Prophage Provide a Safe Haven for Adaptive Exploration in Temperate Viruses

Lindi M. Wahl¹ and Tyler Pattenden

Applied Mathematics, Western University, London, Ontario N6A 5B7, Canada

ABSTRACT Prophage sequences constitute a substantial fraction of the temperate virus gene pool. Although subject to mutational decay, prophage sequences can also be an important source of adaptive mutations for these viral populations. Here we develop a life-history model for temperate viruses, including both the virulent (lytic) and the temperate phases of the life cycle. We then examine the survival of mutations that increase fitness during the lytic phase (attachment rate, burst size), increase fitness in the temperate phase (increasing host survival), or affect transitions between the two phases (integration or induction probability). We find that beneficial mutations are much more likely to survive, ultimately, if they first occur in the prophage state. This conclusion applies even to traits that are only expressed during the lytic phase, and arises due to the substantially lower variance in the offspring distribution during the temperate cycle. This observation, however, is balanced by the fact that many more mutations can be generated during lytic replication. Overall we predict that the prophage state provides a refuge, relatively shielded from genetic drift, in which temperate viruses can explore possible adaptive steps.

KEYWORDS temperate viruses; prophage; fixation probability; adaptation; genetic drift

WHILE both temperate and virulent bacteriophage have the ability to infect and lyse bacterial host cells, producing a large number of viral particles, temperate phage has the option, instead, to integrate its genome into the host cell DNA. The viral genome is then replicated when the host cell, or *lysogen*, replicates, such that both daughter cells carry a copy of the phage sequence. The term *prophage* refers to viral DNA harbored in this way in a bacterial genome, indicating that in future, under stressful conditions for example, the prophage may be induced to produce new phage. Upon induction, the prophage will excise from the bacterial genome and set in motion the process of producing and releasing viral particles, resulting in the death of the host cell (Casjens 2003).

Although the presence of inducible virus in bacterial cells has long been recognized (Bail 1925; Bordet 1925), whole-genome sequencing revealed an unexpected abundance of prophage sequences in bacterial genomes (Canchaya *et al.* 2003; Casjens 2003). A recent analysis of >2000 sequenced

bacterial genomes revealed that about half were lysogens, encoding a total of 2863 prophage sequences (Touchon *et al.* 2016). These results imply that lysogens carry ~3 prophage on average, with some bacteria carrying as many as 15–20 (Casjens 2003; Touchon *et al.* 2016). Although not all prophage sequences correspond to intact, inducible viral genomes, it is clear that bacterial genomes encode a substantial reservoir of viral genetic material. As a back-of-the-envelope approximation, assume temperate viruses in the environment outnumber lysogens by the same ratio that viruses outnumber bacteria on a global scale, that is, 10:1 (Brüssow *et al.* 2004). If lysogens encode three prophage on average, then about one third of the temperate viral gene pool on the earth is in the prophage state at any time.

The genetic degradation of unselected prophage sequences over time is well understood; prophage are temporarily shielded from selection on their lytic function, while the host clearly stands to benefit from the inactivation of unselected prophage. Thus deletions and other deleterious mutations should ultimately accrue and impair function, as evidenced by the large number of defective prophage and degraded prophage elements observed in bacterial genomes (Canchaya *et al.* 2003; Casjens 2003). Several recent studies, however, have shed new light on the adaptive potential of prophage sequences. In a comparison of 85 sequenced enterobacterial

Copyright © 2017 by the Genetics Society of America

doi: <https://doi.org/10.1534/genetics.116.197541>

Manuscript received November 3, 2016; accepted for publication March 10, 2017; published Early Online March 15, 2017.

Supplemental material is available online at www.genetics.org/lookup/suppl/doi:10.1534/genetics.116.197541/-/DC1.

¹Corresponding author: Applied Mathematics, Applied Mathematics, MC255, Western University, London, ON N6A 5B7, Canada. E-mail: lwahl@uwo.ca

genomes, Bobay *et al.* (2014) observed a clear pattern of purifying selection on prophage genes, suggesting rapid inactivation and frequent domestication of these sequences by their hosts. Similarly, Wang *et al.* (2010) demonstrated that inactivated or “cryptic” prophage confer a range of beneficial functions to *Escherichia coli* in adverse environmental conditions. Thus inactivated prophage sequences can clearly adapt to improve host fitness. In addition, adaptive recombination events between prophage and infecting phage have been well documented, particularly in the lactococcal phages, demonstrating the role of prophage as an accessible reservoir of phage genetic material (Moineau *et al.* 1994; Bouchard and Moineau 2000; Durmaz and Klaenhammer 2000; Labrie and Moineau 2007).

Here we address a simpler, but related issue: the potential of prophage sequences to acquire mutations that benefit the phage itself. In particular, we focus attention on *de novo* beneficial mutations that might ultimately contribute adaptively to the phage lineage, if they survive loss when rare. We find that these mutations are far more likely to survive if they first occur in prophage, as opposed to during the lytic replication cycle, even if the mutation affects only the lytic function of the phage. In essence, the “low variance lifestyle” provided by binary fission within the host cell offers a refuge from genetic drift. Thus, we predict that prophage sequences represent not only a substantial fraction of the temperate viral gene pool, but a fraction that may play a disproportionately important role in functional innovation.

Model

In previous work, we have found that the survival of adaptive mutations can depend sensitively on the details of the organism’s life history (Hubbarde *et al.* 2007; Patwa and Wahl 2008; Wahl and Zhu 2015). Thus, our approach begins by developing a life-history model for temperate viruses, including both the lytic and prophage replication cycles. The model considers both free virions and lysogens, tracking which events occur and in what order (e.g., attachment, then lysis, then new virions); however, time is not modeled explicitly and this greatly simplifies the analysis. We will return to issues of timing later in this section.

Figure 1 illustrates the approach. A free virion successfully attaches to a host cell with probability A , or is lost with probability $1 - A$. After attachment, either the virus lysogenizes the host cell with probability P , or enters the lytic cycle with probability $1 - P$. The outcome of lysogeny is a single new lysogen; the outcome of lysis is a burst of B free virions.

Similarly, a lysogen can either be induced to lysis, or can grow and reproduce. The probability that the cell is induced before fission is denoted N . If not induced, the cell undergoes fission and we assume that each daughter cell dies with independent mortality probability M .

We note that the attachment probability A reflects the overall probability of attachment and successful infection of a susceptible host cell. An underlying assumption here is that

Figure 1 Diagram of the life-history model. The life-history events modeled for (A) free phage and (B) lysogens are shown (see *Model* section for details). For the plasmid prophage, the life history of free phage is analogous to (A) while the more complex life history of the prophage is shown in (C).

the population of host cells susceptible to a given phage is large and roughly constant. In addition, the model assumes that the infected host cell is not already a lysogen for the phage in question; this would be true whenever the prophage confers immunity to sister phage.

Wild-type population

To study the adaptation of temperate viruses, we would like to predict the fate of rare beneficial mutations. Thus we consider a wild-type population initially at equilibrium, that is, a wild-type phage population for which, on average, each phage has one surviving offspring. This condition that the wild-type fitness is one (growth rate is zero) can be achieved by balancing the parameter values in a number of ways.

For example, we can assume that both the lytic cycle and lysogenic cycle have one expected offspring per generation. For the lytic cycle, we require the product $AB = 1$. For the lysogenic cycle, we require each bacterium to have one offspring on average, so $M = 0.5$. In this “lysis-lysogeny neutral”

Table 1 Wild-type parameter values

Parameter	Lysogeny advantage $AB < 1$	LL neutral $AB = 1$	Lysis advantage $AB > 1$
A , attachment probability	0.008	0.010	0.012
B , burst size	100	100	100
M , cell mortality	0.450	0.500	0.550
P , probability of lysogeny	0.001	0.001	0.200
N , induction probability	0.091	0.091	0.020

(LL neutral) case, whether an infecting virion completes a lytic or lysogenic replication cycle does not affect fitness, since the expected outcome is the same in both reproductive phases.

In contrast, we can assume that lysogeny, for example, has a slight fitness advantage by setting $M < 0.5$ such that lysogens, on average, have more than one surviving offspring. This corresponds to a growing bacterial population. In this case, to have wild-type phage at equilibrium, we must assume that the lytic cycle has fitness less than one ($AB < 1$). We then set the probabilities P and N such that the wild-type temperate phage, using both phases of the life cycle, has a fitness of one. This “lysogeny advantage” case could correspond to a situation in which susceptible bacteria are rare, and thus lytic reproduction is not able to sustain the phage population.

The final case is to assume that lysis has a fitness advantage and balance this with the assumption that the population of lysogens is decreasing slightly over time. In this situation, $AB > 1$ but $M > 0.5$. This “lysis advantage” case corresponds to a situation in which susceptible bacteria are common and thus the lytic cycle is more profitable than the lysogenic cycle. In nature, a phage lineage might experience periods in each of these parameter regimes, for example, during times when susceptible hosts are relatively rare or abundant. In the work to follow we will test all three of these cases.

Mathematically, this life-history model can be described as a branching process in which a single free virion produces AP lysogens and $A(1-P)B$ free virions, whereas a single lysogen produces NB free virions and $(1-N)[2(1-M)^2 + 2M(1-M)] = 2(1-M)(1-N)$ lysogens. We can thus express the average number of individuals produced by a single wild-type individual as the dominant eigenvalue, λ , of the matrix

$$\mathbf{G} = \begin{bmatrix} A(1-P)B & NB \\ AP & 2(1-M)(1-N) \end{bmatrix}, \quad (1)$$

where $\mathbf{G}$ is the projection matrix, familiar from matrix models in population dynamics and ecology (Caswell 2001). It is straightforward to demonstrate that $\lambda = 1$ when $M = 0.5$ and $AB = 1$. For the wild-type virus in other cases, we choose values of AB , M , and P , and compute N to satisfy the condition $\lambda = 1$. Table 1 gives the parameter values that were used to generate the figures to follow. We take $B = 100$ and choose A such that AB is just less than, equal to, or greater than one.

Likewise, we take M to be just less than, equal to, or greater than 0.5. Refardt and Rainey (2010) report lysogeny probabilities in λ -phage in the range 10^{-4} – 10^{-1} , with most replicates showing $P \approx 10^{-3}$ (we use a larger value in the lysis advantage case for illustrative purposes, since beneficial mutations must reduce this value as described in the next section). The probability of induction for temperate viruses is strongly dependent on environmental conditions and could vary widely; values in Table 1 yield $\lambda = 1$.

Beneficial mutations

We wish to consider the fate of rare adaptive mutations which may first arise either in the free virus or prophage genome. Clearly, mutations that increase the attachment probability or the burst size increase the phage growth rate. If the wild-type strain has life-history parameter x , we will use $\tilde{x}$ to denote the analogous parameter in the mutant. Thus $\tilde{A} > A$ and $\tilde{B} > B$ always reflect beneficial mutations. Similarly, mutations that reduce mortality of the lysogens, $\tilde{M} < M$, are always beneficial to the prophage. The probabilities P and N , however, reflect transitions between the two phases of the temperate life cycle. As mentioned previously, changes to these traits are neutral if both phases of the life cycle have equal fitness. If the lysogenic cycle produces more offspring on average than the lytic cycle, however, increasing P and reducing N will be beneficial to the phage lineage. The reverse will be true if the lytic cycle is favored. We will examine the fate of beneficial mutations in each of these parameter regimes in isolation; the possibility of alternating selection is an avenue for future work.

To compare mutations affecting different traits, we compute the selective advantage, s , of each mutation. If the generation times in the free virus and provirus life cycles are similar, we can simply compute the dominant eigenvalue, $\tilde{\lambda}$, of matrix $\tilde{\mathbf{G}}$, which is the matrix $\mathbf{G}$ with each parameter replaced by the appropriate parameter value for the mutant lineage. The selective advantage is then defined to be $s = \tilde{\lambda} - 1$.

If the generation times in the two possible life cycles are not equal, however, a greater advantage can be accrued by improving fitness in the cycle with the shorter generation time. To accurately compute the selective advantage, we will use a continuous time analog of the branching process described above. In this case, free virus has a constant attachment rate, α , and a clearance rate, c . The average generation time of a

Figure 2 The survival probability of beneficial mutations affecting one of five life-history traits vs. the trait value. Results are shown for beneficial mutations that first occur in a single free phage (dots) or in a single prophage (lines). Wild-type parameter values are as provided for the lysogeny-advantage case in Table 1.

free virion is then $1/(\alpha + c)$ and the attachment probability is $A = \alpha/(\alpha + c)$. Similarly, we define an induction rate ν and fission rate f for lysogens, and the expected generation time for lysogens is $1/(\nu + f)$. The induction probability is $N = \nu/(\nu + f)$. As before, we let M be the probability that each offspring in a fission event independently survives and carries the prophage. These assumptions yield the following system of linear differential equations for the expected population size of free virus V and lysogens L , at time t :

$$\begin{aligned} \frac{dV}{dt} &= \alpha(1 - P)BV + \nu BL - (\alpha + c)V \\ \frac{dL}{dt} &= \alpha PV + fL \left[2(1 - M)^2 + 2M(1 - M) \right] - (\nu + f)L. \end{aligned} \quad (2)$$

This continuous-time model allows us to set different generation times in the two phases of the life history by setting parameters c and f . Parameters α and ν are then calculated to give the appropriate probabilities A and N . To compute the selective advantage of a rare mutation, we compute the largest root of the characteristic equation of system (2), replacing all parameter values by the appropriate parameters for the mutant lineage. Although this does not affect the extinction probabilities in any way, it affects which selective advantage we associate with them. Algebraically, we note that the condition that the dominant eigenvalue of matrix G is unity ($\lambda = 1$) is exactly equivalent to the condition that the largest root of the characteristic equation of system (2) is equal to zero.

Extinction probabilities

The simple branching-process model described here has the advantage that the extinction probabilities can be readily expressed in two implicit equations. We use X_V to denote the extinction probability of a mutant lineage, beginning with a single free virion, while X_L is the extinction probability of a lineage that begins with a single lysogen. We write:

$$X_V = 1 - A + A[PX_L + (1 - P)X_V^B], \quad (3)$$

since the virus lineage is lost if the original virion does not attach (with probability $1 - A$) or if the virion successfully attaches (with probability A), but the lineage thus created is lost. If the phage creates a lysogen, the lineage is lost only if the resulting lineage, starting with the single mutant lysogen, is lost; this occurs by definition with probability X_L . Similarly, we can write

$$X_L = NX_V^B + (1 - N) \left[M^2 + 2M(1 - M)X_L + (1 - M)^2X_L^2 \right] \quad (4)$$

to describe the extinction of a lineage that begins as prophage, yielding two equations in two unknowns. For the wild type, $X_V = X_L = 1$ is the only solution. For beneficial mutations, once again we replace the wild-type parameters in these equations with the appropriate parameters for the mutant

lineage, and find solutions with both X_V and X_L between zero and one. In the results to follow, we plot survival probabilities defined as $\Pi_V = 1 - X_V$ and $\Pi_L = 1 - X_L$, thus illustrating the probability that *de novo* adaptive mutations survive drift when rare.

Data availability

The authors state that all data necessary for confirming the conclusions presented in the article are represented fully within the article.

Results

Figure 2 plots the survival probability vs. the trait value for each of the five traits. For each trait, we note the striking difference between the fate of beneficial mutations that first occur in the prophage state (solid lines) and those that first occur as free phage (dots). While this figure shows results for the lysogeny-advantage parameter regime, these results were repeated in the other regimes (see Supplemental Material, Figures S1 and S2 in [File S1](#)); survival is always much more likely for mutant lineages that begin as prophage.

We also note the inflection points in the survival probabilities for mutations that increase attachment or burst size. These inflections occur at the points $\bar{A}B = 1$ and $A\bar{B} = 1$, respectively, reflecting the fact that increasing either attachment or burst size has little effect on mutant survival, provided the mutant lineage remains in the lysogeny-advantage regime. For larger magnitude changes (changes that exceed the $AB = 1$ threshold), substantial survival probabilities emerge. A similar inflection is seen for mutations that reduce cell mortality in the lysis-advantage regime (Figure S2 in [File S1](#)).

Figure 3 shows the same results, plotted against the selective advantage, s , of each mutation, and thus allowing for comparison across traits. In this figure, we have computed s from the matrix eigenvalue λ , assuming that the generation times in the two phases of the life cycle are similar. Noting the change in the y-axis scale between the two panels, we see once again that mutations first occurring in prophage are roughly two orders of magnitude more likely to survive. For mutations that first occur as free phage, attachment mutations are the most likely to survive, followed by burst size mutations, reflecting the importance of these steps in the life history of free phage. For mutations that first occur as prophage, mutant lineages that reduce host mortality are the most likely to survive, reflecting the importance of sustaining several generations of host cell division while building the mutant lineage. Again, these results are nearly identical for the LL-neutral and lysis-advantage regimes (Figures S3 and S4 in [File S1](#)), although some differences for free phage mutations are observed. Overall, we infer that for prophage, the reduced variance in the offspring distribution substantially shields new lineages from genetic drift.

We can relax the assumption that the generation times are similar, and compare cases in which the free virus has either a shorter or longer generation time than the lysogen. In both

Figure 3 The survival probability of beneficial mutations affecting one of five life-history traits vs. the selective advantage. Results are shown for beneficial mutations that first occur in a single free phage (top panel) or in a single prophage (bottom panel; note change of scale). Wild-type parameter values are as provided for the lysogeny-advantage case in Table 1.

cases the results described above hold, but mutations affecting one life cycle or the other have an enhanced selective advantage, as shown in Figure S5 in [File S1](#).

The survival advantage for prophage mutations may be mitigated when considering the mechanisms underlying novel mutations. For novel mutations occurring in prophage, a new mutation can first occur in one, or at most two, sister cells. In contrast, novel mutations in free phage could first occur as a burst of identical sister phage, if for example a mutation occurs in the “template” phage genome before transcription. In Figure 4, the results shown in Figure 2 are reexamined, assuming the best-case scenario of a mutant lineage initiated by a burst of $B = 100$ sister phage carrying the mutation. In this case, burst size and attachment mutations in free phage are roughly as likely to survive as their counterparts in prophage; the advantage of an initial burst in the lineage balances the negative effects of drift. However, for mutations in the other three life-history traits, prophage lineages remain much more likely to survive.

We note that the eigenvector associated with $\lambda = 1$ in matrix $\mathbf{G}$ gives the stable stage distribution, that is, the

Figure 4 The survival probability of beneficial mutations that first occur in a single prophage (lines) or in an initial burst of B free phage (dots) vs. the trait value. Wild-type parameter values are as provided for the lysogeny-advantage case in Table 1.

proportion of viral genomes in either the free virus or prophage state at equilibrium, assuming the two generation times are similar. For the wild-type parameter values in Table 1, the stable stage distributions are all strongly biased toward free phage (lysogeny advantage, 2.2 lysogens per 100 free phage; lysis advantage, 2.0 lysogens per 100 free phage). This is reasonable since the vast majority of the free phage will fail to attach (A is of order $1/100$), and cannot contribute to a surviving lineage. To make a more meaningful comparison, we should consider the equilibrium ratio of lysogens that will not die [proportion of lysogens times $(1 - M)$] to free phage that will successfully attach (proportion of free phage times A). In this case the equilibrium ratios are more moderate, and are consistent with the advantages accrued in each life cycle (lysogeny advantage, 1.5 lysogens per free phage; lytic advantage, 0.75 lysogens per free phage).

Comparison with lytic phage

A further consideration is how the fate of these mutations in temperate phage might compare with similar mutations in an obligately lytic bacteriophage. To compare these cases, we consider a temperate phage in the lysogeny-advantage parameter regime (Table 1), such that the temperate phage has no advantage *a priori* during the lytic cycle of replication. As a comparison, we consider an obligately lytic phage with burst size $B = 100$ and attachment probability $A = 1/100$, such that both wild-type phage have the same burst size and fitness. We then compare the fate of mutations increasing either B or A for lineages first occurring in (a) free temperate phage, (b) prophage, or (c) the obligately lytic phage. Not surprisingly, mutations first occurring in the lytic phage fare similarly to mutations that first occur in the lytic cycle of the temperate phage, as shown in the left-hand panels of Figure 5. In the panels on the right of Figure 5, we plot the same results vs. the selective advantage, s . In the top panel, we see that for both traits, mutations first occurring in the lytic phage or the lytic cycle of the temperate phage fare similarly. In contrast, if we compare mutations first occurring in prophage to the same mutations occurring in the lytic phage, the prophage mutations are far more likely to survive drift when rare.

The magnitude of this effect is illustrated in Figure 6, in which the ratio of survival probabilities is plotted against the trait value. The same mutation is 50–100 times more likely to survive drift if it first occurs in prophage, as compared to in the genome of an obligately lytic phage. However, all of these results must be tempered by the underlying rates at which new mutations may be generated, which may differ between these life histories. We consider this idea further in the Discussion.

Plasmid prophage

In developing this model, we have assumed that the host cell passes exactly one copy of the prophage to each daughter, thus linking the fate of the prophage with the host cell lineage. While this mode of replication is typical of most prophage in

Figure 5 The survival probability of beneficial mutations affecting attachment probability or burst size vs. trait value (left panels) or selective advantage (right panels). Results are shown for beneficial mutations that first occur in a single free temperate phase (symbols), in a single prophage (solid lines), or in a single obligately lytic phase (dashes in left-hand panels, black symbols on right). Wild-type parameter values for the temperate phase are as provided for the lysogeny-advantage case in Table 1; for the lytic phase, $A = 0.01$ and $B = 100$.

chromosomal DNA, prophages that have been incorporated into plasmid sequences may be gained or lost in more complex processes.

Typical plasmid prophages are moderately large genetic elements and must be kept at low copy numbers to reduce metabolic burden (Casjens *et al.* 2004; Łobocka *et al.* 2004). Due to their low copy number, during partition, one or both daughter cells may inherit a single copy of the prophage. In addition, plasmid prophage employ a number of other strategies to ensure transmission (Sengupta and Austin 2011). For example, cells carrying plasmid prophage instigate higher rates of bacterial conjugation (Sengupta and Austin 2011). In addition, some prophage such as *P1* employ plasmid addiction, such that any daughter cell without a prophage is rendered inviable due to growth inhibition or postsegregation killing (Lehnherr *et al.* 1993; Sengupta and Austin 2011).

To address the special case of plasmid prophage, we modified the model to include key aspects of the plasmid prophage life cycle. Specifically, cells containing a plasmid prophage may undergo either induction into a lytic cycle, conjugation, cell fission, or cell death; the probabilities that each of these events is the first to occur are given by N , J , F , and M , respectively, with the condition $N + J + F + M = 1$. We use copy fidelity, C , to quantify the probability that each daughter cell after fission carries a functional copy of the plasmid prophage. Thus the plasmid prophage may be carried by zero, one, or two daughters with probabilities $(1-C)^2$, $2C(1-C)$, and C^2 , respectively. When $C = 1$, both daughter cells carry the plasmid prophage; whereas when $C = 0$, neither daughter is a carrier. The life-cycle model for the plasmid prophage is illustrated in Figure 1.

Note that daughter cells that contain zero copies of the plasmid may be short lived due to postsegregational killing

mechanisms (Lehnherr *et al.* 1993), however this is immaterial since such cells do not contribute to the prophage lineage. We do not distinguish between daughters carrying one or more copies of the plasmid prophage since more than one copy is unlikely, as described above.

Solving the following two equations in place of Equations 3 and 4, we obtain the extinction probability for beneficial mutations affecting any of these life-history traits:

$$X_{v,pl} = 1 - A + A \left[PX_{l,pl} + (1 - P)X_{v,pl}^B \right],$$

$$X_{l,pl} = NX_{v,pl}^B + M + JX_{l,pl}^2 + F \left(C^2 X_{l,pl}^2 + 2C(1 - C)X_{l,pl} + (1 - C)^2 \right).$$

Finally, the selective advantage of mutations in plasmid prophage is computed using the dominant eigenvalue of the matrix $\mathbf{G}_{pl}$ (see Equation 1 for comparison):

$$\mathbf{G}_{pl} = \begin{bmatrix} A(1 - P)B & NB \\ AP & 2J + F[2C(1 - C) + 2C^2] \end{bmatrix}.$$

We point out that the matrix $\mathbf{G}_{pl}$ is very similar to matrix $\mathbf{G}$, since the probability that a daughter cell dies (M in the previous model) is equivalent, from the point of view of the plasmid, to a daughter cell that does not carry a plasmid copy ($1 - C$ in $\mathbf{G}_{pl}$). The only substantive difference is the addition of the $2J$ term in $\mathbf{G}_{pl}$, allowing an additional mechanism for plasmid prophage to reproduce.

Figure 7 illustrates results analogous to Figure 3, in which the survival probability is plotted against the selective advantage, s , for changes to various life-history traits. As in the previous case, we note the change in the y-axis scale between the two panels and once again see that mutations first occurring in prophage are far more likely to survive. For mutations that occur first in free phage, attachment mutations are most

Figure 6 The survival probability for beneficial mutations that first occur in prophage, divided by the survival probability for the same mutation first occurring in lytic phase vs. the trait value. Wild-type parameter values for the temperate phage are as provided for the lysogeny-advantage case in Table 1; for the lytic phage, $A = 0.01$ and $B = 100$.

likely to survive, followed by burst size mutations, echoing previous results for chromosomal prophage. For mutations that first occur in prophage, a mutation that reduces host mortality is most beneficial. Mutations that increase the probability of conjugation are also very likely to survive, which presumably follows from the fact that conjugation, unlike fission, always creates a new copy of the plasmid prophage in our model.

The fate of mutations affecting fission rates depends on copy fidelity, C , in the wild-type. If $C < 0.5$, as illustrated in Figure 7, plasmids are often lost during fission and the expected number of copies of the plasmid after fission is < 1 ; thus reducing F is beneficial. The reverse is true if $C > 0.5$. In the latter case, mutations affecting F have a high survival probability, as illustrated in Figure S6 in File S1. These results are nearly identical in the LL-neutral and lysis-advantage regimes (Figures S7 and S8 in File S1).

Discussion

Prophage sequences constitute one of the primary sources of genetic diversity in bacterial genomes (Brüssow *et al.* 2004; Canchaya *et al.* 2004), contributing not only to gene sequence diversity but also promoting large-scale genome rearrangements (Canchaya *et al.* 2004; Casjens 2005) and horizontal gene transfer in the host (Brüssow *et al.* 2004). Likewise, prophage contribute to viral diversity through recombination events with other prophage in the host genome, or with infecting phage (Fortier and Sekulovic 2013). Prior to these recombination, reshuffling, or gene transfer events,

however, the role of prophage in generating *de novo* genetic innovations is largely unknown.

Because phage sequences constitute the most diverse and ubiquitous gene pool on the planet, with extremely large population sizes and short generation times, phage arguably represent the “principle reservoir of genetic novelty” (Koonin and Dolja 2013) for all life forms. For example, recent structural comparisons indicate that a novel ancestral protein fold, most likely in a phage tail fiber, gave rise to a large number of extant proteins in both phage and bacteria (Veesler and Cambillau 2011). Similarly, phage can facilitate lateral gene transfer between highly diverged organisms, as evidenced by the recent discovery of arthropod genetic sequences in a bacteriophage (Bordenstein and Bordenstein 2016). While phage are uniquely well suited to an extensive exploration of genotype space, most innovations in nature are ultimately lost to genetic drift. We sought to understand this process and predict which factors affect the survival of adaptive mutations in this important reservoir.

Overall, our results predict that *de novo* adaptive innovations in bacteriophage are far more likely to survive drift if they first occur during prophage replication, as opposed to occurring during either the lytic replication of a temperate phage, or in an obligately lytic phage. This is due to the relatively low variance in a life cycle defined by binary fission. In contrast, high variance in offspring success is inherent in the lytic cycle: for a population at equilibrium with burst size B , on average only one of B daughter phage can successfully attach to and infect a host cell. Thus the probability of immediate extinction of any new lineage is $1 - 1/B$ for the lytic

Figure 7 The survival probability of mutations on plasmid prophage vs. selective advantage. Panels on the left illustrate traits that are in common between plasmid and chromosomal phage; panels on the right are for traits unique to the plasmid life history. Once again we find that mutations that first occur in the plasmid prophage state are far more likely to survive than those first occurring in the free phage state of the plasmid prophage life cycle. Parameter values are: $A = 0.008$, $B = 100$, $M = 0.2614$, $N = 0.15$, $P = 0.10$, $C = 0.45$, $F = 0.20$, and $J = 0.3886$.

phage, but only 1/2 for prophage propagated through cellular fission. Thus, roughly speaking, we expect that prophage mutations are at least $B/2$ times more likely to survive drift.

As mentioned above, this result is mitigated by the expected rate at which adaptive mutations might occur. Mutations arising via error-prone replication steps, in particular, might occur in lytic lineages in proportion to the burst size, while in prophage only one replication event is necessary per generation. Similarly, for a mutation occurring before the viral genome is copied, the lytic phage could release a burst of B mutant offspring simultaneously. In the branching process approach we develop, these two situations are mathematically equivalent, both allowing B times more mutations to be generated during each round of the lytic, as opposed to the temperate, life cycle. In Figure 4 we demonstrate that this factor eliminates the survival advantage of prophage mutations only for attachment and burst size mutations; mutations

affecting other traits remain more likely to survive if they arise in prophage sequences. A further mitigating factor is the generation time: if the generation time in the lytic cycle is much shorter than in the temperate cycle, again more mutations will be generated. While both of these generation times depend on a multitude of factors, it is possible that in extreme cases the cellular fission time could exceed the lytic generation time by a factor of $B/2$, balancing the prophage advantage.

Nonetheless, these results suggest that prophages may have played a disproportionate role in functional innovation for their associated phage lineages. Experimentally, rapid evolution and coevolution in temperate phages has been observed; for example induction rates and the probability of lysogeny rapidly respond to selection (Refardt and Rainey 2010), whereas temperate phage can adapt to host coevolution by exploiting new host receptors (Meyer *et al.* 2012;

Burmeister *et al.* 2016). Virulent and avirulent (*i.e.*, “mostly” lytic and “mostly” temperate) strains of λ -phage have been compared in well-mixed and spatially structured experiments, elegantly mirroring epidemiological scenarios (Berngruber *et al.* 2013, 2015). The evolutionary potential, however, of obligately lytic vs. temperate phage has not yet been explored. If all else could be held equal, would a temperate phage evolve more rapidly with the genes for lysis knocked out, or with the genes for lysogeny knocked out? Our results suggest that genetic drift would strongly favor innovation in the prophage.

Acknowledgment

L.M.W. is funded by the Natural Sciences and Engineering Research Council of Canada.

Literature Cited

- Bail, O., 1925 Der kolistamm 88 von Gildemeister und Herzberg. *Med. Klin.* 21: 1271–1273.
- Berngruber, T. W., R. Froissart, M. Choisy, and S. Gandon, 2013 Evolution of virulence in emerging epidemics. *PLoS Pathog.* 9: e1003209.
- Berngruber, T. W., S. Lion, and S. Gandon, 2015 Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS Pathog.* 11: e1004810.
- Bobay, L.-M., M. Touchon, and E. P. Rocha, 2014 Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci. USA* 111: 12127–12132.
- Bordenstein, S. R., and S. R. Bordenstein, 2016 Eukaryotic association module in phage WO genomes from *Wolbachia*. *Nat. Commun.* 7: 1–10.
- Bordet, J., 1925 Le problème de l'autolyse microbienne transmissible ou du bactériophage. *Ann. Inst. Pasteur (Paris)* 39: 711–763.
- Bouchard, J. D., and S. Moineau, 2000 Homologous recombination between a lactococcal bacteriophage and the chromosome of its host strain. *Virology* 270: 65–75.
- Brüssow, H., C. Canchaya, and W.-D. Hardt, 2004 Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* 68: 560–602.
- Burmeister, A., R. Lenski, and J. Meyer, 2016 Host coevolution alters the adaptive landscape of a virus. *Proc. Biol. Sci.* 283: 20161528.
- Canchaya, C., C. Proux, G. Fournous, A. Bruttin, and H. Brüssow, 2003 Prophage genomics. *Microbiol. Mol. Biol. Rev.* 67: 238–276.
- Canchaya, C., G. Fournous, and H. Brüssow, 2004 The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* 53: 9–18.
- Casjens, S., 2003 Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49: 277–300.
- Casjens, S. R., 2005 Comparative genomics and evolution of the tailed-bacteriophages. *Curr. Opin. Microbiol.* 8: 451–458.
- Casjens, S. R., E. B. Gilcrease, W. M. Huang, K. L. Bunney, M. L. Pedulla *et al.*, 2004 The pKO2 linear plasmid prophage of *Klebsiella oxytoca*. *J. Bacteriol.* 186: 1818–1832.
- Caswell, H., 2001 *Matrix Population Models: Construction, Analysis, and Interpretation*. Sinauer Associates, Sunderland, MA.
- Durmaz, E., and T. R. Klaenhammer, 2000 Genetic analysis of chromosomal regions of *Lactococcus lactis* acquired by recombinant lytic phages. *Appl. Environ. Microbiol.* 66: 895–903.
- Fortier, L.-C., and O. Sekulovic, 2013 Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* 4: 354–365.
- Hubbarde, J. E., G. Wild, and L. M. Wahl, 2007 Fixation probabilities when the generation times are variable: the burst-death model. *Genetics* 176: 1703–1712.
- Koonin, E. V., and V. V. Dolja, 2013 A virocentric perspective on the evolution of life. *Curr. Opin. Virol.* 3: 546–557.
- Labrie, S. J., and S. Moineau, 2007 Abortive infection mechanisms and prophage sequences significantly influence the genetic makeup of emerging lytic lactococcal phages. *J. Bacteriol.* 189: 1482–1487.
- Lehnherr, H., E. Maguin, S. Jafri, and M. B. Yarmolinsky, 1993 Plasmid addiction genes of bacteriophage P1: *doc*, which causes cell death on curing of prophage, and *phd*, which prevents host death when prophage is retained. *J. Mol. Biol.* 233: 414–428.
- Łobocka, M. B., D. J. Rose, G. Plunkett, M. Rusin, A. Samojedny *et al.*, 2004 Genome of bacteriophage P1. *J. Bacteriol.* 186: 7032–7068.
- Meyer, J. R., D. T. Dobias, J. S. Weitz, J. E. Barrick, R. T. Quick *et al.*, 2012 Repeatability and contingency in the evolution of a key innovation in phage lambda. *Science* 335: 428–432.
- Moineau, S., S. Pandian, and T. R. Klaenhammer, 1994 Evolution of a lytic bacteriophage via DNA acquisition from the *Lactococcus lactis* chromosome. *Appl. Environ. Microbiol.* 60: 1832–1841.
- Patwa, Z., and L. M. Wahl, 2008 Fixation probabilities for lytic viruses: the attachment-lysis model. *Genetics* 180: 459–470.
- Refardt, D., and P. B. Rainey, 2010 Tuning a genetic switch: experimental evolution and natural variation of prophage induction. *Evolution* 64: 1086–1097.
- Sengupta, M., and S. Austin, 2011 Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect. Immun.* 79: 2502–2509.
- Touchon, M., A. Bernheim, and E. P. Rocha, 2016 Genetic and life-history traits associated with the distribution of prophages in bacteria. *ISME J.* 10: 2744–2754.
- Veesler, D., and C. Cambillau, 2011 A common evolutionary origin for tailed-bacteriophage functional modules and bacterial machineries. *Microbiol. Mol. Biol. Rev.* 75: 423–433.
- Wahl, L., and A. D. Zhu, 2015 Survival probability of beneficial mutations in bacterial batch culture. *Genetics* 200: 309–320.
- Wang, P., L. Robert, J. Pelletier, W. L. Dang, F. Taddei *et al.*, 2010 Robust growth of *Escherichia coli*. *Curr. Biol.* 20: 1099–1103.

Communicating editor: N. H. Barton