantification kit according to the manufacturer's instructions. DNA was sequenced on the Illumina MiSeq next-generation sequencing platform and reads were preprocessed by GeneWiz. Unicycler[49] was used in short-read mode to resolve the ancestral phage genome sequence. The resulting sequence was used as the reference for resequencing and variant calling of the evolved phage genomes using the *breseq* and *gdtools* analysis pipeline[50] to identify de novo mutations and quantify their relative proportions within the population.

Sanger sequence confirmation of *gp17* mutations. It was observed from the whole genome sequencing data that many evolved phage clones contained multiple mutations in tail fibre protein gp17. To confirm the genotypes of these phage clones we PCR-amplified the tail fibre gene *gp17* from a total of 43 replicate phage clones (for primers, see Supplementary Table 7) and carried out Sanger sequencing (Micromon). We sequenced individual phage clones isolated from whole population of phages adapted on various hosts treatments at the end of an evolution experiment. A total 53 clones were isolated from four different experimental replicates from each host treatment (Supplementary Table 3). Individual clones were selected to include a representative for every host treatment and the corresponding infectivity profile shown by that particular host treatment.

Hypergeometric test for multi-hit genes. We determined whether a given amino acid substitution within gp17 was more likely than chance to be found in phage clones isolated from a particular host. First, we created a list of 12 amino acid sites that were hit at least three times (Fig. 5). We generated hypotheses that any of these 12 sites had evolved a non-synonymous substitution that was enriched specifically in one of ten subgroups (EPEC, BL21, MG1655, REPEC, EPEC–BL21, EPEC–MG1655, MG1655–BL21, EPEC–MG1655–REPEC, EPEC–MG1655–BL21 and EPEC–MG1655–REPEC–BL21). We then compared the frequency of a given mutation across all sequenced phage clones to the frequency of that mutation within each of these subset of phage clones. We determined significant enrichment within the subset of phage clones using a hypergeometric distribution and corrected *P* values (Bonferroni correction) on the basis of the number of tests we carried out (120).

Quantification of trade-offs. To test the deviation from a null hypothesis, we generated the expected distribution of *R* values given no trade-off between the new host and the ancestral host. This was done by correlating the position of a set of randomized ancestral phage (high fitness on EPEC, low fitness on new host) with a set of hypothetical evolved clones with high fitness on EPEC and high fitness on the new host. The bounds of these hypothetical evolved clones were set by measurements of the ancestral phage on the host of isolation (*E. coli* EPEC) and measured values from the new host. To find the distribution of correlation coefficients for our measured phage, we generated random pairing of phage isolated from EPEC and novel host, for the data points from Fig. 3g,h. The observed distribution of *R* values was compared to the expected values using a Kolmogorov–Smirnov test.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Raw sequencing reads used to generate the data in Fig. 4a and Extended Data Fig. 4 have been deposited in GenBank under the Bioproject ID PRJNA673261.

Received: 4 June 2020; Accepted: 12 November 2020;
Published online: 11 January 2021

References

1. Koskella, B. & Brockhurst, M. A. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).
2. Fernandez, L., Rodriguez, A. & Garcia, P. Phage or foe: an insight into the impact of viral predation on microbial communities. *ISME J.* **12**, 1171–1179 (2018).
3. Zimmerman, A. E. et al. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* **18**, 21–34 (2020).
4. Gordillo Altamirano, F. L. & Barr, J. J. Phage therapy in the postantibiotic era. *Clin. Microbiol. Rev*. **32**, e00066-18 (2019).
5. Flores, C. O., Valverde, S. & Weitz, J. S. Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *ISME J.* **7**, 520–532 (2013).
6. Koskella, B. & Meaden, S. Understanding bacteriophage specificity in natural microbial communities. *Viruses* **5**, 806–823 (2013).
7. Olival, K. J. et al. Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
8. Law, R. Optimal life histories under age-specific predation. *Am. Nat.* **114**, 399–417 (1979).
9. Williams, G. C. Natural selection, the costs of reproduction, and a refinement of Lack's principle. *Am. Nat.* **100**, 687–690 (1966).
10. Woolhouse, M. E., Taylor, L. H. & Haydon, D. T. Population biology of multihost pathogens. *Science* **292**, 1109–1112 (2001).
11. MacLean, R. C., Bell, G. & Rainey, P. B. The evolution of a pleiotropic fitness tradeoff in *Pseudomonas fluorescens*. *Proc. Natl Acad. Sci. USA* **101**, 8072–8077 (2004).
12. Zhong, S., Khodursky, A., Dykhuizen, D. E. & Dean, A. M. Evolutionary genomics of ecological specialization. *Proc. Natl Acad. Sci. USA* **101**, 11719–11724 (2004).
13. Bennett, A. F. & Lenski, R. E. An experimental test of evolutionary trade-offs during temperature adaptation. *Proc. Natl Acad. Sci. USA* **104**, 8649–8654 (2007).
14. Lee, M. C., Chou, H. H. & Marx, C. J. Asymmetric, bimodal trade-offs during adaptation of *Methylobacterium* to distinct growth substrates. *Evolution* **63**, 2816–2830 (2009).
15. Bono, L. M., Smith, L. B. Jr., Pfennig, D. W. & Burch, C. L. The emergence of performance trade-offs during local adaptation: insights from experimental evolution. *Mol. Ecol.* **26**, 1720–1733 (2017).
16. Duffy, S., Turner, P. E. & Burch, C. L. Pleiotropic costs of niche expansion in the RNA bacteriophage phi 6. *Genetics* **172**, 751–757 (2006).
17. Duffy, S., Burch, C. L. & Turner, P. E. Evolution of host specificity drives reproductive isolation among RNA viruses. *Evolution* **61**, 2614–2622 (2007).
18. Buckling, A. & Rainey, P. B. Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. Biol. Sci.* **269**, 931–936 (2002).
19. Burch, C. L. & Chao, L. Evolvability of an RNA virus is determined by its mutational neighbourhood. *Nature* **406**, 625–628 (2000).
20. Crill, W. D., Wichman, H. A. & Bull, J. J. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**, 27–37 (2000).
21. Brockhurst, M. A., Buckling, A. & Rainey, P. B. The effect of a bacteriophage on diversification of the opportunistic bacterial pathogen, *Pseudomonas aeruginosa*. *Proc. Biol. Sci.* **272**, 1385–1391 (2005).
22. Gomez, P. & Buckling, A. Bacteria–phage antagonistic coevolution in soil. *Science* **332**, 106–109 (2011).
23. Paterson, S. et al. Antagonistic coevolution accelerates molecular evolution. *Nature* **464**, 275–278 (2010).
24. Pal, C., Macia, M. D., Oliver, A., Schachar, I. & Buckling, A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature* **450**, 1079–1081 (2007).
25. Bono, L. M., Gensel, C. L., Pfennig, D. W. & Burch, C. L. Competition and the origins of novelty: experimental evolution of niche-width expansion in a virus. *Biol. Lett.* **9**, 20120616 (2013).
26. Bono, L. M., Gensel, C. L., Pfennig, D. W. & Burch, C. L. Evolutionary rescue and the coexistence of generalist and specialist competitors: an experimental test. *Proc. Biol. Sci.* **282**, 20151932 (2015).
27. Petrie, K. L. et al. Destabilizing mutations encode nongenetic variation that drives evolutionary innovation. *Science* **359**, 1542–1545 (2018).
28. Dennehy, J. J., Friedenberg, N. A., Holt, R. D. & Turner, P. E. Viral ecology and the maintenance of novel host use. *Am. Nat.* **167**, 429–439 (2006).
29. Heineman, R. H., Springman, R. & Bull, J. J. Optimal foraging by bacteriophages through host avoidance. *Am. Nat.* **171**, E149–E157 (2008).
30. Jerison, E. R., Nguyen, Ba,A. N., Desai, M. M. & Kryazhimskiy, S. Chance and necessity in the pleiotropic consequences of adaptation for budding yeast. *Nat. Ecol. Evol.* **4**, 601–611 (2020).
31. Remold, S. Understanding specialism when the jack of all trades can be the master of all. *Proc. Biol. Sci.* **279**, 4861–4869 (2012).
32. Betts, A., Gray, C., Zelek, M., MacLean, R. C. & King, K. C. High parasite diversity accelerates host adaptation and diversification. *Science* **360**, 907–911 (2018).
33. Betts, A., Rafaluk, C. & King, K. C. Host and parasite evolution in a tangled bank. *Trends Parasitol.* **32**, 863–873 (2016).
34. Iguchi, A. et al. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. *J. Bacteriol.* **191**, 347–354 (2009).
35. Scheuerl, T. et al. Bacterial adaptation is constrained in complex communities. *Nat. Commun.* **11**, 754 (2020).
36. Lanfear, R., Kokko, H. & Eyre-Walker, A. Population size and the rate of evolution. *Trends Ecol. Evol.* **29**, 33–41 (2014).
37. Li, Y., Petrov, D. A. & Sherlock, G. Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation. *Nat. Ecol. Evol.* **3**, 1539–155 (2019).
38. Bachmann, H. et al. Availability of public goods shapes the evolution of competing metabolic strategies. *Proc. Natl Acad. Sci. USA* **110**, 14302–14307 (2013).
39. Keen, E. C. Tradeoffs in bacteriophage life histories. *Bacteriophage* **4**, e28365 (2014).
40. Gonzalez-Garcia, V. A. et al. Conformational changes leading to T7 DNA delivery upon interaction with the bacterial receptor. *J. Biol. Chem.* **290**, 10038–10044 (2015).
41. Holtzman, T. et al. A continuous evolution system for contracting the host range of bacteriophage T7. *Sci. Rep.* **10**, 307 (2020).
42. McDonald, M. J. Microbial experimental evolution—a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep.* **20**, e46992 (2019).
43. Whitlock, M. C. The red queen beats the jack-of-all-trades: the limitations on the evolution of phenotypic plasticity and niche breadth. *Am. Nat.* **148**, S65–S77 (1996).
44. Zhao, J. et al. Characterizing the biology of lytic bacteriophage vB_EaeM_ phiEap-3 infecting multidrug-resistant *Enterobacter aerogenes*. *Front. Microbiol.* **10**, 420 (2019).
45. Buttimer, C. et al. *Pectobacterium atrosepticum* phage vB_PatP_ CB5: a member of the proposed genus '*Phimunavirus*'. *Viruses* **10**, 394 (2018).
46. Van Twest, R. & Kropinski, A. M. Bacteriophage enrichment from water and soil. *Methods Mol. Biol.* **501**, 15–21 (2009).
47. Bonilla, N. et al. Phage on tap—a quick and efficient protocol for the preparation of bacteriophage laboratory stocks. *PeerJ* **4**, e2261 (2016).
48. Jakociune, D. & Moodley, A. A rapid bacteriophage DNA extraction method. *Methods Protoc.* **1**, 27 (2018).
49. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).
50. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).
51. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
52. Stecher, G., Tamura, K. & Kumar, S. Molecular evolutionary genetics analysis (MEGA) for macOS. *Mol. Biol. Evol.* **37**, 1237–1239 (2020).
53. Dereeper, A. et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465–W469 (2008).
54. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
55. Whelan, S. & Goldman, N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* **18**, 691–699 (2001).

Acknowledgements

M.J.M. was supported by ARC Discovery Grant (grant no. DP180102161) and ARC Future Fellowship (grant no. FT170100441).

Author contributions

M.J.M. undertook the conceptualization, methodology, resources and funding acquisition. M.J.M. and J.J.B. supervised the work. M.J.M. and D.G.S. were responsible for project administration. M.J.M. and D.G.S. conducted the formal analysis. M.J.M., D.G.S. and J.J.B. undertook validation of results. D.G.S. was involved in investigation and visualization. L.C.W. provided software and data curation. M.J.M. and D.G.S. prepared the original draft manuscript. J.J.B. then reviewed and edited the manuscript.

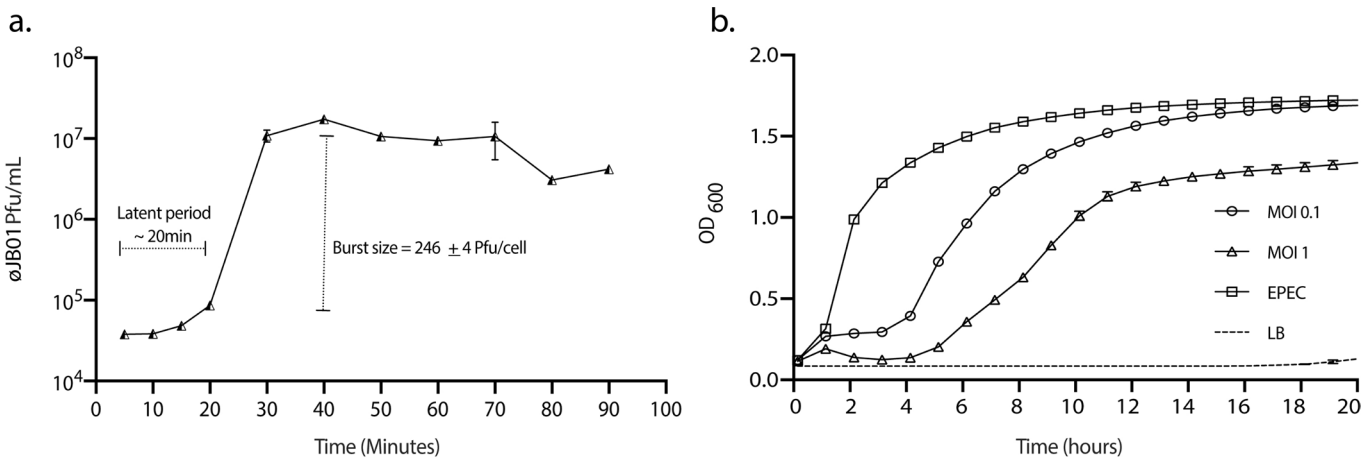
Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41559-020-01364-1.

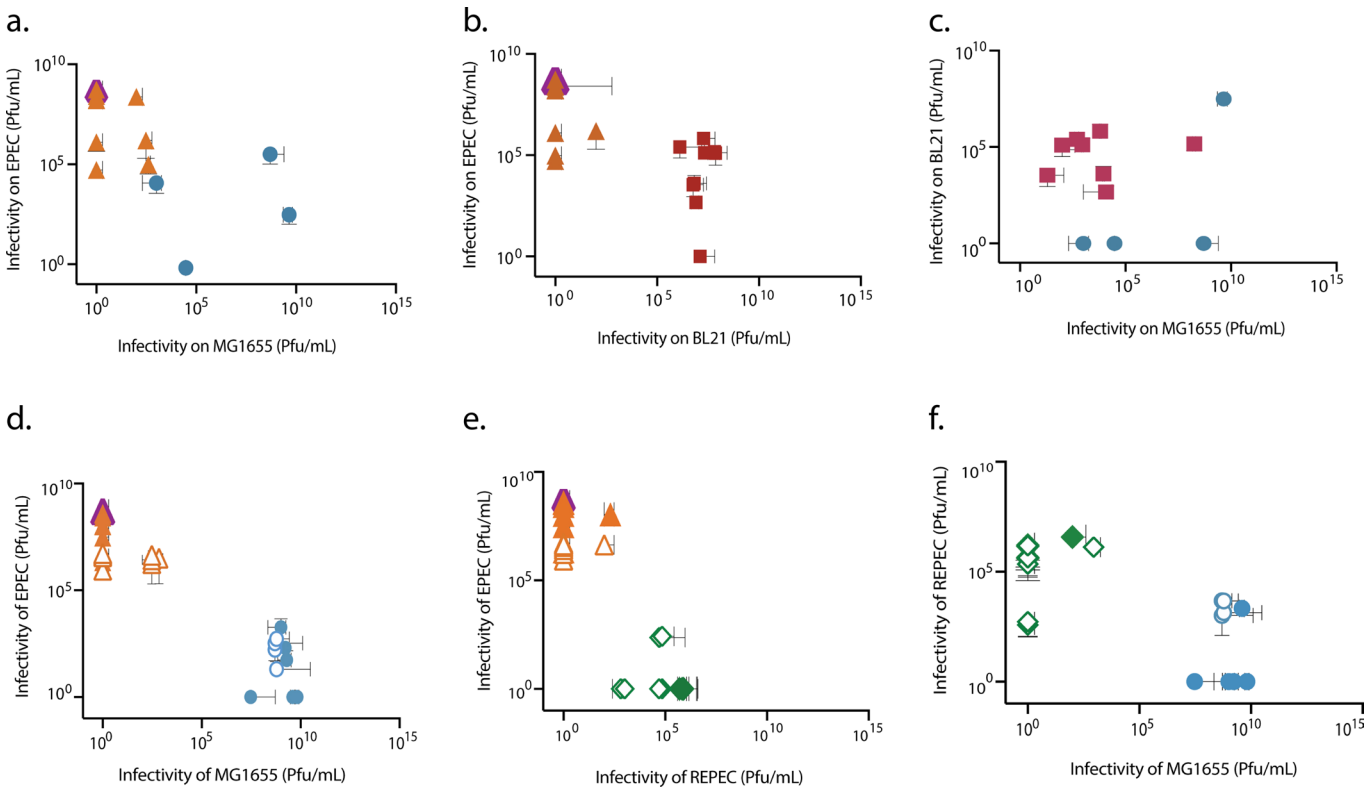
Supplementary information is available for this paper at https://doi.org/10.1038/s41559-020-01364-1.

Correspondence and requests for materials should be addressed to M.J.M.

Peer review information *Nature Ecology & Evolution* thanks Britt Koskella and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



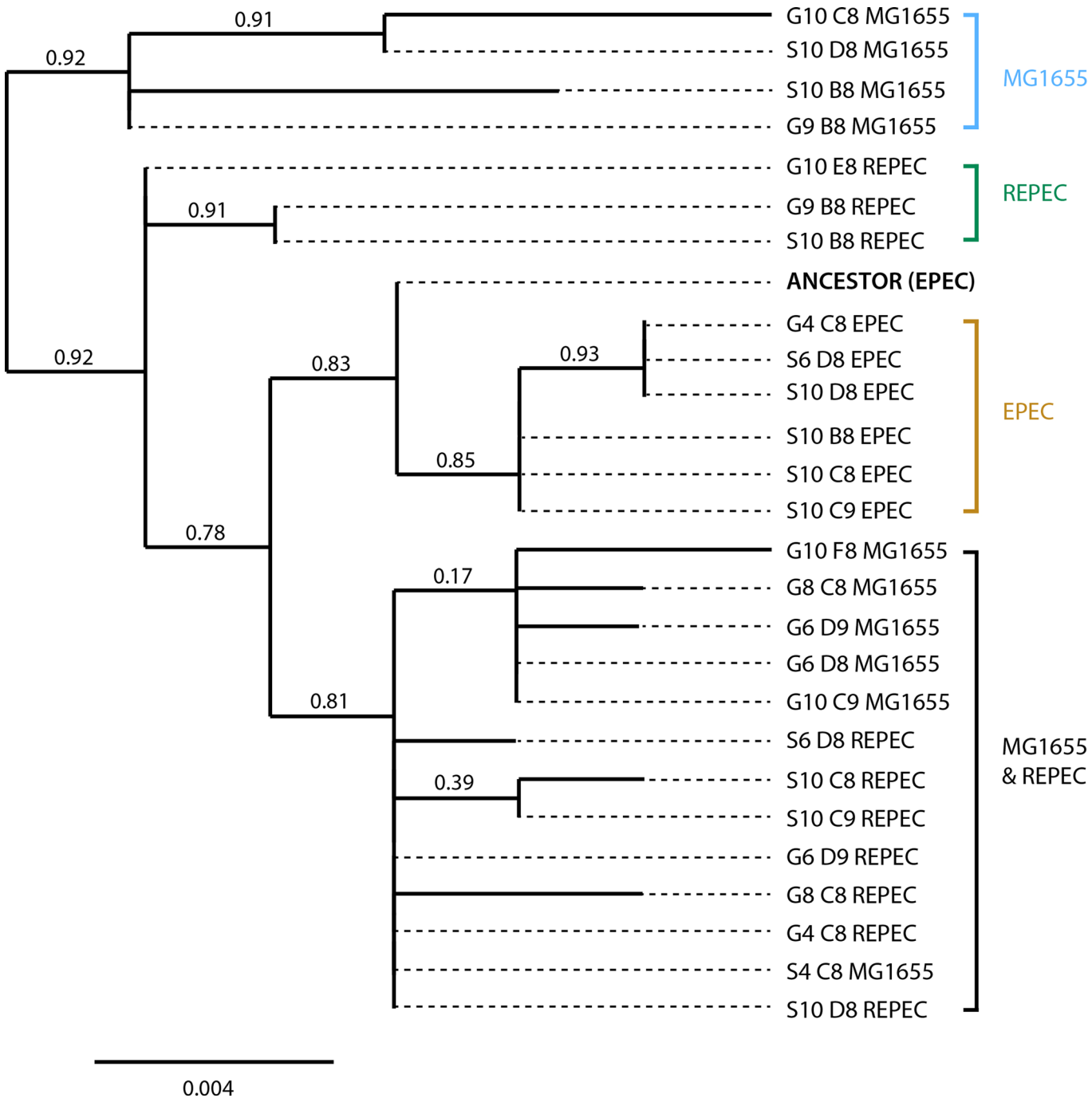
Extended Data Fig. 1 | Growth of E. coli EPEC and phage øJB01 in coculture. The number of øJB01 plaque-forming units were monitored over a period of 90 min in coculture with exponential phase *E. coli* EPEC at MOI 0.1(A). The growth of *E. coli* EPEC was assayed using a plate reader. *E. coli* EPEC was propagated either in monoculture, or in coculture with phage øJB01 added at two different MOIs (1 and 0.1) (B). Error bars represent standard error of the mean (SEM) (n = 3).



Extended Data Fig. 2 | Two-way plots showing data used for simplex plots in Fig. 4. Panels **a-c** correspond to three-host treatment (EPEC, MG1655, BL21). Panels **d-f** correspond to the three-host treatment (EPEC, MG1655, REPEC). Measurements of infectivity (PFU) for phage clones isolated from MG1655 (blue circles), EPEC (orange triangles), BL21(red squares) and REPEC (green diamonds). Filled markers show the measurements for phage clones obtained at the end point of the experiment (Day 10). Open markers show phage clones sampled at early timepoints (day 4 - day 9). Error bars are SEM, n = 3.

```
gp17_T7       MANVIKTVLTYQLDGSNRDFNIPFEYLARKFVVVTLIGVDRKVLTINTDYRFATRTTISL  60
gp17_N30      MANVIKTVLTYQLDGSNRDFNIPFEYLARKFVVVTLIGVDRKVLTINTDYRFATRTTISL  60
gp17_ØJB01    MANVIKTVLTYQLDGSNRDFNIPFEYLARKFVVVTLIGVDRKVLTINTDYRFATRTTISL  60
              ************************************************************

gp17_T7       TKAWGPADGYTTIELRRVTSTTDRLVDFTDGSILRAYDLNVAQIQTMHVAEEARDLTADT  120
gp17_N30      TKAWGPADGYTTIELRRVTSTTDRLVDFTDGSILRAYDLNVAQIQTMHVAEEARDLTADT  120
gp17_ØJB01    TKAWGPADGYTTIELRRVTSTTDRLVDFTDGSILRAYDLNVAQIQTMHVAEEARDLTADT  120
              ************************************************************

gp17_T7       IGVNNDGHLDARGRRIVNLANAVDDRDAVPFGQLKTMNQNSWQARNEALQFRNEAETFRN  180
gp17_N30      IGVNNDGHLDARGRRIVNLANAVDDRDAVPLGQLKVMNQNSWQARNEALQFRNEAETFRN  180
gp17_ØJB01    IGVNNDGHLDARGRRIVNLANAVDDRDAVPFGQLKTMNQNSWQARNEALQFRNEAETFRN  180
              ********************************:****.************************

gp17_T7       QAEGFKNESSTNATNTKQWRDETKGFRDEAKRFKNTAGQYATSAGNSASAAHQSEVNAEN  240
gp17_N30      QAEGFKNESGTNAANTKQWRDETNGFRNEAEQFKNTAGQYATSAGNSATAAHQSEVNAGN  240
gp17_ØJB01    QAEGFKNESGTNAANTKQWRDETQGFRDGAEQLKNTAGQYATSAGNSATAAHQSEVNAEN  240
              *********.***.*********.***::*..:***************.*********.*

gp17_T7       SATASANSAHLAEQQADRAEREADKLENYNGLAGAIDKVDGTNVYWK-GNIHANGRLY-M  298
gp17_N30      SATASADSAASSKQQADRAKQEADKLGNWNALAGTVDSVQGDSVFWKGGGYWVGGPVYQL  300
gp17_ØJB01    SATASANSAHLAEQQADRAEREADKLGNWNALAGTVESVQGDSVFWKGGGYWVGGPAYRL  300
              ******:**....******.:*****.*.*.***.::.*.*..*:** *.. ..*..* :

gp17_T7       TTNG---FDCGQYQQFFGGVTNRYSVMEWGDENGWLMYVQ----------RREWTTAIGG  345
gp17_N30      QTNHAFRFRMGLWEE--GDFLKSDFAMDRPDGIGFIEYAADKTGQHRHQIRLNQLTIAGG  358
gp17_ØJB01    PTNHAFRFRMGIWEE--GDFLKSDFAMDRPDGIGFIGYAADKTGQHRHQIRLNQLTIAGG  358
               **.   *..* ...  *........*:..*..*.. *..         * . .*..**

gp17_T7       NIQLVVNGQIITQGGAMTGQLKLQNGHVLQLESASDKAHYILSKDGNRNN------WYIG  399
gp17_N30      VLESDADDRI---AHAFADQYNTKAPFFQQIHVQSTSSYHPIIKQRTYQSGRWAGCWSIG  415
gp17_ØJB01    VVESDADGLI---AHAFADQYTTKAPFFQQIHVQSTSSYHPIIKQRTYQSGRWAGCWSIG  415
              .... ... *   ..*...*. .......*....*...:...*......     *.**
gp17_T7       RGSDNNNDCTFHSYVHGT-------TLTLKQDYAVVNKHFHVGQAVVATDGNI--QGTKW  450
gp17_N30      TLLTSTSP-SFHVHMKDEGAGDKLWTFNTNGDFRVPGEIWADGGRVVVRNGSFDPEGRVW  474
gp17_ØJB01    TLLTGAHP-SFHIHMKDEGISDKLWTFNTNGDFRAPGEIWADGGRVAVRNGSFDPEGRVW  474
              . .. . . .** :....       *.....*:.:... .. *..*:...*..  .*. *

gp17_T7       GGK----WLDAYLRDSFVAK----SKAWTQVWSGS----AGGGVSVTVSQDLRFRNIWIK  498
gp17_N30      GNVYDTGYLDDWVRRHFIGDVYLGPQSWQQQTGGYQDIIAPAGHVQTGTTA-SNGAV---  530
gp17_ØJB01    GTAYDTGYLDDWIRRHWPGDVYWGPQRWVQKWGGYADIVAGAGEACSGVTADGSGNV---  531
              *      .**...*......    .. * * ..*.    *..*   :. ..   ..:

gp17_T7       CANNSWNFFRTGPDGIYFIASDGGWLRFQIHSNGLGFKNIADSRSVPNAIMVENE  553
gp17_N30      -------------AGIFVRPV---WLGFHN-----GTRRIIASQ-----------  553
gp17_ØJB01    -------------AGAFFRAL---WLGYHN-----GTKRLVASQ-----------  554
                        .*.:. .    **.::.     *.:.:..*.
```

Extended Data Fig. 3 | CLUSTAL multiple sequence alignment of tail fibre protein gp17. The amino acid sequence of *Escherichia* øJB01 phage, gp17 was aligned with homologous sequences from *Escherichia* phage N30 and T7 phage. Identical amino acids are marked with asterisks and non-identical are marked by dots. The gp17 conserved region is from position 1- 250 and the hypervariable region from 300-554.



Extended Data Fig. 4 | Phylogenetic tree based on the gp17 sequences of 26 phage clones from the three-host treatment (EPEC, MG1655, REPEC). Sequences were aligned with MUSCLE (v3.8.31) in MEGA X[51,52]. The phylogenetic tree was reconstructed using the maximum likelihood method as implemented in PhyML (v3.1/3.0)[53,54]. The WAG substitution model[55] was selected assuming an estimated proportion of invariant sites (of 0.951) and 4 gamma-distributed rate categories to account for rate heterogeneity across sites. The gamma shape parameter was estimated directly from the data (gamma=91.073). Reliability of internal branches was assessed using the aLRT test (SH-Like). Final Log-Liklihood: −1839.18. Phage clones labelled by phenotype and day of generation. 'S' or 'G' refers to specialist or generalist according to phenotypic measurements (Figs. 3–4) and the number refers to the day of isolation. C8, D8, C9, D9, E8, F8 refer to the specific three-host experimental population from which the phage clone was isolated. Each phage clone was obtained by plating the whole population lysate onto an individual host (Figs. 3–4). The strain name refers to that host of isolation. For example, 'G10 C8 MG1655' refers to a phage clone with a generalist phenotype isolated from a day 10 population C8, isolated from a plaque on MG1655. Numbers on branches show bootstrapping support for that branch (percent, 100 bootstraps). Strains are clustered based on those clades with greater than 80% support. Dashed lines are for labels and are not part of the tree.

nature research

Corresponding author(s): Michael J. McDonald

Last updated by author(s): Oct 30, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We used Breseq and Gdtools analysis pipeline to identify mutations,
Data analysis	No software was used for the data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Raw sequencing reads used to generate the data in Fig. 4A and Extended Data Figure 4 have been deposited in GenBank under the Bioproject ID: PRJNA673261

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. The number of evolving populations was selected to exceed previous evolution experiments with phage, with the upper limit determined based on the upper number of replicates possible given resource constraints
Data exclusions	No data was excluded from the analyses.
Replication	We repeated out initial evolution experiment with a similar evolution experiment that included some variations in populations size (described in the main text). This repeat experiment confirmed the results of the first experiment.
Randomization	All populations were measured. Clones were selected by independent picking of plaques, there was not sufficient variation in plaque size to bias towards the selection of fit or unfit plaque. All measurements that were compared (eg whole population infectivity assays) were carried out in a single experiment, thereby eliminating variation across treatments due to "batch" effects.
Blinding	Blinding was not explicitly incorporated into the study design.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging