Effects of conjugative plasmids on the ecology and evolution of microbial communities

Olivia Kosterlitz

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2022

Reading Committee:
Benjamin Kerr, Chair
Harvey Bradshaw
Jennifer Nemhauser

Program Authorized to Offer Degree:
Biology

University of Washington

Abstract

Effects of conjugative plasmids on the ecology and evolution of microbial communities

Olivia Kosterlitz

Chair of the Supervisory Committee:
Benjamin Kerr
Department of Biology

CHAPTER 1. To increase our basic understanding of the ecology and evolution of conjugative plasmids, we need reliable estimates of their rate of transfer between bacterial cells. Current assays to measure transfer rate are based on deterministic modeling frameworks. However, some cell numbers in these assays can be very small, making estimates that rely on these numbers prone to noise. Here we take a different approach to estimate plasmid transfer rate, which explicitly embraces this noise. Inspired by the classic fluctuation analysis of Luria and Delbrück, our method is grounded in a stochastic modeling framework. In addition to capturing the random nature of plasmid conjugation, our new methodology, the Luria-Delbrück method ('LDM'), can be used on a diverse set of bacterial systems, including cases for which current approaches are inaccurate. A notable example involves plasmid transfer between different strains or species where the rate that one type of cell donates the plasmid is not equal to the rate at which the other cell type donates. Asymmetry in these rates has the potential to bias or constrain current transfer estimates, thereby limiting our capabilities for estimating transfer in microbial communities. In contrast, the LDM overcomes obstacles of traditional methods by avoiding restrictive assumptions about growth and transfer rates for each population within the assay. Using stochastic simulations and experiments, we show that the LDM has high accuracy and precision for estimation of transfer rates compared to the most widely used methods, which can produce estimates that differ from the LDM estimate by orders of magnitude.

CHAPTER 2. Genes that undergo horizontal gene transfer (HGT) evolve in dramatically different genomic backgrounds as they move between hosts, which is in stark contrast to genes that evolve under strict vertical inheritance. Given the ubiquity of HGT in microbial communities, it is notable that the effects of host-switching on gene evolution have been largely understudied. Here, we present a novel framework to examine the consequences of host switching on gene evolution depending on the existence and form of host-dependent mutational effects. We started exploring the effects of HGT on gene evolution by focusing on a well-known antibiotic resistance gene (encoding a beta-lactamase)

commonly encoded on conjugative plasmids found in Enterobacteriaceae pathogens. By reconstructing the resistance landscape for a small set of mutationally connected alleles in three species (*Escherichia coli, Salmonella enterica, and Klebsiella pneumoniae*), we uncovered that the landscape topography was overwhelmingly aligned with very low levels of host-dependent mutational effects. By simulating gene evolution with and without HGT using the species-specific empirical landscapes, we found that evolutionary outcomes were similar despite HGT. These findings suggest that mobile genes adapting in one species can lead to adaptation in another species. In such a case, vehicles of cross-species HGT enable a distributed form of genetic evolution across a bacterial community, where species can ‘crowdsource’ adaptation from other community members.

Table of Contents

List of figures

List of supplemental figures

List of tables

Acknowledgements

This doctoral thesis is the product of years of my life during which I was more devoted than I have ever been. This loyalty not only produced what I hope are useful contributions to our understanding of microbial population biology, but by maintaining steadfast to my scientific pursuits I grew substantially both professionally and personally. I have learned so many important lessons that I will take into my future endeavors. These lessons will not only help me continue to grow as a scientist but become a more responsible and engaged citizen of the world. I could not have learned these lessons without the amazing mentors, teachers, collaborators, and colleagues I met along the way.

First, I have enormous appreciation for my advisor Benjamin Kerr. Not only is he a brilliant scientist, but his training has shaped my problem-solving skills and scientific philosophy. I can't properly communicate how much his support and confidence has shaped my development, and how it was a light for me during the lowest points in my Ph.D. journey.

In addition, I have so much appreciation to the Kerr laboratory. I was very lucky to be able to join a kind, supportive, and dynamic laboratory. Though the members of the laboratory have changed over the last six years, it has remained a welcoming and accepting learning environment that has supported the various stages of my Ph.D. journey. Thank you to all those who have been a member. I have targeted shout outs for a few individuals that were instrumental in my training given in alphabetical order: Peter Conlin, Katie Dickinson, Reilly Falter, Nathan Grassi, Hannah Jordt, Ryan McGee, Sonia Singhal, Adamaris Muñiz Tirado, Claire Wate, and Bailey Werner.

Next, I must thank Professor Eva Top and the Top laboratory for making me feel like a member of their group, sharing their expertise in plasmid biology, discussing the many facets of my projects, and providing a friendly and supportive atmosphere.

In addition, I thank all my friends with an extra thanks to Abigail Ames, Terra Burton, Caroline Capello, Itzue Solis Caviedes, Katie Dickinson, Reilly Falter, Wes Geisenberger, Ryan Flesch, Hannah Jordt, Andre Kurlovs, Richard Lee, Ryan McGee, Heather Miller, Molly Phillips, Simon Snoeck, and Mo Turner. These individuals went above and beyond to support me during this journey. I will be forever grateful for the memories we have shared.

Lastly, I must thank my family without whom I would not be the person I am today. Especially to my father (Herbert Kosterlitz III), mother (Allison Kosterlitz), and three sisters (Jody, Cassie, and Lea Kosterlitz), who have seen my life at all stages. I am grateful for how much our relationships have grown and matured through the years. With family there is no rule book, but we all do the very best to take care of one another. Knowing that I will have the support of my family in perpetuity has been such a blessing in my life. This has brought me so much comfort during the most challenging periods of this journey.

Dedication

to my grandparents and parents

CHAPTER 1

Estimating the transfer rates of bacterial plasmids with an adapted Luria–Delbrück fluctuation analysis

Olivia Kosterlitz, Adamaris Muñiz Tirado, Claire Wate, Clint Elg, Ivana Bozic, Eva M. Top, Benjamin Kerr

To increase our basic understanding of the ecology and evolution of conjugative plasmids, we need reliable estimates of their rate of transfer between bacterial cells. Current assays to measure transfer rate are based on deterministic modeling frameworks. However, some cell numbers in these assays can be very small, making estimates that rely on these numbers prone to noise. Here we take a different approach to estimate plasmid transfer rate, which explicitly embraces this noise. Inspired by the classic fluctuation analysis of Luria and Delbrück, our method is grounded in a stochastic modeling framework. In addition to capturing the random nature of plasmid conjugation, our new methodology, the Luria-Delbrück method ('LDM'), can be used on a diverse set of bacterial systems, including cases for which current approaches are inaccurate. A notable example involves plasmid transfer between different strains or species where the rate that one type of cell donates the plasmid is not equal to the rate at which the other cell type donates. Asymmetry in these rates has the potential to bias or constrain current transfer estimates, thereby limiting our capabilities for estimating transfer in microbial communities. In contrast, the LDM overcomes obstacles of traditional methods by avoiding restrictive assumptions about growth and transfer rates for each population within the assay. Using stochastic simulations and experiments, we show that the LDM has high accuracy and precision for estimation of transfer rates compared to the most widely used methods, which can produce estimates that differ from the LDM estimate by orders of magnitude.

Introduction

A fundamental rule of heredity involves the passage of genes from parents to their offspring. Bacteria violate this rule of strict vertical inheritance by shuttling DNA between cells through horizontal gene transfer (1, 2). Often the genetic elements being shuttled are plasmids, extrachromosomal DNA molecules that can encode the machinery for their transfer (3). This plasmid transfer process is termed conjugation, in which a plasmid copy is moved from one cell to another upon direct contact. Additionally, plasmids replicate independently inside their host cell to produce multiple copies, which segregate into both offspring upon cell division. Therefore, conjugative plasmids are governed by two modes of inheritance: vertical and horizontal.

This horizontal mode of inheritance makes it possible for non-related cells to exchange genetic material, which includes members of different species (4). In fact, conjugation can occur across vast phylogenetic distances, such that the expansive gene repertoire in the "accessory" genome encoded on conjugative plasmids is shared among many microbial species (5). This ubiquitous genetic exchange reinforces the central role of conjugation in shaping the ecology and evolution of microbial communities (1, 3, 6). Notably, conjugation is a common mechanism facilitating the spread of antimicrobial resistance genes among bacteria and the emergence of multi-drug resistance in clinical pathogens (7–9). To understand how genes, including those of clinical relevance, move within complex bacterial communities, an accurate and precise measure of the rate of conjugation is of the utmost importance.

The basic approach to measure conjugation involves mixing plasmid-containing bacteria, called "donors", with plasmid-free bacteria, called "recipients". As the co-culture incubates, recipients acquire the plasmid from the donor through conjugation, and these transformed recipients are called "transconjugants". Over the course of this "mating assay," the densities of donors, recipients, and transconjugants are tracked over time (D_t, R_t, and T_t, respectively) as the processes of population growth and plasmid transfer occur. To understand how such information is used to calculate the rate of conjugation, we consider an altered version of the foundational Levin *et al*. model (10). In this framework, populations grow exponentially, and recipients become transconjugants via conjugation when they interact with plasmid-bearing cells (i.e., donors or transconjugants). The densities of the populations are described by the following differential equations (the t subscript is dropped from the variables for notational convenience):

$$\frac{dD}{dt} = \psi_D D, \quad [1]$$

$$\frac{dR}{dt} = \psi_R R - \gamma_D DR - \gamma_T TR, \quad [2]$$

$$\frac{dT}{dt} = \psi_T T + \gamma_D DR + \gamma_T TR. \quad [3]$$

In equations [1]-[3], donors, recipients and transconjugants divide at a per-capita rate of ψ_D, ψ_R, and ψ_T, respectively. The parameters γ_D and γ_T measure the rate at which a

recipient acquires a plasmid per unit density of the donor and transconjugant, respectively. Thus, the ψ parameters are population growth rates and the γ parameters are conjugation rates (see Figure 1a). Assuming all the growth rates are equal ($\psi_D = \psi_R = \psi_T = \psi$) and conjugation rates are equal ($\gamma_D = \gamma_T = \gamma$), Simonsen *et. al.* (11) provided an elegant solution to equations [1]-[3] to produce the following estimate for the conjugation rate from donors to recipients (hereafter termed the "donor conjugation rate"):

$$\gamma_D = \psi \ln\left(1 + \frac{T_{\tilde{t}}}{R_{\tilde{t}}}\frac{N_{\tilde{t}}}{D_{\tilde{t}}}\right)\frac{1}{(N_{\tilde{t}} - N_0)}. \quad [4]$$

For a mating assay incubated for a fixed period (hereafter $\tilde{t}$), the initial and final density of all bacteria (N_0 and $N_{\tilde{t}}$, respectively), the final density of each cell population ($D_{\tilde{t}}$, $R_{\tilde{t}}$, and $T_{\tilde{t}}$), and the population growth rate (ψ) are sufficient for an estimate of the conjugation rate.

The Achilles heel of this estimate, as with others, is found in violations of its assumptions. For instance, we label equation [4] as the "Simonsen *et. al.* Identicality Method" estimate (SIM) for the donor conjugation rate because the underlying model assumes all strains are *identical* with regards to growth rates and conjugation rates. However, in natural microbial communities, this identicality assumption is misplaced, especially when the donors and recipients belong to different species. For instance, suppose that the rate of plasmid transfer within a species (i.e., from transconjugants to recipients, which we abbreviate as the "transconjugant conjugation rate") is much higher than between species (i.e., from donors to recipients); that is, $\gamma_T \gg \gamma_D$ (Figure 1b). This elevated within-species conjugation rate (γ_T) will increase the number of transconjugants and consequently inflate the SIM estimate for the cross-species conjugation rate (γ_D) compared to a case where the conjugation rates are equal ($\gamma_T = \gamma_D$, Figure 1c). This Achilles heel is not specific to cross-species scenarios and can occur when estimating conjugation between any cells, including strains of the same species. One approach to minimize the resulting bias is to shorten the incubation time for the assay (12), as estimate bias tends to increase over time (e.g., Figure 1b). However, new problems can arise when using this approach, such as the transconjugant numbers becoming exceedingly low and thus difficult to accurately assess (13). Another approach was introduced by Huisman *et al*. (14), which squarely addressed the SIM identicality assumptions by developing a method to estimate donor conjugation rate when growth and transfer rates differ, thereby enlarging the set of systems amenable to estimation (see SI section 1 for full description of this and other approaches). Nonetheless, this new method can have difficulty with situations in which the donor conjugation rate (γ_D) is substantially lower than the transconjugant rate (γ_T), the example illustrated in Figure 1. Such differences have been reported in multi-species systems (15) and recently several studies have recognized the importance of evaluating the biology of plasmids in microbial communities (7, 16–18). Therefore, a method that provides an accurate estimate despite substantial inequalities in rate parameters is desirable.

Here we derive a novel estimate for conjugation rate by explicitly tracking transconjugant dynamics as a stochastic process (i.e., a continuous time branching

process). This represents a notable deviation from previous approaches that are built upon deterministic frameworks. The random nature of conjugation can lead to substantial variation in the number of transconjugants at the end of a mating assay ($T_{\tilde{t}}$) as this population will often be small. Prior deterministic frameworks rely on this number (e.g., equation [4]), such that transconjugant variation adds problematic noise to the estimate. In contrast, our stochastic approach leverages this noise to produce an estimate (akin to the way Luria and Delbrück estimated mutation rate (19)). In addition, our method allows for unrestricted heterogeneity in growth rates and conjugation rates. Thus, our method fills a gap in the methodological toolkit by allowing unbiased estimation of conjugation rates in a wide variety of strains and species. We used stochastic simulations to validate our estimate and compare its accuracy and precision to other estimates. We developed a protocol for the laboratory by using microtiter plates to rapidly screen many donor-recipient co-cultures for the existence of transconjugants. In addition to its experimental tractability, our protocol circumvents problems that arise in the laboratory that can bias other approaches. Finally, we implemented our method in the laboratory and compared our estimate to the SIM estimate using a *Klebsiella pneumoniae* to *Escherichia coli* cross-species case study with an IncF conjugative plasmid.

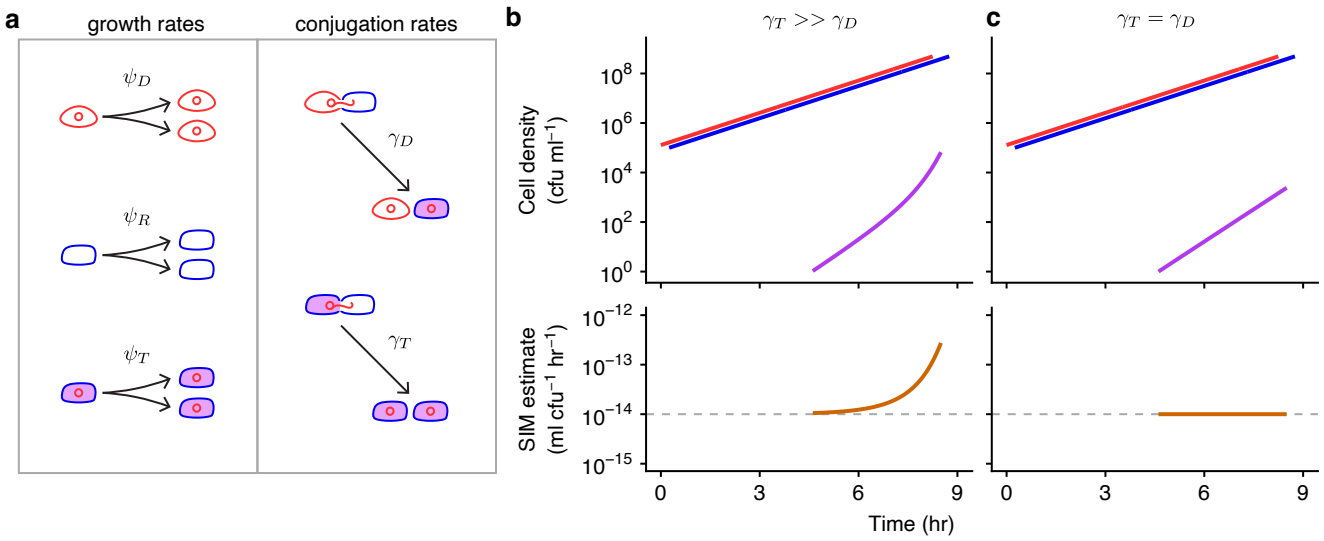


Figure 1 : Basic model parameters and the effects of unequal conjugation rates on the SIM estimate. (a) In this schematic, the conjugative plasmid is a red circle, a donor is a red cell containing the plasmid, a recipient is a blue cell, and a transconjugant is indicated with a purple interior (a blue cell containing a red plasmid). The ψ_D, ψ_R, and ψ_T parameters are donor, recipient, and transconjugant growth rates, respectively, illustrated by one cell dividing into two. The γ_D and γ_T parameters are donor and transconjugant conjugation rates, respectively, shown by conjugation events transforming recipients into transconjugants. (b) When the transconjugant conjugation rate (γ_T) is higher than the donor conjugation rate (γ_D), transconjugants exhibit super-exponential increase (purple curve) while donors and recipients increase exponentially (red and blue lines). The SIM estimate (orange line) increases over time, deviating from the actual donor conjugation rate (gray dashed line). (c) In contrast, when the conjugation rates are equal ($\gamma_T = \gamma_D$),

the transconjugant increase is muted relative to part b (purple line). The SIM assumptions are met, and the estimate is constant and accurate over time (orange line). Equations [1]-[3] were used to produce the top graphs, with $D_0 = R_0 = 10^5$, $T_0 = 0$, $\psi_D = \psi_R = \psi_T = 1$, $\gamma_D = 10^{-14}$, and either $\gamma_T = 10^{-8}$ (in part b) or $\gamma_T = 10^{-14}$ (in part c). The donor and recipient trajectories overlapped but were staggered for visibility. Equation [4] was used to produce the bottom graphs. The code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

Results

A new conjugation rate estimate inspired by the Luria–Delbrück approach.

Previous methods to estimate the rate of conjugation have treated the rise of transconjugants as a deterministic process (i.e., non-random). However, conjugation is inherently a stochastic (i.e., random) process (20). Given that conjugation transforms the genetic state of a cell, we can form an analogy with mutation, which is also a stochastic process that transforms the genetic state of a cell. While mutation transforms a wild-type cell to a mutant, conjugation transforms a recipient cell to a transconjugant.

This analogy inspired us to revisit the way Luria and Delbrück handled the mutational process in their classic paper on the nature of bacterial mutation (19), outlined in Figure 2a-d. For this process, assume that the number of wild-type cells, N_t, is expanding exponentially. Let the rate of mutant formation be given by μ. In Figure 2a, we see that the number of mutants in a growing population increases due to mutation events (highlighted purple cells) and due to faithful reproduction by mutants (non-highlighted purple cells). The rate at which mutants are generated (highlighted purple cells) is μN_t, which grows as the number of wild-type cells increase (Figure 2b). However, the rate of transformation per wild-type cell is the mutation rate μ, which is constant (Figure 2c). Since mutations are random, parallel cultures will vary in the number of mutants depending on if and when mutation events occur. As seen in Figure 2d, for sufficiently small wild-type populations growing over sufficiently small periods, some replicate populations will not contain any mutant cell (gray shading) while other populations exhibit mutants (purple shading). Indeed, the cross-replicate fluctuation in the number of mutants was a critical component of the Luria-Delbrück experiment.

To apply this strategy to estimate the conjugation rate, we can similarly think about an exponentially growing population of recipients (Figure 2e). But now there is another important cell population present (the donors). The transformation of a recipient is simply the generation of a transconjugant (highlighted purple cells) via conjugation with a donor. If we ignore conjugation from transconjugants for the moment, the rate at which transconjugants are generated is $\gamma_D D_t R_t$ (Figure 2f). In contrast to the mutation rate, the rate of transformation per cell is not a constant. Rather, this transformation rate per recipient is $\gamma_D D_t$, which grows with the donor population (Figure 2g). It is as if we are tracking a mutation process where the mutation rate is exponentially increasing. Yet the

rate of transformation per recipient *and* donor is constant, which is the donor conjugation rate γ_D. As with mutation, conjugation is random, which results in a distribution in the number of transconjugants among parallel cultures depending on the time points at which transconjugants arise. As seen in Figure 2h, under certain conditions, some replicate populations will not contain any transconjugant cell (gray shading) while other populations will exhibit transconjugants (purple shading).

Using this analogy, here we describe a new approach for estimating conjugation rate which embraces conjugation as a stochastic process (20). Let the density of donors, recipients, and transconjugants in a well-mixed culture at time t be given by the variables D_t, R_t, and T_t. In all that follows, we will assume that the culture is inoculated with donors and recipients, while transconjugants are initially absent (i.e., $D_0 > 0$, $R_0 > 0$, and $T_0 = 0$). The donor and recipient populations grow according to the following standard exponential growth equations

$$D_t = D_0 e^{\psi_D t}, \quad [5]$$

$$R_t = R_0 e^{\psi_R t}, \quad [6]$$

where ψ_D and ψ_R are the growth rates for donor and recipient cells, respectively. With equations [5] and [6], we are making a few assumptions, which also occur in some of the previous methods (SI Table 3). First, we assume the loss of recipient cells to transformation into transconjugants can be ignored. This assumption is acceptable because, for what follows, the rate of generation of transconjugants per recipient cell (as in Figure 2g, $\gamma_D D_t$) is very small relative to the per capita recipient growth rate (ψ_R). Second, we assume that donors and recipients exhibit deterministic exponential growth. If the initial numbers of donors and recipients are not too small (i.e., $D_0 \gg 0$ and $R_0 \gg 0$) and per capita growth remains constant over the period of interest, then this assumption is reasonable. We note that this assumption does not deny that cell division of donors and recipients are also stochastic processes, but given the large numbers of these cells, a deterministic approximation is appropriate.

On the other hand, the number of transconjugants over the period of interest can be quite small (starting from zero), motivating an explicit stochastic treatment (21). The population growth of transconjugants is modeled using a continuous-time stochastic process. The number of transconjugants, T_t, is a random variable taking on non-negative integer values. In this section, we will assume the culture volume is 1 ml and thus the number of transconjugants is equivalent to the density of transconjugants (per ml). For a very small interval of time, Δt, the current number of transconjugants will either increase by one or remain constant. The probabilities of each possibility are given as follows:

$$\Pr\{T_{t+\Delta t} = T_t + 1\} = \gamma_D D_t R_t \Delta t + \gamma_T T_t R_t \Delta t + \psi_T T_t \Delta t, \quad [7]$$

$$\Pr\{T_{t+\Delta t} = T_t\} = 1 - (\gamma_D D_t R_t + \gamma_T T_t R_t + \psi_T T_t)\Delta t. \quad [8]$$

The three terms on the right-hand side of equation [7] illustrate the processes enabling the transconjugant population to increase. The first term gives the probability that a donor transforms a recipient into a transconjugant via conjugation. The second term gives the probability that a transconjugant transforms a recipient via conjugation. The third term measures the probability that a transconjugant cell divides. Equation [8] is simply the probability that none of these three processes occur.

Given the standard set-up of a mating assay, we focus on a situation where there are initially no transconjugants. Therefore, the only process that can produce the first transconjugant is conjugation of the plasmid from a donor to a recipient. Using equation [8] with $T_t = 0$, we have

$$\Pr\{T_{t+\Delta t} = 0 \mid T_t = 0\} = 1 - \gamma_D D_t R_t \Delta t. \quad [9]$$

We let the probability that we have zero transconjugants at time t be denoted by $p_0(t)$ (i.e., $p_0(t) = \Pr\{ T_t = 0\}$). In SI section 2, we derive the following expression for $p_0(t)$ at time $t = \tilde{t}$:

$$p_0(\tilde{t}) = \exp\left\{\frac{-\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\}. \quad [10]$$

Solving equation [10] for γ_D yields a new measure for the donor conjugation rate:

$$\gamma_D = -\ln p_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right). \quad [11]$$

This expression is similar in form to the mutation rate derived by Luria and Delbrück in their classic paper on the nature of bacterial mutation (19), which is not a coincidence.

In SI section 3, we rederive the Luria-Delbrück result, which can be expressed as

$$\mu = -\ln p_0(\tilde{t})\left(\frac{\psi_N}{N_0\left(e^{\psi_N\tilde{t}} - 1\right)}\right). \quad [12]$$

In the mutational process modeled by Luria and Delbrück, N_0 is the initial wild-type population size, which grows exponentially at rate ψ_N. For Luria and Delbrück, $p_0(\tilde{t})$ refers to the probability of zero mutants at time $\tilde{t}$ (as in a gray-shaded tree in Figure 2d), whereas $p_0(\tilde{t})$ in the conjugation estimate refers to the probability of zero transconjugants (as in a gray-shaded tree in Figure 2h). Comparing equation [12] to equation [11], conjugation can be thought of as a mutation process with initial wild-type population size $D_0 R_0$ that grows at rate $\psi_D + \psi_R$. The structural similarity of the estimates is grounded in a structural similarity of the underlying models; indeed, some of the same assumptions that apply to the mutation process modeled by Luria and Delbrück also apply to the conjugation process modeled here. For instance, the loss of recipients due to plasmid transfer is ignored in the recipient dynamics of the conjugation model (equation [6]) in the same way that the loss of wildtype cells due to mutation is ignored in the wildtype cell dynamics of the mutation model (equation [3.2] in SI section 3), which tends to be a safe assumption when growth greatly outpaces transformation. Furthermore, in the same way that reversions (mutations restoring a wildtype genotype from a mutant) are disregarded in the mutational model, we ignore the possibility of transconjugants (and donors) becoming plasmid-free through segregational loss in the conjugation model. Lastly, as in the original

Luria-Delbrück model, we focus on a pure “birth” process (e.g., once mutants or transconjugants are generated, their numbers do not decrease). In our supplemental sections we explore the impacts of violations to some of these assumptions (e.g., the negligible impact of segregational loss in SI section 4d, and how to correct for an effective loss in transconjugant cell numbers due to a failure of small numbers of transconjugants to establish under experimental conditions in SI section 6d and 7). Given the connections between modeling frameworks and estimate structures, we label the expression in equation [11] as the LDM estimate for donor conjugation rate, where LDM stands for “Luria-Delbrück Method.”

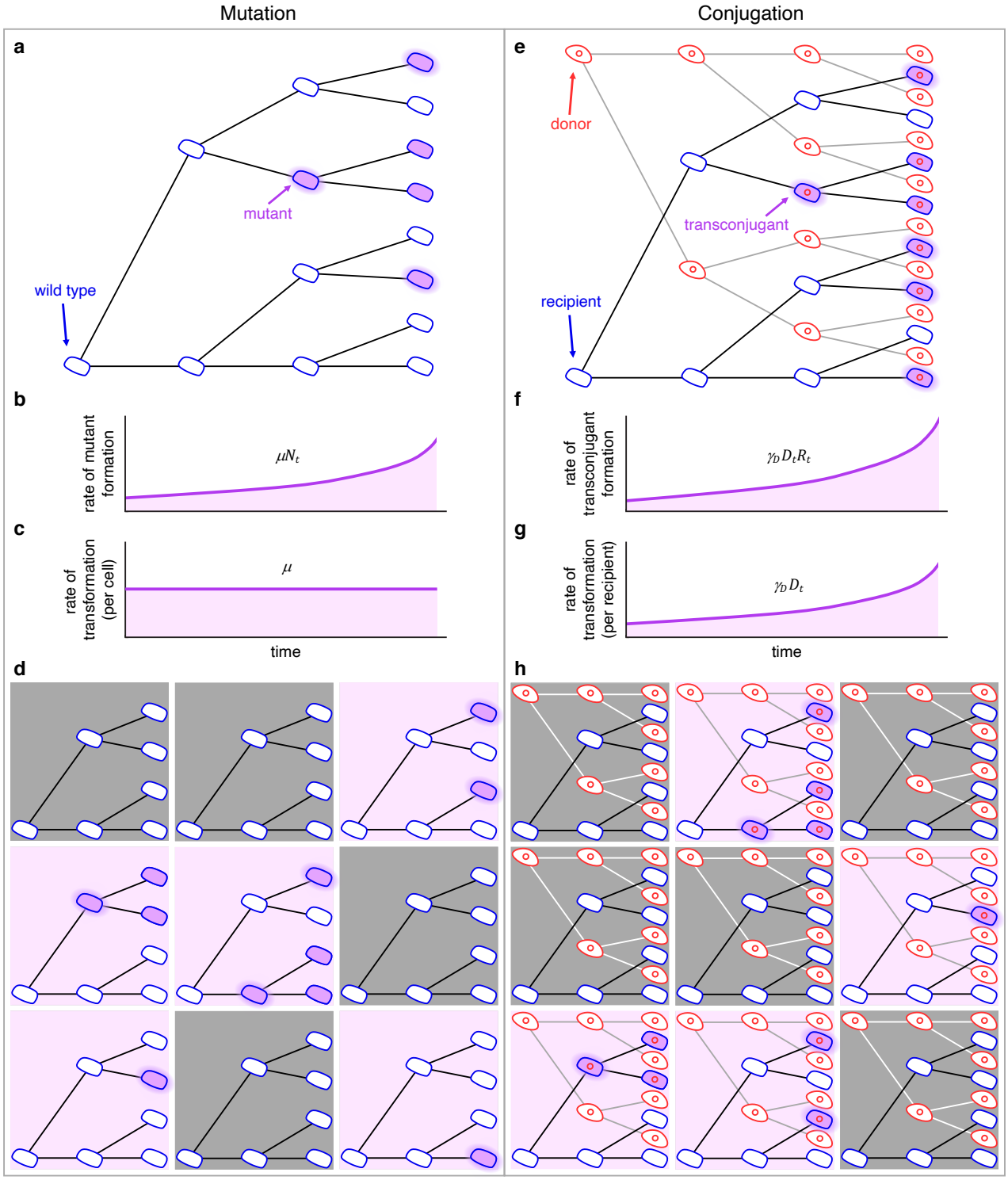


Figure 2 : Schematic comparing the process of mutation (a-d) to the process of conjugation (e-h). (a) In a growing population of wild-type cells, mutants arise (highlighted purple cells) and reproduce (non-highlighted purple cells). (b) The rate at which mutants are generated grows as the number of wild-type cells increases (i.e., μN_t). (c) The rate of transformation per wild-type cell is the mutation rate μ. (d) Wild-type cells growing in 9 separate populations where mutants arise in a portion of the populations (those with purple backgrounds) at different cell divisions. (e) In a growing population of donors and recipients, transconjugants arise (highlighted purple cells) and reproduce (non-highlighted purple cells). (f) The rate at which transconjugants are generated grows

as the numbers of donors and recipients increase (i.e., $\gamma_D D_t R_t$). (g) The rate of transformation per recipient cell grows as the number of donors increases (i.e., $\gamma_D D_t$) where γ_D is the constant conjugation rate parameter. (h) Donor and recipient cells growing in 9 separate populations where transconjugants arise in a portion of the populations (purple backgrounds) at different points in time. For all panels, this is a conceptual figure, and the rates are inflated for illustration purposes.

The Luria-Delbrück method (LDM) has improved accuracy and precision.

To explore the accuracy and precision of the LDM estimate and compare it to the SIM estimate (as well as other estimates, see SI section 4), we used the Gillespie algorithm to stochastically simulate the dynamics of donors, recipients, and transconjugants in a standard mating assay using equations [1]-[3] (Figure 3). Since the mating assay starts without transconjugants, a critical time point (hereafter t^*) is marked by the creation of the first transconjugant cell due to the first conjugation event between a donor and a recipient. Before t^*, the only events occurring are the cell divisions of donors and recipients (Figure 3a). After t^*, all the event types described in Figure 1a can occur. Given that our simulation framework incorporates the stochastic nature of conjugation, t^* will vary among simulated mating assays. One stochastic run of the mating assay constitutes a simulation of the SIM approach. In the laboratory, the standard time point ($\tilde{t}$) used for the SIM estimate is 24 hours, however, a truncated assay ($\tilde{t} < 24$) also produces a non-zero estimate of the conjugation rate as long as the incubation time is greater than t^* (the orange region of Figure 3b and c) and the density of transconjugants is large enough to be detected.

While the SIM estimate uses the density of transconjugants ($T_{\tilde{t}}$), the LDM equation instead involves $p_0(\tilde{t})$, the probability that a population has no transconjugants at the end of the assay. A maximum likelihood estimate for this probability (hereafter $\hat{p}_0(\tilde{t})$) is obtained by calculating the fraction of populations (e.g., 100 parallel simulations were used) that have no transconjugants at the specific incubation time $\tilde{t}$ (top of Figure 3d). Thus, the range of time points to calculate the maximum likelihood estimate ($0 < \hat{p}_0(\tilde{t}) < 1$) will be flanking the average t^* (the brown region of Figure 3d). Because the LDM estimate requires the *absence* of transconjugants in a fraction of populations, while the SIM estimate requires their *presence*, the range of incubation times for the LDM approach will be earlier than the SIM approach.

Even though there is a range of 'valid' incubation times, the accuracy of the SIM estimate can change over time as shown in Figure 3c (same case shown in Figure 1b). In this case, a key modeling assumption of the SIM approach was violated as the transconjugant conjugation rate was much higher than the donor conjugation rate ($\gamma_T \gg \gamma_D$). Consequently, the SIM estimate of the donor conjugation rate was inflated compared to the true value by increasing amounts over time (Figure 3c). In contrast, the LDM estimate under the same scenario had high accuracy and precision over time (Figure 3e).

We explored other parameter settings across various incubation times and the LDM estimate generally performed as well or better than other estimates (SI section 4).

To more systematically explore the effects of heterogeneous growth and conjugation rates on the accuracy and precision of estimating the donor conjugation rate (γ_D), we ran sets of simulations sweeping through values of other parameters ($\psi_D, \psi_R, \psi_T, \text{and } \gamma_T$). An illustrative example of heterogeneous growth occurs when plasmids confer costs or benefits on the fitness of their host. We simulated a range of growth-rate effects on plasmid-containing hosts from large plasmid costs ($\psi_D = \psi_T \ll \psi_R$) to large plasmid benefits ($\psi_D = \psi_T \gg \psi_R$). Relative to the SIM estimate, the LDM estimate had equivalent or higher accuracy and precision across all parameter settings (Figure 4a). To explore inequalities in conjugation rate more comprehensively, we simulated a range of transconjugant conjugation rates from relatively low ($\gamma_T \ll \gamma_D$) to high ($\gamma_T \gg \gamma_D$) values. Once again, the LDM estimate generally outperformed the SIM estimate across this range (Figure 4b). In SI section 4, we explore other parametric combinations along with model extensions, where, overall, the LDM performed better than the SIM approach and other estimates. Given the large number of simulations for these sweeps, we chose parameter values (i.e., D_0, R_0, and γ_D) outside of typical values reported in the literature to reduce the computational burden of the Gillespie algorithm, which occurs when populations sizes become very large. However, the qualitative results were confirmed with a few simulations using parameter settings with more realistic values (SI section 4e).

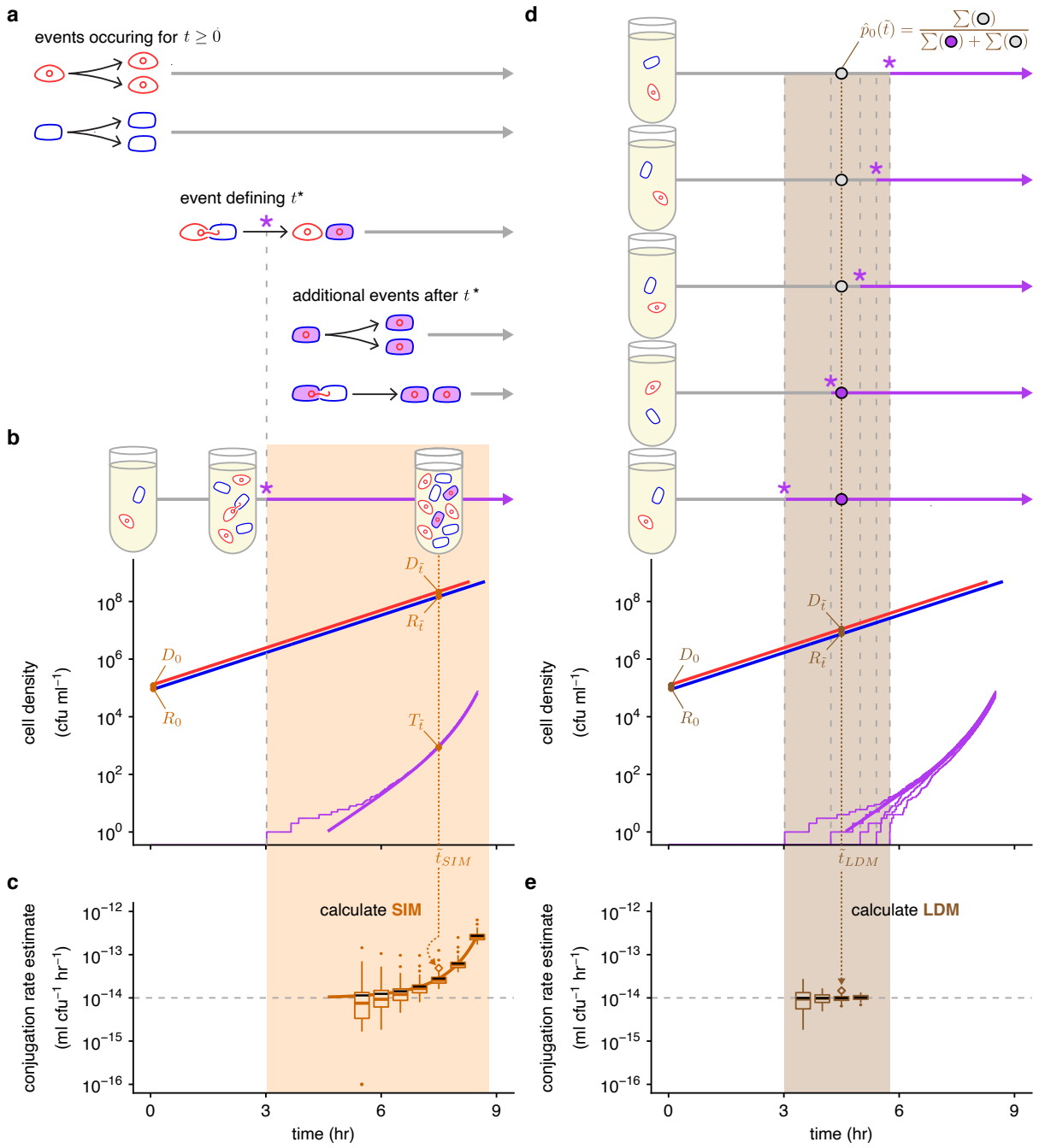


Figure 3 : Overview of stochastic simulation framework and the effects of incubation time on estimating conjugation rate. (a) The mating assay starts ($t = 0$) with donors and recipients; and their populations increase over time. At a critical time (t^*, marked by a purple asterisk), the first transconjugant cell is generated through a conjugation event between a donor and recipient. After t^*, all possible growth and conjugation events can occur (including transconjugant division and conjugation). (b) A stochastic simulation of the equations [1]-[3] shows the donor, recipient, and transconjugant densities (red, blue, and purple thin trajectories, respectively) increasing over time. The deterministic numerical solution of the same equations and parameter settings from Figure 1b is shown for reference (thick lines). We note that for large densities, the stochastic and deterministic trajectories are closely aligned (i.e., the thick

red and blue lines are overlaying their thin counterparts). After a specified incubation time ($\tilde{t}_{\mathrm{SIM}}$, dotted orange line), we measure the densities of the three populations (orange $D_{\tilde{t}}$, $R_{\tilde{t}}$, and $T_{\tilde{t}}$), which can be used to calculate the (c) SIM estimate. (d) Multiple mating assays are needed for the LDM estimate. Here, five stochastic simulations are shown, which display variation in t^*. At a specified incubation time ($\tilde{t}_{\mathrm{LDM}}$, dotted brown line), we determine the number of assay cultures with transconjugants (purple circles, where for a relevant culture i, $t_i^* < \tilde{t}_{\mathrm{LDM}}$) and without (gray circles, where for a relevant culture j, $t_j^* > \tilde{t}_{\mathrm{LDM}}$). These numbers are used to calculate $\hat{p}_0(\tilde{t})$, which, along with the donor and recipient densities (brown D_0, R_0, $D_{\tilde{t}}$ and $R_{\tilde{t}}$) are used for the (e) LDM estimate. The SIM (part c) and LDM (part e) estimates are calculated after different incubation times, where the $\tilde{t}_{\mathrm{SIM}}$ (part b) and $\tilde{t}_{\mathrm{LDM}}$ (part d) are indicated with orange and brown dotted arrows, respectively. The simulated trajectories in parts b and d would correspond to a single SIM or LDM estimate (the diamond points where the arrows terminate). The light orange and brown backgrounds indicate the range of incubation times giving a finite non-zero estimate of donor conjugation rate for the stochastic runs illustrated in parts b and d. In parts c and e, each box represents the distribution from 100 estimates of the donor conjugation rate for a given $\tilde{t}$, spanning from the 25th to 75th percentile. Given the log y-axis, the zero estimates are placed at the bottom of the y-axis range. The whiskers (i.e., vertical lines connected to the box) contain 1.5 times the interquartile range with the caveat that the whiskers were always constrained to the range of the data. The colored line in the box indicates the median. The solid black line indicates the mean. Parameter values are identical to Figure 1b and used throughout. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

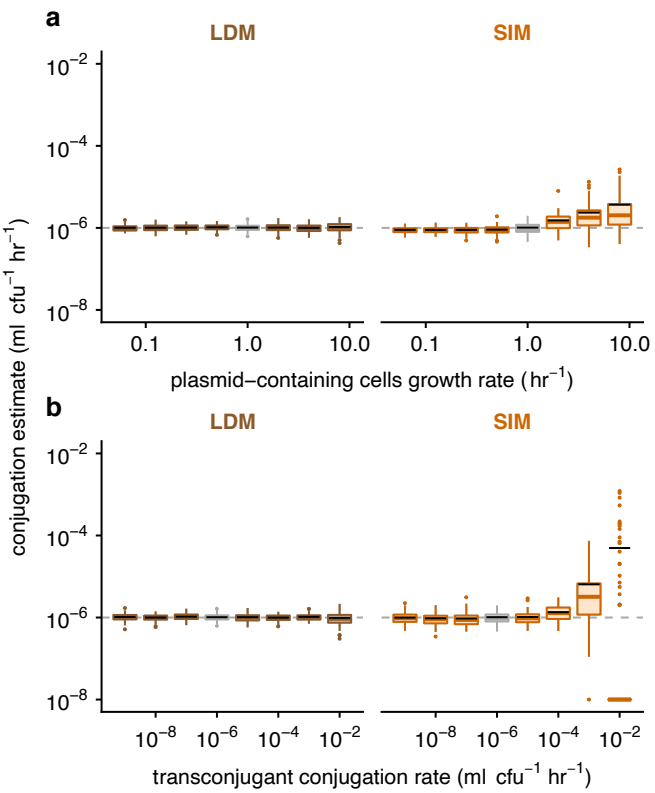


Figure 4 : The effect of parametric heterogeneity on estimating conjugation rate. The Gillespie algorithm was used to simulate population dynamics. 100 estimates of the donor conjugation rate are shown for each parameter (summarized using boxplots with the same graphical convention as in Figure 3). The gray dashed line indicates the true value for the donor conjugation rate (here, 10^{-6}). The boxes in gray indicate the baseline parameter setting, and all colored boxes represent deviation of one or two parameters from baseline. The baseline parameter values were $\psi_D = \psi_R = \psi_T = 1$ and $\gamma_D = \gamma_T = 10^{-6}$. The dynamic variables were initialized with $D_0 = R_0 = 10^2$ and $T_0 = 0$. All incubation times are short but are specific to each parameter setting (see Materials and Methods and SI Table 5). (a) Unequal growth rates were explored over a range of growth rates for the plasmid-bearing strains, namely $\psi_D = \psi_T \in \{0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8\}$. (b) Unequal conjugation rates were probed over a range of transconjugant conjugation rates, namely $\gamma_T \in \{10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$. For the 10^{-2} transconjugant conjugation rate, many of the runs resulted in SIM estimates of zero due to zero transconjugants at the specific incubation time; therefore, the median (colored line) and the box are placed at the bottom of the plot (given that the y-axis is on a log scale). The bulk of the data for this x-value is substantially lower than the mean SIM estimate (black line). The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

New laboratory protocol to implement the LDM.

We developed a general experimental procedure for estimating donor conjugation rate (γ_D) using the LDM approach in the laboratory. The LDM protocol is tractable and can accommodate a wide variety of microbial species and conjugative plasmids by allowing for distinct growth and conjugation rates among donors, recipients, and transconjugants. The basic approach is to inoculate many donor-recipient co-cultures and then, at a time close to the average t^*, add transconjugant-selecting medium (counterselection for donors and recipients) to determine the presence or absence of transconjugant cells in each co-culture.

In SI section 1, we rearrange equation [11] to provide an alternative form to highlight the quantities needed to conduct the LDM assay in the laboratory:

$$\gamma_D = \frac{f}{\tilde{t}}[-\ln \hat{p}_0(\tilde{t})]\frac{\ln D_{\tilde{t}}R_{\tilde{t}} - \ln D_0R_0}{D_{\tilde{t}}R_{\tilde{t}} - D_0R_0}. \quad [13]$$

Similar to previous conjugation estimates, the LDM protocol requires measurement of initial and final densities of donors and recipients (D_0, R_0, $D_{\tilde{t}}$, and $R_{\tilde{t}}$). In addition, the LDM approach requires a fraction of parallel donor-recipient co-cultures to have no transconjugants at the specified incubation time ($\tilde{t}$), which is the maximum likelihood estimate $\hat{p}_0(\tilde{t})$. Lastly, there is a correction factor when the co-culture volume deviates from 1 ml; specifically, f is the reciprocal of the co-culture volume in milliliters (e.g., for a co-culture volume of 100 ml, $f = 1/0.1 = 10$, SI section 5).

Before executing the LDM conjugation assay, an incubation time $\tilde{t}$ and initial density for the donors (D_0) and the recipients (R_0) needs to be chosen so that the probability that transconjugants form ($1 - p_0(\tilde{t})$) is not close to zero or one. We developed a short assay (SI section 6) for screening combinations of incubation time and initial densities to select a *target* incubation time ($\tilde{t}'$) as well as *target* initial densities (D_0' and R_0') where $0 < \hat{p}_0(\tilde{t}) < 1$. Note we add primes to indicate that these are '*targets'* to distinguish D_0, R_0, and $\tilde{t}$ in equation [13], which will be gathered in the conjugation protocol itself. In addition, this pre-assay simultaneously verifies that the LDM modeling assumption of constant growth is satisfied. In our case, this pre-assay revealed several time-density combinations that could have been used. A useful pattern to note is that a higher donor conjugation rate will require shorter incubation times and lower initial densities compared to a lower rate.

For the LDM conjugation assay, we mix exponentially growing populations of donors and recipients, inoculate many co-cultures at the target initial densities in a 96 deep-well plate, and incubate in non-selective growth medium with the specific experimental culture volume ($1/f$ of 1 ml) for the target incubation time (Figure 5). To estimate the initial densities (D_0 and R_0), three co-cultures at the start of the assay are diluted and plated on donor-selecting and recipient-selecting agar plates (Figure 5a). After the incubation time ($\tilde{t}$), final densities ($D_{\tilde{t}}$ and $R_{\tilde{t}}$) are also obtained by dilution-plating from the same co-cultures (Figure 5b). Liquid transconjugant-selecting medium is

subsequently added to the remaining co-cultures (Figure 5c). After a long incubation in the transconjugant-selecting medium, there should be a mixture of turbid and non-turbid wells. A turbid well results from one or more transconjugant cells being present at time $\tilde{t}$ (when transconjugant-selecting medium was added). Therefore, a non-turbid well indicates the absence of transconjugant cells at $\tilde{t}$, since the first conjugation event had not yet occurred ($\tilde{t} < t^*$, Figure 3), although see SI section 6 for a caveat. The proportion of non-turbid cultures is $\hat{p}_0(\tilde{t})$ (Figure 5c). Unlike the traditional Luria–Delbrück method, no plating is required to obtain $\hat{p}_0(\tilde{t})$. With the obtained densities (D_0, R_0, $D_{\tilde{t}}$, and $R_{\tilde{t}}$), the incubation time ($\tilde{t}$), the proportion of transconjugant-free cultures ($\hat{p}_0(\tilde{t})$), and the experimental culture volume correction (f), the LDM estimate for donor conjugation rate (γ_D) can be calculated via equation [13].

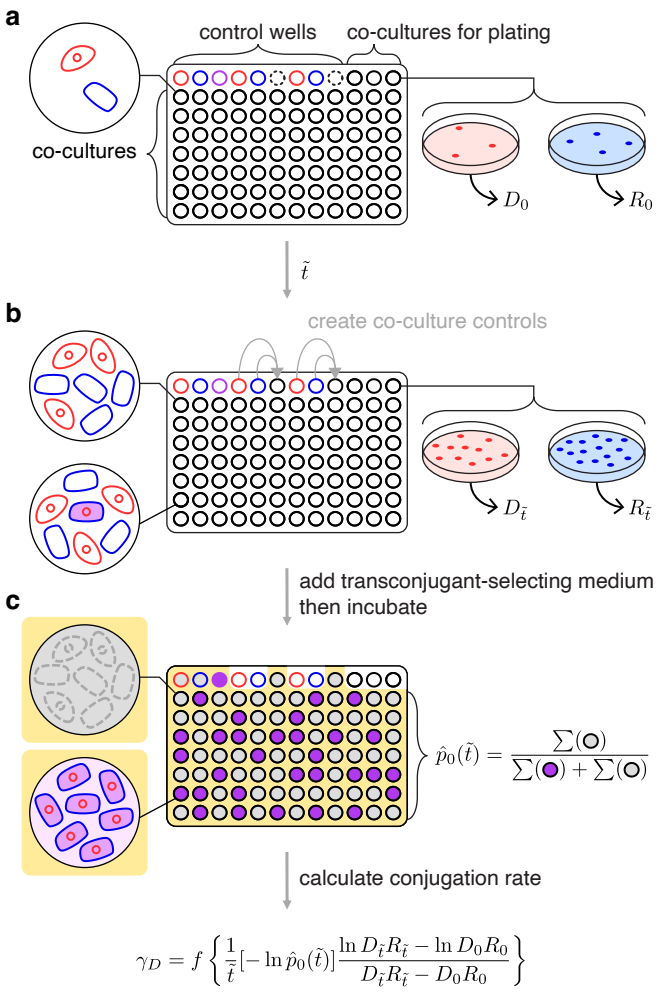


Figure 5 : Overview for executing the LDM conjugation protocol. (a) The wells of a microtiter plate are inoculated with parallel co-cultures (black-bordered circles) at the target initial densities (D_0' and R_0'). In addition, donor, recipient, and transconjugant monocultures serve as controls (red-, blue-, and purple-bordered wells, respectively). Three co-cultures (top-right) are sampled to determine the actual initial densities (D_0 and R_0). Note empty wells (dash-bordered circles) are used later in the assay. (b) After the incubation time ($\tilde{t}$), the same three co-cultures are sampled for final densities ($D_{\tilde{t}}$ and $R_{\tilde{t}}$). In addition, donor and recipient monocultures are mixed into the empty wells (indicated

by grey arrows) to create co-culture controls to verify that diluting with transconjugant-selecting medium effectively prevents conjugation. (c) Subsequently, transconjugant-selecting medium is added to the microtiter plate (indicated by the yellow background) and incubated for a long period. The transconjugant-selecting medium should inhibit donor and recipient growth, leading to non-turbid (gray-filled) donor and recipient control wells, but a turbid (purple-filled) transconjugant control well. In addition, the transconjugant-selecting medium should prevent new conjugation events leading to non-turbid co-culture controls (gray-filled). Focusing on the wells inoculated with parallel co-cultures, the proportion of transconjugant-free (i.e., non-turbid, gray-filled) cultures is $\hat{p}_0(\tilde{t})$. Using this proportion, the actual incubation time ($\tilde{t}$), initial densities (D_0 and R_0), final densities ($D_{\tilde{t}}$ and $R_{\tilde{t}}$), and the experimental culture volume correction (f), the LDM estimate of the donor conjugation rate (γ_D) can be calculated. One microtiter plate yields one LDM estimate.

Cross-species case study.

To empirically test the performance of our assay and our modeling predictions, we initiated a cross-species mating assay between a donor, *Klebsiella pneumoniae* (hereafter 'K') with a conjugative IncF plasmid (hereafter 'pF'), and a plasmid-free recipient, *Escherichia coli* (hereafter 'E'). We denote the donor strain as K(pF), where the host species name is listed first and the plasmid inside the host is given in the parenthesis. E(Ø) denotes the plasmid-free recipient strain. We implemented the LDM and SIM protocols on the same bacterial cultures to compare the laboratory estimates of the cross-species conjugation rate.

The standard SIM protocol involves an incubation of 24 hours. For many bacterial species (including the ones explored here), an incubation time ($\tilde{t}$) of 24 hours will lead to a violation of the assumption of constant growth rates from equations [1]-[3]. However, the original Simonsen et al. study did not actually assume constant growth rates (11). Their model permitted growth rate to vary as a function of resources, but additionally assumed that conjugation rate similarly varied. In other words, the ratio of growth and conjugation rates was assumed to remain constant (SI section 1). Under batch culture conditions, the population growth rates will drop as limiting resources are fully consumed (resulting in a stationary phase). If conjugation rates decrease with resources in a similar fashion and the parametric identicality assumptions hold, the SIM estimate can be used over a full-day incubation. Despite these restrictive underlying assumptions, it is not uncommon for researchers to estimate plasmid transfer rates using the 24-hour SIM protocol without experimental validation of assumptions. We proceeded with the standard SIM protocol to allow a comparison between this popular estimate and our new estimate (resulting from the LDM protocol, which does not rest on the same assumptions).

Our LDM estimate of the cross-species conjugation rate was significantly lower than the standard SIM estimate, by approximately three orders of magnitude (comparison A in Figure 6; t-test, $p<0.05$). This substantial incongruence could be due to a few possible

factors. First, it is possible that the growth and conjugation rates do not change with nutrients in a functionally similar way. While we cannot rule out this possibility, it has been shown for IncF plasmids that both growth and conjugation drop as resources decline to low levels (10), consistent with SIM model assumptions. Second, our growth rate assays (conducted separately from the transfer estimate protocols; see SI section 6b) revealed our cell types have different growth rates (SI Figure 8), thus violating the SIM assumptions. While simulations show there is an effect of these inequalities, the effect size is insufficient to explain the observed difference in comparison A (SI Figure 6). Lastly, it is possible that the within-species conjugation, between the E(pF) transconjugants and E(Ø) recipients, occurs at a substantially higher rate than the cross-species conjugation, between the K(pF) donors and E(Ø) recipients. Our simulations show that this kind of difference in conjugation rates can lead to notable inflation of the SIM estimate, and there is evidence that within-species conjugation rates can be markedly elevated over cross-species rates (15, 22). Thus, this last possibility warranted further investigation.

Next, we performed the within-species mating between *E. coli* strains. The LDM estimate for within-species conjugation rate (within *E. coli*) was higher than the cross-species LDM estimate by almost six orders of magnitude (comparison B in Figure 6; t-test, $p<0.001$), a difference that could explain the inflated SIM estimate. To further explore this explanation, we performed an additional cross-species SIM experiment with a shorter incubation time. In Figure 3c, as the incubation time was shortened, the SIM estimate approached the LDM estimate of the donor conjugation rate. Running the SIM protocol with a truncated incubation period (5 hours) resulted in a significantly lower cross-species conjugation rate estimate relative to the standard SIM estimate (comparison C in Figure 6; t-test, $p<0.05$), a result consistent with the pattern predicted under heterogeneous conjugation rates. Under a scenario with reduced parametric heterogeneity (e.g., $\gamma_D \approx \gamma_T$), we predicted that the SIM and LDM estimate would be similar. Consistent with our prediction, the SIM estimate for the within-species conjugation rate using the truncated SIM protocol was not significantly different than the LDM estimate (comparison D in Figure 6; t-test, $p=0.23$).

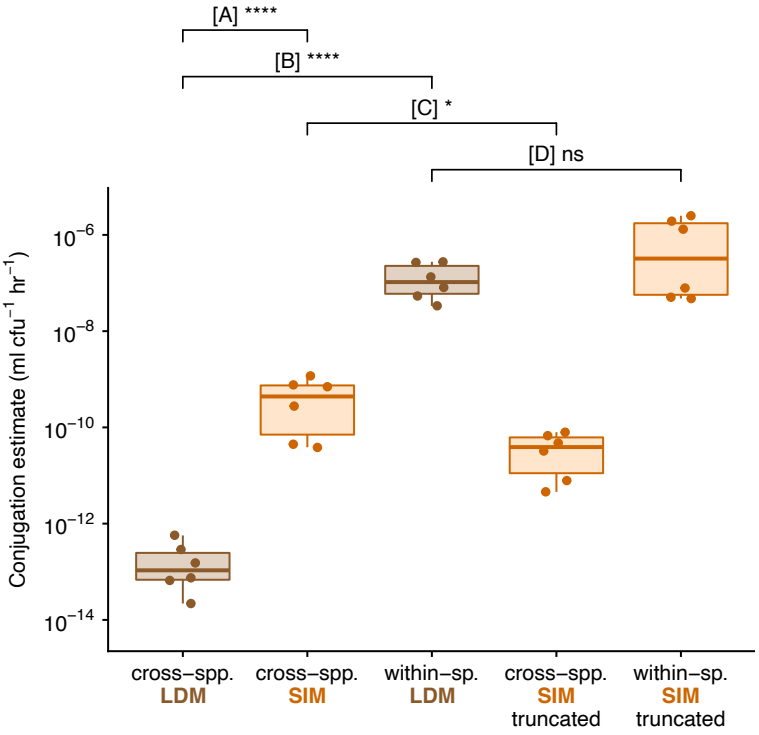


Figure 6 : Experimental estimates for cross-species and within-species conjugation rates. Each box summarizes six replicate estimates by the LDM, SIM or truncated SIM approach, where each data point corresponds to a replicate. We note each of these estimates involved a correction (see Materials and Methods), but the same patterns hold for uncorrected values. [A] compares the LDM and standard SIM approach for a cross-species mating (between *K. pneumoniae* and *E. coli*). [B] compares the cross- and within-species mating using the LDM approach. [C] compares the standard and truncated SIM approach for a cross-species mating. [D] compares the LDM and truncated SIM approach for a within-species mating. The asterisks indicate statistical significance by a t-test (one, three and four asterisks convey p-values in the following ranges: $0.01 < p < 0.05$ and $p < 0.0001$, respectively). The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

Discussion

Conjugation is one of the primary modes of horizontal gene transfer in bacteria, facilitating the movement of genetic material between non-related neighboring cells. In microbial communities, conjugation can lead to the dissemination of genes among distantly related species. Since these genes are often of adaptive significance (e.g., antibiotic resistance), a comprehensive understanding of microbial evolution requires a full account of the process of conjugation. One of the most fundamental aspects of this process is the rate at which it occurs. Here we have presented a new method for estimating the rate of plasmid conjugative transfer from a donor cell to a recipient cell. We derived our LDM estimate using a mathematical approach that captures the

stochastic process of conjugation, which was inspired by the method Luria and Delbrück applied to the process of mutation (19). We explore the connection between mutation and conjugation further in SI section 7. Our new method departs from the mathematical approach for other conjugation rate estimates, which assume underlying deterministic frameworks guiding the dynamics of transconjugants (10, 11, 14). Beyond the incorporation of stochasticity, the model and derivation behind the LDM estimate relaxes assumptions that constrain former approaches, which makes calculating conjugation rates available to a wide range of systems with different plasmid-donor-recipient combinations.

The LDM approach has improved accuracy

The most widely used approaches to estimate conjugation rate are derived from the Levin *et al*. model (10) (discussed in SI section 1) which assumes that all strains grow and conjugate at the same rate ($\psi_D = \psi_R = \psi_T$ and $\gamma_D = \gamma_T$). For instance, the model underlying the SIM approach assumes precisely this kind of homogeneity. These assumptions and constraints are problematic because bacterial growth and conjugation can and do vary (23, 24). Specifically, donors and recipients are often different taxa and contain chromosomal differences that translate to growth or conjugation rate differences ($\psi_D \neq \psi_R$ or $\gamma_D \neq \gamma_T$). Additionally, plasmid carriage can change growth rate substantially (25) and therefore recipients can grow differently from donors ($\psi_R \neq \psi_D$) or transconjugants ($\psi_R \neq \psi_T$). In microbial communities, heterogeneous rates of growth and conjugation are the rule and not the exception. Therefore, a general estimation approach should be robust to this heterogeneity. While the estimates of popular approaches are insensitive to certain forms of heterogeneity, they can also be inaccurate under other forms. In contrast, the LDM estimate remains accurate across a broad range of heterogeneities.

A recent approach by Huisman *et al*. (14) relaxed the assumption of parametric homogeneity, yielding useful revisions to the SIM approach. However, when transconjugants exhibit much larger rates of plasmid transfer than the donors ($\gamma_T \gg \gamma_D$), this new method can become inapplicable. Unfortunately, this kind of difference in conjugation rates is likely not uncommon in microbial communities (15, 26). Indeed, a mating assay involving two species can be thought of as a miniaturized microbial community where cross-species conjugation (between donors and recipients) and within-species conjugation (between transconjugants and recipients) both occur. Both previous work (15, 26) and experimental data from this study (Figure 6) demonstrate that the transconjugant (within-species) conjugation rate can be significantly higher than the donor (cross-species) rate. In addition, a similar difference in conjugation rates can arise from transitory de-repression, a molecular mechanism encoded on the conjugative plasmid that temporarily elevates the conjugation rate of a newly formed transconjugant (10, 27). The LDM approach is robust to these differences because it focuses on the creation of the first transconjugant (an event that must be between a donor and recipient) and ignores subsequent transconjugant dynamics (which is affected by transconjugant transfer). The

LDM method produces an accurate estimate for donor conjugation rate in systems with unequal conjugation rates, whether the differences are taxonomic or molecular in origin.

The LDM approach has improved precision

In addition to improved accuracy, the LDM estimate has advantages in terms of precision. Since conjugation is a stochastic process, the number of transconjugants at any given time is a random variable with a certain distribution. Therefore, estimates that rely on the number of transconjugants (which includes nearly all available methods) or the probability of their absence (the LDM approach) will also fall into distributions. Even in cases where the mean (first moment of the distribution) is close to the actual conjugation rate, the variance (second central moment) may differ among estimates. For the number of parallel co-cultures in our protocol, the LDM estimate had smaller variance compared to other estimates, even under parameter settings where different estimates shared similar accuracy (e.g., Figure 4). This greater precision likely originates from the difference in the distribution of the number of transconjugants ($T_{\tilde{t}}$) and the distribution of the probability of transconjugant absence ($p_0(\tilde{t})$, where variance decreases with the number of co-cultures), something we explore analytically in SI section 8. Beyond the mean and variance, other features of these distributions (i.e., higher moments) may also be important. For certain parameter settings, the estimates relying on transconjugant numbers were asymmetric (the third moment was non-zero). In such cases, a small number of replicate estimates could lead to bias (SI section 4). Typically, a small number of conjugation assays is standard; thus, the general position and shape of these estimate distributions may matter. Over the portion of parameter space that we explored, the LDM distribution facilitated accurate and precise estimates through its position (a mean reflecting the true value) and its shape (a small variance and a low skew).

The LDM approach has implementation advantages

As discussed above, violations of modeling assumptions can lead to significant bias when estimating conjugation rate. Therefore, when implementing a conjugation protocol, the degree to which the experimental system satisfies the relevant assumptions is of prime importance. The most straightforward way to deal with this issue is to experimentally confirm that assumptions hold. For instance, the model underlying the LDM estimate assumes that growth rates of each cell type remain constant throughout the assay. This verification is part of the LDM protocol (see Materials and Methods and SI section 6). We emphasize that confirming the satisfaction of an assumption for one experimental system does not guarantee that the assumption holds for other systems. For example, the model underlying the SIM estimate assumes that growth and conjugation rates respond in a functionally similar way to changes in resources. While this assumption was explored for the IncF plasmid used in the original SIM study (10), other plasmid systems will readily violate it (e.g., some IncP plasmids conjugate during stationary phase after growth has stopped (28)), which can lead to bias in the estimate (SI section 4h). Some approaches do not experimentally verify modeling assumptions as

part of their corresponding protocol, but rather rely on simulated sensitivity analyses showing violations have little to no effect on the estimate (11, 14). For instance, the SIM estimate is robust to relatively small differences in growth rates or conjugation rates (11). Overall, for any conjugation rate estimate, either the underlying assumptions should be validated for the focal experimental system, or a rationale offered for why certain violations by the focal system will not significantly bias the estimate.

Given recent interest in quantitative estimates of conjugation rate (29) and the impacts of model assumption violations (14), there has been a matching interest in altering conjugation protocols such that bias is minimized when violations apply. A common procedural adjustment involves shortening the incubation period of the mating assay because the bias resulting from modeling violations can increase over time (12, 13). For instance, when the transconjugant transfer rate is much higher than the donor rate, shortening the incubation time can mitigate some of the inaccuracies in the SIM estimate (Figure 3 and 6). However, there are a few caveats to this adjustment for estimates that rely on transconjugant density (which includes all common approaches, but not the LDM). First, as the incubation time decreases, the benefits in estimate accuracy come at the expense of costs in estimate precision. Specifically, variation in the timing of the first transconjugant cell appearance (t^* in Figure 3) has a greater impact on estimate variance with earlier incubation times. In part because the LDM approach does not rely on a measurement of transconjugant density, the LDM estimate remains both accurate and precise across various incubation times. Second, as incubation time decreases the transconjugant population can become extremely small and therefore technical problems with measuring an accurate transconjugant density through plating can arise (13). For instance, when the transconjugants are rare in the mating culture, the low dilution factor required for selective agar plating for transconjugants ensures a very high density of donors and recipients are simultaneously plated. Before complete inhibition of the donors and recipients by the transconjugant-selecting medium, conjugation events on the plate can generate additional transconjugants inflating the conjugation rate estimate (7, 13, 30, 31). Recently, a spectrophotometric technique was introduced to avoid selective plating altogether, which addresses this second caveat (13), but not the first. Notably, neither of these two caveats apply to the LDM approach because a binary output (turbid or non-turbid cultures) is used in lieu of measuring transconjugant density. Overall, the LDM protocol is both experimentally streamlined and insensitive to factors that can confound other approaches.

The LDM approach is broadly applicable

In this paper, we have highlighted the possibility that the rate of conjugation may change (substantially) with the identity of the plasmid-bearing cell (32–34). For instance, as a plasmid moves from the original donor strain to the recipient background (forming a transconjugant), the transfer rate can change (i.e., $\gamma_D \neq \gamma_T$). However, the conjugation rate changes with much more than just the identity of the cell holding the plasmid. The rate of transfer can additionally depend on the identity of the recipient as well as

environmental conditions (e.g., level of nutrients, presence of antibiotics, etc.) (35). Thus, there is no *single* conjugation rate "belonging to" a plasmid-bearing strain. Our LDM approach is meant to be a conditional "snapshot," where the conjugation rate depends on conditions of the protocol and the strains used. It is entirely possible to run the LDM approach under different conditions (e.g., changing nutrients) and assess the effect of environmental factors on transfer rate. The donor conjugation rate can be calculated under any condition as long as strain growth rates are constant over the protocol. But the distinguishing feature that gives the LDM method relative breadth of application is that it is robust to a form of conditionality that is tied to the mating assay itself. Specifically, because transconjugants are formed during a mating assay and, like donors, can deliver the plasmid to additional recipients, a form of rate conditionality is an unavoidable possibility for any protocol employing a mating assay. As we have shown (Figure 1, 3, 4, and 6), a difference in transfer rate between donors and transconjugants can make popular estimates inaccurate. However, by focusing on the first transconjugant formed (which only involves the donor and recipient, Figure 6), the LDM sidesteps this conditionality altogether, allowing an unbiased estimate of donor conjugation rate under a user-defined environment.

In conclusion, the LDM offers new possibilities for measuring the conjugation rate for many types of plasmids, species, and environmental conditions. We have presented evidence that supports our method being more accurate and precise than other widely used approaches. Importantly, the LDM eliminates bias caused by relatively high transconjugant conjugation rates, which is not unlikely when the donor and recipient belong to different species. We experimentally explored a case where the transconjugant transfer rate was dramatically higher than the donor rate and found that a standard estimate could inflate the conjugation rate (Figure 6). More generally, violations of model assumptions, intrinsic stochasticity, and implementation constraints can cause problems for currently available approaches. However, an adjustment of the approach Luria and Delbrück used to explore and estimate mutation over 75 years ago can address many of these issues. This new approach greatly expands the ability of experimentalists to accurately measure conjugation rates under the diverse conditions found in natural microbial communities.

Materials and Methods

More detailed information for the mathematical models, simulations, and experiments are provided in the supplementary information. A general LDM protocol is deposited on protocols.io (dx.doi.org/10.17504/protocols.io.e6nvwk812vmk/v3).

Bacterial Strains.

Donor strains included two Enterobacteriaceae species: *Escherichia coli* K-12 BW25113 (36) and *Klebsiella pneumoniae* Kp08 (7). We use the first letter of the genus

(E and K) to refer to these species throughout. The recipient strain is derived from the same isogenic strain as the *E. coli* donor strain but encodes additional chromosomal streptomycin resistance, providing a unique selectable marker to distinguish the donor and recipient hosts in both the cross- (K to E) and within-species (E to E) mating assays. The focal conjugative plasmid was used previously (37): plasmid F'42 from the IncF incompatibility group. A tetracycline resistance gene was cloned into the F'42 plasmid (38) and used as the selectable marker to distinguish plasmid-containing from plasmid-free hosts. This derived plasmid is referred to as 'pF' throughout.

Conjugation Assays.

Strains were inoculated into LB medium from frozen isogenic glycerol stocks and grown for approximately 24 hours. The plasmid-containing cultures were supplemented with 15 μg ml^{-1} tetracycline to maintain the plasmid. The saturated cultures were diluted 100-fold into LB medium to initiate another 24 hours of growth (to acclimate the previously frozen strains to laboratory conditions). The acclimated cultures were then diluted 10,000-fold into LB medium and incubated for 4 hours to ensure the cultures entered exponential growth (SI section 6b). The exponentially growing cultures were diluted by a factor specific to the donor-recipient pair (SI section 6e), mixed at equal volumes, and dispensed into 84 wells of a deep-well microtiter plate at 100 μl per well (Figure 5a solid black-bordered wells in rows 2 - 8, these wells were the co-cultures used to estimate $p_0(\tilde{t})$). In an additional 3 wells (Figure 5a solid black-bordered wells in top row), 130 μl (per well) of the mixture was dispensed and immediately 30 μl was removed to determine the initial densities (D_0 and R_0) via selective plating. An additional 3 wells contained monocultures of the three strains. Specifically, 100 μl of donor, recipient and transconjugant cultures were placed in their own well (Figure 5a red-, blue- and purple-bordered wells, respectively, in the top left). Later in the assay, these monocultures determined if the transconjugant-selecting medium prohibited growth of both donors and recipients, while permitting growth of transconjugants. An additional 4 wells contained diluted monocultures of donors and recipients (2 wells each at 100 μl, Figure 5a red- and blue-bordered wells, respectively, in the top middle). These monocultures were used to create co-cultures (in empty wells, Figure 5a dash-bordered wells) during the assay itself (see below). The deep-well plate was incubated for a pre-determined incubation time $\tilde{t}$ (SI section 6e), after which three events occurred in rapid succession. First, 30 μl was removed from each of the wells used to determine initial densities, to uncover the final densities ($D_{\tilde{t}}$ and $R_{\tilde{t}}$) via selective plating (Figure 5b). We note that densities were calculated from each well then averaged for calculating the LDM estimate. Second, donor and recipient monocultures were mixed at equal volumes into the two empty wells (Figure 5b, gray arrows). At a later point in the assay, these two wells verified that new transconjugants did not form via conjugation after transconjugant-selecting medium was added. Third, 900 μl of transconjugant-selecting medium (7.5 μg ml^{-1} tetracycline and 25 μg ml^{-1} streptomycin; see SI section 6c and SI section 6d) was added to all co-cultures

used to estimate $p_0(\tilde{t})$ as well as relevant control wells (Figure 5c, yellow background). This medium disrupted new conjugation events—immediately by diluting cells then by inhibiting donors and recipients—while simultaneously allowing growth of transconjugants. When designing transconjugant-selecting media, appropriate preliminary and control experiments must be conducted to ensure that the media enables exclusive growth of transconjugants (see SI section 6c). The deep-well plate was incubated for 4 days, and the state of all wells (turbid or non-turbid) was recorded. For both mating assays in this study (i.e., cross- and within-species), this conjugation protocol was repeated 6 times.

For the cross-species mating, the SIM method was executed alongside the LDM method described above. The SIM approach was conducted for two incubation times: a standard 24 hours and a truncated 5 hours. In an additional deep-well plate, 100 µl of the donor-recipient co-culture was dispersed into six wells, split into two groups of three wells each where each group corresponded to a different incubation time. To derive the SIM estimate for each incubation group, 30 µl was removed from each of the three wells in the group at the relevant incubation time ($\tilde{t} = 5$ and $\tilde{t} = 24$) to determine the final donor ($D_{\tilde{t}}$), recipient ($R_{\tilde{t}}$), and transconjugant ($T_{\tilde{t}}$) densities via selective plating. Densities were calculated from each well then averaged for calculating the SIM estimate. This protocol was repeated six times alongside the LDM replicates. For the within-species mating, the SIM method was executed as outlined above for the cross-species mating but using only the truncated SIM method with a 3-hour incubation time. Similar to the LDM protocol, we ran a control to confirm that conjugation did not occur after co-cultures were exposed to transconjugant-selecting medium, but in this case, it was for agar plates instead of liquid medium. Specifically, for the first SIM replicate per incubation time, an additional three donor monocultures and three recipient monocultures were initiated as above. At each incubation time (e.g., 5 and 24 hours for the cross-species mating), three new donor-recipient co-cultures were created in empty wells and, immediately plated on transconjugant-selecting agar at dilutions used to determine transconjugant densities. For this case, no transconjugant colonies formed (indicating that conjugation does not occur on the selective agar plate). We emphasize that this is a necessary step for any new system as post-plating conjugation has been reported (7, 13, 31).

For both the LDM and SIM approaches, the working assumption is that a cell will successfully establish a lineage under the appropriate selective conditions. As one example, a well with a single transconjugant will become turbid after incubation with transconjugant-selective medium. As another example, a donor cell on a donor-selecting agar plate will form a visible colony after incubation. A recent paper (39) has clearly demonstrated that this working assumption needs to be checked. In SI section 6, we offer adjustments to the protocols to improve the chances that this assumption holds. Additionally, we present ways to correct estimates if the assumption does not hold. In Figure 6, we used these corrections (see SI section 6 and SI section 7 for details).

Stochastics simulations.

We used the Gillespie algorithm available in the GillesPy2 open-source Python package for stochastic simulations (40). We specified starting cell densities and parameters and simulated population dynamics using equations [1]-[3] for a set incubation time in a 1 ml culture volume. For each parameter setting, we simulated 10,000 populations and calculated the conjugation rate using the LDM and SIM estimates. Each estimate has different requirements for calculating the conjugation rate (Figure 3). The LDM estimate needs multiple populations to calculate $\hat{p}_0(\tilde{t})$; therefore, for each LDM estimate we reserved 100 independent populations to compute $\hat{p}_0(\tilde{t})$ then one random population in the set of 100 was used to calculate the initial and final cell densities. In other words, the 10,000 populations yielded 100 LDM estimates. In contrast, one simulated population yields one SIM estimate. Therefore, we used the random populations chosen to calculate the densities for each of the 100 LDM estimates to calculate 100 SIM estimates.

For the incubation time sweeps (Figure 3), the conjugation rate was estimated at 30-minute intervals up until the total population size reached 10^9 cfu ml^{-1}. A 30-minute interval was analyzed if at least 90 percent of the estimates were finite and non-zero. Notably, the 30-minute intervals occur over an earlier time range for the LDM estimate then for the SIM estimate due to the different estimate requirements.

To compare across various parameter settings (Figure 4), a single incubation time was chosen per parameter set and type of estimate (see SI Table 5 for the incubation times used). For each parameter setting, the incubation time $\tilde{t}$ for the LDM estimate is set to the average t^*. In addition, the incubation time for the SIM estimate is given by the time point for which an average of 50 transconjugants is reached. This choice resulted in a truncated SIM approach (i.e., $\tilde{t} < 24$). However, any estimate bias from a truncated simulation would be conservative relative to the standard SIM approach.

In SI section 9, we explore through further simulation the impacts of the random effects of dilution, plating, and lineage extinction on the accuracy and precision of the LDM and SIM approaches.

Data and Code Availability.

All generated data and custom software are deposited in a GitHub repository (https://github.com/livkosterlitz/LDM) and archived on Zenodo (https://doi.org/10.5281/zenodo.6677158).

Author contributions

O.K. and B.K. conceived of the presented ideas. B.K. and I.B. developed the theory. O.K. developed the simulations and experimental protocols. C.E. facilitated part of the simulations. O.K., A.M.T., and C.W. performed the experiments. O.K. and B.K. wrote the manuscript. E.M.T. and B.K. supervised the project. All authors discussed the results and contributed to the final manuscript.

Acknowledgements

This work is supported by the National Institute of Allergy and Infectious Diseases Extramural Activities grant no. R01 AI084918 of the National Institutes of Health and Division of Environmental Biology grant no. 2142718 from the National Science Foundation. O.K. is supported by the NSF Graduate Research Fellowship grant no. DGE-1762114. C.E. is supported by the NSF Graduate Research Fellowship grant no. DGE-2019265372. We thank Hannah Jordt and members of the Kerr and Top laboratories for useful suggestions on the manuscript.

CHAPTER 1 supplementary information:

SI section 1 : Overview of approaches to estimate conjugation rate.

SI section 1a : Overview of theoretical frameworks

In this section, we highlight three key methods for estimating conjugation rate. While outlining the theoretical frameworks, we highlight the key distinctions and theoretical assumptions of each approach. Levin *et. al.* (10) introduced a simple mathematical model describing the change in density of donors, recipients and transconjugants over time (given by dynamic variables D_t, R_t, and T_t, respectively). In this model, each population type grows exponentially at the same growth rate ψ. In addition, the transconjugant density increases because of conjugation events both from donors to recipients and from existing transconjugants to recipients at the same conjugation rate γ. The recipient density decreases due to these conjugation events. The densities of these dynamic populations are described by the following differential equations (where the t subscript is dropped from the dynamic variables for notational convenience):

$$\frac{dD}{dt} = \psi D, \qquad [1.1]$$

$$\frac{dR}{dt} = \psi R - \gamma DR - \gamma TR, \qquad [1.2]$$

$$\frac{dT}{dt} = \psi T + \gamma DR + \gamma TR. \qquad [1.3]$$

Equations [1.1]-[1.3] contains four notable assumptions. First, conjugation is described by mass-action kinetics, where conjugation events are proportional to the product of donor and recipient cell densities, which is a reasonable assumption in well-mixed liquid cultures (10). Second, the model assumes a negligible rate of plasmid loss, a process whereby a dividing plasmid-containing cell produces one plasmid-containing daughter cell and one plasmid-free daughter. These first two assumptions exist in all the conjugation rate estimates we discuss. Third, the growth rate is the same for all cell types (i.e., in the language of equations [1]-[3], $\psi_D = \psi_R = \psi_T = \psi$). Fourth, the plasmid conjugation rate is the same from donors to recipients as from transconjugants to recipients (i.e., in the language of equations [1]-[3], $\gamma_D = \gamma_T = \gamma$). More specifically, equations [1.1]-[1.3] are a special case of equations [1]-[3] where growth and conjugation is assumed to be homogeneous across strains.

Popular rate estimation methods solved the set of ordinary differential equations from the Levin *et. al.* model (or a variation) to find an estimate for the conjugation rate γ. The various methods differ by the assumptions used to find the analytical solution. Levin *et. al.* was the first to derive an estimate for the conjugation rate (γ) by making three additional simplifying assumptions. First, the change in cell density of donors due to growth is assumed to be negligible (i.e., $dD/dt \approx 0$). Likewise, the change in cell density

of recipients due to growth and to conjugation (i.e., transformation into transconjugants) is assumed to be negligible (i.e., $dR/dt \approx 0$). Finally, transconjugants are assumed to be rare in the population such that the increase in transconjugant cell density is primarily through plasmid conjugation from donors to recipients (i.e., in equation [1.3], $\gamma DR \gg \psi T + \gamma TR$). All of these assumptions are satisfied if the cell growth rate is zero ($\psi = 0$), the conjugation rate (γ) is small, the system starts without transconjugants ($T_0 = 0$), and the densities of donors and recipients remain much greater than the density of transconjugants for the period under consideration ($D \gg T$ and $R \gg T$). Using these simplifying assumptions, Levin *et. al.* solved for an expression of the conjugation rate in terms of the density of donors, recipients, and transconjugants ($D_{\tilde{t}}$, $R_{\tilde{t}}$, and $T_{\tilde{t}}$, respectively) after a period of incubation $\tilde{t}$ (see the GitHub Appendix I for a few different approaches to the derivation).

$$\gamma_D = \frac{T_{\tilde{t}}}{D_{\tilde{t}} R_{\tilde{t}} \tilde{t}}. \qquad [1.4]$$

We label the expression in equation [1.4] as the "TDR" estimate for the conjugation rate, where the letters in this acronym come from the dynamic variables used in the estimate. Besides the model assumptions of homogenous growth rates and conjugation rates, the most notable assumption used in the TDR derivation is that there is little to no change in the population densities due to growth. Thus, laboratory implementation that respects this assumption can be difficult (see SI section 1b for details). Regardless, TDR is a commonly used estimate (10, 12, 13, 41).

Simonsen *et. al.* derived the other most widely used estimate for conjugation rate γ, which importantly expands application beyond the TDR method by allowing for population growth (11). Indeed, they allowed for the rate of population growth to change with the level of a resource in the environment, adding a dynamic variable for the resource concentration. In addition, the conjugation rate can also change with the resource concentration. The authors focus on a case where both growth and conjugation rates vary with resource concentration according to the Monod function. This choice was informed by experimental results showing that cells enter stationary phase and conjugation ramps down to a negligible level as resources are depleted (10). This pattern occurs for various plasmid incompatibility groups, but not all (42). Simonsen *et. al.* used this updated model to derive an estimate for plasmid conjugation rate (see GitHub Appendix II for the derivation).

$$\gamma_D = \psi \ln\left(1 + \frac{T_{\tilde{t}}}{R_{\tilde{t}}} \frac{N_{\tilde{t}}}{D_{\tilde{t}}}\right) \frac{1}{(N_{\tilde{t}} - N_0)}. \qquad [1.5]$$

We refer to equation [1.5] as the "SIM" estimate for conjugation rate throughout the manuscript, where SIM stands for "Simonsen *et. al.* Identicality Method" since the underlying model assumes that all strains are *identical* with regards to growth, and donors and transconjugants are *identical* with regards to conjugation rate. The SIM estimate involves measuring the density of the initial population (N_0), and the final density of donors ($D_{\tilde{t}}$), recipients ($R_{\tilde{t}}$), transconjugants ($T_{\tilde{t}}$), and the total population ($N_{\tilde{t}}$) after the incubation time $\tilde{t}$. The SIM estimate is popular since it allows for the use of batch culture in the

laboratory (see SI section 1b for details). Thus, it circumvents the constraints of the laboratory implementation of TDR; however, the underlying model holds the same assumptions as before: homogeneous growth rates and conjugation rates.

Huisman *et. al.* recently updated the SIM model, further extending its application by relaxing the assumption of identical growth and transfer rates for all strains (14). Specifically, the authors introduced population specific growth rates for donors, recipients, and transconjugants (ψ_D, ψ_R, and ψ_T, respectively) and population specific conjugation rates for donors and transconjugants (γ_D and γ_T). Huisman *et. al.* made three additional simplifying assumptions. First, conjugation and growth rates are assumed to be constant until resources are depleted, eliminating the additional resource concentration equation added in the SIM approach. Second, the increase in recipients due to growth greatly outpaces the decrease in recipients due to conjugation (i.e., $\psi_R R \gg \gamma_D DR + \gamma_T TR$). Third, the increase in transconjugants due to growth or plasmid conjugation from donors to recipients greatly outpaces the increase in transconjugants due to plasmid conjugation from transconjugants to recipients (i.e., $\psi_T T + \gamma_D DR \gg \gamma_T TR$). These model conditions are reasonable if the system starts with donors and recipients present but transconjugants are absent, the system is tracked for a short period of time $\tilde{t}$, conjugation rates are low relative to growth rates, and the transconjugant conjugation rate (γ_T) is not much higher than the donor conjugation rate (γ_D). With these added assumptions, equations [1.1]-[1.3] can be reformulated as the following *approximate* system of equations:

$$\frac{dD}{dt} = \psi_D D, \qquad [1.6]$$

$$\frac{dR}{dt} = \psi_R R, \qquad [1.7]$$

$$\frac{dT}{dt} = \psi_T T + \gamma_D DR, \qquad [1.8]$$

Huisman *et. al.* used these equations to derive an estimate for the donor conjugation rate

$$\gamma_D = (\psi_D + \psi_R - \psi_T) \frac{T_{\tilde{t}}}{\left(D_{\tilde{t}} R_{\tilde{t}} - D_0 R_0 e^{\psi_T \tilde{t}}\right)}. \qquad [1.9]$$

where different cell types now can have different growth rates (see GitHub Appendix III for the derivation). We term equation [1.9] as the ASM estimate for donor conjugation rate, where ASM stands for “Approximate Simonsen *et. al.* Method”.

For all methods (TDR, SIM, ASM, and LDM), we summarize model variables and parameters in SI Table 1. In addition, all variables used in the conjugation estimates are in SI Table 2. Lastly, all assumptions underlying each estimate are in SI Table 3.

SI Table 1: Variables and parameters used in plasmid dynamic models.

Variable/ Parameter	Description	Relevant Estimate(s)	Units
D	Donor density	TDR, SIM, ASM, LDM	$\frac{\text{cfu}}{\text{ml}}$
R	Recipient density	TDR, SIM, ASM, LDM	
T	Transconjugant density	TDR, SIM, ASM, LDM	
ψ	Growth rate (not population specific)	TDR, SIM	$\frac{1}{\text{hr}}$
ψ_D	Donor growth rate	ASM, LDM	
ψ_R	Recipient growth rate	ASM, LDM	
ψ_T	Transconjugant growth rate	ASM, LDM	
γ	Conjugation rate (not population specific)	TDR, SIM	$\frac{\text{ml}}{\text{cfu} \cdot \text{hr}}$
γ_D	Donor-recipient conjugation rate	ASM, LDM	
γ_T	Transconjugant-recipient conjugation rate	ASM, LDM	

SI Table 2: Variables and parameters used to estimate* conjugation rate

Variable/Parameter	Description	Relevant Estimate	Units
$\tilde{t}$	Incubation time (final sampling time)	TDR, SIM**, ASM, LDM	hr
D_0, R_0	Initial donor and recipient densities	ASM, LDM	$\frac{\text{cfu}}{\text{ml}}$
$D_{\tilde{t}}, R_{\tilde{t}}$	Final donor and recipient densities	TDR, SIM, ASM, LDM	
$T_{\tilde{t}}$	Final transconjugant density	TDR, SIM, ASM	
$N_0, N_{\tilde{t}}$	Initial and final total population density	SIM	
ψ_T	Transconjugant growth rate	ASM	hr^{-1}
$p_0(\tilde{t})$	Probability a population has no transconjugants	LDM	

* The laboratory estimates are used here (see SI section 1b)
** If the SIM assay is conducted on exponentially growing cultures (see SI section 1c) $\tilde{t}$ along with N_0 and $N_{\tilde{t}}$ can be used to estimate ψ (otherwise, an independent estimate of ψ is needed).

SI Table 3: Summary of modeling assumptions.

Assumption	TDR	SIM	ASM	LDM
Conjugation events follow mass-action kinetics	X	X	X	X
The plasmid loss rate of the focal plasmid is zero	X	X	X	X
The cell populations do not change in size due to growth	X			
Processes of conjugation and growth are not resource dependent*	X		X	X
The cell populations grow exponentially (i.e., constant growth rate)			X	X
The growth rate is identical for all cell types	X	X		
The transconjugant conjugation rate is not high relative to the donor conjugation rate	X	X	X	

* The SIM model can incorporate resource-dependent growth and conjugation if (1) growth and transfer rates are homogeneous and (2) the functional form for resource dependence is the same for growth and transfer.

SI section 1b : Alternative laboratory forms for conjugation estimates

Often the conjugation estimates can be re-written into a form of the equation that is more amenable to laboratory implementation. Here we walk through rearranging the equations for a subset of the estimates.

For the SIM estimate, we start with equation [1.5]. If the entire period from $t = 0$ to $t = \tilde{t}$ involves exponential growth, then $N_{\tilde{t}} = N_0 e^{\psi \tilde{t}}$. In such a case, $\psi = (1/\tilde{t}) \ln(N_{\tilde{t}}/N_0)$. We arrive at the alternative laboratory form for SIM

$$\gamma = \frac{1}{\tilde{t}} \left[\ln \left(1 + \frac{T_{\tilde{t}}}{R_{\tilde{t}}} \frac{N_{\tilde{t}}}{D_{\tilde{t}}} \right) \right] \frac{\ln N_{\tilde{t}} - \ln N_0}{N_{\tilde{t}} - N_0}. \qquad [1.10]$$

We note that equation [1.9] is appropriate for some "truncated" versions of the SIM approach, but not generally applicable to the standard overnight version in which the culture does not grow exponentially across the entire assay.

To rearrange the ASM estimate, we start with equation [1.9]. While the equations $\psi_D = (1/\tilde{t}) \ln(D_{\tilde{t}}/D_0)$ and $\psi_R = (1/\tilde{t}) \ln(R_{\tilde{t}}/R_0)$ again provide estimates on donor and recipient growth rates, we cannot express the transconjugant growth rate (ψ_T) as a simple expression of time and initial/final densities of members of the mating culture. However, data from a transconjugant monoculture supply an estimate for this parameter. Thus, we arrive at the laboratory form for ASM

$$\gamma_D = \left\{ \frac{1}{\tilde{t}} (\ln D_{\tilde{t}} R_{\tilde{t}} - \ln D_0 R_0) - \psi_T \right\} \frac{T_{\tilde{t}}}{\left(D_{\tilde{t}} R_{\tilde{t}} - D_0 R_0 e^{\psi_T \tilde{t}} \right)}. \qquad [1.11]$$

To rearrange the LDM estimate, we start with equation [11]. Since $D_{\tilde{t}} = D_0 e^{\psi_D \tilde{t}}$ and $R_{\tilde{t}} = R_0 e^{\psi_R \tilde{t}}$, it is the case that $\psi_D = (1/\tilde{t}) \ln(D_{\tilde{t}}/D_0)$ and $\psi_R = (1/\tilde{t}) \ln(R_{\tilde{t}}/R_0)$. So, we have

$$\gamma_D = \frac{1}{\tilde{t}} \ln p_0(\tilde{t}) \frac{\ln\left(\frac{D_{\tilde{t}}}{D_0}\right) + \ln\left(\frac{R_{\tilde{t}}}{R_0}\right)}{D_0 R_0 - D_{\tilde{t}} R_{\tilde{t}}}$$

After rearrangement, we have

$$\gamma_D = \frac{1}{\tilde{t}} \{-\ln p_0(\tilde{t})\} \frac{\ln(D_{\tilde{t}} R_{\tilde{t}}) - \ln(D_0 R_0)}{D_{\tilde{t}} R_{\tilde{t}} - D_0 R_0}.$$

In the laboratory, we measure an estimate ($\hat{p}_0(\tilde{t})$) of the probability that a population has no transconjugants ($p_0(\tilde{t})$) which is simply the fraction of the populations (i.e., parallel cultures) that have no transconjugants at the incubation time $\tilde{t}$. In addition, if a 1 ml volume is not used for each mating culture (assuming that all cell densities are measured in $\mathrm{cfu/ml}$ units), then we must add a correction factor f (see SI section 5 for details and an example). Thus, we arrive at the laboratory form for the LDM, which is equation [13].

$$\gamma_D = \frac{f}{\tilde{t}} [-\ln \hat{p}_0(\tilde{t})] \frac{\ln D_{\tilde{t}} R_{\tilde{t}} - \ln D_0 R_0}{D_{\tilde{t}} R_{\tilde{t}} - D_0 R_0}$$

SI section 1c : Overview of laboratory implementations

In this section, we compare the laboratory implementations of the various estimates: TDR, SIM, and ASM. Each method is explained either as recommended by its authors or the most simplified protocol to acquire the information for the estimate. For each, we describe proper laboratory implementation for the approaches based on the model and derivation assumptions used to acquire the estimate. Note in this section, we do not explore the assumptions that are violated due to the biological entities being tested (i.e., specific species or plasmids) which can result in violations such as unequal conjugation rates or growth rates. These are explored in the main text and SI section 4 via stochastic simulations. Thus, we focus solely on the parameters under the experimenter's control. For ease of reference, key implementation differences are highlighted in SI Table 4.

The TDR estimate has a simple form (equation [1.4]). Donors and recipients are mixed in non-selective growth medium and incubated for a specified time $\tilde{t}$. Typically, densities after the incubation time are determined using selective plating. The derivation assumes the density in donors and recipients does not change due to growth which sets specific constraints on the implementation of this approach. In the original study, Levin *et. al.* used a chemostat to keep the population constant (10). Other studies shorten the incubation time $\tilde{t}$ such that population growth is negligible and use various laboratory tools to detect the small number of transconjugants (12, 13).

The SIM estimate is not built on an assumption of unchanging population densities. Donor and recipient populations in exponential phase are mixed in non-selective growth medium. The initial population density (N_0) is determined by dilution plating on non-

selective medium. After the mating mixture is incubated (for a period of $\tilde{t}$), the final densities ($D_{\tilde{t}}$, $R_{\tilde{t}}$, $T_{\tilde{t}}$, and $N_{\tilde{t}}$) are determined by dilution plating on selective and non-selective media. To implement SIM as written in equation [1.10] (see SI section 1b), the specified incubation time $\tilde{t}$ must occur well before stationary phase is reached to collect proper data for estimating the population growth rate ($\psi = (\ln N_{\tilde{t}} - \ln N_0)/\tilde{t}$). There is an alternative option for implementing the SIM using equation [1.5]. The donor and recipient populations are mixed and incubated under batch culture conditions (specifically exponential and stationary phase). However, the (maximum) population growth rate (ψ) needs to be determined with two additional samplings from the mixed population at times t_a and t_b, both occurring *within exponential phase:*

$$\psi = \frac{\ln (N_{t_b}/N_{t_a})}{t_b - t_a} \qquad [1.12]$$

The population densities N_{t_a} and N_{t_b} can be estimated either through colony counts from plating or optical density from a spectrophotometer. Either way, the timing of exponential phase is important for this approach and at least some analysis during this phase is required regardless of the implementation strategy.

For the ASM estimate, donor and recipient populations in exponential phase are mixed in non-selective medium. Initial densities (D_0 and R_0) are determined by plating dilutions on the appropriate selective media. After the donor-recipient co-culture incubates for a specified time ($\tilde{t}$), final densities ($D_{\tilde{t}}$, $R_{\tilde{t}}$, and $T_{\tilde{t}}$) are determined by plating dilutions on the appropriate selective media. From the transconjugant-selecting agar plates, a transconjugant clone is incubated in monoculture then sampled twice (at times t_a and t_b) in exponential phase to measure the transconjugant growth rate ($\psi_T = \ln\left(T_{t_b}/T_{t_a}\right)/(t_b - t_a)$). The authors point out a critical consideration for proper implementation of the ASM is the incubation time $\tilde{t}$. Not only is sampling in exponential phase important, but if the incubation time $\tilde{t}$ is too long and passes a critical time (t_{crit}) the approximations used to derive the ASM break down. To avoid overshooting t_{crit}, the authors recommend sampling as soon as measurable transconjugants arise. To determine that the incubation time used was below the critical time (t_{crit}), a second assay is recommended by the authors to measure the transconjugant conjugation rate γ_T, which will determine if the original incubation time $\tilde{t}$ was below t_{crit} for measuring the donor conjugation rate. This second assay would have the transconjugant clone become the donor in the mixture, while a newly marked recipient must be used so that donors and recipients can be distinguished using selective plating.

Each method has aspects of implementation in common. Each one shares the basic approach of mixing donors and recipients over some incubation time $\tilde{t}$. Each estimate requires reliable selectable markers to differentiate donors, recipients, and transconjugants. However, all estimates have some constraints on initial densities and time of measurement. This can occur because the experimenter needs to capture conjugation events (all estimates require this), avoid population growth (TDR), or keep growth exponential (ASM, and at least parts of SIM). Even so, each method has clear distinctions. The most notable is the incubation time $\tilde{t}$ (i.e., the end of the assay). The

TDR method is constrained to conditions where no change in population size due to growth can occur. For SIM, initial and final sampling are not constrained to a particular phase of growth; however, measurement of the growth rate must occur during the exponential growth phase. For ASM, initial sampling is in early exponential phase, and the final sampling needs to occur during a specific time window. In other words, the assay needs to be long enough that measurable transconjugants appear, but short enough so that assumptions are not violated (which can occur if transconjugant density becomes too large).

SI Table 4: Comparison of implementations.

Summary	**TDR**	**SIM**	**ASM**	**LDM**
Assay conditions minimizing the change in density due to growth	X			
Minimize incubation time necessary for producing transconjugants			X	
An incubation time results in a subset of parallel populations having no transconjugants				X
Assay occurs over a period of exponential cell growth		X*	X	X
Assay requires multiple parallel mating cultures to obtain one estimate				X
Assay requires a measurement of transconjugant density	X	X	X	
Assay requires a measurement of population growth rate		X*		
Assay requires a measurement of transconjugant growth rate			X	
* For the SIM assay, either the entire assay is conducted over exponentially growing cultures or an independent estimate for (maximum) population growth rate is needed.				

SI section 2 : Derivation of $p_0(\tilde{t})$ for the LDM estimate

In this section, we will continue to assume an experimental volume of 1 ml for the co-culture such that the density of cells per ml is equivalent to the cell count numerically. We will not explicitly track units in this section, but we deal with the case of an arbitrary experimental volume in SI section 5.

We define $p_n(t)$ to be the probability that there are n transconjugants at time t, where n is a non-negative integer (i.e., $p_n(t) = \Pr\{T_t = n\}$). We focus here on the probability that transconjugants are absent (namely, where $n = 0$) and derive an expression for $p_0(t)$. By definition $p_0(t + \Delta t) = \Pr\{T_{t+\Delta t} = 0\}$. However, $T_{t+\Delta t} = 0$ implies $T_t = 0$, so we can write

$$p_0(t+\Delta t) = \Pr\{(T_{t+\Delta t}=0) \cap (T_t=0)\} = \Pr\{T_{t+\Delta t}=0 \mid T_t=0\}\Pr\{T_t=0\},$$

Given that $p_0(t) = \Pr\{T_t=0\}$, we can use equation [9] to write the following time-increment recursion for $p_0(t)$:

$$p_0(t+\Delta t) = (1-\gamma_D D_t R_t \Delta t) p_0(t).$$

This can be simplified as follows

$$\frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = -\gamma_D D_t R_t p_0(t).$$

Taking the limit as $\Delta t \to 0$ gives

$$\lim_{\Delta t \to 0} \frac{p_0(t+\Delta t) - p_0(t)}{\Delta t} = \frac{dp_0(t)}{dt}.$$

Therefore, we have the following differential equation:

$$\frac{dp_0(t)}{dt} = -\gamma_D D_t R_t p_0(t).$$

We are assuming $D_t = D_0 e^{\psi_D t}$ and $R_t = R_0 e^{\psi_R t}$. We note that these assumptions are reasonable if the densities of donors and recipients are reasonably large and the rate of transconjugant generation per recipient cell ($\gamma_D D_t + \gamma_T T_t$, or if $T_t = 0$, simply $\gamma_D D_t$) remains very small relative to per capita recipient growth rate (ψ_R). Under these assumptions, our differential equation becomes:

$$\frac{dp_0(t)}{dt} = -\gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} p_0(t).$$

We solve this differential equation via separation of variables, integrating from 0 to our incubation time of interest $\tilde{t}$:

$$\int_0^{\tilde{t}} \frac{dp_0(t)}{p_0(t)} = \int_0^{\tilde{t}} -\gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} dt,$$

$$\ln p_0(t)\Big|_0^{\tilde{t}} = \frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R} e^{(\psi_D+\psi_R)t}\Big|_0^{\tilde{t}},$$

$$\ln p_0(\tilde{t}) - \ln p_0(0) = \frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R} e^{(\psi_D+\psi_R)\tilde{t}} - \frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R}.$$

Given that $p_0(0) = 1$,

$$\ln p_0(\tilde{t}) = \frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right),$$

$$p_0(\tilde{t}) = \exp\left\{\frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\},$$

which is equation [10].

SI section 3 : Derivation of mutation rate from the Luria-Delbrück experiment

Here we derive the classic estimate of mutation rate from Luria and Delbrück. We assume that there is a population of wild-type cells that grow according to the following equation:

$$N_t = N_0 e^{\psi_N t}, \qquad [3.1]$$

where N_t is the number of wild type cells at time t and ψ_N is the per capita growth rate. The wild-type population dynamics are assumed to be deterministic (a reasonable assumption if the initial population size is reasonably large, i.e., $N_0 \gg 0$). We are also ignoring the loss of wild-type cells to mutational transformation, but this omission is reasonable if the mutation rate is very small relative to per capita wild-type growth rate.

Let the number of mutants be given by a random variable M_t. This variable takes on non-negative integer values. For a very small interval of time, Δt, the current number of mutants will either increase by one or remain constant. The probabilities of each possibility are given as follows:

$$\Pr\{M_{t+\Delta t} = M_t + 1\} = \mu N_t \Delta t + \psi_M M_t \Delta t, \qquad [3.2]$$

$$\Pr\{M_{t+\Delta t} = M_t\} = 1 - (\mu N_t + \psi_M M_t)\Delta t. \qquad [3.3]$$

The two terms on the right-hand side of equation [3.2] give the ways that a mutant can be generated. The first term measures the probability that a wild-type cell undergoes a mutation (μ is the mutation rate). The second term gives the probability that a mutant cell divides and produces two mutant cells (ψ_M is the mutant growth rate). Equation [3.3] is the probability that neither of these processes occur.

Analogous to the procedure in SI section 2 (with $p_0(t) = \Pr\{M_t = 0\}$):

$$p_0(t + \Delta t) = (1 - \mu N_t \Delta t) p_0(t).$$

By rearranging, taking the limit as $\Delta t \to 0$, and utilizing equation [3.1], we have

$$\frac{dp_0(t)}{dt} = -\mu N_0 e^{\psi_N t} p_0(t).$$

This differential equation can be solved in an analogous way as well

$$\int_0^{\tilde{t}} \frac{dp_0(t)}{p_0(t)} = \int_0^{\tilde{t}} -\mu N_0 e^{\psi_N t} dt,$$

$$\ln p_0(t)|_0^{\tilde{t}} = \frac{-\mu N_0}{\psi_N} e^{\psi_N t}\Big|_0^{\tilde{t}},$$

$$\ln p_0(\tilde{t}) - \ln p_0(0) = \frac{-\mu N_0}{\psi_N} e^{\psi_N \tilde{t}} - \frac{-\mu N_0}{\psi_N}.$$

Because we assume $M_0 = 0$, we must have $p_0(0) = 1$, and

$$\ln p_0(\tilde{t}) = \frac{-\mu N_0}{\psi_N}\left(e^{\psi_N \tilde{t}} - 1\right).$$

Solving for the mutation rate μ

$$\mu = -\ln p_0(\tilde{t}) \frac{\psi_N}{N_0\left(e^{\psi_N \tilde{t}} - 1\right)},$$

which is equation [12]. This equation can also be expressed as:

$$\mu = -\ln p_0(\tilde{t}) \frac{\psi_N}{N_{\tilde{t}} - N_0}.$$

To recover Luria and Delbrück's original formulation, consider a new time variable z defined as follows:

$$z = \frac{t}{t_d/\ln 2},$$

where t_d is the period required for population doubling during exponential growth. Because

$$N_0 e^{\psi_N t} = N_0 2^{t/t_d},$$

it is the case that

$$\psi_N = \frac{1}{t_d/\ln 2}.$$

Therefore, $z = \psi_N t$ and the equation $N_t = N_0 e^{\psi_N t}$ can be expressed as

$$N_z = N_0 e^z$$

Performing the same analysis on this new equation gives the original formulation (their equations [4] and [5], where our μ is given by their "a" and our $\tilde{z}$ is given by their "t"):

$$\mu = \frac{-\ln p_0(\tilde{z})}{N_{\tilde{z}} - N_0}.$$

SI section 4 : Extended Simulation Results

SI section 4a : Extended stochastic simulation methods

To systematically explore the effects of heterogeneous growth and conjugation rates (as well as non-zero rates of plasmid loss) on the accuracy and precision of estimating the donor conjugation rate (γ_D), we developed a stochastic simulation framework using the Gillespie algorithm. We ran sets of simulations sweeping through parameter values. Each simulation examined a biological process (i.e., growth, conjugation) in isolation by manipulating one or two of the relevant parameters. For SI section 4b-d, we used a "baseline" set of parameters ($\psi_D = \psi_R = \psi_T = 1$, and $\gamma_D = \gamma_T = 1 \times 10^{-6}$) and initial densities ($D_0 = R_0 = 1 \times 10^2$ and $T_0 = 0$) unless otherwise indicated. For each initial parameter setting, we simulated 10,000 parallel populations and calculated the conjugation rate using various methods (TDR, SIM, ASM, and LDM). The incubation time selection criteria used for the SIM estimate was also used for the TDR and ASM estimates (see Materials and Methods and SI Table 5). However, given that all our simulated populations *increase in size* over the incubation time, a fundamental assumption of the TDR approach is broken for all the runs (i.e., no change in the density due to growth). The TDR estimate was included to be comprehensive (and illustrate that violation of the no growth assumption leads to systemic bias). Also, we note that we calculated the ASM metric in all scenarios and that in some cases the incubation time $\tilde{t}$ passed the critical time threshold (t_{crit}) where the ASM assumptions break down (see SI section 1a). The ASM estimate was included in all scenarios to be comprehensive and illustrate that implementing the assay after t_{crit} can lead to bias. Given that the chosen incubation time $\tilde{t}$ to evaluate these simulations is early, it highlights that for some parametric combinations proper implementation of the ASM metric is not possible. Given

that the Gillespie algorithm is computationally expensive and the large number of simulations needed to sweep through parameters, we chose low initial densities and high conjugation rates for the baseline condition. In SI section 4e, we demonstrate that the trends shown for the baseline condition are also observed with more realistic parameter values and higher initial densities.

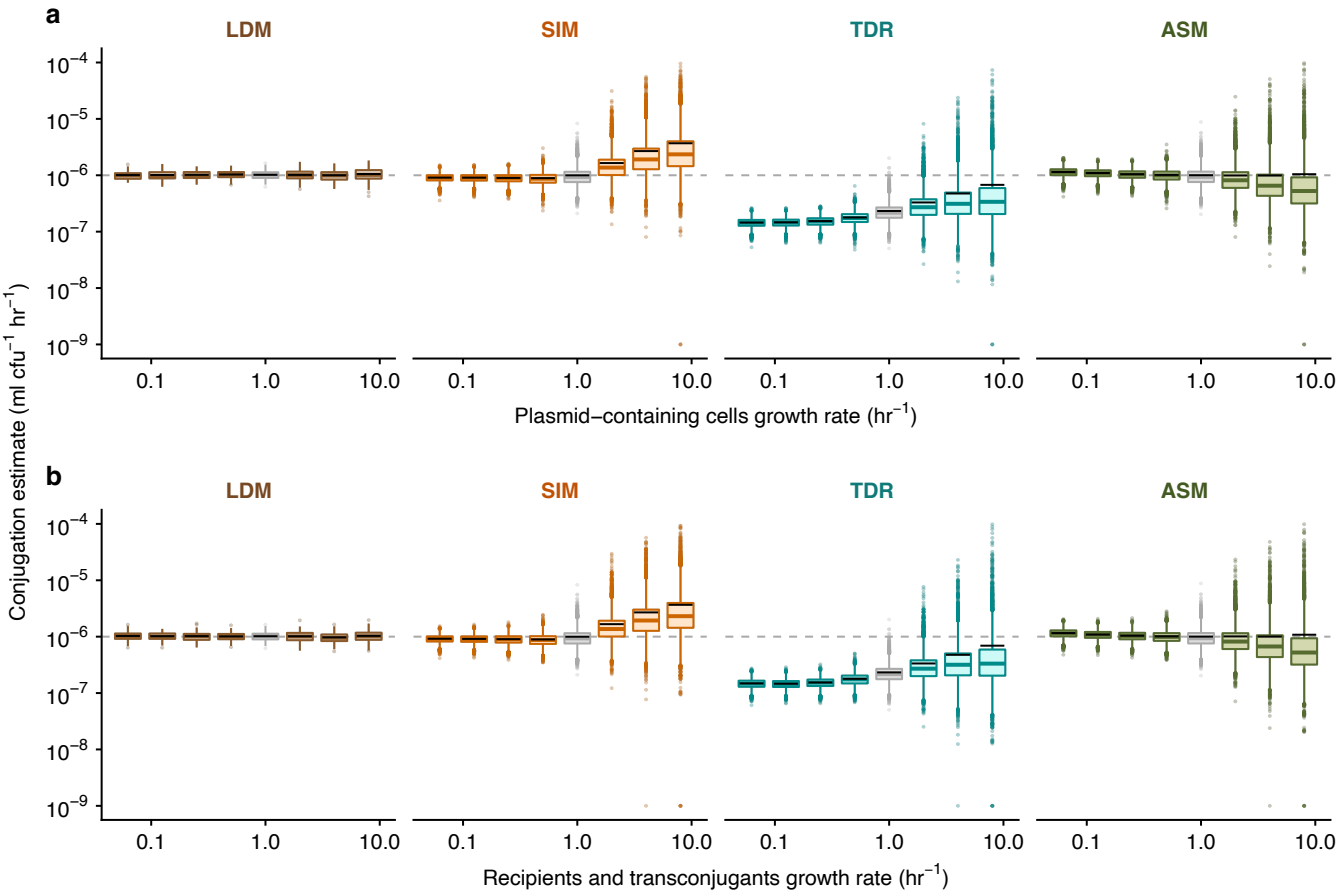
SI Table 5: Specific incubation times ($\tilde{t}$) used in stochastic simulations to compare across parameter settings. Each row lists the relevant figure and the corresponding x-axis value. Time is given in hours. For each parameter setting, the incubation time $\tilde{t}$ for the LDM estimate is set to the average t^*, and for the SIM estimate is given by the time point for which an average of 50 transconjugants is reached.

Figure	**x-axis value**	$\tilde{t}_{\mathrm{LDM}}$	$\tilde{t}_{\mathrm{SIM}}$
Figure 4, SI Figure 1a	0.0625	4.34	7.89
Figure 4, SI Figure 1a	0.125	4.11	7.49
Figure 4, SI Figure 1a	0.25	3.7	6.78
Figure 4, SI Figure 1a	0.5	3.1	5.67
Figure 4, SI Figure 1a	1	2.35	4.27
Figure 4, SI Figure 1a	2	1.61	2.86
Figure 4, SI Figure 1a	4	1.01	1.74
Figure 4, SI Figure 1a	8	0.6	1
Figure 4, SI Figure 3	1×10^9	2.35	4.27
Figure 4, SI Figure 3	1×10^8	2.35	4.27
Figure 4, SI Figure 3	1×10^7	2.35	4.27
Figure 4, SI Figure 3	1×10^6	2.35	4.27
Figure 4, SI Figure 3	1×10^5	2.35	4.25
Figure 4, SI Figure 3	1×10^4	2.33	4.11
Figure 4, SI Figure 3	1×10^3	2.16	3.4
Figure 4, SI Figure 3	1×10^2	1.44	2.02
SI Figure 1b	0.0625	4.35	8.18
SI Figure 1b	0.125	4.11	7.66
SI Figure 1b	0.25	3.71	6.85
SI Figure 1b	0.5	3.1	5.69
SI Figure 1b	1	2.35	4.27
SI Figure 1b	2	1.61	2.86
SI Figure 1b	4	1.01	1.74
SI Figure 1b	8	0.6	0.99
SI Figure 2a	0.0625	3.3	6.4
SI Figure 2a	0.125	3.22	6.23
SI Figure 2a	0.25	3.07	5.89

SI Figure 2a	0.5	2.8	5.26
SI Figure 2a	1	2.35	4.27
SI Figure 2a	2	1.78	3.07
SI Figure 2a	4	1.2	1.99
SI Figure 2a	8	0.74	1.15
SI Figure 2b	0.0625	3.31	6.45
SI Figure 2b	0.125	3.23	6.27
SI Figure 2b	0.25	3.07	5.92
SI Figure 2b	0.5	2.8	5.27
SI Figure 2b	1	2.35	4.27
SI Figure 2b	2	1.78	3.07
SI Figure 2b	4	1.2	1.97
SI Figure 2b	8	0.74	1.15
SI Figure 2c	0.0625	2.64	4.59
SI Figure 2c	0.125	2.62	4.57
SI Figure 2c	0.25	2.59	4.54
SI Figure 2c	0.5	2.52	4.46
SI Figure 2c	1	2.35	4.27
SI Figure 2c	2	1.97	3.62
SI Figure 2c	4	1.34	2.31
SI Figure 2c	8	0.8	1.29
SI Figure 4	0.00001	2.35	4.27
SI Figure 4	0.0001	2.35	4.27
SI Figure 4	0.001	2.35	4.27
SI Figure 4	0.01	2.36	4.29
SI Figure 4	0.1	2.47	4.49

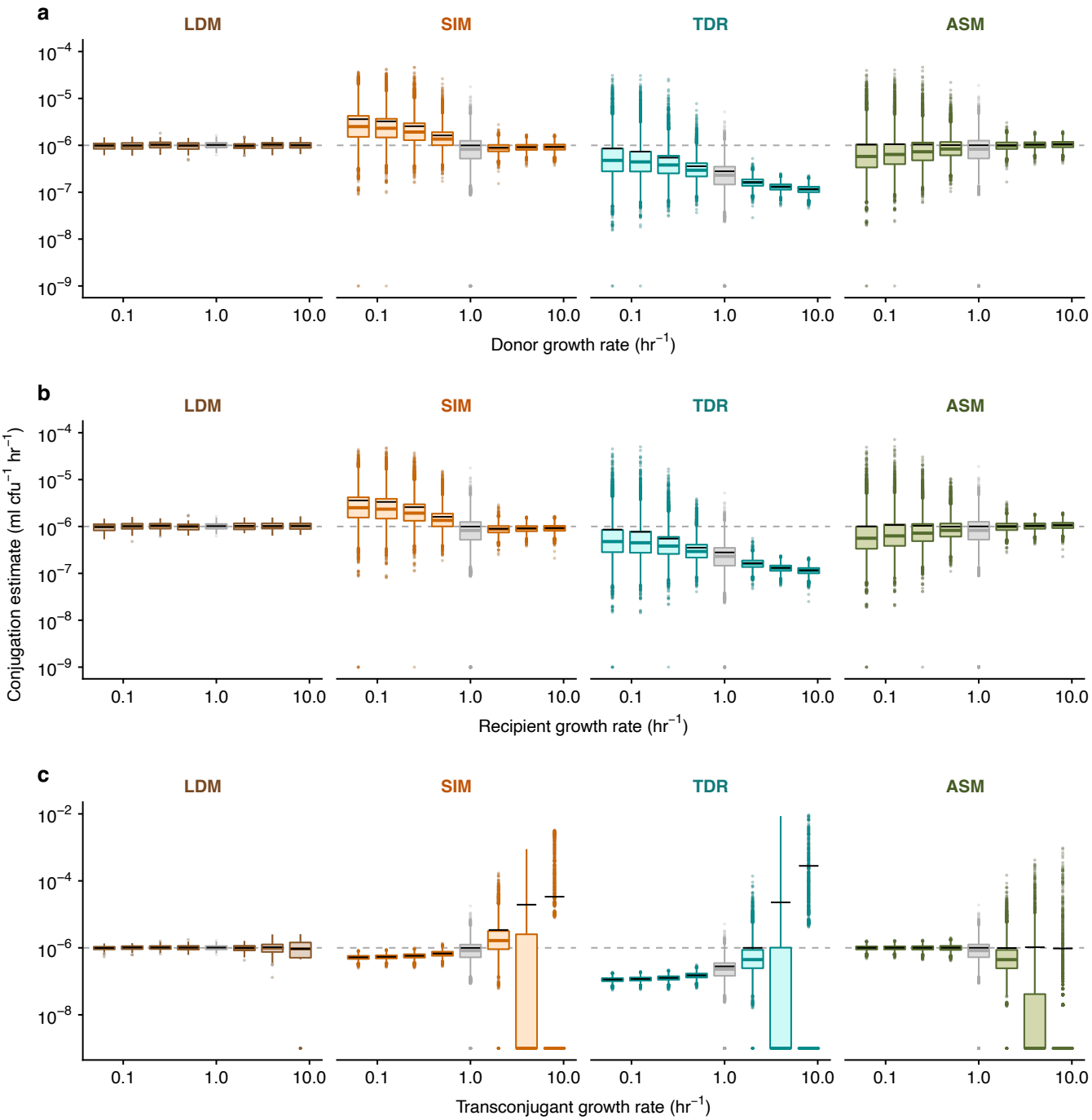
SI section 4b : The effect of unequal growth rates

We expanded the analysis used in Figure 4a by calculating a conjugation rate estimate with two additional estimates, TDR and ASM (SI Figure 1a). We simulated an additional biological scenario (SI Figure 1b) in which growth rates differ due to the host; i.e., where recipients and transconjugants grow faster ($\psi_D < \psi_T = \psi_R$) or slower ($\psi_D > \psi_T = \psi_R$) than the donors. This captures the situation in which the recipient (and therefore transconjugant) is a different strain or species from the donor and differs in growth rate. Like the conclusions drawn from Figure 4 with the effects of plasmid carriage, the LDM exhibited high accuracy and precision relative to other metrics.



SI Figure 1: The effect of heterogenous growth rates on estimating conjugation rate. The Gillespie algorithm was used to simulate population dynamics. 100 estimates of the donor conjugation rate are shown for each parameter combination (summarized using boxplots with the same graphical convention as in Figure 3). The gray dashed line indicates the true value for the donor conjugation rate (here, 10^{-6}). The boxes in gray indicate the baseline parameter setting, and all colored boxes represent deviation of one or two parameters from baseline. The baseline parameter values were $\psi_D = \psi_R = \psi_T = 1$, and $\gamma_D = \gamma_T = 10^{-6}$. The dynamic variables were initialized with $D_0 = R_0 = 10^2$ and $T_0 = 0$. All incubation times are short but are specific to each parameter setting (see Materials and Methods and SI Table 5 for details). The LDM, SIM, TDR, and ASM estimates are in separate plots with estimate specific colors (brown, orange, cyan, and green, respectively). Zero estimates were set to 10^{-9} (the lowest y-value) for plotting on a log axis. (a) Unequal growth rates were explored over a range of growth rates for the plasmid-bearing strains, namely $\psi_D = \psi_T \in \{0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8\}$. (b) Unequal growth rates were explored over a range of growth rates for the recipients and transconjugants, namely $\psi_R = \psi_T \in \{0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8\}$. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

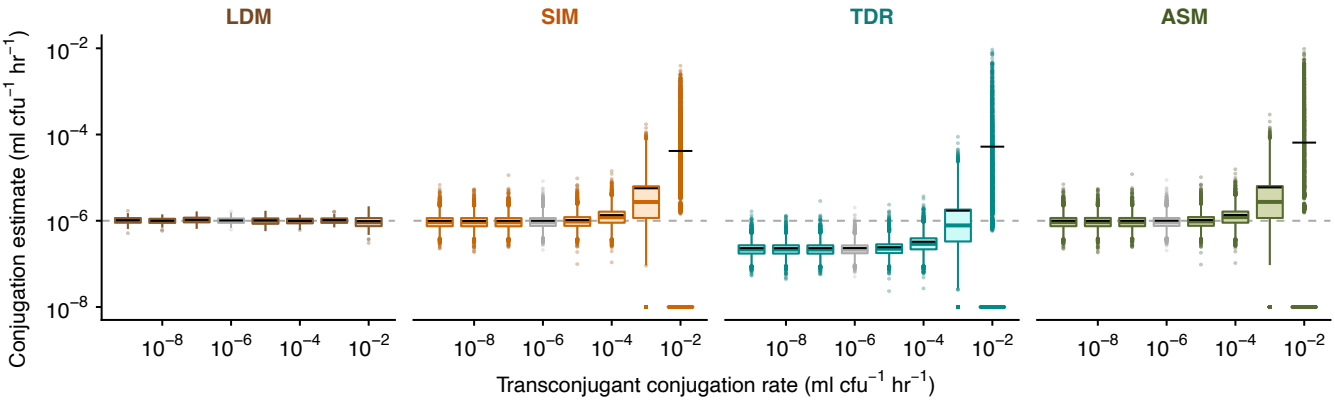
In addition, we explored the effect of a single population (donor, recipient, and transconjugant) growing faster or slower in isolation (SI Figure 2). Notably, a faster transconjugant growth rate led to very large variance with the other metrics (TDR, SIM, and ASM). Therefore, some parameter settings shown in SI Figure 2c have a large proportion of the simulations yielding a zero estimate at the specific incubation time. Given the log-axis, the zero estimates were placed at the lowest y-value for plotting purposes.



SI Figure 2: The effect of population-specific heterogenous growth rates on estimating conjugation rate. Boxplots are using the same graphical representation as SI Figure 1. (a, b, c) Unequal growth rates were explored over a range of growth rates for the donors, recipients, and transconjugants, respectively, namely $\psi_X \in$ {0.0625, 0.125, 0.25, 0.5, 1, 2, 4, 8}. Zero estimates were set to 10^{-9} for plotting on a log axis. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 4c : The effect of unequal conjugation rates

We expanded the analysis used in Figure 4b by calculating a conjugation rate estimate with two additional estimates, TDR and ASM (SI Figure 3). Like the conclusions drawn from Figure 4b with the effects of heterogenous conjugation rate, the LDM exhibited high accuracy and precision relative to other metrics.



SI Figure 3: The effect of heterogenous conjugation rates on estimating conjugation rate. Boxplots are using the same graphical representation as SI Figure 1. Unequal conjugation rates were probed over a range of transconjugant conjugation rates, namely $\gamma_T \in$ {10^{-9}, 10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}}. Zero estimates were set to 10^{-8} for plotting on a log axis. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 4d : The effect of a non-zero plasmid loss rate

We extended the base model (equations [1] - [3]) to include plasmid loss due to improper plasmid segregation. Thus, transconjugants are transformed into plasmid-free recipients due to improper segregation of the plasmid at rate τ_T. The donors are transformed into plasmid-free cells due to improper segregation of the plasmid at rate τ_D. Therefore, the extended model (equations [4.1] - [4.4]) tracks the change in density of a

new population type, plasmid-free *former* donors (F). In total, the extended model describes the change in density of four populations (D, R, T, and F) due to various biological parameters: growth rates (ψ_D, ψ_R, ψ_T, and ψ_F), conjugation rates (γ_{DR}, γ_{TR}, γ_{DF}, and γ_{TF}), and plasmid loss rates (τ_D and τ_T). Importantly, we note that all conjugation rates are dyad-specific (i.e., donor-recipient-specific); therefore, our simulation framework is built to allow all rates to be unique. Since the new population type is a possible plasmid recipient, the subscript on the conjugation rate parameter now indicates the plasmid-bearing cell type and the plasmid-free cell type (e.g., γ_{TF} indicates the conjugation rate between a transconjugant and a plasmid-free *former* donor).

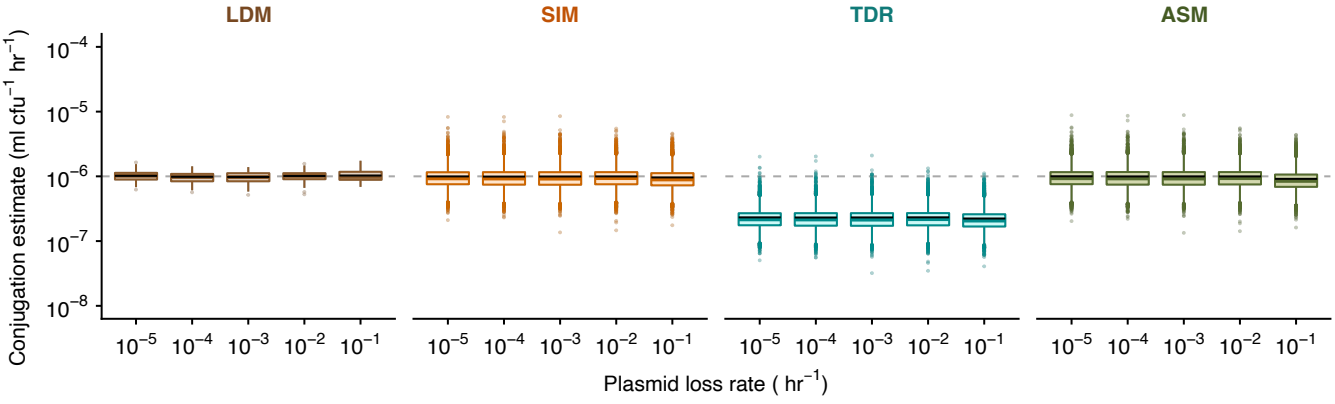
$$\frac{dD}{dt} = \psi_D D + (\gamma_{DF} D + \gamma_{TF} T)F - \tau_D D, \quad [4.1]$$

$$\frac{dR}{dt} = \psi_R R - (\gamma_{DR} D + \gamma_{TR} T)R + \tau_T T, \quad [4.2]$$

$$\frac{dT}{dt} = \psi_T T + (\gamma_{DR} D + \gamma_{TR} T)R - \tau_T T. \quad [4.3]$$

$$\frac{dF}{dt} = \psi_F F - (\gamma_{DF} D + \gamma_{TF} T)F + \tau_D D. \quad [4.4]$$

Plasmid loss due to improper segregation is a common occurrence in plasmid populations and violates a model assumption underlying all the conjugation rate estimates. We simulated a range of plasmid loss rates, ranging from low ($\tau_D = \tau_T = 0.0001$) to high ($\tau_D = \tau_T = 0.1$). The LDM had high accuracy and precision across all parameter settings (SI Figure 4). The effect of plasmid loss was undetectable even for an extremely high loss rate ($\tau_D = \tau_T = 0.1$). Similarly, the effect of plasmid loss was undetectable on the other conjugation estimates compared to their performance with a zero loss rate. Thus, we find that all estimates appear robust with regards to an introduction of plasmid loss.



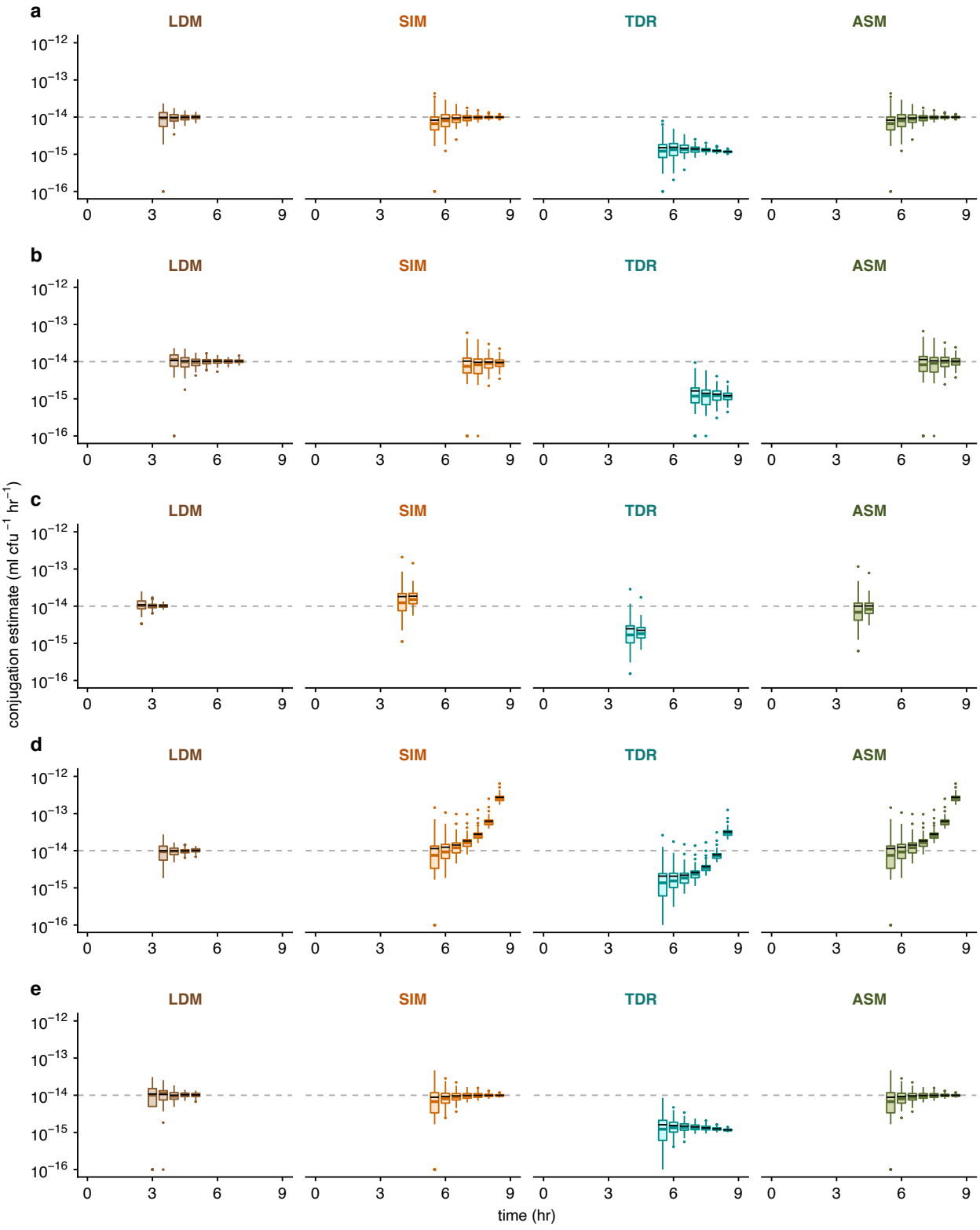
SI Figure 4 : The effect of non-zero plasmid loss rates on estimating conjugation rate. Boxplots are using the same graphical representation as SI Figure 1. We explored improper plasmid segregation by considering a range of plasmid loss rates $\tau_D = \tau_T \in$ {0.00001, 0.0001, 0.001, 0.01, 0.1}. The data and code needed to generate this Figure

can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 4e : The effect of incubation time using realistic parameter settings

We expanded the analysis used in Figure 3 by calculating conjugation rate with two additional estimates, TDR and ASM. In addition, we explored the effects of incubation time in conjunction with other heterogenous parameter settings and a non-zero plasmid loss rate using realistic parameter settings. Given the computational expense of using realistic parameter values and higher initial densities, we explored five parameter combinations and the results are summarized in SI Figure 5. We set more reasonable initial densities of the donors and recipients ($D_0 = R_0 = 1 \times 10^5$) and a conjugation rate that is often reported in the literature ($\gamma_{DR} = \gamma_{DF} = \gamma_{TR} = \gamma_{TF} = 1 \times 10^{-14}$) unless otherwise indicated. The conjugation rate was estimated for each method at 30-minute intervals. For each time interval, we applied estimate-specific filters. For the LDM estimate, a 30-minute interval was shown if at least one parallel population had zero transconjugants. For the other estimates (SIM, TDR, and ASM), the 30-minute interval were shown if at least 90 percent of the simulated populations contained transconjugants at the incubation time.

The LDM estimate had high accuracy over all incubation times for all scenarios with precision increasing through time for the range explored. The other estimates also become more precise over time. However, their greater precision over time was sometimes accompanied by decreased accuracy. We note these inaccuracies recaptured the qualitative patterns revealed in the parameter sweeps. Again, the LDM estimate performed as well or better than other estimates across incubation times.



SI Figure 5 : The effect of incubation time ($\tilde{t}$) on estimating conjugation rate. The Gillespie algorithm with equations [4.1]-[4.4] was used to simulate population dynamics. Donor conjugation rate for each parameter combination was estimated at 30-minute intervals (summarized using boxplots with the same graphical convention as in Figure 3). The gray dashed line indicates the true value for the donor conjugation rate (here, 10^{-14}). The baseline parameter values were $\psi_D = \psi_R = \psi_T = \psi_F = 1$, $\gamma_{DR} = \gamma_{DF} = \gamma_{TR} =$

$\gamma_{TF} = 1 \times 10^{-14}$, and $\tau_D = \tau_T = 0$. The dynamic variables were initialized with $D_0 = R_0 = 10^5$ and $T_0 = F_0 = 0$. The LDM, SIM, TDR, and ASM estimates are in separate plots with estimate-specific colors (brown, orange, cyan, and green, respectively). (a) Baseline parameters were simulated as the non-heterogenous parameter comparison. (b) An unequal growth rate was simulated with $\psi_D = \psi_T = 0.5$. (c) An unequal growth rate was simulated with $\psi_R = \psi_T = 2$. (d) An unequal conjugation rate was simulated with $\gamma_{TR} = 10^{-8}$. (e) A non-zero plasmid loss rate was simulated with $\tau_D = \tau_T = 0.0001$. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 4f : Modified Levin et. al. model with Monod growth and conjugation

To investigate the incongruency observed between the SIM and LDM estimates for the cross-species mating assay in Figure 6, we extend equations [4.1]-[4.4] to incorporate batch culture dynamics by tracking the change in resource concentration:

$$\frac{dD}{dt} = \psi_D(C)D + \gamma_{DF}(C)DF + \gamma_{TF}(C)TF - \tau_D(C)D, \quad [4.5]$$

$$\frac{dR}{dt} = \psi_R(C)R - \gamma_{DR}(C)DR - \gamma_{TR}(C)TR + \tau_T(C)T, \quad [4.6]$$

$$\frac{dT}{dt} = \psi_T(C)T + \gamma_{DR}(C)DR + \gamma_{TR}(C)TR - \tau_T(C)T, \quad [4.7]$$

$$\frac{dF}{dt} = \psi_F(C)F - \gamma_{DF}(C)DF - \gamma_{TF}(C)TF + \tau_D(C)D, \quad [4.8]$$

$$\frac{dC}{dt} = -(\psi_D(C)D + \psi_R(C)R + \psi_T(C)T + \psi_F(C)F)e. \quad [4.9]$$

where e is the amount of resource required to produce a new cell. With the addition of a resource equation, there is an added assumption that growth, conjugation, and plasmid loss are Monod functions of resource concentration C:

$$\psi_X(C) = \psi_{X_{max}}\left(\frac{C}{Q + C}\right), \quad [4.10]$$

$$\gamma_{XY}(C) = \gamma_{XY_{max}}\left(\frac{C}{Q + C}\right), \quad [4.11]$$

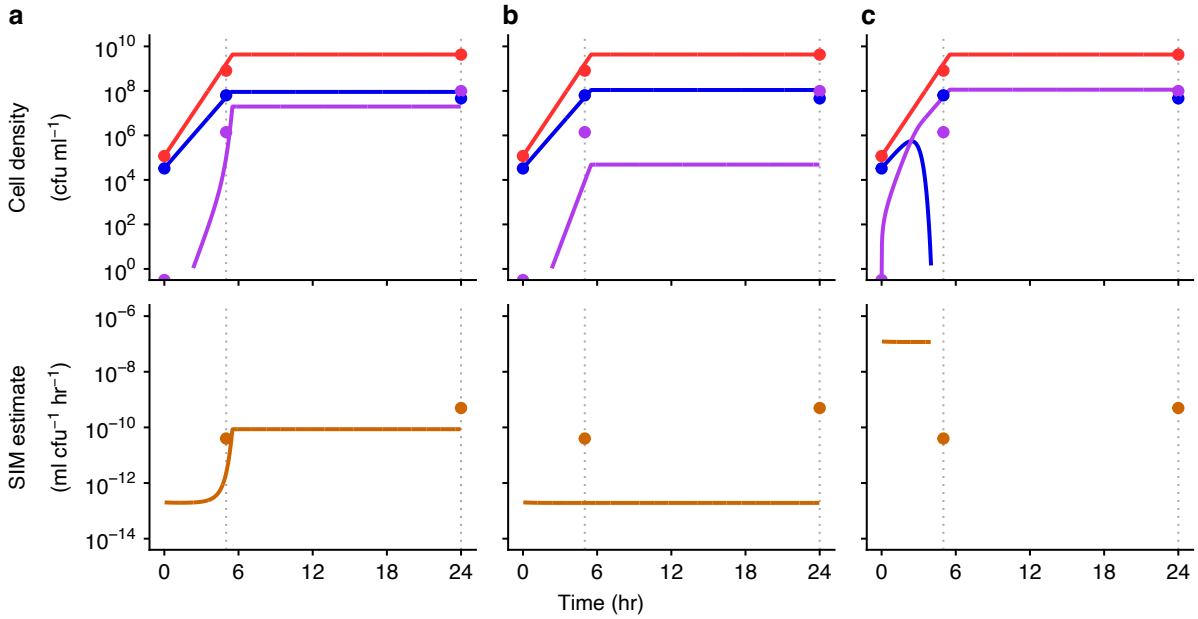
$$\tau_X(C) = \tau_{X_{max}}\left(\frac{C}{Q + C}\right), \quad [4.12]$$

where Q is the half saturation constant, and $\psi_{X_{max}}$, $\gamma_{XY_{max}}$, and $\tau_{X_{max}}$ are the maximum growth, conjugation, and plasmid loss rates for relevant cell types X and Y, respectively. In other words, growth, conjugation, and plasmid loss decline and eventually turn off as resource concentration goes to zero.

SI section 4g : Deterministic simulations with the Monod model using cross-species case study parameters

Here, we used equations [4.5]-[4.12] that incorporate batch culture dynamics to simulate the cross-species case study with the experimental parameters to investigate the incongruency observed between the SIM and LDM estimates for the cross-species mating assay in Figure 6. Most of the parameters were from the average of six experiments (D_0 = 1.17 x 10^5, R_0 = 3.33 x 10^4, ψ_D = 1.91, ψ_R = 1.47, ψ_T = 1.48, γ_{DR} = 1.96 x 10^{-13}, and γ_{TR} = 1.96 x 10^{-7}) with the remaining parameters informed by the 24 hour densities as to mimic the batch culture conditions of the experiment (C_0 = 4.41 x 10^9, Q = 1 x 10^7, and e = 1). We used the numerical solution to calculate the SIM estimate over time.

We compared the numerical solution to the actual experimental measurements from the cross-species experiments. The simulated density and conjugation estimate (SI Figure 6a solid lines) were similar to the average experimental densities and the experimental SIM estimate (SI Figure 6a circle data points). Thus, the experimental LDM estimates for the cross-species (γ_{DR} = 1.96 x 10^{-13}) and within-species (γ_{TR} = 1.96 x 10^{-7}) conjugation rates along with the measured growth rates are sufficient to recapture a relatively inflated experimental SIM estimate. In contrast, a simulation with homogenous conjugation rates using either the cross- or within-species conjugation rate does not closely align with the experimental data (SI Figure 6b and c, respectively). These simulations also demonstrate that the heterogeneity in the measured growth rates is insufficient to produce the mismatch observed in the experimental data (SI Figure 6b and c). This was worth checking given that heterogeneity in growth rates violates a modeling assumption of the SIM approach. This adds further support that the parametric heterogeneity (i.e., $\gamma_D \neq \gamma_T$) in the conjugation rates is the potential cause for the incongruency between the LDM and SIM estimates reported in Figure 6.

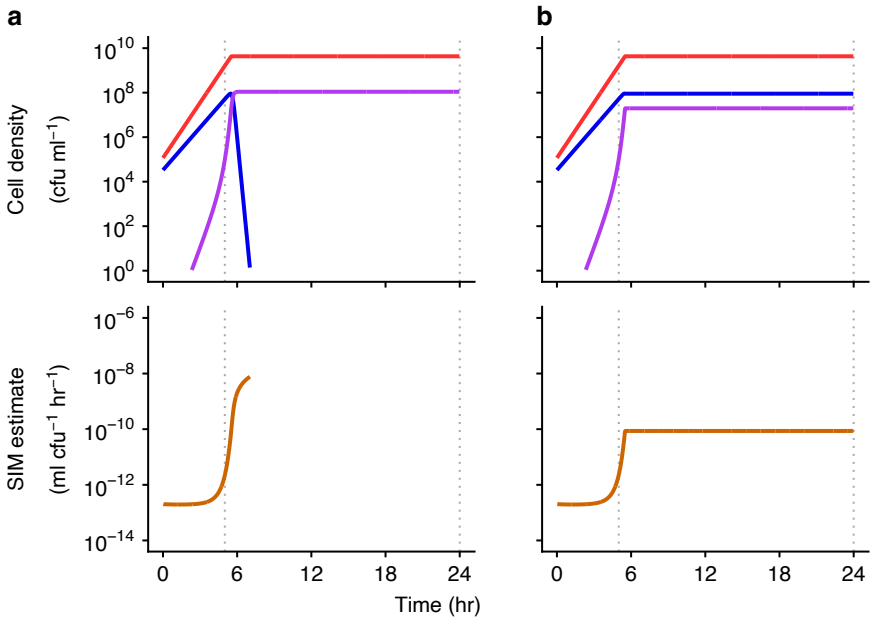


SI Figure 6 : Numerical simulation of extended model with Monod functions using experimental parameters. Deterministic numerical solutions of equations [4.5]-[4.12] showing donor, recipient, and transconjugant densities (red, blue, and purple solid trajectories, respectively) increasing over time using experimental parameter estimates (D_0 = 1.17 x 10^5, R_0 = 3.33 x 10^4, ψ_D = 1.91, ψ_R = 1.47, ψ_T = 1.48, γ_{DR} = 1.96 x 10^{-13}, and γ_{TR} = 1.96 x 10^{-7}) and batch culture parameters (C_0 = 4.41 x 10^9, Q = 1 x 10^7, and e = 1) unless otherwise indicated. The averaged experimental data is overlayed onto each part (circle data points) at both incubation times (grey dotted line at t = 5 and t = 24). (a) The numerical solution with the experimental parameter estimates were close to the experimental measurements. (b) A scenario with homogenous low conjugation rates (γ_{DR} = γ_{TR} = 1.96 x 10^{-13}) deviates markedly from the experimental measurements. (c) A scenario with homogenous high conjugation rates (γ_{DR} = γ_{TR} = 1.96 x 10^{-7}) deviates substantially from the experimental measurements. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 4h : Violation of the Levin et. al. model Monod equation assumptions

In this section, we explored a violation of a modeling assumption in the SIM approach by using a model variation where the functional form of the growth rates and conjugation rates are not proportional. This is relevant given that there are plasmid systems that will readily violate this proportional assumption (e.g., IncP plasmids). Here, we assume that while growth rates follow the Monod equation, conjugation rates are not dependent on resource and remain constant after resources are depleted. We found that using this model and the same experimentally measured parameter values in SI Figure 6

resulted in a higher SIM estimate as the culture enters stationary phase (SI Figure 7a) compared to the scenario where conjugation rates are proportional to growth rates (SI Figure 7b). It is worth noting that by using this new model and these particular parameter values, the recipient pool is completely depleted which coincides with the SIM estimate no longer being a finite, positive value. This differs from SI Figure 7b where the SIM estimate hits an asymptote remaining at a finite, positive value. In this case, the recipient pool is not depleted because in this version of the model (SI section 4f) the conjugation rates approach zero as the resources are depleted. We acknowledge that a violation of the proportional assumption would lead to an inflation of the SIM estimate, which is the same pattern we show in our experimental results in Figure 6. However, we used an IncF plasmid in our experiment which was the plasmid system used in the original SIM study where the experimental results were consistent with a proportional relationship. We note that this analysis is relevant to other plasmid systems where this assumption is known to be violated or has not been experimentally validated.



SI Figure 7: Numerical simulation of a modified model with constant conjugation rates with Monod functions for growth. The same equations and parameters from SI Figure 6 are used throughout unless otherwise indicated. (a) A model modification is made where conjugation rates are no longer proportional to growth rates. Specifically, conjugation rates are constant (i.e., not resource dependent). (b) The same panel in SI Figure 6a for comparison. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 5 : Experimental volume unit conversion using f

In this section, we walk through the addition of f to the LDM estimate. This is important to maintain the typical units $\mathrm{ml}/(\mathrm{h}\cdot\mathrm{cfu})$ used for reporting the conjugation rates. In the original differential equations [1]-[3], the units of the dynamic variables were cfu/ml. If we want to deal with numbers instead of density, the let us define a new volume unit termed the "evu" standing for "experimental volume unit" where we will assume there are f evu's per ml. Focusing on the number of donors in the experiment (which we label $\breve{D}$), we have the following conversion:

$$\breve{D}\left(\frac{\mathrm{cfu}}{\mathrm{evu}}\right)=\frac{D\left(\frac{\mathrm{cfu}}{\mathrm{ml}}\right)}{f\frac{\mathrm{evu}}{\mathrm{ml}}},$$

Focusing on the numerical values (and ignoring the units for what follows), we have

$$\breve{D}=\frac{D}{f},$$
$$\breve{R}=\frac{R}{f},$$
$$\breve{T}=\frac{T}{f}.$$

In our original differential equations, let us multiply both sides of all the differential equations by $1/f$, yielding:

$$\frac{1}{f}\frac{dD}{dt}=\psi_D\frac{1}{f}D,$$
$$\frac{1}{f}\frac{dR}{dt}=\psi_R\frac{1}{f}R-\ (\gamma_D D+\gamma_T T)\frac{1}{f}R,$$
$$\frac{1}{f}\frac{dT}{dt}=\psi_T\frac{1}{f}T+\ \ (\gamma_D D+\gamma_T T)\frac{1}{f}R.$$

This can be reworked as

$$\frac{d\breve{D}}{dt}=\psi_D\breve{D},$$
$$\frac{d\breve{R}}{dt}=\psi_R\breve{R}-\ (\gamma_D D+\gamma_T T)\breve{R},$$
$$\frac{d\breve{T}}{dt}=\psi\breve{T}+(\gamma_D D+\gamma_T T)\breve{R}.$$

It follows that:

$$\frac{d\breve{D}}{dt}=\psi_D\breve{D},$$
$$\frac{d\breve{R}}{dt}=\psi_R\breve{R}-\ \left(f\gamma_D\breve{D}+f\gamma_T\breve{T}\right)\breve{R},$$
$$\frac{d\breve{T}}{dt}=\psi\breve{T}+\left(f\gamma_D\breve{D}+f\gamma_T\breve{T}\right)\breve{R}.$$

If we let

$$\breve{\gamma}_D=f\gamma_D,$$

and

$$\breve{\gamma}_T = f\gamma_T.$$

then the above system becomes

$$\frac{d\breve{D}}{dt} = \psi_D\breve{D},$$

$$\frac{d\breve{R}}{dt} = \psi_R\breve{R} - \left(\breve{\gamma}_D\breve{D} + \breve{\gamma}_T\breve{T}\right)\breve{R},$$

$$\frac{d\breve{T}}{dt} = \psi\breve{T} + \left(\breve{\gamma}_D\breve{D} + \breve{\gamma}_T\breve{T}\right)\breve{R}.$$

This set of equations tracks the number of cells (per evu). Thus, if the above equations were used, then the derivations of the LDM estimate could flow exactly like we show in SI section 2. That is, the following will be correct:

$$\breve{\gamma}_D = -\ln p_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{\breve{D}_0\breve{R}_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)$$

Note, no correction is needed on $p_0(\tilde{t})$ as everything is in terms of numbers, which was how this quantity was derived. Because $\breve{D} = \frac{D}{f}$ and $\breve{R} = \frac{R}{f}$, we can rewrite the above as

$$\breve{\gamma}_D = -\ln p_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{\frac{D_0}{f}\frac{R_0}{f}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)$$

Or:

$$\frac{\breve{\gamma}_D}{f} = f\left\{-\ln p_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)\right\}$$

Because $\gamma_D = \frac{\breve{\gamma}_D}{f}$, we have

$$\gamma_D = f\left\{-\ln p_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)\right\}$$

Note that if our evu was 1 ml, then $f = 1$ and we could use our estimate exactly as written in equation [11]. Generally, we have to correct our original metric by multiplying by f.

SI section 6 : Extended Experimental Methods and Results

SI section 6a : Strains.

Escherichia coli K-12 BW25113 from the Top Lab was used as the ancestor of the three *E. coli* strains in this study. To derive the first strain, *E. coli* BW25113 was grown overnight and plated onto LB agar supplemented with 100 μg ml^{-1} streptomycin. A single streptomycin-resistant colony was selected and used to create an isogenic glycerol stock, *E. coli* K-12 BW25113 str^{R}, to be used as the plasmid-free *E. coli* recipient in this study (hereafter 'E(Ø)').

To derive the second strain, *E. coli* K-12 BW25113 was mixed with a host carrying the focal conjugative plasmid and incubated overnight in LB medium to facilitate plasmid

transfer. The focal plasmid was the modified IncF conjugative plasmid F'42 (hereafter 'pF') in which a tetracycline resistance gene was inserted using lambda red recombination (38). The mixture was plated onto LB agar supplemented with 100 μg ml^{-1} ampicillin (host selection) and 15 μg ml^{-1} tetracycline (pF plasmid selection) to select for *E. coli* K-12 BW25113 host containing the pF plasmid. A single colony was selected and used to create an isogenic glycerol stock to be used as the plasmid-containing *E. coli* donor in this study (hereafter 'E(pF)').

To derive the third strain, E(pF) was mixed with E(Ø) and incubated overnight in growth medium to facilitate plasmid transfer. The mixture was plated onto LB agar supplemented with 100 μg ml^{-1} streptomycin (host selection) and 15 μg ml^{-1} tetracycline (plasmid selection) to select for *E. coli* K-12 BW25113 str^R host containing the pF plasmid. A single colony was selected and used to create an isogenic glycerol stock to be used as a representative isogenic *E. coli* transconjugant in this study, hereafter 'E_T(pF)' where the T subscript is added to distinguish this strain from the plasmid-bearing *E. coli* E(pF) strain, which is susceptible to streptomycin.

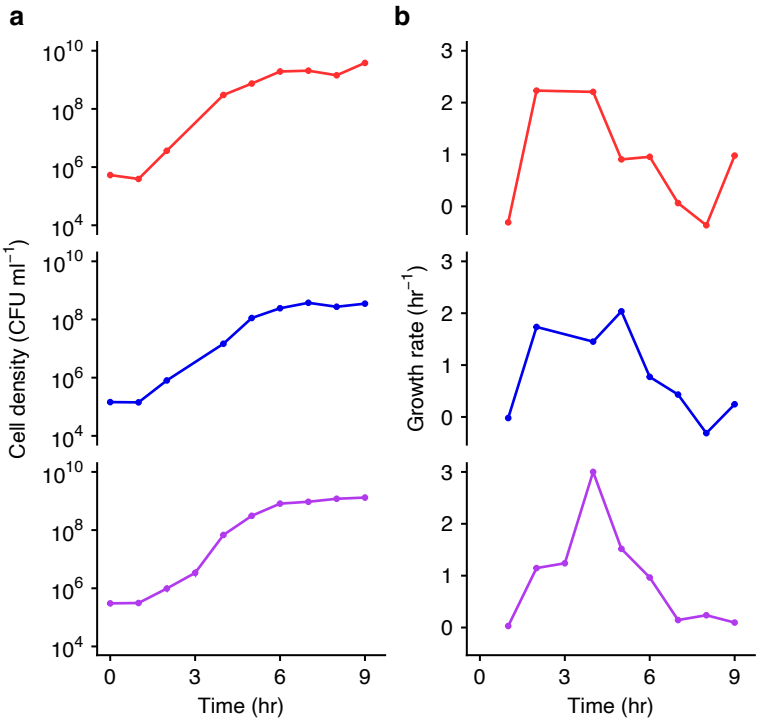
The *Klebsiella pneumoniae* strain Kp08 from Jordt *et. al.* (7) was used as the ancestor for the *K. pneumoniae* strain in this study. Kp08 was grown overnight and plated onto LB agar supplemented with 30 μg ml^{-1} nalidixic acid. A single nalidixic-acid-resistant colony was selected and used to create an isogenic glycerol stock, *K. pneumoniae* Kp08 nal^R. Kp08 nal^R was mixed with E(pF) and incubated overnight in growth medium to facilitate plasmid transfer. The mixture was plated onto LB agar supplemented with 30 μg ml^{-1} nalidixic acid (host selection) and 15 μg ml^{-1} tetracycline (plasmid selection) to select for *K.pneumoniae* Kp08 nal^R host containing the pF plasmid. A single colony was selected and used to create an isogenic glycerol stock to be used as the plasmid-containing *K. pneumoniae* donor in this study (hereafter 'K(pF)'). See SI Table 6 for a quick overview of the strains used in this study.

SI Table 6: The strains used in this study. Antibiotic abbreviations are as follows: tet = tetracycline, str = streptomycin, and nal = nalidixic acid, and the 'R' superscript indicates drug resistance in the strain.

Strain	Host	Plasmid
E(pF)	*E. coli* K-12 BW25113	F'42 tet^R
K(pF)	*K. pneumoniae* Kp08 nal^R	F'42 tet^R
E(Ø)	*E. coli* K-12 BW25113 str^R	None
E_T(pF)	*E. coli* K-12 BW25113 str^R	F'42 tet^R

SI section 6b : Growth rate assays.

The strains (SI Table 6) were inoculated into LB medium from frozen glycerol stocks and grown overnight. The plasmid-containing cultures were supplemented with 15 μg ml^{-1} tetracycline to select for maintenance of the plasmid. The saturated cultures were diluted 100-fold into LB medium to initiate a second 24 hours of growth (in order to acclimate the previously frozen strains to laboratory conditions). The acclimated cultures were then diluted 10,000-fold into LB growth medium and dispersed into 27 wells in a deep-well microtiter plate at a volume of 100 μl per well. Every hour, 30 μl was removed from three wells to determine cell density via selective plating (SI Figure 8a). The three replicate plates were averaged to estimate the cell density at each hour. The growth rates were calculated by taking the slope of each neighboring time point using equation [1.12] (SI Figure 8b). Using the growth rate calculated over time, an incubation time was chosen that coincided with the population growing at or near the maximum growth rate for each strain to ensure bacterial cultures entered the phase of maximal or close to maximal growth rate before the start of the conjugation assay. Thus, the growth rate estimates over time were used solely for determining the pre-assay growth period before the conjugation assay is executed and not to calculate the LDM estimate itself. A pre-assay growth period of 4 hours was used for both donors, E(pF) and K(pF), and the recipient, E(Ø).



SI Figure 8: The change in density and resulting growth rates of the relevant strains. (a) Monocultures of K(pF), E(Ø), and E(pF) (red, blue, and, purple, respectively) were tracked over 9 hours of growth via plating. Bars indicate the standard error of the mean of three replicate cultures, but the standard error was so small in all cases that it is

not visible in the plot. Note that at 3 hours a data point is missing for both K(pF) and E(Ø) due to plating error resulting in zero colonies and therefore no density estimate was available. (b) Using equation [1.12], the growth rates were calculated by taking the slope of a line connecting a focal point and the closest point earlier in time (in part a). This growth rate estimate is plotted at the focal point's time (in part b). The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 6c : Minimum inhibitory concentration (MIC) assays.

The strains (SI Table 6) were grown from glycerol stocks with two overnight incubations as previously described in SI section 6b. The acclimated cultures were diluted 100-fold into LB growth medium and dispersed into a column of wells in a deep-well plate at a volume of 500 µl per well. Then 500 µl of dual-antibiotic medium (streptomycin and tetracycline) was added to each well at increasing concentrations, forming a 2-fold gradient across the column. We note that the ratio of the two antibiotics was kept constant over the gradient. For each strain, this was repeated in three columns. After an overnight incubation, the well with the lowest concentration of the dual antibiotic medium across all replicates with no turbid growth was identified as the strain-specific MIC (SI Table 7). The concentration chosen for the transconjugant-selecting medium must be above the donor and recipient MIC, but below the transconjugant MIC. For this study, we proceeded with 7.5 µg ml^{-1} tet + 25 µg ml^{-1} str.

SI Table 7: The dual-drug gradient MIC for the strains of interest. The antibiotics used in the gradient were specific to the resistance profile of the transconjugant E_T(pF); streptomycin (str) and tetracycline (tet). The MIC data was used to identify the antibiotic concentration for the transconjugant-selecting medium used in both conjugation assays; cross- and within-species.

Strain	Cell type	Str and tet gradient MIC
E(pF)	Donor	1.88 µg ml^{-1} tet + 6.25 µg ml^{-1} str
K(pF)	Donor	1.88 µg ml^{-1} tet + 6.25 µg ml^{-1} str
E(Ø)	Recipient	3.75 µg ml^{-1} tet + 12.5 µg ml^{-1} str
E_T(pF)	Transconjugant	15 µg ml^{-1} tet + 50 µg ml^{-1} str

SI section 6d : Extinction probability assays.

A key component of the LDM conjugation protocol is differentiating parallel donor-recipient co-cultures that contain transconjugants from those that do not. This is done by adding transconjugant-selecting medium prepared at antibiotic concentrations below the MIC of the transconjugant and above the MIC of the donor and recipient. Given the low numbers of transconjugants in the co-cultures, the results from a recent study of Alexander and MacLean (39) have high relevance. First, the authors show that levels of antibiotic below the MIC of the resistant strain are sufficient to decrease the chance of outgrowth with very low cell numbers (e.g., a single cell). In the context of our current study, if the concentration of the transconjugant-selecting medium is too high then co-cultures that contain transconjugants could produce a non-turbid culture because the transconjugant cell(s) fail to establish a lineage. Therefore, to avoid spurious non-turbid wells in the LDM protocol, the probability that a transconjugant cell fails to establish (the transconjugant extinction probability) should ideally be 0 in the transconjugant-selecting medium. Second, the authors show that the presence of a sufficiently dense sensitive cell population in the environment can decrease the extinction probability of the resistant type. In the context of our current study, the presence of donors and recipients in the cultures may decrease the transconjugant extinction probability. Overall, a non-zero transconjugant extinction probability could lead to a biased estimate of the conjugation rate; therefore, it needs to be explicitly checked.

Inspired by the approach of Alexander and MacLean, we developed a similar approach to estimate the extinction probability of a transconjugant cell. First, we assume that a transconjugant cell has zero probability of extinction in antibiotic-free medium. While this assumption may be misplaced, it provides a starting point, and may itself be checked if there are reasons to doubt it holds. Second, we assume that a transconjugant cell has a specific probability of extinction in transconjugant-selecting medium with certain antibiotic concentrations given by the variable x, which is denoted π_x. Third, we assume that the lineage from every transconjugant cell in a population goes extinct independently. Consider a population of transconjugants distributed into many subpopulations containing transconjugant-selecting medium such that the average number of cells per subpopulation is initially T. Assuming an initial Poisson distribution, the fraction of subpopulations that leave zero transconjugant descendants, P_x is:

$$P_x = \sum_{i=0}^{\infty} \frac{e^{-T} T^i}{i!} (\pi_x)^i = \frac{e^{-T}}{e^{-T\pi_x}} \sum_{i=0}^{\infty} \frac{e^{-T\pi_x}}{i!} (T\pi_x)^i = e^{-T(1-\pi_x)}.$$

By our assumption, when considering antibiotic-free medium, which we represent as $x = 0$, we have

$$P_0 = e^{-T(1-\pi_0)} = e^{-T}.$$

Thus, it is the case

$$T = -\ln P_0.$$

Given that

$$P_x = e^{\ln P_0(1-\pi_x)},$$

we have a form to calculate the extinction probability in the transconjugant-selecting medium in the laboratory

$$\pi_x = 1 - \frac{\ln P_x}{\ln P_0}, \qquad [6.1]$$

where P_x is the fraction of non-turbid wells with transconjugant-selecting medium and P_0 is the fraction of non-turbid wells with antibiotic-free medium.

In the laboratory, we used the protocol implemented by Alexander and MacLean to estimate π_x with a few adjustments. Briefly, the transconjugants were diluted (4 x 10^7 fold) and 50 µl aliquots were dispensed into all wells in a deep-well microtiter plate. For the antibiotic-free condition, the wells were filled with LB medium to a final volume of 1 ml. For the transconjugant-selecting condition, the wells were filled with LB medium supplemented with transconjugant-selecting antibiotics (7.5 µg ml^{-1} tet + 25 µg ml^{-1} str, see SI section 6c for details) to a final volume of 1 ml. Both deep-well plates were incubated for 4 days. Using equation [6.1], we calculated a transconjugant extinction probability of 0.95 in the antibiotic concentration used for the transconjugant-selecting medium in this study.

Given that the extinction probability was non-negligible (i.e., $\pi_x > 0$), we ran a subsequent assay to estimate π_x in the presence of sensitive cells (donors and recipients) at approximately the final densities that occur when the transconjugant-selecting medium is added for both mating assays (cross- and within-species, see SI Table 8) reported in this study. This provided a more accurate π_x for correcting the LDM estimate (see SI section 7). In this experiment, the deep-well microtiter plates are prepared the same as above but supplemented with donor and recipient cells at the appropriate densities. As a result, we calculated mating-specific transconjugant extinction probabilities (SI Table 8). These mating-specific transconjugant extinction probabilities (given in SI Table 8) were used to correct the LDM estimate from each experimental replicate using equation [7.1].

Given the non-negligible extinction probability in the selective liquid medium, the extinction probability on the selective agar plates needed to be determined. We ran a subsequent assay to estimate π_x for the donor-, recipient-, and transconjugant-selecting agar plates. Briefly, the monocultures of each strain were diluted (10^{-5} and 10^{-6}) and plated onto antibiotic-free plates and the appropriate selecting plates. We used a slightly altered form for calculating the agar extinction probability

$$\pi_x = 1 - \frac{C_x}{C_0}, \qquad [6.2]$$

where C_x is the number of colonies on the antibiotic-infused plate and C_0 is the number of colonies on the antibiotic-free plate for the same diluted culture. Using equation [6.2], we calculated each strain's extinction probability (see SI Table 9 for the antibiotic concentration used in the selective agar plates in this study). These strain-specific extinction probabilities were used to correct the density estimates from each experiment. We note that correcting the density estimates for the 24-hour data in 3 out of the 6 experiments resulted in negative estimates for the recipient density data. We can explain

the negative estimates as follows. Given the high transconjugant extinction probability on the transconjugant-selecting agar plates (see SI Table 9), the transconjugant density increases after the correction. Indeed, the transconjugant population can become more common than the "estimated" recipient population. We say "estimated" because there are no agar plates that select *only* for recipient cells. Specifically, the "recipient-selecting" agar plates allow for both recipient and transconjugant growth. To determine the recipient density, we subtract the transconjugant density from the density of cells calculated from the "recipient-selecting" agar plate counts. When the transconjugants are more abundant than—or at relatively similar densities to—recipients, the exact recipient density cannot be determined due to its relative scarcity. Specifically, the subtractive plating scheme could result in a negative value. We note that this happens rarely given that transconjugant densities are typically orders of magnitude lower than recipients. In the cases of high conjugation rates and long incubation times, this issue is more likely to arise. If the recipient density went negative after subtraction, then the non-subtracted recipient density was used instead. An overestimate for recipient density leads to an underestimate for the SIM estimate at 24 hours; therefore, the differences between the cross-species LDM and SIM estimates shown in Figure 6 are conservative.

This section highlights the importance of non-zero extinction probabilities in selective conditions in the laboratory. Therefore, the extinction probabilities in selective liquid-medium and selective-agar plates need to be explicitly checked. If the extinction probabilities are indistinguishable from zero in each selective condition used, then the user can proceed, and no adjustments are necessary. However, a non-zero extinction probability is likely and can be a source of bias if not considered. We recommend two solutions. The first is to find a selection condition where the extinction probability is indistinguishable from zero. This option leans on the result from the Alexander and MacLean study which shows that the antibiotic concentration being sufficiently below the MIC of the focal strain can lower the extinction probability to a point that is indistinguishable from zero. We recognize that this solution may not be possible. For instance, the donor and recipient MIC for the transconjugant-selecting condition may be too close to the transconjugant MIC, such that there are no antibiotic concentrations that yield a zero transconjugant extinction probability and still counterselect donors and recipients. In this case, the user would proceed with the second solution where the extinction probabilities are used to compute a corrected estimate. This second solution was used in this study (see SI section 7).

SI Table 8: Mating-specific transconjugant extinction probabilities with transconjugant-selecting liquid medium. The donor and recipient densities were estimated using selective plating and were close to the final densities in the LDM conjugation protocol. Transconjugant-selective medium was prepared at the concentration used throughout the study (7.5 μg ml^{-1} tet + 25 μg ml^{-1} str).

Mating	Donor density	Recipient density	π_x
within-species E(pF) to E(Ø)	5 x 10^4	2 x 10^4	0.95
cross-species K(pF) to E(Ø)	1 x 10^8	7 x 10^6	0.93

SI Table 9: Strain-specific extinction probabilities with selective-agar plates. Donor-, recipient, and transconjugant-selective plates were prepared at concentrations that were used throughout the study (7.5 μg ml^{-1} tet, 25 μg ml^{-1} str, and 7.5 μg ml^{-1} tet + 25 μg ml^{-1} str, respectively).

Strain	Selective-plate type	π_x
E(pF)	Donor	0.30
K(pF)	Donor	0.21
E(Ø)	Recipient	0.55
E_T(pF)	Transconjugant	0.99

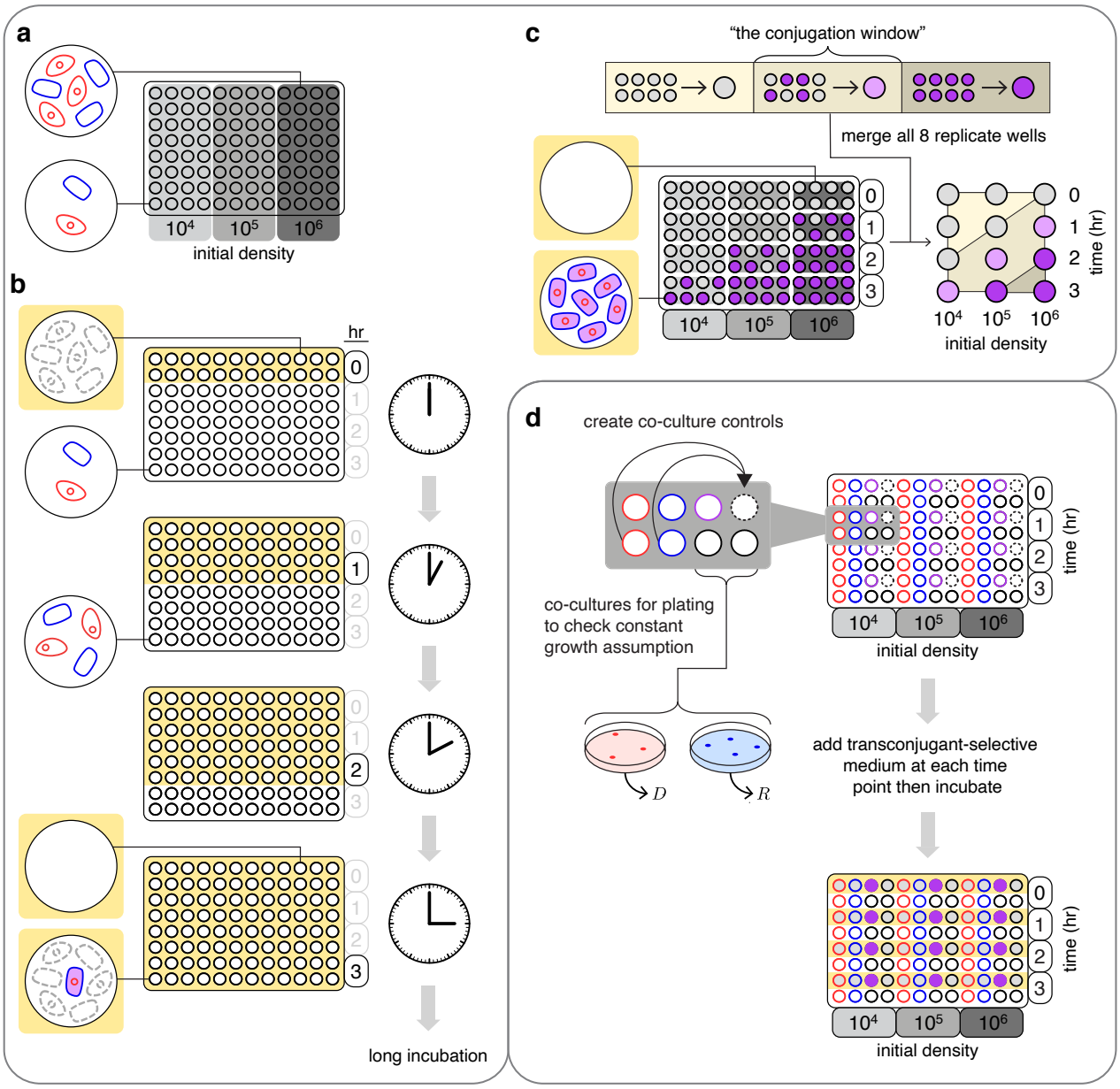
SI section 6e : Choosing an incubation time and initial density for executing the LDM conjugation assay.

To find an incubation time and initial densities for executing the LDM protocol, all strains (SI Table 6) were prepared using the procedure in SI section 6b. We mixed exponentially growing donors and recipients in a large array of parallel co-cultures for a full factorial treatment of three initial densities and four incubation times (SI Figure 9a). We note that the resolution of initial densities and incubation times can be adjusted as needed. This is particularly useful if the conjugation rate is completely unknown. Alternatively, there could be good reasons for longer incubation times such as slow growth rates. For ease of explanation, we illustrate the protocol with a concrete example. Four columns were used for each initial density (10^4, 10^5, and 10^6 cells per ml) where 2 rows were used for each incubation time (0, 1, 2, and 3 hours) resulting in 8 wells per density-time treatment. For each dilution, the exponentially growing donor and recipient cultures were diluted by the specific factor, mixed at equal volumes, and dispensed into

the wells in the corresponding four columns at a volume of 100 µl per well (SI Figure 9a, black-bordered wells). At each incubation time, 900 µl of transconjugant-selecting medium (7.5 µg ml^{-1} tetracycline and 25 µg ml^{-1} streptomycin; see SI section 6c and SI section 6d) was added to each well in the corresponding two rows (SI Figure 9b, yellow-background). After the last time point (t = 3 hours), the deep-well plate was incubated for 4 days. After the long incubation, we assessed the co-cultures within each time-density treatment for presence or absence of transconjugants by recording the turbidity (4 columns x 2 rows = 8 wells; SI Figure 9c). There were three outcomes possible for each time-density treatment: none of the co-cultures have transconjugants (gray-filled dot), all co-cultures have transconjugants (purple-filled dot), or there is both transconjugant-containing and transconjugant-free co-cultures (light-purple dot). The goal is to identify a density-time combination with the last outcome (i.e., both turbid and non-turbid co-cultures). These treatments meet the $\hat{p}_0(\tilde{t})$ condition (i.e., $0 < \hat{p}_0(\tilde{t}) < 1$). As a general expectation, a high donor conjugation rate (γ_D) will require shorter incubation times than a lower rate for a given initial density. For our matings (within- and cross-species), we found multiple density-time combinations that met the $\hat{p}_0(\tilde{t})$ condition. For the within-species mating assay, we chose a 10^3-fold dilution and an incubation time of 1 hour and 15 minutes. For the cross-species mating assay, we chose the 10^3-dilution and a 4-hour incubation time.

Even though multiple density-time treatments met the $\hat{p}_0(\tilde{t})$ condition, the final choice could not be made without the information from the controls. Thus, an additional deep-well plate was created (SI Figure 9d) to accompany the density-time plate containing the co-cultures (SI Figure 9a). This deep-well plate had the same factorial layout for densities (four columns) and incubation times (two rows) except the 8 wells within each treatment are not exclusively co-cultures. 3 of the wells contained monocultures of the three strains. Specifically, 100 µl of donor, recipient and transconjugant cultures were each placed in their own well (SI Figure 9d red-, blue- and purple-bordered wells). At a later point in the assay, these monocultures allowed us to determine that transconjugant-selecting medium prohibited growth of both donors and recipients, while permitting growth of transconjugants at each density-time treatment. An additional 2 wells contained monocultures of donors and recipients which are used to create a co-culture (in an empty well, dash-bordered well) during the assay itself at each incubation time (for each initial density). An additional 2 wells contained 100 µl of donor-recipient co-cultures which were used for selective plating to verify that the donors and recipients maintain a constant growth rate. At each incubation time, three events occurred in rapid succession. First, 30 µl was removed from each of the wells used to determine densities via selective plating. Second, donor and recipient monocultures were mixed at equal volumes into the empty well (SI Figure 9d, indicated by the gray arrows). Importantly, this well served as a control to verify that new transconjugant cells did not form via conjugation after transconjugant-selecting medium was added. Third, 900 µl of transconjugant-selecting medium was added to the first row of wells at the relevant time point (yellow background). The deep-well plate was incubated for 4 days. For the density-

time combinations chosen, the control wells verified that the transconjugant-selecting medium operated as expected. In addition, the selective plating indicated the conditions under which the donors and recipients maintained constant growth.



SI Figure 9: Overview for finding an incubation time and initial densities for executing the LDM. (a) the microtiter plate map designating the placement of the co-cultures over 10-fold increases in initial densities (different shades of gray). For simplicity, donors and recipients are at the same proportion in each co-culture. (b) Using the microtiter plate from part a, transconjugant-selecting medium (yellow-background) is added at each time designated by two rows in the microtiter plate. Two example wells from different density-time combinations are highlighted on the left. In the top example well, transconjugant-selecting medium is added immediately, inhibiting growth of donor and recipient cells (grey dashed cells), and resulting in a non-turbid well as no transconjugants formed. In the bottom example well, the donor and recipient population

in the co-culture grow until transconjugant-selecting medium is added at 3-hours, inhibiting growth of donors and recipients, and permitting growth of the formed transconjugants. (c) After a lengthy incubation of the microtiter plate from part b, there are two well-types in the microtiter plate (bottom-left): transconjugant-containing (purple-filled) and transconjugant-free (gray-filled). For each density-time treatment, the 8 mating wells are considered as a group resulting in one of three outcomes (top): all transconjugant-free wells (gray dot), all transconjugant-containing wells (purple dot), a proportion of both well types (light-purple dot). Any treatment with a light-purple dot represents a viable combination of initial densities (D_0' and R_0') and incubation time ($\tilde{t}'$). (d) The microtiter plate with the control wells is set up with the same factorial layout used in part a except the 8 wells in each density-time treatment are not all co-cultures (black-bordered circles). Donor, recipient, and transconjugant monocultures serve as controls (red-, blue-, and purple-bordered wells, respectively). For the empty well (dash-bordered circles), donor and recipient monocultures are mixed into the empty well (indicated by grey arrows) to create a co-culture control at each time point to verify that diluting with transconjugant-selecting medium effectively prevents conjugation. In addition, the co-cultures are sampled at each time point to uncover densities and determine whether donors and recipients maintain constant growth. Subsequently, transconjugant-selecting medium is added to the microtiter plate at the same times as the microtiter plate in part a. The control wells inoculated with transconjugants should be turbid (purple-filled) while the monocultures with donors and recipients should be non-turbid. In addition, the co-cultures created at each time point for the different initial density treatments should be non-turbid.

SI section 7 : Probability generating function, low-order moments, and failure to establish

The aim of the first part of this section is to explore the connection between mutation and conjugation processes further. In the second part of this section, we derive a general expression for the LDM estimate that incorporates cases when the transconjugant doesn't always establish a successful lineage (i.e., non-zero extinction probability).

Keller and Antal (43) studied a generalization of the process explored by Luria and Delbrück (19). To start, Keller and Antal consider a wildtype population expanding from a single cell as follows:

$$N_t = f(t) = e^{\delta t}.$$

Each wildtype cell generates a mutant cell at a rate ν', which grows as a stochastic birth process with rate α (Keller and Antal studied a supercritical birth-death process, but we will focus on the special case of a pure birth process). In this case, mutants form at a rate $\nu' f(t)$, such that the times of mutant arrival conform to a non-homogeneous Poisson process. We note that if we start with N_0 cells, then mutants form at a rate $N_0 \nu' f(t)$.

Alternatively, we can set $\nu = N_0\nu'$, such that mutants form at a rate $\nu f(t)$, which is the case explored by Keller and Antal.

Keller and Antal derive the probability generating function for the total number of mutants at an arbitrary time:

$$G(z,t) = \exp\left\{\frac{\nu}{\delta}\left(F\left(1,\kappa;1+\kappa;\frac{z}{z-1}e^{-\alpha t}\right) - e^{\delta t}F\left(1,\kappa;1+\kappa;\frac{z}{z-1}\right)\right)\right\},$$

where F is the Gaussian hypergeometric function and $\kappa = \frac{\delta}{\alpha}$.

Our process of interest (the formation and growth of transconjugants) can be seen as an instance of their formulation by making the following substitutions:

$$\delta = \psi_D + \psi_R,$$
$$\nu = \gamma_D D_0 R_0,$$
$$\alpha = \psi_T.$$

With these substitutions, the generating function becomes:

$$G(z,t) = \exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(F\left(1,\frac{\psi_D+\psi_R}{\psi_T};1+\frac{\psi_D+\psi_R}{\psi_T};\frac{z}{z-1}e^{-\psi_T t}\right)\right.\right.$$
$$\left.\left. - e^{(\psi_D+\psi_R)t}F\left(1,\frac{\psi_D+\psi_R}{\psi_T};1+\frac{\psi_D+\psi_R}{\psi_T};\frac{z}{z-1}\right)\right)\right\}.$$

Because

$$G(z,t) = \sum_{n=0}^{\infty} p_n(t) z^n,$$

the probability of zero transconjugants now becomes straightforward (given $F\left(1,\frac{\psi_D+\psi_R}{\psi_T};1+\frac{\psi_D+\psi_R}{\psi_T};0\right) = 1$):

$$p_0(t) = G(0,t) = \exp\left\{\frac{-\gamma_D D_0 R_0}{\psi_D+\psi_R}\left(e^{(\psi_D+\psi_R)t} - 1\right)\right\},$$

which agrees with the result from SI section 2.

Making the appropriate substitutions, we can also write the mean and variance (eqs. 8 and 9 from Keller and Antal) for the transconjugants:

$$E[T_t] = \begin{cases} \gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} t & \text{if } \psi_D + \psi_R = \psi_T \\ \dfrac{\gamma_D D_0 R_0\left(e^{(\psi_D+\psi_R)t} - e^{\psi_T t}\right)}{\psi_D+\psi_R-\psi_T} & \text{if } \psi_D + \psi_R \neq \psi_T \end{cases}$$

$$\mathrm{Var}[T_t]$$
$$= \begin{cases} \dfrac{2\gamma_D D_0 R_0\left(e^{2(\psi_D+\psi_R)t} - e^{(\psi_D+\psi_R)t}\right)}{\psi_D+\psi_R} - \gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} t & \text{if } \psi_D+\psi_R = \psi_T \\ \dfrac{2\gamma_D D_0 R_0\left(e^{(\psi_D+\psi_R)t/2} - e^{(\psi_D+\psi_R)t}\right)}{\psi_D+\psi_R} + 2\gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} t & \text{if } \psi_D+\psi_R = 2\psi_T \\ \gamma_D D_0 R_0 \left\{ \dfrac{2e^{2\psi_T t}(\psi_T - (\psi_D+\psi_R)) - e^{\psi_T t}(2\psi_T - (\psi_D+\psi_R)) + (\psi_D+\psi_R)e^{(\psi_D+\psi_R)t}}{(2\psi_T - (\psi_D+\psi_R))(\psi_T - (\psi_D+\psi_R))} \right\} & \text{otherwise} \end{cases}$$

We provide derivations for these expressions in GitHub Appendix VI. In all cases, the variance grows relative to the mean over time (see GitHub Appendix VII for the derivations).

In our experiment, at time $\tilde{t}$, medium selecting for transconjugants is added to every mating culture. If every transconjugant always establishes a successful lineage, then every mating culture with one or more transconjugant cells at time $\tilde{t}$ will produce a turbid culture after a lengthy incubation. A more realistic scenario would be to assume that every transconjugant cell fails to establish a lineage with some probability, which we call π. If failure to establish occurs independently for each transconjugant, then the probability of a non-turbid culture after incubation (P_{nt}) when selective medium was added at time $\tilde{t}$ is:

$$P_{\mathrm{nt}} = \sum_{n=0}^{\infty} p_n(\tilde{t})\pi^n.$$

However, this is equivalent to an appropriate evaluation of the generating function:

$$P_{\mathrm{nt}} = G(\pi, \tilde{t}).$$

This can be rewritten as

$$P_{\mathrm{nt}} = \exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D+\psi_R}\left(F\left(1, \frac{\psi_D+\psi_R}{\psi_T}; 1+\frac{\psi_D+\psi_R}{\psi_T}; \frac{\pi}{\pi-1}e^{-\psi_T\tilde{t}}\right)\right.\right.$$
$$\left.\left. - e^{(\psi_D+\psi_R)\tilde{t}} F\left(1, \frac{\psi_D+\psi_R}{\psi_T}; 1+\frac{\psi_D+\psi_R}{\psi_T}; \frac{\pi}{\pi-1}\right)\right)\right\}.$$

Solving for γ_D yields

$$\gamma_D = \frac{-\ln(P_{\mathrm{nt}})(\psi_D+\psi_R)}{D_0 R_0}\left(e^{(\psi_D+\psi_R)\tilde{t}} F\left(1, \frac{\psi_D+\psi_R}{\psi_T}; 1+\frac{\psi_D+\psi_R}{\psi_T}; \frac{\pi}{\pi-1}\right)\right.$$
$$\left. - F\left(1, \frac{\psi_D+\psi_R}{\psi_T}; 1+\frac{\psi_D+\psi_R}{\psi_T}; \frac{\pi}{\pi-1}e^{-\psi_T\tilde{t}}\right)\right)^{-1}$$

If the values of D_0 and R_0 are not the total initial numbers, but cell densities $(\mathrm{cfu/ml})$ in some volume for the mating culture (such that there are f experimental volumes per ml) and we wish to measure conjugation rate in units $\mathrm{ml}/(\mathrm{h}\cdot\mathrm{cfu})$, then must add a correction factor (see SI section 5), yielding

$$\gamma_D = f\frac{-\ln\left(P_{\mathrm{nt}}\right)(\psi_D+\psi_R)}{D_0R_0}\Bigg(e^{(\psi_D+\psi_R)\tilde{t}}F\left(1,\frac{\psi_D+\psi_R}{\psi_T};1+\frac{\psi_D+\psi_R}{\psi_T};\frac{\pi}{\pi-1}\right) - F\left(1,\frac{\psi_D+\psi_R}{\psi_T};1+\frac{\psi_D+\psi_R}{\psi_T};\frac{\pi}{\pi-1}e^{-\psi_T\tilde{t}}\right)\Bigg)^{-1} \quad [7.1]$$

First of all, we note that if every transconjugant establishes a lineage (i.e., $\pi = 0$), then $P_{\mathrm{nt}} = p_0(\tilde{t})$ and equation [7.1] reduces to

$$\gamma_D = f\frac{-\ln\left(p_0(\tilde{t})\right)(\psi_D+\psi_R)}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-1\right)},$$

which, using the maximum likelihood estimate for $p_0(\tilde{t})$, can be rewritten as

$$\gamma_D = f\left\{\frac{1}{\tilde{t}}\left[-\ln\hat{p}_0(\tilde{t})\right]\frac{\ln D_{\tilde{t}}R_{\tilde{t}} - \ln D_0R_0}{D_{\tilde{t}}R_{\tilde{t}} - D_0R_0}\right\},$$

and this is simply equation [13].

However, equation [7.1] is the more general expression. In SI section 6d, we discuss a method for estimating π. The maximum likelihood estimate for P_{nt} is the fraction of empty wells in the LDM protocol. Before, we called this $\hat{p}_0(\tilde{t})$, however, when there is positive probability that a transconjugant cell fails to establish (i.e., $\pi > 0$), then generally $P_{\mathrm{nt}} > p_0(\tilde{t})$. Thus, we will denote the maximum likelihood estimate as $\hat{P}_{\mathrm{nt}}$ (the fraction of non-turbid wells).

If we let the density of transconjugants in a monoculture at times 0 and $\tilde{t}$ be T_0^m and $T_{\tilde{t}}^m$, respectively (see SI section 6b) the following is the more general conjugation rate estimate (where all growth rates have been converted into estimated densities):

$$\gamma_D = f\frac{-\ln\left(\hat{P}_{\mathrm{nt}}\right)\varsigma}{\tilde{t}}\Bigg(D_{\tilde{t}}R_{\tilde{t}}F\left(1,\kappa;1+\kappa;\frac{\pi}{\pi-1}\right) - D_0R_0F\left(1,\kappa;1+\kappa;\frac{\pi}{\pi-1}\frac{T_0^m}{T_{\tilde{t}}^m}\right)\Bigg)^{-1} \quad [7.2]$$

with

$$\varsigma = \ln D_{\tilde{t}}R_{\tilde{t}} - \ln D_0R_0,$$

and

$$\kappa = \frac{\varsigma}{\ln T_{\tilde{t}}^m - \ln T_0^m} = \frac{\ln D_{\tilde{t}}R_{\tilde{t}} - \ln D_0R_0}{\ln T_{\tilde{t}}^m - \ln T_0^m}.$$

SI section 8 : Variance in Estimates

Here we will focus on two estimates, ASM and LDM, and ask about their variance (enabling us to compare precision). We will focus exclusively on the contributions to this

variance coming from the stochasticity in the transconjugant numbers (i.e., ignoring contributions coming from assessment of initial and final donor and recipient populations). Details on some of the derivations in this section are given in Github Appendix VII.

We start with the ASM estimate (here we express the estimate in terms of growth rate parameters):

$$\gamma_D = \frac{\psi_D + \psi_R - \psi_T}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)} T_{\tilde{t}}.$$

Because we are only focusing on the contribution of the transconjugant variation, all parameters (initial densities and growth rates will be taken to be fixed). Thus, we can think about the ASM estimate as a random variable Γ_{ASM}, where

$$\Gamma_{\mathrm{ASM}} = c_1 T_{\tilde{t}},$$

where the constant c_1 is

$$c_1 = \frac{\psi_D + \psi_R - \psi_T}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)}.$$

The variance of the ASM estimate is then

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = {c_1}^2 \{\mathrm{var}(T_{\tilde{t}})\}.$$

But we have a closed form expression for $\mathrm{var}(T_{\tilde{t}})$. If $\psi_T \notin \{\psi_D + \psi_R, (\psi_D + \psi_R)/2\}$, we have

$$\mathrm{var}(\Gamma_{\mathrm{ASM}})$$
$$= \frac{\gamma_D(\psi_D + \psi_R - \psi_T)}{D_0 R_0} \left\{ \frac{(\psi_D + \psi_R)e^{(\psi_D+\psi_R)\tilde{t}} + (\psi_D + \psi_R - 2\psi_T)e^{\psi_T \tilde{t}} - (\psi_D + \psi_R - \psi_T)2e^{2\psi_T \tilde{t}}}{(\psi_D + \psi_R - 2\psi_T)\left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)^2} \right\}.$$

The formulas for $\psi_T = \psi_D + \psi_R$ and $2\psi_T = \psi_D + \psi_R$ could also be derived via simple substitution (note, $\lim_{\psi_T \to \psi_D + \psi_R} c_1 = 1/(D_0 R_0 t e^{(\psi_D+\psi_R)\tilde{t}})$. These formulas allow us to project variance in the ASM estimate over time due to transconjugant variation if all parameters are known.

We now turn to the LDM estimate:

$$\gamma_D = -\ln \hat{p}_0(\tilde{t}) \left(\frac{\psi_D + \psi_R}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)} \right).$$

What we actually measure is the number of populations (or wells) that have no transconjugants (call this w) out of the total number of populations (or wells) tracked (call this W). As we show in Github Appendix IV, the maximum likelihood estimate of $p_0(\tilde{t})$ is

$$\hat{p}_0(\tilde{t}) = \frac{w}{\mathrm{W}}.$$

Of course, from experiment to experiment, there will be variance in the number of populations with no transconjugants. Let us consider a random variable F, which represents the fraction of total populations that have no transconjugants. The expectation of F is (we drop the time argument for notational convenience):

$$\mathrm{E}[F] = p_0.$$

The second central moment of F is

$$\mathrm{var}[F] = \frac{p_0(1-p_0)}{\mathrm{W}}.$$

Because we are only focusing on the contribution of the transconjugant variation, all parameters (initial densities and growth rates will be taken to be fixed). Thus, we can think about the LDM estimate as a random variable Γ_{LDM},

$$\Gamma_{\mathrm{LDM}} = c_2 \ln F,$$

where the constant c_2 is

$$c_2 = -\left(\frac{\psi_D + \psi_R}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right).$$

The variance of the LDM estimate is then

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) = {c_2}^2\{\mathrm{var}(\ln F)\}.$$

Here we use a first-order Taylor series approximation for $\ln F$ centered at $\mathrm{E}[F]$:

$$\ln F \approx \frac{F}{\mathrm{E}[F]} + \ln(\mathrm{E}[F]) - 1.$$

And we have

$$\mathrm{var}[\ln F] \approx \frac{1}{\mathrm{W}}\left(\frac{1}{p_0} - 1\right).$$

This approximation will be accurate when the deviation between F and $\mathrm{E}[F]$ is very small (i.e., $\frac{|F-\mathrm{E}[F]|}{\mathrm{E}[F]} \ll 1$). As W (the number of replicate populations in the experiment) gets large, the distribution of F will tighten around $\mathrm{E}[F]$, making the approximation more reasonable.

Now, we have the following expression for p_0 (reintroducing the time argument):

$$p_0(\tilde{t}) = \exp\left\{\frac{-\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\}.$$

Therefore,

$$\mathrm{var}[\ln F_{\tilde{t}}] \approx \frac{1}{\mathrm{W}}\left(\exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\} - 1\right),$$

where we make the time dependence of F clear. Returning to the variance for the LDM estimate,

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{1}{\mathrm{W}}\left(\frac{\psi_D + \psi_R}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)^2 \left(\exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\} - 1\right).$$

If we define

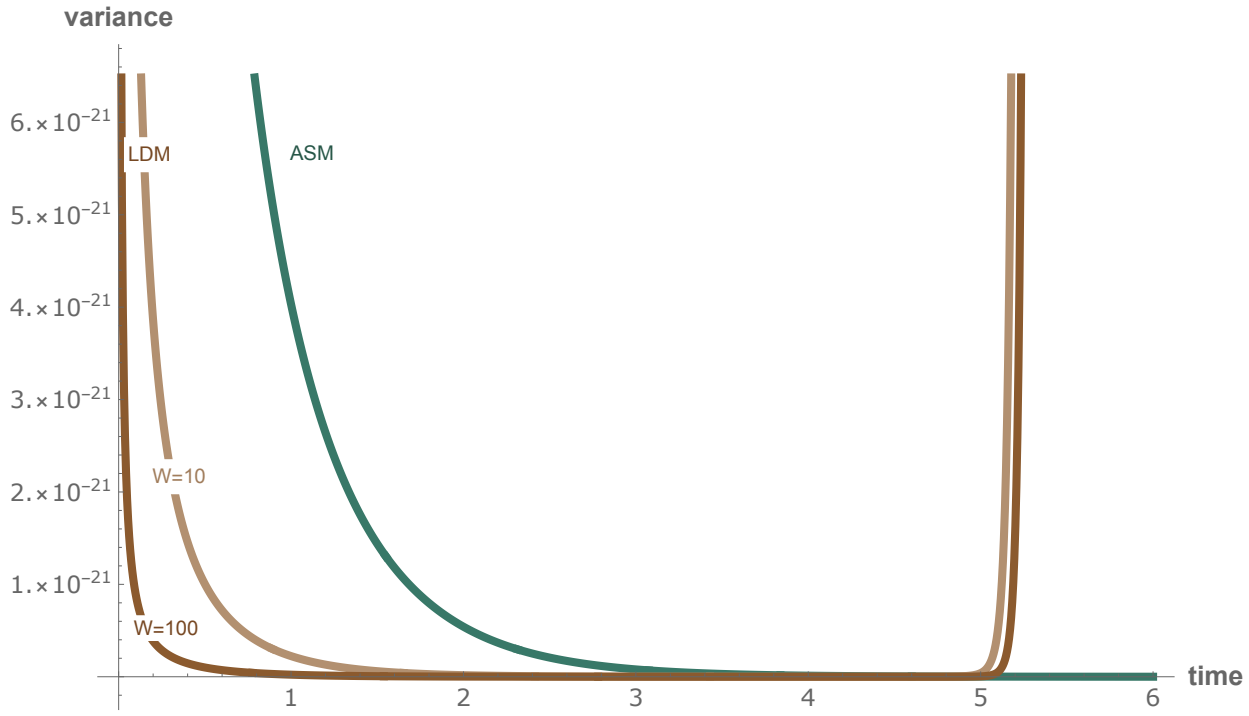
$$\xi_{\tilde{t}} = \frac{\psi_D + \psi_R}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)},$$

then we have

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{{\xi_{\tilde{t}}}^2}{\mathrm{W}}\left(e^{\left(\frac{\gamma_D}{\xi_{\tilde{t}}}\right)} - 1\right).$$

In SI Figure 10, we explore the variances (approximate in the case of LDM) as a function of time. The LDM estimates (for two different numbers of populations) are more precise

(lower variance) for much of the time range. However, if the time gets too high ($\tilde{t} \approx 5$ for the parameter set shown in SI Figure 10), then the LDM variance blows up (while the ASM variance remains very low). In a case like this, the LDM is predicted to be more precise when the time of the assay is sufficiently low. In GitHub Appendix VII, we demonstrate this precision advantage for the LDM estimate mathematically. Also in GitHub Appendix VII, we derive an approximation for the variance for the SIM estimate, which demonstrates that the variances for the SIM and ASM estimates are extremely similar.



SI Figure 10 : The variance of the ASM (green) and LDM (brown) estimates. Different numbers of populations (W) are used for the LDM estimates, as indicated. The parameters used here are $\gamma_D = 10^{-12}$, $D_0 = R_0 = 10^4$, $\psi_D = 1$, and $\psi_R = \psi_T = 1.5$.

As illustrated in SI Figure 10, the variance in the LDM estimate changes with the number of populations (W). How does this number affect the variance in the LDM estimate? Here we use simulations to further explore this question. In SI Figure 11a, we present the variance of LDM estimates as a function of incubation time ($\tilde{t}$) and the number of populations (W). Generally, as the number of populations decreases or as the boundaries of the time interval are approached (where nearly none or all of the populations have transconjugants) the variance in the LDM estimate rises. The exception seems to be for times that are very long, but the low variance is likely a result of having many infinite estimates that are not included in the estimate variance (SI Figure 11b). Both infinite estimates (SI Figure 11b) and zero estimates (SI Figure 11c) are more likely as the number of populations decreases; in other words, the interval of incubation times

producing non-zero finite estimates increases with the number of populations. Generally, the greater the number of populations and the more intermediate the incubation time (e.g., where approximately half of the populations have transconjugants), the lower the variance.

Suppose an experimenter is considering some number of wells (populations) and wants to decide how many estimates to produce. For instance, with 500 wells, the experimenter could decide to run a single LDM assay and obtain a single estimate (with $\mathrm{W} = 500$) or perhaps instead could run 5 assays (with $\mathrm{W} = 100$), 10 assays (with $\mathrm{W} = 50$), 50 assays (with $\mathrm{W} = 10$) or 100 assays (with $\mathrm{W} = 5$) for 5, 10, 50, and 100 estimates, respectively. Does it make a difference to the precision or accuracy to split or lump wells? Here we explore this question through simulation. How do we compare different partitions of wells? Let us consider some total number of wells, call this W^*, and consider some factor of W^*, which we will call W'; i.e., $\mathrm{W}^*/\mathrm{W}' = n$, where n is an integer. Here we will compare a single estimate with W^* wells with the mean of n estimates that each use W' wells. Thus, for SI Figure 11d, each point for $\mathrm{W} = 500$ is a single estimate, where each point for $\mathrm{W} = 5$, $\mathrm{W} = 10$, $\mathrm{W} = 50$, and $\mathrm{W} = 100$ is the mean of 100, 50, 10, and 5 estimates, respectively. With these comparisons in mind, we see two slight effects of different partitioning patterns. First, the variance is a bit higher for the single estimate coming from the largest number of wells. We attribute this shift to the fact that other quantities involved in the estimate (e.g., density of donors and recipients) are only being computed once for each point for $\mathrm{W} = 500$ in SI Figure 11d, whereas these quantities are being computed multiple times for smaller W values, such that anomalous values would tend to get muted as the estimates were averaged. The second effect is a more notable one. We see that as the number of wells per estimate goes down, slight inaccuracies in the estimate start to occur. Why does this happen?

To answer this question, let us consider the LDM estimate:

$$\gamma_D = -\ln p_0(\tilde{t}) \left(\frac{\psi_D + \psi_R}{D_0 R_0 \left(e^{(\psi_D + \psi_R)\tilde{t}} - 1 \right)} \right)$$

The main thing that will be affected by the number of populations is $p_0(\tilde{t})$. Specifically, as W decreases, the variance in the fraction of populations without transconjugants increases. Suppose that we have n LDM estimates under consideration, and for each one a value $\hat{p}_0(\tilde{t})$ is needed. Here we define:

$$\overline{\hat{p}_0(\tilde{t})} = \frac{\sum_{i=1}^{n} \hat{p}_{0,i}(\tilde{t})}{n},$$

where $\hat{p}_{0,i}(\tilde{t})$ is the fraction of populations without transconjugants for the i^{th} estimate. Now, by Jensen's inequality, we have:

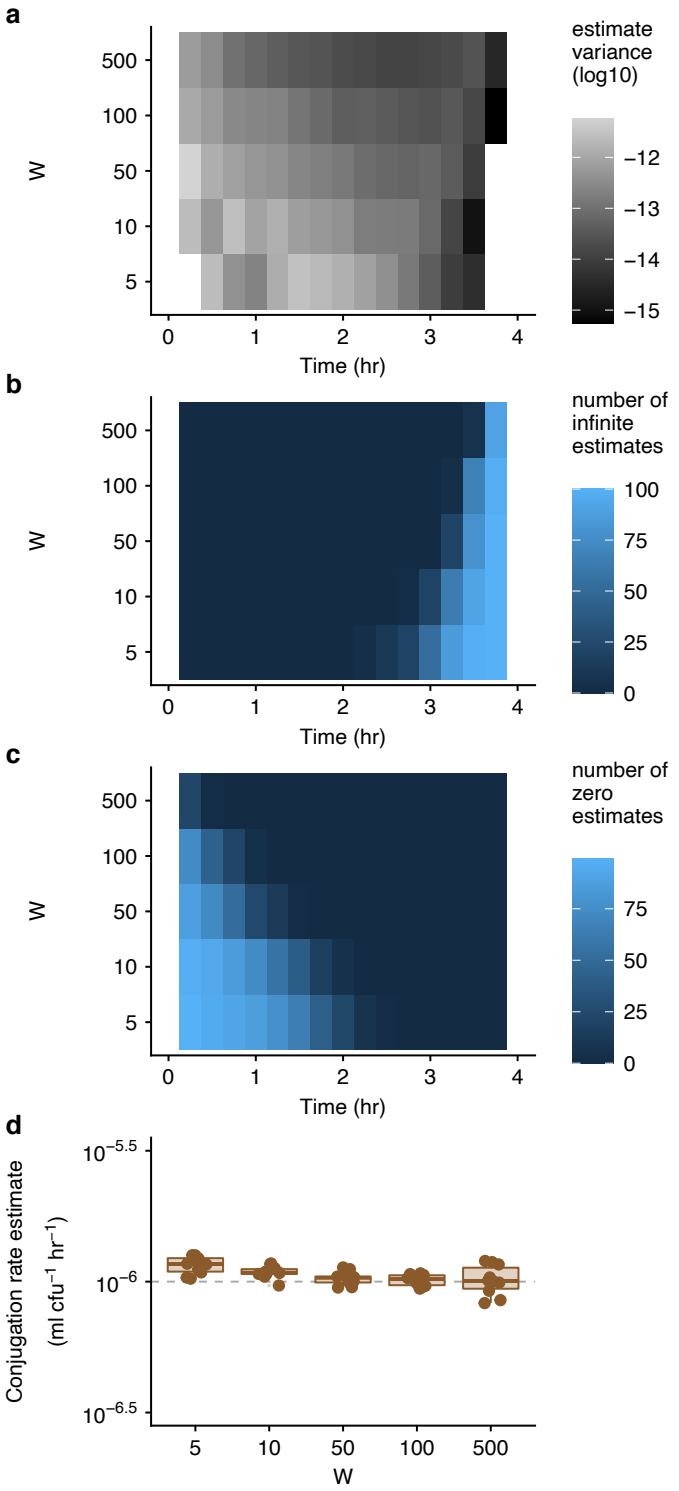
$$-\ln\left\{\frac{\sum_{i=1}^{n}\hat{p}_{0,i}(\tilde{t})}{n}\right\}\left(\frac{\psi_D+\psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-1\right)}\right)<\frac{1}{n}\sum_{i=1}^{n}-\ln\hat{p}_{0,i}(\tilde{t})\left(\frac{\psi_D+\psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-1\right)}\right)$$

$$-\ln\overline{\hat{p}_0(\tilde{t})}\left(\frac{\psi_D+\psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-1\right)}\right)<\frac{1}{n}\sum_{i=1}^{n}-\ln\hat{p}_{0,i}(\tilde{t})\left(\frac{\psi_D+\psi_R}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-1\right)}\right)$$

As W gets large, the value $\hat{p}_0(\tilde{t})$ is close to $\overline{\hat{p}_0(\tilde{t})}$ for smaller W values. Thus, using the terminology from above:

$$\gamma_D[\mathrm{W}^*]<\frac{1}{n}\sum_{i=1}^{n}\gamma_D[\mathrm{W}'_i],$$

where $\gamma_D[\mathrm{W}^*]$ is the conjugation rate for the largest number of wells (W^*), and $\gamma_D[\mathrm{W}'_i]$ is the conjugation rate for the i^{th} assay using a smaller number of wells (W'). Thus, we see that as we partition wells into smaller numbers per estimate, the mean estimate will rise, which is what we see in SI Figure 11d. Consequently, we recommend a reasonably large number of wells in the LDM assay. A number between 50 and 100 appears sufficient to avoid inaccuracy and is also convenient when using a microtiter plate format for populations.



SI Figure 11: The variance of LDM estimates using stochastic simulation. Different number of populations (W) are used for the LDM estimates, as indicated. The parameters used here are the same baseline parameters in SI Figure 1 which were $\psi_D = \psi_R = \psi_T = 1$, and $\gamma_D = \gamma_T = 10^{-6}$. The dynamic variables were initialized with $D_0 = R_0 = 10^2$ and $T_0 = 0$. (a) The variance among the 100 estimates is given at 15-minute intervals where more than 1 out of the 100 calculated estimates produced a finite non-zero value. We ignore infinite estimates in the calculation of the variance. (b) The number of estimates

with an infinite value out of the 100 calculated. (c) The number of estimates with a zero value out of the 100 calculated. (d) A total of 500 populations is partitioned in different ways—split into 100 groups of 5 populations (W=5), 50 groups of 10 populations (W=10), 10 groups of 50 populations (W=50), 5 groups of 100 populations (W=100), or a single group of 500 populations (W=500). Each plotted point is the mean conjugation rate of the rates calculated for each group (where the number of populations within each group vary as indicated by the W value) at a specific incubation time ($\tilde{t} = 2.35$) selected using the criteria described in the Materials and Methods. We ran the partitioning analysis 10 times using a new set of 500 populations. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

SI section 9 : Random effects on estimate accuracy and precision

In this section we explore, through simulation, some of the consequences of other random effects on the LDM and SIM estimates. Some of these effects are a consequence of experimental protocols. For instance, both approaches require dilution and plating in the laboratory to estimate donor and recipient density (and the SIM approach also uses dilution and plating to estimate transconjugant density). Because dilution and plating are subject to random sampling effects, there will be density-estimation errors introduced by these procedures. Other random effects are features of the cells under study. As we describe in SI section 6d and 7, there can be a non-zero probability that any cell will fail to establish a lineage. For instance, a donor cell may fail to form a colony on a plate after incubation on selective medium, or a lone transconjugant cell in a well may fail to yield a turbid culture after incubation in selective medium. Again, there will be stochasticity in the number of cell lineages that go extinct, which will lead to error in calculating key quantities needed for the estimates (even with corrections). Here we explore the consequences of some of these random effects.

Random effects in dilution, plating, and failure to form colonies: We ran our stochastic simulations as before (SI section 4), but instead of using the simulated numbers of cells directly for our estimates, we wrote a dilution-plating subroutine to simulate how cell density would be gauged in the lab. Suppose that a cell population has an actual density of N_0 cells/mL. A 10-fold dilution series is generated recursively by diluting 100mL into 900mL. Thus, the density of cells in the first dilution is:

$$N_{-1} = rv[\text{Poisson}(0.1N_0)]$$

where $rv[d]$ is a random value for a variable with a distribution given by d. The density of cells in the second dilution is:

$$N_{-2} = rv[\text{Poisson}(0.1N_{-1})].$$

More generally, the i^{th} dilution has density:

$$N_{-i} = rv\left[\text{Poisson}\left(0.1N_{-(i-1)}\right)\right]$$

Now 100mL of each dilution in the entire series is plated, where the number of bacterial cells from the i^{th} dilution landing on the plate is:

$$B_{-i} = rv[\text{Poisson}(0.1N_{-i})]$$

Finally, the number of colonies forming (given an extinction probability of π) on the i^{th} dilution plate is:

$$C_{-i} = rv[\text{Binomial}(B_{-i}, 1-\pi)]$$

We pick the dilution plate with the maximum number of colonies in the range between 30 and 300. If every dilution plate is below 30 colonies, we simply use the plate with the maximum number of colonies. For generality, let's suppose we select the i^{th} dilution plate. We compute the cell density of the undiluted culture as:

$$N_{\text{est}} = \frac{C_{-i}}{1-\pi} \times 10^{i+1} \ \frac{\text{cells}}{\text{mL}}$$

Given the random effects of dilution, plating, and cell lineage extinction, it is likely that N_{est} will deviate from the actual cell density N_0.

For the SIM estimate, we use this procedure to generate the density of donors, recipients and transconjugants that are used in the estimate. For the LDM estimate, we use this procedure to generate the density of donors and recipients that are used in the estimate. Also, if the extinction probability of transconjugants in the wells is non-zero, we must also track a monoculture of transconjugants in order to estimate the transconjugant growth rate needed for the LDM correction (equation [7.1]), and we use the above procedure to estimate the transconjugant densities in these cases.

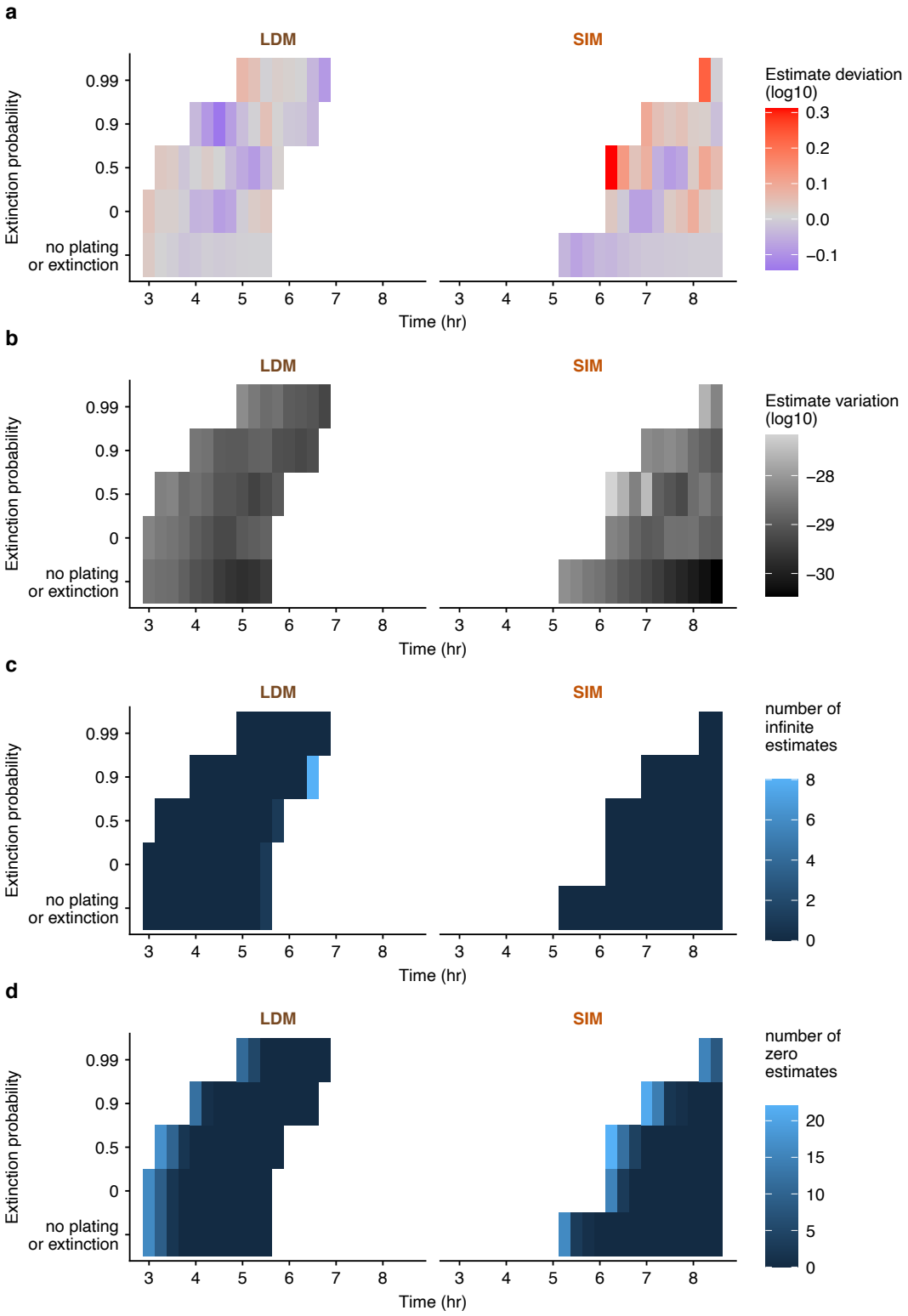
Random effects in wells with transconjugants: However, we also need to calculate the fraction of wells with transconjugant-selecting medium that are not turbid for the LDM estimate. Here the actual simulated number of transconjugants in a given population at the end of the assay is $T_{\tilde{t}}$. The number of lineages that do not go extinct is

$$L_{-i} = rv[\text{Binomial}(T_{\tilde{t}}, 1-\pi)]$$

If $L_{-i} > 0$, then the well is turbid, whereas if $L_{-i} = 0$, then the well is non-turbid. The proportion of non-turbid wells out of a total of W wells (P_{nt}) can then be calculated. If we have this quantity and all the relevant cell densities, we can then use equation [7.2] to calculate the corrected LDM estimate.

Results: We show the results of adding these random effects in SI Figure 12. Each rectangle represents 100 estimates for a combination of the incubation time ($\tilde{t}$) and an extinction probability (π), which, for simplicity, we assume is the same for all cell types both on plates and in wells. For reference, estimates without the random effects of dilution, plating, and extinction are given in the bottom row of each plot. Estimates with the random effects of only dilution and plating can be found in the row with zero extinction probability in each plot. We note that as the extinction probability increases, the end point of the assay must also increase (to obtain sufficient colonies and turbid wells), thus, the range of incubation times shift with this quantity.

As random effects are added, both the LDM and SIM estimates of the donor conjugation rate tend to deviate more from the actual value, but there is not systematic deviation (SI Figure 12a). Not surprisingly, as random effects are added, the variance in estimates rises, but this effect is more pronounced for the SIM estimate (SI Figure 12b). For both approaches, a zero estimate is possible (when there are no transconjugant colonies or no turbid transconjugant wells) and for the LDM estimate an infinite estimate is possible (when all the transconjugant wells are turbid). However, we see these extreme values occur primarily at the boundaries of the time interval for incubation times (SI Figure 12c and 12d).



SI Figure 12 : The random effects of dilution, plating, and failure to establish on the accuracy and variance of the LDM and SIM estimates. Different extinction probabilities are used, as indicated. The parameter values and initial densities are the same as SI Figure 5a which were $\psi_D = \psi_R = \psi_T = 1$ and $\gamma_D = \gamma_T = 1 \times 10^{-14}$. The dynamic variables were initialized with $D_0 = R_0 = 10^5$ and $T_0 = 0$. The scenario with no dilution plating and a zero-extinction probability (the bottom row in each panel) is the data from SI Figure 5a. The mean deviation (a) and variation (b) of each set of estimates is given

at 15-minute time intervals where at least 75 out of the 100 calculate estimates produced a finite non-zero value. (c) The number of infinite estimates out of the 100 calculated in the relevant intervals. (d) The number of estimates with a zero value out of the 100 calculated in the relevant intervals. We note that the Gillespie algorithm is computationally expensive when the densities get very large. Therefore, due to the longer incubation times needed for the SIM, only 100 populations of the 10,000 were simulated through the later time intervals until on average a population density of 1 x 10^9 is reached (i.e., $\tilde{t}$ = 8.5 h). The remaining 9,900 populations, used to compute $\hat{p}_0(\tilde{t})$ for the LDM, were run until an average of 100 transconjugants was reached (i.e., $\tilde{t}$ = 6.9 h). This explains the truncation of the SIM estimates at 8.5 hours and the LDM estimates 6.75 hours, which is most notable in the scenario where the extinction probability is 0.99. The data and code needed to generate this Figure can be found at https://github.com/livkosterlitz/LDM or https://doi.org/10.5281/zenodo.6677158.

CHAPTER 1 appendix:

Appendix I : TDR derivation

We present a few derivations for the TDR method. The derivations differ by the assumptions used to solve for TDR. We use this section to show the conditions where the TDR method can be used. In addition, we use this section to walk through all of the mathematical steps in order to provide an easy reference for comparing the derivations of all four methods (TDR, SIM, ASM, and LDM). The original condition presented by Levin *et. al.* (10) is the last derivation described in this section.

Here we assume that we are considering conditions where population growth is not occurring. These conditions could involve an environment hindering cell division. Alternatively, the environment could promote growth, but we restrict the time period sufficiently so that population change is negligible. Or finally, if all strains possess the same growth rate, the population could be growing over longer periods of time but doing so in a continuous flow-through device (e.g., a chemostat) such that population density remains constant, see ref (10). Because we assume no plasmid loss of the plasmid, the donor population must remain constant

$$D_t = D_0, \qquad \text{[I.a]}$$

for any time t under consideration. Importantly, even though growth is not occurring, conjugation can proceed. We assume $T_0 = 0$, but the population of transconjugants can increase over time. Because every transconjugant was formerly a recipient cell, it must be the case that

$$R_t + T_t = R_0,$$

for any time t under consideration. Therefore,

$$R_t = R_0 - T_{t.} \qquad \text{[I.b]}$$

The dynamics of transconjugants is given by

$$\frac{dT_t}{dt} = \gamma_D D_t R_t + \gamma_T T_t R_t \qquad \text{[I.c]}$$

By substituting terms from equations [I.a] and [I.b]

$$\frac{dT_t}{dt} = \gamma_D D_0 (R_0 - T_t) + \gamma_T T_t (R_0 - T_t),$$

$$\frac{dT_t}{dt} = (\gamma_D D_0 + \gamma_T T_t)(R_0 - T_t),$$

$$\frac{dT_t}{dt} = -\gamma_T \left(T_t + \frac{\gamma_D D_0}{\gamma_T}\right)(T_t - R_0),$$

We can solve this differential equation by a separation of variables:

$$\int_0^{\tilde{t}} \frac{dT_t}{-\gamma_T \left(T_t + \frac{\gamma_D D_0}{\gamma_T}\right)(T_t - R_0)} = \int_0^{\tilde{t}} dt.$$

The following identity is relevant here:

$$\frac{d\left\{\frac{1}{a(b-c)}\ln\frac{x-b}{x-c}\right\}}{dx}=\frac{1}{a(x-b)(x-c)}.$$

Letting $x=T_t$, $a=-\gamma_T$, $b=-\frac{\gamma_D D_0}{\gamma_T}$, and $c=R_0$, we can proceed as follows:

$$\left\{\frac{1}{\gamma_T\left(\frac{\gamma_D D_0}{\gamma_T}+R_0\right)}\ln\frac{T_t+\frac{\gamma_D D_0}{\gamma_T}}{T_t-R_0}\right\}\Bigg|_0^{\tilde{t}}=(t)|_0^{\tilde{t}},$$

$$\frac{1}{\gamma_T\left(\frac{\gamma_D D_0}{\gamma_T}+R_0\right)}\left(\ln\frac{T_{\tilde{t}}+\frac{\gamma_D D_0}{\gamma_T}}{T_{\tilde{t}}-R_0}-\ln\frac{T_0+\frac{\gamma_D D_0}{\gamma_T}}{T_0-R_0}\right)=\tilde{t}.$$

Because $T_0=0$,

$$\frac{1}{\gamma_T\left(\frac{\gamma_D D_0}{\gamma_T}+R_0\right)}\left(\ln\frac{T_{\tilde{t}}+\frac{\gamma_D D_0}{\gamma_T}}{T_{\tilde{t}}-R_0}-\ln\frac{\frac{\gamma_D D_0}{\gamma_T}}{-R_0}\right)=\tilde{t},$$

$$\frac{1}{\gamma_T\left(\frac{\gamma_D D_0}{\gamma_T}+R_0\right)}\left(\ln\frac{-R_0\left(T_{\tilde{t}}+\frac{\gamma_D D_0}{\gamma_T}\right)}{\frac{\gamma_D D_0}{\gamma_T}(T_{\tilde{t}}-R_0)}\right)=\tilde{t},$$

$$\ln\frac{1+\frac{\gamma_T T_{\tilde{t}}}{\gamma_D D_0}}{1-\frac{T_{\tilde{t}}}{R_0}}=(\gamma_D D_0+\gamma_T R_0)\tilde{t}, \qquad \text{[I.d]}$$

$$\frac{1+\frac{\gamma_T T_{\tilde{t}}}{\gamma_D D_0}}{1-\frac{T_{\tilde{t}}}{R_0}}=\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\},$$

$$1+\frac{\gamma_T T_{\tilde{t}}}{\gamma_D D_0}=\left(1-\frac{T_{\tilde{t}}}{R_0}\right)\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\},$$

$$\frac{\gamma_T T_{\tilde{t}}}{\gamma_D D_0}+\frac{T_{\tilde{t}}}{R_0}\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}=\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}-1,$$

$$T_{\tilde{t}}\left[\frac{\gamma_T}{\gamma_D D_0}+\frac{1}{R_0}\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}\right]=\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}-1,$$

$$T_{\tilde{t}}=\frac{R_0(\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}-1)}{\frac{\gamma_T R_0}{\gamma_D D_0}+\exp\{(\gamma_D D_0+\gamma_T R_0)\tilde{t}\}}. \qquad \text{[I.e]}$$

Thus, equation [I.e] is a general solution for the number of transconjugants at any time. If we assume that $\gamma_T=0$, then we can rewrite [I.d] as

$$\ln\frac{1}{1-\frac{T_{\tilde{t}}}{R_0}}=\gamma_D D_0\tilde{t},$$

$$-\ln\left(1-\frac{T_{\tilde{t}}}{R_0}\right)=\gamma_D D_0\tilde{t},$$

$$\gamma_D = \frac{-\ln\left(1-\frac{T_{\tilde{t}}}{R_0}\right)}{D_0\tilde{t}}. \qquad \text{[I.f]}$$

When $R_0 \gg T_{\tilde{t}}$, a first-order Taylor approximation ensures $-\ln\left(1-\frac{T_{\tilde{t}}}{R_0}\right) \approx \frac{T_{\tilde{t}}}{R_0}$, and therefore

$$\gamma_D \approx \frac{\frac{T_{\tilde{t}}}{R_0}}{D_0\tilde{t}},$$

$$\gamma_D \approx \frac{T_{\tilde{t}}}{D_0 R_0 \tilde{t}}.$$

Because $D_{\tilde{t}} = D_0$ (see [I.a]) and $R_{\tilde{t}} \approx R_0$ (when $R_0 \gg T_{\tilde{t}}$ by [I.b]), we have

$$\gamma_D \approx \frac{T_{\tilde{t}}}{D_{\tilde{t}} R_{\tilde{t}} \tilde{t}}.$$

On the other hand, if we assume $\gamma_D = \gamma_T = \gamma$, then we can rewrite [I.d] as

$$\ln\frac{1+\frac{T_{\tilde{t}}}{D_0}}{1-\frac{T_{\tilde{t}}}{R_0}} = \gamma(D_0 + R_0)\tilde{t},$$

$$\gamma = \frac{1}{\tilde{t}(D_0+R_0)}\ln\frac{1+\frac{T_{\tilde{t}}}{D_0}}{1-\frac{T_{\tilde{t}}}{R_0}},$$

$$\gamma = \frac{1}{\tilde{t}(D_0+R_0)}\left\{\ln\left(1+\frac{T_{\tilde{t}}}{D_0}\right) - \ln\left(1-\frac{T_{\tilde{t}}}{R_0}\right)\right\}. \qquad \text{[I.g]}$$

When $D_0 \gg T_{\tilde{t}}$ and $R_0 \gg T_{\tilde{t}}$, first-order Taylor approximations ensure

$$\gamma \approx \frac{1}{\tilde{t}(D_0+R_0)}\left(\frac{T_{\tilde{t}}}{D_0} + \frac{T_{\tilde{t}}}{R_0}\right),$$

$$\gamma \approx \frac{T_{\tilde{t}}}{\tilde{t}(D_0+R_0)}\left(\frac{1}{D_0} + \frac{1}{R_0}\right),$$

$$\gamma \approx \frac{T_{\tilde{t}}}{\tilde{t}(D_0+R_0)}\left(\frac{R_0+D_0}{D_0 R_0}\right),$$

$$\gamma \approx \frac{T_{\tilde{t}}}{D_0 R_0 \tilde{t}}.$$

Because $D_{\tilde{t}} = D_0$ and $R_{\tilde{t}} \approx R_0$ (when $R_0 \gg T_{\tilde{t}}$), we have

$$\gamma \approx \frac{T_{\tilde{t}}}{D_{\tilde{t}} R_{\tilde{t}} \tilde{t}}.$$

So we have general expressions for donor conjugation rate when $\gamma_T = 0$ (equation [I.f]) or when $\gamma_D = \gamma_T$ (equation [I.g]). However, when $D_0 \gg T_t$ and $R_0 \gg T_t$, for all t under consideration, both of these measures are well approximated by equation [1.4]. Here we extend the application of equation [1.4] even further. When $R_0 \gg T_t$, then $R_t \approx R_0$. Let us assume $R_t = R_0$. We will also assume

$$\gamma_D D_0 R_0 \gg \gamma_T T_t R_0, \qquad \text{[I.h]}$$

namely, the rate of formation of transconjugants by donors is much greater than the formation by transconjugants. Of course, if $0 \leq \gamma_T \leq \gamma_D$, then $D_0 \gg T_t$ ensures assumption [I.h]. In general, this inequality is satisfied when $T_0 = 0$, D_0 and R_0 are large, γ_T is not dramatically higher than γ_D, and the period is small. Under assumption [I.h], the dynamics can be well approximated by a simplified version of the differential equation [I.c]. (where the transconjugant conjugation term is gone). This is the differential equation originally solved by Levin *et. al.,*

$$\frac{dT_t}{dt} = \gamma_D D_0 R_0.$$

Because everything on the right-hand-side of the equation is a constant, the solution is straightforward:

$$\int_0^{\tilde{t}} dT_t = \int_0^{\tilde{t}} \gamma_D D_0 R_0 dt,$$
$$(T_{\tilde{t}})|_0^{\tilde{t}} = (\gamma_D D_0 R_0 t)|_0^{\tilde{t}},$$
$$T_{\tilde{t}} - T_0 = \gamma_D D_0 R_0 \tilde{t}.$$

Because $T_0 = 0$,

$$T_{\tilde{t}} = \gamma_D D_0 R_0 \tilde{t},$$
$$\gamma_D = \frac{T_{\tilde{t}}}{D_0 R_0 \tilde{t}}.$$

Because $D_{\tilde{t}} = D_0$ and $R_{\tilde{t}} = R_0$ (by assumption), we have recovered equation [1.4].

Appendix II : SIM derivation

In this section, we walk through the mathematical steps in the Simonsen *et. al.* derivation. We include this derivation as a quick reference to help the reader compare the derivations of all four methods (TDR, SIM, ASM, and LDM). Simonsen *et. al.* modified the model (equations [1]-[3]) by adding a dynamic variable for resource concentration (C). The dynamics of this resource incorporate an additional parameter for the amount of resource needed to produce a new cell (e). Both the growth rate and conjugation rate are taken to be functions of the resource concentration.

$$\frac{dD}{dt} = \psi(C)D, \qquad \text{[II.a]}$$

$$\frac{dR}{dt} = \psi(C)R - \gamma(C)R(D+T), \qquad \text{[II.b]}$$

$$\frac{dT}{dt} = \psi(C)T + \gamma(C)R(D+T), \qquad \text{[II.c]}$$

$$\frac{dC}{dt} = -\psi(C)(R+D+T)e. \qquad \text{[II.d]}$$

The other variables are consistent with their use in equations [1]-[3]. In Simonsen *et. al.*, when resources are depleted, growth and conjugation stop. A Monod function introduces batch culture dynamics (i.e., exponential and stationary phase) making growth and

conjugation both increase in a saturated manner with resource concentration, where the resource concentration yielding ½ the maximal rate is given by the parameter Q. Importantly, conjugation and growth are assumed to have the same functional form:

$$\psi(C) = \frac{\psi_{max} C}{Q + C}, \qquad \text{[II.e]}$$

$$\gamma(C) = \frac{\gamma_{max} C}{Q + C}, \qquad \text{[II.f]}$$

where Q is the half saturation constant.

More generally, let us assume

$$\psi(C) = \psi^{\blacksquare} g(C), \qquad \text{[II.g]}$$

$$\gamma(C) = \gamma^{\blacksquare} g(C), \qquad \text{[II.h]}$$

where $\psi^{\blacksquare}$ and $\gamma^{\blacksquare}$ are constants and $g(C)$ is some function. We see that equations [II.e] and [II.f] are a special case of equations [II.g] and [II.h]. For the most general case, the ratio of growth rate to conjugation rate is a constant:

$$\frac{\psi(C)}{\gamma(C)} = \frac{\psi^{\blacksquare}}{\gamma^{\blacksquare}}.$$

Here we derive the estimation of the conjugation rate parameter $\gamma^{\blacksquare}$ for this general case. Note, we can connect equations [II.a]-[II.d] this to the equations [1]-[3] if we assume $\psi_D = \psi_R = \psi_T = \psi(C) = \psi^{\blacksquare} g(C)$ and $\gamma_D = \gamma_T = \gamma(C) = \gamma^{\blacksquare} g(C)$. Simonsen *et. al.* assume that all strains have the same growth rate, and conjugation rates from donors and transconjugants are the same. Additionally, Simonsen *et. al.* assume no segregative loss. We define $N_t = D_t + R_t + T_t$. Therefore, using equations [II.a]-[II.c], and droping the t subscripts for notational convenience, we have:

$$\frac{dN}{dt} = \frac{dD}{dt} + \frac{dR}{dt} + \frac{dT}{dt},$$

$$\frac{dN}{dt} = \psi(C)D + \psi(C)R - \gamma(C)R(D + T) + \psi(C)T + \gamma(C)R(D + T),$$

$$\frac{dN}{dt} = \psi(C)(D + R + T),$$

$$\frac{dN}{dt} = \psi(C)N. \qquad \text{[II.i]}$$

Letting $X_t = D_t / N_t$, we have the following by the quotient rule:

$$\frac{dX}{dt} = \frac{\frac{dD}{dt} N - \frac{dN}{dt} D}{N^2}$$

$$\frac{dX}{dt} = \frac{\psi(C)DN - \psi(C)ND}{N^2}$$

$$\frac{dX}{dt} = 0$$

Thus, X_t does not change over time (i.e., $X_t = X_0$ for all t).

Lastly, we define a fraction $Y_t = T_t/R_t$. Using equations [II.b] and [II.c],

$$\frac{dY}{dt} = \frac{\frac{dT}{dt}R - \frac{dR}{dt}T}{R^2},$$
$$\frac{dY}{dt} = \frac{\{\psi(C)T + \gamma(C)R(D+T)\}R - \{\psi(C)R - \gamma(C)R(D+T)\}T}{R^2},$$
$$\frac{dY}{dt} = \frac{\psi(C)TR + \gamma(C)(D+T)R^2 - \psi(C)TR + \gamma(C)(D+T)TR}{R^2},$$
$$\frac{dY}{dt} = \frac{\gamma(C)(D+T)R^2 + \gamma(C)(D+T)TR}{R^2},$$
$$\frac{dY}{dt} = \frac{\gamma(C)(D+T)R + \gamma(C)(D+T)T}{R},$$
$$\frac{dY}{dt} = \frac{\gamma(C)(DR + TR + DT + T^2)}{R},$$
$$\frac{dY}{dt} = \gamma(C)\frac{DT + RT + T^2 + DR}{R},$$
$$\frac{dY}{dt} = \gamma(C)\left(\frac{T(D+R+T)}{R} + D\right).$$

Because $N_t = D_t + R_t + T_t$,

$$\frac{dY}{dt} = \gamma(C)\left(\frac{TN}{R} + D\right),$$
$$\frac{dY}{dt} = \gamma(C)N\left(\frac{T}{R} + \frac{D}{N}\right).$$

Because $X_t = D_t/N_t$ and $Y_t = T_t/R_t$,

$$\frac{dY}{dt} = \gamma(C)N(Y+X),$$
$$\left(\frac{1}{Y+X}\right)\frac{dY}{dt} = \gamma(C)N.$$

Equation [II.i] ensures $N = \left(\frac{1}{\psi(C)}\right)\frac{dN}{dt}$, and along with equations [II.g] and [II.h], we have

$$\left(\frac{1}{Y+X}\right)\frac{dY}{dt} = \gamma(C)\left(\frac{1}{\psi(C)}\right)\frac{dN}{dt},$$
$$\left(\frac{1}{Y+X}\right)\frac{dY}{dt} = \gamma^{\blacksquare}g(C)\left(\frac{1}{\psi^{\blacksquare}g(C)}\right)\frac{dN}{dt},$$
$$\psi^{\blacksquare}\left(\frac{1}{Y+X}\right)\frac{dY}{dt} = \gamma^{\blacksquare}\frac{dN}{dt},$$
$$\gamma^{\blacksquare}(dN) = \psi^{\blacksquare}\left(\frac{dY}{Y+X}\right).$$

To emphasize that X is a constant, we write this as X_0 and then integrate both sides,

$$\gamma^{\blacksquare}\int_0^{\tilde{t}} dN = \psi^{\blacksquare}\int_0^{\tilde{t}}\left(\frac{dY}{Y+X_0}\right),$$

Making the time dependence of the variables explicit via subscripts

$$\gamma^{\blacksquare}N_t|_0^{\tilde{t}} = \psi^{\blacksquare}\ln(Y_t + X_0)|_0^{\tilde{t}},$$
$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare}\{\ln(Y_{\tilde{t}} + X_0) - \ln(Y_0 + X_0)\}.$$

Because $X_0 = X_{\tilde{t}}$

$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare}\{\ln(Y_{\tilde{t}} + X_{\tilde{t}}) - \ln(Y_0 + X_{\tilde{t}})\},$$

$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare} \ln\left(\frac{Y_{\tilde{t}} + X_{\tilde{t}}}{Y_0 + X_{\tilde{t}}}\right).$$

Because $X_{\tilde{t}} = D_{\tilde{t}}/N_{\tilde{t}}$ and $Y_{\tilde{t}} = T_{\tilde{t}}/R_{\tilde{t}}$,

$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare} \ln\left(\frac{\frac{T_{\tilde{t}}}{R_{\tilde{t}}} + \frac{D_{\tilde{t}}}{N_{\tilde{t}}}}{\frac{T_0}{R_0} + \frac{D_{\tilde{t}}}{N_{\tilde{t}}}}\right).$$

Because $T_0 = 0$,

$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare} \ln\left(\frac{\frac{T_{\tilde{t}}}{R_{\tilde{t}}} + \frac{D_{\tilde{t}}}{N_{\tilde{t}}}}{\frac{D_{\tilde{t}}}{N_{\tilde{t}}}}\right),$$

$$\gamma^{\blacksquare}(N_{\tilde{t}} - N_0) = \psi^{\blacksquare} \ln\left(\frac{T_{\tilde{t}}}{R_{\tilde{t}}}\frac{N_{\tilde{t}}}{D_{\tilde{t}}} + 1\right),$$

$$\gamma^{\blacksquare} = \psi^{\blacksquare} \ln\left(1 + \frac{T_{\tilde{t}}}{R_{\tilde{t}}}\frac{N_{\tilde{t}}}{D_{\tilde{t}}}\right)\frac{1}{N_{\tilde{t}} - N_0}.$$

If we let $\psi^{\blacksquare} = \psi_{max}$ and $\gamma^{\blacksquare} = \gamma_{max}$ (with $g(C) = C/(Q + C)$) then we have recovered the SIM estimate for conjugation rate (equation [1.5]).

Appendix III : ASM derivation

In this section, we walk through all the mathematical steps in the Huisman *et. al.* (14) derivation. We include this derivation as a quick reference to help the reader compare the derivations of all four methods (TDR, SIM, ASM, and LDM). We start with a system of ordinary differential equations described by Huisman *et. al.,* the Extended Simonsen Model (ESM):

$$\frac{dD}{dt} = \psi_D(C)D, \qquad \text{[III.a]}$$

$$\frac{dR}{dt} = \psi_R(C)R - (\gamma_D(C)D + \gamma_T(C)T)R, \qquad \text{[III.b]}$$

$$\frac{dT}{dt} = \psi_T(C)T + (\gamma_D(C)D + \gamma_T(C)T)R, \qquad \text{[III.c]}$$

$$\frac{dC}{dt} = -(\psi_D(C)D + \psi_R(C)R + \psi_T(C)T)e, \qquad \text{[III.d]}$$

Where ψ subscripts specify population specific growth rates, γ subscripts specify donor and transconjugant specific conjugation rates, and other variables are consistent with equations [III.a]-[III.d]. As with the SIM, growth and conjugation rate are both dependent on resource concentration. When resources are depleted, growth and conjugation stop.

$$\psi_A(C) = \frac{\psi_{A,max}C}{Q + C},$$

$$\gamma_B(C) = \frac{\gamma_{B,max} C}{Q + C},$$

where $A \in \{D, R, T\}$ and $B \in \{D, T\}$.

Under the assumption that growth rate and conjugation rate are constant (where $\psi_A(C) = \psi_{A,max}$ and $\gamma_B(C) = \gamma_{B,max}$), equation [III.d] can be dropped. Although the maximal rates of growth and conjugation are assumed, in what follows, we drop the "max" subscript for notational convenience. This new set of simplified ordinary differential equations is termed the Approximate Simonsen *et. al.* Method ('ASM').

$$\frac{dD}{dt} = \psi_D D, \qquad \text{[III.e]}$$

$$\frac{dR}{dt} = \psi_R R - (\gamma_D D + \gamma_T T) R, \qquad \text{[III.f]}$$

$$\frac{dT}{dt} = \psi_T T + (\gamma_D D + \gamma_T T) R, \qquad \text{[III.g]}$$

If we assume that the recipient population is dominated by growth $\psi_R R \gg (\gamma_D D + \gamma_T T) R$ and the transconjugant population is dominated by growth and conjugation from donors $\psi_T T + \gamma_D D R \gg \gamma_T T R$, then we can replace equations [III.e]-[III.g] with the following set of differential equations, which do a good job approximating the dynamics:

$$\frac{dD}{dt} = \psi_D D, \qquad \text{[III.h]}$$

$$\frac{dR}{dt} = \psi_R R, \qquad \text{[III.i]}$$

$$\frac{dT}{dt} = \psi_T T + \gamma_D D R, \qquad \text{[III.j]}$$

The solutions to differential equations [III.h] and [III.i] are:

$$D_t = D_0 e^{\psi_D t} \qquad \text{[III.k]}$$

$$R_t = R_0 e^{\psi_R t} \qquad \text{[III.l]}$$

Here we will derive the solution for the differential equation [III.j] for transconjugant density T. We know that $D_t = D_0 e^{\psi_D t}$ and $R_t = R_0 e^{\psi_R t}$, therefore

$$\frac{dT}{dt} = \psi_T T + \gamma_D D_0 R_0 e^{(\psi_D + \psi_R) t} \qquad \text{[III.m]}$$

We propose that the transconjugant density can be written as a product of time dependent functions:

$$T_t = u_t v_t \qquad \text{[III.n]}$$

We'll drop the t subscripts for notational ease. By the product rule,

$$\frac{dT}{dt} = u \frac{dv}{dt} + v \frac{du}{dt} \qquad \text{[III.o]}$$

We can rewrite equation [III.m] by plugging in equations [III.n] and [III.o] as follows:

$$u\frac{dv}{dt} + v\frac{du}{dt} = \psi_T uv + \gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t}$$

$$u\frac{dv}{dt} + v\left(\frac{du}{dt} - \psi_T u\right) = \gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t} \qquad \text{[III.p]}$$

We have some freedom to pick u_t as we please. So, let's choose a function such that the second term of the left-hand side of equation [III.p] is zero:

$$\frac{du}{dt} - \psi_T u = 0$$

$$\frac{du}{dt} = \psi_T u$$

The solution to this differential equation is:

$$u_t = u_0 e^{\psi_T t} \qquad \text{[III.q]}$$

where u_0 is a constant.

Thus, we can rewrite equation [III.p] as:

$$u_0 e^{\psi_T t}\frac{dv}{dt} = \gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t}$$

$$u_0\frac{dv}{dt} = \frac{\gamma_D D_0 R_0 e^{(\psi_D+\psi_R)t}}{e^{\psi_T t}}$$

$$u_0\frac{dv}{dt} = \gamma_D D_0 R_0 e^{(\psi_D+\psi_R-\psi_T)t}$$

To solve this equation, we integrate

$$u_0\int dv = \gamma_D D_0 R_0 \int e^{(\psi_D+\psi_R-\psi_T)t}dt$$

$$u_0 v_t = \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T} e^{(\psi_D+\psi_R-\psi_T)t} + c \qquad \text{[III.r]}$$

where c is a constant of the integration. To solve for c, plug in $t = 0$,

$$u_0 v_0 = \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T} e^{(\psi_D+\psi_R-\psi_T)0} + c$$

$$u_0 v_0 = \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T} + c$$

$$c = u_0 v_0 - \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T}$$

Because $T_0 = u_0 v_0$,

$$c = T_0 - \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T}$$

So, we now can find the solution for $v_{\tilde{t}}$ by plugging in our solution for c into equation [III.r]:

$$u_0 v_{\tilde{t}} = \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T} e^{(\psi_D+\psi_R-\psi_T)\tilde{t}} + T_0 - \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T}$$

$$u_0 v_{\tilde{t}} = T_0 + \frac{\gamma_D D_0 R_0}{\psi_D+\psi_R-\psi_T}\left\{e^{(\psi_D+\psi_R-\psi_T)\tilde{t}} - 1\right\}$$

$$v_{\tilde{t}} = \frac{1}{u_0}\left(T_0 + \frac{\gamma_D D_0 R_0}{\psi_D + \psi_R - \psi_T}\left\{e^{(\psi_D+\psi_R-\psi_T)\tilde{t}} - 1\right\}\right) \qquad \text{[III.s]}$$

Because $T_{\tilde{t}} = u_{\tilde{t}} v_{\tilde{t}}$ through substitution of equations [III.q] and [III.s] and, we have

$$T_{\tilde{t}} = \left[u_0 e^{\psi_T \tilde{t}}\right]\left[\frac{1}{u_0}\left(T_0 + \frac{\gamma_D D_0 R_0}{\psi_D + \psi_R - \psi_T}\left\{e^{(\psi_D+\psi_R-\psi_T)\tilde{t}} - 1\right\}\right)\right]$$

$$T_{\tilde{t}} = e^{\psi_T \tilde{t}}\left(T_0 + \frac{\gamma_D D_0 R_0}{\psi_D + \psi_R - \psi_T}\left\{e^{(\psi_D+\psi_R-\psi_T)\tilde{t}} - 1\right\}\right)$$

$$T_{\tilde{t}} = T_0 e^{\psi_T \tilde{t}} + \frac{\gamma_D D_0 R_0}{\psi_D + \psi_R - \psi_T}\left\{e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right\}$$

Because $T_0 = 0$, we arrive at the result in Huisman *et al*.

$$T_{\tilde{t}} = \frac{\gamma_D D_0 R_0}{\psi_D + \psi_R - \psi_T}\left\{e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right\}$$

Given that $D_{\tilde{t}} = D_0 e^{\psi_D \tilde{t}}$ and $R_{\tilde{t}} = R_0 e^{\psi_R \tilde{t}}$, this can be rewritten as

$$T_{\tilde{t}} = \frac{\gamma_D}{\psi_D + \psi_R - \psi_T}\left\{D_{\tilde{t}} R_{\tilde{t}} - D_0 R_0 e^{\psi_T \tilde{t}}\right\}$$

Solving for γ_D gives equation [1.9].

Appendix IV : The LDM MLE derivation

We start by focusing on $p_0(\tilde{t})$, the probability that a population will have no transconjugants at time $\tilde{t}$ (for notational convenience we'll drop the time variable). What we actually measure is the number of independent populations (or wells) that have no transconjugants (call this w) out of the total number of populations (or wells) tracked (call this W). What is our best estimate of p_0, given our data w? That is, what value p_0 maximizes the likelihood function:

$$\mathcal{L}(p_0) = \Pr\{w|p_0\} = \binom{\mathrm{W}}{w}(p_0)^w(1-p_0)^{\mathrm{W}-w}$$

Because $-\ln(x)$ is a monotonic decreasing function, the value p_0 that maximizes $\mathcal{L}(p_0)$ will be the same value that minimizes:

$$\mathrm{L}(p_0) = -\ln\{\mathcal{L}(p_0)\}$$

This can be rewritten as follows:

$$\mathrm{L}(p_0) = -\ln\left\{\binom{\mathrm{W}}{w}(p_0)^w(1-p_0)^{\mathrm{W}-w}\right\}$$

$$\mathrm{L}(p_0) = -\ln\binom{\mathrm{W}}{w} - w\ln p_0 - (\mathrm{W}-w)\ln(1-p_0)$$

To find the maximum likelihood estimate, we find the critical points of L by setting its derivative to zero, or:

$$\frac{d\mathrm{L}}{dp_0} = -\frac{w}{p_0} + \frac{\mathrm{W}-w}{1-p_0} = 0$$

The maximum likelihood estimate for p_0 (which we'll denote $\hat{p}_0$) solves the above equation, or

$$\frac{\mathrm{W}-w}{1-\hat{p}_0} = \frac{w}{\hat{p}_0}$$

$$(\mathrm{W}-w)\hat{p}_0 = w(1-\hat{p}_0)$$
$$\mathrm{W}\hat{p}_0 - w\hat{p}_0 = w - w\hat{p}_0$$
$$\mathrm{W}\hat{p}_0 = w$$
$$\hat{p}_0 = \frac{w}{\mathrm{W}}$$

This answer makes intuitive sense: the most likely estimate for p_0 (the probability that a population has no transconjugants) is simply the fraction of the populations that have no transconjugants.

To double check that this estimate actually corresponds to a *minimum* of $\mathrm{L}(p_0)$, consider the second derivative:

$$\frac{d^2\mathrm{L}}{dp_0{}^2} = \frac{w}{p_0{}^2} + \frac{\mathrm{W}-w}{(1-p_0)^2}$$

Evaluating this derivative at the critical point yields:

$$\left.\frac{d^2\mathrm{L}}{dp_0{}^2}\right|_{p_0=\hat{p}_0} = \frac{w}{\left(\frac{w}{\mathrm{W}}\right)^2} + \frac{\mathrm{W}-w}{\left(\frac{\mathrm{W}-w}{\mathrm{W}}\right)^2}$$
$$\left.\frac{d^2\mathrm{L}}{dp_0{}^2}\right|_{p_0=\hat{p}_0} = \mathrm{W}^2\left(\frac{1}{w} + \frac{1}{\mathrm{W}-w}\right)$$

As long as $0 < w < \mathrm{W}$,

$$\left.\frac{d^2\mathrm{L}}{dp_0{}^2}\right|_{p_0=\hat{p}_0} > 0,$$

which means that L is convex at $\hat{p}_0$ and therefore, this value is a local minimum. In turn, this means that $\hat{p}_0$ is a local maximum for $\mathcal{L}$.

Appendix V : Derivations of the first and second central moments

In this section, we derive the equations for the moments using terminology close to what Keller and Antal (43) used (focusing on the case where $\kappa \notin \{1,2\}$). From SI section 7, $N = e^{\delta t}$ and with $\kappa = \frac{\delta}{\alpha}$, $N^{-1/\kappa} = e^{-\alpha t}$. Keller and Antal denote the number of mutants (analogous to our transconjugants) as B. Also, letting $\xi = \frac{z}{z-1}$ and $\mu = \frac{\nu}{\alpha}$ and dropping the time argument for notational convenience, the probability generating function (PGF) can be written as:

$$G_B(z) = \exp\left\{\frac{N\mu}{\kappa}\left[\frac{1}{N}F\left(1,\kappa;1+\kappa;\xi N^{-\frac{1}{\kappa}}\right) - F(1,\kappa;1+\kappa;\xi)\right]\right\}.$$

The expected value for the number of mutants can be obtained from the PGF in the usual way:

$$E[B] = G_B'(1-) = \lim_{z\to 1-} G_B(z)\frac{N\mu}{1+\kappa}\frac{1}{(z-1)^2}\Sigma,$$

where

$$\Sigma = -N^{-1-\frac{1}{\kappa}} F\left(2,1+\kappa;2+\kappa;\xi N^{-\frac{1}{\kappa}}\right) + F(2,1+\kappa;2+\kappa;\xi).$$

We use the following identity

$$F(a,b;c;z) = (1-z)^{c-a-b} F(c-a,c-b;c;z),$$

to obtain

$$\begin{aligned} F(2,1+\kappa;2+\kappa;\xi) &= (1-\xi)^{-1} F(\kappa,1,2+\kappa,\xi) \\ &= (1-z) F\left(\kappa,1,2+\kappa,\frac{z}{z-1}\right) \end{aligned}$$

Using a related identity

$$F(a,b;c;z) = (1-z)^{-b} F\left(c-a,b;c;\frac{z}{z-1}\right),$$

we get:

$$\begin{aligned} F(2,1+\kappa;2+\kappa;\xi) &= (1-z) F\left(\kappa,1,2+\kappa,\frac{z}{z-1}\right) \\ &= (1-z)^2 F(2,1,2+\kappa,z) \end{aligned}$$

We will use a similar procedure for the first term in Σ, obtaining:

$$\begin{aligned} F(2,1+\kappa;2+\kappa;\xi N^{-1/\kappa}) &= (1-\xi N^{-1/\kappa})^{-1} F(\kappa,1,2+\kappa;\xi N^{-1/\kappa}) \\ &= (1-\xi N^{-1/\kappa})^{-1} F\left(\kappa,1,2+\kappa;\frac{x}{x-1}\right) \\ &= (1-\xi N^{-1/\kappa})^{-1} (1-x) F(2,1,2+\kappa;x) \\ &= \frac{(z-1)^2}{(zN^{-1/\kappa}-z+1)^2} F(2,1;2+\kappa;x) \end{aligned}$$

where $x = \frac{zN^{-1/\kappa}}{zN^{-1/\kappa}-z+1}$.

In sum, we have shown that the first derivative of the PGF can be expressed in the following way:

$$G_B'(z) = G_B(z) \frac{N\mu}{1+\kappa} \left(-\frac{N^{-1-\frac{1}{\kappa}}}{(zN^{-\frac{1}{\kappa}}-z+1)^2} F(2,1;2+\kappa;x) + F(2,1,2+\kappa,z) \right). \quad \text{[V.a]}$$

Using the fact that $\lim_{z\to 1-} G_B(z) = 1$ and the Gauss identity $F(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}$ (which holds if $c > a+b$) we obtain

$$E[B] = G_B'(1-) = \frac{N\mu}{\kappa-1}(-N^{-1+1/\kappa}+1).$$

To calculate the variance for the number of mutants, we note that

$$G_B''(1-) = E[B^2] - E[B],$$

and thus

$$\begin{aligned} \mathrm{Var}[B] &= G_B''(1-) + E[B] - (E[B])^2 \\ &= G_B''(1-) + G_B'(1-) - (G_B'(1-))^2 \end{aligned}$$

From equation [V.a] we can express the first derivative of the PGF as:

$$G_B'(z) = G_B(z) \frac{N\mu}{1+\kappa} \Sigma_1.$$

It follows that

$$\begin{aligned}\mathrm{Var}[B] &= G_B''(1-) + G_B'(1-) - (G_B'(1-))^2 \\ &= \frac{N\mu}{1+\kappa}\lim_{z\to 1-}(\Sigma_1 + \Sigma_1') \\ &= E[B] + \frac{N\mu}{1+\kappa}\lim_{z\to 1-}\Sigma_1'\end{aligned}$$

To calculate the derivative of Σ_1, we note that

$$\Sigma_1 = -\frac{N^{-1-\frac{1}{\kappa}}}{(zN^{-\frac{1}{\kappa}} - z + 1)^2}F(2,1;2+\kappa;x) + F(2,1,2+\kappa,z).$$

We can calculate the derivative of the second term:

$$\frac{\partial F(2,1,2+\kappa,z)}{\partial z} = \frac{2}{2+\kappa}F(3,2;3+\kappa;z),$$

and of the first term:

$$\begin{aligned}&\frac{\partial[-\frac{N^{-1-1/\kappa}}{(zN^{-1/\kappa} - z + 1)^2}F(2,1;2+\kappa;x)]}{\partial z} \\ &= -N^{-1-1/\kappa}\left(-\frac{2(N^{-1/\kappa}-1)}{(N^{-1/\kappa}z - z + 1)^3}F(2,1;2+\kappa;x)\right. \\ &\left. + \frac{2N^{-1/\kappa}}{(N^{-1/\kappa}z - z + 1)^4}\frac{1}{2+\kappa}F(3,2;3+\kappa;x)\right)\end{aligned}$$

Taken together, and using Gauss' equality,

$$\begin{aligned}\lim_{z\to 1-}\Sigma_1' &= 2(N^{-1+1/\kappa} - N^{-1+2/\kappa})\frac{\kappa+1}{\kappa-1} + \frac{2}{2+\kappa}\frac{(\kappa+2)(\kappa+1)}{(\kappa-1)(\kappa-2)}(1 - N^{-1+2/\kappa}) \\ &= 2N^{-1+2/\kappa}\frac{\kappa+1}{\kappa-1}\left(-1 - \frac{1}{\kappa-2}\right) + 2N^{-1+1/\kappa}\frac{\kappa+1}{\kappa-1} + \frac{2(\kappa+1)}{(\kappa-1)(\kappa-2)} \\ &= 2N^{-1+2/\kappa}\frac{\kappa+1}{2-\kappa} + 2N^{-1+1/\kappa}\frac{\kappa+1}{\kappa-1} + \frac{2(\kappa+1)}{(\kappa-1)(\kappa-2)}\end{aligned}$$

Coming back to

$$\mathrm{Var}[B] = E[B] + \frac{N\mu}{1+\kappa}\lim_{z\to 1-}\Sigma_1',$$

we obtain

$$\mathrm{Var}[B] = N\mu\left[\frac{2}{2-\kappa}N^{-1+\frac{2}{\kappa}} + \frac{1}{\kappa-1}N^{-1+\frac{1}{\kappa}} + \frac{\kappa}{(\kappa-1)(\kappa-2)}\right].$$

Making the appropriate substitutions yields our expressions for the mean and variance for our transconjugants in SI section 7.

Appendix VI : Behavior of the variance relative to the mean over time

In this section, we will show that the variance in transconjugant numbers grows relative to the mean over time. To do so, we will need to consider several cases.

Let's consider the ratio of the variance to the mean, which we denote $\rho_t = \frac{\mathrm{Var}[T_t]}{E[T_t]}$. Using the results from SI section 7, we will start by focusing on the case where $\psi_D + \psi_R \neq \psi_T$ and $\psi_D + \psi_R \neq 2\psi_T$:

$$\rho_t = \frac{\gamma_D D_0 R_0 \left\{ \frac{2e^{2\psi_T t}(\psi_T - (\psi_D + \psi_R)) - e^{\psi_T t}(2\psi_T - (\psi_D + \psi_R)) + (\psi_D + \psi_R)e^{(\psi_D + \psi_R)t}}{(2\psi_T - (\psi_D + \psi_R))(\psi_T - (\psi_D + \psi_R))} \right\}}{\frac{\gamma_D D_0 R_0 (e^{(\psi_D + \psi_R)t} - e^{\psi_T t})}{\psi_D + \psi_R - \psi_T}}.$$

This can be simplified as follows:

$$\rho_t = \frac{2e^{2\psi_T t}(\psi_T - (\psi_D + \psi_R)) - e^{\psi_T t}(2\psi_T - (\psi_D + \psi_R)) + (\psi_D + \psi_R)e^{(\psi_D + \psi_R)t}}{(\psi_D + \psi_R - 2\psi_T)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})}.$$

Letting $\kappa = \frac{\psi_D + \psi_R}{\psi_T}$,

$$\rho_t = \frac{2e^{2\psi_T t}(1 - \kappa) - e^{\psi_T t}(2 - \kappa) + \kappa e^{(\psi_D + \psi_R)t}}{(\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})},$$

$$\rho_t = \frac{\kappa e^{(\psi_D + \psi_R)t} - 2(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)e^{\psi_T t}}{(\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})}.$$

To determine the behavior of this ratio over time we take the derivative with respect to time:

$$\frac{d\rho_t}{dt} = \frac{\left\{ \begin{array}{c} \left(\kappa(\psi_D + \psi_R)e^{(\psi_D + \psi_R)t} - 4\psi_T(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)\psi_T e^{\psi_T t}\right)(\kappa - 2)\left(e^{(\psi_D + \psi_R)t} - e^{\psi_T t}\right) \\ -\left(\kappa e^{(\psi_D + \psi_R)t} - 2(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)e^{\psi_T t}\right)(\kappa - 2)\left((\psi_D + \psi_R)e^{(\psi_D + \psi_R)t} - \psi_T e^{\psi_T t}\right) \end{array} \right\}}{\left((\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})\right)^2}$$

$$= \frac{\left\{ \begin{array}{c} \left(\kappa(\psi_D + \psi_R)e^{(\psi_D + \psi_R)t} - 4\psi_T(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)\psi_T e^{\psi_T t}\right)(\kappa - 2)\left(e^{(\psi_D + \psi_R)t}\right) \\ -\left(\kappa(\psi_D + \psi_R)e^{(\psi_D + \psi_R)t} - 4\psi_T(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)\psi_T e^{\psi_T t}\right)(\kappa - 2)\left(e^{\psi_T t}\right) \\ -\left(\kappa e^{(\psi_D + \psi_R)t} - 2(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)e^{\psi_T t}\right)(\kappa - 2)\left((\psi_D + \psi_R)e^{(\psi_D + \psi_R)t}\right) \\ +\left(\kappa e^{(\psi_D + \psi_R)t} - 2(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)e^{\psi_T t}\right)(\kappa - 2)\left(\psi_T e^{\psi_T t}\right) \end{array} \right\}}{\left((\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{\left\{ \begin{array}{c} \left(-4\psi_T(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)\psi_T e^{\psi_T t}\right)(\kappa - 2)\left(e^{(\psi_D + \psi_R)t}\right) \\ -\left(\kappa(\psi_D + \psi_R)e^{(\psi_D + \psi_R)t} - 4\psi_T(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)\psi_T e^{\psi_T t}\right)(\kappa - 2)\left(e^{\psi_T t}\right) \\ -\left(-2(\kappa - 1)e^{2\psi_T t} + (\kappa - 2)e^{\psi_T t}\right)(\kappa - 2)\left((\psi_D + \psi_R)e^{(\psi_D + \psi_R)t}\right) \\ +\left(\kappa e^{(\psi_D + \psi_R)t} - 2(\kappa - 1)e^{2\psi_T t}\right)(\kappa - 2)\left(\psi_T e^{\psi_T t}\right) \end{array} \right\}}{\left((\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{\left\{ \begin{array}{c} (\psi_D + \psi_R - 2\psi_T)2(\kappa - 1)(\kappa - 2)e^{2\psi_T t}\left(e^{(\psi_D + \psi_R)t}\right) \\ -(\psi_D + \psi_R - \psi_T)(\kappa - 2)^2 e^{\psi_T t}\left(e^{(\psi_D + \psi_R)t}\right) \\ -(\psi_D + \psi_R - \psi_T)\kappa(\kappa - 2)e^{\psi_T t}\left(e^{(\psi_D + \psi_R)t}\right) \\ +2\psi_T(\kappa - 1)(\kappa - 2)e^{3\psi_T t} \end{array} \right\}}{\left((\kappa - 2)(e^{(\psi_D + \psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{\begin{Bmatrix}(\psi_D + \psi_R - 2\psi_T)2(\kappa-1)(\kappa-2)e^{2\psi_T t}\left(e^{(\psi_D+\psi_R)t}\right) \\ -(\psi_D + \psi_R - \psi_T)2(\kappa-1)(\kappa-2)e^{\psi_T t}\left(e^{(\psi_D+\psi_R)t}\right) \\ +2\psi_T(\kappa-1)(\kappa-2)e^{3\psi_T t}\end{Bmatrix}}{\left((\kappa-2)(e^{(\psi_D+\psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{(\kappa-1)(\kappa-2)\left\{(\kappa-2)e^{(\psi_D+\psi_R+2\psi_T)t} - (\kappa-1)e^{(\psi_D+\psi_R+\psi_T)t} + e^{3\psi_T t}\right\}}{\left(\frac{1}{2\psi_T}\right)\left((\kappa-2)(e^{(\psi_D+\psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{(\kappa-1)(\kappa-2)e^{\psi_T t}\left\{(\kappa-2)e^{(\psi_D+\psi_R+\psi_T)t} - (\kappa-1)e^{(\psi_D+\psi_R)t} + e^{2\psi_T t}\right\}}{\left(\frac{1}{2\psi_T}\right)\left((\kappa-2)(e^{(\psi_D+\psi_R)t} - e^{\psi_T t})\right)^2},$$

$$\frac{d\rho_t}{dt} = \frac{(\kappa-1)(\kappa-2)\left\{(\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}\right\}}{\left(\frac{e^{-\psi_T t}}{2\psi_T}\right)\left((\kappa-2)(e^{(\psi_D+\psi_R)t} - e^{\psi_T t})\right)^2}.$$

The denominator is always positive, so the sign of this derivative is governed by the numerator

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[(\kappa-1)(\kappa-2)\left\{(\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}\right\}\right].$$

For $\kappa > 2$, which is when the growth rate of the transconjugant is less than the average of the donor and recipient growth rates, we have

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[(\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}\right].$$

For any $t > 0$, for $\kappa = 2$

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[0 - e^{2\psi_T t} + e^{2\psi_T t}\right] = 0.$$

Now letting $g(\kappa) = (\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}$, we have

$$g'(\kappa) = e^{(\kappa+1)\psi_T t} + (\kappa-2)\psi_T t e^{(\kappa+1)\psi_T t} - e^{\kappa\psi_T t} - (\kappa-1)\psi_T t e^{\kappa\psi_T t},$$

$$g'(\kappa) = \{1 + (\kappa-2)\psi_T t\}e^{(\kappa+1)\psi_T t} - \{1 + (\kappa-1)\psi_T t\}e^{\kappa\psi_T t}.$$

Letting $h(x) = \{1 + (\kappa - 1 - x)\psi_T t\}e^{(\kappa+x)\psi_T t}$, we see

$$g'(\kappa) = h(1) - h(0).$$

But we see that

$$h'(x) = -\psi_T t e^{(\kappa+x)\psi_T t} + \{1 + (\kappa-1-x)\psi_T t\}\psi_T t e^{(\kappa+x)\psi_T t},$$

$$h'(x) = \psi_T t e^{(\kappa+x)\psi_T t}(-1 + \{1 + (\kappa-1-x)\psi_T t\}),$$

$$h'(x) = \psi_T t e^{(\kappa+x)\psi_T t}(\kappa-1-x)\psi_T t.$$

If $\kappa > 2$ and $t > 0$, then $h'(x) > 0$ for $0 \le x \le 1$. Since $h(x)$ is a continuous function, we must have

$$h(1) > h(0).$$

This implies that for $\kappa > 2$ and $t > 0$,

$$g'(\kappa) > 0.$$

Since $g(\kappa)$ is also continuous, and since $g(2) = 0$, we now know that for $\kappa > 2$ and $t > 0$,

$$g(\kappa) > 0.$$

Therefore $\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = 1$, or the derivative is positive for $t > 0$ and $\kappa > 2$, which means the variance grows relative to the mean over time.

For $1 < \kappa < 2$,

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[-(\kappa-2)e^{(\kappa+1)\psi_T t} + (\kappa-1)e^{\kappa\psi_T t} - e^{2\psi_T t}\right].$$

Letting $\varphi = -(\kappa-2)$, we know that $0 < \varphi < 1$, and we can rewrite the above

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[\varphi e^{(3-\varphi)\psi_T t} + (1-\varphi)e^{(3-(1+\varphi))\psi_T t} - e^{2\psi_T t}\right].$$

Because $e^{(3-x)\psi_T t}$ is convex $\left(t > 0 \Longrightarrow \frac{d^2 e^{(3-x)\psi_T t}}{dx^2} > 0\right)$, Jensen's inequality ensures

$$\varphi e^{(3-\varphi)\psi_T t} + (1-\varphi)e^{(3-(1+\varphi))\psi_T t} - e^{2\psi_T t} > e^{\left(3-\left(\varphi^2+(1-\varphi)(1+\varphi)\right)\right)\psi_T t} - e^{2\psi_T t},$$

$$\varphi e^{(3-\varphi)\psi_T t} + (1-\varphi)e^{(3-(1+\varphi))\psi_T t} - e^{2\psi_T t} > e^{\left(3-\left(\varphi^2+1-\varphi^2\right)\right)\psi_T t} - e^{2\psi_T t},$$

$$\varphi e^{(3-\varphi)\psi_T t} + (1-\varphi)e^{(3-(1+\varphi))\psi_T t} - e^{2\psi_T t} > e^{2\psi_T t} - e^{2\psi_T t},$$

$$\varphi e^{(3-\varphi)\psi_T t} + (1-\varphi)e^{(3-(1+\varphi))\psi_T t} - e^{2\psi_T t} > 0.$$

Therefore $\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = 1$, or the derivative is positive for $t > 0$ and $1 < \kappa < 2$, which means the variance grows relative to the mean over time.

Next, we turn to $0 < \kappa < 1$, where

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[(\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}\right].$$

For any $t > 0$, for $\kappa = 1$

$$\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = \operatorname{sign}\left[-e^{2\psi_T t} - 0 + e^{2\psi_T t}\right] = 0.$$

Now letting $g(\kappa) = (\kappa-2)e^{(\kappa+1)\psi_T t} - (\kappa-1)e^{\kappa\psi_T t} + e^{2\psi_T t}$, we have from above

$$g'(\kappa) = \{1 + (\kappa-2)\psi_T t\}e^{(\kappa+1)\psi_T t} - \{1 + (\kappa-1)\psi_T t\}e^{\kappa\psi_T t}.$$

Letting $h(x) = \{1 + (\kappa-1-x)\psi_T t\}e^{(\kappa+x)\psi_T t}$, we see

$$g'(\kappa) = h(1) - h(0).$$

But we see from above that

$$h'(x) = \psi_T t e^{(\kappa+x)\psi_T t}(\kappa-1-x)\psi_T t.$$

If $0 < \kappa < 1$ and $t > 0$, then $h'(x) < 0$ for $0 \le x \le 1$. Since $h(x)$ is a continuous function, we must have

$$h(1) < h(0).$$

This implies that for $0 < \kappa < 1$ and $t > 0$,

$$g'(\kappa) < 0.$$

Since $g(\kappa)$ is continuous, and since $g(1) = 0$, we now know that for $0 < \kappa < 1$ and $t > 0$, $g(\kappa) > 0$. Therefore $\operatorname{sign}\left[\frac{d\rho_t}{dt}\right] = 1$, or the derivative is positive for $t > 0$ and $0 < \kappa < 1$, which means the variance grows relative to the mean over time.

Now consider the special case of $\kappa = 2$

$$\rho_t = \frac{\frac{2\gamma_D D_0 R_0\left(e^{\psi_T t} - e^{2\psi_T t}\right)}{2\psi_T} + 2\gamma_D D_0 R_0 e^{2\psi_T t} t}{\frac{\gamma_D D_0 R_0 (e^{2\psi_T t} - e^{\psi_T t})}{2\psi_T - \psi_T}},$$

$$\rho_t = \frac{\left(e^{\psi_T t} - e^{2\psi_T t}\right) + 2\psi_T t e^{2\psi_T t}}{(e^{2\psi_T t} - e^{\psi_T t})},$$

$$\rho_t = \frac{(2\psi_T t - 1)e^{2\psi_T t} + e^{\psi_T t}}{(e^{2\psi_T t} - e^{\psi_T t})}.$$

Taking the derivative with respect to time yields

$$\frac{d\rho_t}{dt} = \frac{\left\{\begin{matrix}\left(2\psi_T e^{2\psi_T t} + 2\psi_T(2\psi_T t - 1)e^{2\psi_T t} + \psi_T e^{\psi_T t}\right)\left(e^{2\psi_T t} - e^{\psi_T t}\right) \\ -\left((2\psi_T t - 1)e^{2\psi_T t} + e^{\psi_T t}\right)\left(2\psi_T e^{2\psi_T t} - \psi_T e^{\psi_T t}\right)\end{matrix}\right\}}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = \psi_T \frac{\left(4\psi_T t e^{2\psi_T t} + e^{\psi_T t}\right)\left(e^{2\psi_T t} - e^{\psi_T t}\right) - \left((2\psi_T t - 1)e^{2\psi_T t} + e^{\psi_T t}\right)\left(2e^{2\psi_T t} - e^{\psi_T t}\right)}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = \psi_T \frac{\left\{\begin{matrix}4\psi_T t e^{4\psi_T t} + e^{3\psi_T t} - 4\psi_T t e^{3\psi_T t} - e^{2\psi_T t} \\ -2(2\psi_T t - 1)e^{4\psi_T t} - 2e^{3\psi_T t} + (2\psi_T t - 1)e^{3\psi_T t} + e^{2\psi_T t}\end{matrix}\right\}}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = \psi_T \frac{\{4\psi_T t - 2(2\psi_T t - 1)\}e^{4\psi_T t} - \{4\psi_T t + 1 - (2\psi_T t - 1)\}e^{3\psi_T t}}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = \psi_T \frac{2e^{4\psi_T t} - \{2\psi_T t + 2\}e^{3\psi_T t}}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = 2\psi_T e^{3\psi_T t} \frac{e^{\psi_T t} - \psi_T t - 1}{(e^{2\psi_T t} - e^{\psi_T t})^2}.$$

Using the Taylor expansion

$$\frac{d\rho_t}{dt} = 2\psi_T e^{3\psi_T t} \frac{\left(1 + \psi_T t + \frac{(\psi_T t)^2}{2!} + \frac{(\psi_T t)^3}{3!} + \frac{(\psi_T t)^4}{4!} \ldots\right) - \psi_T t - 1}{(e^{2\psi_T t} - e^{\psi_T t})^2},$$

$$\frac{d\rho_t}{dt} = 2\psi_T e^{3\psi_T t} \frac{\left(\frac{(\psi_T t)^2}{2!} + \frac{(\psi_T t)^3}{3!} + \frac{(\psi_T t)^4}{4!} \ldots\right)}{(e^{2\psi_T t} - e^{\psi_T t})^2}.$$

Therefore, $\frac{d\rho_t}{dt} > 0$ for $t > 0$ and $\kappa = 2$, which means the variance grows relative to the mean over time.

Finally, consider the special case of $\kappa = 1$

$$\rho_t = \frac{\frac{2\gamma_D D_0 R_0\left(e^{2\psi_T t} - e^{\psi_T t}\right)}{\psi_T} - \gamma_D D_0 R_0 e^{\psi_T t} t}{\gamma_D D_0 R_0 e^{\psi_T t} t},$$

$$\rho_t = \frac{2e^{2\psi_T t} - (2 + \psi_T t)e^{\psi_T t}}{\psi_T t e^{\psi_T t}},$$

$$\rho_t = \frac{2e^{\psi_T t} - (2 + \psi_T t)}{\psi_T t}.$$

Using the Taylor expansion

$$\rho_t = \frac{2\left(1 + \psi_T t + \frac{(\psi_T t)^2}{2!} + \frac{(\psi_T t)^3}{3!} + \frac{(\psi_T t)^4}{4!} \ldots\right) - (2 + \psi_T t)}{\psi_T t},$$

$$\rho_t = \frac{\psi_T t + 2\left(\frac{(\psi_T t)^2}{2!} + \frac{(\psi_T t)^3}{3!} + \frac{(\psi_T t)^4}{4!} \ldots\right)}{\psi_T t},$$

$$\rho_t = 1 + 2\left(\frac{\psi_T t}{2!} + \frac{(\psi_T t)^2}{3!} + \frac{(\psi_T t)^3}{4!} \ldots\right).$$

Once again, the variance grows relative to the mean over time.

Overall, we have shown that for all $\kappa > 0$ and $t > 0$, the ratio of the transconjugant variance to the mean number of transconjugants is amplified over time.

Appendix VII : Derivations for estimate variance

In this appendix, we provide details for the derivation of the variance expressions for the LDM, SIM, and ASM estimates. Starting with the ASM estimate:

$$\gamma_D = \frac{\psi_D + \psi_R - \psi_T}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)} T_{\tilde{t}},$$

which we can think about as a random variable Γ_{ASM}. Specifically,

$$\Gamma_{\mathrm{ASM}} = c_1 T_{\tilde{t}},$$

where the constant c_1 is

$$c_1 = \frac{\psi_D + \psi_R - \psi_T}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)}.$$

The variance of the ASM estimate due to transconjugant variance is then

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = {c_1}^2\{\mathrm{var}(T_{\tilde{t}})\}.$$

If $\psi_T \notin \{\psi_D + \psi_R, (\psi_D + \psi_R)/2\}$, we have:

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = \left(\frac{\psi_D+\psi_R-\psi_T}{D_0R_0\left(e^{(\psi_D+\psi_R)\tilde{t}}-e^{\psi_T\tilde{t}}\right)}\right)^2 \gamma_D D_0 R_0 \left\{\frac{(\psi_D+\psi_R)e^{(\psi_D+\psi_R)\tilde{t}}}{(\psi_D+\psi_R-\psi_T)(\psi_D+\psi_R-2\psi_T)} + \frac{e^{\psi_T\tilde{t}}}{\psi_D+\psi_R-\psi_T} - \frac{2e^{2\psi_T\tilde{t}}}{\psi_D+\psi_R-2\psi_T}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = \frac{\gamma_D}{D_0R_0}\left(\frac{\psi_D+\psi_R-\psi_T}{e^{(\psi_D+\psi_R)\tilde{t}}-e^{\psi_T\tilde{t}}}\right)^2 \left\{\frac{(\psi_D+\psi_R)e^{(\psi_D+\psi_R)\tilde{t}} + (\psi_D+\psi_R-2\psi_T)e^{\psi_T\tilde{t}} - (\psi_D+\psi_R-\psi_T)2e^{2\psi_T\tilde{t}}}{(\psi_D+\psi_R-\psi_T)(\psi_D+\psi_R-2\psi_T)}\right\}$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = \frac{\gamma_D(\psi_D+\psi_R-\psi_T)}{D_0R_0}\left\{\frac{(\psi_D+\psi_R)e^{(\psi_D+\psi_R)\tilde{t}}+(\psi_D+\psi_R-2\psi_T)e^{\psi_T\tilde{t}}-(\psi_D+\psi_R-\psi_T)2e^{2\psi_T\tilde{t}}}{(\psi_D+\psi_R-2\psi_T)\left(e^{(\psi_D+\psi_R)\tilde{t}}-e^{\psi_T\tilde{t}}\right)^2}\right\}.$$

The LDM estimate is expressed as follows:

$$\gamma_D = -\ln \hat{p}_0(\tilde{t})\left(\frac{\psi_D + \psi_R}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right).$$

We measure the number of populations that have no transconjugants (w) out of the total number of populations tracked (W). The maximum likelihood estimate of $p_0(\tilde{t})$ is

$$\hat{p}_0(\tilde{t}) = \frac{w}{\mathrm{W}}.$$

Across experiments, there will be variance in the number of populations with no transconjugants. We define the random variable F to represent the fraction of total populations that have no transconjugants. The expectation of F is:

$$\mathrm{E}[F] = \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} \left(\frac{w}{\mathrm{W}}\right),$$

$$\mathrm{E}[F] = \frac{1}{\mathrm{W}} \sum_{w=1}^{\mathrm{W}} \frac{\mathrm{W}!}{w!\,(\mathrm{W}-w)!} (p_0)^w (1-p_0)^{\mathrm{W}-w}\, w,$$

$$\mathrm{E}[F] = \frac{1}{\mathrm{W}} \sum_{w=1}^{\mathrm{W}} \frac{\mathrm{W}!}{(\mathrm{w}-1)!\,(\mathrm{W}-w)!} (p_0)^w (1-p_0)^{\mathrm{W}-w},$$

$$\mathrm{E}[F] = p_0 \sum_{w=1}^{\mathrm{W}} \frac{(\mathrm{W}-1)!}{(\mathrm{w}-1)!\left((\mathrm{W}-1)-(\mathrm{w}-1)\right)!} (p_0)^{w-1} (1-p_0)^{(\mathrm{W}-1)-(w-1)}.$$

Letting $i = w - 1$, we have

$$\mathrm{E}[F] = p_0 \sum_{i=0}^{\mathrm{W}-1} \frac{(\mathrm{W}-1)!}{i!\,((\mathrm{W}-1)-i)!} (p_0)^i (1-p_0)^{(\mathrm{W}-1)-i}.$$

However, because $\sum_{i=0}^{\mathrm{W}-1} \binom{\mathrm{W}-1}{i} (p_0)^i (1-p_0)^{(\mathrm{W}-1)-i} = 1$, we have

$$\mathrm{E}[F] = p_0,$$

which makes sense.

The variance of F is

$$\mathrm{var}[F] = \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} \left(\frac{w}{\mathrm{W}} - p_0\right)^2,$$

$$\mathrm{var}[F] = \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} \left(\left(\frac{w}{\mathrm{W}}\right)^2 - 2p_0\left(\frac{w}{\mathrm{W}}\right) + (p_0)^2\right),$$

$$\mathrm{var}[F] = \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} \left(\frac{w}{\mathrm{W}}\right)^2 - 2p_0 \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} \left(\frac{w}{\mathrm{W}}\right) + (p_0)^2 \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w},$$

$$\mathrm{var}[F] = \left\{\left(\frac{1}{\mathrm{W}^2}\right) \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} w^2\right\} - 2p_0\mathrm{E}[F] + (p_0)^2,$$

$$\mathrm{var}[F] = \left\{\left(\frac{1}{\mathrm{W}^2}\right) \sum_{w=0}^{\mathrm{W}} \binom{\mathrm{W}}{w} (p_0)^w (1-p_0)^{\mathrm{W}-w} w(w-1+1)\right\} - 2(p_0)^2 + (p_0)^2,$$

$$\mathrm{var}[F] = \left\{\left(\frac{1}{\mathrm{W}^2}\right)\sum_{w=2}^{\mathrm{W}}\frac{\mathrm{W}!}{w!\,(\mathrm{W}-w)!}(p_0)^w(1-p_0)^{\mathrm{W}-w}w(w-1) + \left(\frac{1}{\mathrm{W}^2}\right)\sum_{w=0}^{\mathrm{W}}\binom{\mathrm{W}}{w}(p_0)^w(1-p_0)^{\mathrm{W}-w}w\right\} - (p_0)^2,$$

$$\mathrm{var}[F] = \left\{\left(\frac{1}{\mathrm{W}^2}\right)\sum_{w=2}^{\mathrm{W}}\frac{\mathrm{W}!}{(w-2)!\,(\mathrm{W}-w)!}(p_0)^w(1-p_0)^{\mathrm{W}-w} + \left(\frac{1}{\mathrm{W}}\right)\sum_{w=0}^{\mathrm{W}}\binom{\mathrm{W}}{w}(p_0)^w(1-p_0)^{\mathrm{W}-w}\left(\frac{w}{\mathrm{W}}\right)\right\} - (p_0)^2,$$

$$\mathrm{var}[F] = \left(\frac{\mathrm{W}(\mathrm{W}-1)(p_0)^2}{\mathrm{W}^2}\right)\left\{\sum_{w=2}^{\mathrm{W}}\frac{(\mathrm{W}-2)!}{(w-2)!\,((\mathrm{W}-2)-(w-2))!}(p_0)^{w-2}(1-p_0)^{(\mathrm{W}-2)-(w-2)}\right\} + \frac{p_0}{\mathrm{W}} - (p_0)^2.$$

Letting $j = w - 2$, we have

$$\mathrm{var}[F] = \left(\frac{(\mathrm{W}-1)(p_0)^2}{\mathrm{W}}\right)\left\{\sum_{j=0}^{\mathrm{W}-2}\binom{\mathrm{W}-2}{j}(p_0)^j(1-p_0)^{(\mathrm{W}-2)-j}\right\} + \frac{p_0}{\mathrm{W}} - (p_0)^2,$$

$$\mathrm{var}[F] = \left(\frac{(\mathrm{W}-1)(p_0)^2}{\mathrm{W}}\right) + \frac{p_0}{\mathrm{W}} - (p_0)^2,$$

$$\mathrm{var}[F] = \frac{(\mathrm{W}-1)(p_0)^2 + p_0 - \mathrm{W}(p_0)^2}{\mathrm{W}},$$

$$\mathrm{var}[F] = \frac{-(p_0)^2 + p_0}{\mathrm{W}},$$

$$\mathrm{var}[F] = \frac{p_0(1-p_0)}{\mathrm{W}}.$$

We represent the LDM estimate as a random variable Γ_{LDM},

$$\Gamma_{\mathrm{LDM}} = c_2 \ln F,$$

where the constant c_2 is

$$c_2 = -\left(\frac{\psi_D + \psi_R}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right).$$

The variance of the LDM estimate (due to the variance in transconjugant presence) is

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) = {c_2}^2\{\mathrm{var}(\ln F)\}.$$

Using a first-order Taylor series approximation for $\ln F$ centered at $\mathrm{E}[F]$

$$\ln F \approx \ln(\mathrm{E}[F]) + \frac{1}{\mathrm{E}[F]}(F - \mathrm{E}[F]),$$

$$\ln F \approx \frac{F}{\mathrm{E}[F]} + \ln(\mathrm{E}[F]) - 1.$$

Because $\ln(\mathrm{E}[F]) - 1$ is constant, we have

$$\mathrm{var}[\ln F] \approx \mathrm{var}\left\{\frac{F}{\mathrm{E}[F]}\right\},$$

$$\mathrm{var}[\ln F] \approx \left(\frac{1}{\mathrm{E}[F]}\right)^2 \mathrm{var}[F],$$

$$\mathrm{var}[\ln F] \approx \left(\frac{1}{p_0}\right)^2 \left(\frac{p_0(1-p_0)}{\mathrm{W}}\right),$$

$$\mathrm{var}[\ln F] \approx \frac{1}{\mathrm{W}}\left(\frac{1}{p_0} - 1\right).$$

The following is the expression for $p_0(\tilde{t})$:

$$p_0(\tilde{t}) = \exp\left\{\frac{-\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\}.$$

Reintroducing the time argument in F, and substituting the expression for $p_0(\tilde{t})$ yields

$$\mathrm{var}[\ln F_{\tilde{t}}] \approx \frac{1}{\mathrm{W}}\left(\exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\} - 1\right).$$

And therefore the variance for the LDM estimate is

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{1}{\mathrm{W}}\left(\frac{\psi_D + \psi_R}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}\right)^2 \left(\exp\left\{\frac{\gamma_D D_0 R_0}{\psi_D + \psi_R}\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)\right\} - 1\right).$$

This can be written more compactly as

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{{\xi_{\tilde{t}}}^2}{\mathrm{W}}\left(e^{\left(\frac{\gamma_D}{\xi_{\tilde{t}}}\right)} - 1\right),$$

with

$$\xi_{\tilde{t}} = \frac{\psi_D + \psi_R}{D_0 R_0\left(e^{(\psi_D+\psi_R)\tilde{t}} - 1\right)}.$$

To show mathematically that the LDM estimate is more precise for short times, we will approximate the expressions for the variance when $\tilde{t}$ is very small. This enables us to use a first-order Maclaurin approximation $e^{c\tilde{t}} \approx 1 + c\tilde{t}$. This allows the following approximation of the variance for the ASM estimate:

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D(\psi_D + \psi_R - \psi_T)}{D_0 R_0}\left\{\frac{(\psi_D + \psi_R)(1 + (\psi_D + \psi_R)\tilde{t}) + (\psi_D + \psi_R - 2\psi_T)(1 + \psi_T\tilde{t}) - 2(\psi_D + \psi_R - \psi_T)(1 + 2\psi_T\tilde{t})}{(\psi_D + \psi_R - 2\psi_T)((\psi_D + \psi_R)\tilde{t} - \psi_T\tilde{t})^2}\right\}.$$

We can then simplify this expression:

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D(\psi_D + \psi_R - \psi_T)}{D_0 R_0}\left\{\frac{(\psi_D + \psi_R)(\psi_D + \psi_R)\tilde{t} + (\psi_D + \psi_R - 2\psi_T)\psi_T\tilde{t} - 2(\psi_D + \psi_R - \psi_T)2\psi_T\tilde{t}}{(\psi_D + \psi_R - 2\psi_T)(\psi_D + \psi_R - \psi_T)^2\tilde{t}^2}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D + \psi_R)(\psi_D + \psi_R) + (\psi_D + \psi_R - 2\psi_T)\psi_T - 2(\psi_D + \psi_R - \psi_T)2\psi_T}{(\psi_D + \psi_R - 2\psi_T)(\psi_D + \psi_R - \psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R)+2\psi_T(\psi_D+\psi_R)+(\psi_D+\psi_R-2\psi_T)\psi_T-2(\psi_D+\psi_R-\psi_T)2\psi_T}{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R)+2\psi_T\psi_D+2\psi_T\psi_R+\psi_T\psi_D+\psi_T\psi_R-2{\psi_T}^2-4\psi_T\psi_D-4\psi_T\psi_R+4{\psi_T}^2}{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R)-\psi_T\psi_D-\psi_T\psi_R+2{\psi_T}^2}{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R)-(\psi_D+\psi_R-2\psi_T)\psi_T}{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}\left\{\frac{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}{(\psi_D+\psi_R-2\psi_T)(\psi_D+\psi_R-\psi_T)}\right\},$$

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) \approx \frac{\gamma_D}{D_0 R_0 \tilde{t}}.$$

We now turn to the variance for the LDM (which we note is already an approximation). For very small $\tilde{t}$,

$$\xi_{\tilde{t}} \approx \frac{\psi_D+\psi_R}{D_0 R_0(1+(\psi_D+\psi_R)\tilde{t}-1)},$$

$$\xi_{\tilde{t}} \approx \frac{1}{D_0 R_0 \tilde{t}}.$$

Thus, we have

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{\left(\frac{1}{D_0 R_0 \tilde{t}}\right)^2}{\mathrm{W}}\left(e^{(\gamma_D D_0 R_0 \tilde{t})}-1\right).$$

Using the Maclaurin approximation again yields

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{\left(\frac{1}{D_0 R_0 \tilde{t}}\right)^2}{\mathrm{W}}(1+\gamma_D D_0 R_0 \tilde{t}-1),$$

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) \approx \frac{\gamma_D}{\mathrm{W} D_0 R_0 \tilde{t}}.$$

Our LDM assay requires $\mathrm{W} > 1$. Therefore when $\tilde{t}$ is very small

$$\mathrm{var}(\Gamma_{\mathrm{LDM}}) < \mathrm{var}(\Gamma_{\mathrm{ASM}}).$$

Again, we note the caveat that our estimate for the variance of the LDM estimate was already an approximation.

One can derive the variance for the SIM estimate in a way analogous to the ASM estimate, with the caveat that an approximation is needed (namely one that is similar to the approximation used for the variance of the LDM estimate). We will assume that the SIM estimate is obtained during exponential growth of all populations (which are assumed to grow at the same rate), which will allow us to connect the variance for the SIM to the variance for the ASM. We provide some of the details here.

If we focus solely on the contribution of transconjugant variation to estimate variance, we can represent the SIM estimate as a random variable Γ_{SIM}:

$$\Gamma_{\mathrm{SIM}} = \mathrm{A}\ln(1 + \mathrm{B}T_{\tilde{t}})$$

where the coefficients are treated as the following constants:

$$\mathrm{A} = \frac{\psi}{N_0(e^{\psi\tilde{t}} - 1)},$$

$$\mathrm{B} = \frac{N_0}{D_0 R_0 e^{\psi\tilde{t}}}.$$

The variance of the estimate is then

$$\mathrm{var}(\Gamma_{\mathrm{SIM}}) = \mathrm{A}^2\mathrm{var}[\ln(1 + \mathrm{B}T_{\tilde{t}})]$$

Using the first-order Taylor expansion centered at $E[T_{\tilde{t}}]$:

$$\ln(1 + \mathrm{B}T_{\tilde{t}}) \approx \ln(1 + \mathrm{B}E[T_{\tilde{t}}]) + \frac{B}{1 + \mathrm{B}E[T_{\tilde{t}}]}(T_{\tilde{t}} - E[T_{\tilde{t}}]).$$

And we have

$$\mathrm{var}[\ln(1 + \mathrm{B}T_{\tilde{t}})] \approx \left(\frac{\mathrm{B}}{1 + \mathrm{B}E[T_{\tilde{t}}]}\right)^2 \mathrm{var}(T_{\tilde{t}}).$$

Thus, we have

$$\mathrm{var}(\Gamma_{\mathrm{SIM}}) \approx \left(\frac{\mathrm{AB}}{1 + \mathrm{B}E[T_{\tilde{t}}]}\right)^2 \mathrm{var}(T_{\tilde{t}})$$

The quantity $E[T_{\tilde{t}}]$ is given in SI Section 7 (here we assume $\psi_D = \psi_R = \psi_T = \psi$). Plugging in the expressions for A, B, and $E[T_{\tilde{t}}]$ and simplifying yields,

$$\mathrm{var}(\Gamma_{\mathrm{SIM}}) \approx \frac{1}{\left(1 + \frac{\gamma_D(N_{\tilde{t}} - N_0)}{\psi}\right)^2}\left(\frac{\psi}{D_0 R_0(e^{2\psi\tilde{t}} - e^{\psi\tilde{t}})}\right)^2 \mathrm{var}(T_{\tilde{t}})$$

The expression for variance of the ASM estimate (derived in SI Section 8) is:

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = \left(\frac{\psi_D + \psi_R - \psi_T}{D_0 R_0 \left(e^{(\psi_D+\psi_R)\tilde{t}} - e^{\psi_T \tilde{t}}\right)}\right)^2 \mathrm{var}(T_{\tilde{t}})$$

If $\psi_D = \psi_R = \psi_T = \psi$ (as is assumed for the SIM estimate) we have

$$\mathrm{var}(\Gamma_{\mathrm{ASM}}) = \left(\frac{\psi}{D_0 R_0 \left(e^{2\psi\tilde{t}} - e^{\psi\tilde{t}}\right)}\right)^2 \mathrm{var}(T_{\tilde{t}})$$

Therefore,

$$\mathrm{var}(\Gamma_{\mathrm{SIM}}) \approx \frac{1}{\left(1 + \frac{\gamma_D (N_{\tilde{t}} - N_0)}{\psi}\right)^2} \mathrm{var}(\Gamma_{\mathrm{ASM}})$$

At the time of the end of the assay ($t = \tilde{t}$), the product of donors and recipients ($D_{\tilde{t}} R_{\tilde{t}}$) is in the vicinity of the reciprocal of the conjugation rate ($1/\gamma_D$), but the sum of donors and recipients is much smaller than the reciprocal of the conjugation rate ($D_{\tilde{t}} + R_{\tilde{t}} \ll 1/\gamma_D$). We note $N_t \approx D_t + R_t$. Therefore, for reasonable times in which the assay is ended and a reasonable growth rate:

$$N_{\tilde{t}} - N_0 \ll \frac{\psi}{\gamma_D}$$

In such a case, $1 + \frac{\gamma_D (N_{\tilde{t}} - N_0)}{\psi} \approx 1$, and

$$\mathrm{var}(\Gamma_{\mathrm{SIM}}) \approx \mathrm{var}(\Gamma_{\mathrm{ASM}})$$

Indeed, for an example close to that explored in SI Figure 10 (with $\psi_D = \psi_R = \psi_T = 1$, $D_0 = R_0 = 10^4$, and $\gamma_D = 10^{-12}$), the variances for the SIM and ASM estimates are virtually indistinguishable.

CHAPTER 2

Evolutionary crowdsourcing: interchangeable adaptive progress across bacterial species occurs due to underlying fitness landscape alignment of a horizontally transferred gene

Olivia Kosterlitz, Nathan Grassi, Bailey Werner, Ryan Seamus McGee, Eva M. Top, Benjamin Kerr

Genes that undergo horizontal gene transfer (HGT) evolve in dramatically different genomic backgrounds as they move between hosts, which is in stark contrast to genes that evolve under strict vertical inheritance. Given the ubiquity of HGT in microbial communities, it is notable that the effects of host-switching on gene evolution have been largely understudied. Here, we present a novel framework to examine the consequences of host switching on gene evolution depending on the existence and form of host-dependent mutational effects. We started exploring the effects of HGT on gene evolution by focusing on a well-known antibiotic resistance gene (encoding a beta-lactamase) commonly encoded on conjugative plasmids found in Enterobacteriaceae pathogens. By reconstructing the resistance landscape for a small set of mutationally connected alleles in three species (*Escherichia coli, Salmonella enterica, and Klebsiella pneumoniae*), we uncovered that the landscape topography was overwhelmingly aligned with very low levels of host-dependent mutational effects. By simulating gene evolution with and without HGT using the species-specific empirical landscapes, we found that evolutionary outcomes were similar despite HGT. These findings suggest that mobile genes adapting in one species can lead to adaptation in another species. In such a case, vehicles of cross-species HGT enable a distributed form of genetic evolution across a bacterial community, where species can ‘crowdsource’ adaptation from other community members.

Introduction

In bacterial communities, genes that undergo horizontal transfer can evolve in dramatically different genomic backgrounds as they move between hosts (5). This stands in stark contrast to genes that evolve under strict vertical inheritance, where the genomic backdrop is relatively constant over time. Despite the ubiquity and importance of horizontal gene transfer (HGT), the effect of host-switching on the evolution of genes that undergo HGT (hereafter 'mobile genes') has received little attention.

A mobile gene that is transferred from one host species to another may exhibit changes in the fitness effects of mutations due to the new host's genomic context. If beneficial mutations in one host are also beneficial in other hosts, then one species may effectively "crowdsource" adaptative evolution of the mobile gene. That is, a focal species that transfers a mobile gene to a different species and later reacquires that gene can benefit from adaptive genetic changes in the mobile gene that occurred in the second host. In contrast, if beneficial mutations in one host are detrimental in other hosts, then the crowdsourcing value from HGT decreases. In this case, greater adaptive progress is expected to occur "in house." To determine the adaptive consequences of HGT, we need to understand how the fitness effects of mutations change in sign (beneficial or deleterious) or magnitude depending on the host housing the mobile gene. We term this dependence *host epistasis*, a subcategory of genetic epistasis in which the fitness effects of mutations depend on the whole host genomic background rather than a subset of variable sites (44). If the host genome is likened to an environmental context for the evolving mobile gene, host epistasis could be envisioned as a kind of environmental epistasis (i.e., GxE interaction) (45). Both genetic and environmental epistasis have been shown to profoundly shape adaptive evolution (46, 47). Here, we sought to determine the presence and form of host epistasis and its effects on the evolution of a mobile gene.

A visual metaphor that can aid in understanding the evolutionary effects of host epistasis is a 'fitness landscape' that maps a network of mutationally connected genotypes to fitness. Host epistasis manifests as differences in the landscape topography across hosts. To illustrate different forms of host epistasis and their evolutionary consequences, we focus on a simple example involving two host species and three variant sites in a mobile gene. In Figure 7a, the landscapes of the blue and red hosts are aligned with no instances of sign host epistasis (i.e., the signs of the fitness effects of mutations are the same in the two hosts). In this scenario, HGT between the hosts does not impact the evolutionary end point reached in the red host relative to adaptation without HGT (Figure 7b). These conditions enable evolutionary crowdsourcing where the red host can make use of the transient adaptation in the blue host. In Figure 7c, the landscapes of the two hosts are mirror images of one another, with rampant sign host epistasis. Here, adaptation in the blue host is counterproductive to evolutionary progress in the red host (Figure 7d). This scenario highlights evolutionary "insourcing" where the red host can make more progress by adaptation occurring in house without HGT. A more subtle case is found in Figure 7e where a handful of mutations exhibit sign host epistasis. Specifically, there is a suboptimal fitness peak in the red host landscape that is absent in the blue host

landscape. In this case, adaptive evolution in the blue host explores additional regions of genotype space, and HGT from the blue to red host introduces genetic variation that effectively releases the red host from a suboptimal evolutionary endpoint. This scenario highlights evolutionary “outsourcing” where HGT can qualitatively alter the evolutionary trajectory in the red host relative to adaptation without HGT (Figure 7f). These simple cases illustrate that comparing landscape topographies across hosts is the first step in determining how cross-species HGT may affect mobile gene evolution.

In a previous empirical study, sign host epistasis was uncovered in an antibiotic resistance landscape (48, 49). However, this prior work focused on a small set of alleles in an essential gene primarily encoded in the bacterial chromosome. Surprisingly there has been, to our knowledge, no attention given to *mobile* genes where host epistasis may be most relevant due to HGT facilitating host-switching. Here, we experimentally constructed a portion of a mobile gene’s landscape in three Enterobacteriaceae pathogens. This gene naturally resides on conjugative plasmids in enteric bacteria and encodes a beta-lactamase in which a handful of mutations are known to increase resistance orders of magnitude to various beta-lactam antibiotics (44). Our study aims to reveal the presence and nature of host epistasis and its consequences on the evolution of drug resistance.

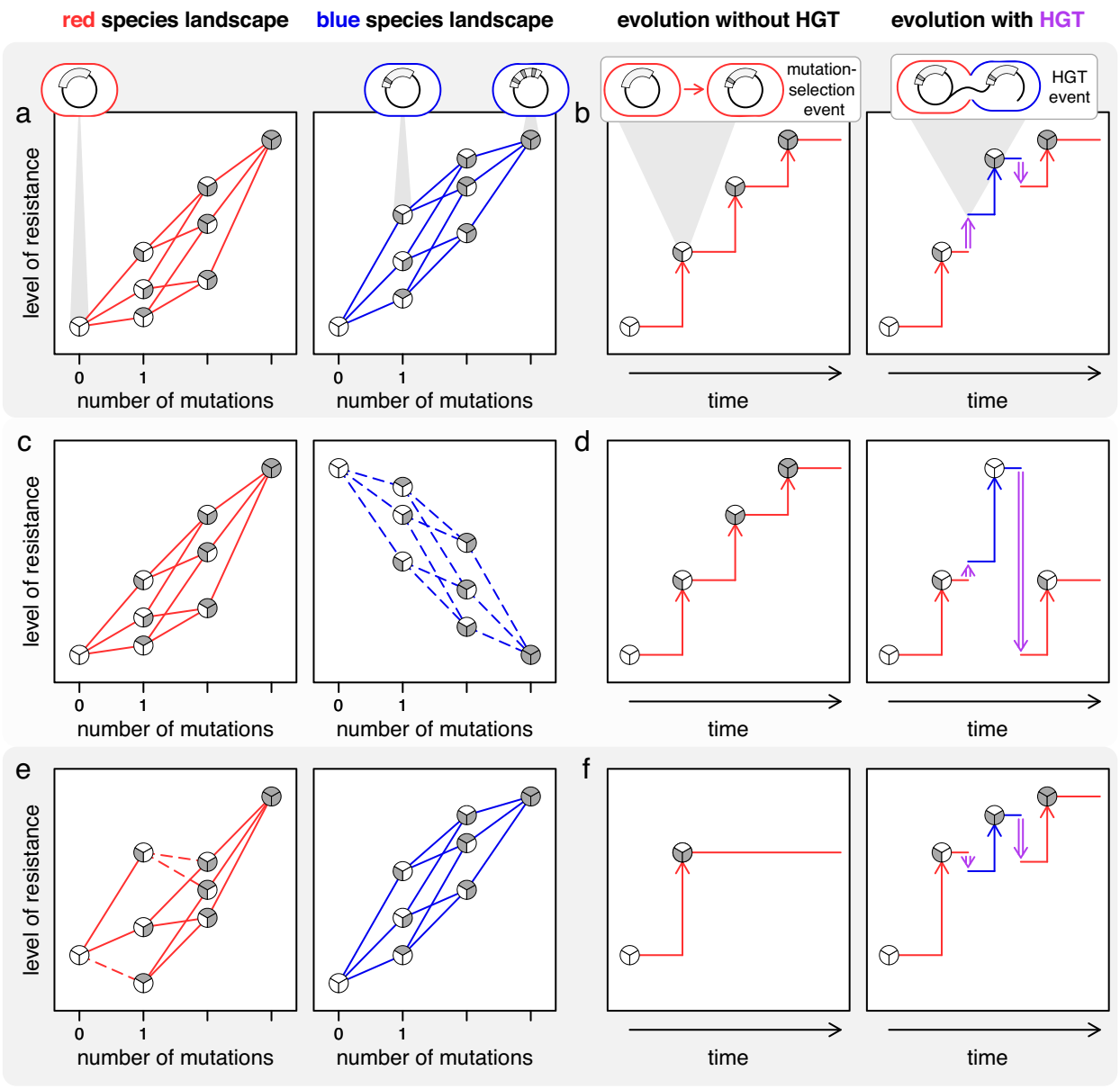


Figure 7: Effect of HGT on mobile gene evolution with hypothetical host-specific landscapes. Here, we consider a simple bi-allelic three-site landscape in two hosts (differentiated by the red and blue colors) for a gene encoded on a mobile genetic element. (a) The adaptive landscape can be visualized by plotting the genotypes level of resistance (taken to be a proxy for fitness) as a function of the number of mutations on a WT background. Each of the $2^3 = 8$ genotypes is represented by a circle dividing into 'wedges' equal to the number of sites (3 in this case) where the evolved variant at a site is shown by shading the wedge (grey shading). The edges (lines where the color matches the host) connect genotypes differing by a single mutation. Here, the landscapes of the red and blue host are well aligned (note there are no mutations having fitness effects of opposite sign). (b) If we assume that selection is strong and mutation is weak, we can represent the fixation of each beneficial mutation (vertical arrows) as a step up in the level of drug resistance. In this represented evolutionary trajectory after three mutational events (right panel), the population reaches the adaptive peak, a genotype from which all

mutations are detrimental. When HGT (vertical purple double-ended arrow) to and from the blue host occurs proceeding and following the second mutational event, the population still reaches the adaptive peak given that the landscapes (part a) of the blue and red host are aligned. This scenario is an example of evolutionary crowdsourcing where the red host can make use of the transient adaptation in the blue host. (c) Here, we show a different example where the landscapes of the two hosts possess rampant sign host epistasis. Thus, mutational steps are beneficial (solid lines) in the red host but are deleterious (dashed lines) in the blue host. (d) This scenario is an example of evolutionary insourcing where transient adaptation in the blue host is counterproductive to the evolutionary progress in the red host. (e) In this last example, the landscapes of the two hosts have only a handful of mutational steps exhibiting sign host epistasis; however, given the location of these mutations, a suboptimal fitness peak occurs in the red host landscape that is absent in the blue host landscape. (f) Evolution in the red host can result in a suboptimal evolutionary endpoint (left panel). However, this scenario highlights evolutionary outsourcing where adaptation in the blue host can effectively release the red host from the suboptimal endpoint.

Experimental Approach

We assessed the host-specific landscape topography of a set of plasmid-borne antibiotic resistance alleles using a high-throughput multiplexed assay. Each allele was mapped to the level of resistance it conferred in the focal host, which is a proxy for the fitness. Together, these mappings formed a set of mutationally connected alleles comprising the “resistance” landscape. Our basic approach to assess the resistance level of each allele in each host consisted of three steps. First, each allele in the set was engineered, tagged with unique barcodes, and transformed into a given host (Figure 8a). Second, all transformants were pooled to create the initial host library and subsequently inoculated and incubated in a series of tubes with an increasing concentration of the relevant antibiotic (Figure 8b). Third, the pre- and post-selection cell count estimates of each allele were used to approximate the growth rate of each allele within each antibiotic concentration (Figure 8c). The estimated growth rates of each allele across the gradient yielded a dose response curve by fitting a log-logistic function. The inflection point of this curve, namely where growth rate is dropping most precipitously, was our measure of its level of resistance. Collectively, the resistance levels for the set of alleles determined the topography of the landscape for the given host (Figure 8d). By implementing this procedure in multiple bacterial species, we compared landscapes across different hosts

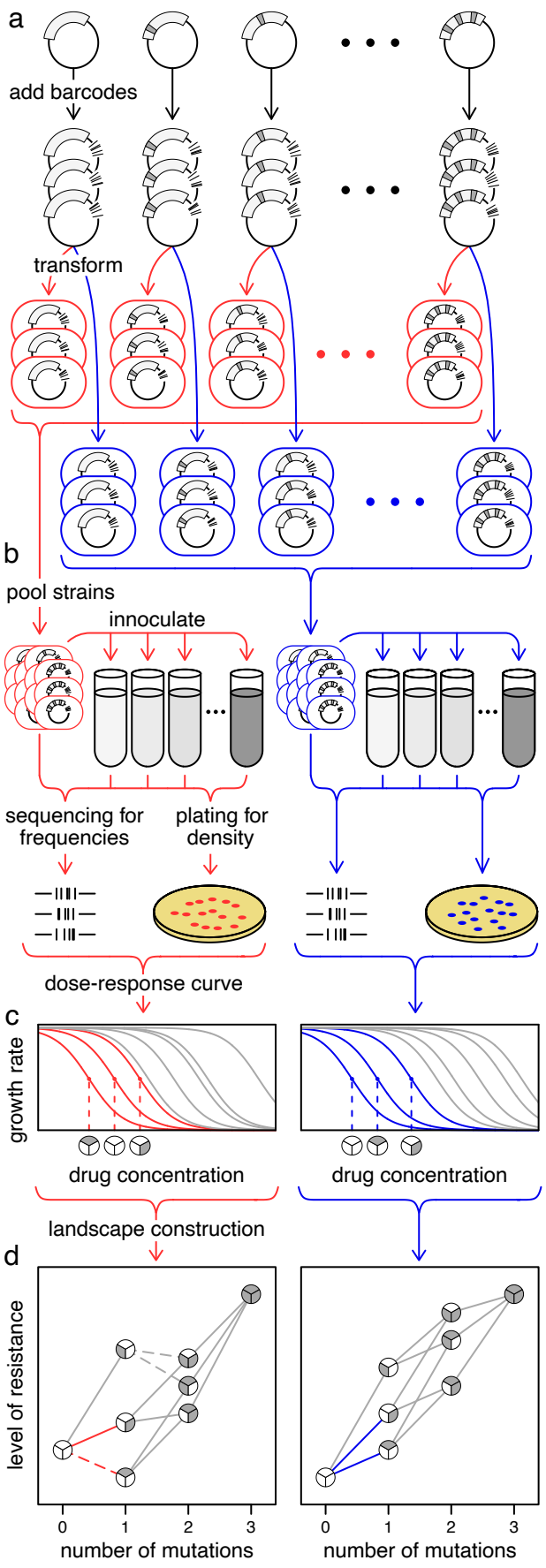


Figure 8: A multiplexed protocol for constructing host-specific landscapes. (a) To construct each allele of interest, mutations (dark grey notches) are introduced into a focal gene (rectangular arc) on a plasmid. To facilitate allele tracking in the experiment, each allele is tagged with three unique barcodes (black notches). Subsequently, the barcoded alleles are transformed into each host species ('red' and 'blue' hosts are shown here). (b) To assess the resistance level of each allele, all transformants within a species are pooled to create the initial bacterial library and inoculated into an antibiotic gradient (the intensity of grey-shaded medium increases with antibiotic concentration). Samples are acquired before and after incubation to determine barcode frequency using deep sequencing and the total population density using dilution plating. (c) Using the product of total population density and barcode frequencies associated with each allele before and after selection at a given concentration, a growth rate can be calculated (allele-concentration specific). For each allele, the estimated growth rates across the antibiotic gradient yields a dose response curve by fitting a log-logistic function where the level of resistance is given by the inflection point of the curve (three alleles are highlighted per host where the resistance of each is given by the dashed vertical line). (d) The landscape topography for each host is given by the collection of the set of alleles' resistance levels (the x-axis values for inflection points in part c). The connections between the three highlighted genotypes from part c are shown in the host specific color.

Results and Discussion

We used our experimental approach to construct the resistance landscape in three enteric species (*Escherichia coli*, *Salmonella enterica,* and *Klebsiella pneumoniae*) for an allele set comprised of all combinations of five particular mutations ($2^5 = 32$ alleles, nodes in Figure 9a) to the TEM-1 allele of the *bla* gene, which encodes a beta-lactamase. Given that these five mutations in combination increased resistance in *E. coli* to cefotaxime (a beta-lactam antibiotic) by five orders of magnitude, Weinreich and colleagues used this allele set to construct the resistance landscape in *E. coli* (44). In this study, we used this allele set to determine the degree of topographic alignment across three enteric species by analyzing the existence and form of host epistasis for each mutation connecting two alleles in our set.

Only 8 out of the possible 80 mutations exhibited sign host epistasis (edges split into red, blue, and yellow connecting distinct nodes in Figure 9a). Furthermore, these 8 mutations changed resistance by small amounts (purple, green, and orange points near zero in Figure 9b) and larger effect mutations exhibited similar increases in resistance across species (brown points in Figure 9b). Thus, we found that the host-specific landscapes were generally aligned (apparent in the structural similarity among the diagrams in Figure 9c).

As illustrated in Figure 7a and b, topographical congruence potentially translates to species effectively crowdsourcing adaptive evolution of a mobile gene from another species. However, the landscape alignment in our three species was not perfect. As seen in Figure 7c, even a handful of mutations exhibiting sign host epistasis can impact the evolutionary trajectory of a mobile gene. To gauge the consequences of the topographical differences we found, we simulated evolution on our empirically determined landscapes (see Materials and Methods for details). Each simulation consisted of several rounds of stochastic mutation (limited to alleles in our set) and selection for maximal resistance (as if all alleles were placed on a drug gradient and the allele growing at the highest concentration was picked). As a baseline, we tracked the average level of resistance (over 1,000 replicates) for genetic evolution within a single host without transfer to another species (Figure 9d, SI Figure 13a and d). To determine the effect of host-switching due to HGT, we simulated evolution over distinct periods where several rounds of stochastic mutation and selection occurred in different species (Figure 3b-c, SI Figure 13b-c, and SI Figure 13e-f). That is, in each period, the mobile gene evolved in a single species for several rounds before a horizontal transfer event moved the gene into a different species, initializing the next period. Specifically, the mobile gene evolved in a different species during a middle period (hereafter the ‘transient’ host) from the species during the first and third periods (hereafter the ‘focal’ host). Despite time evolving in a transient species, the final level of resistance for the same total period of evolution, was statistically indistinguishable (SI Figure 14). This pattern was consistent when HGT occurred between different species at different periods. Thus, our species can effectively crowdsource the evolution of antibiotic resistance in this case.

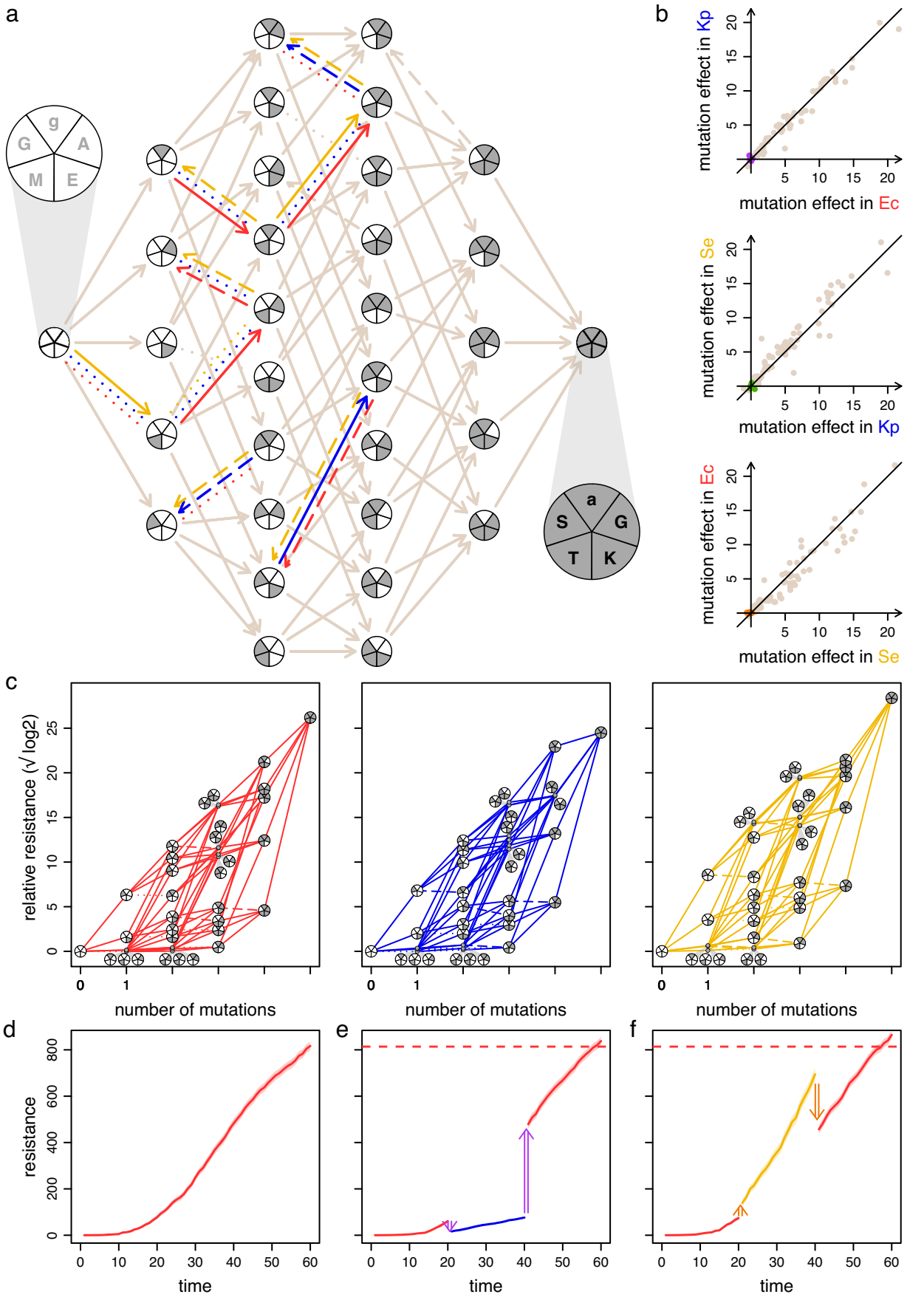

a
g
G
A
M
E
a
S
G
T
K
b
mutation effect in Kp
mutation effect in Ec
mutation effect in Se
mutation effect in Kp
mutation effect in Ec
mutation effect in Se
c
relative resistance (√ log2)
number of mutations
number of mutations
number of mutations
d
e
f
resistance
time
time
time

Figure 9: Multi-host landscapes of a mobile gene reveal evolutionary crowdsourcing. (a) The resistance landscape with five mutations to the TEM-1 allele encoding a beta-lactamase was constructed for three enteric species: *Escherichia coli* (red), *Klebsiella pneumoniae* (blue), and *Salmonella enterica* (yellow). The five mutations are displayed with five wedges in each node where shading indicates the presence of a mutation. In the highlighted larger circle, each mutation is indicated with either a lower- or upper-case letter corresponding to either a single nucleotide polymorphism in the promotor region or an amino acid substitution, respectively. Starting at the 12 o'clock position and moving clockwise on the pie charts, the mutation were g4205a, A42G, E104K, M182T, and G238S. The mutational steps that exhibited sign host epistasis are split into three arrows (one for each host) where the effect in each host (red, blue, and yellow edges) is indicated as beneficial (solid line), neutral (dotted line), or deleterious (dashed line). If the mutational step had the same effect in all three hosts (i.e., no sign host epistasis), then only one edge (brown) with the corresponding effect (solid, dotted, or dashed) is shown. (b) The effect of each mutational step (80 edges in the directed network in part a) on the level of resistance (akin to the slope in part c) are compared across each species pairing. The mutational steps that exhibited sign host epistasis (split arrows in part a) had small effects (purple, green, orange dots near zero in the top, middle, and bottom panel, respectively) compared to the mutational steps exhibiting no host sign epistasis (brown dots). (c) Given the low number of mutational steps that exhibited sign host epistasis and their small effects, the *E. coli, K. pneumoniae,* and *S. enterica* (red, blue, and yellow, respectively) landscapes were largely aligned. (d) Using an evolutionary simulation (see Materials and Methods for details), the average level of resistance (over 1,000 replicates with the standard error given by the shading) increased due to gene evolution in *E. coli.* (e, f) Under the same evolutionary period as part d, a similar evolutionary endpoint is reach in *E. coli* when the gene evolved in a different species during a middle period facilitated by HGT events (double-ended arrows). The dashed red line indicates the average evolutionary endpoint from part d.

For both the landscape reconstruction and evolutionary simulation, there are a few caveats that warrant attention. First, we focused on a small set of mutations in one gene in three closely related species. It is possible that considering additional or alternative mutations in the *bla* gene, a different gene, or different species (e.g., hosts that are more phylogenetically distant) could lead to more or less cross-species topographical alignment. Second, our "fitness" landscapes mapped each allele to its level of resistance, but the fitness of any genotype will be affected by more than resistance (50). For instance, genotypes may exhibit different baseline growth rates (i.e., in drug-free conditions), which may not correlate with drug resistance (51). Third, our evolutionary simulation makes several simplifying assumptions. Specifically, we assumed a series of selective sweeps (where a more resistant mutant replaces its immediate ancestor) punctuated by cross-species HGT. However, multiple genotypes within and across species can compete in

natural bacterial communities, and the outcomes of these competitions (as well as the opportunities for HGT) will be governed by the distribution of the relevant species across a potentially heterogeneous environment (e.g., a multi-species biofilm in a drug gradient). In addition, our simulations glossed over some of the distinguishing features of conjugative plasmids, such as copy number, fitness cost, and basic rates of conjugation and segregational loss. Many of these plasmid features vary with host context (15, 52, 53) and these differences may thereby influence HGT opportunities and competitive outcomes. Considering these limitations, more complex simulation frameworks incorporating factors such as growth differences, clonal interference, environmental heterogeneity, and basic life history traits of plasmids would be a promising direction for future work.

Despite these noted caveats, evolutionary adaptation of our focal mobile gene can be a cosmopolitan affair, in which the progress made in one species readily translates to progress in another. We emphasize that the availability of widespread evolutionary crowdsourcing in a microbial community through HGT will be affected by the prevalence of host sign epistasis, its effect size, and topographical location. Indeed, we subjected the multi-host landscapes of an essential chromosomal gene (SI Figure 15) from a recently published study (48) to our evolutionary simulation framework showing that occurrence of HGT in this case would have hindering (as in Figure 7b) or facilitating (Figure 7c) effects depending on the focal and transient host (SI Figure 16 and SI Figure 17). We find it interesting that the level of sign host epistasis and the accompanying effect sizes for this essential gene are greater than the level for our nonessential gene, which is commonly found on plasmids. Indeed, this pattern highlights a connection to the “complexity hypothesis,” in which the proteins encoded by genes experiencing higher rates of HGT are less connected to other proteins in the cell (54, 55). If fewer intracellular connections translate to fewer opportunities for host-dependencies, it would not be surprising if these more “modular” mobile genes also exhibit less sign host epistasis. However, it will be necessary to construct landscapes for additional genes undergoing different rates of HGT across the same set of species (i.e., controlling for phylogenetic relatedness) to gauge whether evolutionary crowdsourcing is more often an option for genes that experience a high rate of HGT.

In summary, we present a novel framework to examine the molecular evolution of mobile genes – a highly relevant subset of genes evolving with an additional mode of genetic transmission (i.e., horizontal). In this study, we uncovered that for a small set of mutations in a common mobile gene, landscape topography and thus evolutionary outcomes are largely aligned across species. These findings suggest that mobile genes adapting in one species can lead to adaptation for another species. In such a case, conjugative plasmids and other vehicles of cross-species HGT will enable a distributed form of genetic evolution across a bacterial community, where any particular species can crowdsource adaptation from other community members.

Author contributions

O.K. and B.K. conceived of the presented ideas. O.K. and B.K. developed the theory, simulations, and experimental protocols. O.K., N.G., and B.W. performed the experiments. O.K. and B.K. wrote the manuscript. E.M.T. and B.K. supervised the project. All authors discussed the results and contributed to the final manuscript.

Acknowledgements

This work was supported by the National Institute of Allergy and Infectious Diseases Extramural Activities from the National Institutes of Health and Division (grant number. R01 AI084918 to E.M.T. and B.K.), the Environmental Biology Division from the National Science Foundation (grant number 2142718 to B.K. and E.M.T.), and a Graduate Research Fellowship from the National Science Foundation (grant number DGE-1762114 O.K.

CHAPTER 2 supporting material:

Materials and Methods

General reagents.

Unless otherwise noted, all enzymes and related buffers were obtained from New England Biolabs. Plasmid isolation kits were obtained from Qiagen. DNA cleaning and gel extraction kits were obtained from Zymo. Oligonucleotide primers were obtained from Integrated DNA Technologies. Sanger sequencing was conducted by GeneWiz.

Allele construction and barcoding.

To generate the beta-lactamase alleles, we mutated the cloning plasmid, pBR322, using the New England Biolabs Site-Directed Mutagenesis Kit per manufacturer's instructions. The plasmid contains two accessory genes, the focal *bla* gene and a tetracycline resistance gene, *tetA*. Plasmid maintenance was ensured by supplementing the culture medium with 15 μg ml^{-1} tetracycline. TEM-1 was the starting allele for the *bla* gene and used as the template to start the mutagenesis. The first round of cloning created the five single variant alleles. For each variant, a custom pair of primers was designed where the mutation was coded in the forward primer. All primers are listed in SI Table 10. Each engineered variant was isolated and variant sequences were confirmed with Sanger sequencing. All Sanger sequencing primers are listed in SI Table 11. The double, triple, quadruple, and quintuple variant alleles were constructed using the same procedure. All engineered alleles are listed in SI Table 12.

Each beta-lactamase allele was associated with three unique molecular barcodes. To facilitate the engineering of barcodes, we modified the backbone of the pBR322 plasmid before generating the beta-lactamase alleles (described above) by adding two restriction sites, NsiI and NcoI, downstream of the *bla* gene. To barcode each allele, we digested the mutated plasmids with NsiI and NcoI at 37°C for 1 h and the restriction enzymes were heat inactivated at 65°C for 20 min. We isolated the digested vector backbone using a gel extraction kit and purified the DNA. We next prepared the double-stranded barcoded fragments to be inserted by ligation using two oligonucleotides: (1) an oligonucleotide with 18bp random barcode sequences nested between the NsiI and NcoI cut sites to be used in directional cloning, and (2) a shorter priming oligonucleotide containing homology to the barcode oligonucleotide. These two oligonucleotides are listed in SI Table 13. To take the two oligonucleotides into a double-stranded fragment, we mixed 1 μL of each oligonucleotide with 5 μL of CutSmart Buffer and 47.5 μL of ddH_2O and annealed these oligonucleotides together by incubating at 98°C for 3 min followed by a ramping down to 25°C at -0.1°C/s. After annealing, we added 1 μL of Klenow polymerase (exonuclease negative) and 1.65 μL of 1mM dNTPs to make the barcode

fragment double stranded by incubating at 25°C for 15 min, 75°C for 20 min, and then a ramping down to 37°C at -0.1°C/s. We digested the double-stranded fragment using the same enzymes and protocol for digesting the vector backbone described above. The digested barcoded fragment was purified. The digested vector and barcoded fragment were ligated at 21°C for 30 min, the enzymes were heat inactivated at 65°C for 10 min and the circular products were transformed into *E. coli*. To get three barcodes per beta-lactamase allele, we isolated and sequenced the barcode sequences in three colonies with Sanger sequencing. Given the 32 alleles with 3 barcodes per allele, the beta-lactamase library contained 96 engineered plasmids.

To create the heterogenous populations for each host species, we first transformed the 96 engineered plasmids independently into each host, isolated a colony from each transformation, and stored at - 80°C in 15% (v/v) glycerol. For each species, the 96 strains were grown overnight, mixed at equal volumes, and stored at - 80°C in 1 ml aliquots to create library stocks for each species.

Bacterial Strains, Media, and Culture Conditions.

Hosts included three *Enterobacteriaceae* species: *Escherichia coli* B strain REL606 (56), *Klebsiella pneumoniaea* Kp08 (7), and *Salmonella enterica* serovar *typhirium* LT2 (57). We used the first letter of the genus and species name (Ec, Kp, and Se, respectively) to refer to these species throughout. All strains were cultured at 37°C in lysogeny broth (LB).

Pooled competitions assays.

The resistance level conferred by the *bla* alleles was estimated using a modified version of a standard minimum inhibitory concentration assay (58). Briefly, a 1 ml library stock was thawed and added to 50 ml of growth medium supplement with 15 μg ml^{-1} tetracycline and grown overnight. The library was diluted to an initial cell density close to 10^5 cell ml^{-1}. We plated the diluted libraries in triplicate to assess the actual initial cell density which is reported in SI Table 14. To start competitions assays, 2.5 ml of diluted library was inoculated into 41 test tubes with 2.5 ml medium. One test tube contained growth medium without CTX. The other 40 test tubes were supplemented with CTX according to a drug-gradient from 0.00393 up to 2049.37 μg ml^{-1} using $\sqrt{2}$-fold dilutions. After overnight growth in the 41 test tubes, 1ml samples were taken from the test tubes with visible turbidity for library amplification and sequencing. These samples were harvested by centrifugation and the cell pellets were stored at -20°C. Final density was also determined for each of the test tubes with visible turbidity reported in SI Table 15 using dilution plating.

Library amplification and sequencing.

Plasmids were extracted from the cell pellets stored at -20°C and the barcode region was amplified using the primers homologous to the plasmid backbone with the following conditions: 95°C for 3 min, five cycles of 98°C for 15 s, 65°C for 15 s, 72°C for 30 s, and 72°C for 1 min. Amplicons were then purified with AMPure XP beads (Beckman Coulter) at 1:1 ratio. Each sample's purified PCR product was amplified with a unique pair of forward and reverse indexing primers plus SyberGreen with the following PCR conditions on a miniOpticon (Bio-Rad): 95°C for 3 min, fifteen cycles of 98°C for 20 s, 60°C for 15 s, 72°C for 30 s, and 72°C for 2 min. Using the relative fluorescence units, the amplicons were mixed, gel extracted, quantified by Qubit fluorometry, and sequenced on the Illumina NextSeq500 platform by the Microbial Genome Sequencing Center using custom sequencing primers. All PCR and custom Illumina sequencing primers are listed in SI Table 16.

Library sequence analysis.

To determine barcode frequency, the 18bp barcode from each sequenced read was extracted from the FASTQ file and clustered into groups using Bartender (59). The barcode clusters were matched to the barcodes used in this study (listed in SI Table 17). Barcode frequency was calculated by dividing the barcode counts by the total counts in the sample. Next, the approximate growth rate for each barcode was calculated using the estimated initial and final cell densities (SI Table 14 and SI Table 15) along with barcode frequency using equation [10.1] (see SI section 10 for details). Given that we have three barcodes per allele, these served as internal replicates in the competitions. We eliminated the most deviate barcode for each allele by calculating the pair-wise differences between the barcode approximate growth rates, summed the differences across the concentrations, and found the most globally deviant barcode. For each of the remaining two barcodes per allele, we fit a four-parameter log logistic dose-response curve using the drc package in R (60). Given that there was an inflation of the bottom asymptote (i.e., non-zero), we took the lower asymptote parameter average for each species using approximately half of the alleles with the lowest MIC (n = 17). For each barcode, we used these species-specific lower asymptote values to fit a three-parameter log logistic dose-response curve. By fixing the lower asymptote, it established the same "no-growth" baseline across the alleles which improved the curve-fitting. We used the inflection point, namely where growth rate is dropping most precipitously, of the dose-response curve as the proxy for the level of resistance for each allele. In the event the lower asymptote was zero, then our resistance level would be an IC50 value. Since all the barcodes were pooled and thus exposed to the same environments, using the relative inflection points as the proxy for level of resistance was justified.

For the less resistant alleles, the concentration gradient was truncated to the highest concentration where the approximate growth rate of the most resistant allele was unaffected by the antibiotic. For the most resistant allele (g4205a, A42G, E104K, M182T,

G238S) and the second most resistant allele (g4205a, A42G, M182T, G238S), the entire concentration gradient was used for fitting the three-parameter log logistic dose-response curve to obtain the inflection point for these alleles. The three-parameter estimates for the dose-response curve (upper asymptote, steepness, and inflection point) for each barcode-allele-species combination are given in SI Table 10.

To construct the host-specific landscapes (Figure 9), the level of resistance for each neighboring allele (two estimates per allele given in SI Table 10) were compared to determine if a mutational step was beneficial, deleterious, or neutral (solid, dashed, or dotted lines in Figure 9a and c). If the resistance estimates for the single mutant neighbor were all higher than the focal genotype, the mutational step was beneficial. If the estimates for the single mutant neighbor were all lower than the focal genotype, the mutational step was deleterious. If the estimates for the single mutant neighbor overlapped with the focal genotype, the mutational step was neutral. Each mutational step was compared across the three species to calculate the amount of sign host epistasis. A mutational step exhibited sign host epistasis if the effect (beneficial, deleterious, or neutral) of the mutation was different in one of the three species.

Evolutionary Simulations

Each evolutionary simulation comprised of periods of host-specific evolution, in which the focal gene evolved for a specified number of time steps inside a single host species (SI Table 19). An HGT event was defined as a switch in the host of the evolving gene, which occurred at specified time steps (SI Table 19). Therefore, a simulation for a gene evolving in different hosts over time consisted of distinct periods of single-species evolution linked together by HGT events. Within each period, mutation and selection occurred at each time step. Specifically, we considered a set of random mutants along with the ancestral allele where the genotype conferring the highest resistance fixed. We note that at some time steps the most resistant genotype may have been the ancestral allele if no mutants were generated (which becomes more likely as the mutation rate decreases) or if no generated mutant was more resistant than the ancestor (which becomes more likely as fewer mutational neighbors improve resistance). Technically, we simulated this mutation-selection process by computing the probability of genotype j fixing given that genotype i is the current ancestor. This set of probabilities depended on mutation rate, population size, the fitness landscape, and the values of i and j (see SI section 11 for details on calculating the probabilities). If a mutation exhibited sign host epistasis, or a mutation exhibited magnitude host epistasis shifting the rank ordering (based on resistance) of the mutational neighbors, the evolutionary trajectory of a gene could be affected by HGT. Thus, sign or magnitude host epistasis could enable HGT to impact the path and endpoint of adaptive evolution. We note that our simulation approach treats the evolutionary trajectory of a gene as a sequence of single allele states, where the most beneficial allele of those mutationally available fixes after each time step. That is, we did not model the process of one allele displacing another over multiple time steps.

Therefore, differences in allelic replacement rates across species (the time for a selective sweep) or the potential for multiple alleles to coexist (forms of clonal interference) were not incorporated into this approach. One way to think of our simulation is to imagine a population distributed across a drug gradient and the most resistant genotype is selected at every discrete time step to initiate the next round of population growth and placement along the drug gradient, which is a reasonably approximation of some directed evolution schemes (61). The empirically determined host-specific landscapes provided the information about the beneficial mutants available at each focal genotype (number and ranking). Since each genotype had multiple (replicate) estimates for its level of resistance (SI Table 18), at each time step, the resistance for each allele was sampled randomly from the set of estimates. Therefore, for each time step of the simulation, the host landscape could potentially shift; however, these shifts were small given that the variance across the resistance estimates in our assay were generally very low. The set of simulations and parameters used are given in SI Table 19.

CHAPTER 2 supplementary information:

SI section 10 : Derivation of the approximate growth rate

Let c^* be the highest concentration where the genotype with the highest resistance does not see a drop in frequency. That is, for $c > c^*$, the genotype frequency for the most resistant genotype will drop, indicative that the genotype became affected by the antibiotic. In other words, this corresponds to a situation where there is one genotype that hasn't had its calculated growth rate drop. We label this most resistant genotype as g'. We let the growth rate calculated for genotype g at concentration c be given by m_g^c. To calculate this growth rate, we used a time period of 24 hours even though the culture could have been growing for less than 24 hours. We let μ_g^c be the actual growth rate of genotype g at concentration c. Now at concentration c^*, we will assume that for genotype g'

$$\mu_{g'}^{c^*} = m_{g'}^{c^*},$$

which is simply assuming that growth is occurring during the full $T = 24$ hours. We will assume that this growth rate remains constant for lower drug concentrations.

Now, consider some $c < c^*$. We have the following:

$$n_{g'}^c(t_c) = n_{g'}^c(0)e^{\mu_{g'}^{c^*}(t_c)}$$

$$t_c = \frac{1}{\mu_{g'}^{c^*}} \ln \frac{n_{g'}^c(t_c)}{n_{g'}^c(0)}$$

If $n_*^c(t)$ is the total cell count at time t, and $b_g^c(t)$ is the proportion of barcodes associated with genotype g in concentration c at time t, then we have

$$t_c = \frac{1}{\mu_{g'}^{c^*}} \ln \frac{n_*^c(t_c) b_{g'}^c(t_c)}{n_*^c(0) b_{g'}^c(0)}$$

Or, assuming that the number and proportion of cells does not change from t_c to T:

$$t_c = \frac{1}{\mu_{g'}^{c^*}} \ln \frac{n_*^c(T) b_{g'}^c(T)}{n_*^c(0) b_{g'}^c(0)}$$

Now, we can consider any arbitrary genotype:

$$n_g^c(t_c) = n_g^c(0)e^{\mu_g^c(t_c)}$$

$$\mu_g^c = \frac{1}{t_c} \ln \frac{n_*^c(t_c) b_g^c(t_c)}{n_*^c(0) b_g^c(0)}$$

Or, assuming that the number and proportion of cells does not change from t_c to T:

$$\mu_g^c = \left(\frac{\ln \frac{n_*^c(T) b_g^c(T)}{n_*^c(0) b_g^c(0)}}{\ln \frac{n_*^c(T) b_{g'}^c(T)}{n_*^c(0) b_{g'}^c(0)}} \right) \mu_{g'}^{c^*}. \qquad [10.1]$$

SI section 11 : Probabilities that a given mutation is the most resistant under gradient selection

We modeled mutation and selection using a first-order Markov process. First, we describe the probability that the population shifts from the ancestral genotype (hereafter focal genotype) to a single mutant neighboring genotype at the end of a time step. We let μ be the probability a mutation arises in the focal genotype. The probability of mutation is small such that we ignore any instance of two or more distinct mutations and assume that most individuals within the population have the focal genotype (and the rare mutants have only a single mutation). All possible single mutant genotypes are denoted by the set $\mathbf{M} \equiv \{m_1, m_2, m_3, \ldots m_y\}$ (where $y \equiv |\mathbf{M}|$ is the number of mutants). Focusing on one mutant genotype (e.g., m_j), if the focal genotype is copied n time (independently), the probability that there are k mutants with this genotype will be:

$$\pi_k(m_j) = \binom{n}{k} (\mu)^k (1-\mu)^{n-k}$$

The probability that there are one or more m_j individuals is then:

$$1 - \pi_0(m_j) \approx 1 - e^{-\mu n}$$

We denote this probability as $\pi^*(m_j)$.

After mutation, the population undergoes growth followed by selection across an antibiotic gradient. Here we assume that there is enough growth of the population such that if a mutant genotype is present, it is distributed across the entire antibiotic gradient. For a given focal genotype i, and a set of neighboring mutant genotypes labeled $\mathbf{M}$, we denote the genotypes in set $\mathbf{M}$ that are more resistant as $\mathbf{H}_\mathbf{M}(i)$. First, we start with the probability that the most resistant genotype is the focal genotype i. In this case, the probability is:

$$p_i = \prod\nolimits_{m \in \mathbf{H_M}(i)} \left(1 - \pi^*(m)\right)$$

If the focal genotype is the most resistant, then this probability becomes one. If a mutant neighbor is more resistant than the focal genotype, this probability becomes less than one. Generally, as the focal genotype becomes less resistant is becomes less probable for the population to stay at this genotype.

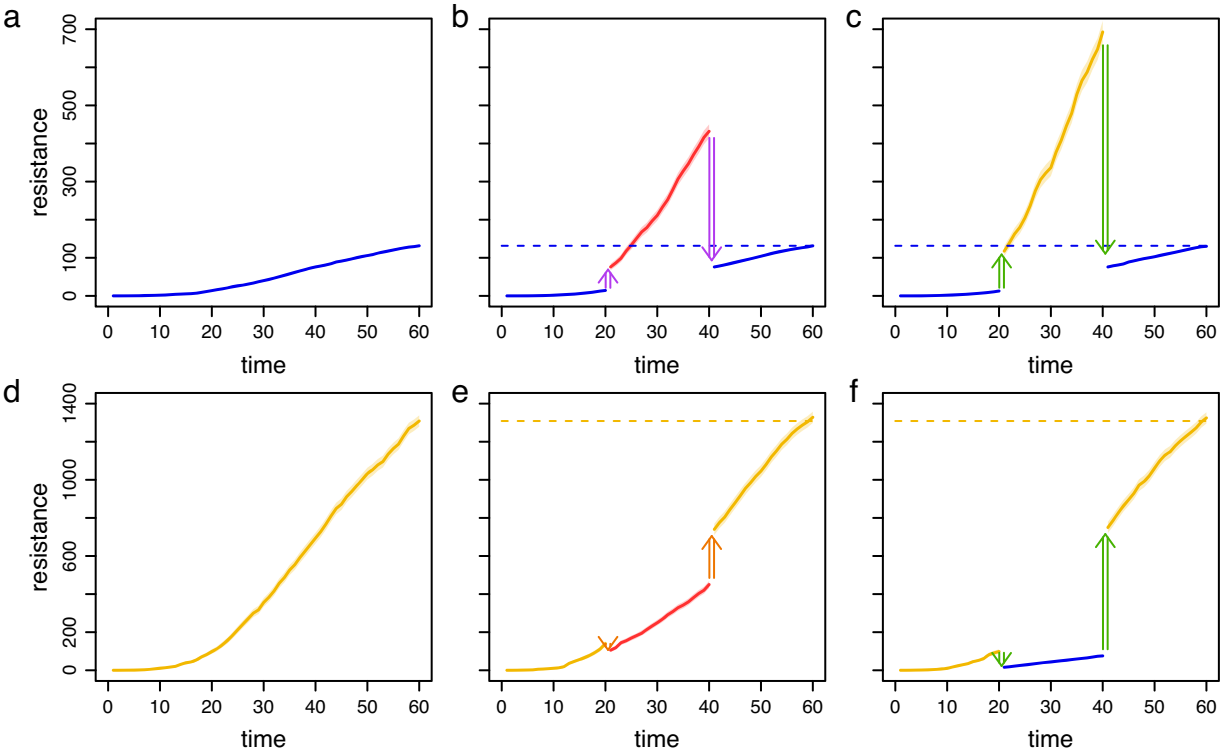
Next, we turn to the probability that the most resistant genotype comes from the mutant set $\mathbf{M}$. Here we need to introduce more notation. For a given genotype g, and a set of genotypes labeled $\mathbf{S}$, we denote the genotypes in $\mathbf{S}$ that are equally resistant to genotype g as $\mathbf{E_S}(g)$. If genotype g is in the set $\mathbf{S}$, we will define $\mathbf{E_S}(g)$ to *not* include genotype g (i.e., $g \notin \mathbf{E_S}(g)$). Therefore, $\mathbf{E_S}(g)$ are genotypes with equivalent resistance to genotype g, *other than* g *itself*. Consider a mutant genotype m_j, which has higher resistance than the focal genotype i. We denote the set of other mutants that have higher and equivalent resistance to genotype m_j as $\mathbf{H_M}(m_j)$ and $\mathbf{E_M}(m_j)$, respectively. The probability that genotype m_j is chosen is given by:

$$p_{m_j} = \pi^*(m_j) \prod_{m \in \mathbf{H_M}(m_j)} (1 - \pi^*(m)) \left[\sum_{\mathbf{S} \subseteq \mathbf{E_M}(m_j)} \frac{\left(\prod_{m' \in \mathbf{S}} \{\pi^*(m')\}\right) \left(\prod_{m'' \in \mathbf{E_M}(m_j) - \mathbf{S}} \{1 - \pi^*(m'')\}\right)}{1 + |\mathbf{S}|} \right]$$

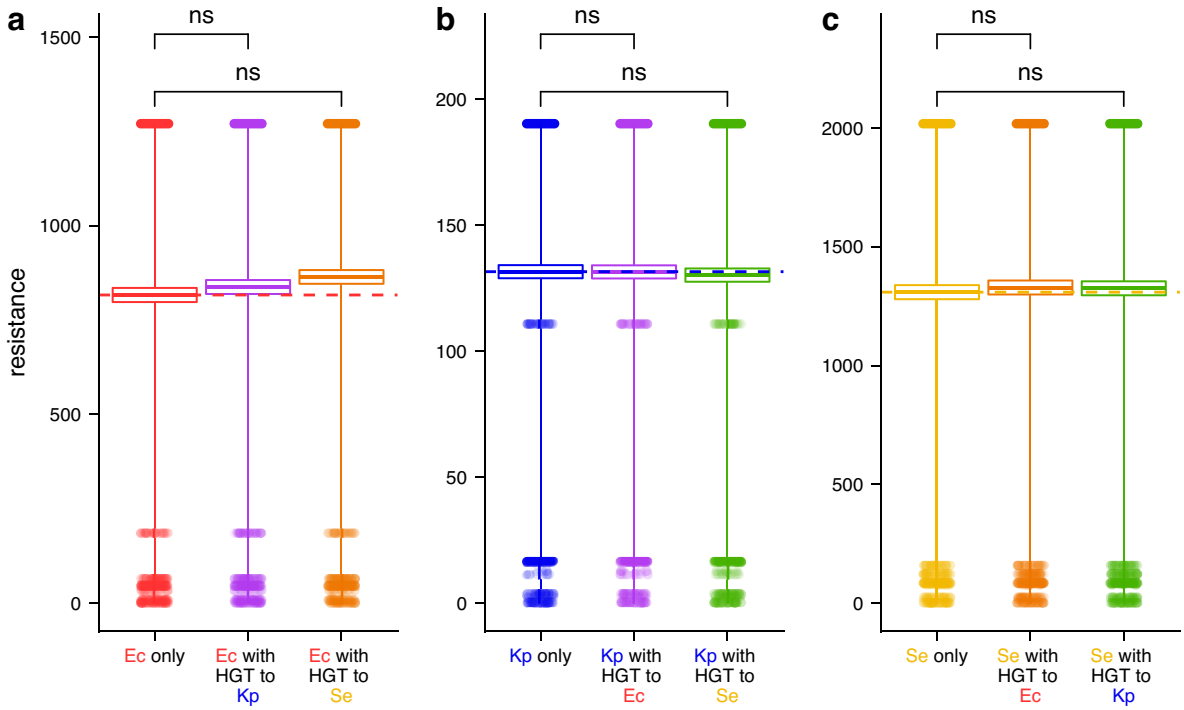
If there are never any ties between genotypes with regards to resistance, then $\left|\mathbf{E_M}(m_j)\right| = 0$ and the above equation simplifies to:

$$p_{m_j} = \pi^*(m_j) \prod_{m \in \mathbf{H_M}(m_j)} (1 - \pi^*(m))$$

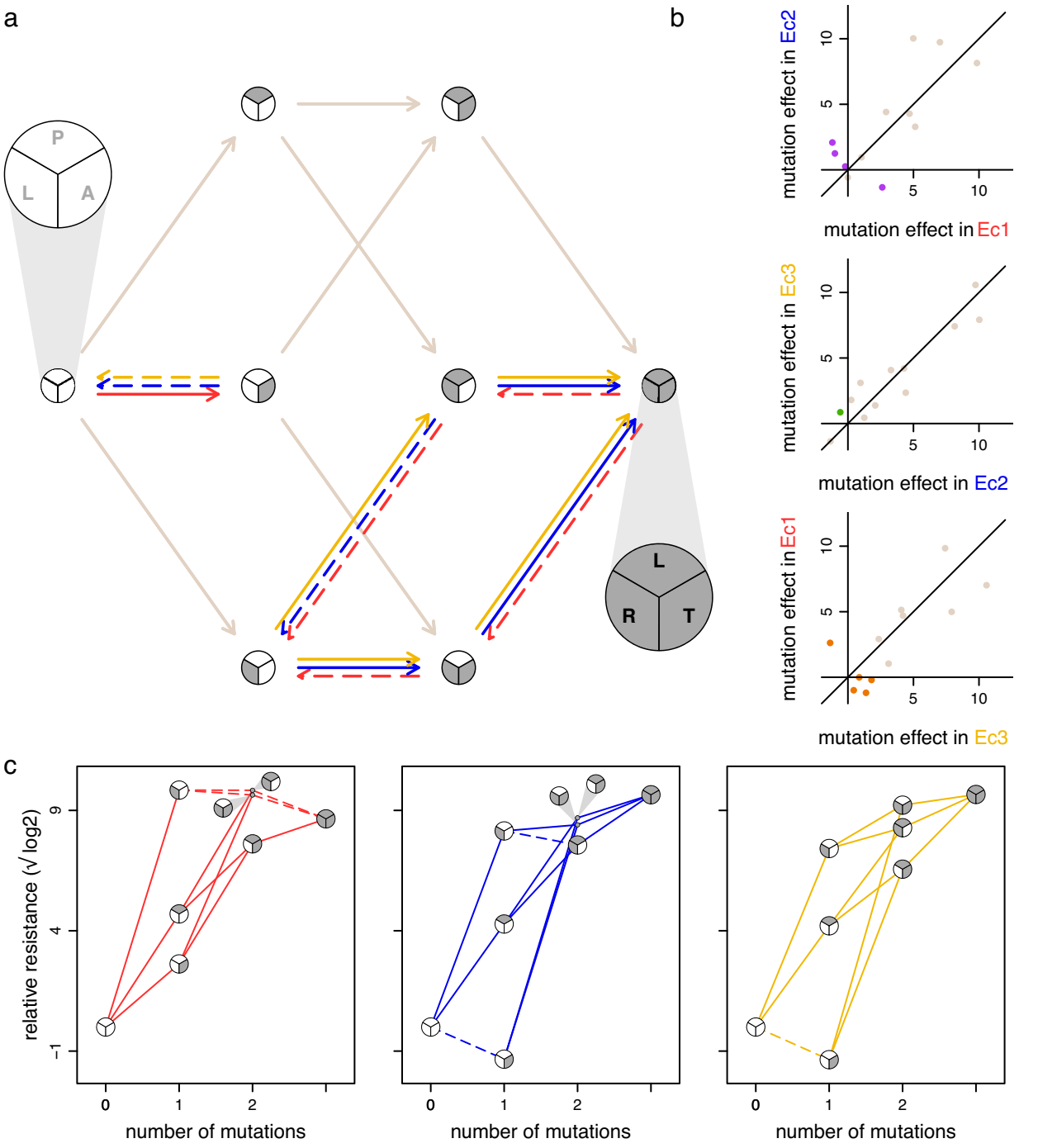
If a particular mutant m_j is the most resistant mutant of the focal genotype ($\left|\mathbf{H_M}(m_j)\right| = 0$) then the probability of picking the mutant is $\pi^*(m_j)$ which depends only on the population size and the mutation rate. More generally, the probability of selecting mutant m_j covaries positively with its ranking in the set (i.e., the more resistant this mutant genotype is relative to the other mutants of the focal genotype, the more likely it is to be selected).



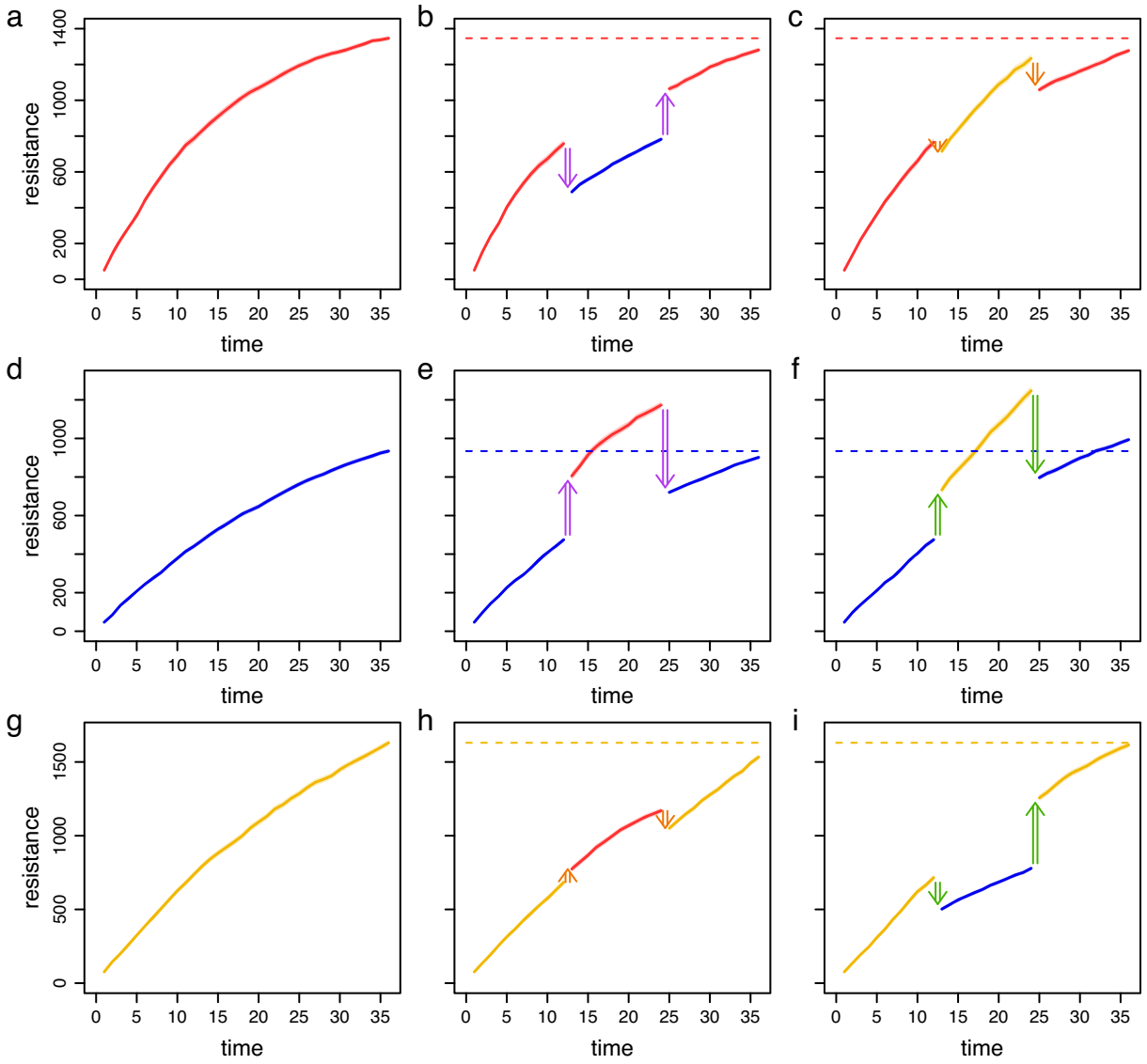
SI Figure 13: Evolutionary simulations in *K. pneumoniae* and *S. enterica* where the mobile gene evolved without (a and d, respectively) and with HGT (b, c, e, and f, respectively). The simulation framework and graphical representation is the same as Figure 9. For both *K. pneumoniae* and *S. enterica,* a similar evolutionary endpoint is reached when the gene evolved in a different species during a middle period facilitated by HGT events.



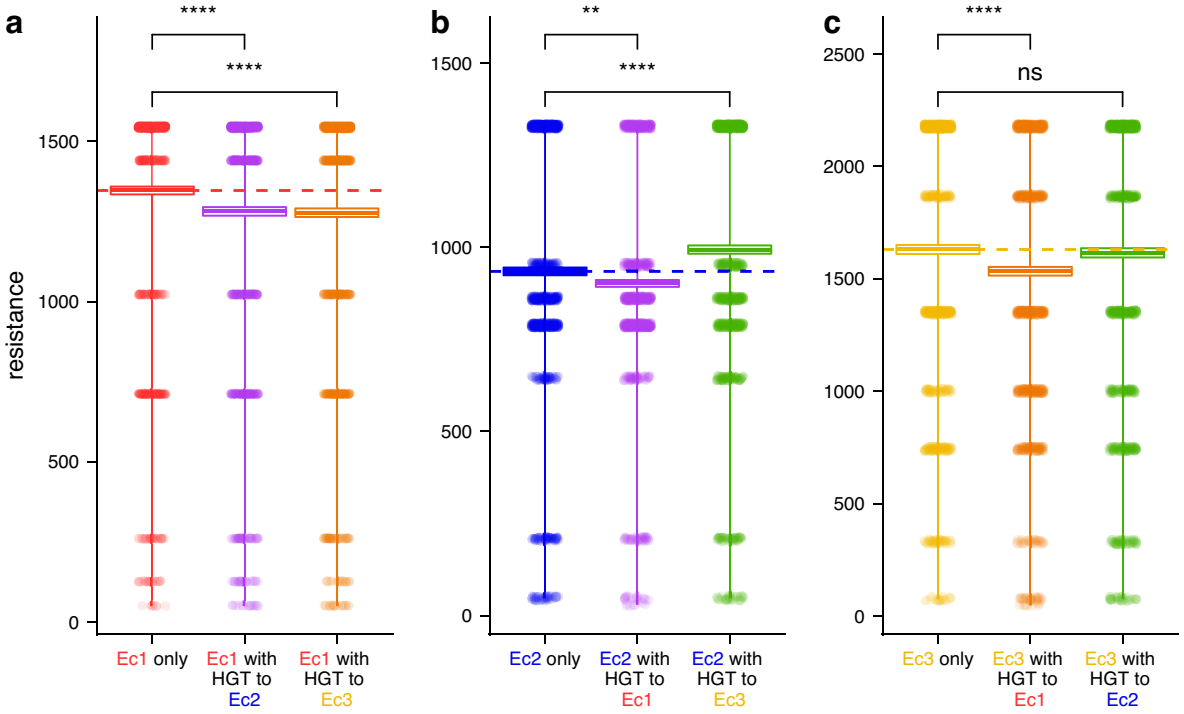
SI Figure 14: Similar evolutionary endpoints are reached when the mobile gene evolved with and without HGT. Each box summarizes the mean and standard error of the endpoint resistance values from 1000 replicates from an evolutionary simulation, which corresponds to the evolutionary endpoints in Figure 9d-f and SI Figure 13. In the baseline simulations (labelled 'Ec only', 'Kp only', and 'Se only'), evolution of the gene occurred in the focal species (*E. coli, K. pneumoniae, and S. enterica*, respectively) for the entire evolutionary period and there is no transfer to another species. In the simulations with HGT, the mobile gene evolved in a transient species during the middle period. All comparisons to the baseline simulations were insignificant by a t-test with a Bonferroni correction (p=0.69, p=0.07, p=1, p=1, p=1, and p=1, a to c from top to bottom respectively).



SI Figure 15: Multi-host landscapes of an essential gene from the study by Guerrero and colleagues (48)**.** The graphical representations are the same as Figure 9. (a) The resistance landscape with three mutations to an *E. coli* allele of DHFR was constructed in three strains of *E. coli* referred to as 'Ec1', 'Ec2', and 'Ec3' (red, blue, and yellow, respectively). In the pie blow ups starting at 12 o'clock, the mutations are P21L, A26T, and L28R. (b) The effect of each mutational step (12 edges in the directed network in part a) on the level of resistance are compared across each host pairing. (c) The host-specific adaptive landscapes visualized by plotting the genotypes level of resistance (relative to the WT background of each host) as a function of the number of mutations.



SI Figure 16: Evolutionary simulations with and without HGT using the DHFR data set from Guerrero and colleagues (48)**.** The simulation framework and graphical representation is the same as Figure 9.



SI Figure 17: Different evolutionary endpoints are reached where the mobile gene evolved with and without HGT. The graphical representation is the same as SI Figure 14. The data corresponds to the evolutionary endpoints in SI Figure 16. The asterisks indicate statistical significance by a t-test with a Bonferroni correction (two and four asterisks convey p-values in the following ranges: $0.001 < p < 0.01$, and $p < 0.0001$, respectively), and 'ns' indicates statistical insignificance ($p=0.71$ in part c).

SI Table 10: Primers used for Site-Directed Mutagenesis. The mutagenic primer is labelled with an asterisk. If an amino acid is being mutated, the codon is underlined. The nucleotide being mutated is **bolded**.

Mutation	Primer	Primer orientation	Sequence (5' -> 3')
g4205a	pOK84	Forward*	**A**AGCGGATACATATTTGAATGTATTTAGAAAAATAA
	pOK85	Reverse	ATGAGACAATAACCCTGATAAATGCTTC
A42G	pOK80	Forward*	T**C**CACCCAACTGATCTTCAGCATCT
	pOK81	Reverse	CGAGTGGGTTACATCGAACTG
E104K	pOK78	Forward*	CT**T**AACCAAGTCATTCTGAGAATAGTGTATG
	pOK79	Reverse	TACTCACCAGTCACAGAAAAGCA
M182T	pOK82	Forward*	C**G**TCGTGGTGTCACGCTCG
	pOK83	Reverse	CCTGCAGCAATGGCAACAACGTTGC
G238S	pOK75	Forward*	CTCAC**T**GGCTCCAGATTTATCAGC
	pOK74	Reverse	CGTGGGTCTCGCGGTATC

SI Table 11 : Primers used for Sanger sequencing.

Primer	Sequence region	Sequence (5' -> 3')
pOK6	280th amino acid to 600 downstream nucleotides	CAGGCAACTATGGATGAACG
pOK10	50th to 286th amino acid	CCTTCCTGTTTTTGCTCACC
pOK38	Promotor region to 185th amino acid	GAGGATGACGATGAGCGCAT

SI Table 12 : Engineered variants using Site-Directed Mutagenesis.

Variant	Number of mutations
g4205a	1
A42G	1
E104K	1
M182T	1
G238S	1
g4205a, A42G	2
g4205a, E104K	2
g4205a, M182T	2
g4205a, G238S	2
A42G, E104K	2
A42G, M182T	2
A42G, G238S	2
E104K, M182T	2
E104K, G238S	2
M182T, G238S	2
g4205a, A42G, E104K	3
g4205a, A42G, M182T	3
g4205a, A42G, G238S	3
g4205a, E104K, M182T	3

g4205a, E104K, G238S	3
g4205a, M182T, G238S	3
A42G, E104K, M182T	3
A42G, E104K, G238S	3
A42G, M182T, G238S	3
E104K, M182T, G238S	3
g4205a, A42G, E104K, M182T	4
g4205a, A42G, E104K, G238S	4
g4205a, A42G, M182T, G238S	4
g4205a, E104K, M182T, G238S	4
A42G, E104K, M182T, G238S	4
g4205a, A42G, E104K, M182T, G238S	5

SI Table 13 : Primers used for creating the barcode fragment. The NcoI and NsiI restriction sites are **bolded.** The homologous nucleotides used for creating the double stranded fragment are underlined.

Primer	Sequence (5' -> 3')
pOK67	CGGACCGCTGGACGTATCTTAGTTTTCTCGAGTAAGATCCAT**CCATG**GTCTGTC ACACCGAGAGGCTAGGCAGTTGCGCGCGTACGNNNNNNNNNNNNNNNNNAC CGGTCCGGTAATCGAACTGGGCGAGACATCCCAGCTTAGCT**ATGCAT**TCACTA GAGGACGCGTGTCCACGTGAAGACATCCCAGCGCTTGA
pOK68	TCAAGCGCTGGGATGTCTTC

SI Table 14 : Initial cell density of each library before selection in CTX.

Species	Number of replicates	Average initial density (cfu ml^{-1})	Standard error
Ec	6	4.07×10^5	2.67×10^4
Kp	6	3.21×10^5	2.99×10^4
Se	6	6.07×10^5	3.41×10^4

SI Table 15 : Final cell density (cfu ml^{-1}) from each library selection. The concentrations where sequencing data was obtained are **bolded** for each species. A portion of the lower concentration was not submitted for sequencing given the resistance level of the ancestral allele, TEM-1. Test tubes that were not turbid (NT) after the 24 h incubation are designated.

CTX concentration (µg ml^{-1})	Ec	Kp	Se
0.0000	**2.44×10^9**	**3.00×10^9**	**2.84×10^9**
0.00393	2.28×10^9	3.88×10^9	1.91×10^9
0.0056	2.32×10^9	3.32×10^9	2.26×10^9
0.0079	1.72×10^9	3.52×10^9	2.39×10^9
0.0112	2.00×10^9	3.22×10^9	2.33×10^9

0.0158	1.42 x 10^9	2.24 x 10^9	2.15 x 10^9
0.0223	1.06 x 10^9	2.88 x 10^9	1.97 x 10^9
0.0315	**1.06 x 10^9**	**2.36 x 10^9**	**2.37 x 10^9**
0.0445	**9.80 x 10^8**	**2.18 x 10^9**	**1.70 x 10^9**
0.0629	**9.80 x 10^8**	**3.48 x 10^9**	**1.92 x 10^9**
0.0889	**1.06 x 10^9**	**3.04 x 10^9**	**2.01 x 10^9**
0.1257	**1.32 x 10^9**	**3.44 x 10^9**	**1.74 x 10^9**
0.1777	**1.32 x 10^9**	**3.88 x 10^9**	**1.55 x 10^9**
0.2513	**1.46 x 10^9**	**3.20 x 10^9**	**2.07 x 10^9**
0.3553	**1.26 x 10^9**	**3.08 x 10^9**	**1.92 x 10^9**
0.5024	**1.50 x 10^9**	**4.00 x 10^9**	**1.91 x 10^9**
0.7104	**1.24 x 10^9**	**3.12 x 10^9**	**1.92 x 10^9**
1.0045	**1.86 x 10^9**	**3.10 x 10^9**	**2.05 x 10^9**
1.42	**1.26 x 10^9**	**2.78 x 10^9**	**1.71 x 10^9**
2.01	**1.32 x 10^9**	**3.04 x 10^9**	**2.74 x 10^9**
2.84	**7.50 x 10^8**	**3.76 x 10^9**	**2.02 x 10^9**
4.02	**1.07 x 10^9**	**3.36 x 10^9**	**1.98 x 10^9**
5.68	**1.08 x 10^9**	**2.68 x 10^9**	**1.44 x 10^9**
8.03	**1.56 x 10^9**	**3.24 x 10^9**	**2.72 x 10^9**
11.35	**1.42 x 10^9**	**3.92 x 10^9**	**1.82 x 10^9**
16.05	**2.54 x 10^9**	**3.40 x 10^9**	**2.47 x 10^9**
22.69	**1.04 x 10^9**	**4.44 x 10^9**	**2.08 x 10^9**
32.08	**1.46 x 10^9**	**3.20 x 10^9**	**1.94 x 10^9**
45.36	**1.36 x 10^9**	**3.68 x 10^9**	**1.41 x 10^9**
64.14	**1.26 x 10^9**	**2.64 x 10^9**	**1.55 x 10^9**
90.69	**1.37 x 10^9**	**1.81 x 10^9**	**1.16 x 10^9**
128.24	**4.80 x 10^9**	**2.63 x 10^9**	**1.58 x 10^9**
181.33	**1.14 x 10^9**	NT	**7.00 x 10^8**
256.4	NT	NT	**1.15 x 10^9**
362.55	NT	NT	**1.92 x 10^8**
512.65	NT	NT	NT
724.89	NT	NT	NT
1024.99	NT	NT	NT
1449.34	NT	NT	NT
2049.37	NT	NT	NT

SI Table 16 : Primers used for library amplification and sequencing. The nucleotides that are homologous to the plasmid are **bolded.** The nucleotides that are homologous to the indexing primers are underlined. The 9bp index used for multiplexing the samples is represented with N nucleotides and are sequence specific depending on the sample.

Primer	Purpose	Sequence (5' -> 3')
pOK55	PCR round 1 forward primer	CCGCGTGATTACGAGTCG**GCAGCAGATTACGCGCAGAA**
pOK56	PCR round 1 reverse primer	GGGTTAGCAAGTGGCAGCCT**AGCGCTGGGATGTCTCG**

	PCR round 2 forward primer	AATGATACGGCGACCACCGAGATCTACACNNNNNNNNNC CGCGTGATTACGAGTCG
	PCR round 2 reverse primer	CAAGCAGAAGACGGCATACGAGATNNNNNNNNGGGTTAG CAAGTGGCAGCCT
pOK57	Custom read 1 primer	TCACACCGAGAGGCTAGGCAGTTGCGCGCGTACG
pOK59	Custom read 2 primer	GCTGGGATGTCTCGCCCAGTTCGATTACCGGACCGGT

SI Table 17 : Genotype to barcode map.

Variant	Barcode 1	Barcode 2	Barcode 3
Wild-type TEM-1	GATGGCCTTTTGCCGGTT	ATTCGTAAACTTCTGGTT	GAGTCGTGCATTGAGTTC
g4205a	GTATTAGTTTTACTTTAG	TTGAGTCCCGAGGGTGGT	TACTTTTAAACTATAAGA
A42G	CAGTGAGTCAGATATCTT	GTAGCCATTTACTTCTGT	TTCGCATTATATTCCGTC
E104K	TTGGCGCCTTCTCTTCTG	GGAGCAGTAGAGGTGGTA	ATAAAGGATGTGACTGTA
M182T	CCTAGCAGCTCGTAAGAG	TCCAAAGGGTGGCACGAG	TGTGATTTACACACGTCC
G238S	ACTGCGTTTTAATATTTT	CCAACTGTAAGCCTATTT	CAGGGTCATACGAGCTTC
g4205a, A42G	CTCATTTTAGACTTCGTT	CATTCATATTAAAGTTTG	CAGGTTTTAGCATATGCC
g4205a, E104K	CGACAACGTATCAAGCTC	TCCGACCATTAAGGGTTA	GCTGGTCCGATCAGATAT
g4205a, M182T	AGGGATCTGGAGTAGGTC	AATGTGCGTTAATAGATT	TATCATAGTGAGTTCCAT
g4205a, G238S	ATAAGGTTTGTTTCCCTG	AATTTAAGTATAGAGGGG	TAGGTTAATTCTCGGTGA
A42G, E104K	ATTAGATTTATTAATATG	GTTTCCTCTAAAGATTTC	TTTTTCCTCCGCTCTGGT
A42G, M182T	CGTACCCCTTGCTGGTGG	AGAACTTGGTAACGGGGC	TTGGGACCTCTTTGGGTA
A42G, G238S	CAGCTGGTTGGTTCTCTA	ATAGTTATTTTGGAACTA	ACTACAGTAATAGTGCAT
E104K, M182T	GCTTCCTTTATTTGTTTA	GCGAGATGATTAGAGAGA	GTGTGAGACGCAGTTTAG
E104K, G238S	TGTATTGGTTAACGTTAC	TAACGCGAGTCGTAATCT	GAGGGTGTGATTAGCAAT
M182T, G238S	GATCCGATGATAGTAGTT	TATCCGTCCTCGGCAGAG	TCCCTAGCATGGATTGGC
g4205a, A42G, E104K	ATAAATATGTGGTCCCTG	CACCATCCTACAACTAAA	AATCATTCAAATCGAAGA
g4205a, A42G, M182T	TTAACACATGGTATTTAC	TTGTTTATTTACCGGACT	AAAGTTTGAGGAATAACG
g4205a, A42G, G238S	AAGTGTTCGCATTGCAAG	ACAGGAGACGGTATCTTT	TTGAAATGCTTTCGGTTA
g4205a, E104K, M182T	AGTCGTGTGGGGGCCTAC	GCGTCCTCGAGTCTTTAC	TTTGGACCACTTTTCTGT
g4205a, E104K, G238S	AACGGTTGGACCGAGCGG	CCATTTGATTTTAAGCTC	TTGCGGAAGTGGTCGTGG
g4205a, M182T, G238S	TTTGTTGCTCTTTCGATG	TCCGGTACGATTACAACG	AACCTAATTCTTAACGGA
A42G, E104K, M182T	GTGCGAAAGCATTACACT	TTGCATAGTTTCATAATA	GCAGTATGCGGAAAAGCT
A42G, E104K, G238S	AGCATGCTCTGCCGAGAA	TCCACGTACATAAATGTT	ACTGCTTAGCAGTTTGTC
A42G, M182T, G238S	GGCAGGTGAATCTACCAG	TTTAGCCCCATCACTAAC	TTTGTGTAGGTACTATCC
E104K, M182T, G238S	CAATATACTCTGTACTAA	CATCTAATTTATTGGGTA	TAGTGCTTGTTCAGGGGT

g4205a, A42G, E104K, M182T	GCATCGCTCTCATGGGTA	CATAGTCACCGGCTAGAT	GAGAGGCTAAGGTGAAAC
g4205a, A42G, E104K, G238S	TATTTGAGTTATTAGTTC	ACGTAAAGTAAGACTTCA	AGTGCAGGTTTAAATACT
g4205a, A42G, M182T, G238S	TTAGCTTGGTTTCTGTCT	GATAGGGTATATTGGCAC	TAGCATCGGGTCAGGGCG
g4205a, E104K, M182T, G238S	ATATTTTACCGTCTTAAA	GCCGCGGCGTGTGTGGTT	GTAGTAGGTGTCTCAGAC
A42G, E104K, M182T, G238S	ACGAACTTTGCTTTCTTT	GAAAACATACGGCGTGGT	AGTCATGGCATTATGAAA
g4205a, A42G, E104K, M182T, G238S	CTGTTTTTGAACTTGAAG	TGAAGCACGTAAACTATC	TTGGTTGCACCAGACATT

SI Table 18: The three-parameter estimates (inflection point, steepness, and upper asymptote) from the dose-response curve fitting for each barcode-allele-species combination.

Species	Shorthand	Barcode	Outlier	Inflection point	Slope	Upper asymptote
Ec	Wild-type TEM-1	GATGGCCTTTTGCCGGTT		0.12	3.25	0.48
Ec	Wild-type TEM-1	GAGTCGTGCATTGAGTTC		0.12	3.67	0.48
Ec	Wild-type TEM-1	ATTCGTAAACTTCTGGTT	x	0.12	3.4	0.47
Ec	g4205a	GTATTAGTTTTACTTTAG		0.13	3.39	0.44
Ec	g4205a	TACTTTTAAACTATAAGA		0.12	3.67	0.44
Ec	g4205a	TTGAGTCCCGAGGGTGGT	x	0.1	4.28	0.4
Ec	A42G	CAGTGAGTCAGATATCTT		0.12	3.74	0.48
Ec	A42G	TTCGCATTATATTCCGTC		0.12	3.72	0.47
Ec	A42G	GTAGCCATTTACTTCTGT	x	0.13	3.23	0.48
Ec	E104K	GGAGCAGTAGAGGTGGTA		0.2	3.53	0.47
Ec	E104K	ATAAAGGATGTGACTGTA		0.21	3.5	0.46
Ec	E104K	TTGGCGCCTTCTCTTCTG	x	0.22	2.59	0.48
Ec	M182T	CCTAGCAGCTCGTAAGAG		0.12	3.37	0.49
Ec	M182T	TCCAAAGGGTGGCACGAG		0.12	3.32	0.49
Ec	M182T	TGTGATTTACACACGTCC	x	0.11	3.35	0.48
Ec	G238S	ACTGCGTTTTAATATTTT		1	3.55	0.42
Ec	G238S	CCAACTGTAAGCCTATTT		1.11	3.33	0.42
Ec	G238S	CAGGGTCATACGAGCTTC	x	1.13	3.24	0.42
Ec	g4205a, A42G	CTCATTTTAGACTTCGTT		0.14	3.56	0.46
Ec	g4205a, A42G	CATTCATATTAAAGTTTG		0.13	3.61	0.45
Ec	g4205a, A42G	CAGGTTTTAGCATATGCC	x	0.14	2.95	0.47
Ec	g4205a, E104K	TCCGACCATTAAGGGTTA		0.27	3.1	0.42
Ec	g4205a, E104K	GCTGGTCCGATCAGATAT		0.28	3.42	0.42

Ec	g4205a, E104K	CGACAACGTATCAAGCTC	x	0.29	2.65	0.43
Ec	g4205a, M182T	AATGTGCGTTAATAGATT		0.13	3.31	0.5
Ec	g4205a, M182T	TATCATAGTGAGTTCCAT		0.13	3.25	0.47
Ec	g4205a, M182T	AGGGATCTGGAGTAGGTC	x	0.13	2.97	0.49
Ec	g4205a, G238S	ATAAGGTTTGTTTCCCTG		1.05	3.6	0.34
Ec	g4205a, G238S	TAGGTTAATTCTCGGTGA		1	3.39	0.34
Ec	g4205a, G238S	AATTTAAGTATAGAGGGG	x	1.05	3.81	0.35
Ec	A42G, E104K	GTTTCCTCTAAAGATTTC		0.44	3.42	0.46
Ec	A42G, E104K	TTTTTCCTCCGCTCTGGT		0.47	2.81	0.47
Ec	A42G, E104K	ATTAGATTTATTAATATG	x	0.37	3.82	0.46
Ec	A42G, M182T	AGAACTTGGTAACGGGGC		0.12	3.37	0.48
Ec	A42G, M182T	TTGGGACCTCTTTGGGTA		0.12	3.48	0.48
Ec	A42G, M182T	CGTACCCCTTGCTGGTGG	x	0.12	2.78	0.51
Ec	A42G, G238S	ATAGTTATTTTGGAACTA		4.2	2.47	0.46
Ec	A42G, G238S	ACTACAGTAATAGTGCAT		4.58	2.26	0.47
Ec	A42G, G238S	CAGCTGGTTGGTTCTCTA	x	5.13	2.09	0.48
Ec	E104K, M182T	GCTTCCTTTATTTGTTTA		0.21	3.34	0.47
Ec	E104K, M182T	GTGTGAGACGCAGTTTAG		0.21	3.44	0.47
Ec	E104K, M182T	GCGAGATGATTAGAGAGA	x	0.18	3.83	0.47
Ec	E104K, G238S	TGTATTGGTTAACGTTAC		7.03	3.12	0.41
Ec	E104K, G238S	GAGGGTGTGATTAGCAAT		7.05	2.68	0.41
Ec	E104K, G238S	TAACGCGAGTCGTAATCT	x	9.26	2.51	0.42
Ec	M182T, G238S	TATCCGTCCTCGGCAGAG		2.7	3.01	0.5
Ec	M182T, G238S	TCCCTAGCATGGATTGGC		2.74	3	0.5
Ec	M182T, G238S	GATCCGATGATAGTAGTT	x	2.69	2.63	0.49
Ec	g4205a, A42G, E104K	ATAAATATGTGGTCCCTG		0.62	2.93	0.45
Ec	g4205a, A42G, E104K	CACCATCCTACAACTAAA		0.65	2.85	0.45
Ec	g4205a, A42G, E104K	AATCATTCAAATCGAAGA	x	0.58	3.42	0.44
Ec	g4205a, A42G, M182T	TTAACACATGGTATTTAC		0.14	2.91	0.48
Ec	g4205a, A42G, M182T	TTGTTTATTTACCGGACT		0.14	3.55	0.49
Ec	g4205a, A42G, M182T	AAAGTTTGAGGAATAACG	x	0.13	3.74	0.47
Ec	g4205a, A42G, G238S	ACAGGAGACGGTATCTTT		6.78	2.49	0.44
Ec	g4205a, A42G, G238S	TTGAAATGCTTTCGGTTA		6.46	2.66	0.44
Ec	g4205a, A42G, G238S	AAGTGTTCGCATTGCAAG	x	7.03	2.42	0.44
Ec	g4205a, E104K, M182T	GCGTCCTCGAGTCTTTAC		0.27	2.95	0.47

Ec	g4205a, E104K, M182T	TTTGGACCACTTTTCTGT		0.27	3.14	0.46
Ec	g4205a, E104K, M182T	AGTCGTGTGGGGGCCTAC	x	0.29	2.59	0.48
Ec	g4205a, E104K, G238S	AACGGTTGGACCGAGCGG		6.69	2.12	0.35
Ec	g4205a, E104K, G238S	TTGCGGAAGTGGTCGTGG		6.24	2.23	0.35
Ec	g4205a, E104K, G238S	CCATTTGATTTTAAGCTC	x	6.31	2.46	0.33
Ec	g4205a, M182T, G238S	TTTGTTGCTCTTTCGATG		4.42	2.01	0.5
Ec	g4205a, M182T, G238S	AACCTAATTCTTAACGGA		4.85	2.46	0.5
Ec	g4205a, M182T, G238S	TCCGGTACGATTACAACG	x	5.27	2.39	0.53
Ec	A42G, E104K, M182T	TTGCATAGTTTCATAATA		0.38	3.52	0.47
Ec	A42G, E104K, M182T	GCAGTATGCGGAAAAGCT		0.39	3.04	0.48
Ec	A42G, E104K, M182T	GTGCGAAAGCATTACACT	x	0.46	3.21	0.5
Ec	A42G, E104K, G238S	TCCACGTACATAAATGTT		32.15	2.07	0.46
Ec	A42G, E104K, G238S	ACTGCTTAGCAGTTTGTC		33.65	2.22	0.46
Ec	A42G, E104K, G238S	AGCATGCTCTGCCGAGAA	x	58.76	3.24	0.45
Ec	A42G, M182T, G238S	GGCAGGTGAATCTACCAG		5.39	2.29	0.49
Ec	A42G, M182T, G238S	TTTGTGTAGGTACTATCC		4.8	2.47	0.48
Ec	A42G, M182T, G238S	TTTAGCCCCATCACTAAC	x	6.21	2.55	0.51
Ec	E104K, M182T, G238S	CAATATACTCTGTACTAA		34.19	1.82	0.47
Ec	E104K, M182T, G238S	CATCTAATTTATTGGGTA		35.52	1.75	0.47
Ec	E104K, M182T, G238S	TAGTGCTTGTTCAGGGGT	x	31	1.8	0.48
Ec	g4205a, A42G, E104K, M182T	GCATCGCTCTCATGGGTA		0.6	2.65	0.48
Ec	g4205a, A42G, E104K, M182T	GAGAGGCTAAGGTGAAAC		0.55	2.76	0.48
Ec	g4205a, A42G, E104K, M182T	CATAGTCACCGGCTAGAT	x	0.67	3.22	0.49
Ec	g4205a, A42G, E104K, G238S	TATTTGAGTTATTAGTTC		45.24	2.41	0.41

Ec	g4205a, A42G, E104K, G238S	AGTGCAGGTTTAAATACT		47.43	2.34	0.41
Ec	g4205a, A42G, E104K, G238S	ACGTAAAGTAAGACTTCA	x	53.16	2.07	0.41
Ec	g4205a, A42G, M182T, G238S	TTAGCTTGGTTTCTGTCT		8.8	2.37	0.49
Ec	g4205a, A42G, M182T, G238S	TAGCATCGGGTCAGGGCG		8.62	2.34	0.49
Ec	g4205a, A42G, M182T, G238S	GATAGGGTATATTGGCAC	x	7.71	2.42	0.49
Ec	g4205a, E104K, M182T, G238S	ATATTTTACCGTCTTAAA		167.29	2.35	0.48
Ec	g4205a, E104K, M182T, G238S	GTAGTAGGTGTCTCAGAC		203.43	2.58	0.48
Ec	g4205a, E104K, M182T, G238S	GCCGCGGCGTGTGTGGTT	x	96.01	1.47	0.48
Ec	A42G, E104K, M182T, G238S	GAAAACATACGGCGTGGT		66.33	1.36	0.48
Ec	A42G, E104K, M182T, G238S	AGTCATGGCATTATGAAA		63.07	1.7	0.47
Ec	A42G, E104K, M182T, G238S	ACGAACTTTGCTTTCTTT	x	107.15	1.45	0.5
Ec	g4205a, A42G, E104K, M182T, G238S	TGAAGCACGTAAACTATC		528.76	1.56	0.48
Ec	g4205a, A42G, E104K, M182T, G238S	TTGGTTGCACCAGACATT		2012.02	0.72	0.48
Ec	g4205a, A42G, E104K, M182T, G238S	CTGTTTTTGAACTTGAAG	x	205.94	1.54	0.48
Kp	Wild-type TEM-1	GATGGCCTTTTGCCGGTT		0.04	5.44	0.56
Kp	Wild-type TEM-1	ATTCGTAAACTTCTGGTT		0.04	7.15	0.56
Kp	Wild-type TEM-1	GAGTCGTGCATTGAGTTC	x	0.04	6.54	0.56
Kp	g4205a	TTGAGTCCCGAGGGTGGT		0.04	7.06	0.55
Kp	g4205a	TACTTTTAAACTATAAGA		0.04	27.18	0.52
Kp	g4205a	GTATTAGTTTTACTTTAG	x	0.04	6.75	0.52
Kp	A42G	GTAGCCATTTACTTCTGT		0.04	7.86	0.51
Kp	A42G	TTCGCATTATATTCCGTC		0.04	6.24	0.53
Kp	A42G	CAGTGAGTCAGATATCTT	x	0.04	8.07	0.54
Kp	E104K	TTGGCGCCTTCTCTTCTG		0.08	5.91	0.55
Kp	E104K	GGAGCAGTAGAGGTGGTA		0.07	5.87	0.57
Kp	E104K	ATAAAGGATGTGACTGTA	x	0.08	5.44	0.54
Kp	M182T	CCTAGCAGCTCGTAAGAG		0.04	7.44	0.55
Kp	M182T	TGTGATTTACACACGTCC		0.04	7.16	0.56
Kp	M182T	TCCAAAGGGTGGCACGAG	x	0.03	7.41	0.51

Kp	G238S	ACTGCGTTTTAATATTTT		0.4	4.5	0.55
Kp	G238S	CCAACTGTAAGCCTATTT		0.41	4.66	0.55
Kp	G238S	CAGGGTCATACGAGCTTC	x	0.56	3.55	0.62
Kp	g4205a, A42G	CTCATTTTAGACTTCGTT		0.05	5.65	0.54
Kp	g4205a, A42G	CATTCATATTAAAGTTTG		0.05	7.4	0.53
Kp	g4205a, A42G	CAGGTTTTAGCATATGCC	x	0.05	6.49	0.55
Kp	g4205a, E104K	CGACAACGTATCAAGCTC		0.11	5.95	0.54
Kp	g4205a, E104K	TCCGACCATTAAGGGTTA		0.11	6.13	0.55
Kp	g4205a, E104K	GCTGGTCCGATCAGATAT	x	0.11	5.81	0.54
Kp	g4205a, M182T	AATGTGCGTTAATAGATT		0.04	11.85	0.52
Kp	g4205a, M182T	TATCATAGTGAGTTCCAT		0.04	6.88	0.54
Kp	g4205a, M182T	AGGGATCTGGAGTAGGTC	x	0.05	5.81	0.64
Kp	g4205a, G238S	ATAAGGTTTGTTTCCCTG		0.38	3.68	0.48
Kp	g4205a, G238S	TAGGTTAATTCTCGGTGA		0.38	4.2	0.48
Kp	g4205a, G238S	AATTTAAGTATAGAGGGG	x	0.38	4.76	0.5
Kp	A42G, E104K	GTTTCCTCTAAAGATTTC		0.24	4.01	0.68
Kp	A42G, E104K	TTTTTCCTCCGCTCTGGT		0.21	6.22	0.58
Kp	A42G, E104K	ATTAGATTTATTAATATG	x	0.14	5.18	0.53
Kp	A42G, M182T	CGTACCCCTTGCTGGTGG		0.04	6.82	0.57
Kp	A42G, M182T	TTGGGACCTCTTTGGGTA		0.04	6.13	0.54
Kp	A42G, M182T	AGAACTTGGTAACGGGGC	x	0.04	8.1	0.53
Kp	A42G, G238S	ATAGTTATTTTGGAACTA		1.79	3.21	0.55
Kp	A42G, G238S	ACTACAGTAATAGTGCAT		2.18	2.14	0.54
Kp	A42G, G238S	CAGCTGGTTGGTTCTCTA	x	1.95	5.41	0.54
Kp	E104K, M182T	GCTTCCTTTATTTGTTTA		0.08	4.26	0.56
Kp	E104K, M182T	GTGTGAGACGCAGTTTAG		0.07	6.52	0.53
Kp	E104K, M182T	GCGAGATGATTAGAGAGA	x	0.06	6.44	0.54
Kp	E104K, G238S	TGTATTGGTTAACGTTAC		2.93	3.18	0.54
Kp	E104K, G238S	TAACGCGAGTCGTAATCT		2.72	4.06	0.55
Kp	E104K, G238S	GAGGGTGTGATTAGCAAT	x	3.15	3.72	0.52
Kp	M182T, G238S	GATCCGATGATAGTAGTT		1.17	4.21	0.55
Kp	M182T, G238S	TATCCGTCCTCGGCAGAG		1.23	4.18	0.56
Kp	M182T, G238S	TCCCTAGCATGGATTGGC	x	1.31	4.56	0.55
Kp	g4205a, A42G, E104K	ATAAATATGTGGTCCCTG		0.28	5.68	0.53
Kp	g4205a, A42G, E104K	AATCATTCAAATCGAAGA		0.27	9.95	0.53
Kp	g4205a, A42G, E104K	CACCATCCTACAACTAAA	x	0.31	5.19	0.54
Kp	g4205a, A42G, M182T	TTAACACATGGTATTTAC		0.05	7.55	0.54
Kp	g4205a, A42G, M182T	AAAGTTTGAGGAATAACG		0.04	7.42	0.54

Kp	g4205a, A42G, M182T	TTGTTTATTTACCGGACT	x	0.06	2.98	0.74
Kp	g4205a, A42G, G238S	ACAGGAGACGGTATCTTT		2.95	3.71	0.55
Kp	g4205a, A42G, G238S	TTGAAATGCTTTCGGTTA		2.84	4.33	0.55
Kp	g4205a, A42G, G238S	AAGTGTTCGCATTGCAAG	x	2.77	6.3	0.51
Kp	g4205a, E104K, M182T	GCGTCCTCGAGTCTTTAC		0.11	5.9	0.54
Kp	g4205a, E104K, M182T	TTTGGACCACTTTTCTGT		0.1	5.58	0.53
Kp	g4205a, E104K, M182T	AGTCGTGTGGGGGCCTAC	x	0.12	6.29	0.54
Kp	g4205a, E104K, G238S	AACGGTTGGACCGAGCGG		3.44	4.11	0.52
Kp	g4205a, E104K, G238S	TTGCGGAAGTGGTCGTGG		3.4	4.48	0.58
Kp	g4205a, E104K, G238S	CCATTTGATTTTAAGCTC	x	2.33	4.63	0.49
Kp	g4205a, M182T, G238S	TTTGTTGCTCTTTCGATG		2.01	3.27	0.54
Kp	g4205a, M182T, G238S	AACCTAATTCTTAACGGA		2.05	3.46	0.55
Kp	g4205a, M182T, G238S	TCCGGTACGATTACAACG	x	2.41	3.51	0.55
Kp	A42G, E104K, M182T	TTGCATAGTTTCATAATA		0.15	5.46	0.55
Kp	A42G, E104K, M182T	GCAGTATGCGGAAAAGCT		0.16	5.34	0.53
Kp	A42G, E104K, M182T	GTGCGAAAGCATTACACT	x	0.16	4.34	0.5
Kp	A42G, E104K, G238S	TCCACGTACATAAATGTT		11.03	3.49	0.55
Kp	A42G, E104K, G238S	ACTGCTTAGCAGTTTGTC		11.42	3.33	0.54
Kp	A42G, E104K, G238S	AGCATGCTCTGCCGAGAA	x	11.9	3.27	0.55
Kp	A42G, M182T, G238S	GGCAGGTGAATCTACCAG		2.32	4.14	0.56
Kp	A42G, M182T, G238S	TTTAGCCCCATCACTAAC		2.37	3.25	0.54
Kp	A42G, M182T, G238S	TTTGTGTAGGTACTATCC	x	2.1	3.07	0.53
Kp	E104K, M182T, G238S	CAATATACTCTGTACTAA		13	3.02	0.55
Kp	E104K, M182T, G238S	TAGTGCTTGTTCAGGGGT		12.1	3.26	0.57

Kp	E104K, M182T, G238S	CATCTAATTTATTGGGTA	x	13.44	3.03	0.5
Kp	g4205a, A42G, E104K, M182T	CATAGTCACCGGCTAGAT		0.26	5.53	0.55
Kp	g4205a, A42G, E104K, M182T	GAGAGGCTAAGGTGAAAC		0.25	4.93	0.53
Kp	g4205a, A42G, E104K, M182T	GCATCGCTCTCATGGGTA	x	0.24	5.31	0.53
Kp	g4205a, A42G, E104K, G238S	TATTTGAGTTATTAGTTC		17.35	2.8	0.53
Kp	g4205a, A42G, E104K, G238S	AGTGCAGGTTTAAATACT		15.66	3.22	0.54
Kp	g4205a, A42G, E104K, G238S	ACGTAAAGTAAGACTTCA	x	15.98	3.21	0.52
Kp	g4205a, A42G, M182T, G238S	TTAGCTTGGTTTCTGTCT		4.18	3.38	0.54
Kp	g4205a, A42G, M182T, G238S	GATAGGGTATATTGGCAC		3.33	4.53	0.54
Kp	g4205a, A42G, M182T, G238S	TAGCATCGGGTCAGGGCG	x	3.39	4.14	0.49
Kp	g4205a, E104K, M182T, G238S	ATATTTTACCGTCTTAAA		128.31	1.21	0.52
Kp	g4205a, E104K, M182T, G238S	GCCGCGGCGTGTGTGGTT		93.22	0.73	0.56
Kp	g4205a, E104K, M182T, G238S	GTAGTAGGTGTCTCAGAC	x	211.57	2.62	0.52
Kp	A42G, E104K, M182T, G238S	ACGAACTTTGCTTTCTTT		16.06	2.85	0.53
Kp	A42G, E104K, M182T, G238S	GAAAACATACGGCGTGGT		15.91	3.08	0.54
Kp	A42G, E104K, M182T, G238S	AGTCATGGCATTATGAAA	x	22.86	2.82	0.53
Kp	g4205a, A42G, E104K, M182T, G238S	TGAAGCACGTAAACTATC		220.64	2.1	0.53
Kp	g4205a, A42G, E104K, M182T, G238S	TTGGTTGCACCAGACATT		159.87	6.58	0.54
Kp	g4205a, A42G, E104K, M182T, G238S	CTGTTTTTGAACTTGAAG	x	13503.68	0.42	0.55
Se	Wild-type TEM-1	ATTCGTAAACTTCTGGTT		0.09	5.47	0.49
Se	Wild-type TEM-1	GAGTCGTGCATTGAGTTC		0.09	6.25	0.47
Se	Wild-type TEM-1	GATGGCCTTTTGCCGGTT	x	0.1	4.79	0.48
Se	g4205a	GTATTAGTTTTACTTTAG		0.12	4.03	0.49
Se	g4205a	TTGAGTCCCGAGGGTGGT		0.12	3.83	0.49
Se	g4205a	TACTTTTAAACTATAAGA	x	0.12	4.02	0.47

Se	A42G	CAGTGAGTCAGATATCTT		0.12	4.11	0.49
Se	A42G	GTAGCCATTTACTTCTGT		0.11	4.62	0.49
Se	A42G	TTCGCATTATATTCCGTC	x	0.1	6.09	0.46
Se	E104K	TTGGCGCCTTCTCTTCTG		0.34	3.49	0.49
Se	E104K	GGAGCAGTAGAGGTGGTA		0.3	3.86	0.49
Se	E104K	ATAAAGGATGTGACTGTA	x	0.32	4.82	0.46
Se	M182T	TCCAAAGGGTGGCACGAG		0.09	5.47	0.49
Se	M182T	TGTGATTTACACACGTCC		0.1	5.14	0.49
Se	M182T	CCTAGCAGCTCGTAAGAG	x	0.1	5.02	0.48
Se	G238S	CCAACTGTAAGCCTATTT		1.77	4.84	0.45
Se	G238S	CAGGGTCATACGAGCTTC		1.87	5.62	0.46
Se	G238S	ACTGCGTTTTAATATTTT	x	1.62	3.74	0.55
Se	g4205a, A42G	CATTCATATTAAAGTTTG		0.16	3.56	0.47
Se	g4205a, A42G	CAGGTTTTAGCATATGCC		0.16	3.01	0.52
Se	g4205a, A42G	CTCATTTTAGACTTCGTT	x	0.16	2.8	0.47
Se	g4205a, E104K	CGACAACGTATCAAGCTC		0.52	4.25	0.48
Se	g4205a, E104K	TCCGACCATTAAGGGTTA		0.5	4.3	0.48
Se	g4205a, E104K	GCTGGTCCGATCAGATAT	x	0.47	3.44	0.47
Se	g4205a, M182T	AGGGATCTGGAGTAGGTC		0.11	4.92	0.45
Se	g4205a, M182T	AATGTGCGTTAATAGATT		0.11	5.29	0.48
Se	g4205a, M182T	TATCATAGTGAGTTCCAT	x	0.13	3.19	0.5
Se	g4205a, G238S	ATAAGGTTTGTTTCCCTG		1.64	4.28	0.41
Se	g4205a, G238S	AATTTAAGTATAGAGGGG		1.73	5.51	0.4
Se	g4205a, G238S	TAGGTTAATTCTCGGTGA	x	1.63	3.84	0.4
Se	A42G, E104K	GTTTCCTCTAAAGATTTC		0.81	3.8	0.48
Se	A42G, E104K	TTTTTCCTCCGCTCTGGT		0.88	3.65	0.48
Se	A42G, E104K	ATTAGATTTATTAATATG	x	0.68	3.97	0.47
Se	A42G, M182T	AGAACTTGGTAACGGGGC		0.1	5.39	0.49
Se	A42G, M182T	TTGGGACCTCTTTGGGTA		0.1	5.3	0.5
Se	A42G, M182T	CGTACCCCTTGCTGGTGG	x	0.11	3.54	0.48
Se	A42G, G238S	CAGCTGGTTGGTTCTCTA		13.9	1.69	0.48
Se	A42G, G238S	ATAGTTATTTTGGAACTA		12.55	1.88	0.51
Se	A42G, G238S	ACTACAGTAATAGTGCAT	x	14.91	1.82	0.53
Se	E104K, M182T	GCGAGATGATTAGAGAGA		0.26	3.85	0.48
Se	E104K, M182T	GTGTGAGACGCAGTTTAG		0.36	3.12	0.49
Se	E104K, M182T	GCTTCCTTTATTTGTTTA	x	0.29	4.03	0.56
Se	E104K, G238S	TAACGCGAGTCGTAATCT		14.37	1.69	0.49
Se	E104K, G238S	GAGGGTGTGATTAGCAAT		14.11	1.7	0.48
Se	E104K, G238S	TGTATTGGTTAACGTTAC	x	15.92	1.63	0.54
Se	M182T, G238S	TATCCGTCCTCGGCAGAG		7.25	1.98	0.49
Se	M182T, G238S	TCCCTAGCATGGATTGGC		8.39	1.92	0.49
Se	M182T, G238S	GATCCGATGATAGTAGTT	x	7.93	1.58	0.51

Se	g4205a, A42G, E104K	ATAAATATGTGGTCCCTG		1.28	3.24	0.49
Se	g4205a, A42G, E104K	CACCATCCTACAACTAAA		1.45	3.07	0.48
Se	g4205a, A42G, E104K	AATCATTCAAATCGAAGA	x	1.37	3.31	0.45
Se	g4205a, A42G, M182T	TTAACACATGGTATTTAC		0.13	3.77	0.46
Se	g4205a, A42G, M182T	AAAGTTTGAGGAATAACG		0.12	4.31	0.47
Se	g4205a, A42G, M182T	TTGTTTATTTACCGGACT	x	0.16	3.22	0.47
Se	g4205a, A42G, G238S	ACAGGAGACGGTATCTTT		18.25	1.96	0.48
Se	g4205a, A42G, G238S	TTGAAATGCTTTCGGTTA		16.28	2.18	0.47
Se	g4205a, A42G, G238S	AAGTGTTCGCATTGCAAG	x	16.88	2.05	0.46
Se	g4205a, E104K, M182T	GCGTCCTCGAGTCTTTAC		0.5	4.3	0.47
Se	g4205a, E104K, M182T	TTTGGACCACTTTTCTGT		0.52	3.85	0.47
Se	g4205a, E104K, M182T	AGTCGTGTGGGGGCCTAC	x	0.59	4.19	0.49
Se	g4205a, E104K, G238S	AACGGTTGGACCGAGCGG		12.12	2.1	0.42
Se	g4205a, E104K, G238S	TTGCGGAAGTGGTCGTGG		12.44	2.21	0.42
Se	g4205a, E104K, G238S	CCATTTGATTTTAAGCTC	x	12.71	2.35	0.41
Se	g4205a, M182T, G238S	TTTGTTGCTCTTTCGATG		13	1.83	0.48
Se	g4205a, M182T, G238S	AACCTAATTCTTAACGGA		11.92	2.25	0.47
Se	g4205a, M182T, G238S	TCCGGTACGATTACAACG	x	14.89	1.91	0.47
Se	A42G, E104K, M182T	TTGCATAGTTTCATAATA		0.74	3.42	0.47
Se	A42G, E104K, M182T	GCAGTATGCGGAAAAGCT		0.75	3.31	0.48
Se	A42G, E104K, M182T	GTGCGAAAGCATTACACT	x	0.73	3.7	0.54
Se	A42G, E104K, G238S	AGCATGCTCTGCCGAGAA		77.61	2.26	0.47
Se	A42G, E104K, G238S	ACTGCTTAGCAGTTTGTC		75.2	2.11	0.48
Se	A42G, E104K, G238S	TCCACGTACATAAATGTT	x	67.54	2.18	0.49

Se	A42G, M182T, G238S	GGCAGGTGAATCTACCAG		17	1.74	0.46
Se	A42G, M182T, G238S	TTTAGCCCCATCACTAAC		16.99	2.03	0.47
Se	A42G, M182T, G238S	TTTGTGTAGGTACTATCC	x	13.99	1.96	0.47
Se	E104K, M182T, G238S	CAATATACTCTGTACTAA		81.03	2.19	0.48
Se	E104K, M182T, G238S	TAGTGCTTGTTCAGGGGT		80.42	2	0.48
Se	E104K, M182T, G238S	CATCTAATTTATTGGGTA	x	84.48	2.06	0.47
Se	g4205a, A42G, E104K, M182T	GCATCGCTCTCATGGGTA		1.27	3.43	0.48
Se	g4205a, A42G, E104K, M182T	GAGAGGCTAAGGTGAAAC		1.13	3.33	0.48
Se	g4205a, A42G, E104K, M182T	CATAGTCACCGGCTAGAT	x	1.42	2.79	0.47
Se	g4205a, A42G, E104K, G238S	ACGTAAAGTAAGACTTCA		86.27	2.31	0.46
Se	g4205a, A42G, E104K, G238S	AGTGCAGGTTTAAATACT		82.6	2.47	0.45
Se	g4205a, A42G, E104K, G238S	TATTTGAGTTATTAGTTC	x	99.11	1.82	0.45
Se	g4205a, A42G, M182T, G238S	TTAGCTTGGTTTCTGTCT		26.49	2.16	0.47
Se	g4205a, A42G, M182T, G238S	GATAGGGTATATTGGCAC		23.65	1.91	0.46
Se	g4205a, A42G, M182T, G238S	TAGCATCGGGTCAGGGCG	x	27.5	1.66	0.52
Se	g4205a, E104K, M182T, G238S	GCCGCGGCGTGTGTGGTT		149.32	2.01	0.49
Se	g4205a, E104K, M182T, G238S	GTAGTAGGTGTCTCAGAC		167.72	3.91	0.48
Se	g4205a, E104K, M182T, G238S	ATATTTTACCGTCTTAAA	x	164.9	2.96	0.45
Se	A42G, E104K, M182T, G238S	GAAAACATACGGCGTGGT		106.57	2.25	0.48
Se	A42G, E104K, M182T, G238S	AGTCATGGCATTATGAAA		136.69	2.39	0.47
Se	A42G, E104K, M182T, G238S	ACGAACTTTGCTTTCTTT	x	27179.24	0.03	0.83
Se	g4205a, A42G, E104K, M182T, G238S	CTGTTTTTGAACTTGAAG		3009.24	0.81	0.48
Se	g4205a, A42G, E104K, M182T, G238S	TTGGTTGCACCAGACATT		1026	1.52	0.46

Se	g4205a, A42G, E104K, M182T, G238S	TGAAGCACGTAAACTATC	x	1162.37	1.42	0.49

SI Table 19 : Specific datasets and parameters used in the evolutionary simulations. The same population size (1,000 individuals) and mutation rate (5 x 10^5) were used in each treatment.

Relevant figure	Focal gene	Focal host	Transient host	Time steps	Time steps with the transient host
Figure 9d, SI Figure 14a	bla	Ec	*None	60	*None
Figure 9e, SI Figure 14a	bla	Ec	Kp	60	21-40
Figure 9f, SI Figure 14a	bla	Ec	Se	60	21-40
SI Figure 13a, SI Figure 14b	bla	Kp	*None	60	*None
SI Figure 13b, SI Figure 14b	bla	Kp	Ec	60	21-40
SI Figure 13c, SI Figure 14b	bla	Kp	Se	60	21-40
SI Figure 13d, SI Figure 14c	bla	Se	*None	60	*None
SI Figure 13e, SI Figure 14c	bla	Se	Ec	60	21-40
SI Figure 13f, SI Figure 14c	bla	Se	Kp	60	21-40
SI Figure 16a, SI Figure 17a	DHFR	Ec1	*None	36	*None
SI Figure 16b, SI Figure 17a	DHFR	Ec1	Ec2	36	13-24
SI Figure 16c, SI Figure 17a	DHFR	Ec1	Ec3	36	13-24
SI Figure 16d, SI Figure 17b	DHFR	Ec2	*None	36	*None
SI Figure 16e, SI Figure 17b	DHFR	Ec2	Ec1	36	13-24
SI Figure 16f, SI Figure 17b	DHFR	Ec2	Ec3	36	13-24
SI Figure 16g, SI Figure 17c	DHFR	Ec3	*None	36	*None
SI Figure 16h, SI Figure 17c	DHFR	Ec3	Ec1	36	13-24
SI Figure 16i, SI Figure 17c	DHFR	Ec3	Ec2	36	13-24

References

1. C. M. Thomas, K. M. Nielsen, Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3**, 711–721 (2005).

2. F. de la Cruz, J. Davies, Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol.* **8**, 128–133 (2000).

3. A. Norman, L. H. Hansen, S. J. Sørensen, Conjugative plasmids: vessels of the communal gene pool. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **364**, 2275–2289 (2009).

4. P. Mazodier, J. Davies, Gene transfer between distantly related bacteria. *Annu. Rev. Genet.* **25**, 147–171 (1991).

5. S. Redondo-Salvo, *et al.*, Pathways for horizontal gene transfer in bacteria revealed by a global map of their plasmids. *Nat. Commun.* **11**, 3602 (2020).

6. A. K. Olesen, *et al.*, IncHI1A plasmids potentially facilitate horizontal flow of antibiotic resistance genes to pathogens in microbial communities of urban residential sewage. *Mol. Ecol.* **31**, 1595–1608 (2022).

7. H. Jordt, *et al.*, Coevolution of host-plasmid pairs facilitates the emergence of novel multidrug resistance. *Nat Ecol Evol* **4**, 863–869 (2020).

8. J. Davies, Inactivation of antibiotics and the dissemination of resistance genes. *Science* **264**, 375–382 (1994).

9. Antimicrobial Resistance Collaborators, Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629–655 (2022).

10. B. R. Levin, F. M. Stewart, V. A. Rice, The kinetics of conjugative plasmid transmission: fit of a simple mass action model. *Plasmid* **2**, 247–260 (1979).

11. L. Simonsen, D. M. Gordon, F. M. Stewart, B. R. Levin, Estimating the rate of plasmid transfer: an end-point method. *J. Gen. Microbiol.* **136**, 2319–2325 (1990).

12. A. J. Lopatkin, *et al.*, Antibiotics as a selective driver for conjugation dynamics. *Nat Microbiol* **1**, 16044 (2016).

13. J. H. Bethke, *et al.*, Environmental and genetic determinants of plasmid mobility in pathogenic Escherichia coli. *Sci Adv* **6**, eaax3173 (2020).

14. J. S. Huisman, *et al.*, Estimating plasmid conjugation rates: A new computational tool and a critical comparison of methods. *Plasmid* **121**, 102627 (2022).

15. T. Dimitriu, L. Marchant, A. Buckling, B. Raymond, Bacteria from natural populations transfer plasmids mostly towards their kin. *Proc. Biol. Sci.* **286**, 20191110 (2019).

16. A. San Millan, Evolution of Plasmid-Mediated Antibiotic Resistance in the Clinical Context. *Trends Microbiol.* **26**, 978–985 (2018).

17. L. Li, *et al.*, Plasmids persist in a microbial community by providing fitness benefit to multiple phylotypes. *ISME J.* **14**, 1170–1181 (2020).

18. J. P. J. Hall, E. Harrison, D. A. Baltrus, Introduction: the secret lives of microbial mobile genetic elements. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **377**, 20200460 (2022).

19. S. E. Luria, M. Delbrück, Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **28**, 491–511 (1943).

20. A. R. Johnsen, N. Kroer, Effects of stress and other environmental factors on horizontal plasmid transfer assessed by direct quantification of discrete transfer events. *FEMS Microbiol. Ecol.* **59**, 718–728 (2007).

21. N. Fedoroff, W. Fontana, Genetic networks. Small numbers of big molecules. *Science* **297**, 1129–1131 (2002).

22. W. Loftie-Eaton, *et al.*, Compensatory mutations improve general permissiveness to antibiotic resistance plasmids. *Nat Ecol Evol* **1**, 1354–1363 (2017).

23. J. B. Alderliesten, *et al.*, Effect of donor-recipient relatedness on the plasmid conjugation frequency: a meta-analysis. *BMC Microbiol.* **20**, 135 (2020).

24. R. J. Sheppard, A. E. Beddis, T. G. Barraclough, The role of hosts, plasmids and environment in determining plasmid transfer rates: A meta-analysis. *Plasmid* **108**, 102489 (2020).

25. C. Dahlberg, L. Chao, Amelioration of the cost of conjugative plasmid carriage in Eschericha coli K12. *Genetics* **165**, 1641–1649 (2003).

26. W. Loftie-Eaton, *et al.*, Contagious Antibiotic Resistance: Plasmid Transfer among Bacterial Residents of the Zebrafish Gut. *Appl. Environ. Microbiol.* **87** (2021).

27. P. D. Lundquist, B. R. Levin, Transitory derepression and the maintenance of conjugative plasmids. *Genetics* **113**, 483–497 (1986).

28. S. Bates, A. M. Cashmore, B. M. Wilkins, IncP plasmids are unusually effective in mediating conjugation of Escherichia coli and Saccharomyces cerevisiae: involvement of the tra2 mating system. *J. Bacteriol.* **180**, 6538–6543 (1998).

29. A. E. Dewar, *et al.*, Plasmids do not consistently stabilize cooperation across bacteria but may promote broad pathogen host-range. *Nat Ecol Evol* **5**, 1624–1636 (2021).

30. K. R. Philipsen, L. E. Christiansen, H. Hasman, H. Madsen, Modelling conjugation with stochastic differential equations. *J. Theor. Biol.* **263**, 134–142 (2010).

31. E. Smit, J. D. van Elsas, Determination of plasmid transfer frequency in soil: Consequences of bacterial mating on selective agar media. *Curr. Microbiol.* **21**, 151–157 (1990).

32. De Gelder Leen, Vandecasteele Frederik P. J., Brown Celeste J., Forney Larry J., Top Eva M., Plasmid Donor Affects Host Range of Promiscuous IncP-1β Plasmid pB10 in an Activated-Sludge Microbial Community. *Appl. Environ. Microbiol.* **71**, 5309–5317 (2005).

33. A. Kottara, J. P. J. Hall, M. A. Brockhurst, The proficiency of the original host species determines community-level plasmid dynamics. *FEMS Microbiol. Ecol.* **97** (2021).

34. A. Kottara, L. Carrilero, E. Harrison, J. P. J. Hall, M. A. Brockhurst, The dilution effect limits plasmid horizontal transmission in multispecies bacterial communities. *Microbiology* **167** (2021).

35. J. D. van Elsas, M. J. Bailey, The ecology of transfer of mobile genetic elements. *FEMS Microbiol. Ecol.* **42**, 187–197 (2002).

36. W. Loftie-Eaton, *et al.*, Evolutionary Paths That Expand Plasmid Host-Range: Implications for Spread of Antibiotic Resistance. *Mol. Biol. Evol.* **33**, 885–897 (2016).

37. R. J. F. Haft, *et al.*, General mutagenesis of F plasmid TraI reveals its role in conjugative regulation. *J. Bacteriol.* **188**, 6346–6353 (2006).

38. H. L. Jordt, "Of E. coli and classrooms: Stories of persistence." (2019).

39. H. K. Alexander, R. C. MacLean, Stochastic bacterial population dynamics restrict the establishment of antibiotic resistance from single cells. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 19455–19464 (2020).

40. J. H. Abel, B. Drawert, A. Hellander, L. R. Petzold, GillesPy: A Python Package for Stochastic Model Building and Simulation. *IEEE Life Sci Lett* **2**, 35–38 (2016).

41. X. Zhong, J. Droesch, R. Fox, E. M. Top, S. M. Krone, On the meaning and estimation of plasmid transfer rates for surface-associated and well-mixed bacterial populations. *J. Theor. Biol.* **294**, 144–152 (2012).

42. T. A. Sysoeva, Y. Kim, J. Rodriguez, A. J. Lopatkin, L. You, Growth-stage-dependent regulation of conjugation. *AIChE J.* **66** (2020).

43. P. Keller, T. Antal, Mutant number distribution in an exponentially growing population. *J. Stat. Mech.* **2015**, P01011 (2015).

44. D. M. Weinreich, N. F. Delaney, M. A. Depristo, D. L. Hartl, Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science* **312**, 111–114 (2006).

45. H. A. Lindsey, J. Gallie, S. Taylor, B. Kerr, Evolutionary rescue from extinction is contingent on a lower rate of environmental change. *Nature* **494**, 463–467 (2013).

46. L. Tan, S. Serene, H. X. Chao, J. Gore, Hidden randomness between fitness landscapes limits reverse evolution. *Phys. Rev. Lett.* **106**, 198102 (2011).

47. C. P. Goulart, *et al.*, Designing antibiotic cycling strategies by determining and understanding local adaptive landscapes. *PLoS One* **8**, e56040 (2013).

48. R. F. Guerrero, S. V. Scarpino, J. V. Rodrigues, D. L. Hartl, C. B. Ogbunugafor, Proteostasis Environment Shapes Higher-Order Epistasis Operating on Antibiotic Resistance. *Genetics* **212**, 565–575 (2019).

49. C. B. Ogbunugafor, M. J. Eppstein, Genetic Background Modifies the Topography of a Fitness Landscape, Influencing the Dynamics of Adaptive Evolution. *IEEE Access* **7**, 113675–113683 (2019).

50. C. B. Ogbunugafor, C. S. Wylie, I. Diakite, D. M. Weinreich, D. L. Hartl, Adaptive Landscape by Environment Interactions Dictate Evolutionary Dynamics in Models of Drug Resistance. *PLoS Comput. Biol.* **12**, e1004710 (2016).

51. S. G. Das, S. O. Direito, B. Waclaw, R. J. Allen, J. Krug, Predictable properties of fitness landscapes induced by adaptational tradeoffs. *Elife* **9** (2020).

52. L. De Gelder, J. M. Ponciano, P. Joyce, E. M. Top, Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* **153**, 452–463 (2007).

53. O. Kosterlitz, *et al.*, Estimating the transfer rates of bacterial plasmids with an adapted Luria-Delbrück fluctuation analysis. *PLoS Biol.* **20**, e3001732 (2022).

54. R. Jain, M. C. Rivera, J. A. Lake, Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 3801–3806 (1999).

55. A. Novick, W. F. Doolittle, Horizontal persistence and the complexity hypothesis. *Biol. Philos.* **35** (2020).

56. H. Jeong, *et al.*, Genome sequences of Escherichia coli B strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 644–652 (2009).

57. M. McClelland, *et al.*, Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. *Nature* **413**, 852–856 (2001).

58. I. Wiegand, K. Hilpert, R. E. W. Hancock, Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175 (2008).

59. L. Zhao, Z. Liu, S. F. Levy, S. Wu, Bartender: a fast and accurate clustering algorithm to count barcode reads. *Bioinformatics* **34**, 739–747 (2018).

60. C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, Dose-Response Analysis Using R. *PLoS One* **10**, e0146021 (2015).

61. M. Barlow, B. G. Hall, Predicting evolutionary potential: in vitro evolution accurately reproduces natural evolution of the tem beta-lactamase. *Genetics* **160**, 823–832 (2002).