A mathematical model of generating a mutant strain of a lysogenic bacteriophage

Cameron Haynes
December 12, 2022

Abstract

Eliminating bacterial infections that develop immunity in response to bacteriophage (henceforth referred to as "phage") therapy requires a therapeutic strategy that evolves in response. Directed evolution of a mutant strain of the therapeutic phage may provide one way of circumventing bacterial immunity. Knowing the selective pressures at work, how should one modify the environment of the phage and its host to accomplish this on a feasible timescale? We develop a mathematical model of a host cell and an infecting lysogenic phage that allows for natural selection, depending on parameter values. In particular, when the induction rate for the lysogen population was 2 per generation, the system approached a steady state where only the original viral strain, along with the susceptible host cell population, was present. However, when the induction rate was lowered to 1 per generation , the system evolved toward a steady state where the host population was driven to extinction by a new strain of phage. This suggests experimental protocols that may aid in generating new mutant strains of lysogenic phages for the purpose of treating difficult infections.

Introduction

As the ubiquitous predators of bacteria, bacteriophages are fundamental in shaping life on Earth. Phages are viruses that infect bacteria with one of two life cycles: the lytic cycle, and the lysogenic cycle. In the lytic cycle, the phage uses its host cell's machinery to rapidly make many copies of itself and then bursts the cell, releasing these new copies to produce yet more infections. The lysogenic cycle employs a more subtle approach, where the phage's genome inserts itself into its host's genome: then, as the host replicates itself and reproduces, the phage is also replicated and passed on. If a host cell that has been infected using the lysogenic cycle comes under threat, the infecting phage can enter the lytic cycle to quickly make more copies of itself and use the host machinery before the host dies: this is called induction, and it will play an important role in this paper. Not all phages can use the lysogenic cycle: if they can, they can be referred to as *lysogenic phage*, and if they can not, they can be referred to as *lytic phage*.

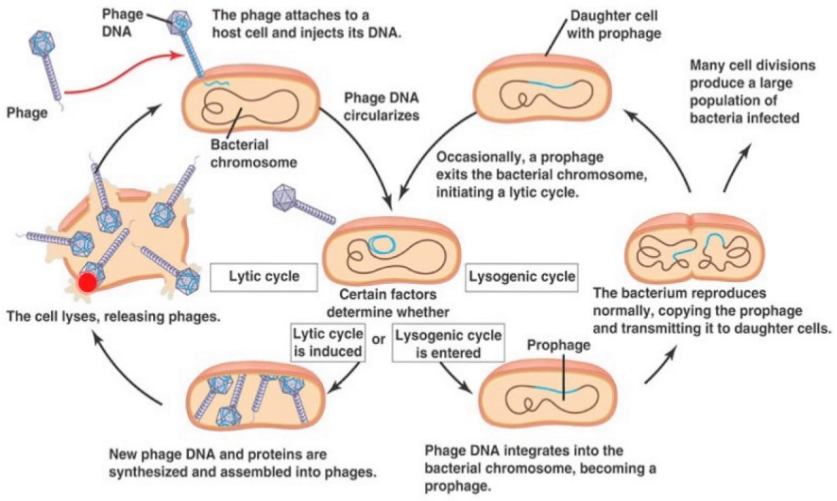


Figure 1: The lytic and lysogenic cycles of a bacteriophage graphically depicted. Note the ability of viruses in the lysogenic cycle to switch to lytic behavior. Figure Credit: https://www.youtube.com/watch?v=QNKlE4IsDME

Phages mold environments as diverse as the ocean and the human gut, and their role in the latter is receiving a lot of attention from researchers. The human gut hosts a very complex mixture of microbial species -- the microbiome -- and the viruses that predate upon them act to keep their populations regulated and in balance. Disease can occur when these fall out of equilibrium due to diet or the use of antibiotics, a devastating instance being infection by an antibiotic-resistant bacterium. Causing troublesome or even life-threatening symptoms, these infections are difficult to treat and pose a staggering threat to global health: it is predicted that the global mortality per year caused by anti-microbial resistant organisms will be 10,000,000 by 2050 [1].

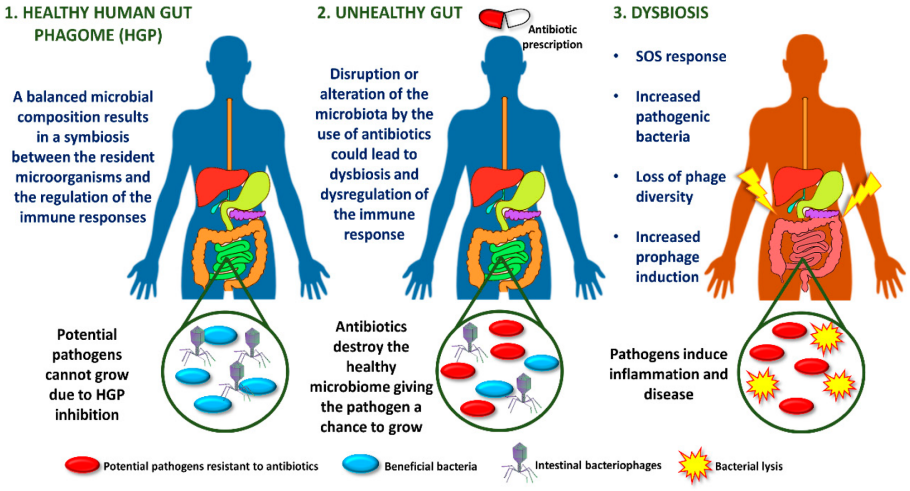


Figure 2: A graphical depiction of the development of dysbiosis in the gut. Once the gut has fallen into a state of dysbiosis, phage are no longer present in the numbers or diversity necessary to keep populations suppressed, and opportunistic pathogens can grow rapidly. Figure Credit: https://www.mdpi.com/2076-2607/8/9/1420

Medical research has recently become interested in the behavior of phage in the human gut. Application of virions capable of destroying antibiotic-resistant hosts seems to be a promising avenue for mitigating these infections and treating them when all other options fail [2]. This phage therapy has proven to be successful in some clinical instances, but it is stymied by the capacity for the evolution of immunity in the host cells. This resistance also occurs naturally, so phages have mechanisms for coevolution toward regaining virulence, but this may take too long for patients critically ill with a persistent infection [2]. Once immunity to phage infection spreads in the problematic population, speeding up this coevolutionary process to arrive a genetically novel strain of phage may help in eliminating the infection on a timescale more conducive to treatment, but the steps that must actually be performed to achieve this "speed up" are not very clear.

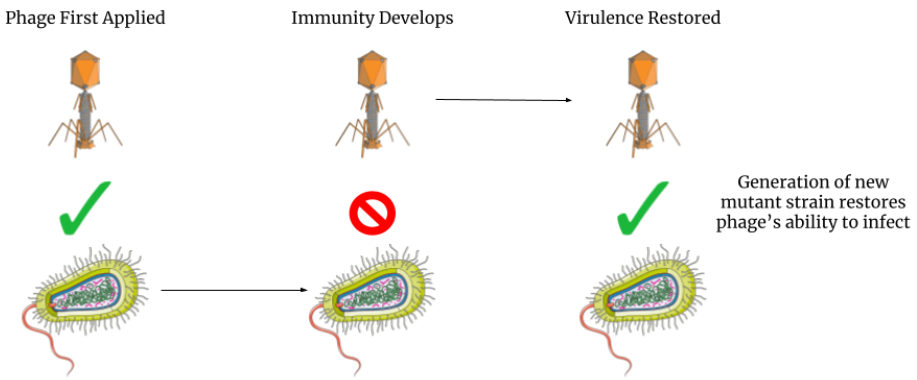


Figure 3: The restoration of virulence through coevolution. When bacterial hosts evolve immunity to a phage's infection, the phage can coevolve to regain the ability to infect. Can we make this faster?

How does a new, genetically distinct, variant of phage arise? For phages capable of lysogeny, this process is well-understood, so we will restrict our attention to lysogenic phage in this paper. To answer the question posed, we must briefly review the mechanisms by which lysogens are immune to superinfection by phages of the same genetic composition.

Cells that have been infected by a phage using the lysogenic cycle are immune to being infected by phage of the same type: this immunity from "superinfection" is due to the structure of the phage's genome. The genes used by the phage -- when induced -- to activate the lytic cycle are all regulated by the same operator. This operator is bound by a repressor protein that is always being expressed by phages while in the lysogenic cycle.

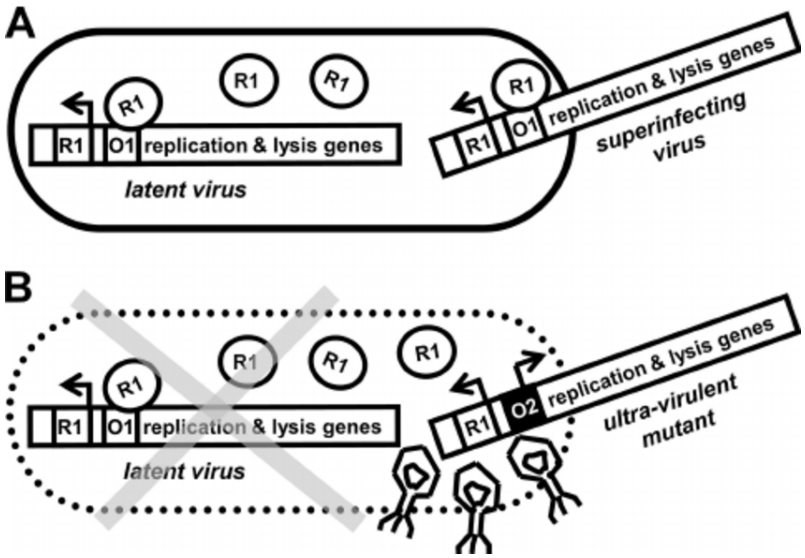


Figure 4: The λ phage operon for replication and lysis, and how it can mutate. The presence of repressor protein inside of all lysogens prevents other phage with the same operator and repressor protein from infecting. However, mutate the operator of the superinfecting phage (or, equivalently, its repressor protein), and it will be able to superinfect with abandon. Figure Credit: https://www.researchgate.net/figure/Superinfection-inhibition-and-its-avoidance-by-ultravirulence-the-example-of-phage_fig1_45365798

This has the indirect effect of preventing other phages from infecting the same cell, because the repressor protein that prevents transcription of the lytic cycle genes also binds to the operators of incoming viruses. Unless, of course, the operator of the superinfecting virus is mutated. In the instance that the operator -- or the repressor protein -- is mutated, the superinfecting virus is not bound by the repressor protein and can thus initiate an infection of its own: however, because the repressor protein can not bind to its operator, it can only use the lytic cycle, making it an ultra-virulent mutant.

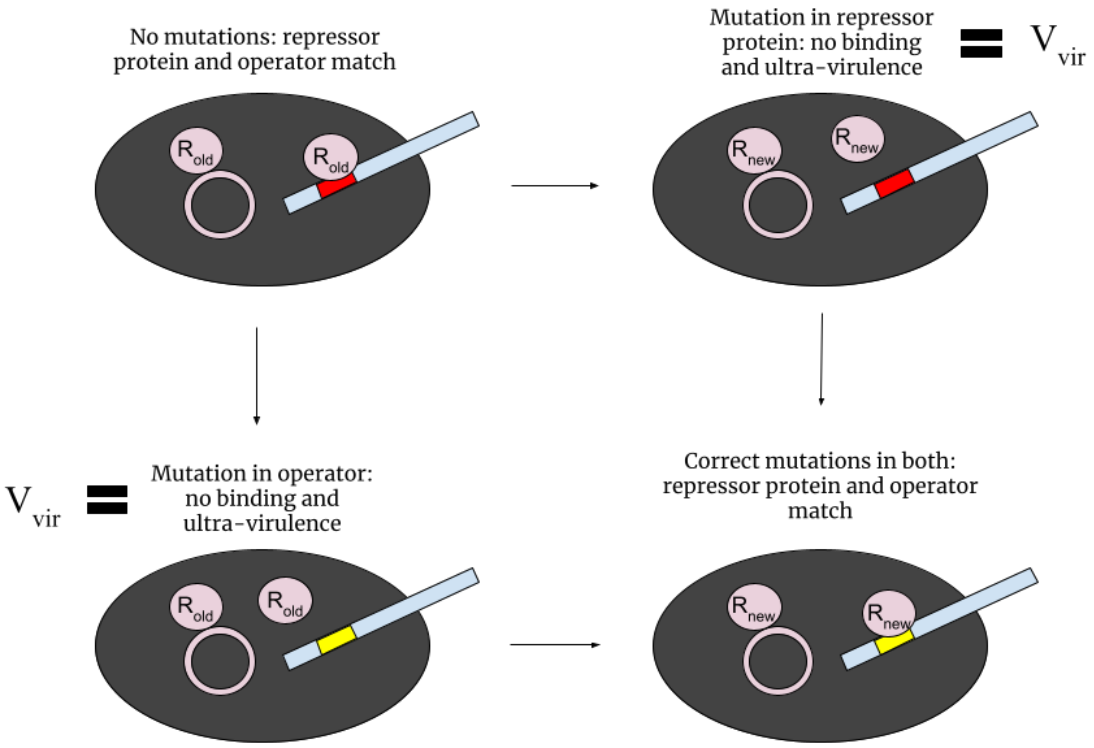


Figure 5: A commutative diagram showing the path from the original lysogenic strain to the new lysogenic strain. When there are no mutations, the repressor and operator match. When one of them has changed, there is no longer a match, and ultra-virulence results. To restore lysogeny, one of two things can occur: the first mutation can revert, or there can be a second mutation to make the repressor protein and operator match again.

Our work begins by adding on to a model, developed by Wahl et al, of a lysogenic phage -- a compartmental model that describes the behavior of the phage, its host, and its corresponding lysogens [3]. Once the model was written down, analysis showed that the outcome falls into one of three categories depending on the values of the model parameters used:

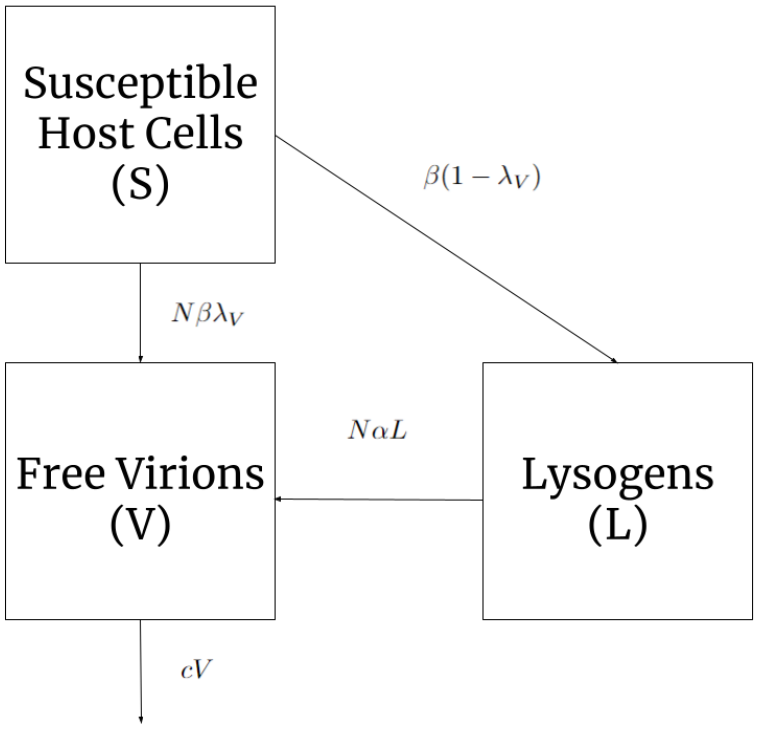
1. Extirpation of the virus and its lysogens
2. Extinction of the host cells and victory for the virus
3. Stable coexistence of the three populations

We want to take the same approach with our model. Taking our cue from Wahl et al., we rephrase the biological question about the generation of a new strain of the virus as a set of questions about the model and its behaviors: what choice of parameter values lead to the

persistence of infection? What choice of parameter values cause the original phage-host-lysogen system to go extinct with a new phage strain taking its place?

In the sections below, we begin by providing a brief review of the model of Wahl et al. and its long-term behavior under different parameter values. Then, we start building up the model by reviewing some of the biology pertinent to the mutational path a lysogenic phage must travel before it can become a genetically distinct, lysogenic strain. After developing the model in full and explaining the equations and their functional forms, we perform simulations to find sets of parameter values that match the scenarios above. Finally, we report on how changing one parameter, the induction rate, causes changes in the qualitative behavior of the model and discuss real-life analogues of changing this parameter to generate a mutant strain.

Mathematical Model Development



Wahl et al.'s model involves three compartments: one for the free virions of a lysogenic phage; another for the host bacterium population; and the last for the lysogens associated with the phage [3]. The model makes a few assumptions, such as

1. Logistic growth for the host and lysogen populations, which makes sense in an environment where resources are limited
2. The environment is well-mixed – this is a simplifying assumption to eliminate spatial effects and may be well justified in a turbulent environment
3. Lytic infections complete instantaneously, producing a new burst of virions – again, this is a simplifying assumption to avoid the complexity of time-delay equations.
4. Virions are washed away at a constant rate, c, but the cells are not washed away

These assumptions result in the following set of three equations, which have been slightly modified.

$$\Sigma = \text{Sum of the Cell Populations}$$
$$\nu = \text{Sum of the Viral Populations}$$
$$\frac{dS}{dt} = \text{Logistic Growth} - \text{Predation}$$
$$\frac{dV}{dt} = \text{Virions from Lytic Infection} + \text{Virions from Induction} - \text{Clearance}$$
$$\frac{dL}{dt} = \text{Logistic Growth} + \text{Lysogens from Infections} - \text{Induction}$$

where S is the host cell population in units of cells, V is the free virion population in units of virions, and L is the lysogen population in units of cells. One can see that the only source of new susceptible host cells is from their own growth, but lysogens are increased through new infections and are free from predation through their immunity to superinfection; however, lysogens must also contend with the decrease due to induction, which is assumed to be proportional to the number of lysogens in the population. These equations find their full form as

$$\Sigma = S + L$$
$$\nu = V$$
$$\frac{dS}{dt} = rS(1 - \frac{\Sigma}{K}) - \beta S\nu$$
$$\frac{dV}{dt} = N\beta(1 - \lambda_V)(\Sigma - L)V + N\alpha L - cV$$
$$\frac{dL}{dt} = rL(1 - \frac{\Sigma}{K}) + \beta\lambda_V(\Sigma - L)V - \alpha L$$

In terms of the parameters, K is the carrying capacity of the environment, β controls the rate of new infections, N is the burst size of each new lytic infection, α is the induction rate of the lysogens, c is the clearance rate at which virions are flushed out of the environment, and λ_V is the probability of a new infection entering the lysogenic cycle. Although there are likely numerous factors that cause small changes in the value of all these parameters, they are all assumed to be constant in this model for the sake of simplicity.

It may seem somewhat silly to use ν instead of V in some of the terms above or $\Sigma - L$ in place of S, but when the full model is shown, with all 8 equations, it makes things more compact and comprehensible to do so. It also gives greater meaning to the terms. For example, the term $N\beta(1 - \lambda_V)(\Sigma - L)V$ says that the number of new virions from the lytic cycle is proportional to

1. The total number of free virions of type V
2. The total number of cells that are not immune to infection by V (i.e those that aren't L)
3. The rate of new infections, β
4. The proportion of infections that are not lysogenic, $1 - \lambda_V$
5. The burst size, N

One doesn't have to know what all the cell types present in the environment are to know the number of new virions to be released: rather, one just has to avoid including the cell types that should be immune.

The functional form $\beta S\nu$ for predation arises from the well-mixed assumption: the more of the cell being predated upon and the more virions are present, the more infections should occur. The decay term $(-\alpha L)$ in the equation for the lysogen compartment has its functional form because the amount of induction occurring in the lysogen population is assumed to be directly proportional to the number of lysogens.

In the context of the problem posed in the Introduction, this model effectively describes the interaction of a lysogenic phage with an infection but it does not allow for mutation and natural selection. As stated in the introduction, when a lysogenic phage evolves into a new strain, it must pass through an ultra-virulent mutant stage. Let's write down the equation for this compartment.

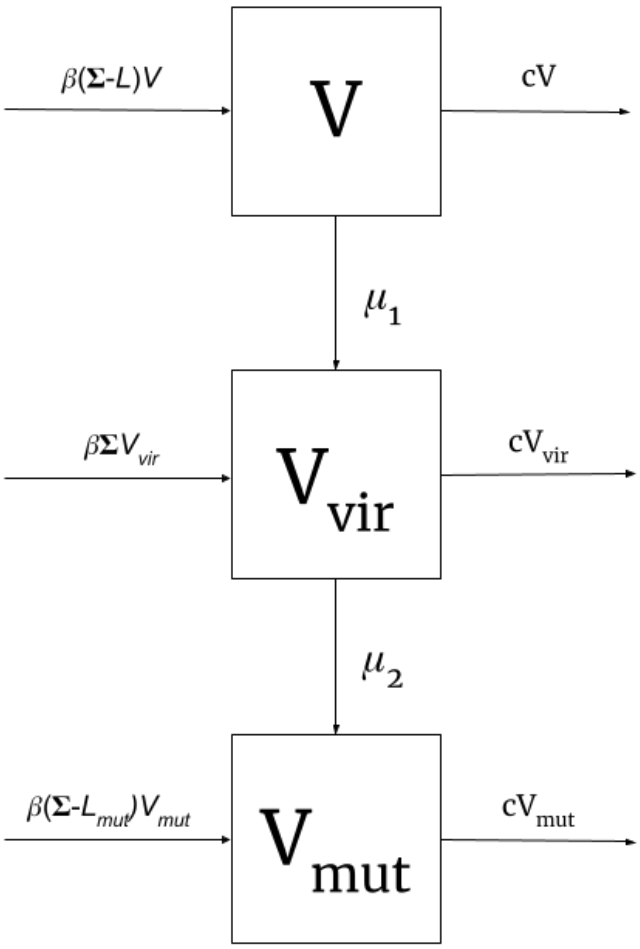
We will keep all of the assumptions from Wahl et al. and build off of their model. What are the key characteristics of the ultra-virulent mutant V_{vir}?

1. It can infect all cell types
2. It has no associated lysogen
3. V_{vir} originally arises from mutating V virions
4. The new mutant strain arises from further mutating V_{vir}

Using the word equation for V in the model of Wahl et al. as a template, we construct one for V_{vir}

$$\frac{dV_{vir}}{dt} = \text{Virions from Lytic Infection} - \text{Clearance} + \text{Mutation}$$

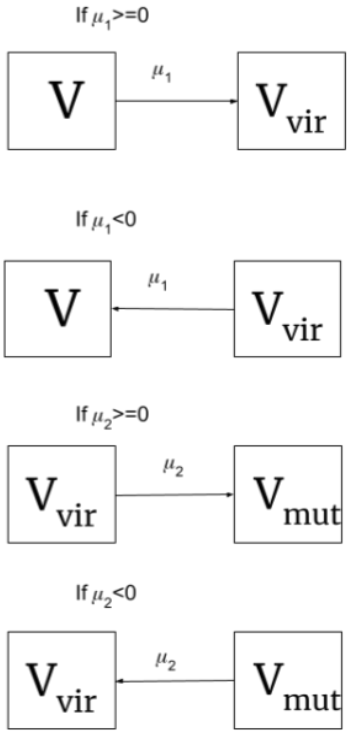
Now, there is no longer any term for induction, but there is a term for mutation. Let's look at a compartmental drawing of V_{vir} to better understand.



Letting V_{mut} be the virions of the new, mutant strain and L_{mut} be the associated lysogens of the new, mutant strain, we see that V_{vir} lay at the crossroads as an intermediate step. μ_1 is a rate which controls the selective pressure for ultra-virulence from the original strain, and μ_2 is a rate that controls the selective pressure for a restoration of lysogeny in the form of a new strain from the ultra-virulent state. Thus, the mutation term in V_{vir} has to include both the inflow from the V compartment and the outflow to the V_{mut} compartment.

Based on the circumstances in the system and the resulting selective pressures, it is possible for μ_1 and μ_2 to become negative. Normally, flow between compartments is handled through something like the following. If I and J are compartments, then if there is flow from I to J, then the equation for I will contain a term like $-\ kI$ and the equation for J will contain kI. However, if k becomes negative, this can cause I to become unboundedly positive in a

self-reinforcing way, and similarly, J can become unboundedly negative. This is not what we want. Instead, if k becomes negative, we want to instead shift the paradigm to say that, instead of a negative flow from I to J, there is just a flow from J to I: that is, there will be a term in the equation for J like $-\ kJ$ and there will be a term in the equation for I like kJ. How to accomplish all of this in one mutation term? The two figures below shed light on the subject.



$$\delta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\chi(x) = \begin{cases} 0 & x \geq 0 \\ 1 & x < 0 \end{cases}$$

The figure on the left is a pictorial reiteration of what was said in the above paragraph. The figure on the right contains the definitions for two functions: δ and χ. They are opposites, and when their argument is positive, any coefficient attached to δ will be fully expressed while any coefficient attached to χ will vanish; similarly, when their argument is negative, any coefficient attached to δ will vanish and any coefficient attached to χ will be fully expressed.

Now we're ready to see the mutation term. For the equation governing V_{vir}, the mutation term is

$$\text{Mutation} = \text{Mutation from } V \text{ to } V_{vir} - \text{Mutation from } V_{vir} \text{ to } V_{mut}$$
$$\text{Mutation} = \mu_1[\delta(\mu_1)V + \chi(\mu_1)V_{vir}] - \mu_2[\delta(\mu_2)V_{vir} + \chi(\mu_2)V_{mut}]$$

Note that, due to how δ and χ have been defined, only one summand in each of the brackets will be present for any given value of μ_1 or μ_2. This ensures that, as discussed above, flow only comes from non-empty compartments, and it goes in the right direction.

With the mutation term in hand and having the word equation template above, we can now write the equation governing V_{vir}.

$$\frac{dV_{vir}}{dt} = N\beta\Sigma V_{vir} - cV_{vir} + \mu_1[\delta(\mu_1)V + \chi(\mu_1)V_{vir}] - \mu_2[\delta(\mu_2)V_{vir} + \chi(\mu_2)V_{mut}]$$

Before moving on to the equations for the V_{mut} and L_{mut} compartments, we must provide the equations that govern μ_1 and μ_2 and the biology to justify them. There are many selective pressures at work in a real phage-host system, but for the sake of having a tractable model, here we assume there to only be two present: the selective pressure for ultra-virulence in order to predate upon lysogens and the selective pressure for sustainability.

As mentioned above, μ_1 is a quantity related to the selective pressure to mutate from the original strain to the ultra-virulent mutant state; what causes this selective pressure? What incentive does the virus have to become ultra-virulent? Virions in the V_{vir} compartment can predate on lysogens as well as the susceptible host population, so the more lysogens there are, the more of an advantage there is for virions in the V_{vir} compartment. Thus, we would expect flow to ultra-virulence to increase when there is a large enough ratio of lysogens to susceptible host cells in the environment, as the dominant food source is then only available to those virions that can superinfect.

We can phrase this mathematical intuition as the following: when $(L + L_{mut}) > \frac{S}{\gamma}$ for some parameter γ, then μ_1 should be increasing. Otherwise, it should be constant or decreasing, as it may be more favorable to not enter the ultra-virulent state when there are few lysogens. An equation that implements this intuition is presented below.

$$\frac{d\mu_1}{dt} = \mu_{scale}[\gamma(L + L_{mut}) - S]$$

μ_{scale} is a scaling parameter to control the size of the values obtained from this equation. When $(L + L_{mut}) > \frac{S}{\gamma}, \frac{d\mu_1}{dt} > 0$, so this equation accomplishes what we set out to do above.

μ_2 is related to the selective pressure in the system for sustainability. When the V_{vir} compartment is large, all cell types are experiencing heavy predation and are thus in risk of being driven to extinction, which would deprive the virions of their food source in the long-term. Therefore, there is a selective pressure for sustainability, which is accomplished much better through lysogeny. Thus, the lower the overall cell population becomes, the greater the pressure for a restoration of lysogeny: in terms of our model, the greater the flow between the V_{vir} compartment and the V_{mut} compartments. An equation that accomplishes this is presented below.

$$\frac{d\mu_2}{dt} = \mu_{scale}[P - \Sigma]$$

P is a critical population threshold parameter that denotes the beginning of "the danger zone" for the phage population. When Σ is above this threshold, there is great advantage to be had in being ultra-virulent and fully lytic, as that will cause the greatest increase in virion numbers, so $\frac{d\mu_2}{dt} < 0$. However, when $\Sigma < P$, the selective pressure for sustainability takes over and μ_2 begins increasing, causing a shift from the ultra-virulent state V_{vir} to the mutant lysogenic strain V_{mut}.

All that remains to be discussed are the equations governing V_{mut} and L_{mut}. As another phage-lysogen pair, they behave very similarly to V and L in the model of Wahl et al. and will have similar equations. However, there are some differences: let's look at the word equations.

$$\frac{dV_{mut}}{dt} = \text{Virions from Lytic Infection} - \text{Clearance} + \text{Virions from Induction} + \text{Mutation}$$
$$\frac{dL_{mut}}{dt} = \text{Logistic Growth} + \text{New Lysogenic Infections} - \text{Induction} - \text{Predation}$$

The differences between these equations and those for V and L in the model of Wahl et al are the presence of a mutation term in V_{mut} and the presence of a predation term in L_{mut}. The mutation term in V_{mut} will just be receiving the μ_2 flow from V_{vir}, so the term itself will just be the opposite of the μ_2 term that appears in the equation governing V_{vir}. The lysogens in L_{mut} can be predated upon by every virus type except for V_{mut}, so that will be expressed using $\nu - V_{mut}$. These reflections suggest the form of these equations.

$$\frac{dV_{mut}}{dt} = N\beta(1-\lambda_\mu)(\Sigma - L_{mut})V_{mut} + N\alpha L_{mut} - cV_{mut} + \mu_2[\delta(\mu_2)V_{vir} + \chi(\mu_2)V_{mut}]$$
$$\frac{dL_{mut}}{dt} = rL_{mut}(1-\frac{\Sigma}{K}) + \beta\lambda_\mu(\Sigma - L_{mut})V_{mut} - \alpha L_{mut} - \beta L_{mut}(\nu - V_{mut})$$

The only new parameter to mention is λ_μ, which is the probability of initiating a lysogenic infection for V_{mut}. It may be that $\lambda_\mu = \lambda_V$, and the probability of lysogeny will not have changed over the course of the evolution of the new strain, but this offers an extra degree of freedom for simulation and experimentation. As mentioned before, these are almost exactly like the equations of Wahl et al. for the derivative of V and L, but the mutation term in V_{mut} – opposite of the corresponding one in the equation governing V_{vir} – and the predation term in L_{mut} make a big difference.

Here are all 8 differential equations of the model together for reference: note that V and L have been updated in the manner of the changes we just made to V_{mut} and L_{mut} to include a mutation term and a predation term to reflect the mutation of V into V_{vir} and the predation upon L by every virus type except for V.

$$\Sigma = S + L + L_{mut}$$
$$\nu = V + V_{vir} + V_{mut}$$
$$\frac{dS}{dt} = rS(1-\frac{\Sigma}{K}) - \beta S\nu$$
$$\frac{dV}{dt} = N\beta(1-\lambda_V)(\Sigma - L)V + N\alpha L - cV - \mu_1[\delta(\mu_1)V + \chi(\mu_1)V_{vir}]$$
$$\frac{dL}{dt} = rL(1-\frac{\Sigma}{K}) + \beta\lambda_V(\Sigma - L)V - \alpha L - \beta L(\nu - V)$$
$$\frac{dV_{mut}}{dt} = N\beta(1-\lambda_\mu)(\Sigma - L_{mut})V_{mut} + N\alpha L_{mut} - cV_{mut} + \mu_2[\delta(\mu_2)V_{vir} + \chi(\mu_2)V_{mut}]$$
$$\frac{dL_{mut}}{dt} = rL_{mut}(1-\frac{\Sigma}{K}) + \beta\lambda_\mu(\Sigma - L_{mut})V_{mut} - \alpha L_{mut} - \beta L_{mut}(\nu - V_{mut})$$
$$\frac{d\mu_1}{dt} = \mu_{scale}[\gamma(L + L_{mut}) - S]$$
$$\frac{d\mu_2}{dt} = \mu_{scale}[P - \Sigma]$$

Analysis and Results

Now that we have a model, we can turn back to the questions originally posed in the Introduction: is there a set of parameter values where the infection and original phage strain

persist and no phage coevolution occurs? Is there a set of parameter values where there is phage coevolution, leading to the presence of a mutant strain and the elimination of the infection? And finally, what happens when these parameter values are changed over a period of time - can the system switch from being attracted from one outcome to the other?

Normally, these questions could be pursued through use of steady state analysis: however, the model developed so far is very difficult to understand analytically, and the fact that μ_1 and μ_2 are variables only complicates this matter. Instead, we will find answers through simulation.

Before trying out the model with different sets of parameter values, let's establish what the initial conditions will be and understand how the selective pressures in the system might guide our choices. We will start each run of the model with every equation except for S, V, and L at 0. That way, the system begins from a state of presumed infection. Then, we will let $S_0 = 1000$ cells, $V_0 = 1$ virion, and $L_0 = 1000$ cells. Simulating the model with many different initial conditions seems to suggest that there is only one stable steady state at a time, so while the dynamics of the system at the beginning depend greatly on the initial condition, the long-term outcome is insensitive to it; thus, any initial condition with these three compartments having non-zero values will likely work.

The two selective pressures modeled in our system of equations are the pressure to circumvent lysogens' immunity to superinfection and the pressure to be sustainable: these are encoded in the governing equations for μ_1 and μ_2, respectively. Because μ_1 starts at 0, only when the lysogen population, L, comes to greatly outnumber the host cell population, S, will μ_1 grow to an appreciable value and force mutation from the original viral strain, V, to the ultra-virulent mutant, V_{vir}. Thus, if L never grows to greatly outnumber S, then no ultra-virulent mutants will ever be produced, and thus, evolution in the system will never occur. So, in order for the host-cell population to persist with no evolution in the phage population, one should look for ways to decrease the number of lysogens without negatively affecting the susceptible host cell population.

$$\frac{d\mu_1}{dt} = \mu_{scale}(\gamma(L + L_{mut}) - S)$$
$$\frac{d\mu_2}{dt} = \mu_{scale}(P - \Sigma)$$

Figure 6: The equations governing changes in μ_1, the flow from the original viral strain V to the ultra-virulent mutant V_{vir}, and μ_2, the flow from the ultra-virulent mutant V_{vir} to the mutant strain V_{mut}. Letting $\gamma = 1$ for simplicity, when the total number of lysogens is greater than the number of susceptible host cells S, μ_1 will increase. When the total cell population Σ decreases below a critical population threshold P, then the rate of flow from ultra-virulent mutants V_{vir} to the new mutant strain V_{mut} will increase.

What parameters are present in the lysogen equations but not in the susceptible host cell equation? The figure below shows these equations together for comparison.

$$\Sigma = S + L + L_{mut}$$
$$\nu = V + V_{vir} + V_{mut}$$
$$\frac{dS}{dt} = rS(1 - \frac{\Sigma}{K}) - \beta S\nu$$
$$\frac{dL}{dt} = rL(1 - \frac{\Sigma}{K}) - \beta L(\nu - V) - \alpha L + \beta\lambda_V(\Sigma - L)V$$
$$\frac{dL_{mut}}{dt} = rL_{mut}(1 - \frac{\Sigma}{K}) - \beta L_{mut}(\nu - V_{mut}) - \alpha L_{mut} + \beta\lambda_\mu(\Sigma - L)V_{mut}$$

Figure 7: The cell-type equations. Note the presence of the induction rate, α, in a decay term in each of the lysogen equations, whereas it is not present in the susceptible host cell equation.

There are a number of differences between the equation governing S and the equation governing the lysogen populations, but the one that shall perhaps serve us well are the terms $-\ \alpha L$ and $-\ \alpha L_{mut}$ present in the equations governing L and L_{mut}, respectively. These represent the loss in the lysogen populations due to induction. If we increase the value of α, then the total number of lysogens should be negatively impacted, but the value of S shouldn't be directly affected. Thus, α seems like a promising parameter that can be tweaked to prevent or allow evolution in the system.

This will be our strategy: we will find a value of α that is high enough that the lysogen populations will never be able to outnumber S – and thus the system will never experience evolution – and we will also find a value of α low enough that the lysogens are able to grow past S, starting the sequence of evolution in the system and eventually driving the host population to extinction. In the first case, we would expect a stable steady state where the susceptible host

population persists and V_{vir}, V_{mut}, and L_{mit} are at zero. In the second, we would expect a stable steady state where S and L are at zero, and V_{mut} and L_{mut} are non-zero.

These are the parameter values used for our simulation.

Parameter	Estimated Value
Timescale	20 minutes (the generational time for E. Coli under optimal conditions)
r	1 / time - each cell reproduces itself in 1 time unit
K	1 x 10^6 cells
β	0.0005 /(time * virions)
N	100 virions [5]
λ_V	0.5
λ_μ	0.25
α	(1 - 2)/time (depending on steady state desired)
c	4 /time
μ_{scale}	1 x 10^{-7} /time
ɣ	1
P	1000

After trying a number of values for α, two values giving distinct qualitative behavior were found.

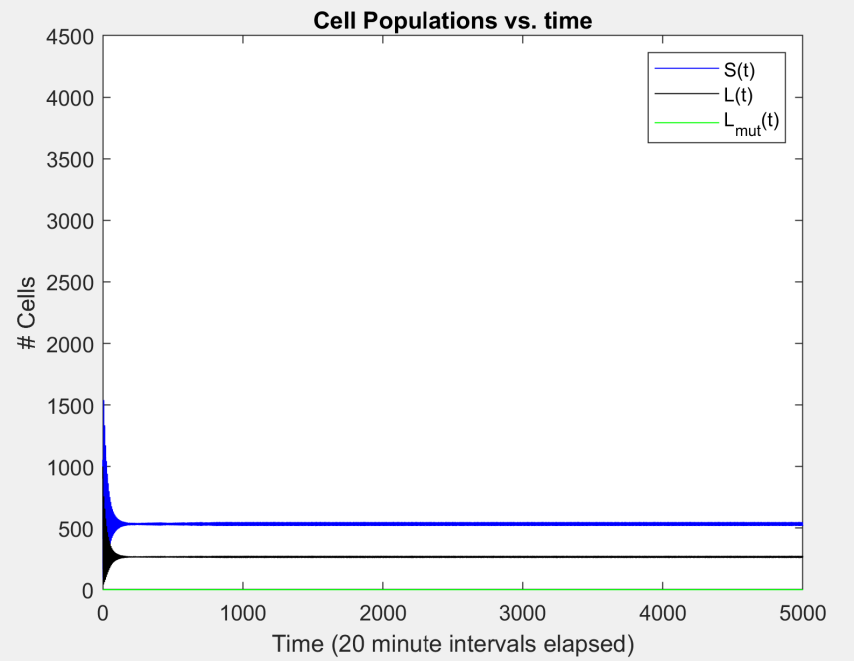


Figure 8: The cell and viral compartments versus time when the system is run with $\alpha = 2$. The high induction rate suppresses the lysogen population, minimizing the selective pressure for ultra-virulence and stopping evolution in the system before it starts.

When $\alpha = 2$, the induction rate is too high for the lysogens to ever outnumber the susceptible host population, and thus, the selective pressure for ultra-virulence is never high enough to initiate evolution in the system. Thus, the susceptible hosts persist.

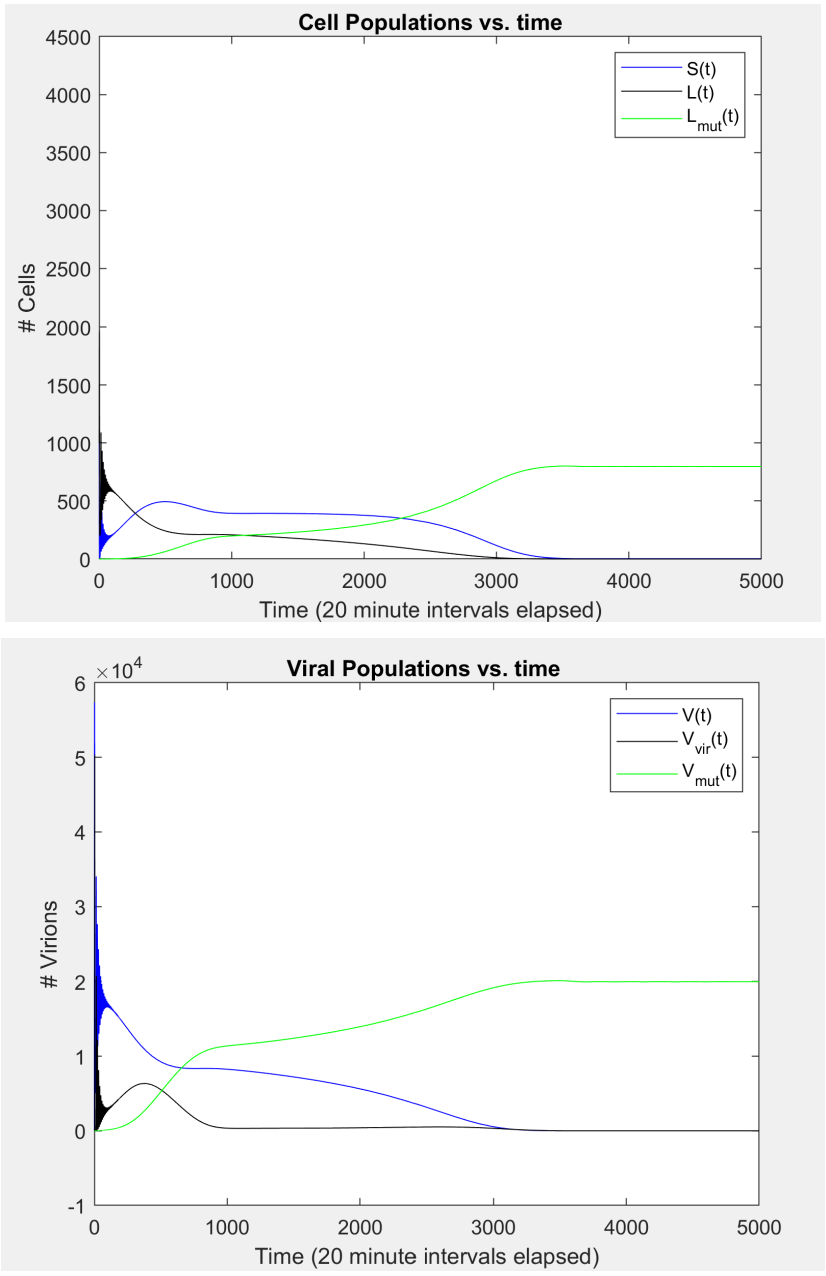


Figure 9: The cell and viral compartments versus time when the system is run with $\alpha = 1$. The low induction rate allows for favored growth of the lysogen population, triggering the selective pressure for ultra-virulence and getting evolution underway. Then, once the population is decreased greatly by the ultra-virulent mutant V_{vir}, there is a strong selective pressure for the new mutant lysogenic strain to arise.

When $\alpha = 1$, the induction rate is low enough for the lysogens to outcompete the susceptible host cells, triggering the selective pressure for ultra-virulence, which opens the gate for evolution to occur in the system. This particular value of α is low enough that the susceptible host

population is completely driven to extinction, but for some intermediate values between $\alpha = 2$ and $\alpha = 1$, there is coexistence between the susceptible host population and the new mutants.

Without a full analysis of the system's steady states and their stability, a full bifurcation diagram is out of reach. However, we can still plot how the steady state values of S, L, and L_{mut} depend on the induction rate α.

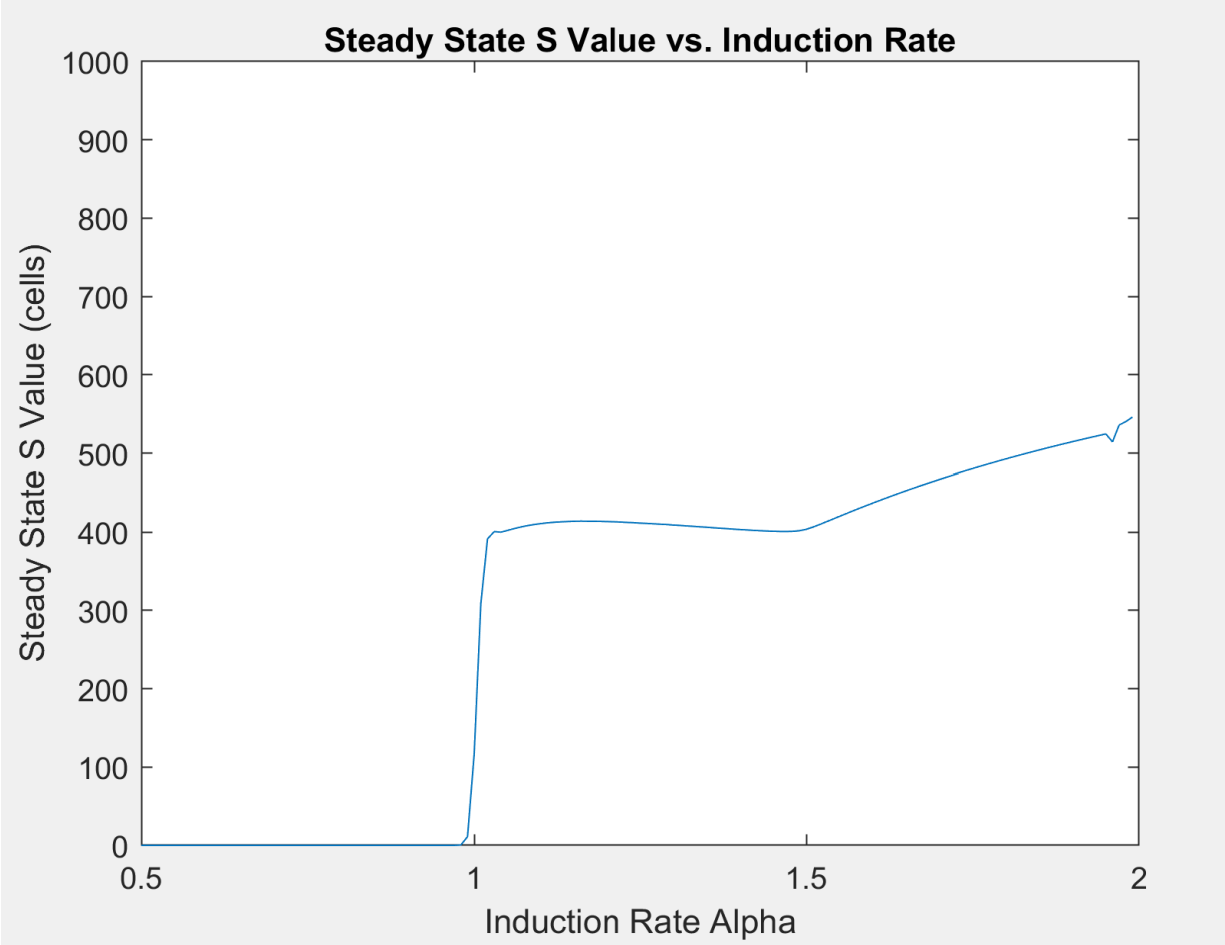


Figure 10: The value of S at steady state versus the induction rate α. When $\alpha > 1$, the susceptible host population is stable, but when α gets close to 1, the susceptible host cells are rapidly overwhelmed by the new mutant strain. Thus, in the context of phage therapy, $\alpha = 1$ would be the target induction rate to reach.

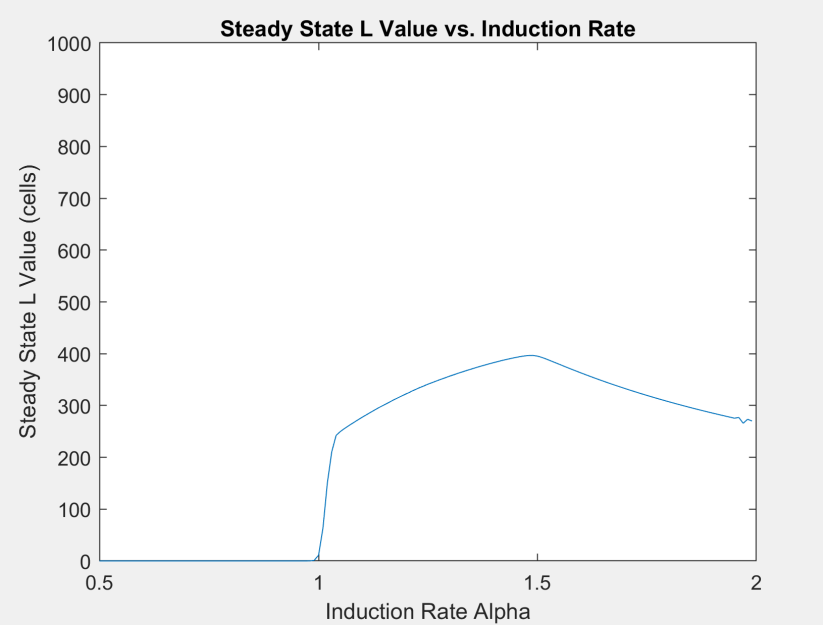


Figure 11: The steady state value of L versus the induction rate α. The population reaches its maximum halfway in-between the two critical α-values used in plotting timecourses above.

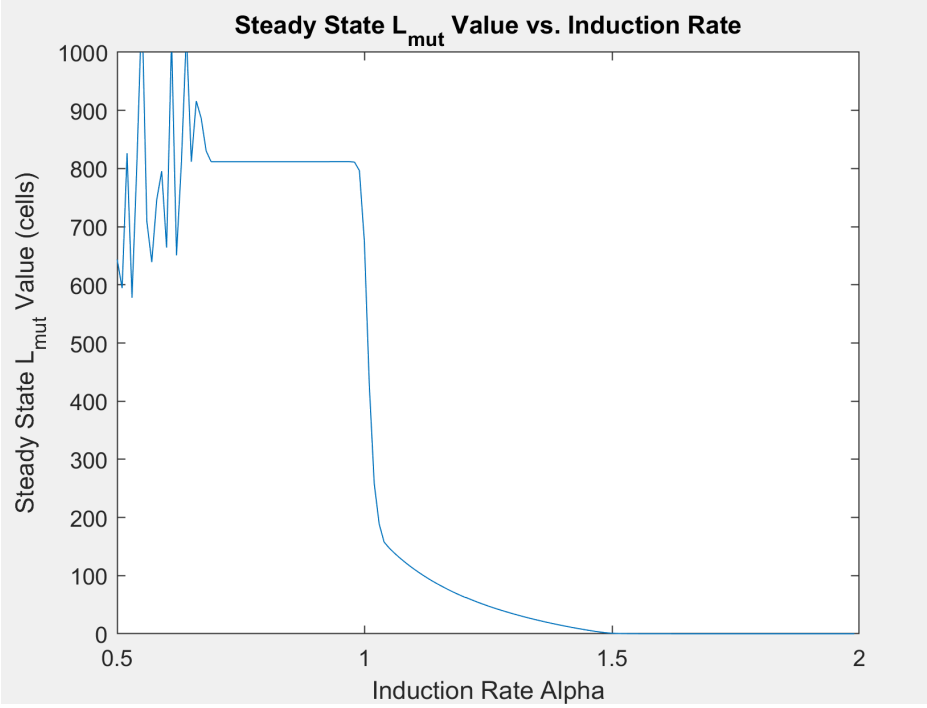


Figure 12: The steady state value of L_{mut} versus the induction rate α. There is a large range, $1 < \alpha < 1.5$ where the new mutant phage strain and its lysogens can coexist with the susceptible host population, but when $\alpha > 1.5$, the system experiences no evolution.

From these graphs, we are able to get some important results: namely, the answers to the questions initially posed about the model. First, $\alpha = 1.5$ is approximately the value which separates the system undergoing evolution from not undergoing evolution. Second, the shift from one kind of behavior to the other is rather rapid: the S and L_{mut} steady state value graphs take the form of plateaus that sharply fall off at $\alpha = 1$. Finally, we see that the model does predict that, if the induction rate α is lowered sufficiently, a new mutant strain should be generated, giving us a general principle to extend into experiment.

Discussion and Conclusions

We started with the question: what causes a genetically distinct strain of lysogenic phage to arise in a population? In developing a mathematical model of the situation, we rephrased into a question we could ask of the model: what sets of parameters will cause a mutant strain to arise and supplant the original as well as the susceptible host population? Through reasoning about the selective pressures inherent in the system, we hypothesized that changes in the induction rate of the lysogen populations could cause a change in the qualitative behavior of the model. Then, in Figures 3 - 7, we saw that hypothesis confirmed, with three different regimes of qualitative behavior depending on the induction rate: $\alpha \leq \sim 1$ led to the elimination of the susceptible host population and the generation of a new mutant strain; $\sim 1 < \alpha < 1.5$ led to the coexistence of the susceptible host population, the original strain, and the mutant strain; and $\alpha \geq 1.5$ saw no evolution in the system.

Thus, if the model is a good analogue of reality, one might expect that sufficiently lowering the induction rate of a lysogenic phage population will lead to the generation of a new mutant strain. Because the induction rate of a lysogen population is related to the stress they experience, a reduction in stress might accomplish this. Increasing nutrient availability, changing the temperature to more optimal levels for the host cell species, or reducing host cell density may all be valid ways to reduce environmental stress on lysogens. This may result in an increase in the effectiveness of phage therapy when an infection is being particularly troublesome: if an effective technique for experimentally generating new strains based on this result can be found, it may help save lives.

Here is one possible way to design a lysogenic phage for phage therapy that has an induction rate that can be raised or lowered as desired. Synthetic biologists have engineered heat-inducible expression systems that will produce proteins as soon as the temperature is raised above 37 ℃: if the protein to be expressed were toxic to the cell, then as long as the temperature is above 37 ℃, the induction rate will be high [4]. Once the system has settled into steady state, if the temperature is lowered below 37 ℃, then the induction rate will be greatly lowered. By editing the genome of a lysogenic phage to incorporate such a heat-inducible expression system

producing a toxic protein, one could experimentally validate the results of this model. On the whole, how to manipulate the parameters present in this model in a real experimental setting is a topic for further investigation and a lot of thought.

One way in which this model connects to evolutionary biology as a whole is through μ_1 and μ_2. These variables essentially quantify the natural selection occurring in the system. Their functional forms were chosen to make sense for this particular context, but in general, one may find that there's a generic equation for this kind of mutational, compartmental flow. If i and j are compartments with fitnesses $\phi(i)$ and $\phi(j)$, then the flow from compartment i to compartment j μ_{ij} is $\frac{d\mu_{ij}}{dt} = k(\phi(i) - \phi(j))$, where k is a parameter to control the strength of the difference. This generalization is nice in that it rests solidly on evolutionary theory and is broadly applicable: however, one challenge would be finding an appropriate fitness function for each modeled system.

Mutation and evolution are always occurring in natural settings: this model is a first step in understanding what factors influence the evolution of lysogenic phage populations and making predictions about their long-term behavior. It shows that only two selective pressures are needed for a lysogenic phage to experience evolution and that researchers might learn to use these selective pressures in a directed evolution context. However, it has many shortcomings that should be addressed by future modeling efforts: it only considers the one kind of immunity to infection that comes from lysogeny; it does not take into account the other selective pressures acting on a phage population and its host in a real setting; the timescales at which the long-term behavior is being evaluated are very long (on the order of months); and it's almost impossible to get analytic results from the model due to its structure. If future models investigated different functional forms and ways to approach this system, it may be possible to derive analytic results as well as discover more general principles about how phages coevolve virulence in response to host immunity. Although it could be frustrating at times due to its large size and inherent complexity, I like the model developed in this paper: the rock-paper-scissors nature of the populations' interactions with each other leads to very interesting behavior, and I think that the way μ_1 and μ_2 have been designed greatly expands what kind of behavior can be seen from a compartmental model. It's rather incredible that we can model reality with 8 lines of text on a sheet of paper, and this was a great introduction to the art of doing so.

References

[1] Uyttebroek S, Chen B, Onsea J, Ruythooren F, Debaveye Y, Devolder D, Spriet I, Depypere M, Wagemans J, Lavigne R, Pirnay JP, Merabishvili M, De Munter P, Peetermans WE, Dupont L, Van Gerven L, Metsemakers WJ. Safety and efficacy of phage therapy in difficult-to-treat infections: a systematic review. Lancet In

[2] Hatfull GF, Dedrick RM, Schooley RT. Phage Therapy for Antibiotic-Resistant Bacterial Infections. Annu Rev Med. 2022 Jan 27;73:197-211. doi: 10.1146/annurev-med-080219-122208. Epub 2021 Aug 24. PMID: 34428079.
fect Dis. 2022 Aug;22(8):e208-e220. doi: 10.1016/S1473-3099(21)00612-5. Epub 2022 Mar 3. PMID: 35248167.

[3] Wahl, L.M., Betti, M.I., Dick, D.W., Pattenden, T. and Puccini, A.J. (2019), Evolutionary stability of the lysis-lysogeny decision: Why be virulent?. Evolution, 73: 92-98.

[4] Valdez-Cruz, N.A., Caspeta, L., Pérez, N.O. *et al.* Production of recombinant proteins in *E. coli* by the heat inducible expression system based on the phage lambda pL and/or pR promoters. *Microb Cell Fact* 9, 18 (2010). https://doi.org/10.1186/1475-2859-9-18