OPEN ACCESS Freely available online

PLOS | PATHOGENS

Evolution of Virulence in Emerging Epidemics

Thomas W. Berngruber[1]*, Rémy Froissart[2,3], Marc Choisy[3,4], Sylvain Gandon[1]

1 Centre d'Ecologie Fonctionnelle et Evolutive (CEFE) - UMR 5175, Montpellier, France, **2** Laboratoire Biologie & Génétique des Interactions Plantes-Parasites (BGPI), UMR 385 INRA-CIRAD-SupAgro, Campus international de Baillarguet, Montpellier, France, **3** Laboratoire Maladies Infectieuses et Vecteurs : Ecologie, Génétique, Evolution et Contrôle (MIVEGEC), UMR 5290 CNRS-IRD-Université de Montpellier I-Université de Montpellier II, Montpellier, France, **4** Oxford University Clinical Research Unit, Hanoi, Vietnam

Abstract

Theory predicts that selection for pathogen virulence and horizontal transmission is highest at the onset of an epidemic but decreases thereafter, as the epidemic depletes the pool of susceptible hosts. We tested this prediction by tracking the competition between the latent bacteriophage λ and its virulent mutant λcI857 throughout experimental epidemics taking place in continuous cultures of *Escherichia coli*. As expected, the virulent λcI857 is strongly favored in the early stage of the epidemic, but loses competition with the latent virus as prevalence increases. We show that the observed transient selection for virulence and horizontal transmission can be fully explained within the framework of evolutionary epidemiology theory. This experimental validation of our predictions is a key step towards a predictive theory for the evolution of virulence in emerging infectious diseases.

Citation: Berngruber TW, Froissart R, Choisy M, Gandon S (2013) Evolution of Virulence in Emerging Epidemics. PLoS Pathog 9(3): e1003209. doi:10.1371/journal.ppat.1003209

Editor: François Balloux, University College London, United Kingdom

Received July 25, 2012; **Accepted** January 9, 2013; **Published** March 14, 2013

Funding: The work was funded by the ANR grant EPICE 07 JCJC 0128, and the ERC grant EVOLEPID 243054 to SG. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: berngruber@gmail.com

Introduction

Understanding and predicting the conditions under which pathogens evolve towards higher levels of virulence (pathogen induced host mortality) is a major challenge in the control of infectious diseases [1,2]. Nevertheless, the theoretical understanding of virulence evolution is often based on several major simplifying assumptions. In particular, the classical adaptive dynamics framework assumes that mutations are rare and thus that evolution occurs on a much slower time scale than epidemiological dynamics [2]. In other words, adaptive dynamics theory relies on the assumption that there is very little amount of genetic variation in the pathogen population and that a single pathogen strain reaches an equilibrium before a new strain arises by mutation. However, ecological and evolutionary time scales may overlap when the amount of genetic variation is high [3–5]. This is the case for many pathogens, and in particular for viruses with large mutation rates. The recurrent introduction of new mutants violates a major assumption of adaptive dynamics since many different strains may compete with each other before the system reaches a new endemic equilibrium [6–10]. Previous theoretical analyses suggest that the outcome of this competition changes strikingly throughout an epidemic; even though selection can act against virulent mutants at the endemic equilibrium, there is a transitory phase during the early stage of the epidemic where the abundance of susceptible hosts can favor the more transmissible and aggressive strains [11–18].

Studying selection on virulence in the field is notoriously difficult because the characterization of the pathogen phenotypes can be obscured by host heterogeneity and healthcare measures. The unambiguous demonstration of the evolution of virulence evolution during an epidemic requires an experimental approach. Two different types of experimental setups can be used [19]. First, in a top-down approach, the evolution of the pathogen population is monitored in different environments (e.g. before and after an epidemic). In this case, making quantitative predictions on the epidemiology and evolution of the pathogen remains out of reach because evolutionary trajectories rely on random mutations occurring during the experiment. In contrast, the bottom-up approach attempts to measure and/or manipulate the initial amount of genetic variation in the pathogen population and try to predict the evolution from this standing genetic variation. Although many stochastic factors like new mutations may alter the quality of the predictions in the long term, this approach may provide good quantitative predictions in the short term. For this reason, we follow the bottom-up approach to analyze the interplay between the epidemiology and the evolution of the bacteriophage λ. To study the dynamics of selection on virulence we monitor the competition of the bacteriophage λ and its virulent mutant λcI857 throughout the development of an epidemic in continuous cultures of *E. coli*. Bacteriophage λ is a typical temperate virus which integrates into the host genome and transmits vertically to daughter cells at cell division. Integration of phage λ into the genome protects the host cell against superinfection of other λ phage particles and this way provides lifelong immunity to superinfection by other λ phage particles [20]. Nevertheless stochastic reactivation of the integrated phage results in lysis and destruction of the host cell, causing pathogen induced host mortality. Lysis of its host prevents vertical transmission but allows the phage to be transmitted horizontally to uninfected susceptible cells. Whereas the non-virulent λ wildtype transmits mostly vertically by dormant integration into the host genome, the virulent mutant λcI857 transmits mostly horizontally by host lysis

Author Summary

Why are some pathogens more virulent than others? Theory predicts that pathogens that 'keep their host alive' can sometimes outcompete virulent pathogens in times when transmission to new susceptible hosts is unlikely. Yet, this prospect of finding a new susceptible host changes itself throughout an epidemic. In the early stage of an epidemic susceptible hosts are abundant and virulent pathogens that invest more into horizontal transmission should win the competition. Later on, the spread of the infection reduces the pool of susceptible hosts and may reverse the selection on virulence. This may favor benign pathogens after the acute phase of the epidemic. We model this transient benefit for virulence and predict both the epidemiology and the evolution of pathogens during an epidemic. To put these predictions to the test we monitor the competition of the temperate bacterial virus λ and its virulent mutant λcI857 in experimental epidemics. Our experimental results agree remarkably well with all our theoretical predictions. This demonstrates the ability of evolutionary epidemiology to predict selection for virulence in an ongoing epidemic.

(see Figure 1). This difference in virulence and transmission mode is the result of a point mutation in the λ virulence repressor protein cI which actively controls the decision to 'kill or not to kill' the host cell [21–23]. This active control of the fate of the infected cell has been shown to respond rapidly to different selection regimes [24]. Studying the competition between such cI variants is particularly relevant to study phage evolution during and epidemic.

In order to predict the competition between the temperate λ and the virulent λcI857 we first measured the effect of the cI857 mutation on several aspects of the viral life-cycle. In particular we focused on the effects of the cI857 mutation on the life-history traits known to be under the direct control of protein cI. We thus measured ϕ the ability to integrate into the genome of its host after infection, and α the spontaneous lysis rate of the lysogenic bacteria for both the λ wildtype and the mutant λcI857 (Figure S3). We used these life-history estimates and other estimates from the literature to parameterize an evolutionary epidemiology model which generated three clear-cut predictions. Our evolutionary experiments confirmed all three predictions and thus demonstrate the predictive power of evolutionary epidemiology theory.

Results

Evolutionary epidemiology theory

We modeled the competition of the temperate bacteriophage λ with its virulent mutant λcI857 throughout the course of an epidemic in chemostat cultures of its bacterial host *E. coli*. To understand and predict the competition dynamics of these two viruses throughout an epidemic we first developed a mathematical model. The epidemiology of phage λ can be described by the following set of ordinary differential equations for the densities of susceptible hosts, S, infected hosts, I, and free viral particles, V:

$$\begin{aligned}\dot{S} &= rS(1-(S+I)/K) - abSV - mS \\ \dot{I} &= \rho I(1-(S+I)/K) + ab\varphi_{\bullet}VS - (\alpha_{\circ}+m)I \\ \dot{V} &= ab(1-\varphi_{\bullet})VSB + \alpha_{\circ}BI - mV - a(S+I)V\end{aligned} \quad (1)$$

where susceptible hosts and infected hosts grow at rate r and ρ, respectively, to a carrying capacity K and die at a background mortality rate m. Infected hosts spontaneously switch to lysis at

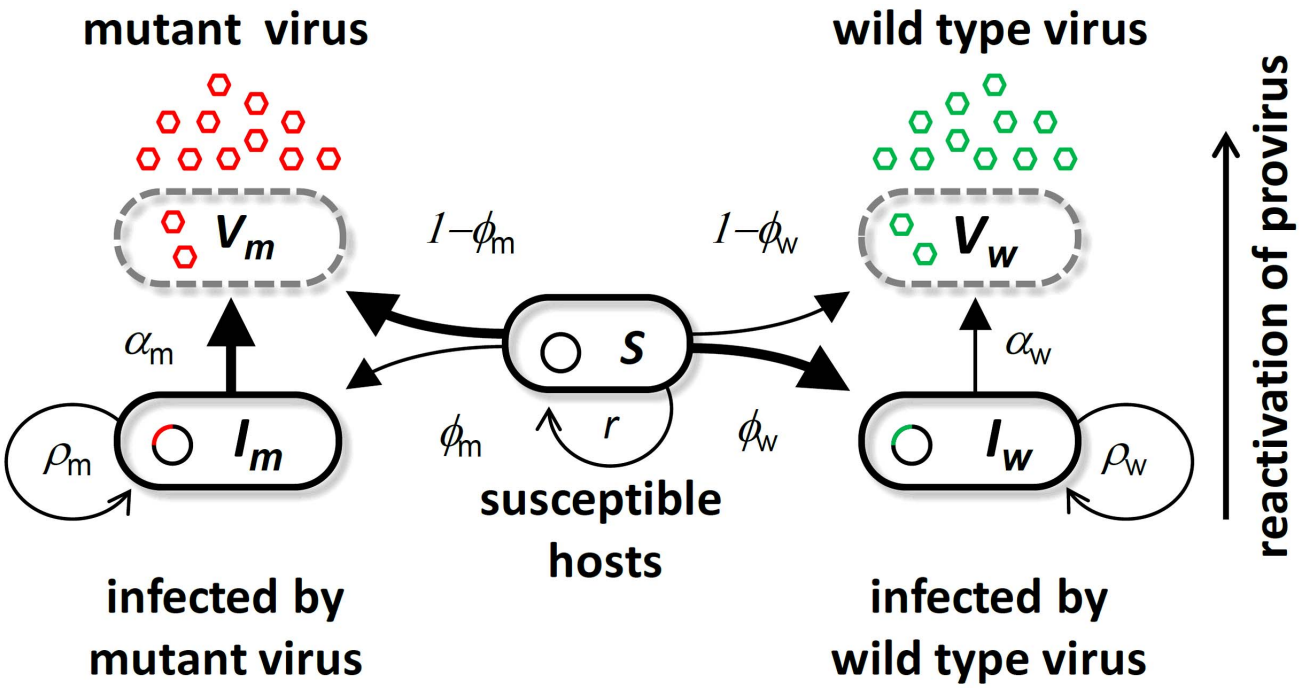

Figure 1. Schematic representation of the bacteriophage λ life cycle. Free viral particles of the wild type virus V_W (green) and the virulent mutant V_M (red) infect susceptible cells S. A proportion of successful infections leads to genome integration at rate ϕ_W and ϕ_M to produce infected cells I_W and I_M or results in cell lysis at rate $1-\phi_W$ and $1-\phi_M$. Infected cells lyse through spontaneous reactivation of the provirus at rate α_W and α_M for I_W and I_M, respectively. (See Table S1 in Text S1 for the definition and the values of all the parameters of this model).
doi:10.1371/journal.ppat.1003209.g001

rate α_o and produce B free viral particles. Viral particles adsorb to bacterial cells at rate a, and inject their genome with probability b. Injected viral genomes either replicate and destroy the host cell with probability $1-\phi_\bullet$ to release B viral particles, or integrate into the host genome with probability $\phi_\bullet$. Integration of the virus into the genome of the host cell excludes superinfection by a second phage particle of the same kind [20]. Open and closed subscripts indicate averaged trait values taken over the provirus and the free virus stage, respectively (see section S1.2 in Text S1). To capture the evolution of viral traits we track the frequency of a strain i among all genome integrated provirus p_i, and among free viral particles q_i as follows (see section S1.2 in Text S1):

$$\dot{p}_i = p_i\left(-\overbrace{(\alpha_i-\alpha_o)}^{virulence}+\overbrace{abS\frac{V}{I}(\phi_i-\phi_\bullet)}^{genome\ integration}\right)+\underbrace{abV(q_i-p_i)\left(\frac{S}{I}\phi_i\right)}_{gene\ flow} \quad (2.1)$$

$$\dot{q}_i = \overbrace{abSBq_i(\phi_\bullet-\phi_\circ)}^{failed\ genome\ integration}+B\frac{I}{V}\left(\overbrace{q_i(\alpha_i-\alpha_\circ)}^{virulence}+\underbrace{\alpha_i(p_i-q_i)}_{gene\ flow}\right) \quad (2.2)$$

The above two equations readily show the different forces that affect the change in frequency of the virulent type in these two stages of the virus life cycle. First, in the provirus stage, the frequency of a virulent mutant decreases because of its increased lysis rate (first term in 2.1) and its lower rate of genome integration (second term in 2.1). But this frequency may increase by gene flow from the free virus stage (the last term in 2.1), because the frequency of the virulent mutant tends to be higher in the free virus stage (see prediction ii below). Second, in the free virus stage, the frequency of a virulent mutant increases because it has a lower rate of genome integration and a higher rate of lysis (first two terms in 2.2). Yet this frequency may decrease by gene flow from the provirus stage (last term in 2.2), because the frequency of the virulent mutant tends to be lower in the provirus stage (see prediction ii below).

Epidemiology, evolution and their interactions are fully integrated in the above five equations. Three general predictions emerge from the analysis of this model (see Figure 2): (i) The virulent mutant initially wins the competition with the wildtype when susceptible hosts are abundant, but the competitive outcome is reversed as soon as the epidemic reaches high prevalence; (ii) The virulent mutant is, at all times, more frequent among free viruses than among proviruses; and (iii) Lower initial prevalences result in a higher increase in virulence during the epidemic.

Experimental evolutionary epidemiology

To test these three predictions, we infected *E. coli* chemostat cultures with a range of initial infection prevalences (between 1% and 100%) and monitored viral competition between λcI857 and λ (in a 1:1 starting ratio at the provirus stage) during the spread of the epidemic. We introduced two fluorescent protein marker colors (CFP and YFP) into λcI857 and λ strains to measure their frequencies during competition. To experimentally control for a small marker color effect (see Table S2.1 in Text S1), we replicated the experiment switching marker colors between the two virus strains. We monitored strain frequencies in the provirus by flow cytometry, and in the free virus by marker specific qPCR (see section S2.1.2 in Text S1 and Figure S4).

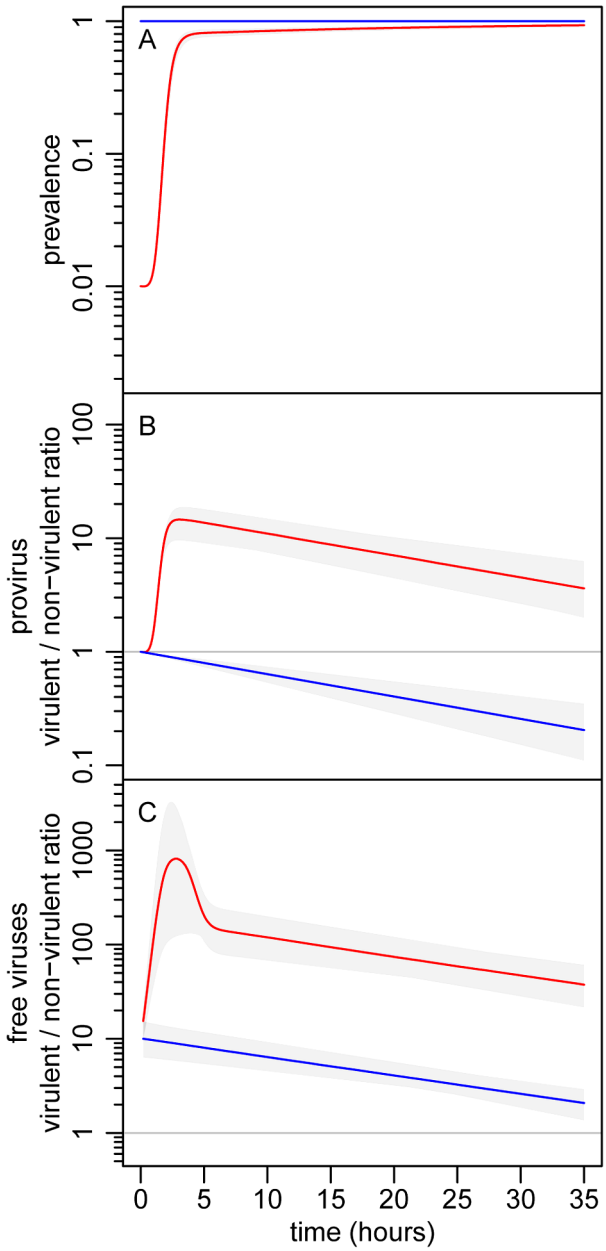

Figure 2. Theoretical evolutionary epidemiology. (A) change in prevalence (proportion of infected bacteria). (B) change in the λcI857/λ ratio in the provirus stage. (C) change in the λcI857/λ ratio in the free virus stage. The initial value of the λcI857/λ ratio in the provirus was 1:1, and two initial prevalence values were considered: 1% (red) and 100% (blue). We ran 10000 simulations of our model (see equations (1) and (2)) allowing some variation over the phenotypic values (ϕ and α) of the two virus strains. The gray envelopes show the range of variation among all simulation runs and colored lines show the median of these simulations (see section S1.3 in Text S1). See Table S1 in **supporting Text S1** for other parameter values.
doi:10.1371/journal.ppat.1003209.g002

In a first experiment, we tracked the change in prevalence (proportion of infected bacteria) and strain frequencies (in both the provirus and the free virus stages of the phage life cycle) by sampling hourly in 8 chemostats (2 marker/virulence combinations, 2 replicates and 2 initial prevalences: 1% and 100%). We performed a second experiment to further investigate the impact

of initial prevalence using 6 chemostats (2 marker/virulence combinations and 3 initial prevalences: 1%, 10% and 99%; see section S2.3 in Text S1).

Starting from an initial prevalence of 1% the epidemic spread rapidly until nearly all hosts were infected, roughly 10 h later (Figure 3A). The virulent λcI857 rapidly outnumbered λ in both the provirus and the free virus compartments. Yet, despite this initial advantage in the first 7 h of the epidemic, the frequency of the virulent λcI857 started to decrease in both compartments thereafter (Figure 3 and Figure 4). This confirms our first theoretical prediction. Furthermore, the frequency of the virulent mutant λcI857 remained higher among free viruses than among proviruses during the entire course of the epidemics (Figure 3 and Figure 4). This confirms our second prediction. Moreover, as expected from our third prediction, at an initial prevalence of 100% the virulent mutant λcI857 lost the competition from the outset of the experiment (Figures 3B,C). The third prediction got additional support from the second experiment where the value of the peak frequency of the virulent mutant decreased with higher initial prevalence (Figure 4 and Figure S6).

Discussion

To demonstrate that epidemiology can affect selection on viral virulence and transmission mode we studied the competition between two viral strains during experimental epidemics in chemostats. The two main life-history traits that govern virulence and transmission in λ (ϕ the ability to integrate into the genome of its host after infection, and α the lysis rate of the lysogenic bacteria) were measured for these two viral strains. We parameterized a model for the competition of several viral strains throughout an epidemic using estimations for the remaining parameter values from other published studies (Table S1 in Text S1). This model was used to generate three qualitative predictions on the epidemiology and the evolution of the bacteriophage λ. Our experimental results agree well with all three theoretical predictions. This demonstrates the predictive power of the bottom-up approach we used to model this system. This study confirms the importance of modeling both epidemiology and evolution to accurately predict the transient evolution of pathogens. In particular, the shift between positive and negative selection on the virulent mutant makes only sense because we took into account the feed-back of epidemiology on the evolution of the virus.

Our theoretical predictions are based on the competition between two pathogen variants in a fully susceptible host population. In this way we focus on the short-term evolutionary dynamics taking place during an epidemic. As pointed out above, the accuracy of these predictions is expected to drop as other mutations in virus or host come into play. In particular, compensatory mutations that reduce phage virulence could alter the ultimate trajectory of λcI857, as it would no longer pay the cost of virulence in the long run. In the present experiment, we did not find evidence of such compensation (Figure S5). Nevertheless, we can readily include compensation in our model to judge its effect on short and long-term dynamics. We found that if we allow for these compensatory mutations the above short term predictions still hold (Figure S1). Another evolutionary route the pathogen could take is to escape superinfection inhibition, which would allow it to gain access to hosts even when all the bacteria are infected [25]. However, in λ the rate of mutation towards such ultravirulent strains has been shown to be very small and is thus unlikely to affect the short-term evolutionary dynamics [26]. In the long term, however, the coevolution between superinfection inhibition and the resistance to superinfection inhibition may play

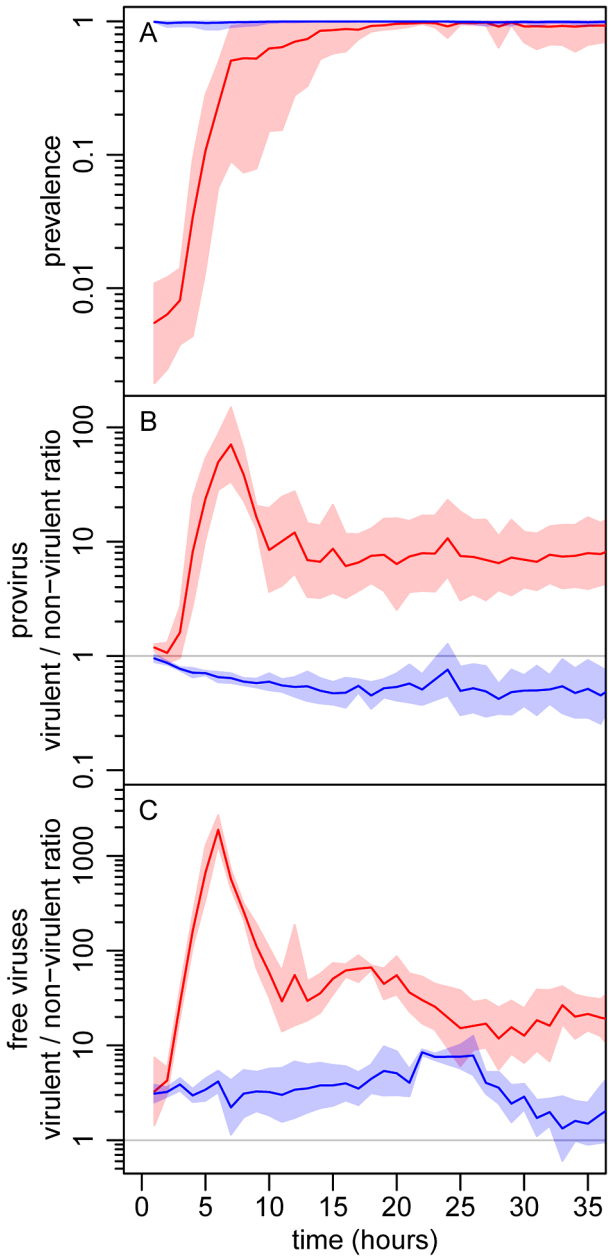

Figure 3. Experimental evolutionary epidemiology. (A) change in prevalence (proportion of infected bacteria). (B) change in the λcI857/λ ratio in the provirus stage. (C) change in the λcI857/λ ratio in the free virus stage. The initial value of the λcI857/λ ratio in the provirus was 1:1, and competition was started from two initial prevalence values: 1% (red) and 100% (blue). The data was obtained from the first experiment. The lines are the mean over four chemostats (2 marker/virulence combinations with 2 independent replicates), and the envelopes show the 95% confidence intervals of the log transformed data.
doi:10.1371/journal.ppat.1003209.g003

a major role for the evolutionary maintenance of viral latency and the emergence of diversity in λ-like phages [27]. In addition, the *E. coli* host cell may also acquire resistance to λ by well known mechanisms [28–29]. We did find some evidence of host resistance evolution but only in the large volume chemostats (50 mL, second experiment) and not before 40 h (Figure S7 and Figure S8), which indicates that the appearance of these mutations is limited by population size and by time. Again, including host resistance in

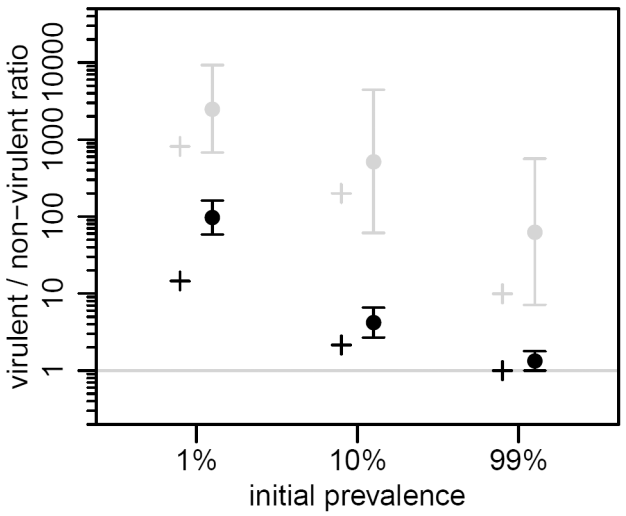

Figure 4. Effect of initial prevalence on transient virulence evolution, in the two life stages of the virus. We plot the maximal value of the λcI857/λ ratio from our theoretical model (crosses, same parameter values as in Figure 2) and from the second experiment (dots and vertical bars are the means and their 95% confidence intervals over two chemostats of the log transformed data, see supplementary information). The λcI857/λ ratio is shown for both the provirus (black) and the free virus (gray) stages. The λcI857/λ ratio is significantly higher among free viruses than among proviruses, and decreases significantly as the initial prevalence increases (maximal λcI857/λ ratio in the first 15 hours of the second experiment, see also Figure S6).
doi:10.1371/journal.ppat.1003209.g004

our model confirms that the above three qualitative predictions hold in the short-term (Figure S2).

Our model assumes that all the parameter values governing phage life cycle remain constant throughout the experiment. The lyzogenisation rate of phage λ, however, is known to vary with the multiplicity of infection (MOI), which is the number of virus entering the bacteria. The higher the MOI, the higher the lysogenization rate [22,30]. For the sake of simplicity we do not consider this effect in our model but additional simulations indicate that it does not alter qualitatively our conclusions (not shown) because both the wild type and the cI857 mutant harbor this phenotypic plasticity [30]. Our model, however, may shed some light on the adaptive nature of the sensitivity of the rate of lysogenisation to the MOI. The MOI provides accurate information of the epidemiological state of the environment. When the MOI is low the number of susceptible hosts is likely to be high and it is a good strategy to lyse and to try infecting new hosts horizontally. In contrast, when the MOI becomes high, it is very unlikely that a free virus particle will encounter a susceptible host. In this case the phage would benefit more from investing into lyzogeny and vertical transmission [31–32]. In other words, the evolution of plasticity may be another evolutionary outcome resulting from the epidemiological feedback during an epidemic.

When do we expect epidemiology to feed back on the evolution of virulence? In our experimental system, this feedback operates because the spread of the virus in the population reduces the density of susceptible hosts. This erodes the benefit of virulence (horizontal transmission) while the cost of virulence (induced host mortality) remains. Note that this qualitative result is robust to variations of the initial frequency of the mutant strain (not shown). In our system, the cost of virulence acts mainly via the reduction of vertical transmission. Yet, in the absence of vertical transmission, a similar pattern is expected when horizontal transmission carries other costs. In many lytic phages these fitness costs result from the trade-off between lysis time and burst size [33]. The virus with short lysis time (and small burst size) may only outcompete the virus with longer lysis time (and larger burst size) at the beginning of the epidemic when the availability of susceptible bacteria is maximal [34]. Similar patterns of transient evolution are also expected in pathogens that transmit throughout the course of the infection. In this case, higher rates of horizontal transmission are often associated with more aggressive host exploitation strategies which reduce host life-span and pathogen's infectious period. Shortened infectious period can result in substantial fitness costs for the virulent pathogens. It is only during the early stages of an epidemic that such virulent strains may outcompete the others [3–6]. Hence, the transient evolution we report in our study is expected whenever there is a fitness cost associated with increased virulence and horizontal transmission.

Epidemic feedback on the evolution of virulence is likely to be widespread and could affect many other pathogens. For example, this feedback may also operate during viral invasion into a multicellular host organism (within-host evolution). During this within-host spread of the infection the availability of susceptible cells is expected to drop. Noteworthy this effect is particularly strong for viruses with superinfection exclusion like herpes- and retroviruses as well as phage λ, where infected cells remain resistant to a second infection and can vertically pass on this immunity to daughter cells [35–36]. A similar evolutionary trajectory is expected in large scale epidemics (between-host evolution) when the spread of the infection reduces the availability of susceptible hosts because both infected and recovered individuals tend to be immune to new infections [36,37]. If multiple infections or superinfections are possible, the evolution of virulence results from selection acting on two different levels: within and between hosts [38–40]. Nevertheless it is possible to modify our model to allow for this additional level of complexity but the evolutionary outcome depends on the relative competitive abilities of different variants which can be obtained in some pathogens [40] and could thus be used to provide quantitative predictions when these two levels of selection are acting on the evolution of virulence.

Our experimental epidemics occur in a small and well-mixed community free of the considerable complications arising from stochasticity, multi-host life cycles, host immunity, input of new mutations and spatial structuring of host populations. Albeit a necessary simplification of a more complex reality, our experiments provide a unique opportunity to understand pathogen evolution during the course of an infection (within-host dynamics). In HIV and HCV, for instance, the replicative fitness and the ability of the virus to resist therapeutic drugs has been shown to change throughout the course of the infection [41–43]. The understanding of such within-host evolution is key to providing more effective therapies that are not necessarily aimed towards eradication of the pathogen but towards patient health improvement [44–45]. On a larger spatial scale, the interplay between epidemiology and evolution that we demonstrate here can have far reaching consequences for emergent epidemics during the spread into a host population (between-host dynamics) [6]. More specifically, we expect drastic changes of virulence between early and later stages of global pandemics, but also between different stages of the pathogen life cycle. Our joint theoretical and experimental approach is a first step towards a framework aiming to forecast both the epidemiological and evolutionary trajectories of pathogens.

Materials and Methods

Fluorescently labeled phage construction

The GFP-Kan cassette from λGFP (from reference [21]) was amplified with primers λ20165F (CGCAGAAGCGGCATCAG-CAA) and λ22543R (GGACAGCAGGCCACTCAATA) and subcloned into pUC18. Subsequently GFP was mutated to CFP and YFP by a quick-change PCR approach with megaprimers amplified from pDH3 and pDH5 (Yeast Resource center, University of Washington, primers: Fwd TGGCCAACACTTGT-CACTAC, Rev AGAAGGACCATGTGGTCTCT). CFP-Kan and YFP-Kan cassettes were integrated into the prophage of K12[λ] and KL740[λcI857] (Yale E.coli Stock Center) by the aid of recombineering plasmid pKM201 (Addgene). Fluorescent lysogens were induced by 10 μM MitomycinC and chloroformed lysate was used to reinfect *E. coli* MG1655.

Life-history of fluorescently marked viral strains (Figure S3)

Our method for the detection of selection on virulence is based on the competition of the non-virulent λ wildtype against the virulent λcI857. In order to verify the expected differences in life-history traits between those strains and to obtain rough parameter estimates for the simulations we measured the viral life-history traits *virus production* (PFU/mL), *genome integration rate* (% lysogenized) and *vertical transmission* (CFU/mL). We determined these traits for all constructed viruses (λcI857CFP, λcI857YFP and λCFP, λYFP) by independent life-history assays prior to competition in the chemostat. The life-history traits were measured by the following three independent assays. (1) Virus production (PFU/mL) (Figure S3A) was determined by growing lysogen cultures to OD600 nm = 0.6 at 30°C and shifting them to 35°C and 38°C for 2 h until lysis occurred. From these lysates, viral titers were determined by qPCR on a Roche LightCycler480 (primers F:5′AATGAAGGCAGGAAGTA3′ R:5′GCTTTCCAT TCCATCGG3′). Viral titers were calculated from a calibration curve based on CP values of a dilution series of a lysate of λvir of known titer (3×109 pfu). (2) Vertical transmission (CFU/mL) (Figure S3A) was measured by diluting lysogen cultures of λCFP, λYFP and λcI857CFP, λcI857YFP to OD600 nm = 0.07 and growing them for 6 h at 35°C and 38°C in eight replicates each in 96-well plates on a Titramax shaker (Heidolph, Germany) at 900 rpm. Every hour OD600 nm was measured in an Infinity200 microplate reader (Tecan, Austria). OD600 nm values were converted to CFU's by a calibration curve which was obtained by plating. (3) Lysogenization rate (Figure S3B) was determined by challenging non-infected E.coli MG1655 with 10^8 PFU/mL free virus particles of λCFP, λYFP, λcI857CFP and λcI857YFP for 24 h. After 24 h, the proportion of lysogenized (fluorescent) cells was determined by flow cytometry.

Chemostat growth conditions

Media, 0.25 x LB, 0.2% Maltose, 10 mM MgSO4 and 5 mM IPTG. Dilution rate 0.8/h with 5 mL chamber volume at 35°C (50 mL chamber volume in the second experiment). Samples were drawn at 1 h intervals and stored at 2°C in 10 mM Na-citrate.

Quantifying competition in the provirus stage by flow cytometry (Figure 2B and 3)

To follow the relative prevalence of each strain in experimental epidemics we distinguished cells infected by CFP and YFP tagged virus through flow cytometry. CFP was detected at 405 nm excitation and 510/50BP emission and YFP was detected at 488 nm excitation and 530/30BP emission on a BD FacsCantoII flowcytometer. We used the FlowJo7 (Tree Star, Inc.) software to apply automatic compensation and gating.

Quantifying competition in the free virus stage by marker specific qPCR (Figure 2C and 3)

To quantify the amount of each strain in the free virus stage we developed CFP and YFP specific qPCR primer sets that match the functional substitution T203Y, which is responsible for the spectral shift from GFP to YFP and V163A and N164H which are used as fluorescence enhancers of CFP (Table S3 in Text S1). Our primers show no non-specific amplification even in large excess of the non-specific template (Figure S4) but high amplification efficiencies (AE) on their specific template (AE = 2.0 for CFP and AE = 1.98 for YFP).

Supporting Information

Figure S1 Theoretical evolutionary epidemiology (analogous to Figure 2) for a modified model which allows for virulence compensation. (A) change in prevalence (proportion of infected bacteria). (B) change in the λcI857/λ ratio in the provirus stage. (C) change in the λcI857/λ ratio in the free virus stage. The virus mutation probability on virulence is $\varepsilon_{ij} = \varepsilon_{ji} = 0.01$. See Table S1 in Text S1 for other parameter values. (Red and blue line: 1% and 10% initial prevalence. Gray envelopes show the range of variation among the 10000 simulation runs and colored lines show their median).
(TIF)

Figure S2 Theoretical evolutionary epidemiology (analogous to Figure 2) for a modified model which allows for mutation towards host resistance. (A) change in prevalence (proportion of infected bacteria). (B) change in the λcI857/λ ratio in the provirus stage. (C) change in the λcI857/λ ratio in the free virus stage. The host mutation probability towards resistance is $\mu = 0.01$, and the cost of resistance is assumed to be $c = 0.01$. See Table S1 in Text S1 for other parameter values. (Legend similar to Figure S1, except black line: frequency of resistant cells).
(TIF)

Figure S3 Life-history of the constructed viral strains λcI857CFP, λcI857YFP and λCFP, λYFP. (**A**) **Horizontal (free phages PFU/mL) and vertical transmission (infected cells CFU/mL).** At 35°C: λcI857CFP and λcI857YFP show significantly higher horizontal transmission and reduced vertical transmission compared to the wildtype constructs λCFP and λYFP. At 38°C, horizontal transmission of λcI857CFP and λcI857YFP is further increased and vertical transmission is further reduced. (**B**) **Genome integration rate (% lysogenized cells at 24 h).** Lysogenization at 35°C is about 6 fold higher for λCFP, λYFP than for the mutants λcI857CFP and λcI857YFP ((A) Crosses, 95% CI. (B) Bars, 95% CI. Gray circles, raw data, see section S2.1.1 in Text S1 for statistical analysis).
(TIF)

Figure S4 Test for cross-specificity of CFP and YFP specific qPCR primers. Primers at 1 μM concentration were tested on dilution series of pure λCFP and λYFP lysates (10^9 to 10^2 pfu in 10-fold steps) as well as a reciprocal mixture of the dilutions series (5×10^8 : 5×10^1 to 5×10^1 : 5×10^8 pfu/mL of λCFP : λYFP). qPCR on the reciprocal lysate mixtures shows no non-specific quantification even with a 10^7 fold excess of the non-specific template (see Table S3 in Text S1 for primers).
(TIF)

Figure S5 Test for the occurrence of mutations that compensate virulence. As a proxy for virulence we calculated viruses produced per infected cell (viruses/cell) for the λcI857 mutant strains and the wildtype strains. The λcI857 produces more viruses/cell and is more virulent than the wildtype as long as the ratio $(\text{viruses/cell})_{\text{mutant}}$ devided by $(\text{viruses/cell})_{\text{wildtype}}$ is larger than 1. Indeed this ratio is significantly larger than 1 throughout the experiment (except t = 25 and 33 h in the 1% treatment). Hence, the λcI857 mutants have remained more virulent than the wildtype even if compensatory mutations might have occurred. (blue area: 1% initial prevalence, red area: 100% initial prevalence, solid line and shading: Mean and 95% CI envelope of the log transformed data from 4 chemostats).
(TIF)

Figure S6 Competition dynamics in the second chemostat experiment with initial prevalence of 1%, 10% and 99%. (A) Prevalence, (B) fitness benefit for the provirus and (C) free virus in the second chemostat experiment. The 1% initial prevalence treatment (blue) leads to the highest benefit of virulence. This benefit of virulence decreases in the 10% and 99% initial prevalence (green and red). The maxima of B and C between t = 0 and t = 15 h were extracted to create Figure 4 in the main text (Solid lines: Mean, shading: 95% CI interval of log transformed data from 2 chemostats pooled by color replicate).
(TIFF)

Figure S7 Invasion of resistant host cells. In the second experiment, resistant, but non-infected, host cells invaded 5 out of 6 chemostats after 40 h and caused a drop in overall prevalence. Chemostat 5 (red solid line) showed no invasion of resistant cells and maintained high prevalence. (Solid lines: λCFP versus λcI857YFP and dotted lines λYFP versus λcI857CFP. Blue, green and red: 1%, 10% and 99% initial prevalence, black numbers correspond to the numbering in Figure S8).
(TIF)

Figure S8 Test for the origin of resistance. We cross-streaked individual colonies from t = 60 h of each chemostat against the indicator strain λKH54h80ΔcI (this strain enters through the FhuA receptor and only lyses non-lysogens). Resistant hosts from chemostats 1, 2, 3, 4 and 6 were sensitive to λKH54h80ΔcI and are therefore non-lysogens and most likely lamB resistance mutants. Since most colonies from chemostat 5 carry a prophage they are not lysed by the indicator strain.
(TIF)

Text S1 Supporting text file containing the theoretical appendix, supporting experiments and statistics.
(PDF)

Acknowledgments

We thank Samuel Alizon, Troy Day, Ana Rivero, Pedro Vale, Pierre-Alexis Gros and Sébastien Lion for discussions and four anonymous reviewers for their constructive comments. We are grateful to Francois St-Pierre for sharing λGFP. We thank Francois Gatchitch for technical assistance and Cedric Mongellaz (Rio Imaging Montpellier) and Philippe Clair (Plateforme qPHD Université Montpellier II) for advice on flow cytometry and qPCR.

Author Contributions

Conceived and designed the experiments: TWB SG. Performed the experiments: TWB. Analyzed the data: TWB MC SG. Contributed reagents/materials/analysis tools: RF. Wrote the paper: TWB SG.

References

1. Frank SA. (1996) Models of parasite virulence. Q Rev Biol 71: 37–78.
2. Dieckmann U, Metz JAJ, Sabelis MW, Sigmund K (2002) Adaptive dynamics of infectious diseases: in pursuit of virulence management. Cambridge: Cambridge University Press. 532 p.
3. Yoshida T, Jones LE, Ellner SP, Fussmann GF, and Hairston Jr NG. (2003). Rapid evolution drives ecological dynamics in a predator-prey system. Nature 424:303–306.
4. Hairston NG, Ellner SP, Geber MA, Yoshida T, Fox JA. (2005). Rapid evolution and the convergence of ecological and evolutionary time. Ecology Letters 8:1114–1127.
5. Bull JJ, Millstein J, Orcutt J, Wichman HA. (2006). Evolutionary feedback mediated through population density, illustrated with viruses in chemostats.Am. Nat. 167:E39–E51.
6. Arista S, Giammanco GM, De Grazia S, Ramirez S, Lo Biundo C, et al. (2006) Heterogeneity and temporal dynamics of evolution of G1 human rotaviruses in a settled population. J Virol 80: 10724–10733.
7. Gallimore CI, Iturriza-Gomara M., Xerry J, Adigwe J, Gray JJ. (2007) Inter-seasonal diversity of norovirus genotypes: Emergence and selection of virus variants. Arch Virol 152: 1295–1303.
8. Ojosnegros S, Beerenwinkel N, Antal T, Nowak MA, Escarmis C, et al. (2010) Competition-colonization dynamics in an RNA virus. Proc Natl Acad Sci USA 107: 2108–2112.
9. Wright CF, Morelli MJ, Thébaud G, Knowles NJ, Herzyk P, et al. (2011) Beyond the consensus: Dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing. J Virol 85: 2266–2275.
10. Vijaykrishna D, Smith GJ, Pybus OG, Zhu H, Bhatt S, et al. (2011) Long-term evolution and transmission dynamics of swine influenza A virus. Nature 473: 519–522.
11. Lenski RE, May RM. (1994) The evolution of virulence in parasites and pathogens: Reconciliation between two competing hypotheses. J Theor Biol 169: 253–265.
12. Day T, Gandon S. (2007) Applying population-genetic models in theoretical evolutionary epidemiology. Ecol Lett 10: 876–888.
13. Bull JJ, Ebert D. (2008) Invasion thresholds and the evolution of nonequilibrium virulence. Evol Appl 1: 172–182.
14. Bolker BM, Nanda A, Shah D. (2010) Transient virulence of emerging pathogens. J R Soc Interface 7: 811–822.
15. Stewart FM, Levin BR. (1984) The population biology of bacterial viruses: why be temperate. Theor Popul Biol 26: 93–117.
16. Day T, Proulx SR. (2004) A General Theory for the Evolutionary Dynamics of Virulence. Am. Nat. 163: E40–E63.
17. Day T, Gandon S. (2006) Insights from Price's equation into evolutionary epidemiology. In: Feng Z Dieckmann U, Levin S, editors. Disease evolution: models, concepts and data analyses. The American Mathematical Society.
18. Gandon S, Day T. (2007) The evolutionary epidemiology of vaccination. J R Soc Interface 4: 803–817.
19. Dykhuizen DE, Dean AM. (2009) Experimental Evolution from the Bottom Up. In: Garland T, Rose MR editors. Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments. Berkley: University of California Press.
20. Ptashne M. (1992) A Genetic Switch: Phage Lambda and Higher Organisms. Oxford: Blackwell Publishers.
21. St-Pierre F, Endy D. (2008) Determination of cell fate selection during phage lambda infection. Proc Natl Acad Sci USA 105: 20705–20710.
22. Zeng L, Skinner SO, Zong C, Sippy J, Feiss M, et al. (2010) Decision making at a subcellular level determines the outcome of bacteriophage infection. Cell 141: 682–691.
23. Sussman R, Jacob F. (1962). Sur un système de répression thermosensible chez le bacteriophage lambda d'Escherichia coli. Comp. Rend. 254:1517–1519.
24. Refardt D, Rainey PB. (2010) Tuning a genetic switch: experimental evolution and natural variation of prophage induction. Evolution 64, 1086–1097.25.
25. Bailone A, Devoret R. (1978) Isolation of ultra-virulent mutants of phage lambda. Virology 84: 547–550.
26. Northrop JH (1968) Appearance of virulent bacteriophages in lysogenic e. coli cultures after prolonged growth in the presence of triethylenemelamine. J Gen Physiol 52: 136–143.
27. Berngruber TW, Weissing FJ, Gandon S. (2010) Inhibition of superinfection and the evolution of viral latency. J Virol 84: 10200–10208.
28. Clément JM, Lepouce E, Marchal C, Hofnung M. (1983) Genetic study of a membrane protein: DNA sequence alterations due to 17 lamB point mutations affecting adsorption of phage lambda. EMBO J 2: 77–80.
29. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, et al. (2012) Repeatability and Contingency in the Evolution of a Key Innovation in Phage Lambda. Science 335: 428–432.
30. Kourilsky P. (1973). Lysogenization by bacteriophage lambda. I. Multiple infection and the lysogenic response. Mol. Gen. Genet. 122:183–195.

31. Avlund M, Dodd IB, Semsey S, Sneppen K, Krishna S. (2009). Why do phage play dice?. J Virol. 83.
32. Joh RI, Weitz JS. (2011). To lyse or not to lyse: transient-mediated stochastic fate determination in cells infected by bacteriophages. PLoS Comput. Biol. 7:e1002006.
33. Bull JJ, Pfennig DW, Wang IN. (2004). Genetic details, optimization and phage life histories. Trends Ecol Evol 19:76–82.
34. Bull JJ (2006). Optimality models of phage life history and parallels in disease evolution. J Theor Biol 241:928–938.
35. Mador N, Panet A,Steiner I. (2002) The latency-associated gene of herpes simplex virus type 1 (HSV-1) interferes with superinfection by HSV-1. J Neurovirology 8: 97–102.
36. Nethe M, Berkhout B, van der Kuyl AC. (2005) Retroviral superinfection resistance. Retrovirology 2: 52.
37. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, et al. (2004) Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. Science 303: 327–332.
38. Nowak MA, May RM. (1994). Superinfection and the evolution of parasite virulence. Proc R Soc. Lond B 255: 81–89.
39. van Baalen M, Sabelis MW. (1995). The dynamics of multiple infection and the evolution of virulence. Am Nat 146: 881–910.
40. de Roode, JC, Pansini R, CheesmanSJ, Helsinki MEH, Huijben S, et al. (2005) Virulence and competitive ability in genetically diverse malaria infections. Proc Natl Acad Sci USA. 102: 7624–7628.
41. Arien KK, Vanham G, Arts EJ. (2007) Is HIV-1 evolving to a less virulent form in humans? Nat Rev Micro 5: 141–151.
42. Troyer RM, Collins KR, Abraha A, Fraundorf E, Moore DM, et al. (2005) Changes in Human Immunodeficiency Virus Type 1 Fitness and Genetic Diversity during Disease Progression. J Virol 79: 9006–9018.
43. Adiwijaya BS, Kieffer TL, Henshaw J, Eisenhauer K, Kimko H, et al. (2012) A viral dynamic model for treatment regimens with directacting antivirals for Chronic Hepatitis C Infection. PLoS Comput Biol 8:e1002339.
44. Read AF, Day T, Huijben S. (2011) The evolution of drug resistance and the curious orthodoxy of aggressive chemotherapy. Proc Natl Acad Sci USA 108: 10871–10877.
45. Alizon S, Luciani F, Regoes RR. (2011) Epidemiological and clinical consequences of within-host evolution. Trends Microbiol. 19: 24–32.

Supplementary information for "Evolution of virulence in emerging epidemics"

T.W. Berngruber, R. Froissart, M. Choisy and S. Gandon

Overview:

S1. Theoretical evolutionary epidemiology

S1.1 Epidemiology

S1.2 Evolution

S1.3 Simulations

S2. Experimental evolutionary epidemiology

S2.1 Methods

S2.1.1 Life-history of fluorescently marked viral strains

S2.1.2 Quantifying competition in the free virus stage by marker specific qPCR

S2.2 First chemostat experiment

S2.2.1 Effects of initial prevalence and marker color on competition

S2.2.2 Test for the occurrence of mutations that compensate virulence

S2.3 Second chemostat experiment

S2.3.1 Competition at initial prevalence 1%, 10% and 99%

S2.3.2 Invasion of resistant host cells

S1. Theoretical evolutionary epidemiology

S1.1 Epidemiology

We derive below a model to describe the epidemiology and the evolution of the temperate bacteriophage λ in a population of fully susceptible *E. coli* bacteria (see **Figure 1,** in the main text for a schematic description of the life cycle). This model can be used to understand the dynamics taking place in a chemostat. In the absence of infection, the bacteria is assumed to reproduce at a rate r, and to die at a rate m (where m refers to the outflow rate of the chemostat). Density dependence is assumed to decrease the fecundity and to limit the bacterial density to the carrying capacity K. We assume that multiple virus strains may circulate. For each strain i we model both the dynamics of the density I_i of infected bacteria (provirus stage) and the density V_i of viral particles circulating in the medium (free virus stage). We assume that lysis of an infected bacteria releases a constant number B (burst size) of virus particles. Free virus may die at a rate m (where, again, m refers to the outflow rate of the chemostat) or adsorb to both infected and uninfected bacteria at a rate a. The adsorbed virus may enter the cell with a probability b, and, with the probability ϕ_i, it may integrate in the bacterial genome of the bacteria. Infected bacteria are reproducing at a rate ρ_i and the virus is vertically transmitted with probability δ (the fidelity of vertical transmission). For the sake of generality, we assume that with probability ϕ_{ij} bacteria infected with strain i can be superinfected with a strain j which replaces strain i. Infected bacteria may lyse when the virus fails to integrate into the bacterial genome, with probability $1-\phi_i$, but also after genome integration at a constant rate α_i (the lysis rate). The virulence of the virus (the mortality of the host induced by the virus) hence depends on both $1-\phi_i$ and α_i. The above described life cycle yields the following system of ordinary differential equations:

$$\begin{aligned}
\dot{S} &= (rS + \rho_\circ(1-\delta)I)(1-N/K) - abSV - mS \\
\dot{I}_i &= \rho_i I_i \delta(1-N/K) + ab\phi_i V_i S - (\alpha_i + m)I_i + abV_i \textstyle\sum_j I_j \phi_{ji} - abI_i \sum_j V_j \phi_{ij} \\
\dot{V}_i &= ab(1-\phi_i)V_i SB + \alpha_i B I_i - mV_i - aNV_i
\end{aligned} \qquad \text{(A1)}$$

The total density of infected bacteria is $I = \sum_i I_i$, and the total density of free virus is $V = \sum_i V_i$, and $N = S + I$. The frequencies of strain i are $p_i = I_i/I$ and $q_i = V_i/V$ in the provirus and in the free-virus stage, respectively. We use the following notations to refer to the value of the phenotypic trait z of the virus averaged over the provirus stage, $z_\circ = \sum_i p_i z_i$, or over the free-virus stage, $z_\bullet = \sum_i q_i z_i$.

The epidemiological dynamics of the total density of the virus (either in the provirus stage or in the free-virus stage) is thus:

$$\begin{aligned}
\dot{S} &= (rS + \rho_\circ(1-\delta)I)(1-N/K) - abSV - mS \\
\dot{I} &= \rho_\circ \delta I(1-N/K) + ab\phi_\bullet SV - (\alpha_\circ + m)I \\
\dot{V} &= ab(1-\phi_\bullet)VSB + \alpha_\circ BI - mV - aNV
\end{aligned} \qquad \text{(A2)}$$

The condition for a resident virus (with phenotypic traits α and ϕ) to generate an epidemic can be derived from the calculation of the basic reproductive ratio using the next-generation-matrix method [46]. The parasite life-cycle can be decomposed into fecundity (matrix F) and mortality (matrix V) components:

$$F = \begin{pmatrix} \rho\delta\lambda & ab\hat{S}\phi \\ \alpha B & ab(1-\phi)\hat{S}B \end{pmatrix}$$
$$V = \begin{pmatrix} \alpha + m & 0 \\ 0 & m + a\hat{S} \end{pmatrix}$$

where $\lambda = (1 - N/K)$ and $\hat{S} = K(r - m)/r$ refers to the density of susceptible bacteria before the introduction of the virus in the chemostat. The matrix F gives the rates at which new individuals appear in the provirus or in the free virus stages. The matrix V gives the rate at which these individuals die. The basic reproduction ratio is the spectral radius of the matrix $F\,V^{-1}$ which is:

$$R_0 = \frac{A+\sqrt{A^2-4abBS(m+aS)(m+\alpha)(\delta\lambda\rho(1-\phi)-\alpha\phi)}}{2(m+aS)(m+\alpha)} \quad \text{(A3)}$$

with $A = m\delta\lambda\rho + aS(\delta\lambda\rho + bB(m+\alpha)(1-\phi))$. The above expression can be readily used to find the parameter values allowing the virus to generate an epidemic in the chemostat (i.e. when $R_0 > 1$).

S1.2 Evolution

To better understand the evolution of the virus we focus next on the dynamics of the frequency of strain i in both the provirus (p_i) and the free-virus (q_i) compartments.

Using $\dot{p}_i = \frac{\dot{I}_i}{I} - p_i\frac{\dot{I}}{I}$, (A1) and (A2) we obtain:

$$\dot{p}_i = p_i\left(\overbrace{(\rho_i - \rho_\circ)\delta(1-N/K)}^{growth} - \overbrace{(\alpha_i - \alpha_\circ)}^{lysis} + \overbrace{abS\frac{V}{I}(\phi_i - \phi_\bullet)}^{genome\ integration} + \overbrace{abV(\phi_{\circ i} - \phi_{i\bullet})}^{superinfections}\right) + \underbrace{abV(q_i - p_i)\left(\frac{S}{I}\phi_i + \phi_{\circ i}\right)}_{\text{gene flow } between\ free\ virus\ and\ pro-virus\ compartment}$$

Similarly, using $\dot{q}_i = \frac{\dot{V}_i}{V} - q_i\frac{\dot{V}}{V}$, (A1) and (A2)we obtain:

$$\dot{q}_i = \overbrace{abSBq_i(\phi_\bullet - \phi_i)}^{failed\ genome\ integration} + B\frac{I}{V}\left(\overbrace{q_i(\alpha_i - \alpha_\circ)}^{lysis} + \underbrace{\alpha_i(p_i - q_i)}_{\text{gene flow } between\ pro-virus\ compartment\ and\ free\ virus}\right)$$

In the main text we consider a simpler scenario where we assume that only two strains are in competition (the avirulent wildtype and the virulent mutant), that infection does not affect the growth rate of the bacteria (i.e. $\rho_i = r$) and that superinfection is not possible (i.e. $\phi_{ij} = 0$). This yields equations (1) and (2). Note that in this model the only difference between the two virus strains occurs in the rate of genome integration and in the rate of lysis, and this is consistent with the properties of the mutant we are using (see **Figure S3** and [22]).

S1.3 Simulations

To generate specific predictions on the epidemiology and evolution of our system we simulated our model using parameter values given in **Table S1** below. Those parameters were chosen to match measures obtained in previous studies as well as our own measurements (see **Figure S3**). We explored the robustness of our theoretical predictions by allowing some variation on all the parameters affected by the mutation (i.e. the virulence phenotype): ϕ_R, ϕ_M, α_R, α_M. To do so we performed 10000 simulation runs, and for each run the values of these four parameters were drawn independently from a normal distribution with a mean and variance given in **Table S1**. In **Figures 2, S1 and S2** we plot all representations of these simulation runs and their median.

We further explored the potential effects of the evolution of bacterial resistance and virulence compensation in the virus using a modified version of the above model. In this new model we assumed that upon reproduction, susceptible bacteria could mutate with probability μ to a new type of bacteria, R, fully resistant to infection by the virus. Because resistance to λ requires the loss of a receptor, we further assumed that the resistance could induce a cost c on fecundity. In addition, we considered that the virus could mutate back and forth between the virulent and the avirulent phenotype. Our experimental method tracks the change in frequency of the fluorescent marker, and not the phenotype. These mutations would thus break the linkage between the marker and the virulence phenotype. To explore the effect of these mutations we allowed the virulence phenotype to change from i to j with probability ε_{ij}, but the tag always remains the same. This yields the following system of equations:

$$\dot{S} = (r(1-\mu)S + \rho_\circ(1-\delta)I)(1-N/K) - abSV - mS$$
$$\dot{R} = (r\mu S + \rho(1-c)R)(1-N/K) - mR$$
$$\dot{I}_{ii} = \big((1-\varepsilon_{ii})\rho_i I_{ii} + \varepsilon_{ji}\rho_j I_{ji}\big)\delta(1-N/K) + ab\phi_i V_{ii}S - (\alpha_i + m)I_{ii}$$
$$\dot{I}_{ji} = \big(\big(1-\varepsilon_{ji}\big)\rho_j I_{ji} + \varepsilon_{ij}\rho_i I_{ii}\big)\delta(1-N/K) + ab\phi_j V_{ji}S - \big(\alpha_j + m\big)I_{ji}$$
$$\dot{V}_{ii} = ab(1-\phi_i)V_{ii}SB + \alpha_i B I_{ii} - mV_{ii} - a(S+I)V_{ii}$$
$$\dot{V}_{ji} = ab\big(1-\phi_j\big)V_{ji}SB + \alpha_j B I_{ji} - mV_{ji} - a(S+I)V_{ji}$$

where I_{ij} refer to the density of bacteria infected by the virus with phenotype i and fluorescent tag j. Similarily V_{ij} refers to the density of free virus with phenotype i and fluorescent tag j. In addition we assume $= \sum_{i,j} I_{ij}$, $N = S + I + R$, $V = \sum_{i,j} V_{ij}$, $z_\circ = \sum_{i,j} \frac{I_{ij}}{I} z_i$ and $z_\bullet = \sum_{i,j} \frac{V_{ij}}{V} z_i$. Because we only considered 2 phenotypes (wildtype and virulent mutant) and 2 tags (the 2 fluorescent markers), this yields a system of 9 ordinary differential equations in total.

The epidemiological dynamics of the total density of the virus (either in the pro-virus stage or in the free-virus stage) is thus:

$$\dot{S} = (r(1-\mu)S + \rho_\circ(1-\delta)I)(1-N/K) - abSV - mS$$
$$\dot{R} = (r\mu S + \rho(1-c)R)(1-N/K) - mR$$
$$\dot{I} = \rho_\circ\delta I(1-N/K) + ab\phi_\bullet SV - (\alpha_\circ + m)I$$
$$\dot{V} = ab(1-\phi_\bullet)VSB + \alpha_\circ BI - mV - a(S+I)V$$

As above, we simulated the model to generate specific predictions on epidemiology and evolution when virulence compensation was possible (**Figure S1**) and when host resistance

was allowed (**Figure S2**). In both cases we show that, although these mutations can affect the medium to long-term dynamics (after 24h), the short-term predictions discussed in the main text still hold.

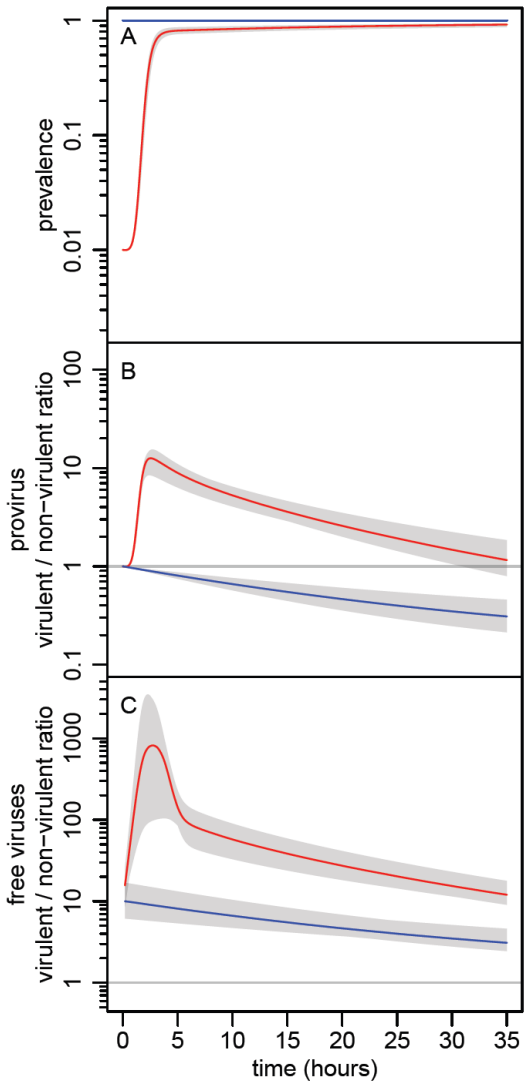

Figure S1. Theoretical evolutionary epidemiology (analogous to Figure 2) for a modified model which allows for virulence compensation. The virus mutation probability on virulence is $\varepsilon_{ij} = \varepsilon_{ji} = 0.01$. See Table S1 for other parameter values. (Red and blue line: 1% and 10% initial prevalence. Gray envelopes show the range of variation among the 10000 simulation runs and colored lines show their median)

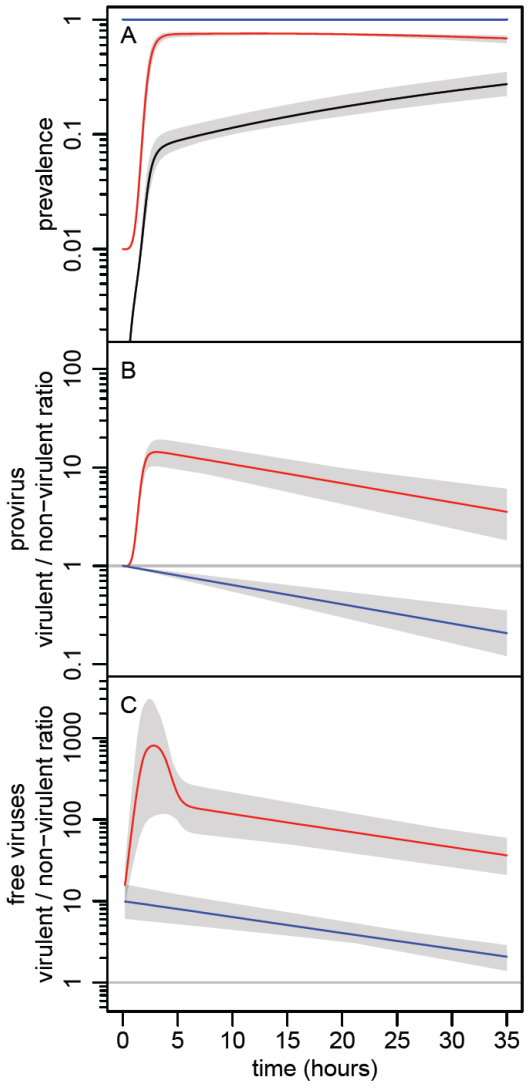

Figure S2. Theoretical evolutionary epidemiology (analogous to Figure 2) for a modified model which allows for mutation towards host resistance. The host mutation probability towards resistance is $\mu = 0.01$, and the cost of resistance is assumed to be $c = 0.01$. See Table S1 for other parameter values. (Legend similar to **S1**, except black line: frequency of resistant cells)

Table S1. Model parameters with their definitions, values and units. The virulence phenotypes (ϕ and α) were sampled in a normal distribution $\mathrm{N}(\bar{z}, \sigma_z)$ with mean $\bar{z}$ and standard deviation σ_z in **Figure 2** and **Figures S1, S2**.

Parameter	Definition	Value	Unit	Experimental estimate	References (main text)
r	growth rate of uninfected cells	1.4	h^{-1}	0.21-1.26	47
ρ	growth rate of infected cells	1.4	h^{-1}	$\rho \sim \mathrm{r}$	48
K	carrying capacity	10^9	$cell$	1-$2 \cdot 10^9$	49
δ	fidelity of vertical transmission	1	-	~ 1	50
B	burst size	200	$virus \cdot cell^{-1}$	9.7-255	51
m	dilution rate	0.75	h^{-1}	0.8	our study
a	adsorption constant	10^{-8}	$h^{-1} \cdot cell^{-1}$	7.9-$59 \cdot 10^{-8}$	52
b	probability of fusion after adsorption	10^{-2}	-		
ϕ_R	probability of genome integration of λ	$\mathrm{N}(0.3, 0.03)$	-	0.6	our study
ϕ_M	probability of genome integration of λcI857	$\mathrm{N}(9 \cdot 10^{-2}, 9 \cdot 10^{-3})$	-	0.09	our study
α_R	rate of reactivation of the λ provirus	$\mathrm{N}(5 \cdot 10^{-3}, 5 \cdot 10^{-4})$	h^{-1}	10^{-4}-10^{-3}	53,54, our study
α_M	rate of reactivation of the λcI857 provirus	$\mathrm{N}(5 \cdot 10^{-2}, 5 \cdot 10^{-3})$	h^{-1}	$>10^{-3}$	54 our study
μ	probability of mutation rate towards resistance	10^{-2}	-	$2.5 \cdot 10^{-3}$	55
c	cost of resistance	10^{-2}	-	$4 \cdot 10^{-3}$-10^{-2}	56
$\varepsilon_{ij} = \varepsilon_{ji}$	probability of mutation that compensate virulence of λcI857	10^{-2}	-	0.0038	55

S2. Experimental evolutionary epidemiology

S2.1 Methods

S2.1.1 Life-history of fluorescently marked viral strains

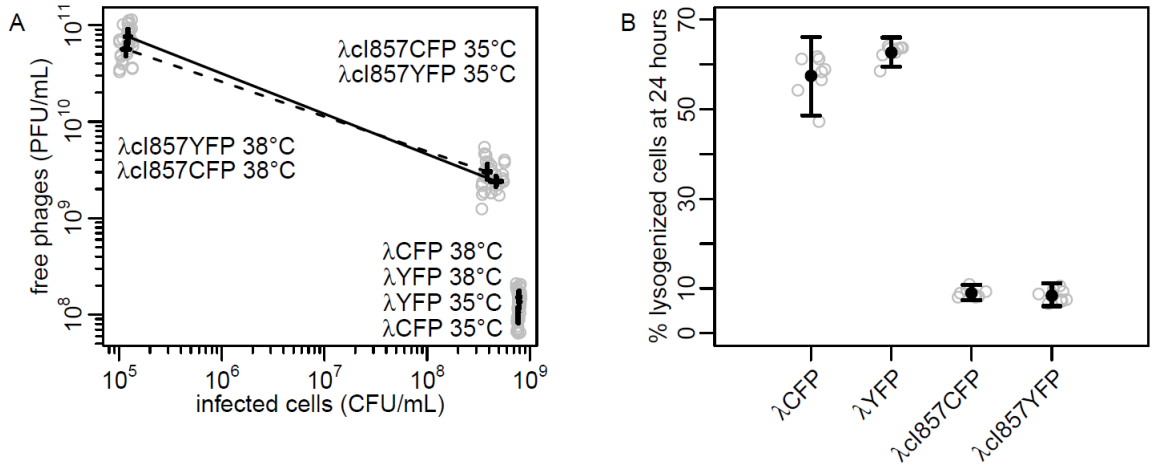

Figure S3. (A) Horizontal (free phages PFU/mL) and vertical transmission (infected cells CFU/mL) and (B) genome integration rate (% lysogenized cells at 24h) of the constructs λcI857CFP, λcI857YFP and λCFP, λYFP. (A) At 35°C, λcI857CFP and λcI857YFP show significantly higher horizontal transmission and reduced vertical transmission compared to the wildtype constructs λCFP and λYFP. At 38°C, horizontal transmission of λcI857CFP and λcI857YFP is further increased and vertical transmission is further reduced. **(B)** Lysogenization at 35°C is about 6 fold higher for λCFP, λYFP than for the mutants λcI857CFP and λcI857YFP ((A) Crosses, 95% CI. (B) Bars, 95% CI. Gray circles, raw data, see text of S2.1.1 for statistical analysis).

By ANOVA, we statistically tested the contribution of the factors *Strain* (λ or λcI857), *Temperature* (35 and 38°C) and *Color* (CFP or YFP) to the life-history traits *PFU*, *CFU* and *Lysogenized*. ANOVA on *PFU* (**Table S2.1**), *CFU* (**Table S2.2**) and *Lysogenized* (**Table S2.3**) revealed that *Strain*, *Temperature* and *Color* significantly affect all three life-history traits (except that lysogenization was only assayed at 35°C and its temperature dependence could not be determined). Even though the effect of *Color* is significant, the magnitude of its effect is several orders of magnitude lower than the effect of *Temperature* and/or *Strain*, as is visible by the percentage of sum of squares explained by each covariate (see **Table S2.1, S2.2 and S2.3, but also Figure S3**). Nevertheless, we experimentally controlled for the potential effect of color by carrying out each competition in 2 marker/virulence combinations (λCFP vs λcI857YFP and λYFP vs λcI857CFP).

Table S2.1. ANOVA for virus production (PFU/mL, ":" refers to interaction)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% of total Sum Sq
Strain	1	4.1770e+22	4.1770e+22	348.7900	< 2.2e-16	31.9
Temperature	1	3.5768e+22	3.5768e+22	298.6682	< 2.2e-16	27.3
Color	1	8.4085e+20	8.4085e+20	7.0213	0.009140	0.6
Color:Strain	1	8.4107e+20	8.4107e+20	7.0231	0.009131	0.6
Color:Temp	1	9.7030e+20	9.7030e+20	8.1022	0.005202	0.7
Strain:Temp	1	3.5662e+22	3.5662e+22	297.7838	< 2.2e-16	27.2
Color:Strain:Temp	1	9.6471e+20	9.6471e+20	8.0555	0.005329	0.7
Residuals	120	1.4371e+22	1.1976e+20			11.0

Table S2.2. ANOVA for vertical transmission (CFU/mL, ":" refers to interaction)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% of total Sum Sq
Strain	1	5.0789e+18	5.0789e+18	4041.1538	< 2e-16	76.2
Temperature	1	6.8333e+17	6.8333e+17	543.7086	< 2e-16	10.3
Color	1	6.7864e+15	6.7864e+15	5.3998	0.02379	0.1
Strain:Color	1	8.9076e+15	8.9076e+15	7.0876	0.01011	0.1
Strain:Temp	1	8.0037e+17	8.0037e+17	636.8400	< 2e-16	12.0
Color:Temp	1	8.5785e+15	8.5785e+15	6.8257	0.01152	0.1
Strain:Color:Temp	1	7.0762e+15	7.0762e+15	5.6304	0.02111	0.1
Residuals	56	7.0380e+16	1.2568e+15			1.1

Table S2.3. ANOVA for lysogenization (% lysogenized, ":" refers to interaction)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% of total Sum Sq
Strain	1	21135.2	21135.2	2860.9730	< 2e-16	98.5
Color	1	44.7	44.7	6.0538	0.02031	0.2
Strain:Color	1	68.5	68.5	9.2691	0.00503	0.3
Residuals	28	206.8	7.4			1.0

S2.1.2 Quantifying competition in the free virus stage by marker specific qPCR

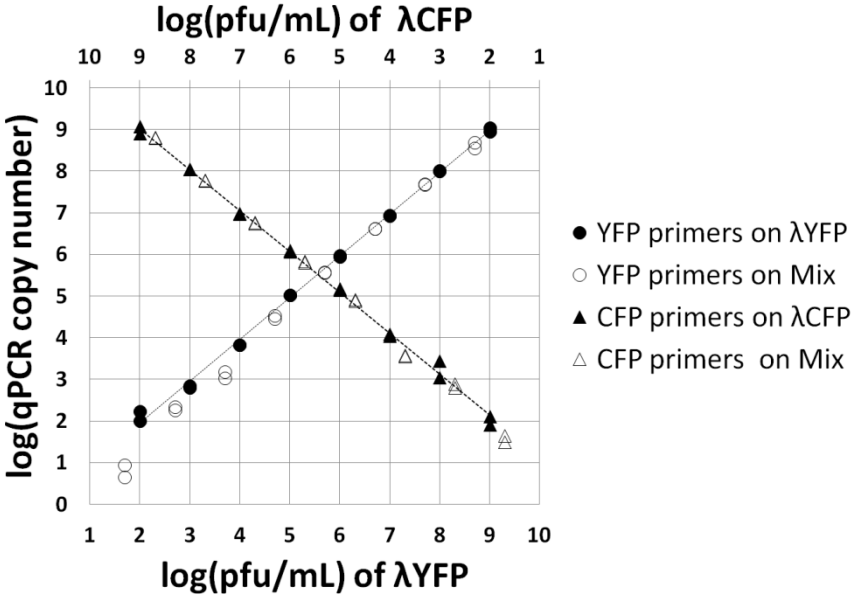

Figure S4. Test for cross-specificity of CFP and YFP specific qPCR primers. Primers at 1µM concentration were tested on dilution series of pure λCFP and λYFP lysates (10^9 to 10^2 pfu in 10-fold steps) as well as a reciprocal mixture of the dilutions series ($5x10^8$: $5x10^1$ to $5x10^1$: $5x10^8$ pfu/mL of λCFP : λYFP). qPCR on the reciprocal lysate mixtures shows no non-specific quantification even with a 10^7 fold excess of the non-specific template (see **Table S3** for primers).

Table S3. CFP and YFP specific primers (CFP and YFP specific nucleotides are in boldface).

Primer name	Specificity (Plasmid of origin)	Sequence
FCFP275	CFP (pDH3)	5'-ACAAAAGAATGGAATCAAAG**C**T**C**A**T**-3'
RCFP390	CFP (pDH3)	5'-CGAAAGGGCAGATTG**TGT**-3'
FYFP275	YFP (pDH5)	5'-ACAAAAGAATGGAATCAAAG**T**T**A**A**C**-3'
RYFP390	YFP (pDH5)	5'-CGAAAGGGCAGATTG**ATA**-3'

S2.2 First chemostat experiment

S2.2.1 Effects of initial prevalence and marker color on competition

To statistically test the effect of the *Initial Prevalence* treatment and *Color* on the competition dynamics, we performed an ANOVA on the data in **Figure 4B,C**. In order to account for repeated measurements we treated time as a random effect. Results show that the *Initial Prevalence* treatment significantly affects competitive dynamics in the provirus (**Table S4.1**) and in the free virus stage (**Table S4.2**). The effect of *Color* is, however, marginally significant (p=0.07) only in the provirus stage. More important, the magnitude of the *Color* effect is 50 times lower than that of the *Initial Prevalence* treatment, as is visible by the percentage of sum of squares explained by each covariate (see **Table S4.1 and S4.2**). Based on this result, we pooled the data from 2 marker/virulence combinations (λCFP vs λcI857YFP and λYFP vs λcI857CFP) in the first experiment (see **Figure 3**) and in the second experiment (see **Figure S6 and Figure 4)**.

Table S4.1. ANOVA table for the effect of initial prevalence and color on competition in the provirus stage with time treated as random effect (Data of Figure 2B, ":" refers to interaction)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% of total Sum Sq
Initial Prevalence	1	17.0562	17.0562	186.9355	< 2e-16	43.55
Color	1	0.3013	0.3013	3.3027	0.07042	0.77
Initial Prevalence:Color	1	0.0004	0.0004	0.0045	0.94644	0.01
Residuals	239	21.8066	0.0912			55.67

Table S4.2. ANOVA table for the effect of initial prevalence and color on competition in the free virus stage (Data of Figure 2C, ":" refers to interaction)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	% of total Sum Sq
Initial Prevalence	1	58.817	58.817	251.6524	< 2.2e-16	34.7
Color	1	0.270	0.270	1.1569	0.282666	0.2
Initial Prevalence:Color	1	2.255	2.255	9.6486	0.002011	1.3
Residuals	463	108.214	0.234			63.8

S2.2.2 Test for the occurrence of mutations that compensate virulence

In the late phase of our experimental epidemics the virulent λcI857YFP and λcI857YFP could have accumulated mutations that compensate the cI857 mutation to reduce the cost of virulence. To test for the occurrence of such compensatory mutations, we calculated the number of free virus particles that are produced per infected cell (viruses/cell). Due to its virulence, the λcI857 mutant is expected to produce more viruses/cell than the λ wildtype and, hence, the ratio $(\text{virus/cell})_{\text{mutant}}$ divided by $(\text{virus/cell})_{\text{wildtype}}$ should be larger than 1. Indeed, this ratio is above 1 throughout most of the experiment (see **Figure S5**). We can therefore conclude that the virulent λcI857CFP and λcI857YFP remained significantly more virulent than λCFP and λYFP throughout most of the experiment even if compensatory mutations might have occurred.

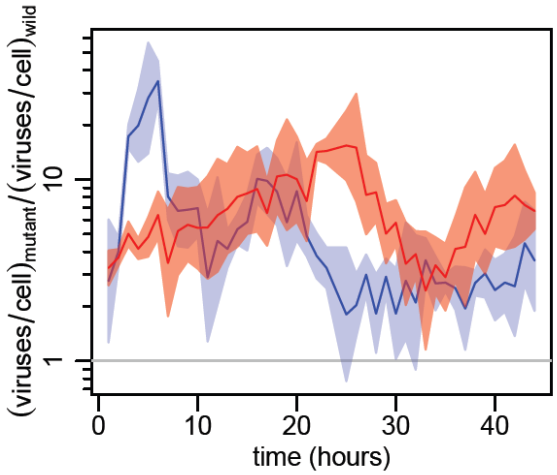

Figure S5: Test for the occurrence of mutations that compensate virulence. As a proxy for virulence we calculated viruses produced per infected cell (viruses/cell) for the λcI857 mutant strains and the wildtype strains. The λcI857 produces more viruses/cell and is more virulent than the wildtype as long as the ratio $(\text{viruses/cell})_{\text{mutant}}$ devided by $(\text{viruses/cell})_{\text{wildtype}}$ is larger than 1. Indeed this ratio is significantly larger than 1 throughout the experiment (except t=25 and 33h in the 1% treatment). Hence, the λcI857 mutants have remained more virulent than the wildtype even if compensatory mutations might have occurred. (blue area: 1% initial prevalence, red area: 100% initial prevalence, solid line and shading: Mean and 95% CI envelope of the log transformed data from 4 chemostats).

S2.3 Second chemostat experiment

S2.3.1 Competition at initial prevalence 1%, 10% and 99%

To further explore the relation between the maximal benefit of virulence and initial prevalence we ran 6 additional chemostat competitions (1%, 10% and 99% initial prevalence each in 2 marker/virulence combinations). The observed competition dynamics in the pro-virus and free-virus stage suggest that the transient benefit of virulence decreases with increasing initial prevalence (see **Figure S6**). We further explore and test this possibility in **Figure 4** by extracting the maximal virulent/non-virulent ratios from the first 15h of **Figure S6** and plotting them directly against initial prevalence. By a linear model on the data of **Figure 4** we statistically tested for the effect of *Initial Prevalence* (1%, 10% and 99%) and *Viral Life-Stage* (provirus and free virus) on maximal virulence (**Table S5**). This analysis shows that maximal virulence significantly decreases between 1% to 10% and 10% to 99% initial prevalence both in the provirus and in the free virus stage. Furthermore, the maximal virulence is significantly higher in the free virus than in the provirus stage for all *Initial Prevalence* treatments and the interaction between *Viral Life-Stage* and *Initial Prevalence* is not significant ($F_{2,6} = 0.6223$, $p = 0.5680598$).

Table S5: Linear Model analysis for the effect of initial prevalence and viral life-stage on the maximal benefit of virulence (Data from Figure 4). ('Intercept' corresponds to the free virus stage at *Initial Prevalence 1%.* Contrasts on *Initial Prevalence* are chosen so that *Initial Prevalence 1%* is compared to *Initial Prevalence 10% (3rd line) and Initial Prevalence 10%* is compared to *Initial Prevalence 99%* (4th line))

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	3.5573	0.2396	14.850	4.17e-07
Viral Life-stage	-1.7281	0.2396	-7.214	9.12e-05
Initial Prevalence 10%	-1.0231	0.2934	-3.487	0.008234
Initial Prevalence 99%	-1.7295	0.2934	-5.895	0.000364

S2.3.2 Invasion of resistant host cells

In the second chemostat experiment we observed a drop in the overall prevalence of fluorescent cells after 40h in chemostats 1,2,3,4 and 6, but not in chemostat 5 (see **FigureS7**). Initially we had 3 alternative explanations for this drop in prevalence: (1) Our chemostats were infected by a bacterium other than *E. coli*, (2) non-fluorescent cells carry a prophage which spontaneously deleted the fluorescent marker, (3) non-fluorescent cells carry no prophage, but have acquired resistance to infection by λ at the lambda receptor, lamB.

To rule out explanations (1) and (2) we cross-streaked colonies from each chemostat (t=60h) against the indicator strain λKH54h80ΔcI. Strain λKH54h80ΔcI infects *E.coli* cells through the FhuA receptor and lyses cells which do not carry a prophage. All colonies from chemostats 1,2,3,4 and 6 were sensitive to the indicator strain (see **Figure S8**). This demonstrates that the invading cells are still *E. coli* and that the invading cells carry no prophage. Since the colonies from chemostat 5 carry a prophage, they are not lysed by the indicator strain. After eliminating explanations (1) and (2) it is therefore most likely that invading cells have acquired resistance at the lambda receptor, lamB. The fact that the invading resistant cells can still be lysed by the indicator strain which enters through FhuA receptor rather than the lamB receptor supports the view that the invading cells have acquired resistance in the original target of phage λ, the lamB receptor.

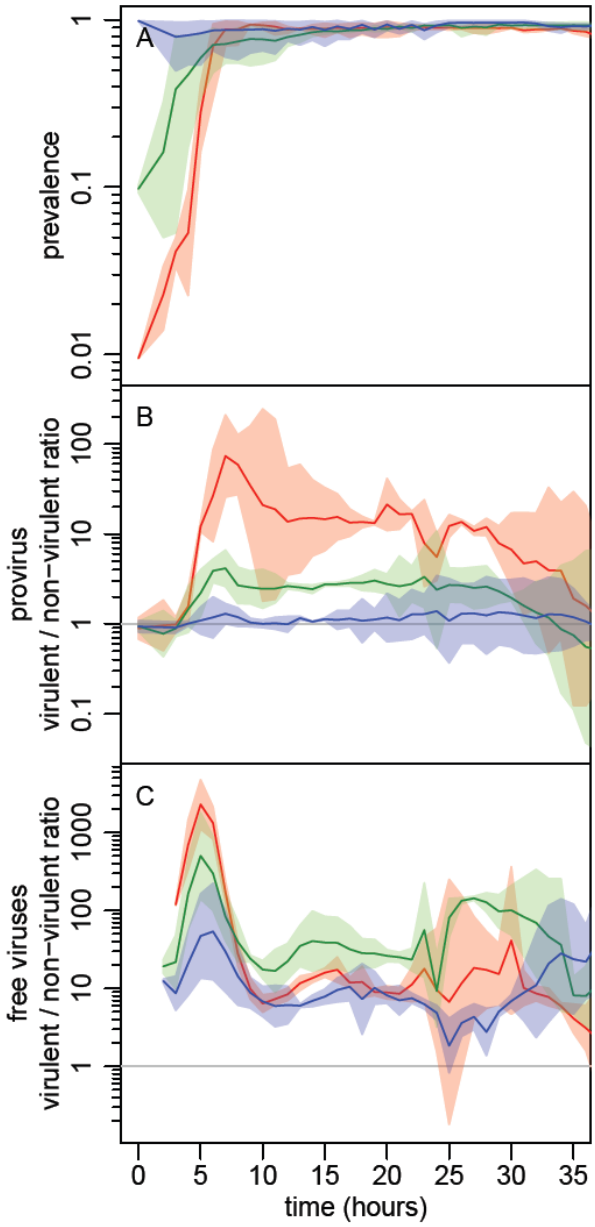

Figure S6: Competition dynamics in the second chemostat experiment with initial prevalence of 1%, 10% and 99% (A) Prevalence, (B) fitness benefit for the provirus and (C) free virus in the second chemostat experiment. The 1% initial prevalence treatment (blue) leads to the highest benefit of virulence. This benefit of virulence decreases in the 10% and 99% initial prevalence (green and red). The maxima of B and C between t=0 and t=15h were extracted to create Figure 4 in the main text (Solid lines: Mean, shading: 95% CI interval of log transformed data from 2 chemostats pooled by color replicate).

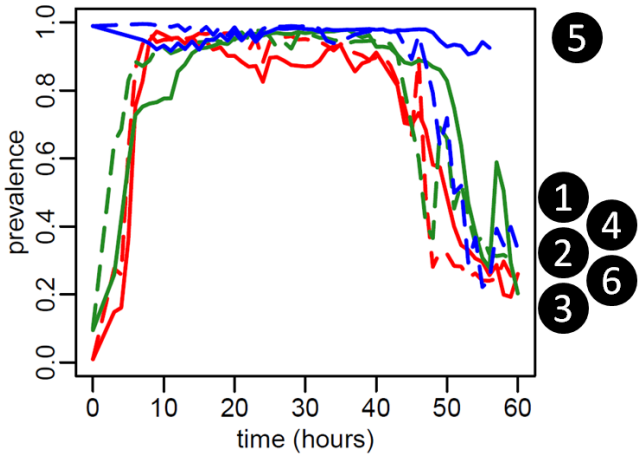

Figure S7: Invasion of resistant host cells. In the second experiment, resistant, but non-infected, host cells invaded 5 out of 6 chemostats after 40h and caused a drop in overall prevalence. Chemostat 5 (red solid line) showed no invasion of resistant cells and maintained high prevalence. (Solid lines: λCFP versus λcI857YFP and dotted lines λYFP versus λcI857CFP. Blue, green and red: 1%, 10% and 99% initial prevalence, black numbers correspond to the numbering in **Figure S8**).

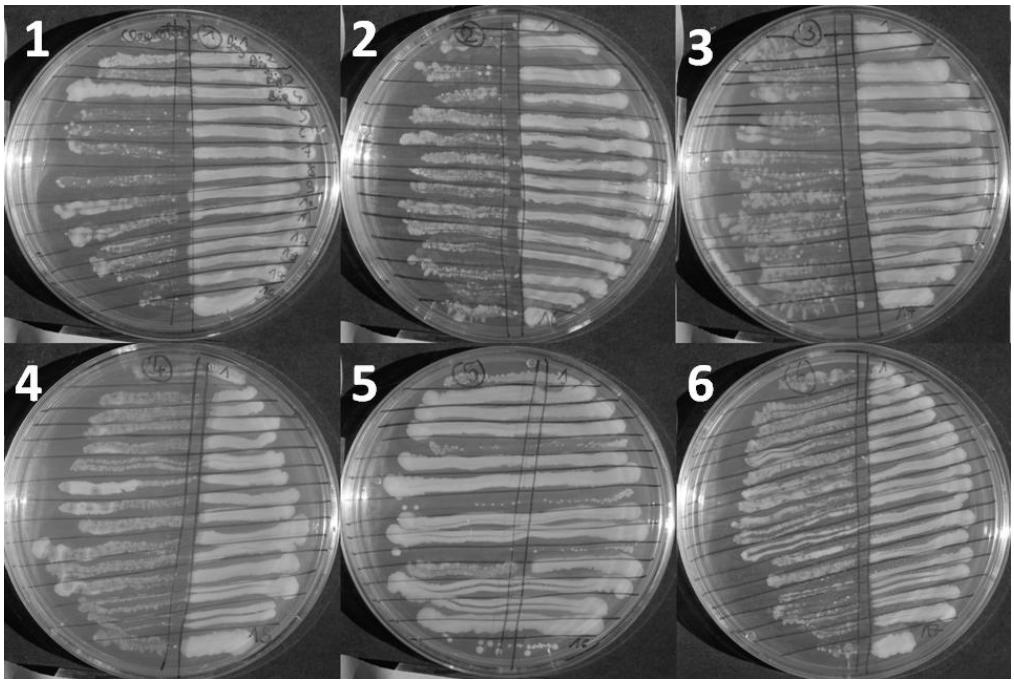

Figure S8: Test for the origin of resistance. We cross-streaked individual colonies from t=60h of each chemostat against the indicator strain λKH54h80ΔcI (this strain enters through the FhuA receptor and only lyses non-lysogens). Resistant hosts from chemostats 1,2,3,4 and 6 were sensitive to λKH54h80ΔcI and are therefore non-lysogens and most likely lamB resistance mutants. Since most colonies from chemostat 5 carry a prophage they are not lysed by the indicator strain.