# Optimality of the spontaneous prophage induction rate

Michael Cortes, Jonathan Krog, Gábor Balázsi

February 10, 2019

**Abstract**

Lysogens are bacterial cells that have survived after being infected by bacterial viruses called bacteriophages. Instead of being killed by the virus, the infected cell survives by integrating the viral DNA into its own genome. This is only possible with "temperate" bacteriophages which do not always lyse their host to reproduce, but sometimes replicate passively using the lysogenic pathway. After an infection resulting in lysogeny, the lysogen continues to grow and divide normally, seemingly unaffected by the integrated viral genome which is now referred to as a prophage. However, the prophage can have an impact on the host's phenotype and overall fitness in certain environments. This makes competition between the lysogen and its nonlysogen counterpart possible because both cells have different genomes and potentially different growth rates. Additionally, the prophages within the lysogens are capable of spontaneously reverting back to the lytic pathway via spontaneous prophage induction (SPI), causing death of the lysogen and the release of new progeny phages. These new phages can then lyse or lysogenize other susceptible nonlysogens, thereby impacting the competition between lysogens and nonlysogens. In a scenario with differing growth rates, it is not clear whether SPI would be beneficial or detrimental to the lysogens since it directly causes cell death but also attacks nonlysogenic competitiors, either lysing or lysogenizing them. In this work we study the evolutionary dynamics of a mixture of lysogens and nonlysogens and derive general conditions on the rate of SPI resulting in lysogens displacing nonlysogens. We show that there exists an optimal SPI rate, and apply the model to bacteriophage $\lambda$. We find that the model can explain why the experimentally measured SPI rate for phage $\lambda$ is so low. We also investigate the impact of stochasticity and conclude that even at low copy numbers the SPI rate can still be fairly low while still providing an advantage to the lysogens.

## 1    Introduction

Bacteriophages (or phages for short) are viruses which infect bacterial cells. By infecting bacteria, the phages can replicate themselves using the host's cellular machinery. The replicated phages can then destroy the host cell and are released into the environment which is known as the lytic or lysis pathway. Lysis is the typical outcome of infection, but some phages can sometimes follow alternative pathway after infection called lysogeny. Phage $\lambda$ is a well-known example for such phages called "temperate" [20]. In the lysogenic pathway, the phages integrate their viral DNA into the host's genome. The host cell, now called a lysogen,

1

is seemingly unaffected and continues to grow and divide normally. Typically this lysogenic state is very stable and the cell remains a lysogen after many cell divisions. Sometimes the lysogen can revert back to the lytic pathway in a process called prophage induction [18]. Prophage induction can be triggered by various environmental factors, such as DNA damage by UV radiation, via induction of the host's SOS response. Prophage induction can also occur spontaneously in a process termed spontaneous prophage induction or SPI [18]. SPI is likely due to spontaneous accumulation of DNA damage initiating the host's SOS response that occurs during cell replication [5, 16, 18, 21].

When a bacterial cell is converted to a lysogen, its genome now contains viral genes. This means that the lysogens genome is different from the nonlysogen's genome, and as a result it is possible that natural selection can occur between the lysogens and nonlysogens. The integrated viral genes can be expressed within the lysogen and affect the lysogen's phenotype and fitness. For many lysogenized cells, the lysogen strain has decreased sensitivity to antibiotics, increased biofilm formation, and virulence [2, 14, 15, 26, 27] due to the integrated viral genes. The additional viral genes could also slow cell growth because expression of extra viral genes could deplete host cell resources that promote cell division [22, 23]. For example, it has been shown that multiple phage $\lambda$ infections reduce host cell transcription and translation [8, 13, 25]. Taken together, these facts imply that lysogens could outcompete nonlysogens or vice versa because of their differing growth rates or fitnesses in certain environments. Thus, in the long run the lysogens could be naturally selected over the nonlysogens or vice versa. Evidence for competitive advantage of lysogens over their nonlysogenized counterparts comes from experiments showing that mixtures of lysogens and lysogen-cured bacteria tend towards a state in which the lysogen starts to displace the lysogen-cured strain [3, 7].

In a mixture of only lysogens and nonlysogens, SPI events could cause some lysogens to lyse, thereby releasing phages into the mixture. The phages can then infect the nonlysogens either killing them via lysis or converting them into lysogens. The lysogens are not affected by additional infections because they are immune to superinfection. Thus, SPI could give the lysogens a competitive advantage, increasing the chances they will be naturally selected. On the other hand, SPI can also be costly to lysogenic fitness because SPI is concomitant with cell death. If there are no nonlysogens left to convert, SPI simply adds to the intrinsic lysogen death rate. It has been suggested that SPI enables lysogens to spread viral DNA within the bacterial population by steady lytic-killing and lysogenic-conversion of the nonlysogens [18]. However, from the perspective of the lysogen, it remains unknown precisely how high or low should the SPI rate be for the lysogens to be naturally selected over the nonlysogens? In other words, does there exist an optimal SPI rate enabling the natural selection of the lysogens over the nonlysogens? If so, then how does this optimal SPI rate depend, if at all, on ecological factors?

In this work, we use replicator dynamics and simulation [19] to gain a theoretical understanding of how SPI influences the competition between lysogens and nonlysogens under general parameter sets, and to identify an optimal SPI rate enabling the lysogens to be naturally selected over the nonlysogens, which is comparable with experimental values. Our model is intended to be general such that it can theoretically describe the competition between any lysogen and its nonlysogen counterpart so long as the lysogen can undergo SPI. We quantitatively compare our model to the lysogens from bacteriophage $\lambda$ infecting $E.$ $coli$ because many of the model's parameters for the phage $\lambda$ - $E.$ $coli$ system have been

2

measured previously [20]. Our major finding is that the SPI rate should be as low as possible but still greater than a lower bound which we derive analytically. For phage $\lambda$ we find that the experimentally measured SPI rate matches our model's prediction. We also derive expressions to describe how stochasticity affects the competition between nonlysogens and lysogens undergoing SPI.

## 2 The LUV model

To understand how SPI plays into the natural selection of lysogenic cells over their uninfected counterparts, we developed a model summarized by equation 1 - 5. This model simulates evolutionary population dynamics using replicator equations [19] with carrying capacity $K \approx 10^6$ and in time units of per cell generation ($\approx$ 30 minutes per cell [1]). Since the lysogenic genome contains viral genes, there are many ways in which lysogenized cells could be different from their nonlysogenized counterparts. These differences could influence host cell physiology by directly affecting the intrinsic cell growth rate via expression of viral genes integrated within the host's genome. Therefore, in our model we assume that the lysogens, $L$, grow at a rate $r$ and the nonlysogens, $U$, grow at a rate $g$. The model simulates lysogenic and nonlysogenic cells growing in the same environment with carrying capacity $K$ (which reflects nutrient and space limitations). The lysogens undergo SPI at a rate $\sigma$, and the resulting free viruses, $V$, infect the nonlysogens $U$ to induce $\alpha$ decision-making events per cell per phage per generation. A proportion $p$ of these phage-cell encounters enter the lysogenic pathway, and the remaining proportion $1 - p$ enter the lytic pathway. When the lytic pathway occurs, $b$ free viruses are released given by the burst size of the phage. The viruses degrade at rate $\gamma$.

$$\dot{L} = rL - \sigma L + p\alpha UV - \phi L \tag{1}$$

$$\dot{U} = gU - \alpha UV - \phi U \tag{2}$$

$$\dot{V} = (1-p)b\alpha UV - \gamma V + b\sigma L \tag{3}$$

$$\phi = \frac{(r - \sigma)L + gU - \alpha UV(1-p)}{K} \tag{4}$$

$$K = L + U \tag{5}$$

The parameter variables are summarized in detail in Table 1 along with the parameter values for the phage $\lambda$ - *E. coli* system. The burst size of phage $\lambda$ is given by $b \approx 150$. The probability of lysogeny is known to be host cell volume and MOI dependent, ranging from 20% to 80% [9, 10, 24, 28], but we take $p \approx 0.3$. Cell growth rates are determined by cell doubling times ranging from 20 mins to 40 mins [1, 29], so we take $g = 1$ cell division per cell per generation. Similarly, we take $r$ to be $\approx 1$ cell per generation, but we investigate different $r$ values in the range $0.9 \leq r \leq 1$ to determine how the growth-cost due to lysogeny affects the dynamics. Phage degradation occurs on the order of 0.1 per day per phage [6],

3

which converts to $\gamma \approx 0.001$ per generation per phage. The parameter $\alpha$ measures the rate at which nonlysogens are infected and either lysed or lysogenized per generation per phage per cell. This includes the phage infection/adsorption rate and the rate at which infected cells develop lytically or lysogenically per generation. We estimated $\alpha$ from the phage adsorption experiment in [4,17] by fitting a small mass-action model of the phage-cell infection dynamics to the raw data. This $\alpha$ estimate allowed fixation events to occur on the timescale of a few days, similar to the experimental results in [7] which compete phage $\lambda$ lysogens against their nonlysogenic counterparts. Increasing $\alpha$ by a factor of 10 resulted in fixation events occurring on the timescale of 10s of days, but the overall behavior of the system remained identical. The intrinsic SPI rate of phage $\lambda$ in a $recA^-$ background is $< 10^{-8}$, but in a $recA^+$ background the SPI rate is estimated to be about $10^{-6}$ to $10^{-7}$ per generation [11, 12, 29]. Since $recA^-$ mutants probably don't exist in nature, we use the $recA^+$ SPI rate because natural selection likely acts on the $recA^+$ strain and not the $recA^-$ strain.

Table 1: LUV model parameters

| Parameter | Symbol | Value |
|---|---|---|
| Lysogen replication rate | $r$ | $0.9 \leq r \leq 1$ per gen |
| Nonlysogen replication rate | $g$ | 1 per gen |
| Carrying capacity | $K$ | $\geq 10^6$ |
| Phage-cell binding/infection rate | $\alpha$ | $\approx 10^{-7}$ per phage per cell per gen |
| Phage burst size | $b$ | $\approx 150$ |
| Lysogenization rate | $p$ | $\approx 0.30$ |
| Phage degradation rate | $\gamma$ | $\approx 0.001$ per gen |
| SPI rate | $\sigma$ | $\approx 10^{-6}$ per gen |

We can take advantage of the differing time scales of the reactions to impose some conditions (inequalities) on the parameters. From our parameter estimates, we find that the quantity $\alpha bK/\gamma \approx 1.5 \cdot 10^4$ which is quite large. The numerator $\nu_1 = \alpha bK$ is essentially the rate (per generation) at which a population of $K$ nonlysogens are lysed or lysogenized given a single phage lytic event released $b$ infecting phages. The denominator $\nu_2 = \gamma$ is the rate at which a single phage degrades per generation. Since phage degradation is on the order of days and lysis-lysogeny decision-events occur on the order of hours, it is no surprise that the ratio $\nu_1/\nu_2$ is quite large using our literature estimates. Thus, we assume that $\alpha bK/\gamma \gg 1$ throughout our study which should hold even if there is some uncertainty in the parameter estimates since this quantity is so large compared to 1. We also assume that $r > \sigma$ because if $\sigma > r$ the net growth rate of the lysogens would be negative indicating an unsustainable population. The lysogen's genome is the same as the nonlysogen's genome except that it contains more genes from the integrated viral genome (the genes from the prophage). Maintenance of the lysogenic state generally requires continued gene expression of a key viral repressor protein to block entrance into the lytic pathway, and in the specific case of phage $\lambda$ this protein is the CI repressor. It is possible that the cell incurs some growth-rate cost by maintaining expression of this viral repressor. Experiments competing $\lambda$ lysogens against their nonlysogenic counterparts showed that their growth rates are not statistically different, suggesting that $r = g$ [7]. However, it is still possible that $r$ could be

4

different from $g$ by a very small amount that could be difficult to detect. For example, this could happen if $g-r = 0.01$ or less. From these observations we assume that the nonlysogenic growth rate $g$ is at least $r$ or otherwise at least slightly larger. These inequalities will be important later on when analyzing this model. Note that since $g \geq r$ then $g > \sigma$ as well. Also note that if $r > g$ then the lysogens would displace the nonlysogens even without SPI in the long run, and the inclusion of SPI when $r > g$ would only accelerate the process. Since this observation is somewhat trivial, we do not specifically investigate it in this work. However, we do prove this analytically later on.

With the model and its parameters specified, we numerically simulated the LUV model using the Runge Kutta integration scheme (see Fig. 1) for the parameter values in Table 1 to gain an understanding of the role of SPI in the competition between lysogens and nonlysogens. The initial conditions we chose were $L(t = 0) = 0.01K$ and $U = 0.99K$ which implies that the lysogens initially comprise 1% of the population. Since $K \geq 10^6$ the initial values of $L$ and $U$ were at least $10^4$ which allows us to well-approximate the behavior of this system using deterministic models. We investigate stochastic models later on in this work given the fact that the low SPI rate can still introduce stochasticity into the model. The simulations show that this system tends towards a state in which the lysogens completely displace the nonlysogens in the long run even though the lysogenic growth rate is lower $(g > r)$. This suggests that SPI allows the lysogens to subvert the nonlysogens even if the lysogenic state is costly to cell growth. This result is consistent with experiment observations in which lysogen strains displaced lysogen-cured strains when growth together [3, 7, 18]. Essentially, a low but steady rate of lysogens undergoing SPI allows the released phages to either kill off the nonlysogen strain or convert them to the lysogenic state.
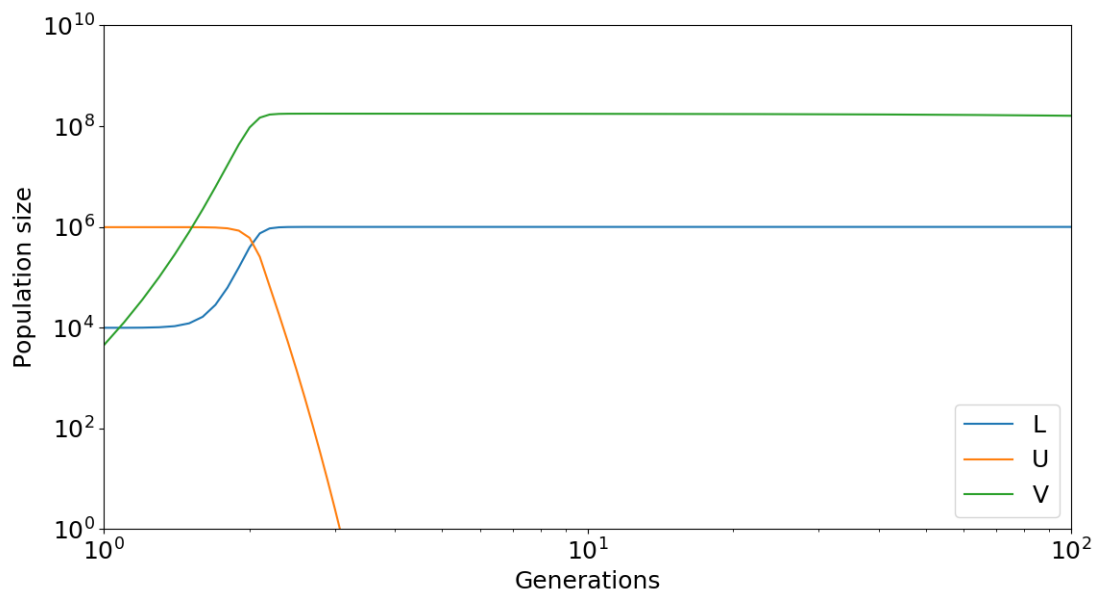


Figure 1: LUV model simulated using parameters from Table 1. Specifically, we set $r = 0.99$ and $\sigma = 10^{-6}$. We get the same results if we vary the model's parameters.

# 3   Lower bound on SPI rate for lysogenic advantage

To determine the exact bounds on the SPI rate necessary for the lysogens to displace the nonlysogens, we analytically solved our model (equations 1 - 5) at steady state to obtain its equilibrium points and then tested their stability by calculating the eigenvalues of the system's associated Jacobian matrix. At first glance, this seems like a very difficult task due to the nonlinearity of the model, but fortunately the model admits analytical solutions for its equilibrium points. We used Mathematica to check all calculations and derivations.

First, we set $\dot{V} = 0$ in equation 3 and solve for $V$. This gives us the following relationship.

$$V = \frac{b\sigma L}{\gamma - b(1-p)\alpha U} \tag{6}$$

Next, we set $\dot{U} = 0$ using equation 2.

$$0 = U[g - \alpha V - \phi] \tag{7}$$

The only way for 7 to be zero is if either $U = 0$ or if $[g - \alpha V - \phi] = 0$. We will analyze both cases.

If $U = 0$ it immediately requires that $L = K$ by constraint equation 5. This then immediately implies that $V = b\sigma K/\gamma$ by equation 6. Thus, the point $P_1^{LUV} = (L, U, V) = (K, 0, b\sigma K/\gamma)$ is one equilibrium point of our system.

Otherwise, if $g - \alpha V - \phi = 0$ then we can plug in 4 for $\phi$. The relation $g - \alpha V - \phi = 0$ immediately implies that equation 8 must hold.

$$(g - r + \sigma)(K - U)(\gamma - b(1-p)\alpha U) = \alpha b\sigma (K - U)(K - U(1-p)) \tag{8}$$

We can immediately note that $U = K$ is a possible solution of equation 8. Setting $U = K$ implies $L = 0$ by the constraint 5. Then, using equation 6 we immediately see that $V = 0$ if $U = K$. Thus, the second equilibrium point of our system is given by $P_2^{LUV} = (L, U, V) = (0, K, 0)$.

If $U \neq K$, then we can cancel the factor $(K - U)$ and solve for $U$ to obtain $U = \gamma (g - r - \sigma (\alpha bK/\gamma - 1))/\alpha b(1-p)(g - r)$. We can then use this value for $U$ to calculate $L$ and $V$ using $L = K - U$ and equation 6. This solution set defines equilibrium point 3, which we denote using $P_3^{LUV}$.

The equilibrium points of our system are summarized in Table 2. To summarize, $P_1^{LUV}$ describes a state in which the lysogens have displaced the nonlysogens, $P_2^{LUV}$ describes a state in which the nonlysogens have displaced the lysogens, and $P_3^{LUV}$ describes a state in which both the lysogens and nonlysogens coexist. Given these equilibrium points, we seek to determine the conditions necessary for the lysogens to displace the nonlysogens.

## 3.1   Stability of equilibrium points for LUV model

To determine the conditions required for the lysogens to displace the nonlysogens, we investigated the stability of the LUV model's equilibrium points (Table 2) by computing the eigenvalues of our model's associated Jacobian matrix $J^{LUV}$.

Table 2: LUV model equilibrium points

| Equilibrium point | L | U | V |
|---|---|---|---|
| $P_1^{LUV}$ | $K$ | $0$ | $b\sigma K/\gamma$ |
| $P_2^{LUV}$ | $0$ | $K$ | $0$ |
| $P_3^{LUV}$ | $K - \frac{g-r-\sigma(\alpha bK/\gamma-1)}{(1-p)(g-r)\alpha b/\gamma}$ | $\frac{g-r-\sigma(\alpha bK/\gamma-1)}{(1-p)(g-r)\alpha b/\gamma}$ | $\frac{(g-r+\sigma)-\alpha bK((g+r)(1-p)-\sigma)/\gamma}{\alpha(1+\alpha bK)(1-p)/\gamma}$ |

$$J^{LUV} = \begin{pmatrix} \partial\dot{L}/\partial L & \partial\dot{L}/\partial U & \partial\dot{L}/\partial V \\ \partial\dot{U}/\partial L & \partial\dot{U}/\partial U & \partial\dot{U}/\partial V \\ \partial\dot{V}/\partial L & \partial\dot{V}/\partial U & \partial\dot{V}/\partial V \end{pmatrix} \quad (9)$$

We calculate the partial derivatives in $J^{LUV}$ using equations defining the LUV model (equations 1 - 5) and we define $J_i^{LUV}$ to be $J^{LUV}$ evaluated at equilibrium point $P_i^{LUV}$.

First, we analyze the stability of equilibrium point $P_1^{LUV}$ from Table 2, namely $(L, U, V) = (K, 0, b\sigma K/\gamma)$. This point characterizes a state of the system in which the lysogens have displaced the nonlysogens. Using these values for $L$, $U$, and $V$ we calculate $J_1^{LUV}$.

$$J_1^{LUV} = \begin{pmatrix} -r+\sigma & -g+\alpha bK\sigma/\gamma & 0 \\ 0 & g-r+\sigma-\alpha bK\sigma/\gamma & 0 \\ b\sigma & \alpha b^2 K(1-p)\sigma/\gamma & -\gamma \end{pmatrix} \quad (10)$$

The eigenvalues of $J_1^{LUV}$ are given by calculating the determinant $\det(J_1^{LUV} - \lambda I)$ where $I$ is the identity matrix. This results in the following characterisic polynomial with eigenvalue $\lambda$ (equation 11).

$$0 = (-r+\sigma-\lambda)\left(g-r-\sigma(\alpha bK/\gamma-1)-\lambda\right)(-\gamma-\lambda) \quad (11)$$

The eigenvalues are the values of $\lambda$ which solve equation 11. We see that these values are $\lambda_1 = -r+\sigma$, $\lambda_2 = g-r-\sigma(\alpha bK/\gamma-1)$, and $\lambda_3 = -\gamma$. The equilibrium point $(K, 0, bK\sigma/\gamma)$ is a stable point of the system if all eigenvalues are negative. Since $\gamma > 0$ we know that $\lambda_3$ is negative. Since $\sigma < r$ we have that $\lambda_1$ is negative since $\lambda_1 = -(r-\sigma) < 0$. The eigenvalue $\lambda_2$ is negative if and only if $\sigma > \sigma_{LB}$ with $\sigma_{LB}$ given by equation 12.

$$\sigma_{LB} = \frac{g-r}{\alpha bK/\gamma - 1} \quad (12)$$

Using literature estimates for the parameter values from Table 1, and specifically with $r = 0.99$ and $\sigma = 10^{-6}$, we estimate $\sigma_{LB} \approx 6 \cdot 10^{-7}$. Thus, $\sigma$ must be larger than this lower bound (defined as $\sigma_{LB}$) in order to establish $(K, 0, bK\sigma/\gamma)$ as a stable equilibrium point. This is very close to the experimentally measured value between $10^{-6}$ or $10^{-7}$ per generation for phage $\lambda$ measured most recently [29].

Recall that we assume $g \geq r$ so there are two possibilities, either $g = r$ or $g > r$. If $g = r$ then the lysogenic state does not hurt cell growth rate and, as a consequence, any $\sigma > \sigma_{LB} = 0$ would cause the point $(K, 0, bK\sigma/\gamma)$ to be a steady equilibrium point. Furthermore, if $g = r$ then critical point $P_3^{LUV}$ no longer exists (the system goes from having 3 equilibrium points to only 2). Otherwise, if $g > r$ then the factor $g - r$ is greater than 0

7

and hence makes the lower bound $\sigma_{LB}$ nonzero. Using the values in Table 1 we estimated that if $g - r \approx 0.01$ per generation then $\sigma_{LB} \approx 6 \cdot 10^{-7}$ per generation. If the lysogenic growth-cost was even smaller (e.g. if $g - r$ is smaller), then $\sigma_{LB} < 10^{-7}$. For example, with $g = 0.999$ then $\sigma_{LB} = 6 \cdot 10^{-8}$. The literature estimate for $\sigma^*$ is $10^{-6}$ [29] which is $\geq \sigma_{LB}$ even in the case that $r = 0.99$. This literature estimate is roughly 10 times larger than this lower bound, which means $\sigma^* > \sigma_{LB}$, implying that equilibrium point $P_1^{LUV}$ is a stable state and equilibrium point $P_3^{LUV}$ does not exist. Note that if $g - r$ was smaller (e.g. the lysogenic and nonlysogenic growth rates were very similar, which is what we expect) or if $K$ was larger (say $K = 10^9$), this lower bound would be even smaller causing $\sigma_{LB}$ to tend towards 0. Thus, the literature estimate of the phage $\lambda$ SPI rate $\sigma^*$ is likely much larger than the lower bound $\sigma_{LB}$. The important thing to note here is that this lower bound is easily a very small number due to the size of the constants in the denominator. Although we assume $g \geq r$, for completeness we also examine what happens when $r > g$. In this case the lower bound $\sigma_{LB}$ is negative. Thus, if $r > g$ then the lysogens will displace the nonlysogens even if $\sigma = 0$ because $0 > \sigma_{LB}$.

Next, we examine equilibrium point $P_2^{LUV} = (0, K, 0)$. This point describes a state in which the nonlysogens have displaced the lysogens. The stability of point 2 is assessed by calculating the eigenvalues of $J_2^{LUV}$ matrix. The Jacobian matrix for $(L, U, V) = (0, K, 0)$ is given in equation 13.

$$J_2^{LUV} = \begin{pmatrix} -g + r - \sigma & 0 & K\alpha p \\ \sigma - r & -g & -K\alpha p \\ b\sigma & 0 & K\alpha b(1 - p) - \gamma \end{pmatrix} \tag{13}$$

The eigenvalues of $J_2^{LUV}$ are obtained by calculating its determinant, which results in a characteristic polynomial (equation 14).

$$0 = (-g - \lambda)[\lambda^2 - \lambda(-g + r - \sigma + K\alpha b(1 - p) - \gamma) \\ + (-g + r - \sigma)(K\alpha b(1 - p) - \gamma) - K\alpha p b\sigma] \tag{14}$$

The eigenvalues are then obtained in a straightforward manner. First, $\lambda_1 = -g < 0$. For $\lambda_2$ and $\lambda_3$, we see that these are the roots of the quadratic inside the square brackets in equation 14. The roots of this quadratic take the form $\lambda = q \pm \sqrt{q^2 + \varepsilon}$ with $q = (-g - \gamma + \alpha b K(1 - p) + r - \sigma)/2$ and $\varepsilon = ((\alpha b K(1 - p) - \gamma)(g - r) + (\alpha b K - \gamma)\sigma))$. It is easy to show that if $\varepsilon > 0$ then one of the solutions $\lambda > 0$. In this case, we have $\varepsilon > 0$ because every factor and term in the $\varepsilon$ expression is positive. We have that $(\alpha b K/\gamma - 1) > 0$, $(g - r) \geq 0$, $\gamma > 0$, and $\sigma > 0$. Since $\alpha b K/\gamma \gg 1$ we also have that $(1 - p)\alpha b K/\gamma > 1$. Thus, equilibrium point $P_2^{LUV} = (0, K, 0)$ is an unstable equilibrium point of this system.

Finally, we examine the stability of $P_3^{LUV}$ which describes a state of the system in which both the lysogens and nonlysogens coexist. First, note that this equilibrium point does not exist if $\sigma$ is larger than the lower bound $\sigma_{LB}$ in 12. Thus, if inequality 12 holds then we do not need to analyze point 3's stability. Since the literature estimate $\sigma^* > \sigma_{LB}$ we hypothesize that this equilibrium point indeed does not exist, but for completeness we examine its stability here.

If $\sigma < \sigma_{LB}$ then both $P_1^{LUV}$ and $P_2^{LUV}$ are unstable. If a steady state indeed does exist under the condition $\sigma < \sigma_{LB}$ then it must be true that $P_3^{LUV}$ becomes the (only) steady

state of the system. The Jacobian matrix for this equilibrium point, $J_3^{LUV}$, is a 3 by 3 matrix with extremely complicated expressions for its elements due to the fact that the expressions for $L$, $U$, and $V$ for $P_3^{LUV}$ are quite complicated themselves. Therefore, we do not show the full Jacobian here, but we numerically evaluate its eigenvalues by using parameter values from Table 1 except we set $\sigma = \theta \cdot \sigma_{LB}$ so that inequality 12 does not hold, which means $\theta < 1$. We discovered that for all $\theta < 1$ the eigenvalues of $J_3^{LUV}$ all had strictly negative real parts, implying that this becomes a steady state of the system if $\sigma < \sigma_{LB}$. Therefore, if $\sigma$ is extremely small (so that the SPI rate is lower than $\sigma LB$) then the steady state of the system is one in which both the lysogens and nonlysogens coexist.

Overall, we have shown that lysogens will displace nonlysogens even if the lysogenic state is costly to cell growth provided that $\sigma$ is at least $(g-r)/[K\alpha b/\gamma - 1]$. We have also shown that the true SPI rate of phage $\lambda$ is likely greater than this lower bound, suggesting that the SPI rate has evolved so that lysogens can displace nonlysogens. However, this analysis only tells us the lower bound on $\sigma$ required for lysogens to have the advantage. That is, $\sigma$ is allowed to be any number greater than $\sigma_{LB}$. Thus, it is natural to ask if a larger SPI rate is advantageous.

## 4 The LUV2 model

In the previous section we showed that if $r > \sigma > (g-r)/[K\alpha b/\gamma - 1] = \sigma_{LB}$ then a mixture of lysogenic and nonlysogenic cells will tend towards a state in which the nonlysogens are displaced in favor of the lysogens. Note that this inequality places no strong restriction on the exact value of $\sigma$. The only requirement is that it must be strictly greater than $\sigma_{LB}$ and must be less than the intrinsic cell replication rate $r$. To understand what value of $\sigma$ may be optimal, we adjusted the LUV model so that it included two lysogen strains, $L_1$ and $L_2$, which differ only in their SPI rates $\sigma_1$ and $\sigma_2$ respectively. These lysogens will compete with each other and with the nonlysogenic cells $U$. These lysogen strains also produce their own phages, denoted $V_1$ and $V_2$, respectively. All other parameters and interactions between the two lysogen strains and the nonlysogens is kept the same so that we ensure we are only comparing different SPI rates. The overall idea here is that we want to identify an SPI rate which would allow a lysogen strain to outcompete another lysogen strain while also being able to displace the nonlysogens. The new model is summarized by equations 15 - 21.

$$\dot{L}_1 = rL_1 - \sigma_1 L_1 + p\alpha U V_1 - \phi L_1 \tag{15}$$

$$\dot{L}_2 = rL_2 - \sigma_2 L_2 + p\alpha U V_2 - \phi L_2 \tag{16}$$

$$\dot{U} = gU - \alpha U (V_1 + V_2) - \phi U \tag{17}$$

$$\dot{V}_1 = (1-p)b\alpha U V_1 - \gamma V_1 + b\sigma_1 L_1 \tag{18}$$

$$\dot{V}_2 = (1-p)b\alpha U V_2 - \gamma V_2 + b\sigma_2 L_2 \tag{19}$$

$$\phi = \frac{1}{K} \left[ (r - \sigma_1) L_1 + (r - \sigma_2) L_2 + gU - \alpha U (V_1 + V_2)(1 - p) \right] \tag{20}$$

$$K = L_1 + L_2 + U \tag{21}$$

In this model, $L_1$ denotes the number of lysogen strain 1 with SPI rate $\sigma_1$ and $L_2$ denotes the number of lysogen strain 2 with SPI rate $\sigma_2$. The variables $V_1$ and $V_2$ are the phages emitted from lysogen strains 1 and 2, respectively, via the SPI process. All parameters are the same as in the previous model. Without loss of generality, we assume that $\sigma_1 > \sigma_2$ so that lysogen strain 1 has a higher SPI rate than lysogen strain 2. Note that here we are also assuming the same inequalities hold from analyzing the LUV model, namely $\alpha bK/\gamma > 1$, $g - r \geq 0$, and $r > \sigma_1$. Since $\sigma_1 > \sigma_2$ we also have $r > \sigma_2$.

To gain an understanding of the behavior of this system, we simulated it using the parameter values in Table 1. We $r = 0.99$ and $\sigma_1$ to be 10 times greater than $\sigma_2$ with $\sigma_2 = 10^{-6}$. In Fig 2 we show a sample trajectory of the system using initial conditions $L_1(t = 0) = 0.005K$, $L_2(t = 0) = 0.005K$, and $U(t = 0) = 0.99K$ so that the lysogens initially only comprise 1% of the population, and this 1% is split evenly among the two lysogen strains. We see that early on, $L_1$ has a clear advantage in the sense that it greatly outnumbers both $L_2$ and the nonlysogens $U$. However, in the long run we see that $L_2$ actually starts to overtake $L_1$, and that this switch occurs after $U$ decreases towards 0.
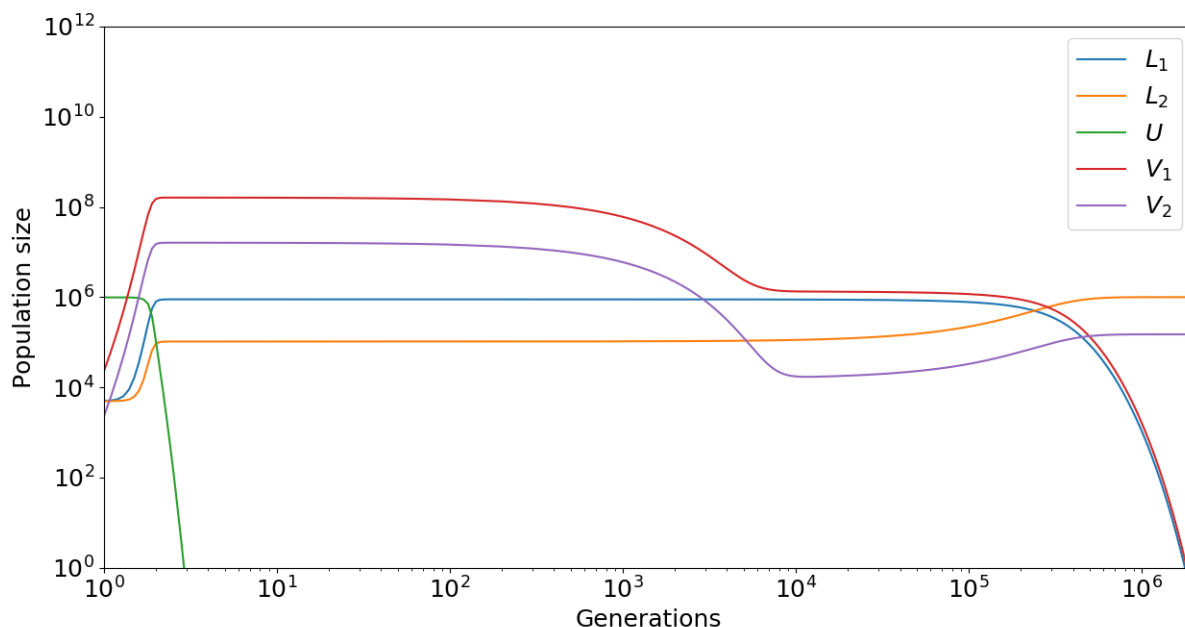


Figure 2: LUV2 model simulation using default parameters from Table 1. We used $r = 0.99$, $\sigma_2 = 10^{-6}$, and $\sigma_1 = 10^{-5}$.

Numerical simulations of various parameter sets showed similar behavior, with $L_2$ always displacing both $L_1$ and $U$ in the long run. This suggests that lower SPI rates are favorable in the long run. Intuitively, we can reason that once the nonlysogens go extinct there is no

10

more benefit to having a higher SPI rate. Thus, we can conclude that if both lysogen strains $L_1$ and $L_2$ survive after $U$ reaches 0, then the lysogen with the lower SPI rate will be selected for in the long run.

# 5  Slower SPI rates are advantageous

To prove that the lysogenic strain with the lower SPI rate is the one which is favored by natural selection, we calculate the equilibrium points of our model and assess their stability. We impose the requirement that each of these lysogen strains be able to outcompete nonlysogens if separately competed with them, which implies that both $\sigma_1 > \sigma_{LB}$ and $\sigma_2 > \sigma_{LB}$. As in the LUV model, we first begin with setting the free phage rate equations to zero (equations 18 and 19) i.e. $\dot{V}_1 = \dot{V}_2 = 0$. This results in the following relationships for $V_1$ and $V_2$.

$$V_1 = \frac{b\sigma_1 L_1}{\gamma - b(1-p)\alpha U} \tag{22}$$

$$V_2 = \frac{b\sigma_2 L_2}{\gamma - b(1-p)\alpha U} \tag{23}$$

Next, we set equation 17 to 0 so that $\dot{U} = 0$. This implies that:

$$0 = U\left[g - \alpha(V_1 + V_2) - \phi\right] \tag{24}$$

Equation 24 can be solved if either $U = 0$ or if $g - \alpha(V_1 + V_2) - \phi = 0$. We will examine both cases. If $U = 0$ then we immediately see that $V_1$ and $V_2$ are determined by their respective lysogen strain levels. That is, if $U = 0$ we have that $V_1 = b\sigma_1 L_1/\gamma$ and $V_2 = b\sigma_2 L_2/\gamma$. Thus, we focus on $L_1$ and $L_2$ next. If $U = 0$ then $K = L_1 + L_2$ by the constraint 21. This means that $L_1 = K - L_2$. We can plug this into the rate equation for $\dot{L}_1$ (equation 15) and set $\dot{L}_1 = 0$. After simplifying, this results in an equation with only 1 unknown variable, namely $L_2$.

$$0 = L_2\left[1 - \frac{L_2}{K}\right] \tag{25}$$

The two solutions to this equation are $L_2 = 0$ or $L_2 = K$. Along with the constraint (equation 21), we know that $(L_1, L_2) = (K, 0)$ and $(L_1, L_2) = (0, K)$ are both solutions of equation 25. Using equations 6 and 23, we can write down two equilibrium points of our system $(L_1, L_2, U, V_1, V_2)$, namely $P_1^{LUV2} = (K, 0, 0, b\sigma_1 K/\gamma, 0)$ and $P_2^{LUV2} = (0, K, 0, 0, b\sigma_2 K/\gamma)$.

If instead $g - \alpha(V_1 + V_2) - \phi = 0$ then we can combine this equation with equations 22, 23, and 20 to show that the following equation must hold.

$$\frac{L_1(g - r + \sigma_1) + L_2(g - r + \sigma_2)}{\alpha\left[K - (1-p)U\right]} = \frac{b\sigma_1 L_1 + b\sigma_2 L_2}{\gamma - b(1-p)\alpha U} \tag{26}$$

We can immediately see that if $L_1 = L_2 = 0$ then equation 26 holds. If $L_1 = L_2 = 0$ then $U = K$ by the constraint 21. Furthermore, if $L_1$ and $L_2$ are both zero then we also have

that $V_1 = V_2 = 0$ by equations 6 and 23. Thus, the third equilibrium point of the system is given by $P_3^{LUV2} = (0, 0, K, 0, 0)$.

If both $L_1$ and $L_2$ are not zero, then we need to solve the equation 26 for both $L_1$, $L_2$, and $U$ along with equations 15 - 21. However, this is very difficult to do by hand. Thus, we use Mathematica here to derive the additional solutions. It turns out that there are three additional solutions, which we denote using $P_4^{LUV2}$, $P_5^{LUV2}$, and $P_6^{LUV2}$. However, some of these solutions are extremely complicated expressions, so we do not show them in their full detail. These additional solutions describe states of the system in which $L_1$, $L_2$, and $U$ coexist in the long run. Hence, these solutions are not of particular interest since our goal is to determine the conditions required for a lysogen strain to displace all others and the nonlysogens.

For equilibrium points $P_4^{LUV2}$ and $P_5^{LUV2}$, we find that the expression for $U$ has the same form, namely $U = [\gamma(g - r + \sigma_j) - \alpha bK\sigma_j)] / (\alpha b(1-p)(g-r))$ with $j = 1$ for point 4 and $j = 2$ for point 5. In order for these equilibrium points to be unique (i.e. different from points 1,2 and 3) we must have that $0 < U < K$. We find that $U > 0$ if and only if $\sigma_j < (g-r)/[K\alpha b/\gamma - 1] = \sigma_{LB}$. However, this contradicts our initial assumption that both $\sigma_1, \sigma_2 > \sigma_{LB}$. Indeed, $\sigma_1 > \sigma_{LB}$ and $\sigma_2 > \sigma_{LB}$ implies that $U < 0$ which is unphysical since population counts cannot be negative. Therefore, equilibrium points $P_4^{LUV2}$ and $P_5^{LUV2}$ do not exist under our parameter regime.

Finally, we analyze the equilibrium point $P_6^{LUV2}$ for which expressions for $L_1$, $L_2$, and $U$ are as follows:

$$
\begin{aligned}
U &= \frac{\gamma}{\alpha b} \\
L_1 &= \frac{\gamma p(g - r) - \gamma\sigma_2 (\alpha bK/\gamma - 1)}{\alpha b (\sigma_1 - \sigma_2)} \\
L_2 &= \frac{-\gamma p(g - r) + \gamma\sigma_1 (\alpha bK/\gamma - 1)}{\alpha b (\sigma_1 - \sigma_2)}
\end{aligned}
\tag{27}
$$

Using the fact that $(\sigma_1 - \sigma_2) > 0$ we can easily show that $L_1, L_2 > 0$ if and only if $\sigma_1 > p(g-r)/[K\alpha b/\gamma - 1]$ and $\sigma_2 < p(g-r)/[K\alpha b/\gamma - 1]$. These conditions are consistent with $\sigma_1 > \sigma_2$, but they contradict the fact that $\sigma_2 > \sigma_{LB}$. To see this, note that $\sigma_2 < p(g-r)/[K\alpha b/\gamma - 1] < (g-r)/[K\alpha b/\gamma - 1] = \sigma_{LB}$ since $0 < p < 1$. Indeed, $\sigma_2 > \sigma_{LB}$ implies that $L_1 < 0$ which is unphysical. Therefore, our original requirement that $\sigma_1, \sigma_2 > \sigma_{LB}$ causes equilibrium point $P_6^{LUV2}$ to not exist.

All equilibrium points of the LUV2 model which exist under our parameter regime are summarized in Table 3. Note that these equilibrium points are similar to the ones we found for the LUV model.

## 5.1 Stability analysis of LUV2 equilibrium points

To understand what conditions enable one lysogen strain to outcompete the other, we analyzed the stability of our model's equilibrium points by calculating the eigenvalues of the Jacobian matrix $J^{LUV2}$ for each equilibrium point in Table 3.

Table 3: LUV2 model equilibrium points

| equilibrium point | $L_1$ | $L_2$ | $U$ | $V_1$ | $V_2$ |
|---|---|---|---|---|---|
| equilibrium point 1 | $K$ | 0 | 0 | $b\sigma_1 K/\gamma$ | 0 |
| equilibrium point 2 | 0 | $K$ | 0 | 0 | $b\sigma_2 K/\gamma$ |
| equilibrium point 3 | 0 | 0 | $K$ | 0 | 0 |

$$J^{LUV2} = \begin{pmatrix} \partial\dot{L}_1/\partial L_1 & \partial\dot{L}_1/\partial L_2 & \partial\dot{L}_1/\partial U & \partial\dot{L}_1/\partial V_1 & \partial\dot{L}_1/\partial V_2 \\ \partial\dot{L}_2/\partial L_1 & \partial\dot{L}_2/\partial L_2 & \partial\dot{L}_2/\partial U & \partial\dot{L}_2/\partial V_1 & \partial\dot{L}_2/\partial V_2 \\ \partial\dot{U}/\partial L_1 & \partial\dot{U}/\partial L_2 & \partial\dot{U}/\partial U & \partial\dot{U}/\partial V_1 & \partial\dot{U}/\partial V_2 \\ \partial\dot{V}_1/\partial L_1 & \partial\dot{V}_1/\partial L_2 & \partial\dot{V}_1/\partial U & \partial\dot{V}_1/\partial V_1 & \partial\dot{V}_1/\partial V_2 \\ \partial\dot{V}_2/\partial L_1 & \partial\dot{V}_2/\partial L_2 & \partial\dot{V}_2/\partial U & \partial\dot{V}_2/\partial V_1 & \partial\dot{V}_2/\partial V_2 \end{pmatrix} \tag{28}$$

For $P_1^{LUV2} = (L_1, L_2, U, V_1, V_2) = (K, 0, 0, Kb\sigma_1/\gamma, 0)$, the Jacobian matrix is given by 29. This equilibrium point describes a state in which lysogen strain $L_1$ (with the higher SPI rate) displaces both the nonlysogens $U$ and lysogen strain $L_2$.

$$J_1^{LUV2} = \begin{pmatrix} -r + \sigma_1 & -r + \sigma_2 & -g + (\alpha b K \sigma_1)/\gamma & 0 & 0 \\ 0 & \sigma_1 - \sigma_2 & 0 & 0 & 0 \\ 0 & 0 & g - r + \sigma_1 - (\alpha b K \sigma_1)/\gamma & 0 & 0 \\ b\sigma_1 & 0 & -((\alpha b^2 K(-1+p)\sigma_1)/\gamma) & -\gamma & 0 \\ 0 & b\sigma_2 & 0 & 0 & -\gamma \end{pmatrix} \tag{29}$$

The eigenvalues of $J_1^{LUV2}$ are determined by solving the equation $\det(J_1^{LUV2} - \lambda I) = 0$ which results in the following characteristic polynomial.

$$0 = ((\lambda + \gamma)^2(\lambda + r - \sigma_1)(\lambda - g + r - \sigma_1 + \alpha b K \sigma_1/\gamma)(\lambda - \sigma_1 + \sigma_2)) \tag{30}$$

The eigenvalues are then easily obtained by setting each factor to 0 and solving for $\lambda$. The eigenvalues are $\lambda_1 = -\gamma$, $\lambda_2 = -\gamma$, $\lambda_3 = -r + \sigma_1$, $\lambda_4 = g - r - \sigma_1(\alpha b K/\gamma - 1)$, and $\lambda_5 = \sigma_1 - \sigma_2$. Without inspecting each eigenvalue, we can immediately conclude $P_2^{LUV2}$ is an unstable equilibrium point because eigenvalue $\lambda_5 > 0$ since $\sigma_1 > \sigma_2$. Thus, the system will not tend to a state in which lysogen strain $L_1$ displaces lysogen strain $L_2$ and the nonlysogens $U$.

Next, we analyze equilibrium point $P_2^{LUV2} = (0, K, 0, 0, Kb\sigma_2/\gamma)$. This point describes a state in which lysogen strain $L_2$ (with the lower SPI rate) displaces both the nonlysogens $U$ and lysogen strain $L_1$. The corresponding Jacobian is given by $J_2^{LUV2}$ in equation 31.

$$J_2^{LUV2} = \begin{pmatrix} -\sigma_1 + \sigma_2 & 0 & 0 & 0 & 0 \\ -r + \sigma_1 & -r + \sigma_2 & -g + (\alpha b K \sigma_2)/\gamma & 0 & 0 \\ 0 & 0 & g - r + \sigma_2 - (\alpha b K \sigma_2)/\gamma & 0 & 0 \\ b\sigma_1 & 0 & 0 & -\gamma & 0 \\ 0 & b\sigma_2 & -((\alpha b^2 K(-1+p)\sigma_2)/\gamma) & 0 & -\gamma \end{pmatrix} \tag{31}$$

The corresponding characteristic equation is given by:

13

$$0 = (\gamma + \lambda)^2 (\lambda + r - \sigma_2)(\lambda + \sigma_1 - \sigma_2)(\lambda - g + r - \sigma_2 + \alpha bK\sigma_2/\gamma) \tag{32}$$

The eigenvalues are given by $\lambda_1, \lambda_2 = -\gamma$, $\lambda_3 = -r + \sigma_2$, $\lambda_4 = -\sigma_1 + \sigma_2$, and $\lambda_5 = (g - r) - \sigma_2 (\alpha bK/\gamma - 1)$. Eigenvalues 1 - 4 are obviously negative since $\gamma > 0$, $\sigma_1 > \sigma_2$, and $\sigma_1, \sigma_2 < r$. The inequality $\sigma_2 > \sigma_{LB}$ can be rewritten as $0 > (g - r) - \sigma_2 (\alpha bK/\gamma - 1)$ which shows that $0 > \lambda_5$. Therefore, all eigenvalues are negative and therefore equilibrium point $P_2^{LUV2} = (0, K, 0, 0, Kb\sigma_2/\gamma)$ is a stable equilibrium point of the system. Thus, this system will tend towards a state in which lysogen strain $L_2$ displaces lysogen strain $L_1$ and the nonlysogens $U$.

Finally, we examine equilibrium point $P_3^{LUV2} = (0, 0, K, 0, 0)$ which describes a state in which both lysogen strain $L_1$ and $L_2$ are displaced by the nonlysogens $U$. The Jacobian for this equilibrium point is given by $J_3^{LUV2}$.

$$J_3^{LUV2} = \begin{pmatrix} -g + r - \sigma_1 & 0 & 0 & \alpha Kp & 0 \\ 0 & -g + r - \sigma_2 & 0 & 0 & \alpha Kp \\ -r + \sigma_1 & -r + \sigma_2 & -g & -\alpha Kp & -\alpha Kp \\ b\sigma_1 & 0 & 0 & -\gamma - \alpha bK(-1 + p) & 0 \\ 0 & b\sigma_2 & 0 & 0 & -\gamma + \alpha bK(1 - p) \end{pmatrix} \tag{33}$$

The characteristic equation is given by:

$$0 = (g + \lambda)$$
$$\times \left[ \lambda^2 + \lambda(\gamma + g - r + \sigma_1 - \alpha bK(1 - p)) + \gamma(g - r + \sigma_1) - \alpha bK((1 - p)(g - r) + \sigma_1) \right]$$
$$\times \left[ \lambda^2 + \lambda(\gamma + g - r + \sigma_2 - \alpha bK(1 - p)) + \gamma(g - r + \sigma_2) - \alpha bK((1 - p)(g - r) + \sigma_2) \right]$$
$$\tag{34}$$

The first eigenvalue is given by $(g + \lambda) = 0 \Rightarrow \lambda_1 = -g < 0$. The remaining eigenvalues are given by solving the quadratic equations within both square brackets $[\cdot]$. These eigenvalues are $\lambda_{2,3} = q_1 \pm \sqrt{q_1^2 + \varepsilon_1}$ and $\lambda_{4,5} = q_2 \pm \sqrt{q_2^2 + \varepsilon_2}$, with $q_j = (-g - \gamma + \alpha bK(1 - p) + r - \sigma_j)/2$ and $\varepsilon_j = -g\gamma + \alpha bK((g - r)(1 - p) + \sigma_j) + \gamma r - \gamma \sigma_j$ for $j = 1, 2$. If $\varepsilon_j > 0$ then it must be true that at least one of the eigenvalues is positive, implying that this equilibrium point is not a stable steady state. It is straightforward to show that $\varepsilon_j < 0$ only if we have $\sigma_j < -(g - r)\left[\alpha bK(1 - p)/\gamma - 1\right] / \left[\alpha bK/\gamma - 1\right] < 0$. However, we require $\sigma_j > \sigma_{LB} > 0 \Rightarrow \varepsilon_j > 0$ and hence this equilibrium point is unstable. Thus, the system will not tend towards a state in which the nonlysogens dipsplace both lysogen strains $L_1$ and $L_2$.

Overall these results collectively show that $L_2$ will displace both $L_1$ and $U$ if $\sigma_{LB} < \sigma_2 < \sigma_1$. The upper half of the inequality $\sigma_2 < \sigma_1$ gives $L_2$ a competitive advantage against $L_1$, and the lower half of the inequality $\sigma_{LB} < \sigma_2$ gives $L_2$ the competitive advantage against the nonlysogens. Thus, a lower SPI rate is advantageous over a higher SPI rate in the long run.

14

# 6 Stochastic factors influencing selection of SPI rate

To understand how stochastic factors could influence the competition between lysogens and nonlysogens, we simulated the LUV model with some modifications. Since the SPI rate is fairly small, at low lysogen counts (e.g. $L \approx 100$ or less) it is possible for several generations to pass before a single SPI reaction takes place. This is difficult to capture in the LUV model because it is simulated using ODEs which are deterministic. In this deterministic setting, some SPI reactions occur from the start of the dynamics since the net rate of SPI, $L \cdot \sigma$, is $> 0$ at $t = 0$. Thus, we modified our LUV model so that the SPI reactions would occur stochastically during the dynamics. Over an small time interval $\Delta t$ over which we integrate, the probability that a single lysogen undergoes SPI is $\sigma \cdot \Delta t$. For $L$ lysogens, the probability that $\ell \leq L$ of them undergo SPI in $\Delta t$ is given by a binomial distribution as in equation 35 with the average number of SPI events in a time interval $\Delta t$ given by $\bar{\ell} = L\sigma\Delta t$.

$$Pr(\ell) = \binom{L}{\ell} [\sigma\Delta t]^L [1 - \sigma\Delta t]^{L-\ell} \tag{35}$$

In Figure 3 we show 2 sample simulations of our model at an initial lysogen count of $L_0 = L(0) = 0.001 \cdot K$. In the right panel we observe the lysogens being slowly displaced by the nonlysogens (since $r < g$) but when an SPI reaction occurs this "saves" the lysogens and they end up taking over the population. However, it can also happen that the SPI reaction does not occur at all. If an SPI reaction fails to fire early enough (as in the left panel), the probability that an SPI reaction will fire later approaches 0. This is because the probability of an SPI reaction is proportional to $L \cdot \sigma$, and since $\sigma$ is already a very low number ($\approx 10^{-6}$) having low $L$ makes this probability approach 0. Thus, if stochasticity plays a major role (e.g. at low population counts) and if $r < g$, then the total SPI rate, $L \cdot \sigma$, must not be too low if the lysogens are to displace the nonlysogens.

At high initial lysogen counts, the lysogens almost always displaced the nonlysogens because an SPI event almost always occurs. On the other hand, when the initial lysogen count is low (e.g. 10 or 100) then they often were displaced by the nonlysogens. A similar trend was observed for $\sigma$, namely that high $\sigma$ allowed the lysogens to displace the nonlysogens more frequently relative to the case of a very low $\sigma$ value. Additionally, if the lysogenic growth-cost was small (e.g. $g \approx r$), then the lysogens were able to displace the nonlysogens with greater probability over a much wider range of $L(0)$ and $\sigma$ values. To systematically understand the effect of $L(0)$, $\sigma$, and $r$ on the probability that the lysogens displace the nonlysogens, we estimated the probability of lysogen-sweeping events from repeated runs of the stochastic simulation. In Figure 4 we show the probability that $L = K, U = 0$ over a range of $L(0)$ and $\sigma$ values for $r = 0.9$ and $r = 0.999$. Generally we see that at low $L(0)$ and $\sigma$ values the lysogens usually cannot displace the nonlysogens. On the other hand, at higher $L(0)$ and $\sigma$ the lysogens were always able to more probably displace the nonlysogens. For intermediate values of $L(0)$ and $\sigma$, the probability that the lysogens displace the nonlysogens ranges from 0 to 1. But as $r$ approaches $g$ (right panel with $r = 0.999$), we see that the lysogens can displace the nonlysogens over a much wider range of $L(0)$ and $\sigma$ values, even if $L(0)$ and $\sigma$ are low. When $g - r = 0$, the lysogens always displace the nonlysogens even in stochastic simulations. This is because the two populations remain at their initial values until the first SPI reaction occurs, no matter how long it takes. Thus, if the lysogenic growth-cost
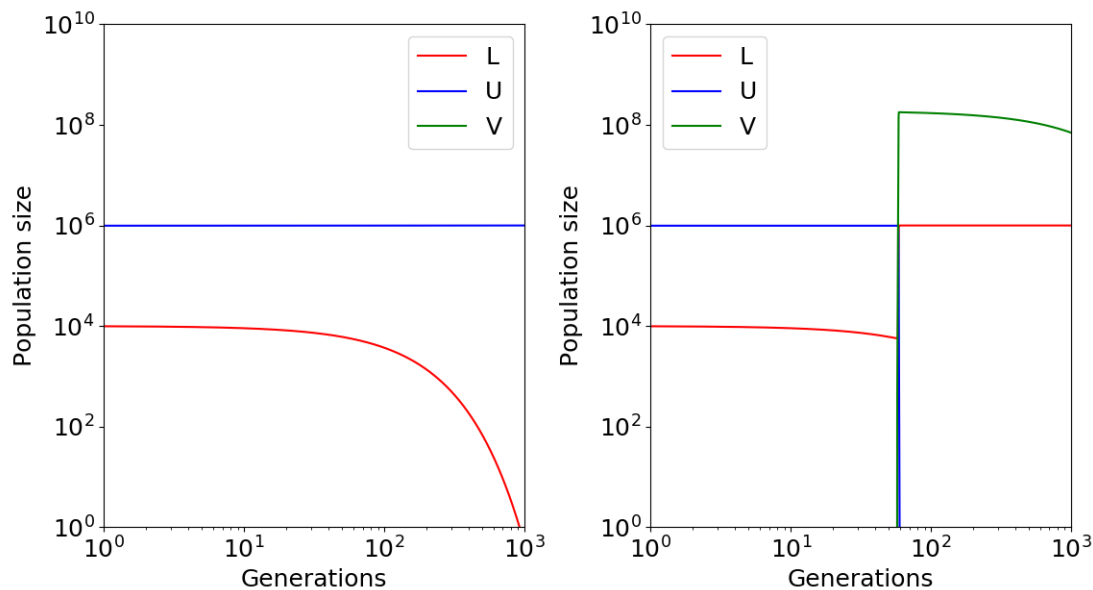
Figure 3: Stochastic LUV model simulations using the parameters from Table 1. In the left panel, the nonlysogens displaced the lysogens because no SPI reaction occurred. In the right panel an SPI reaction occurred early enough for the lysogens to displace the nonlysogens. We used parameter values from Table 1 with $L(0) = 0.001 \cdot K = 1000$. We also set $r = 0.99$, $\sigma_1 = 10^{-5}$, and $\sigma_2 = 10^{-6}$.

is small or zero, we expect that the lysogens will generally displace the nonlysogens with high probability.

In order for the lysogens to displace the nonlysogens, the SPI reaction must occur before the time $T$ at which the lysogen population drops to 0. To understand how this time $T$ depends on $g - r$ and $L_0$, we analytically calculated the latest time an SPI reaction must occur by solving the LUV model (equations 15 - 5) with $\sigma = 0$ for $L(t)$. The result is given as in equation 36.

$$L(t) = \frac{K}{1 + \left(\frac{K-L_0}{L_0}\right) e^{(g-r)t}} \tag{36}$$

The latest time $T$ that an SPI reaction can occur and still allow the lysogens to displace the nonlysogens can be calculated by setting $L(T) = 1$ and solving for $T$. We set $L(T)$ to 1 because by time $T$ there is only 1 lysogen left which can undergo SPI. The approximation in 37 is quite accurate since $K$ is many orders of magnitude greater than 1, as is $K/L_0$.

$$T = \frac{1}{(r-g)} \ln \frac{K/L_0 - 1}{K - 1} \approx \frac{1}{g-r} \ln L_0 \tag{37}$$

If an SPI reaction occurs by time $T$, then the lysogens will displace the nonlysogens. This establishes some additional constraints on the value of $\sigma$ necessary for the lysogens to displace the nonlysogens, namely that if there is a lysogenic growth cost then the SPI rate
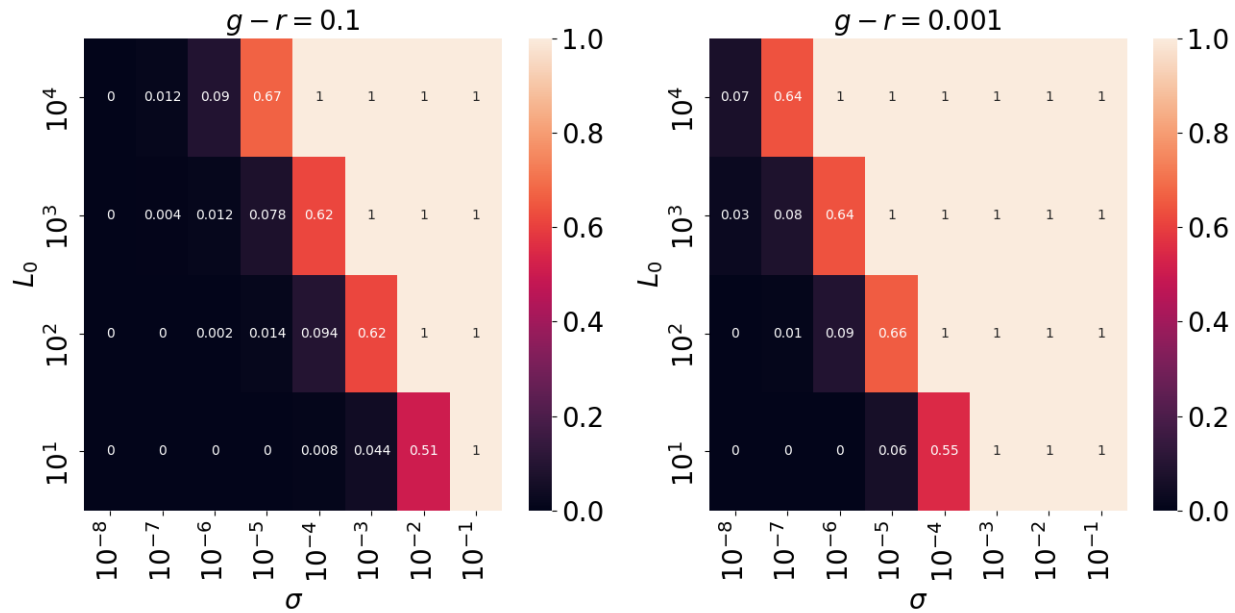
16

Figure 4: We calculated the fraction of simulations resulting in lysogens displacing the nonlysogens for different values of $L(0)$ and $\sigma$ using the stochastic LUV model. The left panel was simulated at $r = 0.9$ and the right panel was simulated at $r = 0.999$. The rest of the parameters were estimated from Table 1

cannot be arbitrarily low in a stochastic setting. In Figure 5 we show the $T$ value at different values of $r - g$ and $L_0$. We see that $T$ gradually increases with $L_0$, implying that higher initial lysogen count gives the lysogens more time to undergo an SPI event. However, we see that as $r$ becomes closer to $g$ the value of $T$ increases dramatically. This means that the lysogens have a much longer time to initiate an SPI event when there is a small lysogenic growth rate cost. If the lysogenic growth-cost is arbitrarily small, then $r \approx g \Rightarrow T \to \infty$, implying that eventually an SPI reaction will occur and the lysogens will displace the nonlysogens with probability 1.

To understand how the probability, $\Psi$, that the lysogens will displace the nonlysogens depends on general parameter values (e.g. not just when $r \approx g$), we calculated the probability that at least 1 SPI reaction occurs within time $T$ by modeling our stochastic simulation analytically. First, we split up the time interval $0 \le t \le T$ into small subintervals denoted by $\Delta t_k$ of fixed size $\Delta t$ of which there are $N = T/\Delta t$. Within each of these intervals, the number of lysogens is denoted by $L_k$. The probability of at least 1 SPI reaction over $0 \le t \le T$ is equal to 1 minus the probability that 0 SPI reactions take place over $0 \le t \le T$. This is equivalent to 1 minus the probability that 0 SPI reactions occur within all $N = T/\Delta t$ subintervals $\Delta t_k$. In each subinterval, the probability of 0 SPI reactions is $[1 - \sigma \Delta t]^{L_k}$. The total number of subintervals is $N = T/\Delta t$, which implies that the probability of 0 SPI reactions over the entire interval $0 \le t \le T$ is $\prod_{k=1}^{N} [1 - \sigma \Delta t]^{L_k}$. This can be rewritten as $[1 - \sigma \Delta t]^{\Sigma L_k}$ where the sum $\Sigma$ ranges over the $L_k$ from $k = 1$ to $k = N = T/\Delta t$. Using $\Sigma L_k = N\bar{L}$ with $\bar{L} =$ time average of $L_k$, we find that $\Psi = (1 - \sigma \Delta t)^{\bar{L}T/\Delta t}$. As $\Delta t \to 0$ one

17

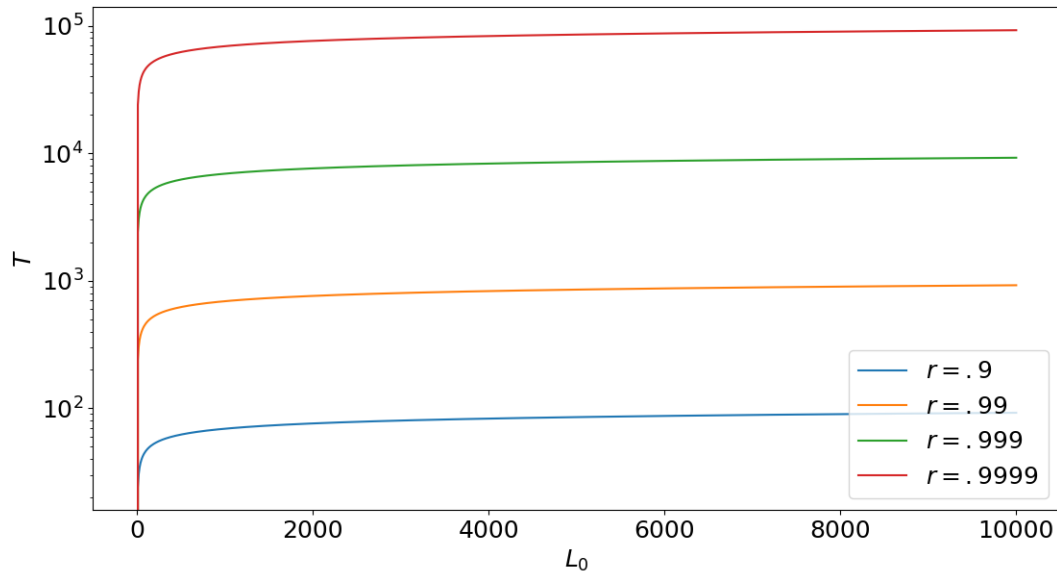Figure 5: The $T$ value calculated at different values of $r$ and $L_0$ using $K = 10^6$.

can easily show that $\Psi$ is given by equation 38.

$$\Psi = 1 - e^{-\sigma T \bar{L}} \tag{38}$$

Since we know $L(t)$ over time $T$ from equation 36, we can calculate $\bar{L} = \frac{1}{T} \int_0^T L(t)dt$. Applying straightforward integration techniques, we find that $\bar{L}$ is given by equation 39.

$$\bar{L} = \frac{-K}{T(g-r)} \ln \left[ 1 - \frac{\left(1 - e^{-(g-r)T}\right)}{K/L_0} \right] \tag{39}$$

We can then determine a final expression for $\Psi$ by combining equations 38, 37, and 39 to derive equation 40.

$$\Psi = 1 - \left[ \frac{K-1}{K-L_0} \right]^{-\sigma K/(g-r)} \tag{40}$$

Our expression for $\Psi$ matches the results of our stochastic simulation remarkably well as shown in Figure 6 using the same parameter values from Table 1. From this expression it is clear that the probability that the lysogens displace the nonlysogens increases with $L_0$, $\sigma$, and $K$. There is also a strong dependence on $g - r$. If the lysogenic-growth cost is very small then $g - r \to 0$ and, as a result, $\Psi \to 1 - 0 = 1$. Overall, the lysogens can maximize their probability of displacing nonlysogens by maximizing the quantity $\sigma/(g - r)$. The value of $\sigma$ can be arbitrarily small so long as the lysogenic growth-cost (e.g. $g - r$) is correspondingly small (which it is expected to be). These insights from our analytical models agree with our stochastic simulations. It is worth noting that when $r \approx g$ the simulations are extremely time consuming, especially at low $L_0$, because both the lysogen and nonlysogen population

18

remains approximately constant in time until finally an SPI event inevitably occurs since $\sigma > 0$. However, from our analytical results we can immediately see the effect of various parameters on $P$ without simulation. We take advantage of this by using $\Psi$ to calculate the probability the lysogens displace the nonlysogens over a range of $L_0$ and $\sigma$ values for $g - r = 10^{-5}$. Using the stochastic LUV model for the same calculation would have taken 10s of hours or longer.
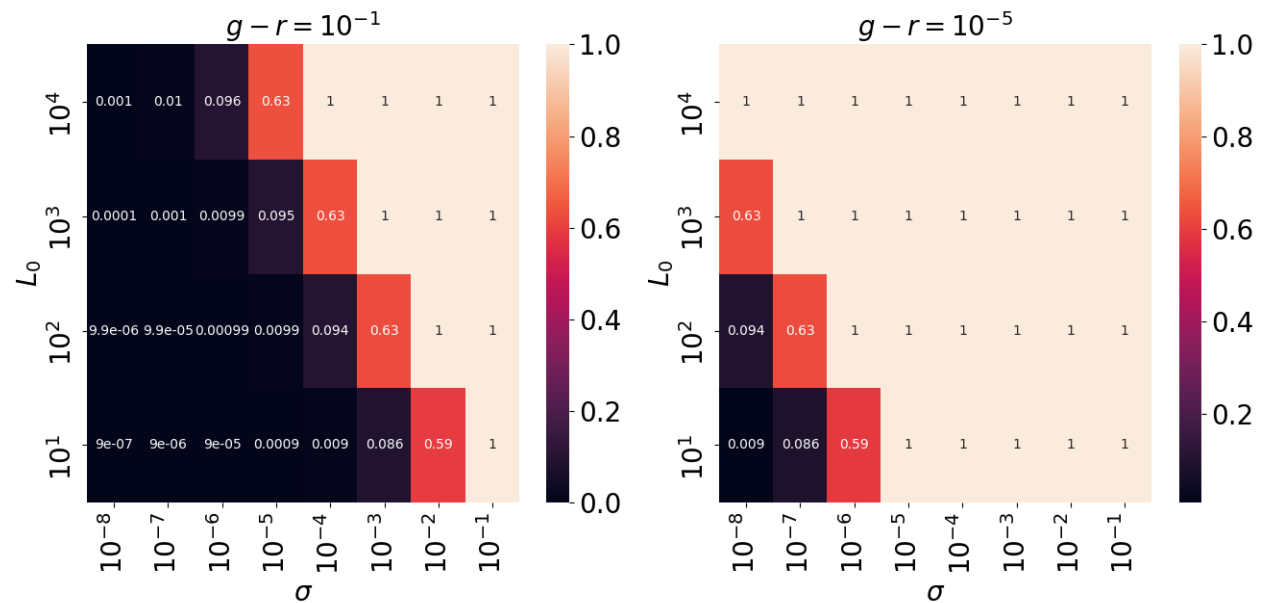


Figure 6: Probability that the lysogens displace the nonlysogens using the $\Psi$ function. Left panel matches with stochastic simulation results in Figure 4, and right panel shows results when $g - r$ is very small (e.g. $10^{-5}$) the lysogens are more likely to displace the nonlysogens for any $L_0$ and $\sigma$.

To understand how the stochastic evolutionary dynamics occurs in the case that there is a single initial lysogen, we created a separate model because our prior models don't adequately capture the case when $L_0 = 1$. In this case, when $t = 0$ we have that $V = 0$, $U = K - 1$, and $L = 1$ so the dynamics of $L$ can be described by $\dot{L} = (r - \sigma - g)[1 - 1/K] \approx (r - \sigma - g)$ using 1. The approximation holds since $1/K$ is very small compared to 1. This expression for $\dot{L}$ shows that the initial rate of change for $L$ is made of two parts. The $g - r$ term corresponds to the competitive growth of the lysogens vs the nonlysogens. The second term $\sigma$ corresponds to the SPI reaction. If $g > r$ then there are only two possibilities for the next reaction, namely i) either the nonlysogens will displace the nonlysogens by outgrowing them or ii) an SPI event will occur and the lysogens will displace the nonlysogens. The time to the next reaction in typical Gillespie fashion is given by $\tau = -\ln[1 - R]/(g - r + \sigma)$ with $R$ as a random number from a uniform distribution ranging from 0 to 1. The probability that the next reaction at time $\tau$ is the SPI reaction is simply given by the proportion $\sigma/(g - r + \sigma)$ and can be written as in equation 41.

19

$$\Psi_{\{L_0=1\}} = \frac{\sigma/(g-r)}{1+\sigma/(g-r)} \tag{41}$$

From equation 41 we clearly see that in the case $L_0 = 1$ the probability the lysogens displace the nonlysogens approaches 1 as $r \to g$. Thus, at little to no lysogenic growth-cost, the lysogens can displace the nonlysogens with high probability even when there is only 1 initial lysogen. As in equation 40, the higher the ratio $\sigma/(g-r)$ is, the more likely the lysogens displace the nonlysogens. It is easy to show that $\Psi_{\{L_0=1\}} \geq 0.5$ if $\sigma \geq g-r$. Thus, with $\sigma = 10^{-5}$ the lysogens have a 50% chance of winning if the lysogenic growth-cost is on the order of $10^{-5}$. If $g = r$ then in the model both $L$ and $U$ remain at their initial population sizes until an SPI reaction fires, and so the lysogens win with probability $\Psi = 1$ in this case. This is easily shown by letting $g \to r$ and taking a limit of equation 41.

To understand how stochastic factors influence the competition between two lysogen strains, we modified the LUV2 model to be stochastic in the same way as in the LUV model. The insights we gained were similar to the case of the LUV model with one extra detail. As shown in Figure 7, we see that so long as either one of the lysogens undergoes SPI then $L_2$ will always displace $L_1$ in the long run. This is because when the SPI event initially occurs, the nonlysogens start to be killed off, causing empty space to appear for both $L_1$ and $L_2$ to both grow into. Once both $L_1$ and $L_2$ fill this empty space and the nonlysogens disappear, it is simply a competition between only $L_1$ and $L_2$. From this point onward it is clear that $L_2$ has the advantage because its effective growth rate $r - \sigma_2$ is slightly greater than the effective growth rate $r - \sigma_1$ of $L_1$ (since $\sigma_2 < \sigma_1$). The $V_1$ curve jumps up and down a bit towards the end because the stochasticity of SPI reactions becomes more apparent as $L_1$ drops to lower levels. Notice that $L_1$ underwent an SPI reaction very early on (marked by the increase in $V_1$) while $L_2$'s first SPI event occured much later on at $\approx 10$ generations. Even though $L_2$ was late to SPI, it still wins in the long run since its effective growth rate is higher.

# 7 Discussion

In this work we explore how SPI influences the natural selection among lysogens and nonlysogens, and also between lysogen variants with different SPI rates. We find that there exists a lower bound $\sigma_{LB} \approx 10^{-6}$ (or smaller) above which lysogens are expected to displace the nonlysogens for general parameter values. Furthermore, we find that lower SPI rates are naturally selected, because although higher SPI rates confer an early advantage they are costly to growth in the long run. These results collectively show that natural selection should push the SPI rate down towards $\sigma_{LB}$, and that this can be viewed as an optimal SPI rate in terms of natural selection. It is plausible that there does exist a lysogenic growth-cost (i.e. $g - r > 0$) but it is very likely that this growth-cost is very small such that $g - r \approx 10^{-n}$ for some large integer $n$. Thus, there likely is a lower bound given by $\sigma_{LB}$ that is strictly greater than 0, but it is likely to be very small.

If stochasticity plays a major role in the dynamics, then our same conclusions hold in a probabilistic sense, with lysogenic fixation becoming increasingly probable at higher SPI rates. Thus, stochastic effects on natural selection would tend to favor larger SPI rates. In the stochastic setting, however, the SPI rate is scaled by the lysogenic growth-cost, implying
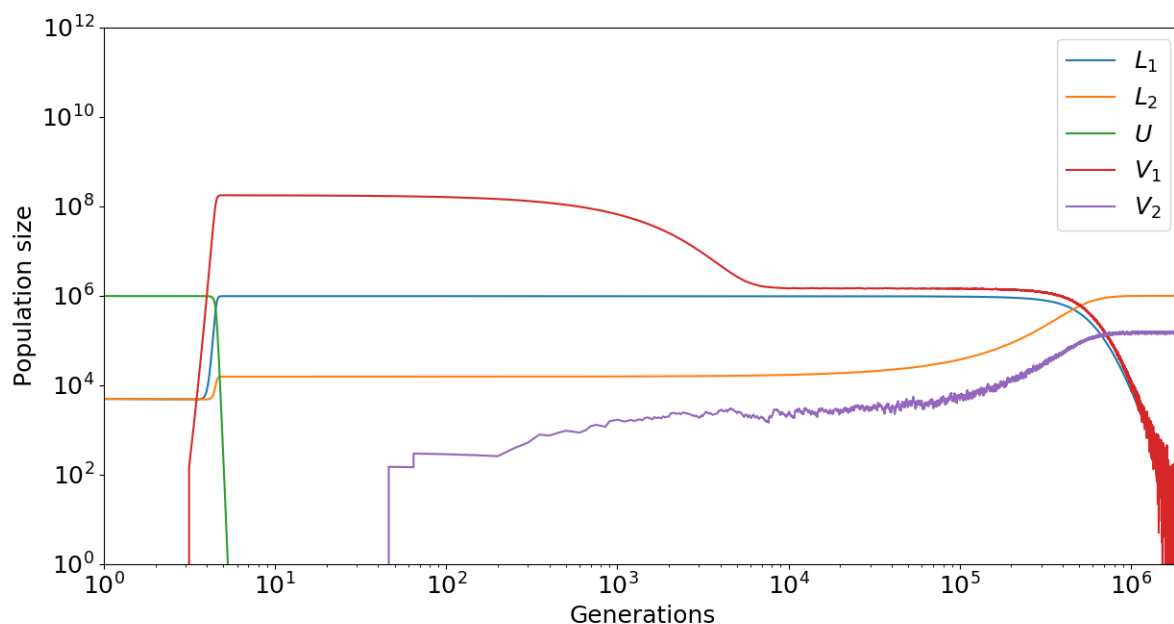
20

Figure 7: A sample simulation from the stochastic LUV2 model. Early on $L_1$ manages to have an SPI reaction, causing $V_1$ to rise and $U$ to be killed off. During this, $L_2$ begins to grow since there is now free space to do so. It does not have its own SPI event until about 10 generations. Eventually, the relatively higher SPI rate of $L_1$ causes it to slowly but surely drop to 0 as $L_2$ rises towards the carrying capacity $K = 10^6$. This simulation was run with $r = 0.99$, $\sigma_1 = 10^{-5}$, and $\sigma_2 = 10^{-6}$. We observed the same qualitative behavior for other parameter values.

that at very low lysogenic growth-cost the SPI rate can remain arbitrarily low while still maintaining a high probability of lysogenic fixation. In the extreme case that there is only 1 initial lysogens, the probability they will fix in the population is $\approx 1$ if the lysogenic growth-cost is on the order of $10^{-6}$ at the naturally occuring SPI rate $\sigma^* \approx 10^{-6}$.

Our estimate on the lower bound of the SPI rate is on the order of $10^{-6}$ or $10^{-7}$ (or smaller according to $g - r$) lysogens per generation, which matches experimental estimates of the $recA^+$ SPI rate [11, 12, 29]. We show that deterministic evolutionary forces are pushing the SPI rate down towards $\sigma_{LB}$, but stochastic factors tend to push it higher only if $g - r$ is not small. Otherwise, if $g - r$ is very small then $\sigma$ can also be very small and still allow natural selection of lysogens over nonlysogens with high probability. Our theory suggests that phage $\lambda$ may be sitting at or near this optimal SPI rate.

Our results demonstrate how mathematical modeling can be effectively used to analyze evolutionary scenarios and explore the consequences of ecological constraints and interactions among competing species. Our theory shows how by sacrificing a little bit of growth rate, a species can gain a competitive advantage over another. This has connections to bacterial persistence, sporulation, and other cell growth strategies in which a small subpopulation switches to a growth-reduced but phenotypically different state. It will be interesting to validate our model further by applying it to other phages and their hosts such as P1, T4,

21

and Mu once all the model's parameters are measured for these phages.

# 8    Acknowledgements

# References

[1] Uri Alon. *An introduction to systems biology: design principles of biological circuits.* Chapman and Hall/CRC, 2006.

[2] Joseph Bondy-Denomy and Alan R Davidson. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. *Journal of Microbiology*, 52(3):235–242, 2014.

[3] Lionello Bossi, Juan A Fuentes, Guido Mora, and Nara Figueroa-Bossi. Prophage contribution to bacterial population dynamics. *Journal of bacteriology*, 185(21):6467–6471, 2003.

[4] Michael G Cortes, Jimmy T Trinh, Lanying Zeng, and Gábor Balázsi. Late-arriving signals contribute less to cell-fate decisions. *Biophysical journal*, 113(9):2110–2120, 2017.

[5] Michael M Cox, Myron F Goodman, Kenneth N Kreuzer, David J Sherratt, Steven J Sandler, and Kenneth J Marians. The importance of repairing stalled replication forks. *Nature*, 404(6773):37, 2000.

[6] Marianne De Paepe and François Taddei. Viruses' life history: towards a mechanistic basis of a trade-off between survival and reproduction among phages. *PLoS biology*, 4(7):e193, 2006.

[7] João Alves Gama, Ana Maria Reis, Iolanda Domingues, Helena Mendes-Soares, Ana Margarida Matos, and Francisco Dionisio. Temperate bacterial viruses as double-edged swords in bacterial warfare. *PloS one*, 8(3):e59043, 2013.

[8] William V Howes. Protein synthesis in escherichia coli: a phage-mediated interruption of translation. *Biochimica et biophysica acta*, 103(4):711, 1965.

[9] P Kourilsky and A Knapp. Lysogenization by bacteriophage lambda: Iii.-multiplicity dependent phenomena occuring upon infection by lambda. *Biochimie*, 56(11-12):1517–1523, 1975.

[10] Philippe Kourilsky. Lysogenization by bacteriophage lambda. *Molecular and General Genetics MGG*, 122(2):183–195, 1973.

[11] John W Little and Christine B Michalowski. Stability and instability in the lysogenic state of phage lambda. *Journal of bacteriology*, 192(22):6064–6076, 2010.

[12] John W Little, Donald P Shepley, and David W Wert. Robustness of a gene regulatory circuit. *The EMBO journal*, 18(15):4299–4307, 1999.

[13] Xiaoqiu Liu, Huifeng Jiang, Zhenglong Gu, and Jeffrey W Roberts. High-resolution view of bacteriophage lambda gene expression by ribosome profiling. *Proceedings of the National Academy of Sciences*, 110(29):11928–11933, 2013.

[14] Jonathan Livny and David I Friedman. Characterizing spontaneous induction of stx encoding phages using a selectable reporter system. *Molecular microbiology*, 51(6):1691–1704, 2004.

[15] Joanna M Los, Marcin Los, Alicja Wegrzyn, and Grzegorz Wegrzyn. Altruism of shiga toxin-producing escherichia coli: recent hypothesis versus experimental results. *Frontiers in cellular and infection microbiology*, 2:166, 2013.

[16] Jesse D McCool, Edward Long, Joseph F Petrosino, Hilary A Sandler, Susan M Rosenberg, and Steven J Sandler. Measurement of sos expression in individual escherichia coli k-12 cells using fluorescence microscopy. *Molecular microbiology*, 53(5):1343–1357, 2004.

[17] R Moldovan, E Chapman-McQuiston, and XL Wu. On kinetics of phage adsorption. *Biophysical journal*, 93(1):303–315, 2007.

[18] Arun M Nanda, Kai Thormann, and Julia Frunzke. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. *Journal of bacteriology*, 197(3):410–419, 2015.

[19] Martin A Nowak. *Evolutionary dynamics*. Harvard University Press, 2006.

[20] Amos B Oppenheim, Oren Kobiler, Joel Stavans, Donald L Court, and Sankar Adhya. Switches in bacteriophage lambda development. *Annu. Rev. Genet.*, 39:409–429, 2005.

[21] Jeanine M Pennington and Susan M Rosenberg. Spontaneous dna breakage in single living escherichia coli cells. *Nature genetics*, 39(6):797, 2007.

[22] Yili Qian, Hsin-Ho Huang, José I Jiménez, and Domitilla Del Vecchio. Resource competition shapes the response of genetic circuits. *ACS synthetic biology*, 6(7):1263–1272, 2017.

[23] Tatenda Shopera, Lian He, Tolutola Oyetunde, Yinjie J Tang, and Tae Seok Moon. Decoupling resource-coupled gene expression in living cells. *ACS synthetic biology*, 6(8):1596–1604, 2017.

[24] François St-Pierre and Drew Endy. Determination of cell fate selection during phage lambda infection. *Proceedings of the National Academy of Sciences*, 105(52):20705–20710, 2008.

[25] Mario Terzi and Cyrus Levinthal. Effects of $\lambda$-phage infection on bacterial synthesis. *Journal of molecular biology*, 26(3):525–535, 1967.

[26] Xiaoxue Wang, Younghoon Kim, Qun Ma, Seok Hoon Hong, Karina Pokusaeva, Joseph M Sturino, and Thomas K Wood. Cryptic prophages help bacteria cope with adverse environments. *Nature communications*, 1:147, 2010.

[27] Xuefang Xu, Sean P McAteer, Jai J Tree, Darren J Shaw, Eliza BK Wolfson, Scott A Beatson, Andrew J Roe, Lesley J Allison, Margo E Chase-Topping, Arvind Mahajan, et al. Lysogeny with shiga toxin 2-encoding bacteriophages represses type iii secretion in enterohemorrhagic escherichia coli. *PLoS pathogens*, 8(5):e1002672, 2012.

[28] Lanying Zeng, Samuel O Skinner, Chenghang Zong, Jean Sippy, Michael Feiss, and Ido Golding. Decision making at a subcellular level determines the outcome of bacteriophage infection. *Cell*, 141(4):682–691, 2010.

[29] Chenghang Zong, Lok-hang So, Leonardo A Sepúlveda, Samuel O Skinner, and Ido Golding. Lysogen stability is determined by the frequency of activity bursts from the fate-determining gene. *Molecular systems biology*, 6(1):440, 2010.