Bacteria-Phage Models with a Focus on Prophage as a Genetic Reservoir

Alina Nadeem, The University of Western Ontario

Supervisor: Dr. Lindi M. Wahl, The University of Western Ontario

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Applied Mathematics

© Alina Nadeem 2016

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Applied Mathematics Commons

Recommended Citation

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.
Abstract

Temperate bacteriophages have the ability to incorporate their genetic material in the host’s DNA, which may be utilized by later generations of phage to overcome the host’s receptor-based defences. This effect of temperance can have major implications for the long-term survival of the phages as well as on bacteria-phage community evolution. To study the impact of prophage on microbial communities we have developed models simulating lytic and lysogenic infection and host and phage coevolution with a focus on prophage-phage recombination. Our results show that recombination can be crucial for the phage to survive host diversification, and a higher incidence of lysogeny as opposed to lysis may favor the phage population in the long run. Moreover, depending on the nature of interaction between the hosts and phages, phage-prophage recombination may promote overall diversity in the phage-host community or push the hosts towards ‘waves’ of innovation.

Keywords: Lysogeny, recombination, bacteriophage, prophage, coevolution.
Co-Authorship Statement

The work in chapter 3 is in preparation for submission to Evolution as “Prophage as a genetic reservoir: promoting diversity and driving innovation in the host community.” by Alina Nadeem and Lindi M. Wahl.

Part of the work in chapter 2 has been submitted for publication as “The impact of prophage on the equilibria and stability of phage and host”, by Pei Yu, Alina Nadeem and Lindi M. Wahl, under revision for the Journal of Non-Linear Science. I originally completed the equilibrium calculations shown in section 2.2.1 independently, but the expressions given in this document are based on the work of Dr. Pei Yu in the aforementioned journal article.
Acknowlegements

First and foremost I would like to express my heartfelt gratitude towards my supervisor Dr. Lindi M. Wahl for her continuous encouragement, guidance, positivism and belief in my abilities. Her incredibly patient attitude and immense knowledge of her field never ceases to impress me. Without her support, this thesis would not have been possible. I would also like to thank my parents, my brother Ahmed and my friend Deepakshi for always being there for me.
## Contents

Abstract ii  
Co-Authorship Statement iii  
Acknowledgements iv  
List of Figures vii  
List of Tables viii  
List of Appendices ix  
List of Abbreviations, Symbols, and Nomenclature x  

1 Introduction 1  
1.1 Viral Modes of Replication: Lysis vs. Lysogeny 1  
1.2 Bacteria-Phage Coevolution and Recombination 2  
1.3 Previous Studies on Bacteria-Phage Systems 4  

2 Models with lysis, lysogeny and recombination 16  
2.1 Introduction 16  
2.2 Bacteria-Phage model with 2 receptor types 17  
2.2.1 Dimensionless Model and Possible Equilibria 21  
2.2.2 Numerical Simulations 24  
2.2.3 Optimal Rate of Lysogeny $p$ 25
List of Figures

2.1 Trivial and Infection-Free Equilibria of System (2.4)-(2.6) . . . . . . . . . . . . 39
2.2 Diversification Equilibrium of System (2.4)-(2.6) . . . . . . . . . . . . . . . . 39
2.3 Stable Endemic Equilibrium and Stable Limit Cycles from System (2.4)-(2.6) . 40
2.4 The Effects of Recombination for System (2.4)-(2.6) . . . . . . . . . . . . . . . 41
2.5 Phage Population vs Lysogeny Probability for System (2.4)-(2.6) . . . . . . . . 41
2.6 Dynamics of Host and Phage Populations for System (2.14)-(2.19) . . . . . . . 42
2.7 Hosts and Phages with $P_J$ Dominating Initially for System (2.14)-(2.19) . . . 43
2.8 Hosts and Phages with $P_K$ Dominating Initially for System (2.14)-(2.19). . . 43
2.9 System (2.14)-(2.19) with mostly $K$ receptors and without recombination . . 44
2.10 System (2.14)-(2.19) with mostly $K$ receptors with recombination . . . . . 45
2.11 System (2.14)-(2.19) with different receptor proportions and without recombi- nation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46
2.12 System (2.14)-(2.19) with different receptor proportions and with recombination 47

3.1 One-to-one Infection Pattern without Recombination. . . . . . . . . . . . . . . 59
3.2 One-to-one Infection Pattern with Recombination. . . . . . . . . . . . . . . . 60
3.3 Nested Infection Pattern with Phage Mutation and Recombination. . . . . 60

B.1 Prophage Acquisition by Host Cells. . . . . . . . . . . . . . . . . . . . . . . . . . 77
B.2 One-to-one Infection with Relaxed Initial Conditions. . . . . . . . . . . . . . 78
B.3 Long-term Dynamics with Nested Infection . . . . . . . . . . . . . . . . . . . 79
## List of Tables

| A.1 | Parameters for System (2.4)-(2.6) | 72 |
| A.2 | Parameters for System (2.14)-(2.19) | 73 |
| A.3 | Parameters for One-to-one and Nested Infection Models | 73 |
List of Appendices

Appendix A Parameter Values for All Systems. ............................. 72
Appendix B Supporting Material for Chapter 3 ............................ 74
## List of Abbreviations, Symbols, and Nomenclature

### Model 1

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$\frac{\lambda}{\delta}$</td>
</tr>
<tr>
<td>$B$</td>
<td>$\frac{\mu}{\delta}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Rate of infection of hosts by phages</td>
</tr>
<tr>
<td>$D$</td>
<td>$\frac{d}{\delta}$</td>
</tr>
<tr>
<td>$d$</td>
<td>Clearance rate of phages</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Death rate of infected host cells</td>
</tr>
<tr>
<td>$F$</td>
<td>$\frac{f(1-r)(1-p)\beta K}{\delta^2}$</td>
</tr>
<tr>
<td>$f$</td>
<td>Lysis rate of phages</td>
</tr>
<tr>
<td>$H_{xy}$</td>
<td>Host cell population with $x$ prophage and $y$ receptors</td>
</tr>
<tr>
<td>$I_{xy}$</td>
<td>Infected host cells with $x$ prophage and $y$ phage type infecting them</td>
</tr>
<tr>
<td>$C$</td>
<td>Carrying capacity of uninfected and infected host cells</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Birth rate of uninfected host cells</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate of uninfected host cells</td>
</tr>
<tr>
<td>$N$</td>
<td>Total number of uninfected and infected host cells of all types</td>
</tr>
<tr>
<td>$P_x$</td>
<td>Phage with $x$-type tail fibre</td>
</tr>
</tbody>
</table>
\( p \) Probability of lysogeny

\[ R = \frac{fr(1-p)\beta K}{2\delta^2} \]

\( r \) Recombination rate for phage-prophage recombination

**Model 2**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_x )</td>
<td>Rate of infection of hosts by phages with tail fibre of type ( x )</td>
</tr>
<tr>
<td>( d )</td>
<td>Clearance rate of phages</td>
</tr>
<tr>
<td>( \delta )</td>
<td>Death rate of infected host cells</td>
</tr>
<tr>
<td>( f )</td>
<td>Lysis rate of phages</td>
</tr>
<tr>
<td>( H_x )</td>
<td>Host cell population with ( x ) prophage</td>
</tr>
<tr>
<td>( I_{xy} )</td>
<td>Infected host cells with ( x ) prophage and ( y ) phage type infecting them</td>
</tr>
<tr>
<td>( C )</td>
<td>Carrying capacity of uninfected and infected host cells</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>Birth rate of uninfected host cells</td>
</tr>
<tr>
<td>( \mu )</td>
<td>Death rate of uninfected host cells</td>
</tr>
<tr>
<td>( N )</td>
<td>Total number of uninfected and infected host cells of all types</td>
</tr>
<tr>
<td>( P_x )</td>
<td>Phage with ( x )-type tail fibre</td>
</tr>
<tr>
<td>( p )</td>
<td>Probability of lysogeny</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Proportion of ( J ) receptors on the cell surface</td>
</tr>
<tr>
<td>( r )</td>
<td>Recombination rate for phage-prophage recombination</td>
</tr>
</tbody>
</table>

**Model 3**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_x )</td>
<td>Rate of infection of hosts by phages with tail fibre of type ( x )</td>
</tr>
<tr>
<td>( d )</td>
<td>Clearance rate of phages</td>
</tr>
</tbody>
</table>
\( \delta \) Death rate of infected host cells

\( f \) Lysis rate of phages

\( \mathcal{H} \) Total host cells of all types

\( H_x \) Host cell population with \( x \) receptor

\( \mathcal{H}_x \) Total host cells with \( x \) receptor

\( j, k, l, m, n \) Types of prophage

\( C \) Carrying capacity of uninfected and infected host cells

\( \lambda_x \) Birth rate of uninfected host cells with \( x \) receptor

\( M \) Expected number of mutations in time \( \Delta t \)

\( \mu \) Death rate of uninfected host cells

\( P_x \) Phage with \( x \)-type tail fibre

\( p \) Probability of lysogeny

\( r \) Recombination rate for phage-prophage recombination

\( s_\beta \) Cost of adaptation for the phages

\( s_\lambda \) Cost of resistance faced by hosts

\( u_H \) Bacterial mutation rate

\( u_P \) Phage mutation rate

\( v_H \) Back mutation rate for host cells
Chapter 1

Introduction

1.1 Viral Modes of Replication: Lysis vs. Lysogeny

Bacteria and their predators, the bacteriophages (also known as phages), form a model system for the study of evolution and coevolution due to their abundance in nature [7], fast evolutionary dynamics [44, 42, 55, 18, 36] and ease of culturing in a laboratory setting (reviewed in [30, 27, 39]). Bacteriophages may be the most abundant life forms on the planet with a total population size approaching $10^{31}$, outnumbering their hosts approximately 10 to 1 [7], and producing as many as $10^{24}$ infections/second worldwide [66, 50]. Infection of a bacterium by a phage, which involves the phage injecting its genetic material into the host following binding to a surface receptor, may result in two different scenarios. Once the infecting phage has injected its genetic material, it may begin a lytic cycle of replication, where it takes over the host’s machinery to produce a large number of virions that are released from the infected cell, destroying the cell in the process. On the other hand the phage may undertake lysogeny, in which the phage genetic material is integrated in the host chromosome and is replicated each time the host cell divides. The latter scenario leaves the bacterial cell unharmed and the integrated phage DNA is called a prophage. In the absence of debilitating mutations, a prophage has the ability to instigate lysis of the host cell at a later time, through a process called induction in which the genes
for lysis are no longer repressed. Bacterial cells carrying prophage are called lysogens and these may induce and lyse either spontaneously or in response to certain environmental signals [20, 25, 52].

1.2 Bacteria-Phage Coevolution and Recombination

The proportion of bacterial genomes that carry prophage is very high [17], and typically multiple prophage sequences can be identified in these lysogens [23]. For example, sequenced *Escherichia coli* strains harbor between seven and twenty distinct prophages [25]. However, due to relaxed selection on lytic function, many identified prophage sequences are likely remnants of ancestor prophage that have degraded by mutation and are no longer inducible to produce competent phage [21, 23]. Recent evidence for purifying selection in prophage sequences[11], as well as direct fitness measures [61], have confirmed that these cryptic prophage can nonetheless play important roles in host cell fitness.

Phage and prophage participate in both homologous and non-homologous recombination [38, 17]. Botstein proposed that phage sequences are composed of functional cassettes (“modules”) and evolve through the exchange of these modules or groups of modules [13]. Although it is now understood that phage evolution is a more complex and varied phenomenon, certain phage genes (particularly the virulence genes) are flanked on one side by a promoter region and on the opposite side by a terminator. This organization forms a discrete autonomous genetic element called “moron” and encourages gene transfer between phages [33]. More interestingly, the highly mosaic nature of phage sequences and the aforementioned abundance of prophage greatly increase the chances of recombination between an infecting phage and a prophage integrated within the infected cell’s genome. This phenomenon has been observed repeatedly in the lactococcal phages, for example [48, 14, 29, 40]; in one carefully studied instance, phage ul36 recombined 79% of its genetic material with prophage found in the host it was infecting. This enabled the phage to acquire immunity to 2 different anti-phage mechanisms being
1.2. Bacteria-Phage Coevolution and Recombination

employed by the host [40]. Another striking example is that of phage Lambda, which in one study extended its target range from just the LamB E. coli receptor, to both LamB and OmpF. Sequencing results indicate that phage Lambda achieved this transition through a new tail fibre sequence, obtained by recombination with a prophage or prophage remnant in the host cell (J. Meyer, personal communication). The new tail fibre allowed the phage to attach to the host through a novel receptor and overcome the host’s defenses. These instances clearly demonstrate that prophage sequences, harboring a genetic reservoir of phage genes within the host cell, can have a powerful impact on bacteria-phage dynamics.

Although the phage may be able to find ways to overcome the host’s defences, in many cases, bacteria and phage interactions result in a coevolutionary struggle, with several cycles of bacterial defenses and phage counter-defenses [27, 39]. There has been some debate about the nature of this coevolution in the scientific community. Rodin and Ratner [53, 54] claimed in 1983 that the cycles of bacteria and phage coevolution may be endless, as they are locally directed but globally undirected. Directional selection implies that coevolution leads to the development of greater resistance ranges. On the other hand undirected selection favours different resistance ranges but the organisms resulting from each mutational step are not superior to the previous ones [18]. In 1984 Lenski [41] believed that phages that evolve to infect resistant bacterial strains have an extended rather than an altered host range, so essentially the coevolution is directed. He asserted that the phage will eventually and inevitably lose out in the race as their evolutionary potential is limited relative to that of the bacteria. Mutations that allow bacteria to evade the phages are much easier to come by than those that allow phages to adapt to the resistant bacteria. However, in 2002 Buckling and Rainey [18]-going back to the hypothesis of Rodin and Ratner-demonstrated long-term coevolution between bacteria and phage that lasted for several cycles. Their study concluded that selection for these populations may be directional in lab settings and fluctuating (undirected) in nature. The phages did not seem to face any fundamental constraints in terms of their coevolutionary potential. In the cases where the bacteria evolved much faster than the phages, one of the reasons for the faster evolution of
bacteria was explained to be their much larger populations. Another line of reasoning proposed by [49] was that coevolution leads to mutator bacteria which evolve so rapidly that they can out-run the phages in the evolutionary race.

More recent research attention has focused on looking at bacteria-phage systems as infection networks between virus and host types rather than isolated coupled interactions [65]. The simplest pattern, the “matching alleles” (MA) model, assumes that each virus type can infect one host type, while each host type is susceptible to one virus [31]. In contrast, the “gene for gene” (GfG) model assumes that hosts successively evolve greater degrees of resistance (resistance to more phage types) whereas viruses coevolve by increasing their host range (infecting more host types) [10], resulting in a nested network of infection. Red Queen dynamics are often observed in MA models, in which negative frequency dependent selection favors rare genotypes; thus resistance and infectivity polymorphisms are maintained as virus and host types oscillate through periods of abundance and rarity. GfG models, however, typically display arms race dynamics, resulting in the replacement of one type by a superior competitor through selective sweeps [27]. In recent experimental work, Hall et al. [36] found that bacteria-phage coevolution follows an evolutionary arms race that later switches to Red Queen dynamics as the costs of further resistance and range expansion become too high [27, 39]. Theoretical studies based on these models will be discussed in the next section.

1.3 Previous Studies on Bacteria-Phage Systems

The intriguing dynamics of the coevolution of bacteria and phage have led to a wide range of experimental studies, as reviewed in some depth by Dennehy [27], Koskella, and Brockhurst [16, 39]. As pointed out by these authors, experimental coevolution studies to date have focused on lytic bacteriophage, and empirical work addressing temperate phage coevolution is comparatively sparse ([39], but see [52] for one-sided evolution of a temperate phage). A number of studies have concluded that coevolution may promote phage and bacterial diversity on
1.3. Previous Studies on Bacteria-Phage Systems

a phenotypic, genotypic and community level [39]. This may follow naturally because of rare genotypes having an advantage over more abundant types [35], or due to the costs and benefits of specificity and generalism [45].

Theoretical studies of bacteria-phage coevolution have a rich history, with one of the seminal works in bacteria-phage modeling being done by Campbell in 1961 [20] who analyzed a simple host-phage model based on sensitive and resistant host populations, and virulent phages. Campbell found that the system could go to infection-free or endemic equilibria based on the growth rate of the phages. He concluded that in the presence of various bacterial species with varying fitness, phages that infect the more fit host not only have a greater chance of survival, but will also help to offset the fitness advantage that this host has over the others. Campbell claimed that bacteria can mutate to become resistant to the phage, and mutant phages can subsequently evolve to attack these newly resistant bacteria. The resistant bacteria may not disturb an existing host-phage equilibrium if they have a lower fitness relative to the original host; they may cause the phage-sensitive population and the phage to be eliminated if they are more fit than the sensitive bacteria. The persistence of temperate phage was justified by pointing out the advantages they may confer to the host in the form of lysogenic conversion and transduction.

An equally simple model consisting of uninfected and infected hosts and phages was analysed by Bremermann [15]. His equations demonstrated that coexistence was possible and stable as long as the carrying capacity of the environment was high enough to support it. If the carrying capacity was to fall below the threshold found by Bremermann, the system was shown to move towards an infection-free state.

A later paper by Levin, Stewart and Chao [44] built on the work done by Campbell by also considering the nature of the habitat and the dependence between prey growth and resource availability. The authors assumed a homogeneous environment, where resource was being supplied and removed at constant rates, and infection rates depended on the relative densities of predator and prey populations. They concluded that the number of host types should be greater than or equal to the number of phage types, and the total number of phage and re-
source types should be larger in number than the host types for biologically relevant equilibria to be found. In the case with just one host, one resource and one phage type, coexistence was possible only if the host population was large enough to sustain the phage – in line with Bremermann. When there were two hosts, one phage and one resource, the model reconfirmed the results found by Campbell. If the phage-sensitive host could achieve a high enough population (to sustain the phage population) in the presence of the phage-resistant host then all three would coexist. The authors derived a minimum fitness threshold for the bacteria to be able to compete for the resource and not go extinct. An important conclusion made by the paper was that partial resistance to the phage increased the chance of survival of the resistant host as compared to complete resistance. Heterogeneity of the environment seemed to favour stability of the system but was not necessary. The authors followed up their theoretical work with practical experiments on different strains of virulent phage T2 and E. coli bacteria, and found most of their results to agree with theory. One of the caveats of the model was that some of the empirically estimated parameters fell in the range in which the model predicted eventual extinction, even though experiments showed stability. The paper was a step forward in the study of bacteria and phage interactions as it attempted to build a realistic model and account for the diversity of phage-bacteria habitats.

While Levin, Stewart and Chao demonstrated the presence of coexistence equilibria in bacteria and phage systems, Lenski and Levin investigated whether there could be co-evolutionary stable states resulting from bacteria and phage interactions. The authors argued that repeated waves of evolution between these organisms must come to an end at some point. The model developed by Lenski and Levin consisted of equations for uninfected and infected bacteria, as well as the primary resource and the phages. It also allowed invasion by resistant bacteria and host range phages that came about through point mutations in the members of the two populations. One important assumption was that the mutant phages had an extended host-range rather than an altered one. Hence they could infect the wild type and mutant host cells. Another assumption was that resistance to phage came with a cost in terms of fitness for
the bacteria. The order in which the mutants appeared was important. If the mutant phages appeared after the hosts developed resistance, they had an advantage in terms of the greater number of hosts available to infect. However, if they appeared before the resistant hosts then they lost their advantage and could even be at a disadvantage. The main conclusion of the authors was that some sort of a coevolutionary stable state will be reached between the evolving host and phage populations (with a clear ‘winner’). This is because the evolutionary potential of the phage was limited – in particular developing the ability to infect new host types involves very specific changes in configuration. The host on the other hand only needs to regulate the expression of its surface receptors which was thought to be more easily achieved. The evolution of hosts may also be limited by the fact that hosts cannot dispose of any receptors that perform essential functions without seriously harming their own fitness, and so such alterations may not survive selection. The authors supplemented their mathematical work with some experimental results based on communities of \textit{E. coli} and virulent T-even and T-odd phages. These experiments confirmed the assumptions about reduced fitness of hosts that develop resistance to phage and showed that phage mutants for all resistant bacteria types do not always develop. They also found that selective constraints may help maintain a population of phage-sensitive bacteria that allow the wild type phage to survive even after mutants have developed. Hence the authors concluded that coevolution between bacteria and phage was limited and more likely to end in favour of the hosts.

The task of explaining the reasons behind the long-term of bacteria-phage coexistence was picked up by Schrag and Mittler [55], who aimed to develop a mathematical model that could accurately predict the high stability of bacteria and phage communities observed in experiments. They presented several hypotheses to explain the long-term maintenance of bacteria-phage systems and attempted to test them through theory and experiments. These were the numerical refuge hypothesis, the endless arms race hypothesis, the physiological refuge hypothesis and the spatial refuge hypothesis. Of these, they found the spatial refuge hypothesis to be the most useful for their purpose. This hypothesis suggested that sensitive bacteria would
clinging to the walls of the experimental vessel and thus the walls provided a spatial refuge for the sensitive cells. Phages cultured in these vessels survived on these wall populations. An improved version of the mathematical model from the works of Levin and Lenski with more accurate parameter values was also presented. The model largely failed to predict the long term stability seen in laboratory settings when it was used to study continuous culture dynamics but was successful in predicting the results for serial transfer settings. The authors also conducted experiments with T1X-phage, *Lambda*-phage and *E. coli* bacteria. Continuous culture experiments showed long-term persistence but serial transfer models resulted in eventual extinction of the phages. Spatial heterogeneity of the environment played a major role in making the bacteria-phage interaction stable, but the authors felt that the other factors mentioned above might also have contributed to the stability to some extent.

One reason for long-term stability that Schrag and Mittler overlooked perhaps was the temperance of phages. However, temperance was taken into account by Stewart and Levin [60] when analysing the reasons viruses maintain a temperate mode of replication even though it seems that the virulent mode would better serve their purpose of rapid replication. The authors examined a model based on temperate and virulent phage populations and hosts that were either sensitive, resistant or lysogens. They considered several combinations of these populations but the case that interests us most consisted of temperate and virulent phages, lysogens and sensitive cells with a possibility of invasion by resistant bacteria. In this scenario, the model predicted that if the resistant bacteria did not invade, the virulent phage might i) eliminate the lysogens and temperate phage and coexist with the sensitive cells, ii) achieve a stable equilibrium in the presence of the temperate phages and lysogens or iii) be eliminated by the lysogens completely. If an invading resistant cell population had a higher fitness level than the lysogens and sensitive bacteria, it would eliminate both of these hosts and their respective virus populations (temperate and virulent). On the other hand, with a growth rate disadvantage, the resistant cells would be able to coexist with the other populations. If a stable coexistence scenario (from those mentioned above) had been established between the phages and host cells, the presence
of resistant cells would not upset their balance. Along with their mathematical predictions, the authors also highlighted some biological explanations for the continued existence of temperance in phages. Temperate phages maintain more stable (long lasting) relationships with their hosts which may ensure phage survival through periods of scarcity in bacterial populations. Viral DNA from temperate phages may give the host super-infection immunity and assist in recombination of genes between hosts in ways that enhance host fitness. Consequently, there are benefits for the bacteria to keep prophage and for the phages to be temperate. By investigating the temperance of phages, the authors brought to light the importance of lysogeny to the dynamics of bacteria and virus populations. Even though they hinted at the importance of prophage to bacteria and viruses, the model itself ignored the importance of prophage (and recombination) from the coevolutionary perspective. Also, it only looked at one cycle of bacterial defenses and did not consider further phage counter-defenses. No major conclusions about the shape of the population dynamics following such coevolutionary cycles were made.

Unlike Stewrt and Levin (who did not extend their model beyond one cycle of coevolution), Weitz, investigated the long-term coevolutionary trajectories of bacteria and bacteriophage by developing a coupled evolutionary and ecological model [64]. He sought out the reasons behind the vast genetic diversity that exists in certain species, with a focus on bacteria and phages, where it was likely that selection would have resulted in most of the less fit mutants being eliminated. The most important assumptions of the model were that resistance to phage comes with reduced fitness for the host and any particular tail fibre for the phage would adsorb to different receptors with differing efficiency. The model derived certain conditions under which multi-species coexistence was possible. The attacking strategy of the phages caused the hosts to diversify, which in turn led to the phage diversifying as well. Several different sequences of coevolutionary branching were seen to be possible at each stage which somewhat explained the widespread genetic diversity in microbial systems. The paper demonstrated that coexistence of several different mutant strains of bacteria and phage is possible through analytical results and numerical simulation.
After several models came to similar conclusions about the equilibria of host-phage systems [42, 55, 44, 20, 15], Weitz and Dushoff [63] proposed a model that derived some alternate stable states. The novel feature of their model was that the phage were less able to lyse the host cells as the host population reached its carrying capacity. The model was based on virulent phages and hosts. Several works were cited as evidence of such behaviour of hosts and phages in nature [1, 47]. The system had 4 fixed points: complete extinction, host only and two coexistence outcomes. The authors found that a range of values for the initial population density of phages resulted in coexistence. The time at which the virus was introduced into the bacterial culture also altered the final outcomes: when virus was added too early, the host population was insufficient to sustain it; when the phages were added too late, the host population was too close to its carrying capacity and could not be invaded. Virus addition at the right time led to coexistence. If the phages and host were going to coexist, their initial dynamics showed oscillations. In the cases where phages went extinct, a boom-bust pattern of phage population was seen. A modified version of the model was considered in which host mortality only resulted from lysis. This version reconfirmed the results of the original model. Although a conservation law for the system existed, it was proved that coexistence was not the only possible outcome. The relative densities of the hosts and phages were found to be related through a power law. Some of the shortcomings of the model stated in the paper were that it does not incorporate a latency period for the phages, it does not account for resources explicitly and it neglects the lysogenic phase of infection altogether. The authors claimed that more attention needed to be paid to predator-prey dynamics in the microbial world as the results found may differ greatly from those for macro-ecological systems.

Models by Bohannan and Lenski [12] and Weitz [65] also presented two different approaches to bacteria-phage modeling. In [12], the authors compared prey-dependent (PD) and ratio-dependent (RD) host-phage models to investigate which ones were better in terms of predicting empirical results. PD models are based on the assumption that the predation rate of the predator depends directly on the density of the prey, whereas RD models assume that
the rate of predation is proportional to the ratio of predators to prey. RD models were thought by some to be better at capturing the heterogeneity that is found in phage-bacteria interactions because of their intrinsic properties, whereas others researchers have suggested that PD models that used explicit equations to model heterogeneity would be more accurate. The authors found that PD models very accurately predicted the qualitative effects of resource enrichment on phage-host systems and did a better job of predicting experimental results following invasion by phage-resistant bacteria as well. However, quantitatively their results were inferior to those predicted by the RD models, perhaps because PD models were not designed to account for spatial heterogeneity found in the environment. Hence overall, the results were in favour of PD models but in some cases RD models were preferred because of simplicity. The paper by Weitz et al [65] was an attempt to examine high diversity phage and bacteria systems as networks rather than as isolated coupled interactions (as was previously done). They believed that as cross-infection was very common in these communities, studying which phages target which bacteria can lead to important findings about how these interactions affect ecosystems. The main aim of the network models was to determine which infection patterns may not simply arise by chance but are drivers of ecological systems. Although the paper described several infection networks, nestedness, modularity and intermediate patterns (known as diffuse coevolution patterns) were most commonly observed in biological studies. The study concluded that different infection networks can lead to different responses in ecological communities to environmental changes such as resource enrichment. For example, the modular pattern led to increased diversity because of resource enrichment, whereas nested networks exhibited a more unimodal relationship with resource supply. Patterns may also determine the level of antagonism against or even benefits to bacteria in complex communities. The paper also mentioned some advanced experimental and modeling techniques that may be used to understand infection patterns in highly diverse communities of bacteria and phage in the future.

A mathematically elegant continuum between GfG and MA models has been suggested by Agrawal and Lively [2] and recently analysed in detail by Song et al. The model by Agrawal
and Lively [2] is built on the idea that MA and GfG models are at two extreme ends of a continuum describing host and phage interaction patterns. Their results show that MA based models have oscillations, whereas GfG models do not show oscillatory behaviour unless they assume resistance and infectivity costs. Hence costs of virulence and resistance reduce the distinctions between MA and GfG dynamics. As for host recombination, Agrawal and Lively found that it led to higher host fitness for the MA half of the continuum but not necessarily for more GfG like models. Hence GfG type systems may not select for recombination. The model by Song et al. [58] improved on the work by Agrawal and Lively by adding the feature of fluctuating populations through Lotka-Volterra dynamics. Numerical simulations and analytical calculations both show that for constant population sizes the populations oscillate with a single frequency. With varying population sizes another oscillating frequency arises as we move along the MA-GfG continuum. For pure MA models, however, a single frequency exists regardless of the population structure. The models in [2] and [58] were the first to consider multiple host and phage populations interacting with each other and focus on the different results emerging from different infection patterns. However, they do not take into account the powerful impact that phage-prophage recombination can have on these dynamics.

Other than models focused on understanding the biological properties of host-phage systems, some models have also been based on their interesting mathematical properties. The article by Beretta and Kuang [3] analysed the equilibria of marine bacteria and bacteriophages by varying the viral replication factor. Analysis showed that the system could go to the infection-free, endemic equilibrium or positive limit cycles depending on a threshold for the replication factor. The results of the model agreed with experiments in the sense that an endemic was possible for a large range of values of the viral replication factor and lab experiments show oscillatory behavior of the bacteria and phage populations. In [6], the authors improved upon their previous work to develop a model that included a latency period for the virus, and hence a delay differential equation for the infected cells. Two other novel features of this model were a source of mortality for the infected cells in addition to lysis and a constant input of phages.
from the surroundings other than those being produced by the epidemic itself. The latency period placed an additional restriction on the stability of the infection-free and endemic equilibria other than the viral replication factor threshold seen in the previous model. With the input of phage from the surroundings, the reproductive ratio lost its meaning as the phages could never go extinct. A new phage-only equilibrium emerged. All other results were the same as the previous study. Similarly, [4] modelled a bacteria and phage system with delay differential equations in a chemostat setting. Three equilibria were shown to exist: the trivial equilibrium which had no constraints on its existence, the infection-free equilibrium which existed when certain conditions on bacterial growth rate and chemostat washout rate were met and the endemic equilibrium which faced constraints of growth rate, washout rate and latency period for its existence. Generally delays tend to destabilize the positive equilibria but in this case, the endemic equilibrium was stable for very large or small delays and only unstable in an in-between range marked by two threshold values. Close to these threshold delays, the system exhibited periodic solutions. The stability of the infection-free and trivial equilibria was found to be unaffected by the delay.

A few models also considered the effects of introducing reaction diffusion terms and stochastic perturbations into the typical predator-prey models of bacteria and phage systems. Gourley and Kuang came up with a delay reaction diffusion model for phage-host dynamics [34]. The aim of the study was to see how density-dependent phage mortality affects the spatial and temporal dynamics of marine bacteria. The authors found that the dynamics were very sensitive to the phage mortality rate. Another conclusion was that the spreading speed of the infection depended largely on the diffusivity of the bacteria and to a lesser extent on the diffusivity of the virus particles. Beretta et. al investigated the outcome of introducing stochastic perturbations in a Campbell-like host phage model with a latency period [5]. Their model consisted of susceptible bacteria, infected bacteria and phages, interacting through lysis. They found that the stability of the phage-free equilibrium as well as the endemic equilibrium depended on a combination of latency period and lysis rate. Stochastic perturbations were introduced in the
growth rate of bacteria and the inflow/outflow of bacteria from the environment. With the latter perturbation only, the model was found to be robust; there were not many changes in the value of virus population. With perturbations in the both the growth rate and the inflow/outflow rate, the model was found to be weak; the virus population was no longer stable. The paper by Carletti [24] further analysed the model in [3] for robustness under stochastic perturbations. The author introduced white noise-like perturbations in the endemic equilibrium of the model for the range in which the equilibrium was stable and feasible. Results showed that below a certain threshold value for the lysis rate, the equilibrium remained stable under stochastic perturbations. However, for a lysis rate above that threshold, the equilibrium became unstable.

In summary, the above mentioned studies consider different topics such as the equilibria of host-phage systems, the long-term persistence of phage and phage-host coexistence. The coevolution of host and phage has received considerable attention; the effect of resource enrichment on phage-bacteria communities and factors accounting for high genetic diversity of the predator and prey populations have also been considered. Other works have investigated the mathematical properties of phage-host systems based on lytic infection, incorporating delay (caused by viral latency period), using reaction-diffusion dynamics and under the effect of stochastic perturbations. Two recent studies have also addressed alternate stable states of host-phage systems under specific biological conditions and phage-host infection networks. However, very few of these studies consider lysogeny and none address recombination between phage and prophage.

Our work aims to fill this gap by looking at the impact of recombination with prophage sequences on host-phage dynamics and coevolution. Do prophages promote or hinder the co-evolutionary race, or are they silent observers? How does the likelihood of lysogeny affect the outcomes of host and phage interactions? How does the infection pattern between multiple types of host and phage influence the outcome? In the study that follows, we develop several versions of a basic model of bacteria and phage interactions, taking into account the ability of phage to archive genetic material in the host’s DNA, which may be accessed by later
phage through recombination. We look at the fate of these phage-host systems under different conditions. Since the combinatorial nature of these models renders most results analytically unwieldy, numerical methods and simulations will be used to determine the fate of the bacteria and phage within a reasonable parameter space.

Our work predicts that both bacteria and phage populations will exhibit oscillatory behaviour. Recombination can play a key role in maintaining phage populations in parameter regimes that would otherwise favour the bacteria. Diversification can be an effective weapon for the hosts to evade a phage infection. Moreover, as long as it is below a certain threshold, a higher probability of lysogeny may be better for the phage than a lower one. Bacterial hosts tend to accumulate different types of prophage sequences within their genomes over time. (These prophage sequences are likely to acquire deleterious mutations over time that make them non-inducible but they may still provide novel genetic ‘spare-parts’ for the infecting phages). As a result prophage recombination can drive “waves of innovation” in the host cell population. Depending on the prevalent infection pattern, recombination may also support the existence of multiple phage types promoting diversity in the phage-host ecosystem.
Chapter 2

Models with lysis, lysogeny and recombination

2.1 Introduction

In general, phages infect host cells through receptor proteins expressed on the surface of the host cell. The phage can only infect the host if it has a corresponding protein that will bind to the host receptor. In tailed phages, for example, the phage has a tail fibre with affinity to the host cell receptor. In this chapter we give two different versions of the host-phage interaction model, both of which have an S-I-V type structure and are based on systems of non-linear ordinary differential equations. They follow the population densities of uninfected host cells, infected host cells and phages that may infect or lysogenize the uninfected hosts. As we are primarily interested in infection dynamics, we focus on genome sequences that confer host binding ability to the phage. For example, our models would follow sequences encoding the tail fibres of the lambdoid phages. We do not assume that the entire prophage genome is inducible, thus prophage sequences that have undergone substantial mutational decay are still included in the models, as long as the subsequence mediating adsorption remains functional. Likewise, we do not assume that the prophage confers immunity to other lambdoid phages.
Although immunity may have important implications for host-phage coevolution, shifting the balance from parasite toward mutualist [39], our goal in this contribution is to isolate the effects of prophage recombination on the coevolutionary outcome.

The main source of variation between the two models is that in the first version each host can express a single type of receptor only. In the more complex but realistic second version, each host cell can express multiple receptor types in different proportions. Both models predict cycling of the bacteria and phage populations and the importance of recombination to the survival of the phage. The first version of the model shows that the phage population thrives when lysogeny becomes more likely, and the second version demonstrates that host cells tend to accumulate different prophage types within their DNA over time.

2.2 Bacteria-Phage model with 2 receptor types

For simplicity of analysis, in the first version of our model each host has one type of receptor on its surface, either type $J$ or type $K$; each phage likewise has either a $J$- or $K$-type tail fibre, determining the receptor through which it can attach, and thus which host cell population it can infect. The phage population densities are denoted $P_J$ and $P_K$.

Once a phage has bound to the host, the host may become an infected host with probability $(1-p)$ or a lysogen with probability $p$. Lysogens are considered uninfected cells, but acquire the prophage corresponding to the type of phage that lysogenized them. Each type of uninfected host cell population can thus be written as $H_{yz}$ where $y \in \{O, J, K, JK\}$ represents the type of prophage the cell is carrying (if any) and $z \in \{J, K\}$ represents the type of receptor it is expressing. Here subscript $O$ refers to the absence of prophage. This yields a total of eight types of uninfected host cells.

If an uninfected host becomes infected, it falls into one of the eight infected population types denoted by $I_{yz}$, where $y \in \{O, J, K, JK\}$ shows the type of prophage the infected host is carrying, and $z \in \{J, K\}$ defines which type of virus has infected it. We use $N$ to denote the total
density of all uninfected and infected cells, \( N = \sum H_{yz} + \sum I_{yz} \).

The uninfected bacterial population grows logistically with maximum growth rate \( \lambda \) and carrying capacity \( C \); all infected and uninfected host cells contribute to this carrying capacity. The death rate of uninfected host cells is denoted \( \mu \). The parameter \( \beta \) gives the infection rate, assuming mass action kinetics, for both \( P_J \) and \( P_K \). Finally, the infected host cells do not reproduce but their death rate is given by \( \delta \).

An infected cell that was infected by a particular type of phage produces phage of the same type, assuming it does not undergo recombination. A unique feature of this model, however, is that infected cells that carry prophage may undergo recombination to produce phage of a type other than the one infecting it. For example, a cell of type \( H_{JK} \) carries prophage \( J \) and expresses receptor \( K \). This cell can only be infected by \( P_K \), and the resulting \( I_{JK} \) cell typically produces viral offspring of type \( P_K \). However with probability \( r \), recombination occurs between the infecting phage and the prophage within the bacterial DNA. In this case, there is a chance that the infected cell \( I_{JK} \) will produce viral offspring of type \( P_J \).

We assume that infected cells produce new phage at a constant rate \( f \), thus strictly speaking our model captures budding, not lysis in which a burst of phage is released simultaneously [37]. The clearance rate for the phage is \( d \). Finally, over evolutionary time, rare mutations allow the evolution of both host receptors and phage tail fibres (see [39] for a recent review). This allows for transitions: rare \( H_{JJ} \) individuals, for example, will diversify through mutation to become \( H_{JK} \), expressing an alternate receptor. Likewise \( P_J \) can diversify to produce \( P_K \) individuals, a process that can require up to four rare mutational substitutions [46].

These stochastic events could be modelled by including deterministic mutation terms in the population dynamics, however this approach has the serious drawback that rare mutational types are continually generated at low densities – densities corresponding to a fraction of a cell or virus. Including these deterministic approximations in the model can substantially alter the predicted dynamics on realistic time scales. Thus, we take a different approach. We assume that mutation terms are negligible in determining the population dynamics, equilibrium states and
stability. However we acknowledge that over evolutionary time scales, these rare transitions do occur, and thus we allow for the existence of populations such as $H_{JK}$, which may have been produced through the diversification of $H_{JJ}$. In a later section of this chapter, we address the implications of diversification by exploring the impact of these rare mutational transitions on the population dynamics.

Based on the parameters, notation and dynamics described above, the equations for the uninfected host cells can be written as:

\[
\begin{align*}
\dot{H}_{OJ} &= \lambda H_{OJ} \left(1 - \frac{N}{C}\right) - \mu H_{OJ} - \beta P_J H_{OJ}, \\
\dot{H}_{OK} &= \lambda H_{OK} \left(1 - \frac{N}{C}\right) - \mu H_{OK} - \beta P_K H_{OK}, \\
\dot{H}_{JJ} &= \lambda H_{JJ} \left(1 - \frac{N}{C}\right) - \mu H_{JJ} - \beta P_J H_{JJ} + p\beta P_J H_{OJ} + p\beta P_J H_{JJ}, \\
\dot{H}_{KK} &= \lambda H_{KK} \left(1 - \frac{N}{C}\right) - \mu H_{KK} - \beta P_K H_{KK} + p\beta P_K H_{OK} + p\beta P_K H_{KK}, \\
\dot{H}_{KJ} &= \lambda H_{KJ} \left(1 - \frac{N}{C}\right) - \mu H_{KJ} - \beta P_J H_{KJ}, \\
\dot{H}_{JK} &= \lambda H_{JK} \left(1 - \frac{N}{C}\right) - \mu H_{JK} - \beta P_K H_{JK}, \\
\dot{H}_{JKJ} &= \lambda H_{JKJ} \left(1 - \frac{N}{C}\right) - \mu H_{JKJ} - \beta P_J H_{JKJ} + p\beta P_J H_{KJ} + p\beta P_J H_{JKJ}, \\
\dot{H}_{JKK} &= \lambda H_{JKK} \left(1 - \frac{N}{C}\right) - \mu H_{JKK} - \beta P_K H_{JKK} + p\beta P_K H_{JK} + p\beta P_K H_{JKK},
\end{align*}
\] (2.1)

where the dot represents differentiation with respect to time $t$. 

Similarly, the equations for the infected host cells are:

\[ \dot{I}_{OJ} = (1 - p)\beta P_J H_{OJ} - \delta I_{OJ}, \]
\[ \dot{I}_{OK} = (1 - p)\beta P_K H_{OK} - \delta I_{OK}, \]
\[ \dot{I}_{JJ} = (1 - p)\beta P_J H_{JJ} - \delta I_{JJ}, \]
\[ \dot{I}_{KK} = (1 - p)\beta P_K H_{KK} - \delta I_{KK}, \]
\[ \dot{I}_{KJ} = (1 - p)\beta P_J H_{KJ} - \delta I_{KJ}, \]
\[ \dot{I}_{JK} = (1 - p)\beta P_K H_{JK} - \delta I_{JK}, \]
\[ \dot{I}_{JKJ} = (1 - p)\beta P_J H_{JKJ} - \delta I_{JKJ}, \]
\[ \dot{I}_{JKK} = (1 - p)\beta P_K H_{JKK} - \delta I_{JKK}. \] (2.2)

The two phage populations can be represented mathematically through the following equations:

\[ \dot{P}_J = (I_{OJ} + I_{KJ} + I_{JJ} + I_{JKJ})(1 - r)f + (I_{JK} + \frac{1}{2}I_{JKK} + \frac{1}{2}I_{JKJ} + I_{JJ})rf - dP_J, \] \[ \dot{P}_K = (I_{OK} + I_{JK} + I_{KK} + I_{JKK})(1 - r)f + (I_{KJ} + \frac{1}{2}I_{JKJ} + \frac{1}{2}I_{JKK} + I_{KK})rf - dP_K. \] \[ (2.3) \]

The \( \frac{1}{2} \) accompanying the \( I_{JKK} \) and \( I_{JKJ} \) terms in the last two equations reflects the assumption that when recombination occurs in an infected cell containing 2 types of prophage, the probability that the infecting phage will recombine with the correct type of prophage to give a different tail fibre is \( \frac{1}{2} \).

Using the primary experimental literature for \textit{E. coli} and phage \textit{Lambda}, we have determined realistic parameter values for the system above, and are using this newly developed model to investigate the complex dynamics of temperate phage and prophage along several lines of inquiry. For the analytical work presented in this subsection, however, system (2.1)-(2.3) can be reduced by noting that if phage populations are non-zero, over time host cells acquire, but never lose, prophage. Due to the rare evolutionary transitions (mutations) described above, all the host cell types acquire both types of prophage in their DNA. For the analysis of
equilibria and stability, therefore, we can neglect the transient dynamics of prophage acquisition, and focus on the populations remaining at long times. We thus consider the following reduced model.

The host cell population has been reduced to just two types that have either a \( J \) or a \( K \) receptor and both types of prophage sequences. Hence we get the equations

\[
\dot{H}_{JK} = \lambda H_{JK} \left( 1 - \frac{N}{c} \right) - \mu H_{JK} - \beta P_J H_{JK} + p \beta P_J H_{JK},
\]

\[
\dot{H}_{JK} = \lambda H_{KK} \left( 1 - \frac{N}{c} \right) - \mu H_{KK} - \beta P_K H_{KK} + p \beta P_K H_{KK},
\]

for the uninfected host cells. The two types of infected host population arising from the infection of the uninfected hosts can be represented by the equations

\[
\dot{I}_{JK} = (1 - p) \beta P_J H_{JK} - \delta I_{JK},
\]

\[
\dot{I}_{JK} = (1 - p) \beta P_K H_{KK} - \delta I_{KK}.
\]

The phage equations have lost the extra population terms and can be written in their reduced form as follows:

\[
\dot{P}_J = I_{JK}(1 - r)f + \frac{1}{2}(I_{KK} + I_{JK})rf - dP_J,
\]

\[
\dot{P}_K = I_{KK}(1 - r)f + \frac{1}{2}(I_{KK} + I_{JK})rf - dP_K.
\]

In the reduced system (2.4)-(2.6) \( N \) is now simply \( N = H_{JK} + H_{KK} + I_{JK} + I_{KK} \). The parameters \( \lambda, \mu, \beta, p, \delta, f, r \) and \( d \) take positive real values, and \( 0 < p < 1, 0 < r < 1 \). Realistic parameter values, estimated from the experimental literature, are provided with references in Table A.1 of Appendix A.

### 2.2.1 Dimensionless Model and Possible Equilibria

In order to simplify the analysis in the following sections, we first apply scaling on the state variables, the parameters and time to obtain a dimensionless model. To achieve this, introduc-
ing the following scaling:

\[ H_{JJK} = C x_1, \quad H_{JKK} = C x_2, \quad I_{JJK} = C x_3, \quad I_{JKK} = C x_4, \]

\[ P_J = \frac{\delta}{(1-p)\delta} x_5, \quad P_K = \frac{\delta}{(1-p)\delta} x_6, \quad \text{and} \quad \tau = \delta t, \]

into the simple 6-dimensional model yields the dimensionless system (where we still use the dot to indicate differentiation with respect to the new time \( \tau \) for brevity):

\[ \dot{x}_1 = x_1 [A(1 - x_1 - x_2 - x_3 - x_4) - B - x_5], \]
\[ \dot{x}_2 = x_2 [A(1 - x_1 - x_2 - x_3 - x_4) - B - x_6], \]
\[ \dot{x}_3 = x_1 x_5 - x_3, \]
\[ \dot{x}_4 = x_2 x_6 - x_4, \]
\[ \dot{x}_5 = (F + R) x_3 + R x_4 - D x_5, \]
\[ \dot{x}_6 = R x_3 + (F + R) x_4 - D x_6, \]

where the new parameters \( A, B, D, F \) and \( R \) take positive values, defined by

\[ A = \frac{\lambda}{\delta}, \quad B = \frac{\mu}{\delta}, \quad D = \frac{d}{\delta}, \quad F = \frac{f(1-r)(1-p)\beta C}{\delta^2}, \quad R = \frac{f(1-p)\beta C}{2\delta^2}. \]

With the parameter values given in Table A.1, these new dimensionless parameters take the typical values: \( A = 1.0033, B = 0.0374, D = 0.509, F = 2.9164(1-p), R = 0.0001458(1-p). \)

For example, taking \( p = 0.5 \), we have \( F = 1.4582 \) and \( R = 0.00007292 \).

To find the equilibrium solutions of (2.4)-(2.6), we first find four possible groups from the first two equations of (2.8) (i.e., setting \( \dot{x}_1 = \dot{x}_2 = 0 \)):

(i) \( x_1 = x_2 = 0, \)

(ii) \( x_1 = 1 - \frac{B}{A} - x_3 - x_4 - \frac{x_5}{A} = 0, \quad x_2 = 0, \)

(iii) \( x_1 = 0, \quad x_2 = 1 - \frac{B}{A} - x_3 - x_4 - \frac{x_6}{A} = 0, \)
and

\[
\begin{aligned}
\text{(iv) } & \begin{cases}
A(1 - x_1 - x_2 - x_3 - x_4) - B - x_5 = 0, \\
A(1 - x_1 - x_2 - x_3 - x_4) - B - x_6 = 0,
\end{cases} \\
\implies x_5 = x_6.
\end{aligned}
\]

Obviously, Group (i) gives the equilibrium solution \( E_0: (0, 0, 0, 0, 0, 0) \). For Group (ii), \( x_2 = 0 \) generates \( x_4 = 0 \), \( x_5 = \frac{F + R}{D} x_3 \) and \( x_6 = \frac{R}{D} x_3 \). Then, the third equation of (2.8) yields \( 0 = x_1 x_5 - x_3 = \frac{F + R}{D} x_1 x_3 - x_3 (\frac{F + R}{D} x_1 - 1) \), which results in either \( x_3 = 0 \), or \( x_1 = \frac{D}{F + R} \). \( x_3 = 0 \) in turn yields \( x_2 = x_4 = x_5 = x_6 = 0 \), leading to an equilibrium solution: \( E_{1a}: (1 - \frac{B}{A}, 0, 0, 0, 0, 0) \).

If \( x_1 = \frac{D}{F + R} \), then we have \( 1 - \frac{B}{A} = x_3 - \frac{F + R}{AD} x_3 = \frac{D}{F + R} \), from which we obtain \( x_3 = \frac{D (F + R) (A - B) - AD}{(F + R)(F + R + AD)} \), and so \( x_5 = \frac{(F + R) (A - B) - AD}{F + R + AD} \) and \( x_6 = \frac{R (F + R) (A - B) - AD}{(F + R)(F + R + AD)} \), giving an equilibrium solution:

\[
\begin{aligned}
E_2: \left( \frac{D}{F + R}, 0, \frac{D (F + R) (A - B) - AD}{(F + R)(F + R + AD)}, 0, \frac{R (F + R) (A - B) - AD}{(F + R)(F + R + AD)}, \frac{(F + R) (A - B) - AD}{F + R + AD} \right). \\
\end{aligned}
\]

A similar analysis on Group (iii) gives two equilibrium solutions: \( E_{1b}: (0, 1 - \frac{B}{A}, 0, 0, 0, 0) \), and

\[
\begin{aligned}
E_3: \left( 0, \frac{D}{F + R}, \frac{D (F + R) (A - B) - AD}{(F + R)(F + R + AD)}, \frac{R (F + R) (A - B) - AD}{(F + R)(F + R + AD)}, \frac{(F + R) (A - B) - AD}{F + R + AD} \right). \\
\end{aligned}
\]

For Group (iv), we have \( x_6 = x_5 \), and need solve \( A(1 - x_1 - x_2 - x_3 - x_4) - B - x_5 = 0 \) together with the remaining four equations in (2.8). Due to \( x_6 = x_5 \), we obtain \( x_3 = x_4 = \frac{D}{F + 2R} x_5 \), and then we get \( x_1 x_5 = \frac{D}{F + 2R} x_5 \) and \( x_2 x_5 = \frac{D}{F + 2R} x_5 \), which results in either \( x_5 = 0 \), or \( x_1 = x_2 = \frac{D}{F + 2R} \). If \( x_5 = 0 \), we then have \( x_3 = x_4 = x_5 = x_6 = 0 \) and \( x_1 + x_2 = 1 - \frac{B}{A} \). This yields an equilibrium line segment:

\[
\begin{aligned}
E_1: \ (x_1, x_2, 0, 0, 0, 0), \ \text{satisfying } x_1 + x_2 = 1 - \frac{B}{A}, \ x_1 \geq 0, \ x_2 \geq 0. \\
\end{aligned}
\]

It is easy to see that the equilibrium solutions \( E_{1a} \) and \( E_{1b} \) are special cases of the equilibrium \( E_1 \) (\( E_{1a} \) and \( E_{1b} \) are the two end points of \( E_1 \)). Hence, in the following, we shall include the \( E_{1a} \) and \( E_{1b} \) into the discussion of \( E_1 \). While for the second case in Group (iv), namely
\( x_5 \neq 0 \), we have a positive equilibrium, given by

\[
E_4 : \left( \frac{D}{F+2R}, \frac{D}{F+2R}, \frac{D[(F+2R)(A-B)-2AD]}{(F+2R)(F+2R+2AD)}, \frac{D[(F+2R)(A-B)-2AD]}{(F+2R)(F+2R+2AD)}, \frac{(F+2R)(A-B)-2AD}{F+2R+2AD}, \frac{(F+2R)(A-B)-2AD}{F+2R+2AD} \right).
\]

(2.13)

2.2.2 Numerical Simulations

To illustrate and confirm these analytical results, simulations for system (2.4)-(2.6) were performed by numerical integration (ODE45 package in MATLAB). Parameter values were as provided in Table A.1, with the exception of \( \mu \), which was varied to get the model to exhibit different equilibria through changes in the composite parameter \( B = \mu/\delta \). The case in the absence of recombination was also tested for, which implies \( r = 0 \) and therefore \( R = 0 \) in the non-dimensionalized model. Results were plotted as population densities versus time. Unless otherwise noted, initial conditions include both types of uninfected host cells and both phage types, thus simulating the invasion dynamics as well as equilibrium conditions.

In Figure 2.1(a) \( B = 1.02 > A \), such that the system approaches the trivial equilibrium. When \( B \) is reduced slightly to \( B = 0.9 \), the system converges to equilibrium \( E_1 \); only uninfected host cells persist. This is illustrated in Figure 2.1(b).

Figure 2.2 illustrates the predicted effect of diversification. In this simulation we take \( B = 0.5 \). When the initial conditions only include host cells with receptor \( J \) (neither \( H_{JKK} \) nor \( I_{JKK} \) is present), the system converges to equilibrium \( E_{2/3} \), which is unstable in the full system but stable when \( H_{JKK} = I_{JKK} = 0 \). Thus, when the complementary host cell population \( H_{JKK} \) is added, the equilibrium loses stability. The system converges to the stable equilibrium \( E_1 \) of the full system, and both phage populations decay to zero. Thus the diversification of the host cell population from a single type to multiple types can potentially drive both phage populations to extinction.

Figure 2.3 illustrates the cases \( B = 0.3028 \) and \( B = 0.25 \). We observe stable convergence to equilibrium \( E_4 \), with all populations present (panels (a) and (b)), and a stable limit cycle around
2.2. Bacteria-Phage model with 2 receptor types

E_4 (panels (c) and (d)) as expected. We note that at the parameter values we have chosen from the literature, the parameter regime for stable equilibria is quite narrow, and thus our results predict that oscillations would be commonly observed in natural phage-host systems.

Finally, we numerically investigated the effect of recombination, r. Starting in the regime \( B < 0.25 \), we investigate a situation in which both types of phage are initially present, and the host cells diversify to escape the phage. Thus we begin with type J host cells, and both viral types, but at a later time introduce type K host cells.

In Figure 2.4(a), the resulting dynamics are shown in the absence of recombination. We see that phage \( P_K \) goes extinct early in the simulation because of the lack of type K host cells. Later, when host cells with K receptors are introduced, these host cells compete with and reduce the population of type J host cells. The uninfected \( H_{JKJ} \) population is no longer sufficiently large to maintain phage, and ultimately phage J is also unable to survive. The system converges to \( E_1 \).

In contrast, Figure 2.4(b) illustrates the same results in the presence of recombination. Before the introduction of the type K host cells, type K phage is present at low levels due to recombination with prophage in the host cell genome. When hosts with K receptors are introduced, the system approaches a limit cycle in which all populations are present. Thus recombination protects the phage populations from extinction.

2.2.3 Optimal Rate of Lysogeny p

Several studies have claimed that temperance is a strategy used by the phage to survive in times when the bacterial population is too low to sustain them otherwise \[60, 55\]. Berngruber et al. have found that lysis is very high at the beginning of an epidemic but as the infection progresses, lysis starts giving way to lysogeny as the microbial community becomes saturated with the phages \[8\]. Recent experimental work has shown the probability of lysogeny to lie in the range 0.1 to 0.6 depending on the level of virulence of the virus \[8, 9\]. In this section we investigate whether there could be an optimal value of lysogenization rate for the long-term
propagation of the phages and if that optimal value lies within the range found experimentally.

System (2.4)-(2.6) was numerically integrated for 31 different values of $p$ between 0 and 1 to find a value that would result in the highest total final population of phages. The number 31 was selected arbitrarily because it seemed to give a good approximation of the entire range. The simulations were run for a time period of 500 days each time to allow most of the bacteria to become infected. The value of $B$ was fixed at 0.037, both types of host $H_{JKI}$ and $H_{JKK}$ and phages $P_J$ and $P_K$ were present initially but the initial population of infected cells was kept at zero. All other parameter values were taken as given in Table A.1 with the exception of carrying capacity $C$, which was taken to be $10^8$. This increase was necessary to maintain the reproductive ratio at a suitable value when multiplying the infection rate of the phage with larger values of $p$. We plotted the final $P_J$ and $P_K$ populations against $p$. The result is given in Figure 2.5. It is clear from the graph that high values of $p$ between 0.3 and 0.9 are actually better for the long-term (approximated by 500 days) survival of the phage. As $p$ rises from 0 to about 0.9, we see an increase in the final populations of the phages. This range overlaps with the experimental range of 0.1 to 0.6 but the optimal value of lysogenization rate is about 0.9 which is even higher than the experimental estimates found. Clearly the precise value of the optimum will depend on parameter values, however of greatest interest here is that an optimum, greater than zero, exists. This result is fascinating because it sheds light on a potential long-term infection strategy that the phages may favor in order to be more successful at infecting a bacterial community. It seems that at a lower infection rate translates into a larger overall host cell population which can support a larger number of phages in the long-run. Another effect that might be coming into play is that lysogeny fuels recombination between infecting phage and prophage. This allows the phage to diversify and survive better in situations where host diversification lowers the phage population (seen in Figure 2.4). Hence our model not only reconfirms the theories of phage survival tactics presented in [60, 55, 8] but also gives an estimate of an optimal lysogenization rate for a particular set of parameter values. Future experimental work may help to corroborate or find a better estimate for the value we have
Although this version of the model simplifies many aspects of the underlying biology, the clearest direction for future work was to relax the assumption that host cells express only one receptor type, as host cells are known to up- and down-regulate various receptors in response to phage pressure [46]. This is what we did for the next part of our analysis of phage-host systems.

## 2.3 Bacteria-Phage Model with Host cells having 2 receptor types in different proportions

For the second version of our host-phage model, the mode of infection of the host by a phage is the same as that in the previous version, i.e. the phage attaches to the host using a surface receptor. The main distinguishing factor between the two versions is that in this version, each host has two kinds of receptors on its surface, $J$ and $K$, and one type of host population can be distinguished from another based on the proportion of each type of receptor it expresses. In contrast, in the previous version, each host type could only express either a $J$- or a $K$-type receptor on its surface. We believe that relaxing this assumption brings our model closer to biological reality without overcomplicating the analysis, as bacterial cells found in nature can have several different types of receptors on their surfaces. A phage can have either a $J$- or a $K$-type tail fibre, determining which surface receptor it will attach to, and also which hosts will be most vulnerable to it.

The probability of an infection resulting in lysis is again $(1 - p)$ whereas the probability of lysogeny is given by $p$. Lysogens are still categorized as uninfected cells. Based on the types of prophage they are carrying (or not carrying), the hosts can be divided into four categories: i) hosts without prophage ($H_0$), ii) hosts with $J$ prophage ($H_J$), iii) hosts with $K$ prophage ($H_K$) and iv) hosts with $J$ and $K$ prophage ($H_{JK}$), where the subscript denotes the type of prophage being carried. If an uninfected host becomes infected, it can fall into one of the eight
infected population types denoted by $I_{yz}$, where $y \in \{O, J, K, JK\}$ shows the type of prophage the infected host is carrying (if any), and $z \in \{J, K\}$ defines which type of phage has infected it. As in the first version of the model, there are two phage populations, namely $P_J$ and $P_K$.

The bacteria have a logistic growth term with growth rate $\lambda$, and death term $\mu$. All host cell populations are functions of two continuous variables, denoted $H_y(t, \rho)$. Here $t$ is time, while the variable $\rho \in [0, 1]$ reflects the proportion of total receptors on the host’s surface that are of type $J$. Hence the proportion of $K$ receptors is given by $(1 - \rho)$. We assume that each host cell has an equal number of total receptors. For the purpose of our numerical simulation, $\rho$ was given a vector of equally spaced values between 0 and 1. $\beta_J$ and $\beta_K$ are the infectious rates for $P_J$ and $P_K$ respectively and are directly proportional to the number of each type of receptor. Thus $\beta_J(\rho) = \beta \rho$ and $\beta_K(\rho) = \beta(1 - \rho)$. The infected host cells do not reproduce but their death rate is $\delta$. The carrying capacity for all the hosts (uninfected and infected) is $C$. $N$ is the total population size of all uninfected and infected cells, integrated over all values of $\rho$, $N = \int_{\rho=0}^{\rho=1} \left( \sum H_y + \sum I_{yz} \right)$.

The lysis of infected cells produces new phages at the rate $f$. An infected cell that was infected by a particular type of phage typically lyses to produce phages of that type. However, recombination does allow progeny of a different type to be produced as long as that type of prophage is being carried by the infected bacterium. The probability of recombination is $r$. The clearance rate for the phage is $d$.

An important point to understand here is that the host population consists of host cells with different surface receptor proportions, which need to be kept track of for the purpose of further analysis. Hence the host equations to follow vary with time as well as $\rho$. However, the host cells are all interacting with the same virus pool simultaneously, so $\rho$ cannot be treated strictly as another parameter. The equations for the host populations (differing on the basis of the type of prophage they are carrying) can be written as follows:
2.3. **Bacteria-Phage Model with Host cells having 2 receptor types in different proportions**

\[
\frac{\partial H_0(t,\rho)}{\partial t} = \lambda H_0(t,\rho) \left(1 - \frac{N}{C}\right) - \mu H_0(t,\rho) - \beta_J(\rho) P_J H_0(t,\rho) - \beta_K(\rho) P_K H_0(t,\rho),
\]

(2.14)

\[
\frac{\partial H_J(t,\rho)}{\partial t} = \lambda H_J(t,\rho) \left(1 - \frac{N}{C}\right) - \mu H_J(t,\rho) - \beta_J(\rho) P_J H_J(t,\rho)
- \beta_K(\rho) P_K H_J(t,\rho) + p\beta_J(\rho) P_J H_J(t,\rho) + p\beta_J(\rho) P_J H_0(t,\rho),
\]

(2.15)

\[
\frac{\partial H_K(t,\rho)}{\partial t} = \lambda H_K(t,\rho) \left(1 - \frac{N}{C}\right) - \mu H_K(t,\rho) - \beta_J(\rho) P_J H_K(t,\rho)
- \beta_K(\rho) P_K H_K(t,\rho) + p\beta_K(\rho) P_K H_K(t,\rho) + p\beta_K(\rho) P_K H_J(t,\rho).
\]

(2.16)

\[
\frac{\partial H_{JK}(t,\rho)}{\partial t} = \lambda H_{JK}(t,\rho) \left(1 - \frac{N}{C}\right) - \mu H_{JK}(t,\rho) - \beta_J(\rho) P_J H_{JK}(t,\rho) - \beta_K(\rho) P_K H_{JK}(t,\rho)
+ p\beta_J(\rho) P_J H_{JK}(t,\rho) + p\beta_K(\rho) P_K H_{JK}(t,\rho) + p\beta_J(\rho) P_J H_J(t,\rho) + p\beta_K(\rho) P_K H_J(t,\rho).
\]

(2.17)

The infected cell populations are differentiated amongst on the basis of the prophage they carry, as well as the type of phage infecting them. The proportion of different receptors is not kept track of for the infected cells as they are no longer vulnerable to infection. Hence we get
the following infected cell population equations:

\[
\begin{align*}
\frac{dI_{OJ}}{dt} &= (1 - p)\beta_J P_J H_O - \delta I_{OJ}, \\
\frac{dI_{OK}}{dt} &= (1 - p)\beta_K P_K H_O - \delta I_{OK}, \\
\frac{dI_{JJ}}{dt} &= (1 - p)\beta_J P_J H_J - \delta I_{JJ}, \\
\frac{dI_{JK}}{dt} &= (1 - p)\beta_K P_K H_J - \delta I_{JK}, \\
\frac{dI_{KK}}{dt} &= (1 - p)\beta_K P_K H_K - \delta I_{KK}, \\
\frac{dI_{JJK}}{dt} &= (1 - p)\beta_J P_J H_{JK} - \delta I_{JJK}, \\
\frac{dI_{JKJ}}{dt} &= (1 - p)\beta_J P_J H_{JK} - \delta I_{JKJ}, \\
\frac{dI_{JKK}}{dt} &= (1 - p)\beta_K P_K H_{JK} - \delta I_{JKK}.
\end{align*}
\]  

(2.18)

The equations for the phage populations are:

\[
\begin{align*}
\frac{dP_J}{dt} &= (I_{OJ} + I_{KK} + I_{JJK})(1 - r)f + (I_{JK} + \frac{1}{2}I_{JJK} + \frac{1}{2}I_{JKJ})r f - dP_J, \\
\frac{dP_K}{dt} &= (I_{OK} + I_{JK} + I_{JKK})(1 - r)f + (I_{KJ} + \frac{1}{2}I_{KJJ} + \frac{1}{2}I_{JJK} + I_{KK})r f - dP_K.
\end{align*}
\]  

(2.19)

2.3.1 Results

We solved system (2.14)-(2.19) numerically using the ODE45 package in MATLAB. The results are illustrated in the figures in this section. Some of the resulting figures are 3 dimensional with time on the x axis, the proportion of receptors \( \rho \) on the y axis and the population density of hosts or phages on the z axis. Sometimes the x and y axes look interchanges on the 3 dimensional graphs. This is just because the plot have been rotated to give a clear view of the population dynamics. For the two dimensional figures, the x axis shows time and the y axis shows the population density. The simulations were run for 60 days. All parameter values used are given in Table A.2 with relevant references. The value used for probability of lysogeny \( p \) for this model is 0.2 instead of 0.5 used earlier. Experimental results found the actual value of
2.3. Bacteria-Phage Model with Host cells having 2 receptor types in different proportions

\( p \) to lie between 0.1 and 0.6 [8]. We picked a more conservative value of 0.2 here as a higher value of \( p \) would only help to enhance the effects of temperance and strengthen our case (refer to section 2.2.3 for an example). The equations were numerically integrated for 11 equally spaced values of \( \rho \) between and including 0 and 1. The variable \( \rho \) has been defined so that \( \rho = 0 \) means that the cell has only \( J \) receptors, \( \rho = 1 \) means that it has only \( K \) receptors and \( \rho = 0.5 \) means that the cell has equal proportions of both \( J \) and \( K \) receptors on its surface.

Figure 2.6 shows the time course of system (2.14)-(2.19) starting with only \( H_O \) and both types of phage \( P_J \) and \( P_K \). Figure 2.6a shows the changes in the \( H_O \) population over time. Although at time 0 all the cells are of type \( H_O \), within the first 5 days, other host populations also start to rise and \( H_O \) declines rapidly to 0 across all values of \( \rho \). Figure 2.6b shows more interesting behaviour. The population of \( H_J \) starts at 0 but begins to rise. The gradient of the increase is greater when \( \rho \) is greater, i.e. when the majority of receptors are of \( J \) type. Following the sharp increase, there is a decline in the population for all values of \( \rho \) except \( \rho = 1 \). The host populations show a decline because host cells acquire both types of prophage and become \( H_{JK} \). However, hosts with \( \rho = 1 \) only have \( J \) receptors and are able to obtain only \( J \) prophage. The graph of \( H_K \) in Figure 2.6c is very similar to that of \( H_J \). It looks like a reflection of the latter in the \( y \) axis. The population starts from zero and rises to its peak and the rise is faster for lower values of \( \rho \). It then starts to drop until it becomes zero again for all values of \( \rho \) other than 0. Thus it is obvious that host cells with more \( J \) receptors are more likely to acquire \( J \), whereas hosts with more \( K \) receptors are more likely to acquire \( K \) prophage. Figure 2.6d shows that the population of \( H_{JK} \) starts at 0 and almost achieves the carrying capacity level of population by the end of the simulation time. The increase in the population is equal for all values of \( \rho \) except \( \rho = 0 \) and \( \rho = 1 \). At the two extreme values of \( \rho \), the population of \( H_{JK} \) is 0 because all the host cells at \( \rho = 0 \) have only \( K \) receptors and hence cannot acquire \( J \) prophage; similarly all the host cells at \( \rho = 1 \) have only \( J \) receptors and hence can only get \( J \) prophage. However, all the host cells within the interval \( \rho \in (0,1) \) eventually acquire both types of prophage, as they come into contact with both \( P_J \) and \( P_K \) over
several generations. It is important to mention that we do not explicitly account for the loss of prophage in the model equations. As mentioned in the introduction of this chapter, this is because our definition of prophage includes complete as well as partial prophage sequences as long as they have enough functional modules to successfully take part in adsorption to a host cell (references given in the introduction). Figure 2.6e shows that the phage population rises and reaches a stable equilibrium in less than 10 days. Although there are variations in the different host populations, these do not affect the phage populations as the total population of all types of hosts remains near carrying capacity, providing the phages with a constant source of susceptible host cells.

For some of the following simulations, we will assume that the hosts and phages have been interacting for some time and have acquired both $J$ and $K$ prophage. And hence we reduce the 14-dimensional system (2.14)-(2.19) to the 5-dimensional system based on $H_{JK}$, $I_{JK}$, $I_{J}$, $P_{J}$ and $P_{K}$. In such cases, the other host and phage populations will not be initialized. All other parameters will be taken as described in Table A.2 unless specified otherwise.

As mentioned in the introduction in chapter 1, the mechanism employed by the hosts to evade the phages is to express fewer and fewer of the receptors that the phages are using to adsorb to the host. As a response, phages that adsorb to a different receptor type start doing better and inevitably become the dominant phage population in the community. In this case, the bacteria may again defend themselves by down-regulating this new more vulnerable receptor, followed by another change in target receptor by the resilient phages. Hence the different host and phage populations may continue to cycle based on frequency-dependent selection. We wanted to see if our current model can predict the existence of these cycles. We ran a simulation with the majority of phages being of type $P_{J}$ and very few of type $P_{K}$ initially. The results are shown in Figures 2.7 and 2.8.

For Figure 2.7 we began with a large number of $P_{J}$ in the phage population (and very few $P_{K}$), making the hosts with $J$ receptors more likely to get infected. As a result we see in Figure 2.7a there is a decline in the host population with mostly $J$ receptors around day 1, matched
by a sharp increase in the hosts that have mostly $K$ receptors. This is because the large $P_J$ population infects and lyses a large number of hosts. The hosts that are partially immune to $P_J$ end up being at an advantage, and increase in number to take over the resources left behind by their less immune counterparts. However, with more hosts that are vulnerable to $P_K$ in the community, the population of $P_K$ increases and the hosts with mostly $K$ receptors decline in number. The oscillations in the phage populations can be seen in Figure 2.7b. Our simulations show that this cycling of populations continues for a long time hinting at the existence of a stable limit cycle.

Figure 2.8 shows the results when the simulation is initiated with a large $P_K$ population and a negligible $P_J$ population. The resulting dynamics are the mirror image of those seen in Figure 2.7 with the host population with mostly $K$ receptors declining first and the hosts with mostly $J$ receptors following suit. Both the host and phage populations continue to oscillate as seen in Figures 2.8a and 2.8b. One intriguing feature of the dynamics of the host populations, which is also clearly visible in the graphs, is that when there is a peak on one end of the $\rho$ axis, there is a valley on the opposite end. This effect comes about due to the change in the proportion of different receptors being expressed and how that affects the susceptibility of the host cells (as explained above). For example if the phage population is being dominated by $P_K$ at any time step, then we expect the population of hosts with mostly $K$ receptors to be low as a consequence. At the same time, the population of hosts with mostly $J$ receptors and hence very few $K$ receptors will be high because of their partial immunity to $P_K$. Another interesting feature of the host-phage dynamics is that is the oscillatory behaviour of the populations. The previous version of this model analysed in section 2.2 also showed that the host and phage populations are likely to cycle, and our current version of the model confirms this result. As described by [27], these oscillations are a common feature of matching allele models and therefore do not come as a surprise.

In order to test if the recombination-related results shown for system (2.4)-(2.6) in Figure 2.4 also hold for the current model, we simulated 2 different cases. In the first case depicted
in Figures 2.9 and 2.10, the initial host populations consisted of $H_0$ cells with a majority of $K$ receptors ($\rho = 0.1$ to $\rho = 0.6$) and the initial phage population was only made up of phage $P_J$; a few hosts with both prophage types were present. Without recombination, the hosts were able to evade the phages in this case by expressing very few $J$ receptors. The phage population went extinct. With the help of recombination, however, the phage population diversified from $P_J$ to $P_K$ and eroded the advantage gained by the host from expressing fewer $K$ receptors. Both phage populations thrived.

In the second case illustrated in Figures 2.11 and 2.12, the simulation was initiated with $H_0$ host cells with a range of different receptor $J$ proportions from $\rho = 0.1$ to $\rho = 0.9$. Once again the phage population consisted of only $P_J$ but some host cells carrying $K$ prophage were also present. When recombination was not possible, the hosts expressed the minimum possible number of $J$ receptors and $P_J$ went extinct before $P_K$ could invade. On the other hand, with recombination, both the phage populations survive and continue to oscillate with decreasing amplitude. Most host cells also acquire both types of prophage and the host population exhibits oscillations with populations at extreme values of $\rho$ increasing and and decreasing alternatively.

The above cases show that even after modifying the assumption that host cells can express only one type of receptor, recombination plays a major role in deciding the fate of host-phage dynamics. Hence the effect of recombination shown in our results is not sensitive to the assumption of one receptor per host cell. Our results agree with biological scenarios in which host defensive tactics consist of altering receptor proportions, and recombination helps the phage to counter such host defenses [46].

### 2.4 Discussion

In conclusion we interpret the implications of the results of our models. For the first version, numerical simulations show that when the death rate is very high, the host cell population cannot sustain itself and the trivial equilibrium is stable. At intermediate death rates, the host
cells are able to sustain themselves, and the two types can coexist at any ratio, such that the population densities sum to a constant. However, the host cell density must be even higher to allow the phage populations to invade before going extinct.

An intriguing prediction emerges from this study. In Figure 2.2, we illustrate a situation in which initially only one type of host, $H_{JK1}$, exists. We allow both types of prophage to exist by recombination, although this does not affect the result. If we take $B = 0.5$ the system converges to equilibrium $E_{2/3}$; the host population sustains the $P_J$ population, $P_K$ is produced at a low level by recombination, and a mix of $H_{JKJ}$ and $I_{JK1}$ survive at equilibrium. Now however, if the host diversifies and is able to produce $H_{JKK}$ individuals, the system will converge to equilibrium $E_1$, which includes only the two uninfected host cell populations. Thus, by diversifying into two distinct populations, the host drives the phage populations to extinction.

This prediction holds even if the phage can diversify equally fast; both types of phage are present and nonetheless they do not persist. The underlying issue in this example is that each type of phage requires a certain minimum density of susceptible host cells, such that the basic reproductive ratio for that phage type exceeds one. When only one type of host cell exists, that host type can grow to the carrying capacity, and sustain the corresponding phage type; recombination will stably maintain the other phage. However at the same parameter values, if two host cell types coexist, neither has sufficient density to maintain their phage predators. Expanding this to a real-world situation with many host and phage types, we predict that host cell populations that are able to diversify, such that only a subset of cells are susceptible to a specific phage type, may be able to drive several phage types to extinction. One caveat is that this phenomenon is only possible in the possibly limited parameter range $0.3030 < B < 0.6531$.

A related prediction highlights the effect of recombination, as illustrated in Figure 2.4(b). In this scenario, we take parameter values in the region $B < 0.3030$, and again consider a situation in which initially only one type of host cell exists. In this region, the host cell population stably maintains the corresponding phage population. We then introduce the second host cell population by diversification. If recombination is possible, the system converges to $E_4$, and
we observe all six populations at equilibrium or in an oscillatory pattern. However this result critically depends on recombination. If we set the recombination rate to zero, the system converges to $E_1$ when the second host cell type is introduced. Thus once again, diversification of the host population can drive the viral population to extinction, but in this parameter regime extinction is only possible if recombination does not occur. In other words, the ability to recombine with prophage in the host genome is critically important to the phage population, and can save the phage from extinction. The condition $B < 0.3030$ means we expect this scenario to be relevant to a wide range of host-phage systems.

Finally, another unexpected result that builds upon some of the previous findings is that a high (between 0.3 and 0.9) value of the probability of lysogeny $p$ may favour a higher phage population in the long run. The value of $p$ that maximizes the long-term reproductive success of the phage population lies around 0.9 (for the parameter values we use here) as seen in Figure 2.5. As defined in our model, an increase in $p$ automatically lowers the lytic infection rate of the phages. It appears that when facing a reduced infection rate, the host cells are able to propagate better and thus more of them are available for the phages to infect in the long run. Hence lysogeny presents a kind of a delayed attack strategy for the phages and attacking at the right time is most beneficial for the phage population. It must be noted that taking a larger number of values of $p \in (0, 1)$ will help to further refine the value found for optimal probability, however the main result here is not the quantitative value of the optimum, but that the optimum exists. In particular, the value of the optimal probability is dependent upon the $R_0$ for the model. If the $R_0$ is too low the phage will not be able to invade the population of hosts successfully, making the question of the long-term survival meaningless.

The second version of our model can also be used to gain some interesting insights. Host cell genomes have been found to carry between 7 and 20 different prophages simultaneously [25] and our results show how a particular host cell may be able to acquire multiple prophages over the course of just a few days. Figure 2.6 illustrates how the entire host population starts out as $H_0$ but all host cells, except for those with only one type of receptor, end up as $H_{JK}$. 
This is because lysogeny results in the phage DNA being integrated within the host DNA for long periods of time [17]. Since the host cells have receptors that can be attacked by $P_J$ and $P_K$, the host cells acquire both $J$ and $K$ prophage to become $H_{JK}$. There is a possibility that this mechanism may be the reason why host cells in nature have been found to have multiple prophages. Generally bacterial cells express many different receptors, several of which can be used by the phages as points of entry. If exposed to different types of phage it seems reasonable that the hosts will acquire different types of prophage.

Our findings also indicate that we can expect the phage and host populations to cycle between the different types that exist. Figures 2.7 and 2.8 show that regardless of which type of phage initially exists, when one of the phage populations is dominant, the host and phage populations will begin to rise and fall alternatively. The fates of the different populations are so closely related to each other that a rise in one is necessarily followed by a fall in another and this cycling appears to continue indefinitely in the chosen parameter regime. Moreover, the cases depicted in Figures 2.9, 2.10, 2.11 and 2.12 also illustrate that recombination can be essential to the survival of the phage when the host population defends itself by expressing fewer of the receptors being targeted by the phages. In cases where recombination is possible, the phage populations $P_J$ and $P_K$ maintain each other when there is a scarcity of susceptible hosts. Whereas in the absence of recombination, the infection gets eliminated soon after the host cells down-regulate the target receptors.

These observations reconfirm the predictions arising from system (2.4)-(2.6) about frequency-dependent oscillations in the host and phage populations; previous studies have claimed that bacteria-phage dynamics may consist of selective ‘sweeps’ (in which the more fit phages and bacteria replace the less fit ones), followed by negative frequency-dependent selection that continues indefinitely [36]. The graphs from system (2.4)-(2.6) and system (2.14)-(2.19) seem to corroborate the existence of these negative frequency-dependent selective pressures, arising from the diversity in host and phage populations, and resulting in Red Queen dynamics. The results also reaffirm that recombination can keep phage populations from going extinct through
periods of dearth in susceptible hosts. Since these results are observable through both the models discussed in this chapter despite their different designs, they are more likely to be inherent features of the host-phage dynamics rather than artifacts of the model design.

Overall, the work we present for these models highlights the importance of prophage and recombination to the equilibria of phage-host systems. Since our simple models can simulate the Red Queen dynamics found in phage-host systems, it seems likely that a more sophisticated version of these models, that takes into account mutation of phages and bacteria, may also be able to confirm or contradict the existence of arm’s race dynamics that leads to selective sweeps. This appeared as an obvious avenue of subsequent research that has been analysed in the following chapter of this document.
2.4. Discussion

Figure 2.1: (a) Time course of system (2.4)-(2.6) with $B = 1.02$. The system approaches the trivial equilibrium with all populations going extinct. Other initial conditions were $H_{JKJ} = \frac{2C}{3}$, $H_{JKK} = \frac{C}{3}$, $I_{JKJ} = I_{JKK} = 0$ and $P_J = 2 \times 10^5$, $P_K = 3 \times 10^5$. (b) Time course of system (2.4)-(2.6) with $B = 0.9$. In this parameter regime, the system converges to $E_1$ with only uninfected host cells remaining. Other initial conditions were $H_{JKJ} = \frac{2C}{3}$, $H_{JKK} = \frac{C}{3}$, $I_{JKJ} = I_{JKK} = 0$ and $P_J = 2 \times 10^5$, $P_K = 3 \times 10^5$.

Figure 2.2: Time course of system (2.4)-(2.6) with $B = 0.5$ and initial conditions $H_{JKJ} = \frac{2C}{3}$, $H_{JKK} = 0$, $I_{JKJ} = \frac{C}{3}$, $I_{JKK} = 0$ and $P_J = 2\times 10^5$, $P_K = 3 \times 10^5$. The system initially converges to equilibrium $E_{2/3}$. At time $t = 10$, we introduce $H_{JKK} = 100$, and the system rapidly converges to $E_1$, eliminating the phage populations.
Figure 2.3: Panels (a) and (b) show the time course for system (2.4)-(2.6) for $B = 0.3028$, with initial conditions $H_{JKJ} = H_{JKK} = \frac{C}{3}$, $I_{JKJ} = I_{JKK} = 0$ and $P_J = P_K = 2 \times 10^5$. The system converges to the stable equilibrium $E_4$. Panels (c) and (d) show the time course for $B = 0.25$, with initial conditions $H_{JKJ} = \frac{2C}{3}$, $H_{JKK} = \frac{C}{3}$, $I_{JKJ} = I_{JKK} = 0$ and $P_J = 2 \times 10^5$, $P_K = 3 \times 10^5$. We observe a stable limit cycle.
2.4. Discussion

Figure 2.4: (a) Extinction of phage in the absence of recombination. A time course is shown for system (2.4)-(2.6) for $B = 0.037$, and initial conditions $H_{JK} = \frac{2C}{3}$, $H_{KK} = 0$, $I_{JK} = I_{KK} = 0$ and $P_J = 2 \times 10^5$, $P_K = 3 \times 10^5$. The parameter $r$ was set to zero to analyse the absence of recombination. At time $t = 10$, $H_{KK} = 100$ is introduced (the host diversifies). Note that phage $P_K$ goes extinct at early times due to the lack of host cells, while $P_J$ goes extinct once the new type of host is introduced. Only uninfected host cells persist. (b) Survival of phage in the presence of recombination. Parameter values and initial conditions are as described for Figure 2.4(a), with the exception that $r = 0.0001$. Recombination preserves the phage population from extinction.

Figure 2.5: Variation in the final phage populations with $p$ for system (2.4)-(2.6). Initial conditions were $B = 0.037$, $H_{JK} = \frac{2C}{3}$, $H_{KK} = \frac{C}{3}$, $I_{JK} = I_{KK} = 0$ and $P_J = 2 \times 10^5$, $P_K = 3 \times 10^5$. The graph shows the result for the final population of the phages at the end of 500 days for 31 different values of $p$ between 0 and 1. The phage population grows with the increase in $p$ between 0.3 and 0.9, reaching its maximum at 0.9. Increasing $p$ beyond 0.9 causes the infection rate to become too low and hence the phage population starts to decrease until it reaches 0 at $p = 1$. 
Figure 2.6: Figures show the time course of the 4 different host populations in for varying proportions of $J$ and $K$ receptors based on system (2.14)-(2.19). Initial conditions were $H_O = 10000/11$, $H_J = H_K = H_{JK} = 0$ and $P_J = 3000$, $P_K = 3000$. The figures show that the entire host population is initially concentrated under $H_O$ as expected. Almost immediately after, (b) shows how host cells with a majority of $J$ receptors acquire $J$ prophage. Whereas (c) shows how hosts with mostly $K$ receptors acquire $K$ prophage. Eventually all hosts with $\rho \in (0, 1)$ acquire both $J$ and $K$ prophage and become part of the $H_{JK}$ population as seen in (d). (e) shows that the two phage populations increase and become stable at about $10^8$ virions. This is due to the stable supply of susceptible hosts available for them to infect.
2.4. Discussion

Figure 2.7: Figures show the time course of the host and phage populations for system (2.14)-(2.19) for different values of $\rho$ starting with more $P_J$. Initial conditions were $H_{JK} = 10000/11$, $H_O = H_J = H_K = 0$ and $P_J = 3000, P_K = 2$. (a) shows that initially all host populations reach a stable size, but after a significant increase in the total phage population hosts with more $J$ receptors decline whereas as hosts with more $K$ receptors show a sharp increase. This difference in behaviour results from the fact that $P_J$ increases first and targets hosts through $J$ receptors and $P_K$ increases later and targets the hosts through $K$ receptors as shown in (b).

Figure 2.8: Figures show the time course of the host and phage populations for system (2.14)-(2.19) with varying proportions of $J$ and $K$ receptors. Initial conditions were $H_{JK} = 10000/11$, $H_O = H_J = H_K = 0$ and $P_J = 3, P_K = 2000$, i.e. we start with more $P_K$. (a) shows that after the host populations become somewhat stable initially, there is a rise in the host cells with more $J$ receptors this time, and a decline in the numbers of hosts with more $K$ receptors. As can be seen from (b) this is because $P_K$ increase in number first in this case, followed by $P_J$. 
Figure 2.9: Figures show the time course of the different host and phage populations for proportions of $J$ and $K$ receptors between $\rho = 0.1$ to $\rho = 0.5$. Simulations are based on system (2.14)-(2.19). Initial conditions are $H_O = 3000$, $H_{JK} = 1$ (for each value of $\rho$), $P_J = 3000$ and $P_K = 0$. $\beta_K = 0.8 \times \beta_J$ to reduce the fitness of $P_K$ as compared to $P_J$ and $r = 0$. The graphs depict that the host populations will become concentrated at the lowest value of $\rho$ causing $P_J$ to go extinct. There is also a drift from $H_O$ to $H_{JK}$. The populations of $H_K$ are not shown as there are no hosts with just $K$ prophage. Without recombination the bacteria successfully evades the phage.
2.4. Discussion

Figure 2.10: Figures show the time course of the different host and phage populations for proportions of $J$ and $K$ receptors between $\rho = 0.1$ to $\rho = 0.5$. Simulations are based on system (2.14)-(2.19) and recombination is allowed to take place. Initial conditions are $H_O = 3000$, $H_{JK} = 1$ (for each value of $\rho$), $P_J = 3000$ and $P_K = 0$. $\beta_K = 0.8 \times \beta_J$ to reduce the fitness of $P_K$ as compared to $P_J$ and $r = 10^{-6}$. The graphs depict that the host populations drift from $H_O$ to $H_{JK}$, they do not become concentrated at the lower values of $\rho$. Instead they are spread over all different proportions of receptors with the largest numbers being found at $\rho = 0.5$. Both host and phage populations show oscillations. Due to recombination the phages manage to survive the bacteria’s evasion tactics.
Figure 2.11: Figures show the time course of the different host and phage populations for proportions of $J$ and $K$ receptors between $\rho = 0.1$ to $\rho = 0.9$. Simulations are based on system (2.14)-(2.19) and recombination is turned off. Initial conditions are $H_0 = 3000$, $H_{JK} = 1$ (for each value of $\rho$), $P_J = 3000$ and $P_K = 0$. $\beta_K = 0.8 \times \beta_J$ to reduce the fitness of $P_K$ as compared to $P_J$ and $r = 0$. The host populations drift to the lower values of $\rho$ to evade $P_J$. As a result $P_J$ goes extinct. In the absence of recombination the host cells successfully escape infection by expressing fewer $J$ receptors.
2.4. Discussion

Figure 2.12: Figures show the time course of the different host and phage populations for proportions of $J$ and $K$ receptors between $\rho = 0.1$ to $\rho = 0.9$. Simulations are based on system (2.14)-(2.19) and recombination is allowed to take place. Initial conditions are $H_O = 3000$, $H_{JK} = 1$ (for each value of $\rho$), $P_J = 3000$ and $P_K = 0$. $\beta_K = 0.8 \times \beta_J$ to reduce the fitness of $P_K$ as compared to $P_J$ and $r = 10^{-6}$. The host cells quickly acquire both types of prophage to become $H_{JK}$ and both the host and phage populations exhibit oscillations. Recombination keeps both $P_J$ and $P_K$ alive through host diversification.
Chapter 3

Models with point mutations and different infection patterns

3.1 Introduction

In this chapter we consider a model that is an extension of the models given in the previous chapters; it takes into account the ability of the hosts and phages to undergo evolutionary changes through point mutations and considers the different infection patterns that may be found in phage-bacteria infection networks. The model presented in this chapter may be generalized to include more than just 2 host/phage types (making it distinct from the last two models in this respect). The analysis of this model was focused on determining whether adding to the complexity of the model (by allowing the phage and bacteria to mutate and defining specific infection patterns) would result in different dynamics than previously seen.

The model predicts that prophage sequences may play a key role in maintaining the phage population in situations that would otherwise favour host cell resistance. In addition, prophage recombination facilitates the existence of multiple phage types, thus promoting diverse coexistence in the phage-host ecosystem. Finally, because the host carries an archive of previous phage strategies, prophage recombination can drive waves of innovation in the host cell popu-
3.2 Bacteria-Phage Model with Mutation, Recombination and Infection Patterns

Overall, our model follows a large population of host cells, many of which have prophage or prophage remnants in their genomes. As we are primarily interested in infection dynamics, we focus on genome sequences that confer adsorption ability to the phage. For example, our model would follow sequences encoding the tail fibres of the lambdoid phages. We do not assume that the entire prophage genome is inducible, thus prophage sequences that have undergone substantial mutational decay are still included in the model, as long as the sequence mediating adsorption remains functional. Likewise, we do not assume that the prophage confers immunity to other lambdoid phages. Although immunity may have important implications for host-phage coevolution, shifting the balance from parasite toward mutualist [39], our goal in this contribution is to isolate the effects of prophage recombination on the coevolutionary outcome. The model considers both host cell and phage population dynamics. Since phage target receptors on the surface of bacterial cells for attachment and entry, the model includes host cells with \( R \) different receptor types, and in its simplest form, each host cell expresses only one type of receptor. Corresponding to each host type are \( R \) phage types with different affinities for each bacterial receptor.

Each host cell also has the ability to carry either no prophage or any combination of prophage from the \( R \) phage types. For example when \( R = 5 \), the densities of host cells are denoted \( H_{x,jknm} \), where \( x \in J, K, L, M, N \) represents the receptor the host cell is expressing and \( j, k, l, m, n \) may be 0 or 1 and indicate which type(s) of prophage the host cell is carrying. Thus, for example, the density of hosts expressing receptor \( K \) and carrying prophage \( J \) and \( L \) would be denoted \( H_{K,10100} \). Similarly, phage densities are represented by \( P_x \), where \( x \in J, K, L, M, N \). We use either a nested or one-to-one infection pattern, described below, to determine which
hosts are susceptible to which phage types.

The phage adsorb to bacteria at a rate proportional to the densities of phage and bacteria. Following adsorption, the phage undertake one of two different modes of infection, lysogeny with probability $p$ and lysis with probability $(1 - p)$. Because we do not assume that the prophage remains inducible or confers immunity, when lysogeny occurs the host cell simply acquires the prophage of the infecting phage and remains in the susceptible class of bacteria. In contrast if lytic replication occurs, the infected host cell produces a burst of virions and is removed from the population. The type of virions produced by the host cell depends on the infecting phage type, and the prophage it carries. With probability $(1 - r)$, the phage released at lysis will be of the same type as the infecting phage. However, with probability $r$, recombination between the infecting phage and a pre-existing prophage will take place, and in this case the virions produced will be of the same type as the prophage sequence the infecting phage is recombining with.

Bacterial populations grow logistically with maximum growth rate $\lambda_x$ and carrying capacity $C$. Mutations are introduced with a constant probability per replication. A mutation for the host changes the dominant receptor it expresses. Forward mutations occur at rate $u_H$, with $H_J$ mutating to become $H_K$, then $H_L$, $H_M$ and $H_N$. With each forward mutation, the bacterial growth rate is reduced by a cost of resistance denoted by $s_i$, consistent with observations of reduced fitness in phage-resistant hosts (for example [19]). Thus $\lambda_K = \lambda_J(1 - s_i)$, while $\lambda_L = \lambda_J(1 - 2s_i)$ etc. Back mutation restores fitness but occurs at a reduced rate $v_H$. The adsorption rate is $\beta_x$ and the bacterial death rate is $\mu$. When lysis occurs, a burst of $f$ viral offspring are instantaneously produced; the lysis time is assumed negligible. Mutation can also occur in the phage population, altering or extending the host range of the mutant phages as described below. With each mutation, the infection rate of the phage can likewise be reduced by $s_\beta$, the cost of adaptation, yielding for example $\beta_K = \beta_J(1 - s_\beta)$. This is consistent with empirical evidence for reduced growth rate in phages with a wider host range [51]. Since recent evidence suggests that phage acquire an extended host range at rates on the order of
10^{-10} per replication (see Parameter Values in section 3.3), only forward mutation at rate \( u_p \) is considered, such that \( P_J \) mutates to produce \( P_K \), etc. The clearance rate of the phage is given by \( d \). Numerically, we have also investigated cases in which the costs of resistance and mutation are varied and/or back mutation for phage is allowed; qualitative outcomes are not sensitive to these choices. Since the prophage population includes both inducible and cryptic prophage remnants, we assume that the excision or loss rate for the prophage is negligible. This assumption could be relaxed in future work.

In the results to follow, we explore two common infection patterns describing the susceptibility of host types to phage types. In both models, phage \( P_J \) can only infect host cells with the \( J \) receptor, \( H_J \). Thus host cells can escape phage pressure through mutation to become \( H_K \), and the phage \( P_J \) require a mutational step, becoming \( P_K \), in order to infect the new host type. In the one-to-one or MA infection model [10], phage \( P_K \) can infect only \( H_K \), and have lost the ability to infect \( H_J \). In contrast, in the nested or GiG infection model [31], phage mutations augment the host range, and the ability to infect previous hosts is retained. Thus phage \( P_K \) can infect both \( H_K \) and \( H_J \) in the nested model. (For an overview of infection models including empirical support, see [32, 65], for recent work reconciling these models in a single framework, see [2, 58].)

We couple a set of ordinary differential equations describing the deterministic population dynamics to a stochastic simulation describing the influx of rare mutations. For the one-to-one infection pattern, the general form of the deterministic model equations is illustrated below. Here, we give several example equations; the full system that was used to generate the figures to follow is provided in Appendix B. For instance the equations of host cells carrying no
prophage, J prophage and J and K prophage can be written as

\[
\begin{align*}
\dot{H}_{J,00000} &= H_{J,00000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \beta_J P_J \right], \\
\dot{H}_{J,10000} &= H_{J,10000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \beta_J P_J \right] + p \beta_J H_{J,00000}, \\
\dot{H}_{J,11000} &= H_{J,11000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \beta_J P_J \right] + p \beta_J H_{J,01000}
\end{align*}
\]

(3.1)

The equation for \( P_J \) in a one-to-one system can be written as

\[
\dot{P}_J = (1 - p) f \beta_J (1 - r) P_J H_J - d P_J + (1 - p) f r \left[ \sum_x \beta_x P_x \frac{\sum_{klmn} H_{x,1klnn}}{1 + k + l + m + n} \right],
\]

(3.2)

where the sum over \( x \) is taken for \( x \in \{ J, K, L, M, N \} \).

We note that the denominator in the last term in the equation (3.2) for \( P_J \) corrects for the probability of recombination with any given prophage, when there are several prophage types in the host cell’s genome. This probability is given by \( \frac{1}{n} \), where \( n \) is the total number of prophage types carried.

For the nested infection pattern, the deterministic equations for the uninfected cells are as follows:

\[
\begin{align*}
\dot{H}_{J,00000} &= H_{J,00000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \sum_x \beta_x P_x \right], \\
\dot{H}_{J,10000} &= H_{J,10000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \sum_x \beta_x P_x \right] + p \beta_J P_J H_{J,00000}, \\
\dot{H}_{J,11000} &= H_{J,11000} \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - \sum_x \beta_x P_x \right] + p \beta_J P_J H_{J,01000} + p \beta_K P_K H_{J,01000}
\end{align*}
\]

(3.3)
Similarly the equations for the phage populations $P_J$ and $P_L$ can be written as:

\[
\dot{P}_J = (1 - p)f\beta_J(1 - r)P_JH_J - dP_J + (1 - p)f \left[ \sum_x \beta_x P_x \frac{\sum_{klmn \in 0,1} H_x,klmn}{1 + k + l + m + n} \right],
\]

\[
\dot{P}_L = (1 - p)f\beta_L(1 - r)P_L[H_J + H_K + H_L] - dP_L + (1 - p)f \left[ \sum_x \beta_x P_x \frac{\sum_{jkln \in 0,1} H_x,jkln}{1 + j + k + m + n} \right],
\]

where summations over $x$ are taken for $x \in \{J, K, L, M, N\}$. Again, the full set of equations used to generate the figures below is available in Appendix B.

To obtain numerical results, we integrate the full system of equations (ODE45 package in MATLAB), using parameter values as given in Table A.3 in Appendix A, and taking $N = 5$ distinct viral/prophage types. Finally, rare mutations are included using a semi-stochastic approach implemented through impulsive differential equations. The numerical integration of the systems above is halted every $\Delta t$ time units, and we compute the total number of replications that occurred, for either phage or host, in time interval $[t - \Delta t, t]$. Multiplying by the mutation rates per replication, $u_H$, $v_H$ and $u_P$, we compute the expected number of mutations in time interval $\Delta t$. If this expected value, $M$, is greater than five, we subtract this number from the source population and add it to the mutated population, for example, subtracting $M$ from $H_{J,10000}$ and adding to $H_{K,10000}$. If $M$ is less than five, we subtract/add mutant individuals in the same way, but the number of mutations is randomly generated from a Poisson distribution with mean $M$. Thus mutations are treated stochastically when rare, and deterministically when sufficiently frequent. In the results presented below, we take $\Delta t = 1$ day. Note that through this forward and backward mutation process, all host types are possible, that is, a host expressing receptor $L$ may have any combination of prophage, not just prophage $L$.

### 3.3 Parameter Values

Realistic parameter values taken from relevant literature are provided in Table A.3 in Appendix A. We take a minimum fission time for host cells of about 36 minutes ($(\ln 2)/27$ days) [56],
when the host cell population is far from carrying capacity. When at carrying capacity, cell division is reduced and host cells live for about 1 day (taking into account a cell death of 2.5% per generation for 40 generations per day) [61, 22]. The virus is cleared/denatured with an average lifetime of 1.7 hours (1/13.9 days) [26]; in nature this value varies widely for example with temperature. Infected cells produce a burst size of approximately 150 virions [26], and the infection rate is set such that in a population of completely susceptible cells, 20 new cells would be infected on average per infected cell ($R_0 = 20$). The mutation rates in our model reflect the emergence of phage resistance or expanded host ranges, not point mutations. We take the bacterial mutation rate to be between $10^{-7}$ and $10^{-8}$ per replication, based on the observed rate at which experimental “non-mutator” lines of *E. coli* developed resistance to phage T5 (see Figure 1c in [28]). Resistance entails a fitness cost of $s_\lambda = 0.02$; qualitative results are not sensitive to this parameter value. We assume back mutation restores fitness and is two orders of magnitude less likely. The phage mutation rate is calculated based on the observed rate at which phage *Lambda* developed the ability to attach to the *OmpF* receptor in replicate experiments; the overall rate at which these phage variants emerged was $1.25 \times 10^{-10}$ per replication (note that four base-pair mutations are involved in this transition) [46]. Phage mutants suffer a fitness loss of $s_\beta = 0.02$, however once again we note that qualitative results are not sensitive to this parameter. Prophage recombination probabilities have been reported in the range $10^{-6}$ to $10^{-8}$ [14, 40]; we take a value of $r = 10^{-6}$. In the results illustrated below, we consider a bacterial population with total carrying capacity of $10^8$ cells; this would correspond for example to $10$ ml at $5 \times 10^7$ cells/ml [43]. The prophage acquisition probability for phage *Lambda* on *E. coli* is expected to lie within the range of 0.1 to 0.6 [8, 9]; we use the more conservative value of $p = 0.2$ as a higher $p$ would only amplify the effects of prophage integration and help to make our case. To explore dynamics in the absence of recombination, the mutation rates and the recombination probability is set to zero in some of the results to follow.
3.4 Results

In this contribution, we are interested in the effects of the genetic reservoir of prophage and its influence on subsequent coevolutionary dynamics. Starting from a host cell population with a mix of prophage sequences, we illustrate in the Appendix B that through mutation and back mutation, host cells of each receptor type eventually accrue prophage of all types. Moreover in chapter 2 it was shown through the analysis of system (2.14)-(2.19) that in natural settings, host cells can be expected to accumulate different prophage types within a few generations through their interactions with different phages. We thus focus our attention on host cells carrying the full reservoir of prophage sequences, starting with initial conditions at which all host cells carry the full complement of prophage. This assumption simplifies the model considerably, reducing the number of host populations to five; we now use $H_x$ to refer to the population $H_{x,11111}$.

The results we present are clearly sensitive to the balance between mutation rates and the carrying capacity, reflecting the accessibility of the relevant mutations. Our goal is not to explore all possible behaviours in this model, but rather to illustrate the effects prophage recombination can induce in a reasonable regime. We therefore focus on finding a population size for which the outcome of coevolution matches long-standing theoretical and empirical predictions: a few brief coevolutionary steps followed by the emergence of a completely resistant bacterial strain [59]. Taking this as the default parameter set, we can address the effects of prophage recombination in isolation.

3.4.1 One-to-One Infection

In the one-to-one or matching alleles infection pattern, each host is susceptible to only one phage type, and each phage type can infect only one host. In the absence of mutation and recombination ($u_H = v_H = u_P = r = 0$), starting with host $H_J$ and phage $P_J$, both host and phage populations exhibit oscillations that quickly converge to a stable equilibrium (not shown). Keeping $r = 0$ but allowing both hosts and phage to evolve through mutation gives
more complex outcomes, as illustrated for example in Figure 3.1. Here the host diversifies within the first few days from $H_J \rightarrow H_K$ and eventually from $H_K \rightarrow H_L$ (left panel). Although in some simulation runs the phage successfully makes one or more mutational steps, eventually the host escapes the phage and drives the phage population to extinction, as seen in the right panel of the figure.

As mentioned above, this result is consistent with long-standing predictions of coevolutionary dynamics [42], in which the probability that the phage acquires the specific set of mutations required to target a new host receptor is low, relative to the ability of the host to down-regulate one surface receptor. For example, in the case of phage Lambda, 4 out of 5 specific mutations, all in the same gene, were required to infect $E. coli$ through a new receptor [46]. Other studies also support the view that bacterial resistance may develop more rapidly than phage coevolution ([49, 62], see [27]) for review), as suggested on theoretical grounds because of an inherent asymmetry in evolvability [41, 42].

This prediction changes, however, if the infecting phage has the additional ability to evolve through recombination with prophage in the host’s chromosome ($r > 0$), as illustrated in Figure 3.2. Here we see the classic Red Queen dynamics emerging, as successive host types arise via mutation, and the phage populations adapt in turn, avoiding extinction. In the one-to-one infection scenario, all host and phage types coexist in a diverse ecosystem by the end of the simulation, showing sustained and complex oscillations. These results clearly demonstrate the advantage of recombination to the phage population; in none of the parameter cases or simulation runs were the hosts able to “outrun” the phage if recombination with prophage was allowed. Thus lysogeny and recombination help to ensure the long-term survival of the phages, as expected, in situations that would otherwise favour the host. More interesting is the result that prophage indirectly support the diversity of the host population. We return to this idea in the Discussion (section 3.5).
3.4.2 Nested Infection

In a nested or gene-for-gene infection pattern, the phage develop the ability to infect bacteria through novel receptors, without losing the ability to infect ancestral hosts. In other words, we consider bacteria that become resistant to all previous phage populations through each successive mutation, and phages that can extend their host range to include one more host type with each successive evolutionary step.

In the absence of recombination \((r = 0)\), results for the nested and one-to-one infection patterns are indistinguishable in the parameter regime we explored (data not shown). When phage can evolve through both mutation and recombination \((r > 0)\), however, complex dynamics ensue, as shown in Figure 3.3. Here the host population \(H_J\) diversifies to produce \(H_K\), and the other host types emerge in turn. The phages also evolve rapidly and all phage populations emerge through recombination. In contrast with the one-to-one case, hosts \(H_J\) and \(H_K\) are driven to extinction when their successors emerge, presumably because each of these host types can be infected by multiple phage populations. At the end of the coevolutionary trajectory, all phage populations persist, being constantly regenerated through recombination with prophage sequences in the host.

Once again, in this scenario, prophage recombination substantially alters the evolutionary outcome. In a parameter regime that otherwise favours host cell evolution, prophage sequences help to ensure the survival of the phage populations. The effect of prophage on the evolution of the hosts is also striking: the host populations are pushed toward innovation. More innovative hosts survive in the long-run, while less adapted hosts are driven to extinction despite their fitness advantage.

In Figure 3.3, we illustrate the coevolutionary dynamics up until the point when all five possible host cell innovations have emerged. The dynamics after this point, when both hosts and phage have “run out” of adaptive steps, are consistent with the experimental results described by Hall et al. [36], in that cyclic frequency-dependent dynamics ensue (see Appendix B). The mix of host and phage types that persist depends on the details of the parameter regime, in
particular, on the balance of the costs of innovation and the force of infection.

3.5 Discussion

Bacteria carry an wide array of prophage sequences, many of which become “domesticated” over evolutionary timescales, conferring adaptive benefits to the host [61, 11]. Since infecting phage can access this archive of potentially useful genetic material through recombination, however, prophage sequences also provide a shortcut across the adaptive landscape, allowing phage to acquire new functions in a single step, rather than exploring multiple point mutations.

Previous work has suggested that bacterial hosts, with their greater evolutionary potential, may have the upper hand in the coevolutionary arm’s race [41, 18]. Our results demonstrate that in realistic parameter regimes that would otherwise favour host evolution, recombination with prophage can be a powerful asset to temperate phage populations, ensuring their survival. In this exceptional situation, when hosts carry an arsenal of genetic code for their predators, driving the predators to extinction becomes increasingly difficult.

While this result has not been previously elucidated, it is not surprising. Of further interest are the unanticipated effects of prophage sequences, and their accessibility by infecting phage, on host evolution.

With a one-to-one infection pattern, the evolutionary dynamics, including prophage, results in a diversification of hosts. Because prophage recombination ensures the continued existence of multiple phage types, a diverse coexistence of multiple phage and host populations is predicted. This degree of diversity is not predicted, in this parameter regime, in the absence of prophage recombination. While recombination between host cell genomes may also be facilitated by the presence of homologous prophage sequences [17], our results predict that prophage recombination may be a further factor that indirectly maintains host cell diversity.

In the case of a nested infection pattern (phage mutations that increase host range), prophage recombination leaves ancestral host populations at a relative disadvantage compared to newly
Figure 3.1: A one-to-one infection pattern in the absence of prophage recombination \((r = 0)\). From initial populations \(H_J = 500\) and \(P_J = 1000\), host type \(H_K\) emerges by mutation (left panel) and is resistant to \(P_J\). At realistic mutation rates, \(P_J\) is driven to extinction (right panel) before \(P_K\) emerges.

evolved strains. Thus prophage recombination coupled with nested infections forces the host to innovate, as hosts that are susceptible to fewer phage strains have greater chances of survival. Prophage recombination is critical to the "waves of innovation" predicted in our model (see Figure 3.3), because recombination ensures that previous phage types can always be regenerated. In essence, when the host carries an archive of previous phage strategies, the host is pushed ever forward in adaptation.

Microbes share complex, ever-changing relationships with each other, and in particular the exchange of genetic information occurs at many levels. Our goal in this chapter was to isolate one aspect of this complex picture, the effect of prophage recombination. The qualitative results described here – that recombination with prophage sequences can promote diversity and drive innovation in the host cell population – are generally robust in the models described in chapter 2 as well. However, steps toward further realism such as partial resistance, modular infection patterns or distinct prophage genes for adsorption, lysis and immunity have not yet been explored. All of these improvements will help to describe and understand the tremendous diversity observed in host-phage relationships in nature.
Figure 3.2: A one-to-one infection pattern with prophage recombination ($r = 10^{-6}$). Initial populations are $H_J = 500$ and $P_J = 1000$. On the left, new host types emerge by successive mutations from an initial population of $H_J$. New phage types also emerge through recombination (right panel), allowing the phages to survive diversification by the host. A diverse coexistence of phage and host types results.

Figure 3.3: A nested infection pattern with prophage recombination ($r = 10^{-6}$). Initial conditions are $H_J = 500$ and $P_J = 1000$. The host cell population shows “innovation waves” (left panel) as new host types successively emerge, driving previous host types to extinction due to their greater susceptibility to phage. All phage types emerge and survive due to recombination (right panel). Prophage recombination thus drives innovation in the hosts.
Chapter 4

Conclusion

In summary, this thesis sought to address the question of the importance of lysogeny and prophage sequences to both bacteria and phage in their interactions with each other. Our models and analysis focused on recombination between the infecting phage and the prophage in the host cell as a possible means of phage evolution; our results showed how the evolutionary benefit to the phage from recombination changed host-phage interactions. In this concluding chapter we will reiterate why our research question is interesting and important by giving the problem some context. We will also discuss the implications of some of our results and also suggest some direction for future research.

In rare cases, infecting phages can recombine with prophage incorporated in the host’s genome by their temperate predecessors and utilize the new information to overcome the host’s receptor-based defenses. Although we know of some benefits that prophage confer to their hosts such as super-infection immunity and increased virulence, we believe that there is much more left to be discovered about the effects that these extra bits of phage DNA can have on the interactions of bacteria and phage. As for the phages, despite natural selection virulent phage have failed to eliminate their slowly replicating temperate counterparts. Lysogeny may prove to be a successful alternative strategy to lysis in times of resource scarcity. However, the lack of research on the effects of phage temperance on phage-bacteria dynamics posits the possibility
that there is more to the picture. The topic of prophage-based recombination is an important one to address as recent biological evidence has shown that it may have a significant impact on the outcome of phage and bacteria coevolution.

In light of the research question, the results from system (2.4)-(2.6) show that recombination increases the chances of survival of the phages in conditions where it was likely that they would be eliminated by the hosts. In certain parameter regimes, diversification of the host population definitively drives the phages to extinction, but this parameter regime was found to be very limited. The more likely outcome of the scenario was that the phages would diversify to the same level as the hosts and the different phage types would maintain each other through recombination when there were insufficient hosts available for the infection to persist. Both the hosts and phage populations exhibited sustained oscillations following the initial diversification. Similar sustained oscillations were found in the results from system (2.14)-(2.19). Prophage-phage recombination was again seen as a successful survival technique for the phages in response to host diversification. In system (2.14)-(2.19), we removed the limitation that each host cell can have only one receptor. Yet the simulation results confirm the behaviours seen from system (2.4)-(2.6), showing that the aforementioned limitation does not adversely affect the qualitative nature of the results from the simpler model.

Our observations from the results of systems (2.4)-(2.6) and (2.14)-(2.19) indicate how prophage can be beneficial to the phages and detrimental to the host cells. The more sophisticated model in chapter 3, however, helped reveal how prophage integration helps the hosts as well. In microbial communities with one-to-one infection, coevolution accompanied by recombination leads to greater overall diversity in the host population. With the nested infection pattern, recombination keeps several phage types alive in the community, giving the hosts incentive to innovate to escape the phages. This, as we saw, leads to ‘waves of innovation’, in which the ancestral host type is successively replaced with a new one with greater immunity to phage infection. Another finding that is specific to this model is that sustained oscillations in the populations are a feature of one-to-one bacteria-phage interaction, whereas
a nested infection pattern results in a single dominant host and phage type ‘winning’ once the evolutionary cycles have come to an end. This result resonates with the observations of Hall et al. [36]. So recombination coupled with one-to-one infection promotes diversity for the hosts (and phages), whereas recombination and nested infection results in a greater degree of innovation than would be observed otherwise. Overall it seems that both hosts and phages stand to gain in different ways from lysogenic infection followed by recombination.

Although this study found some intriguing results, the complexity and dynamic nature of bacteria-phage communities warrants further investigation of the effects of temperance and recombination. One possible future direction for research could be to include factors such as super-infection immunity granted to the hosts due to prophage acquisition, prophage induction and mutational loss of prophage. It is likely that these factors will alter the current picture of dynamics to some extent. Another direction would be to combine some of the features of system (2.14)-(2.19) with the last model and have proportions of different receptor types and different infection patterns. This would help to better capture the complexity of real-world bacteria and phage communities.


## Appendix A

### Parameter Values for All Systems.

Table A.1: Parameters for System (2.4)-(2.6)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Values</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Birth rate of host cells</td>
<td>27.36 day$^{-1}$</td>
<td>[56]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate of host cells</td>
<td>1 day$^{-1}$</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Death rate of infected cells</td>
<td>27.27 day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$d$</td>
<td>Virus clearance rate</td>
<td>13.88 day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$p$</td>
<td>Probability with which virus becomes prophage</td>
<td>0.5</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>$r$</td>
<td>Probability of recombination with prophage</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>Infection rate</td>
<td>$1.5 \times 10^{-6}$ virus$^{-1}$ day$^{-1}$</td>
<td>[64]</td>
</tr>
<tr>
<td>$f$</td>
<td>Rate at which infected cells produce virus</td>
<td>144.6 virus cell$^{-1}$ day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$C$</td>
<td>Carrying capacity of host cells</td>
<td>$1 \times 10^7$ cell ml$^{-1}$</td>
<td></td>
</tr>
</tbody>
</table>
### Table A.2: Parameters for System (2.14)-(2.19)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Values</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Birth rate of host cells</td>
<td>27.36 day$^{-1}$</td>
<td>[56]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate of host cells</td>
<td>1 day$^{-1}$</td>
<td>[61]</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Death rate of infected cells</td>
<td>27.27 day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$d$</td>
<td>Virus clearance rate</td>
<td>13.88 day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$p$</td>
<td>Probability with which virus becomes prophage</td>
<td>0.2</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>$r$</td>
<td>Probability of recombinant prophage</td>
<td>$10^{-6}$</td>
<td>[14, 40]</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Infection rate</td>
<td>$2.4 \times 10^{-7}$ virus$^{-1}$ day$^{-1}$</td>
<td>see text</td>
</tr>
<tr>
<td>$f$</td>
<td>Rate at which infected cells produce virus</td>
<td>144.6 virus cell$^{-1}$ day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$C$</td>
<td>Carrying capacity of host cells</td>
<td>$1 \times 10^8$ cell ml$^{-1}$</td>
<td>[43]</td>
</tr>
</tbody>
</table>

### Table A.3: Parameters for One-to-one and Nested Infection Models

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
<th>Values</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda$</td>
<td>Birth rate of host cells</td>
<td>27.36 day$^{-1}$</td>
<td>[56]</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate of host cells</td>
<td>1 day$^{-1}$</td>
<td>[61]</td>
</tr>
<tr>
<td>$d$</td>
<td>Virus clearance rate</td>
<td>13.88 day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$p$</td>
<td>Probability with which virus becomes prophage</td>
<td>0.2</td>
<td>[8, 9]</td>
</tr>
<tr>
<td>$r$</td>
<td>Probability of recombinant prophage</td>
<td>$10^{-6}$</td>
<td>[14, 40]</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Infection rate</td>
<td>$2.4 \times 10^{-8}$ virus$^{-1}$ day$^{-1}$</td>
<td>see text</td>
</tr>
<tr>
<td>$f$</td>
<td>Rate at which infected cells produce virus</td>
<td>144.6 virus cell$^{-1}$ day$^{-1}$</td>
<td>[26]</td>
</tr>
<tr>
<td>$s_{\lambda}$</td>
<td>Cost of resistance</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>$s_{\beta}$</td>
<td>Cost of adaptation</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>$u_H$</td>
<td>Bacterial mutation rate</td>
<td>$5.4 \times 10^{-8}$ per replication</td>
<td>[57]</td>
</tr>
<tr>
<td>$v_H$</td>
<td>Bacterial back mutation rate</td>
<td>$5.4 \times 10^{-10}$ per replication</td>
<td></td>
</tr>
<tr>
<td>$u_P$</td>
<td>Phage mutation rate</td>
<td>$1.25 \times 10^{-10}$ per replication</td>
<td>[46]</td>
</tr>
<tr>
<td>$C$</td>
<td>Carrying capacity of host cells</td>
<td>$1 \times 10^8$ cell ml$^{-1}$</td>
<td>[43]</td>
</tr>
</tbody>
</table>
Appendix B

Supporting Material for Chapter 3

B.1 Complete Set of Equations for Systems (3.1)-(3.4)

Below are the complete sets of model equations in the simplified case in which all host cells
contain all prophage, where $H_X$ denotes $H_{X,11111}$.

B.1.1 One-to-One Infection

$$\dot{H}_J = H_J \left[ \lambda_J \left( 1 - \frac{H}{C} \right) - \mu - (1 - p)\beta_J P_J \right]$$

$$\dot{H}_K = H_K \left[ \lambda_K \left( 1 - \frac{H}{C} \right) - \mu - (1 - p)\beta_K P_K \right]$$

$$\dot{H}_L = H_L \left[ \lambda_L \left( 1 - \frac{H}{C} \right) - \mu - (1 - p)\beta_L P_L \right]$$

$$\dot{H}_M = H_M \left[ \lambda_L \left( 1 - \frac{H}{C} \right) - \mu - (1 - p)\beta_N P_N \right]$$

$$\dot{H}_N = H_N \left[ \lambda_L \left( 1 - \frac{H}{C} \right) - \mu - (1 - p)\beta_M P_M \right]$$

(B.1)
\( \dot{P}_J = (1-p)f\beta_J(1-r)P_JH_J - dP_J + \frac{1}{3}(1-p)f r R \)

\( \dot{P}_K = (1-p)f\beta_K(1-r)P_KH_K - dP_K + \frac{1}{3}(1-p)f r R \)

\( \dot{P}_L = (1-p)f\beta_L(1-r)P_LH_L - dP_L + \frac{1}{3}(1-p)f r R \)  \hspace{1cm} \text{(B.2)}

\( \dot{P}_M = (1-p)f\beta_M(1-r)P_MH_M - dP_M + \frac{1}{3}(1-p)f r R \)

\( \dot{P}_N = (1-p)f\beta_N(1-r)P_NH_N - dP_N + \frac{1}{3}(1-p)f r R \)

where \( R = \beta_JP_JH_J + \beta_KP_KH_K + \beta_LP_LH_L + \beta_MP_MH_M + \beta_NP_NH_N \).

**B.1.2 Nested Infection**

\[ \dot{H}_J = H_J \left[ \lambda_J \left( 1 - \frac{r}{C} \right) - \mu - (1-p) \sum_{x=j}^n \beta_x P_x \right] \]

\[ \dot{H}_K = H_K \left[ \lambda_K \left( 1 - \frac{r}{C} \right) - \mu - (1-p) \sum_{x=k}^n \beta_x P_x \right] \]

\[ \dot{H}_L = H_L \left[ \lambda_L \left( 1 - \frac{r}{C} \right) - \mu - (1-p) \sum_{x=l}^n \beta_x P_x \right] \]  \hspace{1cm} \text{(B.3)}

\[ \dot{H}_M = H_M \left[ \lambda_M \left( 1 - \frac{r}{C} \right) - \mu - (1-p) \sum_{x=m}^n \beta_x P_x \right] \]

\[ \dot{H}_N = H_N \left[ \lambda_N \left( 1 - \frac{r}{C} \right) - \mu - (1-p) \sum_{x=n}^n \beta_x P_x \right] \]

\( \dot{P}_J = (1-p)f\beta_J(1-r)P_JH_J - dP_J + \frac{1}{3}(1-p)f r S \)

\( \dot{P}_K = (1-p)f\beta_K(1-r)P_KH_K - dP_K + \frac{1}{3}(1-p)f r S \)

\( \dot{P}_L = (1-p)f\beta_L(1-r)P_LH_L - dP_L + \frac{1}{3}(1-p)f r S \)  \hspace{1cm} \text{(B.4)}

\( \dot{P}_M = (1-p)f\beta_M(1-r)P_MH_M - dP_M + \frac{1}{3}(1-p)f r S \)

\( \dot{P}_N = (1-p)f\beta_N(1-r)P_NH_N - dP_N + \frac{1}{3}(1-p)f r S \)

where

\[ S = \beta_JP_J \sum_{x=j}^J H_x + \beta_KP_K \sum_{x=j}^K H_x + \beta_LP_L \sum_{x=j}^L H_x + \beta_MP_M \sum_{x=j}^M H_x + \beta_NP_N \sum_{x=j}^N H_x \]  \hspace{1cm} \text{(B.5)}
B.2 Prophage acquisition by host cell populations

To test the assumption that host cells interacting with different phage types eventually pick up all types of prophage sequences, we calculated the average number of prophage sequences per host cell in both one-to-one and nested infection networks. Starting with $H_J$ cells carrying different combinations of prophage sequences, we observed the acquisition of prophage sequences by the hosts over a time period of 50 days. The results are plotted in Figure B.1. With the one-to-one infection pattern we see in Figure B.1a that initially $H_J$ cells carry on average 1 prophage sequence per cell, a value that increases almost immediately to 2. As new host types emerge, the average number of prophage sequences increases monotonically. By day 50, populations $H_M$ and $H_N$ each carry 5 prophage sequences on average. Due to the stochastic nature of the simulation, there were instances in which only $H_N$ accrued all 5 prophage types by day 50. However, over evolutionary times it is expected that the total number of prophage sequences carried by all types of host cells would increase to 5. This is because acquiring viral prophage represents a neutral genetic change that is continually reintroduced and never lost, thus it will eventually fix in the host population. Further mutational steps (backwards or forwards) can only increase the number of prophage sequences per host cell.

With the nested infection network, initially all host cells are of type $H_J$ with one prophage sequence on average. As seen in Figure B.1b, all five host cell types very quickly accrue all five prophage sequences in this case. Once again the total number of acquired prophage may be less than 5 in some cases due to the probabilistic nature of the model, but the most common outcome of the simulation was that all 5 prophage types were acquired by the hosts over a period less than 50 days. These results motivated our study of a reduced model, in which all hosts carry all types of prophage, as an approximation to the coevolutionary dynamics at longer timescales.
B.3. Results from relaxed initial conditions for the one-to-one model

Figure B.1: Changes in the average number of prophage sequences for different host types in one-to-one and nested infection networks. Panel B.1a shows that starting with $H_J$ cells which have an average of 1 prophage sequence per host, the number of prophage sequences increases by 1 with each mutational step. $H_M$ and $H_N$ have accrued 5 prophage sequences each on average by day 50. Panel B.1b shows that initially $H_J$ has about 1 prophage sequence on average. Each subsequent host population starts with 1 more prophage sequence compared to its ancestor. Eventually only $H_M$ and $H_N$ survive, each with 5 prophage sequences on average in their genome.

B.3 Results from relaxed initial conditions for the one-to-one model

The initial conditions given in the main text assume that all host types begin with all 5 prophage sequences, however depending on the back mutation rate, this initial condition may only be achieved at long evolutionary times for the one-to-one infection model. In this section we demonstrate that qualitatively, the results from the one-to-one infection model are not sensitive to this condition. Figure B.2a shows the results in the case in which initially the $H_J$ population has all 5 types of prophage, but only one prophage sequence per cell. In other words, the initial population of $H_J$ is made up of $H_{J,00001}, H_{J,00010}, H_{J,00100}, H_{J,01000}$ and $H_{J,10000}$ cells in different proportions. Despite the difference in initial conditions, the population dynamics are almost identical to Figure 3.3 in the main text.
B.4 Long-term dynamics resulting from nested infection pattern

Figure B.3 shows the results of a long-term simulation, continued after the coevolutionary cycles have ended because no more innovations are possible in the model. In the left-panel we see that only $H_M$ and $H_N$ survive the ‘waves of innovation’, as they are resistant to the largest range of phages. $H_M$ is dominating the host cell population as it has a higher fitness than $H_N$. The phage population is made up of $P_L$, $P_M$ and $P_N$ (right-panel). $P_N$, having the largest host range, becomes the dominant phage type followed by $P_M$ and $P_L$ respectively. It seems that in the long-run, nested infection dynamics result in certain host and phage types surviving, depending on the collective force of phage infection and the cost of adaptation of the hosts and phages. Within the populations that survive, certain hosts and phages tend to do better than others; this is again a consequence of the different host ranges (for phages) and adaptation costs (for the hosts). These dynamics reflect the artificial feature of the model that both hosts and phage have a limited number of possible adaptive steps, and compensatory mutations are
B.4. Long-term dynamics resulting from nested infection pattern

Figure B.3: Time course of the system (B.3)-(B.4) in the long-run. Initial conditions are $H_J = 500$ and $P_J = 1000$. In this case only $H_M$ and $H_N$ survive and $H_M$ becomes the dominant host type as seen in Figure B.3a. As for the phages, Figure B.3b shows that $P_L$, $P_M$ and $P_N$ survive with $P_N$ being the dominant phage type.

not possible.
Curriculum Vitae

Name: Alina Nadeem

Post-Secondary Education and Degrees:
- Bachelor of Science in Mathematics, Lahore University of Management Sciences, 2010-2014.

Related Work Experience:
- Teaching Assistant, The University of Western Ontario, 2014 - 2016.

Publications:

Nadeem, A. and Wahl, L.M.
Prophage as a genetic reservoir: promoting diversity and driving innovation in the host community. *Under Preparation for Evolution.*

Yu, P., Nadeem, A. and Wahl, L.M.
The impact of prophage on the equilibria and stability of phage and host. *Under review for publication by the Journal of Nonlinear Science*