Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 19(6): 5482–5508.
DOI: 10.3934/mbe.2022257
Received: 08 January 2022
Revised: 14 February 2022
Accepted: 07 March 2022
Published: 28 March 2022

Research article

The gossip paradox: Why do bacteria share genes?

Alastair D. Jamieson-Lane[1,2] **and Bernd Blasius**[2,*]

[1] Department of Mathematics, University of Auckland, Auckland, 1010, New Zealand

[2] Institute for Chemistry and Biology of the Marine Environment, Carl von Ossietzky, Universität Oldenburg, Oldenburg, 26129, Germany.
Helmholtz Institute for Functional Marine Biodiversity, Carl von Ossietzky, Universität Oldenburg, Oldenburg, 26129, Germany.

* **Correspondence:** Email: blasius@icbm.de.

Abstract: Bacteria, in contrast to eukaryotic cells, contain two types of genes: chromosomal genes that are fixed to the cell, and plasmids, smaller loops of DNA capable of being passed from one cell to another. The sharing of plasmid genes between individual bacteria and between bacterial lineages has contributed vastly to bacterial evolution, allowing specialized traits to 'jump ship' between one lineage or species and the next. The benefits of this generosity from the point of view of both recipient cell and plasmid are generally understood: plasmids receive new hosts and ride out selective sweeps across the population, recipient cells gain new traits (such as antibiotic resistance). Explaining this behavior from the point of view of donor cells is substantially more difficult. Donor cells pay a fitness cost in order to share plasmids, and run the risk of sharing advantageous genes with their competition and rendering their own lineage redundant, while seemingly receiving no benefit in return. Using both compartment based models and agent based simulations we demonstrate that 'secretive' genes which restrict horizontal gene transfer are favored over a wide range of models and parameter values, even when sharing carries no direct cost. 'Generous' chromosomal genes which are more permissive of plasmid transfer are found to have neutral fitness at best, and are generally disfavored by selection. Our findings lead to a peculiar paradox: given the obvious benefits of keeping secrets, why do bacteria share information so freely?

Keywords: plasmid; horizontal gene transfer; evolutionary game theory; public good; agent based model; evolutionary dynamics; Conjugation

1. Introduction

Evolution, that blind process of inheritance and selection, is both a process of constant refinement and optimization, and also a process of adaption; while in stable environments the species best able to make efficient use of available resources may drive others to extinction, in more variable environments it is those species most able to change and adapt which will flourish.

This tension between efficiency and adaptability can be seen most clearly in bacterial genomes, where 'core' genes are stored on bacterial chromosomes, while 'accessory' genes are commonly found on mobile gene elements, such as plasmids: small stable loops of DNA, independent from the host chromosome.

Core genes code for critical metabolic pathways, cell division and motility, while accessory genes code for more context specific capabilities – for example resistance to heavy metals [1], uncommon metabolic pathways [2], virulence factors [3], or antibiotic resistance [4]. Genes stored on mobile gene elements can be lost, rearranged, and most importantly for our purposes, passed from one bacteria to another via the process of Horizontal Gene Transfer (HGT).

Research into HGT over the past decades has fundamentally altered our understanding of bacterial evolution. Rather than a gradual accumulation of mutations within a single continuous clonal lineage, HGT allows entire gene modules to be transferred from one genome to another, fulfilling much the same role as sexual recombination. Bacterial genomes are less a tapestry which must be altered one thread at a time, but instead a patchwork quilt, with some estimates [5] suggesting that upwards of 80% of bacterial genes have undergone HGT at some point in their history. While the historical impact of HGT is enough to earn great scientific and philosophical interest, the role of HGT in the spread of antibiotic resistance and bacterial virulence factors [6–8] makes understanding of HGT an urgent practical concern from the point of view of modern medicine.

HGT can take one of several pathways (see Figure 1). It can take place via **transformation**: one bacterial cell taking up genetic material from its environment after the death of some other cell [9]. The process of **transduction** takes place via an invading viruses. By chance, a virus may accidentally package host DNA inside a viral capsid before lysing a cell, eventually leading to DNA from one host being inserted into the next. Most strikingly of all, accessory genes can be shared via bacterial **conjugation**, in which one cell bridges the gap to another, duplicates its plasmids, and funnels these copies across (see Eberhard's classical review [10], or more recently Thomas and Nielson [11]).

Conjugation is extensively regulated, both from the point of view of the donor and recipient. On the side of the donor cell, plasmid bound quorum sensing genes determine the appropriate density of donor cells [12] and intricate restriction-modification mechanisms (what might be thought of as the 'bacterial immune system') act to suppress foreign gene elements [13], including conjugation apparatus. On the side of the recipient cell, individual plasmids exclude incoming plasmids from the same 'incompatibility group', as these are likely to disrupt plasmid replication and regulatory processes [14]. Conjugation rate depends both on the plasmid in question, but also on the bacterial background these plasmids inhabit [15].

Bacterial conjugation is, for many reasons, somewhat odd. In particular, it involves one bacteria supplying another bacteria with potentially beneficial genetic material, often at reasonable expense to itself [16]. From the point of view of the plasmid genes being shared, this expense makes sense- the time and resources involved create more copies of the plasmid and allow it to ride over selective

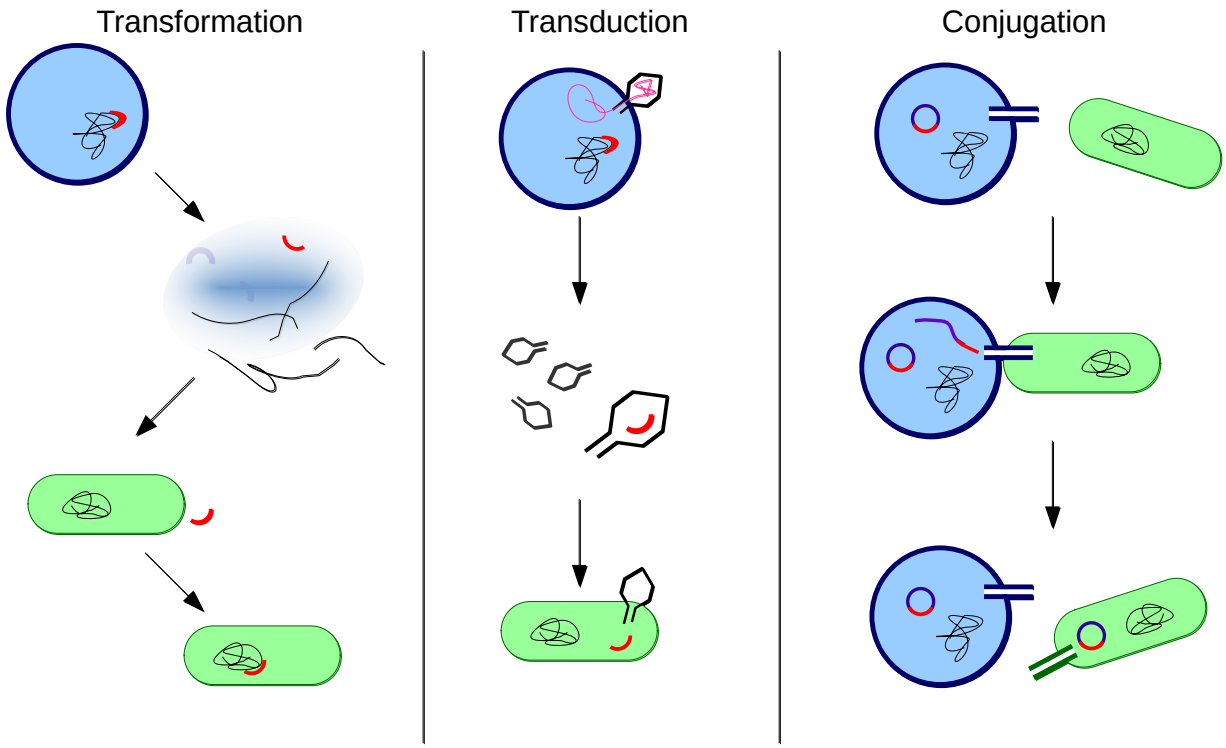

Figure 1. Modes of Horizontal Gene Transfer. Left: HGT via transformation. After a cell dies, the lysing cell spills fragments of genes into the environment. Later, some other cell collects one of these fragments, incorporating the gene fragment into its genome. Middle: HGT via transduction. A phage infects a cell, hijacking the cellular machinery in order to produce further phages. In the process, one of the new phages is accidentally constructed containing a piece of the original cell's DNA, rather than the phages. Later, this defective phage inserts the DNA fragment into a new cell. Right: HGT via conjugation. A cell contains DNA on a plasmid. This plasmid codes for bacterial conjugation, leading to the construction of a 'pilus'. The donor cell binds to another bacteria, duplicates the plasmid DNA, and sends it through the pilus to its new host. Afterwards, the two cells go their separate ways.

sweeps [17] or spread to new species or environments [18]. From the point of view of chromosomal genes in the recipient cell, the benefits are also clear; if the plasmid is beneficial the receiving bacteria and its descendants gain in fitness for many generations to come. Given the sometimes heavy burden imposed by newly arrived plasmid genes [19–21], the cost of accepting new plasmids can not be assumed to be trivial. Acceptance of an arriving plasmid can be viewed as an act of 'trust', but not one of 'generosity' (at least, to the extent that such intentionalistic language can be applied to bacterial behavior). This question of 'how plasmid hosts resolve conflicts with incoming gene elements' has been studied by McGinty et al. [22], and will not be our focus here.

While the benefits of HGT to the plasmid are clear, and the benefit to the receiving cell is at least somewhat understandable, from the perspective of the chromosomal genes in the donor cell, conjugation would appear to be strictly harmful. Not only does conjugation take significant time and risk infection by certain types of phages [23], in sharing advantageous DNA, a bacteria runs the risk of not only helping a rival in the present, but of generating a superior lineage and rendering its own descendants obsolete.

While previous authors have studied the peculiar evolutionary forces resulting from HGT from the point of view of bacterial plasmids, asking such questions as 'how are deleterous genes not lost?' and 'why are beneficial genes not incorporated into chromosomal genomes more directly?' [24], very few works consider HGT from the point of view of those chromosomal genes left behind. Recent work by Dimitriu et al. [25, 26] explores these issues both experimentally and analytically. While their experimental results are excellent, a recent correction to their mathematical model [27] indicates that selection for conjugation may be far more fragile than previously assumed.

Our goal in what follows is to investigate and demonstrate the tension between the interests of chromosomal and plasmid bound genes with respect to 'preferred' conjugation rate. We use both compartment based and agent based models to investigate the long term evolutionary outcome for chromosomal gene capable of either increasing or decreasing a cells conjugation rate. Such 'modifier genes' have been studied previously in the context of mutation, migration, and diploid recombination [28–30], where it can be generally shown that, at equilibrium, evolution selects for modifier genes that minimize change (often refered to as 'the reduction principle') [31, 32]. Unfortunately, these past results are not directly applicable to the current context, as here we study a modifier gene that modulates change in *neighboring* organisms, rather than in the organism bearing the gene in question.

Our work here is in some sense a close mirror to Koraimann et al.'s 2014 paper [12]. Koraimann et al. assume that information sharing is beneficial and discuss the metabolic costs associated with it, and the many regulatory mechanisms employed to minimize these costs. Here we will instead assume negligible metabolic cost, and instead ask the question: when is information sharing beneficial? What conjugation rate might we expect cells to tend towards when completely uninhibited by other complicating factors?

The structure of the paper is as follows: in section 2.1 we make use of a deterministic model to study the evolutionary dynamics in the simple case of one plasmid and two possible conjugation rates; we find that bacterial strains which suppress plasmid conjugation out compete their more generous competitors. We extend this model in section 2.2 in order to consider longer evolutionary time spans. In section 2.3 we construct a more detailed agent based model, and through simulation demonstrate how genes that limit plasmid conjugation gain a long term evolutionary advantage over a wide parameter range, even when accounting for spatial structure. In section 3 we build further upon this model, and

discuss the similarities and differences between HGT and previous evolutionary 'paradoxes'; namely the emergence of altruism and the evolution of sex. In both cases we find that classical resolutions to these past paradoxes fail to stabilize plasmid sharing.

2. Materials and methods

2.1. A single selective sweep

Our goal throughout this article will be to explore the evolutionary pressures experienced by chromosomal genes and how they interact with mobile plasmid genes. While the course of evolution is governed by many factors, in this section our goal is to construct the simplest possible deterministic model in order to study the interaction between mobile and 'static' genes, and build our intuition, before moving on to more complex modelling approaches.

Consider a population of bacteria, each of which is either positive ($\oplus$) or negative ($\ominus$) for some beneficial plasmid. In addition, each bacterium possess either a generous (G) or secretive (S) gene on their chromosome, which share plasmids via conjugation at rate c, or at some drastically reduced rate ϵc. Generous cells are not assumed to be transfer capable at all time, instead the constant conjugation rate c can be thought of as the effective conjugation rate after averaging over time: generous cells are those which share plasmids when circumstance (such as cell density) allow [12]. Secretive cells do not conjugate, even in circumstances where conjugation is convenient and easy. The small parameter ϵc accounts for both imperfect conjugation suppression on behalf of our chromosome, and also 'plasmid leakage' via the unregulated HGT processes of transduction or transformation [*]. Plasmid positive bacteria have some fitness $f_\oplus$, while plasmid negative bacteria have fitness $f_\ominus < f_\oplus$. See figure 2 for a schematic of this set up.

Written as chemical equations we have:

$$G_\pm \underset{f_\pm}{\longrightarrow} 2G_\pm, \qquad S_\pm \underset{f_\pm}{\longrightarrow} 2S_\pm, \tag{2.1}$$

$$G_\oplus + S_\ominus \underset{c}{\rightarrow} G_\oplus + S_\oplus, \qquad G_\oplus + G_\ominus \underset{c}{\rightarrow} 2G_\oplus \tag{2.2}$$

$$S_\oplus + G_\ominus \underset{\epsilon c}{\longrightarrow} S_\oplus + G_\oplus, \qquad S_\oplus + S_\ominus \underset{\epsilon c}{\longrightarrow} 2S_\oplus \tag{2.3}$$

We assume a 'death rate' such that total population is held constant; $G_\oplus + G_\ominus + S_\oplus + S_\ominus = 1$ at all times.

Written as a series of differential equations we have:

$$\begin{aligned} \dot{G}_\oplus &= (f_\oplus - \bar{f})G_\oplus + (cG_\oplus + \epsilon cS_\oplus)G_\ominus, \\ \dot{S}_\oplus &= (f_\oplus - \bar{f})S_\oplus + (cG_\oplus + \epsilon cS_\oplus)S_\ominus, \\ \dot{G}_\ominus &= (f_\ominus - \bar{f})G_\ominus - (cG_\oplus + \epsilon cS_\oplus)G_\ominus, \\ \dot{S}_\ominus &= (f_\ominus - \bar{f})S_\ominus - (cG_\oplus + \epsilon cS_\oplus)S_\ominus. \end{aligned} \tag{2.4}$$

Here $\bar{f} = f_\oplus(G_\oplus + S_\oplus) + f_\ominus(G_\ominus + S_\ominus)$ refers to the average fitness of the current population. The first term in each derivative refers to changes in population due to population growth and

[*] G and S can be thought of as a modifier genotype, controlling a bacterium's propensity to *donate* plasmids to others; all bacteria are equally capable of receiving plasmids. While a suitable change of variables could be used to place these equations more explicitly in the classical 'modifier gene' framework [28, 31], doing so would require entangling reproduction and recombination, resulting in non-neutral modifier genes and taking us outside the scope of previous analysis. We present a more algebraically convenient change of variables later in the article.

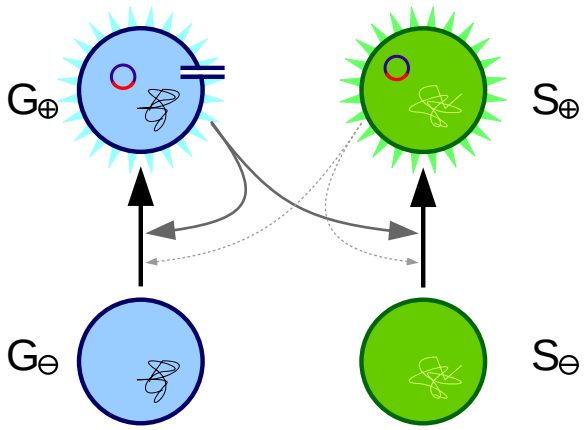

Figure 2. Our four 'variants', $G_\oplus, G_\ominus, S_\oplus$ and $S_\ominus$. 'Generous' variants are depicted in blue (left), 'secretive' in green (right). Plasmids are actively spread by $G_\oplus$ type bacteria, who possess both the plasmid and the chromosomal gene inclined to spread it. $S_\oplus$ type bacteria share the plasmid at a much lower rate, potentially close to zero. All variants replicate according to their fitness, with plasmid carrying cells assumed to have $f_\oplus > f_\ominus$. The case of burdensome plasmids, or plasmids with varying fitness, we ignore for the time being.

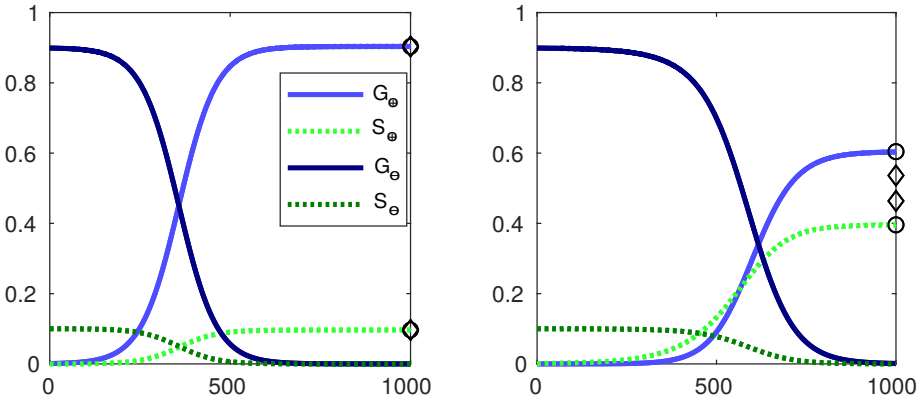

Figure 3. Characteristic anatomy of a selective sweep. Figure shows typical numerical solutions of equations 2.4 for the situation of an incoming plasmid arriving and become ubiquitous in the population. As it does so, the balance between generous and secretive genes shifts. (Left) Initial plasmid-positive population is $G_\oplus = 10^{-3}, S_\oplus = 0, G_\ominus \approx 0.9, S_\ominus = 0.1$. (Right) Initial plasmid positive population is $G_\oplus = 0, S_\oplus = 10^{-3}$. In the right hand panel, black diamonds represent the predicted final values using the approximation eq. 2.19, while circles represent predictions made by solving eq. 2.16 using Newton's method. In the left hand panel, these values are indistinguishable and the symbols are placed one on top of the other. Overall, when starting with a generous mutant (left), the levels of generous individuals increases from 0.899 to 0.9035, a 0.5% increase. When starting with a secretive mutant (right), the secretive population increases from 0.100 to 0.3955, a four-fold increase. In both cases, we assume parameter values $c = 0.01$, $\epsilon = 0.01$, $f_\oplus = 1.01$, $f_\ominus = 1$,with initial values $G_\ominus \approx 0.9, S_\ominus = 0.1$.

death/displacement. The second term in each derivative gives the demographic change caused by plasmid conjugation.

Suppose, we add a single beneficial plasmid to an otherwise plasmid negative population. This plasmid may find itself in either a generous or secretive bacteria. Due to its higher fitness, the plasmid is going to 'sweep' through, becoming ubiquitous in the population (a so called "selective sweep"). An illustration of two such selective sweeps is given in figure 3. Given the interaction between plasmids and chromosomes, we wish to determine how the balance between generous and secretive chromosomes changes as the new plasmid reaches saturation in the population. Stated mathematically, we wish to determine $G_\oplus(\infty)$ and $S_\oplus(\infty)$ for a given set of parameters and initial conditions.

While such systems are often described in terms of the abundance of individual genes, and the linkage disequilibrium between particular genes [33], in this case a somewhat different change of variables proves to be more helpful. Consider the population of mutants, M, which are always plasmid positive, and either all generous or all secretive; M represents the original mutant bacteria and all its direct descendants. We also consider a "Wild type" population W, which can be plasmid positive $W_\oplus$ or plasmid negative $W_\ominus$. The wild type population consists of some mixture of generous and secretive chromosomal genes. Because $\frac{d}{dt}(G_\ominus/S_\ominus) = 0$ we know that the relative proportions of generous and secretive genes remains fixed for $W_\ominus$. Because $\dot{W}_\oplus$ grows proportional to $W_\oplus$ in its first term, and $W_\ominus$ in its second, and because $W_\oplus$ initially has the same G/S ratio as $W_\ominus$, it must inevitably continue with this same ratio throughout the entire course of the ODE system. Thus, the relative ratio of generous and secretive chromosomal genes are fixed in the wild type population. See figure 4 for an illustration of this change of variables.

Stated algebraically, we have:

$$\begin{aligned} G_\oplus &= \rho W_\oplus + \chi M, \\ S_\oplus &= (1-\rho)W_\oplus + (1-\chi)M, \\ G_\ominus &= \rho W_\ominus, \\ S_\ominus &= (1-\rho)W_\ominus. \end{aligned} \tag{2.5}$$

Where $\rho = \frac{G_\ominus(0)}{G_\ominus(0)+S_\ominus(0)}$, and $\chi = \frac{G_\oplus(0)}{G_\oplus(0)+S_\oplus(0)}$. ρ takes values anywhere in the range $[0,1]$, while χ will always be equal to either 0 or 1. This partitioning of the population is algebraic rather than physical, but allows equation 2.4 to be simplified to:

$$\begin{aligned} \dot{M} &= (f_\oplus - \bar{f})M, \\ \dot{W}_\ominus &= (f_\ominus - \bar{f})W_\ominus - (c_M M + c_W W_\oplus)W_\ominus, \\ \dot{W}_\oplus &= (f_\oplus - \bar{f})W_\oplus + (c_M M + c_W W_\oplus)W_\ominus. \end{aligned} \tag{2.6}$$

Here c_M and c_W represent the conjugation rates of the mutant and wild type populations, and are given by $c_M = \chi c + (1-\chi)\epsilon c$ and $c_W = \rho c + (1-\rho)\epsilon c$. As before, the total population is held constant, $M + W_\ominus + W_\oplus = 1$, and $\bar{f}$ is the mean fitness of the current population, $\bar{f} = f_\oplus M + f_\oplus W_\oplus + f_\ominus W_\ominus$.

Combining the definition of $\bar{f}$ with population conservation, we find that $(f_\oplus - \bar{f}) = (f_\oplus - f_\ominus)W_\ominus = \Delta f W_\ominus$ and $(f_\ominus - \bar{f}) = (f_\ominus - f_\oplus)(1 - W_\ominus) = -\Delta f(1 - W_\ominus)$. Similar use of the total population equation gives

$$\dot{M} = \Delta f W_\ominus M, \tag{2.7}$$

$$\dot{W}_\ominus = -\Delta f(1 - W_\ominus)W_\ominus - (c_M M + c_W(1 - M - W_\ominus))\, W_\ominus, \tag{2.8}$$

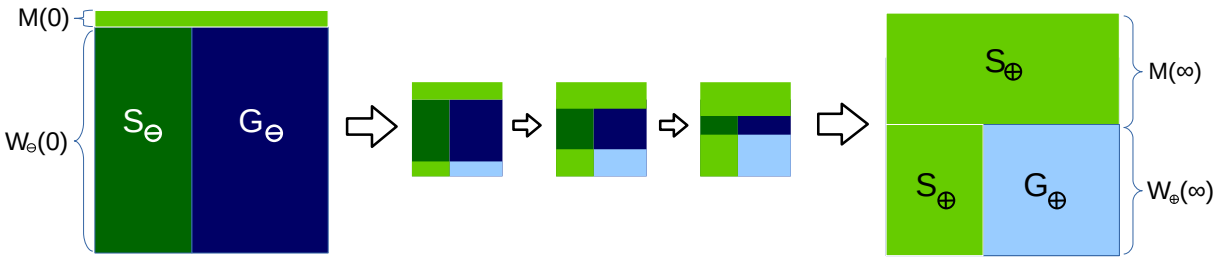

Figure 4. A graphical illustration of the change of variables. Presence/absence of beneficial plasmid is denoted by bright/dark colours. Secretive/generous chromosomes are denoted by green and blue (respectively). Over time descendants of the initial mutant displace the wild type population while simultaneously HGT supplies the wild type population with the beneficial plasmid, leaving the relative abundance of the chromosomal genes in the wild type population unchanged.

$$= -(\Delta f + c_W)(1 - W_\ominus)W_\ominus - (c_M - c_W)MW_\ominus. \tag{2.9}$$

Dividing the equation 2.9 by 2.7 eliminates time dependence:

$$\frac{\dot{W}_\ominus}{\dot{M}} = -\frac{(\Delta f + c_W)(1 - W_\ominus)}{\Delta f M} - \frac{c_M - c_W}{\Delta f}. \tag{2.10}$$

The above can be rearranged to give a D'Alembert equation [34] of the form :

$$W_\ominus = \frac{(W'_\ominus(M) + \beta)M}{\alpha} + 1, \tag{2.11}$$

where $\alpha = (\Delta f + c_W)/\Delta f$, $\beta = (c_M - c_W)/(\Delta f)$ and $W'_\ominus$ is the derivative of $W_\ominus$ with respect to M (as opposed to t). This equation permits solutions of the form:

$$W_\ominus(M) = 1 + kM^\alpha + \frac{\beta M}{\alpha - 1}, \tag{2.12}$$

where k is a constant of integration. Given the initial conditions are $M(0) = M_0 \ll 1$, $W_\oplus = 0$, $W_\ominus = 1 - M_0$. This leads to:

$$W_\ominus(M_0) = 1 - M_0 = 1 + kM_0^\alpha + \frac{\beta M_0}{\alpha - 1}, \tag{2.13}$$

$$\left(-1 - \frac{\beta}{\alpha - 1}\right)M_0 = kM_0^\alpha, \tag{2.14}$$

$$k = -\left(1 + \frac{\beta}{\alpha - 1}\right)M_0^{1-\alpha}. \tag{2.15}$$

The final mutant population is found by noting that $W_\ominus \to 0$ as time proceeds, and hence

$$W_\ominus(M_\infty) = 0 = 1 + kM_\infty^\alpha + \frac{\beta M_\infty}{\alpha - 1}. \tag{2.16}$$

If we assume that M_0 and M_∞ are of the same order of magnitude, then we see that for $M_0 \ll 1$, the term $\frac{\beta M_\infty}{\alpha-1}$ is negligible compared to the +1 term unless $\alpha \approx 1$. Hence

$$M_\infty \approx (-1/k)^{1/\alpha}, \tag{2.17}$$

$$\approx \left(1 + \frac{\beta}{\alpha - 1}\right)^{-1/\alpha} M_0^{1-1/\alpha}, \tag{2.18}$$

$$\approx \left(\frac{c_M}{c_W}\right)^{\frac{-\Delta f}{\Delta f + c_W}} M_0^{\frac{c_W}{\Delta f + c_W}}. \tag{2.19}$$

In cases where this formula suggests M_∞ is large enough such that the approximation $M_\infty \ll 1$ fails, it is possible to instead use Newton's method to solve eq. 2.16 numerically. Predictions using both methods are given in figure 3, indicated by circles and diamonds on the right hand wall of each plot.

For any beneficial plasmid, we can now determine the eventual population share that a mutant will have after one selective sweep, based on its initial fitness advantage, Δf, its propensity to conjugate and share plasmids, c_M, and the rate at which the (potentially mixed) wild type population shares plasmids c_W. If we consider costly plasmids within this model ($\Delta f < 0$), there are two possibilities to consider. If $\Delta f + c_W < 0$, then the plasmid spreads too slowly and fitness effects drive it to extinction: $M \to 0$, $W_\ominus \to 1, W_\oplus \to 0$, regardless of the identity of the mutant. If the plasmid spreads sufficiently quickly ($c_W > -\Delta f$), it will act as a parasite- reaching high saturation in the population despite its costly nature. In this case, eq. 2.19 is applicable. We see that, for costly plasmids, "generous" bacteria have a very slight advantage, but this advantage pales in comparison to the disadvantage they experience when sharing beneficial plasmids (figure 5).

2.2. Evolutionary trajectory

In the previous section, we considered demographic shifts as the result of a single selective sweep. Let us now move on to consider how the balance of chromosomal genes evolves over *many* selective sweeps. Let s_i refer to the fraction of the population containing S-type genes at the start of the i^{th} selective sweep. Each time a beneficial plasmid arrives (through mutation or immigration), it is acquired by an individual, 'the mutant', which will be secretive with probability s_i and generous with probability $1 - s_i$. If the mutant is secretive, then by the end of the evolutionary sweep, secretive individuals will occupy $s_{i+1} = M_\infty + (1 - M_\infty)s_i$ of the population. In contrast, if a generous individual is chosen the final population will have $s_{i+1} = (1 - M_\infty)s_i$. Note that M_∞ is calculated using (among other things) the conjugation rate of the mutant (see eq. 2.19). We are likely to see very different final states depending on the secrecy/generosity of the initial plasmid carrier.

Figure 6 shows lineages with low conjugation rate taking over a population; this take over is initially very slow (taking thousands of evolutionary sweeps), accelerating as secretive genes take over and overall conjugation rate is reduced.

In a population with a generally high rate of conjugation ($c_W \approx \Delta f$), almost any non zero c_M will quickly lead to plasmids spreading through the population, hence the advantage enjoyed by secretive lineages is minor. If conjugation rates in a population are already low secretive chromosomal genes enjoy a far greater advantage, reaching high population density after only a few selective sweeps. Regardless of the exact parameters however, given sufficient time, secrecy is favored.

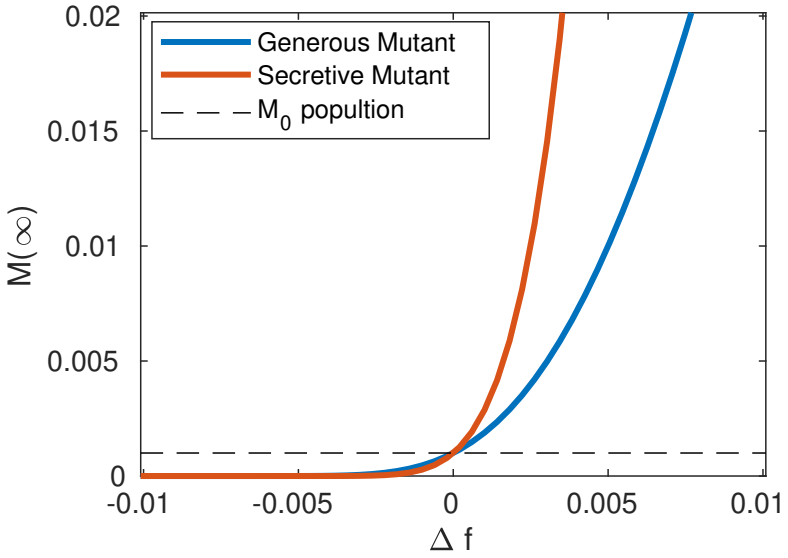

Figure 5. Graph of M_∞, as approximated using equation 2.19. For this example, we consider $c_W = 10^{-2}$, $M_0 = 10^{-3}$. When plasmids arrive in a 'generous' bacteria, we assume $c_M = c_W = 10^{-2}$, when they arrive in a 'secretive' bacteria, we assume $c_M = 10^{-4}$. For $\Delta f < 0$, secretive mutants have a slight disadvantage. For $\delta f > 0$ they have a significant advantage. Due to the simplicity of this exploratory model, there are a number of cases we don't consider: in particular, fitness that varies in time, or fitness that varies based on the genetic background of the host.

2.3. An agent based model

The simplified evolutionary dynamics considered in the section 2.1 indicate that chromosomal genes which repress plasmid spread gain an evolutionary advantage (however slight) over long evolutionary time frames. While useful for the sake of analytic tractability and building intuition, this simplistic model ignores many important features, several of which may contribute to stabilizing plasmid sharing. In order to test the robustness of the previous findings, let us now consider a more detailed agent based model. We will start by constructing a 'baseline' model, and then will consider a number of extensions.

Consider a population of bacteria on an N by N grid with periodic boundary conditions. Each grid cell is either empty (X), or contains a single bacterium which is either generous (G), secretive (S). In nature, plasmids exist in a number of 'incompatibility groups' [35–37] - plasmids belonging to the same incompatibility group make use of the same regulatory proteins, and thus interfere with one anothers' reproduction, driving one another to extinction whenever they exist within the same cell. Plasmids from different incompatibility groups co-exists with no such interference. In our model, we assume that each living bacteria possess one of $k = 4$ possible plasmids from each of q incompatibility groups. If there are (for example) $q = 3$ incompatibility groups, the state of a single grid cell is given by a vector $\mathbf{v}$; for example $\mathbf{v} = \{G, 2, 1, 3\}$ or $\mathbf{v} = \{S, 1, 1, 2\}$.

Individual cells die at some death rate $d(\mathbf{v}, t)$, and reproduce at some reproduction rate $r(\mathbf{v}, t)$, both of which depend on the plasmids contained within the cell, and the current time. Dying cells will transition to the state $\mathbf{v} = \{X, 0, 0, 0\}$. Reproducing cells will select a direction at random, and duplicate

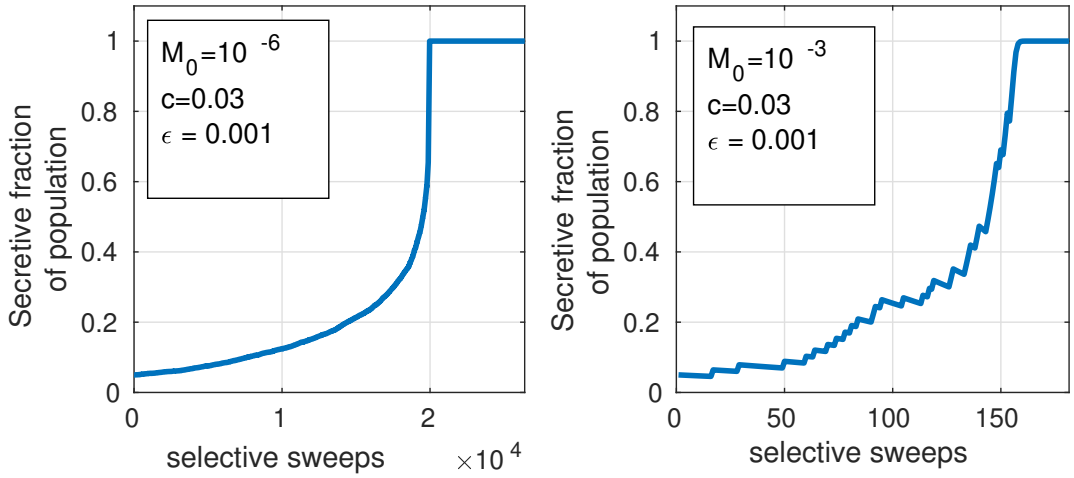

Figure 6. With each selective sweep the balance of the population changes. Over many sweeps, this leads to the fixation of lineages which are less inclined to share plasmids. However, for large population sizes (Left) this generally takes a very very long time, indicating that this effect is probably entirely overpowered compared to the many many other effects influencing evolution. (Right) we repeat the same, but with smaller population size/larger $M(0)$, so as to make the individual jumps more visible.

their state into the adjacent cell in the chosen direction, assuming the target cell is empty. If the selected cell is occupied, replication fails; this gives a crude representation of competition.

Plasmid mutation within cells occurs at some low constant rate $m = 10^{-5}$; cells undergoing mutation select a plasmid group randomly, from 1 to q, and set its value to some random value from 1 to k. This mutation process is included so as to prevent plasmid extinction and ensure a minimal level of variance in the population.

Finally, cells conjugate at a rate $c(\mathbf{v})$; for the purposes of the current exploration, conjugation rate depends only on chromosomal genes, with default rates $c(G) = 0.1$, $c(S) = 10^{-3}$ and $c(X) = 0$. A conjugating cell selects a neighboring grid cell uniformly at random, and replaces one of the neighbors plasmids with its own value. If the selected grid cell is empty, conjugation fails. While it is well known that many plasmids possess surface exclusion mechanisms in order to prevent invasion by incompatible plasmids [14], these exclusion mechanisms are only partially effective [38]. For the purposes of the current investigation, we assume that $c(\mathbf{v})$ is the conjugation rate after taking preexisting plasmid exclusion mechanisms into account.

All events are assumed to be exponentially distributed, and we simulate the entire system using Gillespie's Algorithm [39]. Simulation code is avaliable via Github [40]. A schematic illustration of this model is given in figure 7.

2.3.1. Base case

The agent based model described above is rather detailed. For the sake of concreteness, let us begin by considering the case where each bacteria contains only a single plasmid ($q = 1$), and this plasmid comes in one of $k = 4$ varieties. We assume a grid size of $N = 64$, that is to say 64 cells wide, and

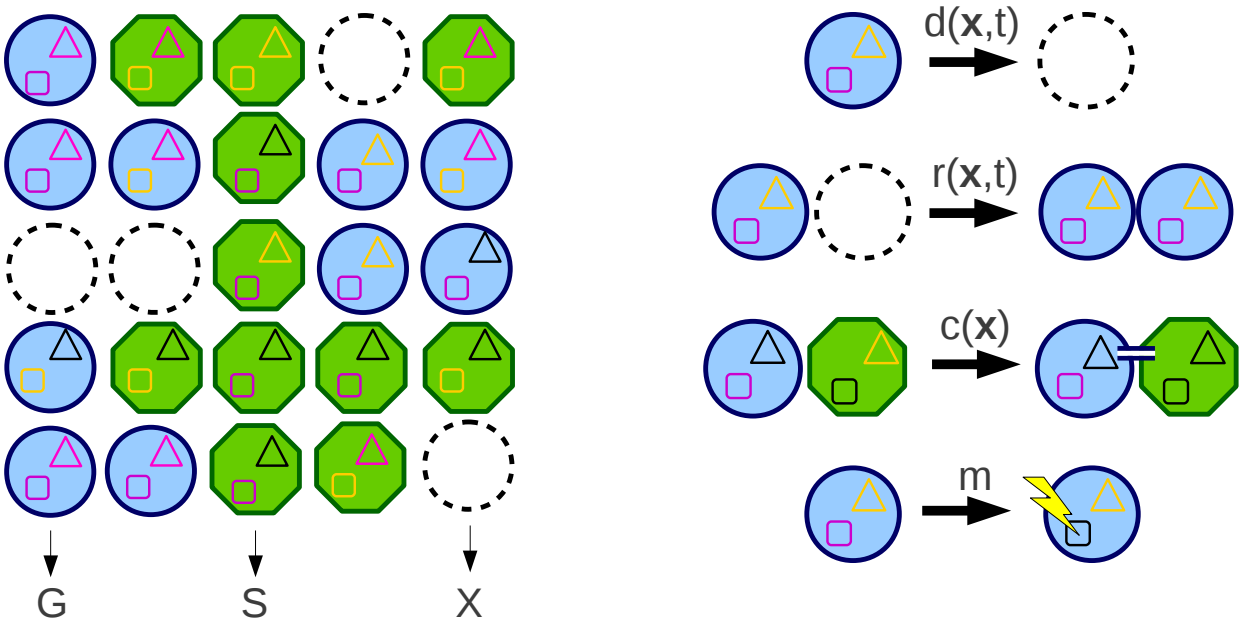

Figure 7. (Left) An N by N grid ($N = 5$) with periodic boundary conditions containing 'generous' type bacteria (blue circles), secretive bacteria (green octagons) and empty space. In this example, each bacteria contains a plasmid belonging to $q = 2$ incompatibility groups (one group represented by triangles, the other by squares). Each plasmid comes in one of $k = 3$ different variants (variants distinguished by colour). (Right) The system evolves according to four mechanisms: death, birth, conjugation and mutation.

64 cells deep. Each grid cell initially has a 54% chance of containing generous bacteria, 6% secretive bacteria and 40% empty space. The plasmid type for each bacteria in the initial population is selected uniformly at random.

At time $t = 0$, the simulation selects $d(\mathbf{v}, t)$ and $r(\mathbf{v}, t)$ uniformly at random in the ranges $[0.35, 0.42]$ and $[1, 1.3]$ respectively, for each plasmid state. At time $t = \tau = 50$, and every τ minutes thereafter, new death and replication rates are selected for each plasmid state. We refer to each such time window as an 'epoch' and refer to τ as the 'epoch time'. This approach gives a rudimentary approximation of the manner in which plasmid fitness changes with variable environmental conditions. Because death and reproduction rates are selected at random, we are implicitly relaxing our previous assumption that plasmids are beneficial. At any given time, a randomly selected plasmid may either improve or reduce a hosts fitness relative to the rest of the population.

Illustrations of the systems state throughout one simulation run are given in figure 8. Each simulation is allowed to run until either $t = 20,000$, or one of G or S is driven to extinction.

We run 100 simulations using the parameter values described. In 47 out of 100 simulations, despite their low initial population, secretive chromosomes reach fixation in the population, driving generous chromosomes to extinction. This rate is nearly five times higher than we would expect for a neutral

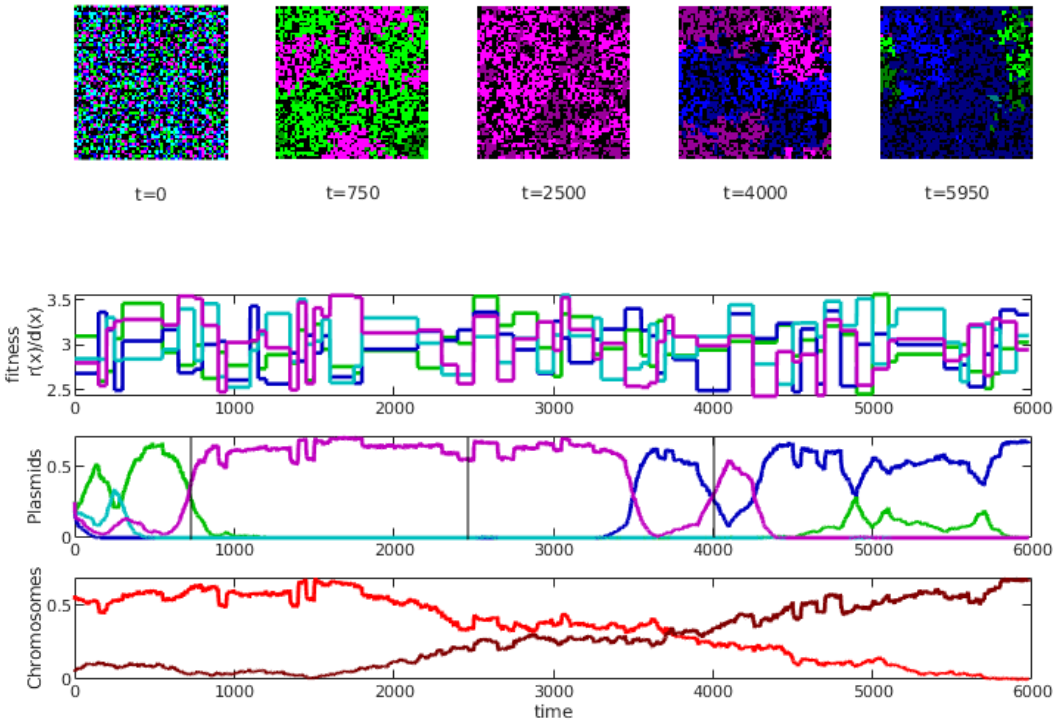

Figure 8. (Top row), illustrations of the full system state at a variety of times. Grid is 64×64 cells. Pixel colours give the identity of the corresponding plasmid, with bright colored nodes being generous, and darker nodes secretive. (Graphs, top) Every τ time steps, new $r(\mathbf{v})$ and $d(\mathbf{v})$ values are selected, leading to different 'fitness' values $r(\mathbf{v})/d(\mathbf{v})$ for each of the four plasmid types. (Graphs, middle) The abundance of each plasmid strain rises and falls throughout the simulation. (Graphs, bottom) The abundance of generous and secretive chromosomal genes (bright and dark, respectively). Over time, secrecy is favored.

mutation, which would reach fixation with 10% probability (as they represent 10% of the initial non-empty population). Generous chromosomes reach fixation in 50 of 100 simulations. In cases where secretive chromosomes reach fixation in the population, the average fixation time is $t_{fix} = 4336$ (86.7 epochs). The average time until secretive chromosomes reach extinction is $t_{ext} = 2328$ (46.6 epochs). What these results indicate is that simple 'network reciprocity' (discussed later) is insufficient to stabilize conjugation, at least, not for the simple 2d lattice considered here.

2.3.2. Exploring parameter space, and environmental heterogeneity

With the base case now covered, let us now consider a range of different simulation conditions. Our aim in what follows is to determine the robustness of the previous results, and determine which aspects of the system (if any) affect the fixation probability of secretive genes. In all simulations, we assume $N = 64$, and initially grid cells are empty with 40% probability, generous with 54% probability and secretive with 6% probability. Table 1 summarizes the results of this section.

We start with two 'sanity tests'. First, we consider a 'control condition', in which $c(S) = c(G) = 0.1$.

In this neutral case we expect a fixation probability for S precisely equal to the initial population fraction, $\frac{6}{60}$. In practice, we observe 12% fixation, 81% extinction, and 7% timeout (Scenario # B, table 1). This is consistent with what we would expect by chance. Next, in order to test sensitivity to simulation geometry, we reuse the same parameters as the base case simulation, but allow bacteria to conjugate and replicate diagonally, such that each cell is adjacent to 8 others rather than 4 (scenario C). In this case we observe 40% fixation, 60% extinction; a slightly lower fixation rate than the base case, but still substantially higher than would be expected given neutral selection. This indicates that our results are not overly sensitive to our choice of geometry. Both fixation and extinction occur roughly 20% faster than in the base case.

With these basic tests out of the way, let us now explore variations of the more biologically meaningful parameters. We consider increasing/decreasing epoch times to $\tau = 250$ and $\tau = 5$ (corresponds to considering a more or less stable environment), decreasing/increasing conjugation rates to $c(S) = 10^{-5}$ and $c(S) = 10^{-2}$ (stricter and looser plasmid suppression, respectively), and setting the mutation rate a factor of 30 higher or lower(scenarios DEFGHI on table 1). In *all* cases, fixation probability for S-type genes remains higher than the 10% we would expect for neutral selection. For the case of weak plasmid suppression, we observe 32% fixation, the remaining scenarios all have fixation rate higher than 40%.

We also consider the case of either increasing death rate to $d(\mathbf{v}) = [0.75, 0.90]$, dramatically reducing population density, or decreasing it to $d(\mathbf{v}) = [0.05, 0.06]$, drastically increasing population density. In the former case, the fixation probability of S type chromosones drops to 11% – close to what we would expect in the neutral case. This suggests that when populations are sparse enough to render plasmid conjugation rare and ineffective, plasmid suppressing genes have no evolutionary effect, either positive or negative (scenario J). In the low death rate/high density case, the most common result of simulations is timeout; no meaningful conclusions can be drawn in this parameter regime (scenario K). If we simplify the model by assuming full population density, and that all death events are immediately followed by a corresponding birth event (treating the system as a death-birth Moran process on a graph [41,42]) simulation speed can be significantly increased, and the timeout condition at $t = 20,000$ can be removed. In this case fixation occurs in 46% of cases, close to the base rate. When fixation does occur, the average time period is $t_{fix} \approx 44,000$. Extinction occurs by $t_{ext} \approx 17,400$ on average. Taken together, these simulations indicate that the advantage of secretive type chromosomal genes is exceptionally robust to changes in parameter values. If we desire a system that will stabilize sharing of plasmids, simple changes to parameter values are insufficient.

In natural settings, plasmids frequently contain resistance [4] and metabolic [2] genes adapted for specific environments. For this reason we might expect plasmids to have particular evolutionary importance in and around transition regions or boundaries between different environments [43]. In order to explore the effects of environmental variability, we extend the basic model to one in which each grid cell is assigned an 'environment' parameter (either A or B). Death and reproduction functions are replaced with environment specific functions $r_A(\mathbf{v}), r_B(\mathbf{v}), d_A(\mathbf{v}), d_B(\mathbf{v})$; with each epoch, the rate functions associated with either A or B (but not both) are changed. In order to explore the effects of variable environment, we consider four different environmental geometries: block, checkerboard, random and gradient (see figure 9). Death and replication rates in the gradient geometry are given as position dependent linear combinations of those found in A and B (hence $r(\mathbf{v}) = xr_A(\mathbf{v}) + (1 - x)r_B(\mathbf{v})$ for some position dependent x in the range [0,1]). Regardless of the environmental condition consid-

ered, fixation is observed to occur in roughly 40 – 50% of all cases (scenarios KLMN). Hence, it would appear that environmental heterogeneity is *also* insufficient to stabilize plasmid sharing.

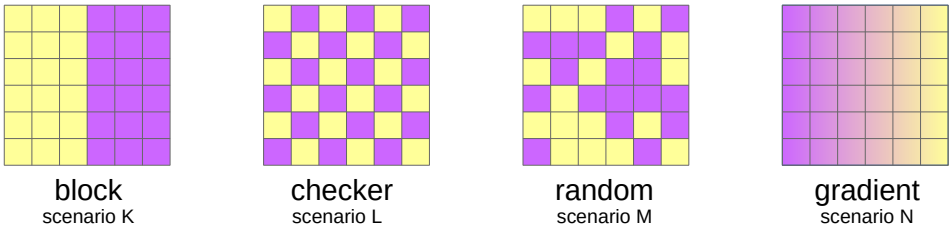

Figure 9. Four possible environmental geometries; from left to right we have 'block', 'checkerboard','random' and 'gradient' type geometries. For each block, checkerboard and random conditions, cells experience one of two death and reproduction rates ($r_a(\mathbf{v})$ or $r_b(\mathbf{v})$) depending on the color of their grid cell. This allows us to simulate the behavior of cells around boundary regions (of various shapes). Block geometry assigns environment A to all cells in the left half of the simulation and environment B to the right half. Checkerboard assigns opposite environment conditions to adjacent cells. In the random condition each grid cell has a 50% chance of having each environment type, each grid cell is determined independently. In the gradient condition, $r(\mathbf{v})$ and $d(\mathbf{v})$ are a linear combinations of the death and reproductive rates that would be experienced in 'pure' environments.

3. Discussion

The paradox we pose here, namely 'how is conjugation maintained?' is by no means the first paradox in the study of evolutionary dynamics, nor is it likely to be the last. Two paradoxes of the past, namely, the evolution of sex, and the evolution of altruism (also called the paradox of cooperation), bare striking resemblance to the current conundrum, as does the so called 'reduction of modification' principle. Let us take a brief detour to examine these principles and paradoxes, their resolutions, and the similarities and contrasts to the question currently under study.

3.1. Can recombination preserve HGT?

In the study of the evolution of sex, simple modelling suggests that asexual mutants would possess two major advantages over variants that reproduce sexually. Firstly, all members of the asexual species are able to reproduce, in contrast, for sexual species, only the female population can reproduce (the cost of males) [44]. At the individual level, asexual individuals are able to pass on 100% of their genes to each offspring, while sexual individuals pass on only 50% of their genes (the cost of meiosis) [45, 46]. Classically, this paradox is resolved by considering 'the evolution of evolvability' [47]. For asexual species, if beneficial mutants arise independently in two different lineages, one will inevitably drive the other to extinction, a process known as clonal interference. Clonal interference severely limits the speed of evolution, especially for large population size [48]. In contrast, sexual recombination causes beneficial genes to accumulate. Mating allows genes initially present in separate lineages to come together in a single organism, both advantages can be retained, and the rate of evolution scales with population size [49, 50].

Similar to sex, HGT allows genes to accumulate, and recombine. In contrast to sex, in which all genes have (approximately) equal chance of transfer and recombination, for bacterial conjugation, plasmid bound genes and other mobile gene elements gain the benefits of recombination, while core chromosomal genes do not. This leads to the question; are the benefits of recombination enough to stabilize plasmid sharing in a bacterial population? In order to examine this we consider a collection of cells each containing $q = 3$ plasmids, each from a different incompatibility group. For this simulation, we assume that individual conjugation events transport only a single plasmid, and thus, at the boundary between two clonal lineages, conjugation will quickly lead to a wide variety of different plasmid combinations, if cells are generous. This effect will be significantly reduced for secretive variants, leading to increased clonal interference.

In order to study the multi-plasmid case, we must first determine how $r(\mathbf{v})$ is defined as a function of the three plasmid values v_1, v_2, v_3. We consider three possible cases, in order of increasing complexity: in the first case (scenario Q), we assume that $r(\mathbf{v})$ for each possible plasmid combination is selected in the range $[1, 1.3]$ entirely independently, and that new $r(\mathbf{v})$ are selected in each epoch. The fitness of the combination $[G, 1, 1, 1]$ and that of $[G, 1, 1, 2]$ are entirely independent of one another. One can think of this as being a 'complex' genespace, the value of each plasmid varies based on the presence and absence of other plasmids. In the second case (scenario R), $r(\mathbf{v})$ is formed as a linear combination $r(\mathbf{v}) = 1 + 0.1r_1(v_1) + 0.1r_2(v_2) + 0.1r_3(v_3)$. In this case $r_i(v_i)$ are selected in the range $(0, 1)$. Each plasmid contributes to fecundity independently, and bacteria can succeed by 'optimizing' for each plasmid incompatibility group independently. Finally, we consider the hybrid case, in which half of the fitness is determined via each of the previous two methods (scenario S). In order to observe the benefits of recombination and the cost of clonal interference, mutants must arrive frequently enough such that multiple mutant strains are competing at any one time. In order to achieve this, we assume $N = 128$ and $m = 10^{-4}$ for this batch of simulations.

In all three cases we observed significantly more fixation events than the control case (Scenario B), $45\%, 47\%$ and 51%, respectively. Fixation and extinction times are comparable to the base case (Scenario A). Taken together these results indicate that recombination is insufficient to maintain plasmid sharing (at least at the scale simulated here). Based on what we have observed here, selection on the static chromosomal gene is entirely unaffected by combinatorial effects between plasmids (or, at most, such effects are drowned by noise).

3.2. Is HGT a public good?

If the advantages of recombination are not enough to stabilize plasmid sharing, let us turn our attention to another evolutionary paradox, and its resolutions. The second major paradox in the study of evolutionary dynamics is the emergence and maintenance of cooperation: slime molds cooperate in order to form stalks and bud [51], wolves hunt in packs [52], meerkats keep watch, and humans cooperate on vast and complex scales spanning the entire globe [53] and beyond [54]. Cooperation is powerful, and yet at every turn, it is the organism that most benefits its own lineage that evolution will select. The lazy wolf, the cowardly meerkat, the spore that forces itself into the fruiting body rather than the stalk. Evolution is a game with winners and losers, not one that is played for fun. Given the ever present advantage of selfish behavior, how then is cooperation maintained? — this is the question posed in the study of cooperation and public goods.

Numerous answers to this question have been proposed [55], and it has been suggested that HGT

may well be supported by many of the same mechanisms [25,26]. Much like cooperative and altruistic behavior, HGT involves one individual paying some 'cost' for the benefit of another; how this benefit is distributed between the receiving cell and the plasmid itself is unknown, and likely to vary wildly between contexts. Unlike classical public goods interactions however, HGT involves the transfer not of resources, but of *information*. Classical public goods are shaped by evolution, and act on the scale of individual organisms and communities. In contrast evolution both shapes, and is shaped by, HGT. Gene transfer uses the same language and acts on the same time scale as evolution itself. Because transferred genes have the potential to benefit the recipient for generations to come it is unclear what time horizon to use when trying to measure the 'value' of the genes given, or the 'cost' which the donor must pay.

With these differences and similarities in mind, let us now consider a variety of mechanisms that have been proposed for stabilizing cooperation, and how they might apply in the context of HGT. We note that there remains some debate in the literature as to the level of overlap between these mechanisms [56, 57]. Here, we make no claims as to the distinctness or similarities between alternative evolutionary mechanisms, and instead simply err on the side of inclusivity whenever doing so appears physically appropriate in the bacterial context.

The most well known and well understood mechanism for stabilizing cooperation at evolutionary time scales is kin selection [58]. Kin selection suggests that when the benefits of cooperation are disproportionately directed to ones relatives, cooperation can be stabilized by evolution. Cooperative genes persist because they inevitably end up helping *other copies* of those same cooperative genes. This is in some sense similar to parental care, albeit on a wider scale. In the context of HGT, kin selection plays a strange role: those individuals most closely related to a plasmid donor are also those most likely to *already possess* a given plasmid, and hence gain no benefit. This is the crux of our recent discussion with Tatiana et al. [26, 27, 59]. This relatedness issue may be partially offset: if a plasmid is well adapted to a donor cell, it is likely to be well adapted to a closely related recipient [16, 60], so any kin that somehow do not already possess the focal plasmid will benefit disproportionately from acquiring it. The kin selection hypothesis is further hampered in the bacterial context, as there is little evidence that bacteria are able to track their kin (beyond spatial proximity), unlike vertebrate animals. Given that our previous simulations already include a spatial component, it would appear kin selection via spatial association is ineffective.

A second resolution that has been proposed to the paradox of cooperation is reciprocity; that is to say, an individual who pays a cost today may well be on the receiving end of generosity tomorrow. Reciprocity comes in a variety of forms. Direct reciprocity involves two parties directly benefiting one another, such as plants supplying root fungi with sugars in exchange for key nutrients [61,62]. Indirect reciprocity takes place when individuals provide benefit too, and gain benefit from, a wider community; "I am willing to help you because I trust that someone else will help me" [63, 64]. Finally, 'network' reciprocity involves repeated interactions between neighboring individuals; these individuals are also likely to be related to the donor. Neither direct, indirect nor network reciprocity would appear relevant in the context of bacterial conjugation. To the best of our current knowledge, bacteria are incapable of tracking the complex reputational networks required for indirect reciprocity, nor are they selective in who they donate to. Because export of plasmids depends on a variety of protein complexes within the donating cell [65], plasmid donation is an explicitly 'single directional' process, and hence direct reciprocity would also appear unlikely. Network reciprocity is already baked into the agent based

model we are using, and has proven ineffective at preventing the invasion of secretive chromosomal genes.

One final resolution to the paradox of cooperation is the hypothesis of 'group selection'. Under this model, individuals are separated into M groups, and while selfish genes have the advantage on an individual level within each group, groups with a greater proportion of cooperative individuals are more likely to spread and divide than groups containing more selfish individuals. Unlike reciprocity and kin selection, this hypothesis is particularly suited to the microscopic world of bacteria. In order to test group selection in the context of HGT, we consider an alternative simulation, in which each column of the grid is considered one 'group', and individuals may reproduce and conjugate to any cell within their column. Group selection can be implemented in a variety of different ways; for a classical review, see Wade 1978 [66]). For the sake of this initial exploration, we mimic Traulsen and Nowak's minimal model [67]; cells duplicate within a group according to their fitness, expanding into the free space available, until eventually the group reaches size N. At this stage, with each duplication event, the group will either split with probability w, or a randomly selected group member will perish with probability $1 - w$. When a group splits, one of the other groups is selected at random and is eradicated, and each member of the splitting group either migrates to the new space, or stays put (50% probability of each outcome). Unlike previous simulation scenarios, death is not explicitly modeled, and is treated as a downstream consequence of either group splitting or individual replication. Conjugation is allowed to occur within groups, but not between groups; for the sake of conjugation, all groups are considered well mixed.

In terms of simulation, we consider two separate scenarios. In both cases, we assume $w = 10^{-4}$ (rare group splitting). For scenario U, we assume initial conditions 54% generous, 6% secretive (as previously). This results in generous chromosomes overtaking the population 91% of the time, and secretive chromosomes reaching fixation 9% of the time, a result in line with neutral selection. In order to investigate this further, we also consider the same simulation with 'balanced' initial conditions with initial populations of both generous and secretive chromosomes at 30% (scenario V). Once again we find results consistent with neutral selection: secretive genes reach fixation 52% of the time, and are driven to extinction 48% of the time. When plasmids are unable to transmit between groups, the advantage of hoarding them is neutralized, and fixation probability reflects initial population fractions almost perfectly. We do not, however, observe selection *against* secretive chromosomal genes.

3.3. Reduction principle

One possible interpretation of the above results is that we have happened across another example of the previously studied 'reduction of modification' principle [28, 29, 31, 68]. The reduction principle states that genes that reduce 'change' (for example, reducing mutation, recombination or migration) are selected for across a broad range of modelling assumptions. While potentially relevant, similarities between the generalized reduction principle and our current observations would appear to be at best superficial: here we have studied haploid populations far from equilibrium, under the effects of genes that regulate changes in *neighboring* organisms. Previous modifier gene studies have focused primarily on diploid populations close to equilibrium, in which modification effected the focal individual itself. While S and G can be viewed as 'modifier genes' in the classical sense [28] using the appropriate change of variables, doing so leads G and S to have non-neutral fitness effects.

The mechanisms driving reduction of modification in each context appear to be strikingly different

as well. When modifier genes control mutation, recombination or migration, reduction of modification acts to minimize the creation of inefficient gene combinations and maximize growth; evolution acts to exploit a peak in fitness space as efficiently as possible. In the context of HGT, reducing conjugation can be seen more as denying benefits to ones competition. In so doing, the overall health and growth rate of the population is reduced rather than maximized.

While the overall results may be the same, it would appear, at least for the time being, that reduction of conjugation represents a phenomena distinct from the previously studied 'reduction of (self) modification'.

3.4. Overview

As is clear from these experiments, the selective advantage of conjugation suppressing genes is robust across a wide range of parameter regimes, including regimes that have previously been found to stabilize cooperation or sexual reproduction. Fixation probability is only reduced back to the 10% probability expected for neutral evolution in circumstances where plasmid transmission is in some sense disabled (group isolation, sparse populations). Generous chromosomes are not found to have an evolutionary advantage for *any* of the scenarios considered.

We also observe that once fixation of secretive chromosomes is likely, fixation probability itself is relatively unaffected by parameter values. The general uniformity in fixation probability across scenarios would appear to indicate that extinction probability is governed almost entirely by local processes, and that larger scale and longer time frame effects (environment, mutation, epoch time) have little to no effect on the local dynamics. Given that extinction (when it occurs) happens 2-3 times faster than fixation, this may be indicative that fixation dynamics of secretive chromosomes are governed primarily by the probability of 'early extinction'. Once early extinction is avoided, fixation occurs with high probability. The one exception to this 'uniform fixation probability' is scenario F, in which we consider $c(S) = 10^{-2}$. Significantly weakened HGT suppression would appear to reduce the evolutionary advantage of secretive chromosomal genes.

See table 1 for a summary of these results.

4. Conclusions

Horizontal gene transfer amongst microorganisms is a significant contributor to the vast complexity and variety of life we see in the world today. It allows for the sharing of resistance [1,4] novel metabolic pathways [2], and virulence factors [3], and is critical both to our understanding of life, and our forays into medicine. HGT is also something of a mystery, complicating our understanding of the evolutionary tree of life. Plasmids have been described as 'paradoxical' in the litereature [24], with Harrison and Brockhurst observing that costly plasmids should be purged via purifying selection, while beneficial plasmid genes would be expected to integrate into the chromosome, rendering the plasmid bound copy redundant. In this work, we present another paradox, namely "why is the *sharing* of plasmids not selected against?"

While past models have explored the dynamics of HGT and plasmid conjugation from the point of view of the plasmids themselves [15, 17, 69–73], this is only half the evolutionary story. Here we have instead focused on the dynamics of chromosomal DNA, exploring a number of models that explicitly model evolution of both plasmids and chromosomes, and the interaction between the two.

Table 1. Table of results of simulations. For each scenario, we give a brief description, fixation and extinction probability (of secretive chromosomal genes), and the mean time to fixation or extinction. By default, initial population densities are 6% secretive bacteria, 54% generous (with the exception of scenario V, where we have 30%, 30%). Because not all simulations reach fixation before time out, probabilities do not add to 100%. Aside from scenario L, all scenarios have a time cut off of $t = 20,000$.

Scenario	Simulation description	Prob. fix.	Prob. ext.	mean t_{fix}	mean t_{ext}
A	Base case	47%	50%	4336	2328
B	Control condition, $c(S) = c(G) = 0.1$.	12%	81%	17182	3062
C	Diagonal travel	40%	60%	3606	1740
D	Epoch time $\tau = 5$	47%	53%	4776	2216
E	Epoch time $\tau = 250$	44%	55%	4488	2007
F	Weak Plasmid Restriction $c(S) = 10^{-2}$	32%	68%	4314	1507
G	Strong Plasmid Restriction $c(S) = 10^{-5}$	46%	51%	4299	1778
H	High Mutation rate ($\times 30$)	45%	55%	4878	1816
I	Low mutation rate ($\div 30$)	43%	56%	4081	2470
J	Sparse population	11%	89%	1595	440
K	Dense population	0%	42%	-	11687
L	Dense population (Death-birth process)	46%	54%	44433	17400
M	Environment: 2 blocks	41%	58%	4826	2565
N	Environment: Checkerboard	43%	56%	4836	1555
O	Environment: Random	47%	51%	4462	2021
P	Environment: Gradient	41%	59%	4205	1791
Q	Multiplasmid: linear sum $f = f_a(v_a) + f_b(v_b) + f_c(v_c)$	45%	49%	4600	2007
R	Multiplasmid: Independent $f = f_{abc}(v_a, v_b, v_c)$	51%	46%	4630	1540
S	Multiplasmid: Hybrid case $f = f_{abc}(\mathbf{v}) + f_a(v_a) + f_b(v_b) + f_c(v_c)$	44%	51%	4452	1810
T	Multiplasmid: $q = 3, k = 25$	47%	52%	3106	1213
U	Group selection.	9%	91%	7721	2438
V	Group selection, balanced initial conditions	52%	48%	5400	5850

We are able to show that for a wide array of modelling assumptions and parameter values, chromosomal genes that restrict bacterial conjugation rates have an evolutionary advantage, at least, on the cellular scale considered here. With each change of environment or selective sweep, individual cells which hoard advantageous plasmids to themselves fare better than those which share freely; this occurs despite ignoring the non-trivial costs [20, 23, 74] of conjugation entirely. The central mechanism behind this appears to be simple competition, playing out across multiple generations: those bacteria which share plasmids readily increase the fitness of their competition, not only in the present, but for many generations to come. In contrast, repressing conjugation allows a bacterial lineage to exploit the the advantage granted by any beneficial plasmid they come across.

To understand and interpret our results, we can compare HGT to a number of previously studied phenomena. As suggested by Dimitriu el al. [25], HGT can be seen as a costly public good, albeit one that involves the transfer of information rather than resources. Unlike direct resource sharing, useful genes, once given, may assist the recipient for many hundreds of generations to come, acting on the same time scale as evolution itself. There are also complicating factors in terms how HGT interacts with kin selection: sharing plasmid DNA with kin is likely to be ineffective, as those most related to a donor organism are also those most likely to *already* have any given plasmid. The importance of HGT is in some sense *defined* by its ability to cross vast evolutionary distances. In our simulations, we find that simulations designed in order to encourage selection for public goods fail to promote HGT via conjugation.

Rather than viewing HGT as a public good, an alternative interpretation might view genetic mixing via HGT as more similar to sexual recombination. Unlike sexual recombination, where all genes benefit from recombination (approximately) equally, plasmid conjugation comes with an intrinsic asymmetry, between plasmid genes (which spread and recombine), and chromosomal genes (which don't). This asymmetry leads to conflict between genes stored on these two separate loci, particularly with regard to preferred conjugation rate. Notably, we find that initial explorations with simulations designed to maximize the advantage of recombination fail to select for plasmid conjugation. Further modelling in this area is necessary.

So where does this leave us? Horizontal gene transfer and conjugation *are* ubiquitous in biology. Experiments, such as those conducted by Dimitriu el al. [26] *do* indicate that plasmid sharing is favored in experimental set ups which select for cooperation. As always, where modelling and observation disagree, it is the model which is lacking. There is no point shaking our fists at nature for disobeying the equations. Rather, our hope in this paper is to point out that bacterial generosity in the sharing of plasmids can not be taken for granted, especially when viewed from the point of view of individual chromosomal genes, which may benefit significantly from restricting plasmid spread.

HGT via conjugation is similar to a *number* of previously studied phenomena, but comparisons to these phenomena are at best analogue, and at worst misleading. Conjugation is its own phenomena, distinct from those studied in the past. Regulatory mechanisms controlling bacterial conjugation serve two masters, and it is easy to produce circumstances where chromosomal and plasmid bound genes experience very different evolutionary pressures.

In terms of resolving this paradox, the work here has been exploratory rather than exhaustive, and there remain a significant number of avenues to explore. Here we study only the evolution of chromosomal genes, and ignore plasmid-chromosome co-evolution; it is entirely plausible that fast evolution of plasmids may be enough to escape any 'regulatory restriction' chromosomal genes might impose.

The cost of restricting conjugation itself may be high enough such that any such genes will be quickly purged from the 'core genome' of a species as non-essential. Here we have considered models with limited 'space', in some sense dominated by competition; considering models where plasmids are able to enlarge the ecological niche, or allow spread into new regions could potentially give very different results. In addition, in our simplified models here we have treated conjugation as being constant over time; bacterial conjugation *in vivo* is known to be a tightly regulated process, with conjugation being up or down regulated by a variety of quorem sensing mechanisms [12]. More detailed modelling of these regulatory mechanisms may lead to differing results.

When plasmids code for genes promoting public goods (such as siderophore production [75] or virulence factors [76, 77]) then sharing of plasmids may be directly beneficial to the donating cell. What concentration of mutualistic plasmid genes would be needed to stabilize HGT via this mechanism is as yet unknown. Our discussion here has been very much focused on the evolution at the scale of a cell. It seems possible that alternative models of group selection that better consider evolution on the scale of entire bacterial communities may demonstrate not observed in the simple models considered here. We may yet find justification for reconsidering kin selection, or some other hypothesis that we have discounted as improbable in the current context.

Regardless of the exact models considered by future researchers, it is our hope to encourage discussion, and draw attention to the largely neglected role of chromosomal genes in the study of bacterial conjugation. As demonstrated by those smallest of creatures that fill the world around us, with free sharing and recombination of information, answers to near any problem can be found.

Acknowledgments

We gratefully acknowledge discussion and feedback from Hildegard Uecker and Arne Traulsen, who provided valuable feedback on an early draft of this note, contributing to both clarity, context and focus. Funding for this research was provided by the School of Medicine and Health Science, Carl von Ossietzky Universität Oldenburg, https://uol.de/en/medicine.

Conflict of interest

The authors declare no conflict of interest.

References

1. A. O. Summers, S. Silver, Mercury Resistance in a Plasmid-Bearing Strain of Escherichia coli, *J. Bacteriol.*, **112** (1972), 1228–1236. https://doi.org/10.1128/jb.112.3.1228-1236.1972
2. Z. Shao, H. Zhao, H. Zhao, DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways, *Nucleic Acids Res.*, **37** (2009), e16–e16. https://doi.org/10.1093/nar/gkn991
3. T. J. Johnson, L. K. Nolan, Pathogenomics of the virulence plasmids of escherichia coli, *Microbiol. Mol. Biol. Rev.*, **73** (2009), 750–774. https://dx.doi.org/10.1128%2FMMBR.00015-09
4. P. M. Bennett, Plasmid encoded antibiotic resistance: acquisition and transfer of

antibiotic resistance genes in bacteria, *Br. J. Pharmacol.*, **153** (2008), S347–S357. https://dx.doi.org/10.1038%2Fsj.bjp.0707607

5. T. Dagan, Y. Artzy-Randrup, W. Martin, Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution, *Proc. Natl. Acad. Sci. U.S.A.*, **105** (2008), 10039–10044. https://doi.org/10.1073/pnas.0800679105
6. X. Yang, E. Wai-Chi Chan, R. Zhang, S. Chen, A conjugative plasmid that augments virulence in klebsiella pneumoniae, *Nat. Microbiol.*, **4** (2019), 2039–2043. https://doi.org/10.1038/s41564-019-0566-7
7. K. Chen, E. W. C. Chan, S. Chen, Evolution and transmission of a conjugative plasmid encoding both ciprofloxacin and ceftriaxone resistance in Salmonella, *Emerg. Microbes Infect.*, **8** (2019), 396–403. https://doi.org/10.1080/22221751.2019.1585965
8. S. Peter, M. Bosio, C. Gross, D. Bezdan, J. Gutierrez, P. Oberhettinger, et. al., Tracking of antibiotic resistance transfer and rapid plasmid evolution in a hospital setting by nanopore sequencing, *mSphere*, **5**. https://doi.org/10.1128/mSphere.00525-20
9. C. M. Johnson, A. D. Grossman, Integrative and conjugative elements (ICEs): What they do and how they work, *Annu. Rev. Genet.*, **49** (2015), 577–601. https://doi.org/10.1146/annurev-genet-112414-055018
10. W. G. Eberhard, Evolution in bacterial plasmids and levels of selection, *Q. Rev. Biol.*, **65** (1990), 3–22. https://doi.org/10.1086/416582
11. C. M. Thomas, K. M. Nielsen, Mechanisms of, and barriers to, horizontal gene transfer between bacteria, *Nat. Rev. Microbiol.*, **3** (2005), 711–721. https://doi.org/10.1038/nrmicro1234
12. G. Koraimann, M. A. Wagner, Social behavior and decision making in bacterial conjugation, *Front. Cell. Infect. Microbiol.*, **4**. https://dx.doi.org/10.3389%2Ffcimb.2014.00054
13. P. H. Oliveira, M. Touchon, E. P. C. Rocha, Regulation of genetic flux between bacteria by restriction modification systems, *Proc. Natl. Acad. Sci. U.S.A.*, **113** (2016), 5658–5663. https://doi.org/10.1073/pnas.1603257113
14. M. P. Garciláin-Barcia, F. de la Cruz, Why is entry exclusion an essential feature of conjugative plasmids?, *Plasmid*, **60** (2008), 1–18. https://doi.org/10.1016/j.plasmid.2008.03.002
15. J. P. J. Hall, A. J. Wood, E. Harrison, M. A. Brockhurst, Source-sink plasmid transfer dynamics maintain gene mobility in soil bacterial communities, *Proc. Natl. Acad. Sci. U.S.A.*, 201600974. https://doi.org/10.1073/pnas.1600974113
16. C. Dahlberg, L. Chao, Amelioration of the cost of conjugative plasmid carriage in Eschericha coli K12, *Genetics*, **165** (2003), 1641–1649. https://doi.org/10.1093/genetics/165.4.1641
17. C. T. Bergstrom, M. Lipsitch, B. R. Levin, Natural selection, infectious transfer and the existence conditions for bacterial plasmids, *Genetics*, **155** (2000), 1505–1519. https://doi.org/10.1093/genetics/155.4.1505
18. M. J. Bottery, A. J. Wood, M. A. Brockhurst, Adaptive modulation of antibiotic resistance through intragenomic coevolution, *Nat. Ecol. Evol.*, **1** (2017), 1364–1369. http://dx.doi.org/10.1038/s41559-017-0242-3

19. K. Trautwein, S. E. Will, R. Hulsch, U. Maschmann, K. Wiegmann, M. Hensler, et. al., Native plasmids restrict growth of phaeobacter inhibens DSM 17395: Energetic costs of plasmids assessed by quantitative physiological analyses, *Environ. Microbiol.*, **18** (2016), 4817–4829. https://doi.org/10.1111/1462-2920.13381

20. A. San Millan, C. Maclean, Fitness costs of plasmids: A limit to plasmid transmission, *Microbiol. Spectr.*, **5**. https://doi.org/10.1128/microbiolspec.mtbp-0016-2017

21. D. A. Baltrus, Exploring the costs of horizontal gene transfer, *Trends Ecol.*, **28** (2013), 489–495. https://doi.org/10.1016/j.tree.2013.04.002

22. S. A. Mc Ginty, D. J. Rankin, The evolution of conflict resolution between plasmids and their bacterial hosts, *Evolution*, **66** (2012), 1662–1670.

23. J. Colom, D. Batista, A. Baig, Y. Tang, S. Liu, F. Yuan, et. al., Sex pilus specific bacteriophage to drive bacterial population towards antibiotic sensitivity, *Sci. Rep.*, **9** (2019), 1–11. https://www.nature.com/articles/s41598-019-48483-9

24. E. Harrison, M. A. Brockhurst, Plasmid-mediated horizontal gene transfer is a coevolutionary process, *Trends Microbiol.*, **20** (2012), 262–267. https://doi.org/10.1016/j.tim.2012.04.003

25. T. Dimitriu, C. Lotton, J. Bénard-Capelle, D. Misevic, S. P. Brown, A. B. Lindner, et. al., Genetic information transfer promotes cooperation in bacteria, *Proc. Natl. Acad. Sci. U.S.A.*, **111** (2014), 11103–11108. https://doi.org/10.1073/pnas.1406840111

26. T. Dimitriu, D. Misevic, C. Lotton, S. P. Brown, A. B. Lindner, F. Taddei, Indirect fitness benefits enable the spread of host genes promoting costly transfer of beneficial plasmids, *PLoS Biol.*, **14** (2016), e1002478. https://doi.org/10.1371/journal.pbio.1002478

27. A. Jamieson-Lane, B. Blasius, Comment on "indirect fitness benefits enable the spread of host genes promoting costly transfer of beneficial plasmids", *PLoS Biol.*, **19** (2021), e3001449. https://doi.org/10.1371/journal.pbio.3001449

28. U. Liberman, M. W. Feldman, Modifiers of mutation rate: A general reduction principle, *Theor. Popul. Biol.*, **30** (1986), 125–142. https://doi.org/10.1016/0040-5809(86)90028-6

29. U. Liberman, M. W. Feldman, The reduction principle for genetic modifiers of the migration rate, in *Mathematical Evolutionary Theory*, Princeton University Press, 2014, 111–137.

30. U. Liberman, M. W. Feldman, A general reduction principle for genetic modifiers of recombination, *Theor. Popul. Biol.*, **30** (1986), 341–371. https://doi.org/10.1016/0040-5809(86)90040-7

31. L. Altenberg, M. W. Feldman, Selection, generalized transmission and the evolution of modifier genes. I. the reduction principle, *Genetics*, **117** (1987), 559–572. https://doi.org/10.1093/genetics/117.3.559

32. L. Altenberg, U. Liberman, M. W. Feldman, Unified reduction principle for the evolution of mutation, migration, and recombination, *Proc. Natl. Acad. Sci. U.S.A.*, **114** (2017), E2392–E2400. https://doi.org/10.1073/pnas.1619655114

33. S. P. Otto, T. Day, *A biologist's guide to mathematical modeling in ecology and evolution*, Princeton University Press, 2011.

34. E. L. Ince, *Ordinary differential equations*, Dover publishing, 1956.

35. C. M. Thomas, Plasmid Incompatibility, in *Molecular Life Sciences: An Encyclopedic Reference*(ed. E. Bell), Springer, New York, 2021, 1–3. https://doi.org/10.1007/978-1-4614-6436-$5_5$65 – 2

36. J. Cullum, P. Broda, Rate of segregation due to plasmid incompatibility, *Genet. Res.*, **33** (1979), 61–79, https://doi.org/10.1017/S0016672300018176

37. J. R. Scott, Regulation of plasmid replication, *Microbiol. Rev.*, **48** (1984), 1–23. https://dx.doi.org/10.1016%2Fb978-0-12-048850-6.50006-5

38. C. Gago-Córdoba, J. Val-Calvo, A. Miguel-Arribas, E. Serrano, P. K. Singh, D. Abia, et.al., Surface exclusion revisited: function related to differential expression of the surface exclusion system of bacillus subtilis plasmid pLS20, *Front. Microbiol.*, **10**. https://doi.org/10.3389/fmicb.2019.01502

39. D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions, *J. Phys. Chem.*, **81** (1977), 2340–2361. https://doi.org/10.1021/j100540a008

40. A. Jamieson-Lane, alastair-JL/HGTparadox, 2020, https://github.com/alastair-JL/HGTparadox

41. P. a. P. Moran, Random processes in genetics, *Math. Proc. Camb. Philos. Soc.*, **54** (1958), 60–71. https://doi.org/10.1017/S0305004100033193

42. E. Lieberman, C. Hauert, M. A. Nowak, Evolutionary dynamics on graphs, *Nature*, **433** (2005), 312–316. https://doi.org/10.1038/nature03204

43. R. Hermsen, J. B. Deris, T. Hwa, On the rapidity of antibiotic resistance evolution facilitated by a concentration gradient, *Proc. Natl. Acad. Sci. U.S.A.*, **109** (2012), 10775–10780. https://doi.org/10.1073/pnas.1117716109

44. J. Maynard Smith, What use is sex?, *J. Theor. Biol.*, **30** (1971), 319–335. https://doi.org/10.1016/0022-5193(71)90058-0

45. G. C. Williams, *Sex and evolution*, (MPB-8), Volume 8, Princeton University Press, 2020.

46. R. Williams, Introduction, in *Politics and Technology*(ed. R. Williams), Studies in Comparative Politics, Macmillan Education UK, London, 1971, 7–10. https://doi.org/10.1007/978-1-349-01385-2_1

47. S. F. Institute, C. G. Langton, Artificial life, volume I: *Proceedings of an interdisciplinary workshop on synthesis and simulation of living systems*, Westview Press, Redwood City, Calif, 1989.

48. P. J. Gerrish, R. E. Lenski, The fate of competing beneficial mutations in an asexual population, *Genetica*, **102** (1998), 127. http://dx.doi.org/10.1023/A:1017067816551

49. S. P. Otto, N. H. Barton, The evolution of recombination: Removing the limits to natural selection, *Genetics*, **147** (1997), 879–906. https://doi.org/10.1093/genetics/147.2.879

50. N. Colegrave, Sex releases the speed limit on evolution, *Nature*, **420** (2002), 664–666. https://doi.org/10.1038/nature01191

51. E. A. Ostrowski, Enforcing cooperation in the social amoebae, *Curr. Biol.*, **29** (2019), R474–R484. http://dx.doi.org/10.1016/j.cub.2019.04.022

52. G. Cordoni, E. Palagi, Back to the future: A glance over wolf social behavior to understand dog-human relationship, *Animals*, **9**. https://doi.org/10.3390/ani9110991

53. T. Verma, N. a. M. Araújo and H. J. Herrmann, Revealing the structure of the world airline network, *Sci. Rep.*, **4** (2014), 5638. http://dx.doi.org/10.1038/srep05638

54. J. Bell, *The interstellar age: Inside the forty-year voyager mission*, First edition edition, Dutton, New York, New York, 2015.

55. M. A. Nowak, Five rules for the evolution of cooperation, *Science*, **314** (2006), 1560–1563. https://doi.org/10.1126/science.1133755

56. L. Lehmann, L. Keller, S. West, D. Roze, Group selection and kin selection: Two concepts but one process, *Proc. Natl. Acad. Sci. U.S.A.*, **104** (2007), 6736–6739. https://doi.org/10.1073/pnas.0700662104

57. J. Birch, S. Okasha, Kin selection and its critics, *BioScience*, **65** (2015), 22–32. https://doi.org/10.1093/biosci/biu196

58. J. B. S. Haldane, The causes of evolution, London: Longmans, Green, 1932. Available from: `http://archive.org/details/causesofevolutio00hald_0`

59. T. Dimitriu, D. Misevic, A. B. Lindner, F. Taddei, S. P. Brown, Bacteria can be selected to help beneficial plasmids spread, *PLOS Biol.*, **19** (2021), e3001489. https://doi.org/10.1371/journal.pbio.3001489

60. T. Stalder, L. M. Rogers, C. Renfrow, H. Yano, Z. Smith, E. M. Top, Emerging patterns of plasmid-host coevolution that stabilize antibiotic resistance, *Sci. Rep.*, **7** (2017), 4853. https://doi.org/10.1038/s41598-017-04662-0

61. T. Kiers, M. Duhamel, Y. Beesetty, J. Mensah, O. Franken, E. Verbruggen, et. al., Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis, *Science*, **333** (2011), 880–882. https://doi.org/10.1126/science.1208473

62. S. Hortal, K. L. Plett, J. M. Plett, T. Cresswell, M. Johansen, E. Pendall, et. al., Role of plant-fungal nutrient trading and host control in determining the competitive success of ectomycorrhizal fungi, *ISME J.*, **11** (2017), 2666–2676. https://doi.org/10.1038/ismej.2017.116

63. M. A. Nowak, K. Sigmund, Evolution of indirect reciprocity, *Nature*, **437** (2005), 1291–1298. https://doi.org/gdfdgrf

64. H. Ohtsuki, Y. Iwasa, How should we define goodness?–reputation dynamics in indirect reciprocity, *J. Theor. Biol.*, **231** (2004), 107–120. https://doi.org/10.1016/j.jtbi.2004.06.005

65. D. Sun, Pull in and push out: Mechanisms of horizontal gene transfer in bacteria, *Front. Microbiol.*, **9** (2018). https://doi.org/10.3389/fmicb.2018.02154

66. M. J. Wade, A critical review of the models of group selection, *Q Rev Biol*, **53** (1978), 101–114. https://www.journals.uchicago.edu/doi/abs/10.1086/410450

67. A. Traulsen, M. A. Nowak, Evolution of cooperation by multilevel selection, *Proc. Natl. Acad. Sci. U.S.A.*, **103** (2006), 10952–10955. https://doi.org/10.1073/pnas.0602530103

68. L. Altenberg, An evolutionary reduction principle for mutation rates at multiple loci, *Bull. Math. Biol.*, **73** (2011), 1227–1270. https://doi.org/10.1007/s11538-010-9557-9

69. F. M. Stewart, B. R. Levin, The population biology of bacterial plasmids: A priori conditions for the existence of conjugationally transmitted factors, *Genetics*, **87** (1977), 209–228. https://doi.org/10.1093/genetics/87.2.209

70. S. J. Tazzyman, S. Bonhoeffer, Fixation probability of mobile genetic elements such as plasmids, *Theor. Popul. Biol.*, **90** (2013), 49–55. https://doi.org/10.1016/j.tpb.2013.09.012

71. L. N. Lili, N. F. Britton, E. J. Feil, The persistence of parasitic plasmids, *Genetics*, **177** (2007), 399–405. https://doi.org/10.1534/genetics.107.077420

72. A. D. Halleran, E. Flores-Bautista, R. M. Murray, Quantitative characterization of random partitioning in the evolution of plasmid-encoded traits, 2019, (Unpublished work) http://resolver.caltech.edu/CaltechAUTHORS:20190402-080939441

73. N. v. d. Hoeven, A mathematical model for the co-existence of incompatible, conjugative plasmids in individual bacteria of a bacterial population, *J. Theor. Biol.*, **110** (1984), 411–423. https://doi.org/10.1016/s0022-5193(84)80183-6

74. A. Ilangovan, S. Connery, G. Waksman, Structural biology of the gram-negative bacterial conjugation systems, *Trends Microbiol.*, **23** (2015), 301–310. https://doi.org/10.1016/j.tim.2015.02.012

75. W. Lee, M. van Baalen, V. A. A. Jansen, Siderophore production and the evolution of investment in a public good: an adaptive dynamics approach to kin selection, *J. Theor. Biol.*, **388** (2016), 61–71. https://doi.org/10.1016/j.jtbi.2015.09.038

76. J. Smith, The social evolution of bacterial pathogenesis, *Proc. Royal Soc. B*, **268** (2001), 61–69. https://doi.org/10.1098/rspb.2000.1330

77. T. Nogueira, D. J. Rankin, M. Touchon, F. Taddei, S. P. Brown, E. P. C. Rocha, Horizontal gene transfer of the secretome drives the evolution of bacterial cooperation and virulence, *Curr. Biol.*, **19** (2009), 1683–1691. https://doi.org/10.1016/j.cub.2009.08.056

AIMS Press