The evolution of plasmid transfer rate in bacteria and its effect on plasmid persistence

Richard J. Sheppard

Timothy G. Barraclough

Vincent A. A. Jansen

1. Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot SL57PY, UK;
2. Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, UK;
3. School of Biological Sciences, Royal Holloway University of London, Egham, TW20 0EX, UK.

* Corresponding author; e-mail: richard.sheppard11@imperial.ac.uk.

Manuscript elements: Figure 1, figure 2, figure 3, table 1, table 2, online appendices A (including figure A1), B, C (including figure C1), D (including figure D1), E, F (including figure F1), G (including figure G1 and G2), H, I (including figures I1, I2, I3 and I4) and J.

Keywords: Horizontal gene transfer, conjugation, plasmid transfer rate, evolution.

Manuscript type: Article.

Prepared using the suggested \LaTeX template for Am. Nat.
Abstract

Plasmids are extrachromosomal segments of DNA that can transfer genes between bacterial cells. Many plasmid genes benefit bacteria but cause harm to human health by granting antibiotic resistance to pathogens. Transfer rate is a key parameter for predicting plasmid dynamics, but observed rates are highly variable and the effects of selective forces on their evolution are unclear. We apply evolutionary analysis to plasmid conjugation models to investigate selective pressures affecting plasmid transfer rate, emphasizing host versus plasmid control, the costs of plasmid transfer, and the role of recipient cells. Our analyses show that plasmid determined transfer rates can be predicted with three parameters (host growth rate, plasmid loss rate and the cost of plasmid transfer on growth) under some conditions. We also show that low frequency genetic variation in transfer rate can accumulate, facilitating rapid adaptation to changing conditions. Furthermore, reduced transfer rates due to host control have limited effects on plasmid prevalence until low enough to prevent plasmid persistence. These results provide a framework to predict plasmid transfer rate evolution in different environments and demonstrate the limited impact of host mechanisms to control the costs incurred when plasmids are present.
Introduction

Bacteria reproduce clonally but they possess a wide variety of mechanisms for transferring DNA between cells independently of reproduction (Ochman et al., 2000; Trevors, 1999). These mechanisms enable the transfer of genes that have a big impact on human populations, such as antibiotic resistance, detoxification, and pathogenesis genes (Bahl et al., 2009). Consequently, it is important to understand the mechanisms and dynamics of horizontal gene transfer within and between bacterial populations.

Of the three main mechanisms of horizontal transfer (transformation, conjugation and transduction), conjugation has attracted particular attention because of its impact on the spread of antibiotic resistance (Von Wintersdorff et al., 2016). Conjugation entails the transfer of segments of extrachromosomal DNA, called plasmids, that carry genes that benefit the host cell in some environments (e.g. antibiotic resistance genes, Bahl et al., 2009; Tazzyman and Bonhoeffer, 2015), while conferring costs to the host (e.g. through plasmid metabolism, replication, disruption, pilus formation and phage vulnerability, Lopatkin et al., 2017). These costs can lead to fitness conflicts between the host cell and the plasmid when in environments where the plasmid genes are not beneficial. Conjugative plasmids contain the genes necessary to initiate plasmid transfer (Frost and Koraimann, 2010; Smit et al., 1998) and are frequently described as self-replicating and transferring autonomously (del Solar et al., 1998; Turner et al., 2014) as selfish genetic elements (Rensing et al., 2002; Werren, 2011). However, host cells often also carry genes that affect replication and/or transfer of the plasmid (Frost et al., 2005). The multiplicity of factors affecting plasmid behaviour makes it hard to predict and manage the spread of plasmids carrying antibiotic resistance and other traits that affect human well-being. Questions such as what determines the frequency of plasmids in the population and how do costly plasmids persist when selective conditions are absent remain difficult to answer. The main complication is that fitness of a gene depends on whether it is transmitted by the host cell or by a selfish genetic element. To answer questions about how selection works on such genes we will therefore derive the fitness of such
genes from the perspective of a gene under control of the host, and of a gene under control of
the plasmid.

Modelling is a useful way to tease apart the effects of factors in complex systems and has been
used widely in plasmid biology, ever since classic work by Stewart and Levin (1977) established
the first mathematical model of plasmid transfer dynamics in a chemostat. They essentially
showed that, in theory, plasmids can persist when the transfer rate exceeds the effects of the cost
of the plasmid on growth and plasmid loss. Subsequent work has extended the model to tackle
a wide range of plasmid behaviour including plasmid mobilisation (Levin, 1980; Levin and Rice,
1980), partition and post-segregational killing systems (Hsu and Waltman, 1997; Lauffenburger,
1985; Mongold, 1992; Rankin et al., 2012; Seo and Bailey, 1985), copy number (Paulsson and
Ehrenberg, 1998), transitory derepression (Fernandez-Lopez et al., 2014; Lundquist and Levin,
1986; Raz and Tannenbaum, 2014) and social goods plasmids (Dimitriu, 2014; Dimitriu et al.,
2014; Mc Ginty et al., 2013, 2011), among many others. Some of these include estimates of
experimentally measured key parameters to demonstrate biological realism (Lauffenburger, 1985;
Levin et al., 1979; Lopatkin et al., 2017), with applications in a variety of environments (e.g.
agricultural waste, broilers, chemostats, Baker et al., 2016; Fischer et al., 2014; Svara and Rankin,
2011).

Although the work of Stewart and Levin (1977) identified the key parameters that shape
plasmid dynamics, a key open question is what determines the value of fundamental traits such
as transfer rate in the first place. In particular, for traits that are encoded by plasmid and/or host
genetics, what causes particular values to evolve? Recent models have explored the evolution
of plasmid cost and demonstrated its amelioration in theory and practice (Harrison et al., 2016;
Loftie-Eaton et al., 2017, 2016; Porse et al., 2016; Zwanzig et al., 2019). Other traits require the
same degree of attention, in particular transfer rate.

The evolution of transfer rates was first explored in the context of host-parasite modelling
(Cressler et al., 2016; Kribs-Zaleta, 2014; Lipsitch and Levin, 1997; Lipsitch et al., 1996; Misevic
et al., 2013) which is appropriate when plasmids are viewed as infectious replicators, exerting a
cost on their host cells. These models identified conditions favouring vertical versus horizontal transmission (e.g. low transfer rate plasmids are favoured in high growth rate cells) but lacked detailed application to plasmid systems, especially the mutual benefits that arise in selective conditions. While later evolutionary models focused more explicitly on plasmid transfer (Atsmon-Raz et al., 2015; Dimitriu et al., 2016; Haft et al., 2009; Hall et al., 2017; Porse et al., 2016; Raz and Tannenbaum, 2014; Turner et al., 1998), they did not provide a general mathematical analysis of parameter evolution or solutions for plasmid transfer rate following selection. Similarly, empirical studies report variation in transfer rates over many orders of magnitude ($10^{-20}$-$10^{-6}$), and identify statistical correlates of high versus low transfer rate (e.g. derepression, media type, host differences), but mechanistic explanations for why transfer rates vary remain scarce (Sheppard et al., 2020).

Theoretical explorations into the evolutionary forces affecting plasmid transfer are challenging due to the aforementioned plasmid-host mutualisms (Carroll and Wong, 2018; Dimitriu et al., 2016) and conflicts in selective and non-selective conditions (Kottara et al., 2018). These forces can separately affect transfer genes found on the plasmid, donor and recipient that collectively determine the observed transfer rates (Kozlowicz et al., 2006; McAnulla et al., 2007). In particular, there is empirical evidence that cells receiving the plasmids affect the rate of transfer (i.e. different recipient strains cause different transfer rates, Reniero et al., 1992; Sansonetti et al., 1980), but this has rarely been explored theoretically.

Analysis of transfer rate evolution also needs to evaluate all possible sources of costs on the host. Traditionally, models have considered costs of carrying a plasmid that were independent of the transfer rate (Harrison et al., 2016), but there are multiple ways that plasmids can incur a cost on the host cells. For example, plasmids can confer metabolic costs to the host due to plasmid replication and gene expression, including the expression of plasmid transfer and pilus production genes, and additional costs through greater susceptibility to infection by phage (Dimitriu et al., 2019; Jalasvuori et al., 2011; Porse et al., 2016; Reinhard et al., 2013; San Millan and MacLean, 2017; Turner et al., 1998). Some plasmids, however, do not demonstrate the same neg-
ative fitness effects often caused by high rates of transfer (Shapiro and Turner, 2014; Turner et al., 2014). For investigating the control of transfer rate, the key theoretical distinction is whether plasmid cost depends on the transfer rate or not – if it does, this could provide a check on the evolution of high transfer rates.

Here, we present new theory on the evolution of transfer rates that considers the perspective of the different partners in the system and a wider range of cost types that can arise. We explore how evolution shapes and determines plasmid transfer rates using adaptive dynamics theory, focussing on the costs of plasmid presence, transfer and host-plasmid conflicts. We also investigate how the selected rates of transfer and host-plasmid conflicts affect plasmid prevalence in bacterial populations. Published data and parameter estimates from the well-studied R1 plasmid in E. coli (Haft et al., 2009) are used as a case study to place model predictions in realistic parameter space and to estimate new parameters proposed here. We aimed to keep our model as simple as possible by building on earlier chemostat models, while including some variants to explore further complexity (e.g. superinfection, the ability of a plasmid-bearing donor to infect other donor cells, Smith, 2011). The potential impact of our model simplifications in relation to the incredible complexity of plasmid biology in nature are considered in the discussion.

Models and analyses

The model, based on the bacterial conjugation model by Stewart and Levin (1977), specifies the growth of an evenly mixed, homogeneous, single-species bacterial population with plasmid-free (density, \(N\)) and plasmid-bearing (density, \(N_P\)) cells (variables and parameters listed in table 1 and a schematic diagram given in Appendix A, Fig A1).

\[
\frac{dN}{dt} = kNS - DN - \gamma_NN_PN + \tau k_P N_P S
\]
\[
\frac{dN_P}{dt} = (1 - \tau)k_P N_P S - D N_P + \gamma_N N_P N
\]
\[
\frac{dS}{dt} = D(S_0 - S) - yk(N + N_P)S
\]  

(3)

Cells grow utilising a single substrate (concentration, \(S\)), which flows into the chemostat at a dilution rate (\(D\)), from a source with a constant concentration (\(S_0\)). Chemostat models assume a well mixed population and continuous inflow of resources/outflow of cells/resources, and are used here for ease of analysis, consistent with previous studies of bacterial dynamics (Stewart and Levin, 1977). We discuss the likely effects of alternative growth conditions on our results below. Plasmid free and plasmid bearing cells grow at rates of \(k\) and \(k_P\), respectively, which differ by the net cost or benefit given by the plasmid. We model the difference in two ways: as a flat cost/benefit of plasmid presence, due to metabolic activity, or a cost proportionate to the rate of transfer due to the costs of transfer (San Millan and MacLean, 2017, metabolic and transfer related costs are considered together in Appendix I). Cell growth is directly proportionate to substrate concentration for simplicity (Monod kinetics are also considered in Appendix I) scaled by a yield coefficient (\(y\)). Cells and substrate are lost from the chemostat through dilution at rate \(D\). Plasmids are lost from donor cells through unequal segregation during cell division at a constant rate (\(\tau\)), which is accordingly proportional to the growth of \(N_P\). Plasmid transfer (\(\gamma N\)) occurs at a rate proportional to the mass action product of the interaction of \(N\) and \(N_P\), with a coefficient assumed to be constant. This model deviates from Stewart and Levin (1977) in that donors and recipients have the same resource consumption rate (\(k\)), despite the different growth rates, and it is assumed that resources are reallocated away from growth to plasmid maintenance and transfer. We also model plasmid loss as a proportion of donor cell growth rather than as a proportion of donor cells. These changes are made to provide greater biological realism, emphasising the energy requirements for plasmid gene expression (San Millan and MacLean, 2017) and the mechanism of plasmid loss during cell division (Haft et al., 2009). Additional parameters and processes are further described in the specific models where they appear.

Using this population model to define our system, we then calculate the fitness of an entity with a mutant trait in terms of its ability to invade the equilibrated resident population when
initially rare. We use the rate at which rare mutants invade the resident population as the measure for fitness (Brännström et al., 2013; Metz et al., 1992), described mathematically by the sign of the differential equation showing the change in mutant population size with time. Specifically, we consider the fitness of invading mutant plasmids \( N_M \) and hosts \( M \), in turn, that alter the transfer rate \( \gamma_M \). Two versions were considered: in model 1, the plasmid has a constant cost or benefit to the host growth rate that is independent of the plasmid transfer rate (i.e. metabolic and replicative effects); whereas in model 2, the cost or benefit of the plasmid is proportional to plasmid transfer. When considering the invasion of a mutant host we also draw out the roles of the donor and recipient in the invasion. For each invasion scenario, the models were analysed mathematically by hand and using Mathematica (Wolfram Research Inc., 2020) to identify the parameter conditions that permit the invasion of the mutant from rare. We then bring together the results of the two models to investigate the impacts of host-plasmid conflicts on plasmid prevalence and persistence.

While the text is kept mathematically abstract for generality, results are plotted with estimated parameter values of the R1 conjugative plasmid in E. coli measured in lab experiments and reported in Haft et al. (2009), with recipient growth rate: \( k = 1.459 \text{ h}^{-1} \), loss rate: \( \tau = 10^{-4} \), substrate concentration: \( S_0 = 200 \mu\text{g ml}^{-1} \), and yield coefficient: \( y = 8 \cdot 10^{-8} \mu\text{g cell}^{-1} \). The data from this study were chosen because they include the majority of the required parameters for a single plasmid, including repressed and derepressed growth and transfer rates that enable the calculation and of metabolic and transfer costs. Dilution rate \( D \) was not available and was set to 0.001 ml h\(^{-1}\). This allowed us to visualise functions in a potentially realistic parameter space, increasing the application of the results.
Table 1: Model variables, key parameters and metrics. Values taken from Haft et al. (2009).

<table>
<thead>
<tr>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$ (\mu g , ml^{-1})</td>
<td>Substrate concentration.</td>
</tr>
<tr>
<td>$N, N_P, N_M$ (\text{cells} , ml^{-1})</td>
<td>Wild-type recipient and donor (wild-type plasmid: $P$, mutant plasmid: $M$) cell densities.</td>
</tr>
<tr>
<td>$M, M_P$ (\text{cells} , ml^{-1})</td>
<td>Mutant recipient and donor (wild-type plasmid: $P$) cell densities.</td>
</tr>
<tr>
<td>$S_0$ (200 , \mu g , ml^{-1})</td>
<td>Substrate stock concentration.</td>
</tr>
<tr>
<td>$D$ (0.001 , ml , h^{-1})</td>
<td>Dilution rate.</td>
</tr>
<tr>
<td>$k$ (1.459 , h^{-1})</td>
<td>Recipient growth rate.</td>
</tr>
<tr>
<td>$k_P$ (1.405 , h^{-1})</td>
<td>Donor growth rate (repressed).</td>
</tr>
<tr>
<td>$y$ (8 \cdot 10^{-8} , \mu g , cell^{-1})</td>
<td>Yield coefficient.</td>
</tr>
<tr>
<td>$\tau$ (10^{-4})</td>
<td>Plasmid loss rate.</td>
</tr>
<tr>
<td>$\gamma_{NN}, \gamma_{NN}$ (4.4 \cdot 10^{-12} , ml , cell^{-1} , h^{-1})</td>
<td>Wild-type plasmid transfer rate (repressed).</td>
</tr>
<tr>
<td>$k_M$ (h^{-1})</td>
<td>Mutant donor growth rate.</td>
</tr>
<tr>
<td>$\gamma_{MM}, \gamma_{MM}$ (ml , cell^{-1} , h^{-1})</td>
<td>Mutant plasmid transfer rate.</td>
</tr>
<tr>
<td>$\gamma_{NM}$ (ml , cell^{-1} , h^{-1})</td>
<td>Transfer rate from wild-type to mutant.</td>
</tr>
<tr>
<td>$\gamma_{MN}$ (ml , cell^{-1} , h^{-1})</td>
<td>Transfer rate from mutant to wild-type.</td>
</tr>
<tr>
<td>$b$ (4.61 \cdot 10^7 , cell , ml^{-1})</td>
<td>Transfer rate cost coefficient.</td>
</tr>
<tr>
<td>$\gamma_{sel}$ (ml , cell^{-1} , h^{-1})</td>
<td>Selected transfer rate.</td>
</tr>
<tr>
<td>$\gamma_{min}$ (ml , cell^{-1} , h^{-1})</td>
<td>Minimum transfer rate for plasmid persistence.</td>
</tr>
<tr>
<td>$\gamma_{90}$ (ml , cell^{-1} , h^{-1})</td>
<td>90% plasmid prevalence transfer rate.</td>
</tr>
<tr>
<td>$P_{Prev}$ %</td>
<td>Plasmid prevalence.</td>
</tr>
</tbody>
</table>
Model 1 - Plasmid costs independent of transfer rate

Mutant plasmid invasion ($N_M$)

We assume that each cell can only contain one plasmid, either the wild-type (at cell density $N_P$) or the mutant (at cell density $N_M$). Cells bearing the mutant plasmid are identical to cells that carry the wildtype, except in the rate of plasmid transfer ($\gamma_M$) and change density as follows:

$$\frac{dN_M}{dt} = (1 - \tau)k_P N_M S^* - DN_M + \gamma_M N_M N^*$$  \hspace{1cm} (4)

The fitness of the mutant plasmid ($W_{N_M}$) is equal to the per capita change in density $N_M$ from rare when the system has reached equilibrium. Rearranging (see Appendix B) gives the invasion success (fitness) of the plasmid:

$$W_{N_M} = \frac{1}{N_M} \frac{dN_M}{dt} = N^* (\gamma_M - \gamma_N)$$  \hspace{1cm} (5)

A mutant plasmid can invade therefore when $W_{N_M} > 0$, which for positive solutions requires $N^* > 0$ and $\gamma_M > \gamma_N$. The first condition indicates that in order for a mutant plasmid to invade the population must contain recipients. Plasmid transfer cannot occur in the absence of recipients, and therefore selective pressures do not affect the rate of transfer. The second condition states that the mutant can only invade when its transfer rate is higher than the wildtype. Evolutionary pressures on the plasmid will, therefore, select only for an increase in the rate of transfer, regardless of other circumstances.

Mutant host cell invasion ($M, M_P$)

Mutant host cells can be plasmid-free (density $M$) or plasmid-bearing (density $M_P$), and their dynamics are identical to $N$ and $N_P$ respectively, except for the rate of plasmid transfer ($\gamma_M$). The equations are also updated to emphasise the roles of donors and recipients in determining transfer (see Appendix C), where the first and second subscripts are the donor and recipient.
respectively \( \gamma_N = \gamma_{NN} \), \( \gamma_M = \gamma_{MM} \), and transfer from wild-type to mutant is included \( \gamma_{NM} \). Transfer from mutant to wild-type does not affect the mutant invasion under these assumptions and is not included. The resulting mutant growth equations are:

\[
\frac{dM}{dt} = kMS - DM - \gamma_{MM}MPM - \gamma_{NM}NPM + \tau k_PM_S \\
\frac{dMP}{dt} = (1 - \tau)k_PM_S - DMP + \gamma_{MM}MPM + \gamma_{NM}NPM
\] (6) (7)

The rarity of mutant interactions allows simplification and linearization of the system of equations allowing the calculation of the fitness equation (see Appendix C):

\[
W_M = (\gamma_{NM} - \gamma_{NN})N^*_P(kPS^* - D)
\] (8)

A mutant host cell can invade when two conditions are met: 1. \( N^*_P > 0 \), and 2. \( (\gamma_{NM} - \gamma_{NN})(kPS^* - D) > 0 \). The first condition requires the presence of plasmids for mutant host cells to invade, necessary for the evolution of plasmid transfer rate. Without plasmids, the mutant is indistinguishable from the wild-type, and the plasmid transfer rate is meaningless. The second condition requires that the product of two terms be positive, which can be achieved in two ways: 1. when \( \gamma_{NM} - \gamma_{NN} > 0 \) and \( kPS^* - D > 0 \), and 2. when \( \gamma_{NM} - \gamma_{NN} < 0 \) and \( kPS^* - D < 0 \).

The first term describes the relationship between the original transfer rate \( \gamma_{NN} \) and the rate of transfer from the wild-type to the mutant \( \gamma_{NM} \). \( \gamma_{NM} \) is determined by the genotype of \( N \) as a donor and \( M \) as a recipient, and mutants with higher or lower transfer rate can invade depending on whether \( kPS^* - D \) is positive or negative. Further analysis of the sign of the term \( kPS^* - D \) reveals identity with the sign of the effect (positive or negative) of the plasmid on the host’s growth (Appendix C, Fig. C1).

\[
\text{sign}(kPS^* - D) = \text{sign}(k - k)
\] (9)

This indicates that mutant cells with a transfer rate lower than the wild-type can invade when
$k PS^* - D$ is negative: when the plasmid is costly ($k_P < k$, Fig. 1A). In contrast, mutant cells with a transfer rate higher than the wild-type can invade when $k PS^* - D$ is positive: when the plasmid is beneficial ($k_P > k$, Fig. 1B). In summary, when the plasmid is beneficial the selection drives an increase in plasmid transfer rate, and when the plasmid is costly selection drives a decrease in the transfer rate by selecting for host cells with a higher/lower propensity of transfer.

![Figure 1: Mutant fitness ($W_M$) as mutant transfer rate changes ($\gamma_M$), where the plasmid is costly (A, $k_P = 1.4$ h$^{-1}$) and beneficial (B, $k_P = 0.15$ h$^{-1}$). Parameter assignments as default.](image)

Importantly, the change in transfer rate comes as a result of a mutation in $\gamma_{NM}$, not $\gamma_{MM}$ or $\gamma_{MN}$, i.e. where the mutant host cell is a recipient, not a donor. This makes biological sense: it will be unlikely for a rare mutant host cell to encounter a fellow mutant host cell in a well mixed environment. The vast majority of the mutant host encounters will be with wild type cells. From the point of view of a mutant host, the only interactions that matter here are the ones where a plasmid is received, because the plasmid costs and benefits are conferred in reception of a plasmid, not in donation. According to these model assumptions, plasmid donation does not affect the host fitness. Therefore the fitness is determined by $\gamma_{NM} - \gamma_{NN}$: the difference
between plasmid receipt in a rare mutant host compared to receipt in a wild type host in that same population. This indicates that the trait under selection is $\nu_M$, the reception propensity of the mutant recipient, which in this model is the mode by which adaptation of host-controlled transfer rate occurs. In practice, this adaptation requires variation in the recipient propensity and can only evolve when mechanistic and evolutionary constraints allow. Studies report variable transfer rates among recipients with the same donor (Sheppard et al., 2020), but it is unclear whether this is due to recipient control or donor discrimination, and whether this control is exerted by plasmid or host genes. Potential recipient control mechanisms can include mutations that affect mating complex formation (e.g. through spatial structure and cell aggregation, Reniero et al., 1992), plasmid removal following transfer (e.g. through restriction enzymes, Jiang et al., 2013; Levin, 2010) and through pheromones/quorum sensing (Chatterjee et al., 2013; Dunny, 2007). While donor effects are clearly substantial and important (De Gelder et al., 2005; Frost and Koraimann, 2010), our results suggest that care should be taken in interpreting donor effects on transfer rates as being controlled by the bacterial host and such interpretations should also explain the selective advantage of such control mechanisms. In other respects, the results of model 1 conform with the idea that high transfer rates are always favourable for the plasmid, irrespective of conditions, whereas a high transfer rate is only selected in hosts when plasmids are beneficial and plasmid-free recipients are present in the population. While the assumption that transfer is not costly for the donor cell is unrealistic (San Millan and MacLean, 2017), and is explored in model 2, these results reveal selection pressures affecting recipient cells that can lead to recipient-led control of plasmid transfer.

**Model 2 - Plasmid costs depend on transfer rate**

While donor and recipient growth rates were previously independent of other parameters, the growth rate is now assumed to decline with increasing transfer rate in proportion to a transfer-cost coefficient $b$. 


\[ k_P = k - b\gamma_N \]  \hspace{1cm} (10)

This means that plasmid cost is only determined through plasmid transfer and that plasmids cannot be beneficial under these assumptions as this would require a negative transfer rate. Because host and plasmid interests only differ when plasmids are costly, we focus on this scenario from now on. We do not explicitly model selection on the host transfer rate using model 2 because the result is too complex and does not simplify enough to be intuitive using these methods, but use an alternative way to explore host-plasmid conflict in the next section.

An empirical estimate of plasmid cost coefficient \( b \) was calculated using the transfer and growth parameters of repressed \((\gamma_N, k_P)\) and derepressed \((\gamma_{N_D}, k_{P_D})\) plasmids reported in Haft et al. (2009), assuming a linear relationship between donor growth rates and plasmid transfer rates. This estimate is used along with the other Haft et al. (2009) parameters to draw visual representations of the results in figures, and to provide a speculative parameter space to put those results into a meaningful context.

\[ b = \frac{k_P - k_{P_D}}{\gamma_{N_D} - \gamma_N} = \frac{1.405 - 1.23}{3.8 \cdot 10^{-9} - 4.4 \cdot 10^{-12}} = 4.61 \cdot 10^7 \text{cell ml}^{-1} \] \hspace{1cm} (11)

Analysis of the system of equations reveals only one real, non trivial solution (see Appendix D for details) that is used in the following sections, not given in this paper due to its length.

**Mutant plasmid invasion \((N_M)\)**

The equation for the invading mutant differing only in transfer rate \((\gamma_M)\) was modified as follows:

\[
\frac{dN_M}{dt} = (1 - \tau)k_M N_M S^* - DN_M + \gamma_M N_M N^* \] \hspace{1cm} (12)

where

\[ k_M = k - b\gamma_M \] \hspace{1cm} (13)
Rearranging (see Appendix E) gives the per capita fitness equation:

\[
W_{NM} = (\gamma_M - \gamma_N) \frac{kN^* - bD}{k_P}
\]

(14)

Mutant plasmids cannot invade when \( W_M \leq 0 \). This means that a mutant with a higher or lower transfer rate can invade when \( kN^* - bD \) is positive or negative, respectively (Fig. 2). \( k_P \) must always be positive for donor viability and biological realism. Selection therefore acts on transfer rate towards the point where the equilibrium density satisfies:

\[
N^* = \frac{bD}{k}
\]

(15)

(see Fig. 2).

Figure 2: Equilibrium density of recipients, \( N^* \), as a function of transfer rate, \( \gamma_N \), demonstrating the effects of selection on \( \gamma_N \). The figure identifies \( bD/k \) (the value of the equilibrium for \( N^* \) above which there is selection for larger \( \gamma_N \), and below which for smaller \( \gamma_N \)). Selection leads to \( \gamma_{sel} \), with arrows indicating the direction of selection.

We name the transfer rate at which this occurs \( \gamma_{sel} \). When this condition is used in conjunction with the real, non trivial solution previously established (Appendix D) we find:
\[
\gamma_{sel} = \frac{kt}{b} \left( \frac{kS_0(1 - \tau) - D}{kS_0(1 - \tau) - D(1 + yb)} \right)
\]  

(16)

Using the methods described by Geritz et al. (1998, see Appendix F) we find that \( \gamma_{sel} \) is always convergence stable, meaning that selection drives the transfer rate towards \( \gamma_{sel} \), and has neutral evolutionary stability. For the linear trade-off that we have used (Eq. 13), this means that once the plasmid population has attained transfer rate \( \gamma_{sel} \), mutants with different transfer rates will be able to reach appreciable frequencies by drift leading to the accumulation of diversity in the transfer rate. This is because selection pressures acting on plasmid transfer rate become weaker when the recipient density is in balance with the other parameters (when \( N^* \) approaches \( bD/k \), \( W_{NM} \) approaches 0). When mutant plasmids persist and increase in density through drift the recipient density changes (either increasing or decreasing) increasing the strength of selection on the plasmid population again. These results have implications for measuring transfer rates, as mutants with a wide range of transfer rates would be present at low frequency within the population and can constitute a reservoir of standing variation on which selection could quickly act in changing conditions.

The effect of each parameter on \( \gamma_{sel} \) is not immediately apparent due to the complexity of the equation (16). By varying the parameters over a realistic parameter space, around the Haft et al. (2009) data, we were able to identify two kind of behaviour (Appendix G, Fig. G1, G2). The first behaviour is when \( D(1 + yb) \) is small compared with \( kS_0(1 - \tau) \) that enables the simplification of \( \gamma_{sel} \) to

\[
\gamma_{sel} \approx \frac{kt}{b}
\]  

(17)

This simplification occurs when using the parameters given by Haft et al. (2009) and a low dilution rate as inputs. \( \gamma_{sel} \) therefore increases as the growth and loss rates increase and decrease as the transfer cost coefficient increases (Appendix G, Fig. G1). With the empirical estimates for \( \tau, k \) and \( b \) from Haft et al. (2009) the predicted value for \( \gamma_{sel} \) is \( 3.165 \cdot 10^{-12} \), which is remarkably close to the experimental estimate for repressed plasmid transfer rate of \( 4.4 \cdot 10^{-12} \).
The second behaviour occurs when the simplification cannot be made (i.e. when \( D, y, \) and \( b \) are high in reference to \( k, S_0 \) and \( \tau \), Appendix G, Fig. G2). In these conditions \( \gamma_{sel} \) increases exponentially as each parameter becomes less conducive to plasmid persistence, until the denominator approaches 0: \( kS_0(1 - \tau) = D(1 + yb) \). For some parameters, this reverses the relationship so that increasing \( b \) and decreasing \( k \) begin to increase \( \gamma_{sel} \) rather than decreasing \( \gamma_{sel} \) as found previously (compare Appendix G, Fig. G1 with G2, row 1, \( k \) and \( b \)).

When costs depend on transfer rate, the selection of transfer rate no longer drives it infinitely higher, but instead towards an intermediate value that balances transmission with the costs incurred on the host. We also see that the equation for this intermediate value simplifies under conditions of low dilution rate, among others.

**The effects of \( \gamma_N \) on plasmid prevalence**

We now investigate the effect of host-plasmid conflicts on plasmid prevalence in non-selective conditions. We previously found that selection acting on the host drives transfer rate down in non-selective conditions until the plasmid goes extinct (model 1), while selection acting on the plasmid drives the plasmid transfer rate to \( \gamma_{sel} \) when the cost is dependent on transfer rate (model 2). Assuming that both these selection pressures are at play, it is reasonable to expect that the combined selection pressures result in a value of \( \gamma_N \) that is between the transfer rate resulting in plasmid extinction (i.e. the value that should be optimum from the host’s perspective, which we name \( \gamma_{min} \)) and \( \gamma_{sel} \) due to host-plasmid conflict. We therefore investigate the effects of transfer rates on plasmid prevalence in this range. Specifically, we ask how much would the host need to reduce transfer rate in order to impact the plasmid prevalence in the population?

We start by using model 2 as a foundation for calculating plasmid prevalence \( (P_{\text{Prev}}, \text{calculated in Mathematica, Wolfram Research Inc., 2020}) \) and \( \gamma_{min} \) (see Appendix H for details):

\[
P_{\text{Prev}} = \frac{N_p^*}{N^* + N_p^*} = 1 - \frac{bDy}{kS_0(1 - \tau) - D} \quad \gamma_{min} = \frac{kDy\tau}{kS_0 - D(yb(1 - \tau) + 1)}
\]  

(18)
Similarly to $\gamma_{sel}$, these equations exhibit two kinds of behaviour as parameters vary (Appendix G, Fig. G1, G2). The first behaviour occurs when some parameters ($D, y, b$) are low compared with others ($k, S_0$) allowing simplification to:

$$P_{prev} \approx 1 \quad \gamma_{min} \approx \frac{Dy\tau}{S_0}$$  \hspace{1cm} (19)

We immediately see that plasmid prevalence is likely to be at saturation ($\approx 1$) in these simplifying conditions (i.e. when $bDy$ is small compared with $kS_0(1 - \tau) - D$). Also, we see the simplified $\gamma_{min}$ increase as yield coefficient, dilution and loss rate increase, and as substrate concentration decreases (Appendix G, Fig. G1). Although there is a slight variation between the points at which these three metrics simplify, the main factors which allow these simplifications are low dilution, yield and plasmid cost, relative to higher bacterial growth and substrate concentration. Under appropriate conditions (e.g. low dilution rate, chemostat environment) these simplifications may provide a quick and easy way to estimate the expected plasmid transfer rate when the plasmid is in control of transfer, and by how much that exceeds the minimum rate required for plasmid persistence.

When the simplification cannot be made $\gamma_{min}$ increases exponentially, similarly to $\gamma_{sel}$, although reaching a slightly different limit: $kS_0 = D(yb(1 - \tau) + 1)$. The difference between the limits of each metric mean that the values of $\gamma_{sel}$ and $\gamma_{min}$ diverge as the system moves into the exponential part where they may have previously been converging (Appendix G, Fig. G2). A small change in any parameter can result in a large difference in all three metrics when in this exponential part, limiting the precision of any applications of the model results. The exponential increase is where we also observe substantial decrease in plasmid prevalence over a relatively small parameter space (see Appendix G, Fig. G2, row 4: $\tau, b, D, y$), indicating a narrow region where sub-saturation levels of plasmid prevalence can be found when the system is at equilibrium.

To investigate the region where plasmid prevalence undergoes this sharp decreases, we now introduce a fourth metric: $\gamma_{90}$ to indicate the transfer rate at 90% plasmid prevalence and which
can be compared with the other metrics. First, $\gamma_{90}$, $\gamma_{sel}$ and $\gamma_{min}$ were plotted over an increasing dilution rate to map the change from simple to complex exponential behaviour (Fig. 3). $\gamma_{90}$ was about an order of magnitude higher than $\gamma_{min}$, and the two metrics increased proportionately with each other for the majority of the parameter space investigated. $\gamma_{sel}$ was largely unaffected by changes in $D$ (as expected from Eq. 17), which means that the difference between $\gamma_{sel}$ and $\gamma_{90}$ can potentially be several orders of magnitude dependent on the specific parameter values. This prospective difference between $\gamma_{sel}$ and $\gamma_{90}$ indicates that the host may need to reduce transfer rate considerably to affect plasmid prevalence substantially. When the metrics reach the exponentially increasing part of their curves, $\gamma_{90}$ increases above $\gamma_{sel}$, indicating that a transfer rate of $\gamma_{sel}$ produces an intermediate plasmid prevalence ($< 90\%$), suggesting that intermediate plasmid prevalence is only achieved when the transfer rate is within an order of magnitude of $\gamma_{min}$.

![Figure 3: Variation in the plasmid selected transfer rate ($\gamma_{sel}$), the host selected transfer rate ($\gamma_{min}$), and the transfer rate resulting in 90% prevalence for the plasmid ($\gamma_{90}$) as dilution rate increases. Parameter assignments as default except $b = 4.61 \cdot 10^8$ cell ml$^{-1}$, $y = 8 \cdot 10^{-6}$ µg cell$^{-1}$.](image)

We then compared $\gamma_{90}$ with $\gamma_{sel}$ and $\gamma_{min}$ across a range of reasonable parameter combina-
tions (see Table 2 legend for details). The dilution rate was set at $10^{-6}$ ml h$^{-1}$ to enable metric simplifications to occur across the majority of the parameter space. Results which did not include a positive $\gamma_{sel}$ or where $\gamma_{sel} < \gamma_{min}$, and results where the algorithm used to estimate $\gamma_{90}$ failed due to numerical rounding errors were filtered out.

Table 2: The average differences between the plasmid selected rate of transfer ($\gamma_{sel}$) and the transfer rate that determines 90% plasmid prevalence ($\gamma_{90}$), and between $\gamma_{90}$ and the host selected rate of transfer ($\gamma_{min}$) with standard deviations.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{sel}-\gamma_{90}$</td>
<td>5.186</td>
<td>2.570</td>
<td>-1.127</td>
<td>12.601</td>
</tr>
<tr>
<td>$\gamma_{90}-\gamma_{min}$</td>
<td>1.001</td>
<td>0.007</td>
<td>0.995</td>
<td>1.232</td>
</tr>
</tbody>
</table>

Note: Calculations were made over combinations of a range of parameters at increments of an order of magnitude unless otherwise stated: growth rate ($k$: $2 \times 10^{-4}$ to 2), plasmid loss rate ($\tau$, $10^{-7}$ to $10^{-1}$, increments of 2 orders of magnitude), plasmid cost coefficient ($b$, $10^5$ to $10^9$), stock substrate concentration ($S_0$, 0.2 to 200, increments of half an order of magnitude), yield coefficient ($y$, $10^{-10}$ to $10^{-6}$). Dilution rate ($D$) set to $10^{-6}$ to enable simplifications to occur.

Regardless of parameter combinations, there was consistently a difference of about one order of magnitude between $\gamma_{min}$ and $\gamma_{90}$, contrasting with the several orders of magnitude between $\gamma_{90}$ and $\gamma_{sel}$ (Table 2). The consistent difference suggests that, similarly to $\gamma_{min}$, $\gamma_{90}$ is approximately proportional to $\tau Dy / S_0$, adjusted by a factor of 10. The relatively small difference further means that when there is a large difference between $\gamma_{sel}$ and $\gamma_{min}$ the host would need to reduce the rate of transfer by several orders of magnitude before plasmid prevalence is substantially affected. For some parameter combinations $\gamma_{sel}$ was lower than $\gamma_{90}$, indicating intermediate plas-
mid prevalence at $\gamma_{sel}$ due to the difference of less than an order of magnitude between $\gamma_{sel}$ and $\gamma_{min}$. In summary, the results show that a decrease in transfer rate due to host-plasmid conflicts will result in little change in plasmid prevalence until the rate of transfer is within an order of magnitude of the minimum transfer rate. In many cases, the host would need to decrease the rate of transfer substantially before plasmid prevalence is affected.

Variations on the model results

Three further variants of the model were used to investigate plasmid metabolic and transfer costs simultaneously, and the effect of a more complex (Monod) growth function (Appendix I). The plasmid metabolic cost ($a$) was estimated at $0.054 \text{ h}^{-1}$ using the Haft et al. (2009) data, smaller than the costs incurred from derepressed rates of transfer (estimated at $0.175 \text{ h}^{-1}$), and far higher than transfer costs at repressed transfer rates ($0.0002 \text{ h}^{-1}$). These variations did not qualitatively alter the results. In addition, a version of the model was created in which a plasmid can transfer into a cell already containing a plasmid (superinfection). The results from this version were found to be broadly consistent with the original results, with only minor changes to fitness equations or with the changes cancelling each other out (Appendix J).

Discussion

To investigate the evolutionary pressures that affect plasmid transfer rate we constructed two mathematical models that were analysed using adaptive dynamics and invasion analyses. We calculated the fitness for plasmid and host genes that control transfer rate to demonstrate the potential for host-plasmid conflict to emerge over transfer rate. Our results show how transfer rate might be considered as an emergent property of underlying plasmid-, host-, and recipient-encoded traits.

We considered models with different cost functions for plasmids transfer. The models gave very different results, highlighting the importance of the plasmid cost function for realistic evo-
olutionary outcomes (model 1: fixed replicative/metabolic plasmid cost, model 2: transfer rate dependent plasmid cost). Model 1 predicts that plasmid-controlled transfer always selects for a higher transfer rate, as long as recipients are present, whereas host-controlled transfer can select for either increasing or decreasing transfer rate, leading to host-plasmid conflicts when the plasmid gives a net cost. Analysis showed that selection on host-controlled transfer rate acts on recipient propensity to receive plasmids, rather than on donor propensity. These findings are limited by the assumptions of the model that plasmid cost is not linked with transfer rate, and thus conflict with the results of model 2 that include the costs of plasmid donation. Analyses of model 2 show that plasmid-controlled transfer rate converges to an equilibrium that is convergence stable but has neutral ESS-stability, enabling standing variation in transfer rates at low frequencies. Transfer rates between this selected rate of transfer and transfer rates that cause plasmid extinction were explored to investigate the effects of host-plasmid conflicts in non-selective conditions and showed that the host must evolve mechanisms that reduce the transfer rate substantially in order to reduce plasmid prevalence. We now discuss the implications of these results in turn for interpreting plasmid dynamics in nature.

**Predicting transfer rates assuming plasmid control**

The prediction of the first model that selection always drives a higher plasmid transfer rate when the plasmid is in control is inconsistent with the literature, where low plasmid transfer rates are frequently observed experimentally (Sheppard et al., 2020) and plasmids often contain transfer-repression genes (Fernandez-Lopez et al., 2014; Haft et al., 2009; Lundquist and Levin, 1986). This inconsistency is due to the absence of transfer-dependent plasmid cost in the initial model. The addition of variable plasmid costs connected to the transfer rate (shown in model 2) demonstrate that selection on plasmid-controlled transfer rate finds an equilibrium. Equilibria balancing transmission and cost are well-known in parasite-host modelling (Cressler et al., 2016; Gibson et al., 2015; Lipsitch et al., 1996; Magalon et al., 2010; Turner et al., 1998), and the results of our models demonstrate the importance of accurate representation of plasmid cost relationships.
An interesting result from model 2 is that the selected equilibrium transfer rate has neutral ESS stability, enabling standing variation in transfer rates to persist at low frequency. If true, this variation would provide a resource for selection to act on under changing conditions, increasing the potential for rapid plasmid adaptation. The variation in transfer rates might also complicate experiments attempting to measure plasmid transfer rate, where a single clone may not be representative of the dominant behaviour. Other models have identified conditions where multiple plasmids with different transfer rates may invade and coexist based on the trade-offs in transfer rate and plasmid cost (when a low transfer, low cost plasmid prevents invasion of an incompatible high-transfer, high-cost plasmid in some cells, Lipsitch and Levin, 1997; van den Bosch et al., 2010; van der Hoeven, 1984; Van der Hoeven, 1986).

The simplified equation for the selected transfer rate ($\gamma_{\text{sel}}$) in model 2 shows that plasmid-controlled transfer rate could be predictable from only three parameters (bacterial growth rate, plasmid loss rate, plasmid transfer cost coefficient), approximately predicting the experimentally measured repressed transfer rate from the Haft et al. (2009) data. While the equation could provide a useful and simple metric to predict the plasmid-controlled selected rate of transfer, the conditions where this metric would be directly applicable are extremely limited, requiring the coevolution of host and plasmid in a chemostat system where the plasmid has complete control over its transfer. Chemostat models assume that there is constant inflow of resources and outflow of resources/waste and cells and therefore only apply directly to industrial systems where this is true. These conditions also favour fast cell growth at low substrate concentrations and keep cells in active growth phase, contrasting with batch-transfer models that assume periodic arrival of resources followed by dilution, selecting for cells that can survive lean periods and dispersal. Furthermore, due to the complexity of bacterial communities, the long-term evolutionary pressures become disrupted as the plasmid invades other hosts. While direct applications of the model results may be limited, the principles may be broadly applied to similar natural systems with flowing resources (e.g. rivers, the animal gut), although the extent to which this can be
done is unknown and requires further investigation, including appropriate parameter sets and results. While plenty of data exists for some of the parameters (e.g. bacterial growth rate, transfer rate) there are few data for others such as transfer cost coefficient or plasmid loss rate, and even fewer examples where all required parameters are found for a single plasmid in an environment (Sheppard et al., 2020).

Considering each parameter in the equation for $\gamma_{sel}$ in turn, we predict that an increase in bacterial growth rate increases the selected rate of transfer when the system has a low dilution rate. This is contrary to some bacterial-plasmid/host-parasite models which find that high growth rates promote lower transfer/transmission and a reduction in virulence (Lipsitch et al., 1996; Magalon et al., 2010; Turner et al., 1998). These models show that an increase in host growth rate increases the parasite growth success from vertical transmission. The parasite must mediate between the potential benefits and costs of its horizontal transmission, which vary between parasite type and are also dependent on recipient density/opportunity for successful transmission (Turner et al., 1998). In our model, however, constant plasmid costs (and corresponding transfer rates) are proportionately higher for strains with a low growth rate and mean that, as growth rate decreases, selection pressures on the plasmid drive the transfer rate down to facilitate plasmid survival. Conversely, an increase in overall growth rate permits the maintenance of plasmids with higher costs with corresponding higher transfer rates. Furthermore, based on the adaptive dynamic modelling here, low cost-low transfer rate strategies, such as those seen in previous models, may not be evolutionarily stable and may still be susceptible to invasion by plasmids with higher transfer rates despite the increased host costs.

Our predictions also superficially contradict results from two experimental studies on host-parasite systems, which indicate that high host growth rates promote lower horizontal transmission and a reduction in virulence (Dusi et al., 2015; Magalon et al., 2010). These experiments, however, did not alter the intrinsic bacterial growth rates, instead providing increased opportunity for cell growth by increasing the dilution rate, which is another variable in our model. The experimental results are therefore consistent with our model, which predicts that the relationship
between host growth rate and $\gamma_{sel}$ reverses with increasing dilution rates. This reversal occurs as conditions become less favourable for plasmid maintenance (e.g. due to increased dilution rate, low host growth rate) and where the plasmid can only be maintained by an increasingly high $\gamma_{sel}$.

These predictions could help to understand how plasmid transfer and prevalence vary across environments. Bacterial growth rates are well described in the literature and easy to measure in the laboratory, but vary considerably with environmental conditions (e.g. substrate type, concentration, temperature, pH). While many lab strains have a high growth rate, the doubling time of strains in harsher environments can be as long as hundreds of hours in rivers (Hendricks, 1972), and hundreds of days in soil (Gibson et al., 2018; Harris and Paul, 1994). Our model predicts that plasmid transfer rates should vary systematically between these different environments. Future work could test these predictions and bring improved understanding for management of plasmid-mediated resistance across varied environments such as hospital outflows, agricultural waste and soil.

The next parameter is the rate of loss of plasmid during cell division. While loss rates are expected to vary over several orders of magnitude due to various mechanisms plasmids employ to improve fidelity (e.g. partition or post-segregational killing genes, Bahl et al., 2009), they are notoriously difficult to measure (Lau et al., 2013). While plasmid presence is easily observed through selective markers on the plasmid, it is much more difficult to detect the absence of the selective marker as the plasmid is lost. Also, as plasmids are lost, loss rate itself can be difficult to separate from differences in donor and recipient growth rate, and further confounded by retransfer. Some methods have been designed to overcome these problems, but tend to be labour intensive and may still over or underestimate rates of loss (Lau et al., 2013). It is also appropriate to consider the effects of plasmid copy number on the results, not included in the scope of this study. Increased copy number, while increasing the metabolic burden on the host, decreases the risk of plasmid loss and co-resident plasmid interactions may also affect transfer rates (Gama et al., 2017), and these factors must be included in future models to investigate these effects.
comprehensively. Reliability of these methods and measurements have yet to be adequately resolved, and potentially limit the use of plasmid loss rate in the estimation of selected transfer rate.

The cost of transfer emerges in our model as a key parameter, but it has been less well studied than other parameters. This work may be the first time that the cost of plasmid transfer has been estimated from real data, but it is integral to understanding transfer rate evolution. The coefficient is particularly difficult to measure because it requires multiple transfer rates per plasmid to estimate the effect of transfer on plasmid cost and subsequent growth. The plasmid data used in this study included transfer and growth rates from repressed and derepressed strains of the same plasmid, which enabled direct comparison (Haft et al., 2009), but this may not be as easy to establish in other plasmids. It is extremely unclear how the coefficient varies among plasmids, and between hosts, and any assumptions are at this point speculative, including the linear relationship we assumed between cost and transfer.

The other parameters (yield coefficient, dilution rate and substrate concentration) are also important to consider when the simplified prediction of plasmid-controlled transfer rate does not apply. The yield coefficient is likely to be high in the majority of cases, assuming that the Haft et al. (2009) estimations are representative. Substrate concentration and dilution are likely to be far more variable and environment-specific. For example, hospital and agricultural waste are likely to have high substrate concentration but could vary in dilution depending on how the waste is treated. Dilution rate can also vary in water and soil systems depending on the speed of the flow and soil composition. Fast-flowing rivers or waste treatment plants may facilitate selection for high rates of plasmid transfer, although a high enough dilution rate would completely prevent plasmid maintenance. In many static environments dilution does not occur, and dilution can be interpreted as a death rate, although cell death does not necessarily mean that the genetic material leaves the system. Genetic material can remain in the system following cell death, and may be adopted by surrounding cells in transformation, contributing to horizontal gene transfer, and complicating our results. Our simplified theory might help to predict the
consequences of different management actions of conditions on plasmid prevalence, although the complexity of systems not considered in our models make specific applications less likely.

Our models are based on single infections. We explored the effect of multiple infections in a variant of our model (Appendix J) and found that our results are robust under superinfection. Superinfection of donor cells leads to within host competition and through this to selection for large transfer rates, if the trait is controlled by the plasmid. When the trait is controlled by the host, superinfection makes no difference. This is in agreement with experimental results which in which superinfection increases within host competition (Smith, 2011). We did not explore the effect of co-infection which is much more challenging to model correctly (Alizon, 2013) and beyond the scope of this study. Coinfection can lead to an increased copy number and inclusive fitness effects. This would be an interesting avenue for subsequent studies that build on these results.

Estimating parameter values in both laboratory and field settings remains a challenging area. In practice, plasmids are frequently transferred into naïve strains preceding transfer experiments in the laboratory (Sheppard et al., 2020), limiting the adaptation of plasmid and host to each other and to the environment. Observed transfer rates may additionally be affected by exposure to selective agents that occurs in the construction of donor strains prior to mating experiments, or during the experiments themselves, potentially increasing the transfer rate in some plasmids (although this is not found consistently, Lopatkin et al., 2016). Many transfer rate estimations also assume even mixing of cells, which can be unrealistic due to cell aggregation that can alter outcomes. These aspects must be considered when comparing theoretical and empirical values.

Consequences of host-plasmid conflicts

The calculated rate of plasmid transfer ($\gamma_{\text{sel}}$) assumes that plasmids adapt to their environment and control the rate of plasmid transfer without the influence of host-control mechanisms, and could serve as a base-rate for quantification of host control effects in non-selective conditions. If the observed rate matches the calculated rate of plasmid-controlled transfer it would indicate
plasmid-control, whereas a lower observed rate would support adaptation of the host to reduce the rate of transfer.

Our model only considers the evolution of transfer rate. Several studies show that coadaptation of plasmid and host results in a reduction of cost of the plasmid on the host growth rate (Loftie-Eaton et al., 2016; San Millan, 2018; Zwanzig et al., 2019). An extension to model 2 (Appendix I) separates two kinds of plasmid cost upon which selection could act and estimates the effect of each from real data - namely costs that are independent or dependent on transfer rate in turn. The parameters we estimated show that the independent costs are low (0.054 h\(^{-1}\)) compared with the costs caused by derepressed rates of transfer, indicating that the majority of plasmid cost of high transferring plasmids can be ameliorated through the reduction of transfer rate. If these parameter estimates are representative across different plasmids, the reduction of transfer may therefore be the primary source of plasmid cost amelioration. Experimental studies showing an evolving transfer rate yield conflicting results. For example, some studies show a decrease in plasmid costs with little change in transfer rate, citing optimisation of the plasmid to its hosts and the resolution of inefficiencies as causes for cost reduction (Dahlberg and Chao, 2003; Loftie-Eaton et al., 2016). Other plasmids do not have a measurable cost (Fischer et al., 2014; Wein et al., 2019), and do not necessarily maintain themselves primarily through transfer, opting for strategies which improve vertical transmission above horizontal (Hall et al., 2017).

One interesting finding from model 1 was that, under host control, selection occurs on the plasmid reception propensity of the recipient, and this conveys an important fitness advantage. The literature tends to focus on plasmids or donors as infectious agents in control of plasmid transfer (Bergstrom et al., 2000; Hall et al., 2017; McAnulla et al., 2007), as they frequently contain genes for initiation and control of plasmid transfer (Frost and Koraimann, 2010; Smit et al., 1998). Our results demonstrate the selection pressures on recipients that could lead to the evolution of recipient control mechanisms. While some studies show that recipients affect plasmid transfer rate (Reniero et al., 1992), there are few examples in the literature that have established mechanisms of recipient control (Chatterjee et al., 2013; Sansonetti et al., 1980). Evolution of
reception propensity may be more difficult to observe due to the stronger impact of other pressures (e.g. reduction of donor propensity due to the costs of transfer) and the difficulties in the evolution of recipient entry exclusion mechanisms (Pérez-Mendoza and de la Cruz, 2009). The results also highlight that the presence of recipients is necessary for the evolution of transfer rate, at least in cases where plasmid transfer does not occur between cells already carrying a plasmid. While this is an important consideration due to selection against recipients in the presence of antibiotics, for instance, the complete removal of recipients is likely to be uncommon in natural environments. Many plasmids confer social goods (e.g. environmental detoxification) which enable heterogenous donor-recipient populations (Rankin et al., 2011). The presence of biofilms can also protect recipient cells from the effects of selective agents (Penesyan et al., 2015). Environments also tend not to exhibit selection homogeneously and antibiotics can be only partially pervasive (Bahl et al., 2009). Donor cells may also be able to infect other donors (superinfection, Smith, 2011), which provides an opportunity for transfer and subsequent selection on transfer rates (Gandon et al., 2002), although with little benefit to either donor or recipient (Smith, 2012) and can be prevented by entry exclusion mechanisms (Hülter et al., 2017).

**Predicting plasmid prevalence**

Under the majority of conditions that we considered, plasmid prevalence was found to be high when plasmids are in control of transfer. Reductions in transfer rate due to host-control are unlikely to reduce plasmid prevalence until the transfer rate is close to the minimum transfer rate for plasmid persistence, which may be several orders of magnitude below the plasmid selected transfer rate. A host may therefore need to reduce the transfer rate by several orders of magnitude to have a substantial effect on plasmid prevalence, and a corresponding effect of the burden of the plasmid on the population. Furthermore, there is only a very narrow window of parameters within which values of prevalence between extinction and near-saturation occur. This prompts the question of why intermediate plasmid prevalence is frequently observed experimentally (Fan et al., 2019; Fox et al., 2008; Kottara et al., 2018; Lilley and Bailey, 1997, 2002). Some intermediate
plasmid prevalences may be the result of the absence of adaptation in newly formed host-plasmid combinations, and further studies investigating the evolution of these parameters in controlled and natural environments would be valuable.

The models make several assumptions which limit their application in favour of simplicity. Many of these missing aspects can be considered in future models to improve realism. For example, biofilms shape plasmid-host dynamics through the population structures they facilitate, and specific modelling is needed to assess selection on plasmid transfer rate in these environments (Beaudoin et al., 1998; Merkey et al., 2011). Phages confer an additional facet of cost, because cells are particularly vulnerable to some phages during pilus production for transfer, and which could be explored in conjunction with the costs previous described (Dionisio et al., 2005; Harrison et al., 2015; Wan and Goddard, 2012). The evolution of rates of transfer in plasmid repression/derepression systems is particularly interesting and can lead to very divergent repressed and derepressed transfer rates several orders of magnitude apart (Haft et al., 2009). Other interesting aspects of transfer rate control include competence switches and specialisation. Some plasmids only become transfer competent when conditions are optimal (e.g. high density of recipients) while in other populations only a fraction of cells in a population become transfer competent. These mechanisms use quorum sensing to reduce the costs of transfer while maximising benefits (Koraimann and Wagner, 2014; McAnulla et al., 2007; Refardt and Rainey, 2010). The model could be extended in future to include and investigate these factors and their impact on the results.

**Statement of Authorship**

All authors were jointly involved in project conceptualization, methods development, model analysis and writing. Timothy Barraclough and Vincent Jansen co-supervised and were responsible for funding acquisition.


References Cited Only in the Online Enhancements