Understanding the evolution of multiple drug resistance in structured populations

David V. McLeod*[1] and Sylvain Gandon†[1]

[1]Centre D'Ecologie Fonctionnelle & Evolutive, CNRS, Montpellier, France

Abstract

The evolution of multidrug resistance (MDR) is a pressing public health concern. Yet many aspects, such as the role played by population structure, remain poorly understood. Here we argue that studying MDR evolution by focusing upon the dynamical equations for linkage disequilibrium (LD) can greatly simplify the calculations, generate more insight, and provide a unified framework for understanding the role of population structure. We demonstrate how a general epidemiological model of MDR evolution can be recast in terms of the LD equations. These equations reveal how the different forces generating and propagating LD operate in a dynamical setting at both the metapopulation and population level. We then apply these insights to show how the LD perspective: (i) provides a simple interpretative framework for transient evolutionary dynamics, (ii) explains equilibrium patterns of MDR, and (iii) can be used to assess the MDR consequences of different drug prescription strategies.

Keywords: antibiotic resistance, multidrug resistance, linkage disequilibrium, evolutionary epidemiology

*david.mcleod@cefe.cnrs.fr
†sylvain.gandon@cefe.cnrs.fr

Introduction

Antibiotic resistance is one of the biggest current public health problems, with antibiotic resistant infections responsible for tens of thousands of deaths annually [1]. Of particular concern is the evolution of *multidrug resistant* (MDR) pathogens, that is, pathogens resistant to multiple classes of antibiotics. Despite its importance, understanding the evolution of MDR remains an ongoing challenge, as it is typically not captured by our understanding of the evolution of single drug resistance [for which there is a large body of theory; e.g., 2–6]. For instance, suppose we have two drugs, A and B, and that a fraction p_{AB} of infections caused by the pathogen of interest are resistant to both drugs. To understand MDR evolution, we need to understand what determines the frequency p_{AB}. If p_A and p_B are the frequency of infections resistant to drug A and B, and D denotes any non-random association between resistance to drugs A and B, then

$$p_{AB} = p_A p_B + D. \tag{1}$$

If $D = 0$, then the evolution of resistance to each drug is independent, and so multiple drugs will not qualitatively alter the evolutionary dynamics of single drug resistance. However, whenever $D \neq 0$, understanding the fitness costs and benefits of resistance to each drug in isolation is insufficient to understand the evolution of MDR, because doing so will not tell us what factors govern the propagation of D. Thus the challenge of understanding MDR evolution can be recast as understanding the dynamics of D. As it turns out, the quantity D is referred to as *linkage disequilibrium* (LD), and it has been extensively studied in population genetics [e.g., 7–12], particularly as it relates to population structure [13–16]. However, there has been little attempt to apply these insights to MDR evolution; often the dynamics of doubly-resistant infections are neglected to simplify the analysis of the dynamics of single drug resistance [e.g., 3, 5, 17].

Here we consider a simple epidemiological model of a primarily asymptomatically carried pathogen (e.g., *Staphylococcus* spp. or *Enterococcus* spp.) in a structured host population. We show how this model relates to general dynamical equations for LD [18], in turn revealing the role of population structure in MDR evolution. We then use these equations to show how analyzing problems from the LD perspective: (i) provides a straightforward framework for understanding transient evolutionary dynamics, which we use to explain patterns of MDR in *Streptococcus pneumoniae*; (ii) reveals the evolutionary logic underlying patterns of MDR at equilibrium, which we use to reexamine a recent paper on MDR evolution [19]; and (iii) provides insight on the consequences different drug prescription strategies have on MDR, which we apply to a hospital-community setting.

Model

In what follows we will introduce and analyze a model of MDR evolution. We will highlight the most important aspects here while providing more extensive details in the Supplementary Material.

Consider an asymptomatically carried pathogen in a metapopulation consisting of N host populations. Focus upon population X. Let S^X and I^X_{ij} denote the density of susceptible hosts and ij-infections, respectively, at time t, where i indicates if the infection is resistant ($i = A$) or not ($i = a$) to drug A and j indicates if the infection is resistant ($j = B$) or not ($j = b$) to drug B. Susceptible hosts contract ij-infections at a per-capita rate $\beta^X_{ij} I^X_{ij}$, where β^X_{ij} is a rate constant,

while ij-infections are naturally cleared at a per-capita rate α_{ij}^X. Hosts are treated with drugs A, B, or both in combination, at per-capita rates τ_A^X, τ_B^X, and τ_{AB}^X, respectively. Hosts move from population X to X' at a per-capita rate $m^{X\to X'}$. Transmission between infected hosts leads to superinfection with probability σ, in which either strain is equally likely to be the victor. Finally, individual infections acquire allele ℓ through either mutation or recombination (during superinfection) at per-capita rates μ_ℓ^X and ρ_ℓ^X, respectively (note that ρ_ℓ^X depends upon infection densities).

From these epidemiological assumptions the change in ij-infections in population X can be written as the sum of four processes

$$\frac{\mathrm{d}I_{ij}^X}{\mathrm{d}t} = \overbrace{(r^X + \mathbf{1}_A s_A^X + \mathbf{1}_B s_B^X + \mathbf{1}_{AB} s_{AB}^X)}^{\text{per-capita growth}} I_{ij}^X + \overbrace{\Delta\mu_{ij}^X}^{\text{mutation}} + \overbrace{\Delta\rho_{ij}^X}^{\text{recombination}} + \overbrace{\sum_{k=1}^{N}\left(m^{k\to X} I_{ij}^k - m^{X\to k} I_{ij}^X\right)}^{\text{migration}}, \quad (2)$$

where $\mathbf{1}_\ell$ is equal to 1 if the ij-infection has allele(s) ℓ and 0 otherwise and $\Delta\mu_{ij}^X$ and $\Delta\rho_{ij}^X$ denote the net change in ij-infections due to mutation and recombination (Fig. 1). To facilitate comparison with previous results, we have broken the per-capita growth term into four components: the 'baseline' per-capita growth rate, r^X, the (additive) selection coefficients for resistance to drugs A and B, s_A^X and s_B^X, and any epistatic interactions, s_{AB}^X. These latter terms have the standard interpretation. If $s_A^X > 0$ (resp. $s_B^X > 0$), then resistance to drug A (resp. B) is selected for. If $s_{AB}^X > 0$, there is positive epistasis, and the per-capita growth rate of doubly-resistant infections is greater than would be expected by consideration of the per-capita growth rate of singly-resistant infections. Thus although equation (2) is derived from a specific model, the partitioning is very general and applies to many epidemiological scenarios.

While system (2) contains all of the information necessary to analyze MDR evolution, as currently written it is particularly opaque for providing insight. Therefore we would like to transform it to a form which brings to the forefront the different factors that promote or impede MDR evolution; the way to do this is by focusing upon the dynamical equations for linkage disequilibrium (LD). However, the inclusion of multiple populations means that doing so is not as simple as equation (1) would suggest since there are different scales at which LD and MDR can be measured. As the scale which is of most interest will depend upon the specifics of the problem, in what follows we will consider MDR evolution at both the population- and metapopulation-level.

Population-level multidrug resistance

To understand MDR evolution in a given population, say X, we need to understand the dynamics of the frequency of infections resistant to drug A and B, p_A^X and p_B^X, and the dynamics of population LD, D^X. First, consider the dynamics of p_A^X (mutatis mutandis p_B^X). Using equation

(2), it is straightforward to compute

$$\frac{\mathrm{d}p_A^X}{\mathrm{d}t} = \overbrace{s_A^X p_A^X(1-p_A^X)}^{\text{direct selection}} + \overbrace{s_B^X D^X}^{\text{indirect selection}} + \overbrace{s_{AB}^X p_A^X(1-p_A^X)\frac{p_{AB}^X}{p_A^X}}^{\text{epistasis}}$$
$$+ \underbrace{(\mu_A^X+\rho_A^X)(1-p_A^X)-(\mu_a^X+\rho_a^X)p_A^X}_{\text{mutation and recombination}} - \underbrace{\sum_{k=1}^{N} m^{k\to X}\frac{I^k}{I^X}(p_A^X-p_A^k)}_{\text{migration}}. \quad (3)$$

where I^X is the total density of infections in population X. A related formulation to equation (3) can be found in [18] [see also 11].

Equation (3) is partitioned into recognizable quantities. First, if resistance to drug A is selectively advantageous, $s_A^X > 0$, then drug A resistance will increase due to direct selection whose strength is dictated by the genetic variance at the locus, $p_A^X(1-p_A^X)$ [18]. Second, if doubly-resistant infections are overrepresented in the population, $D^X > 0$, and resistance to drug B is selected for, $s_B^X > 0$, then drug A resistance will increase due to indirect selection upon resistance to drug B. Third, if epistasis is positive, $s_{AB}^X > 0$, and there is genetic variance at the locus, drug A resistance will increase due to the disproportionate growth of doubly-resistant infections. Fourth, mutation and recombination will increase drug A resistance when there is a mutation or recombination bias towards gain of drug A resistance, $\mu_A^X > \mu_a^X$ or $\rho_A^X > \rho_a^X$, and the frequency of infections sensitive to drug A exceeds the frequency of infections resistant to drug A, $1-p_A^X > p_A^X$. Finally, migration acts to reduce differences between populations.

It follows that drug B treatment alters the predicted dynamics of resistance to drug A via two main effects: (i) the influence of epistasis and (ii) indirect selection on resistance to drug B mediated through the presence of LD ($D^X \neq 0$). Thus consider the dynamics of D^X,

$$\frac{\mathrm{d}D^X}{\mathrm{d}t} = \overbrace{(s_A^X - s^X + s_B^X - s^X)D^X}^{\text{selection}} - \overbrace{(\mu^X+\rho^X)D^X}^{\text{mutation and recombination}}$$
$$+ \underbrace{s_{AB}^X p_{AB}^X p_{ab}^X}_{\text{epistasis}} - \underbrace{\sum_{k=1}^{N} m^{k\to X}\frac{I^k}{I^X}\left(D^X - D^k - (p_A^X-p_A^k)(p_B^X-p_B^k)\right)}_{\text{migration}}, \quad (4)$$

where $s^X = p_A^X s_A^X + p_B^X s_B^X + p_{AB}^X s_{AB}^X$ is the average selection for resistance, and μ^X and ρ^X are the total per-capita rates of mutation and recombination, respectively (e.g., $\mu^X = \mu_a^X + \mu_A^X + \mu_b^X + \mu_B^X$).

Equation (4) is partitioned into four key processes. First, excess selection for resistance to drug A (resp. B), $s_A^X - s^X$, can cause pre-existing LD ($D^X \neq 0$) to increase or decrease. For example, if $s_A^X > s^X$ and $D^X > 0$ then LD will increase. This is because drug A resistant infections are fitter than the average resistant infection and so will increase in frequency. Since $D^X > 0$, it is more likely this increase will occur in doubly-resistant infections, thereby increasing D^X. Second, mutation and recombination removes any LD present at a rate proportional to the LD [11, 12]. Third, epistasis generates same-sign LD, that is, positive epistasis, $s_{AB}^X > 0$, leads to MDR overrepresentation, $D^X > 0$ [7, 8, 20, 21]. Positive epistasis could occur if double-resistance costs are less than expected [22–24] or drugs are prescribed in combination [18, 25].

Migration is the final term of equation (4) and reveals how the metapopulation structure affects population LD. Like epistasis, migration does not require pre-existing LD to operate on

LD [9, 13–15, 26]. In particular, LD in population X will be generated whenever the frequencies of resistance to drugs A and B differ between population X and any other connected population, say X'. If both types of resistance are more common in one population than the other, $(p_A^X - p_A^{X'})(p_B^X - p_B^{X'}) > 0$, then migration will generate positive LD in both populations, $D^X > 0$ and $D^{X'} > 0$. If instead drug A resistance is more prevalent in one population, while drug B resistance is more prevalent in the other, migration will generate negative LD in both populations.

Notice the presence of the multiplier I^k/I^X in the final term of equation (4). If the populations have roughly the same density of infections, then this term is unimportant. However, when one population, say X', has much fewer total infections than population X, $I^{X'} \ll I^X$, the term $I^X/I^{X'}$ will be very large whereas $I^{X'}/I^X$ will be very small. Consequently, the ability of migration to propagate LD will be greater in population X' than X, and so all else being equal we would predict the population with a lower density of infections will have a greater magnitude of LD than the population with a higher density of infections.

The next insight shows the importance of also taking into account equation (3). In particular, if we only inspected the migration term of equation (4) we might conclude that as the per-capita migration rate, $m^{k\to X}$, increases, so too will the ability of migration to propagate LD. However, the magnitude of population LD is actually maximized at intermediate migration rates (Fig. 2). The reason is because the quantity $m^{k\to X}$ has two effects. On the one hand, it directly multiplies the migration term in equation (4) thereby magnifying migration's potential role in LD build-up, while on the other hand, it also balances infection frequencies between populations (equation (3)), which in turn will reduce the magnitude of $(p_A^X - p_A^k)(p_B^X - p_B^k)$ in equation (4). These conflicting forces mean the magnitude of population LD tends to be maximized when migration is neither too infrequent nor too frequent (Fig. 2).

Metapopulation-level multidrug resistance

Now what happens to LD and MDR evolution at the metapopulation-level? Metapopulation LD, or total LD, can be defined in terms of the population variables as

$$D_{\text{tot}} \equiv D + \text{cov}(p_A, p_B), \tag{5}$$

that is, D_{tot} is the sum of the average population LD, D, and the spatial covariance between resistance to drugs A and B. As our goal is to understand how population structure shapes the dynamics of p_A, p_B, and D_{tot}, for clarity we will split the terms in the dynamical equations into two groups. In the first group are those terms (or processes) which are always operating, irrespective of population structure, and so can be expressed in terms of the metapopulation-level variables p_ℓ and D_{tot}. The second group consists of those processes which only operate if the populations differ (i.e., there is spatial heterogeneity). It is the latter group which is crucial to understanding how population structure shapes population MDR, and so will be our focus here.

With this in mind, the change in frequency of infections resistant to drug A (mutatis mutan-

dis drug B) can be written

$$\frac{\mathrm{d}p_A}{\mathrm{d}t} = \overbrace{s_A p_A(1-p_A)}^{\substack{\text{direct}\\\text{selection}}} + \overbrace{s_B D_{\text{tot}}}^{\substack{\text{indirect}\\\text{selection}}} + \overbrace{s_{AB} p_A(1-p_A)\frac{p_{AB}}{p_A}}^{\text{epistasis}}$$
$$+ \underbrace{(\mu_A+\rho_A)(1-p_A) - (\mu_a+\rho_a)p_A}_{\text{mutation and recombination}} + \underbrace{\text{cov}(r, p_A)}_{\substack{\textbf{heterogeneity}\text{ in}\\\text{'baseline' growth}}} + \underbrace{p_B \text{cov}\left(s_B, \frac{p_{AB}}{p_B}\right)}_{\substack{\textbf{heterogeneity}\text{ in}\\\text{indirect selection}}}, \quad (6)$$

where s_ℓ, r, μ_ℓ, ρ_ℓ are the average of their respective population quantities. The first four terms in equation (6) are the metapopulation-level analogues of the first four terms in equation (3); since they share the same interpretation, we do not discuss them further here. The last two terms, however, arise due to spatial heterogeneity in 'baseline' growth and selection and so are the consequence of population structure.

First, spatial heterogeneity arises through differences in the 'baseline' per-capita growth (i.e., $r^X \neq r^{X'}$) coupled with differences in the frequencies of drug A resistant infections (e.g., $p_A^X \neq p_A^{X'}$). In particular, more productive (larger r^X) populations will have a disproportionate effect on the change in drug A resistance. For example, if more productive populations also have a greater frequency of drug A resistance, then heterogeneity increases the population frequency of drug A resistance. Heterogeneity in baseline growth could arise through a variety of mechanisms, such as availability of susceptible hosts, treatment rates differences, or pathogen traits (e.g., transmissibility and duration of carriage).

Second, spatial heterogeneity arises through differences in indirect selection for resistance to drug B (i.e., $s_B^X \neq s_B^{X'}$) coupled with differences in the probability that drug B resistant infections are also doubly-resistant (i.e., $p_{AB}^X/p_B^X \neq p_{AB}^{X'}/p_B^{X'}$). In particular, populations experiencing greater selection for resistance to one drug will have a disproportionate effect on the change in frequency of infections resistant to the other drug, whenever populations differ in frequency of doubly-resistant infections. As an example, if populations experiencing stronger selection for drug B resistance also have a greater probability of drug B-resistant infections being doubly-resistant, heterogeneity in indirect selection increases the frequency of drug A resistance in the metapopulation.

Next, the dynamics of metapopulation, or total, LD can be written as

$$\frac{\mathrm{d}D_{\text{tot}}}{\mathrm{d}t} = \overbrace{(s_A - s + s_B - s)D_{\text{tot}}}^{\text{selection}} - \overbrace{(\mu+\rho)D_{\text{tot}}}^{\substack{\text{mutation and}\\\text{recombination}}} + \overbrace{s_{AB}p_{ab}p_{AB}}^{\text{epistasis}}$$
$$+ \underbrace{\text{cov}(r, D) + \text{coskew}(r, p_A, p_B)}_{\textbf{heterogeneity}\text{ in 'baseline' growth}} + \sum_{\ell\in\{A,B\}} \underbrace{(1-p_\ell)p_\ell \text{cov}\left(s_\ell, \frac{p_{AB}}{p_\ell}\right)}_{\textbf{heterogeneity}\text{ in resistance selection}}, \quad (7)$$

where coskew(r, p_A, p_B) is the coskewness between r, p_A, and p_B and we have assumed population differences in mutation and recombination are negligible (see Sup. Mat. 1.2). The first three terms in equation (7) are the metapopulation level analogues of the first three terms of equation (4) and so share the same interpretation. The last two terms, however, arise due to spatial heterogeneity in 'baseline' growth and selection.

First, spatial heterogeneity arises through differences in the 'baseline' per-capita growth in doubly-resistant infections ($r^X \neq r^{X'}$) coupled with heterogeneities in LD ($D^X \neq D^{X'}$) or resistance frequencies (the coskewness term). The logic of the first term is clear: when population LD differs, more productive populations will disproprotionately contribute to total LD. For the second term, more productive populations with higher frequencies of resistance will tend to produce more doubly-resistant infections; although this need not directly effect population LD, it will disproportionately contribute to total LD.

Second, spatial heterogeneity arises through differences in selection for resistance ($s_\ell^X \neq s_\ell^{X'}$) coupled with differences in the proportion of drug ℓ resistant infections that are doubly-resistant ($p_{AB}^X / p_\ell^X \neq p_{AB}^{X'} / p_\ell^{X'}$). The logic here is that populations experiencing stronger selection for resistance are more likely to see an increase in resistant infections. If this increase occurs disproportionately in doubly-resistant infections, then from equation (1) total LD will increase, whereas if this increase occurs disproportionately in singly-resistant infections, total LD will decrease. The magnitude of this effect is scaled by $p_\ell(1-p_\ell)$ since selection cannot operate without genetic variation. As before, in the absence of population LD, then provided populations experiencing stronger selection for resistance to one drug also have a greater frequency of infections resistant to the other drug, total LD will increase. This could occur if, for example, some populations experience greater treatment rates.

As a final note, observe that in constrast to equation (4), in equation (7) the per-capita migration rates $m^{k\to X}$ are nowhere to be found. The reason for this is intuitive: as migration does not affect the total density of infecteds, nor the resistance status of an infection, it will not change the quantities p_{AB}, p_A, or p_B, and so cannot change total LD. As a consequence, migration only affects total LD indirectly by reducing differences in infection frequency between populations, thereby dampening the magnitude (and hence effect) of $\mathrm{cov}(r, p_\ell)$, $\mathrm{coskew}(r, p_A, p_B)$, and $\mathrm{cov}(s_\ell, p_{AB}/p_\ell)$ in equation (7). It follows that, all else being equal, D_{tot} is a decreasing function of the per-capita migration rate, and so is maximized when migration is infrequent (Fig. 2).

Modeling the dynamics of LD: why bother?

To this point we have focused upon developing the LD perspective to provide a conceptual understanding of MDR evolution in structured populations. However, framing the LD perspective in terms of general quantities has meant this conceptual understanding is somewhat abstract. What we now wish to demonstrate, through the consideration of three scenarios, is how the LD perspective can be used to tackle practical problems. In the first scenario, we show how the LD perspective allows for a straightforward understanding of transient dynamics, and apply this insight to explain patterns of MDR observed in *Streptococcus pneumoniae*. In the second scenario, we show how the LD perspective helps us generalize a recent paper on equilibrium patterns of MDR. In the third scenario we show how the LD perspective generates practical insight into designing drug prescription strategies across populations, with a focus upon a hospital-community setting.

LD perspective explains transient patterns of MDR

In many circumstances we are interested in the transient dynamics of MDR, whether it be to either understand selective sweeps [27, 28], or processes which unfold over sufficiently long time so as to appear in equilibrium [29], or anything in between. However, transient dynamics are more complex than equilibrium processes, and so pose a challenge. In certain circumstances, approximations can simplify the analysis. For example, if selection is sufficiently weak and recombination frequent, then the LD dynamics occur rapidly relative to changes in allele frequencies, and so a quasi-linkage equilibrium approximation can be used [27, 30, 31]. Yet what about situations in which there are no readily available approximations? In these cases, to understand what is (transiently) occurring requires consideration of the dynamical equations (4) and (7). Here we show how transient dynamics coupled with epistasis can explain the patterns of MDR observed in *Streptococcus pneumoniae* [32].

Understanding the patterns of MDR observed in *S. pneumoniae* was first tackled in an important recent paper by [19], using a metapopulation model in which each population represents a different serotype maintained by serotype-specific host immunity [19, 33–35]. In the analysis of [19], they focused upon a metapopulation at equilibrium, and compared their predictions for total (metapopulation) LD and MDR to that of the Maela data set of [32]. However, at equilibrium, the model of [19] predicts each serotype will be in linkage equilibrium, $D^X = 0$, whereas examination of the Maela data reveals that although variation between serotypes accounts for some of the total LD, there also exists significant, unexplained serotype LD (Fig. 3). Can transient dynamics explain this presence of serotype LD?

To explore this possibility, we first need to establish a scenario in which the transient dynamics unfold. In particular, consider a metapopulation initially treated with drug A at sufficiently high rates such that resistance is selected for. At some point ($t = 500$ in Fig. 4), drug B is 'discovered' and is prescribed to patients, while owing to its reduced efficacy, prescription of drug A declines. The increase in drug B prescription means that for many serotypes, resistance to drug B is now favoured, $s_B^X > 0$, and so we should expect drug B resistance to rise in frequency in the metapopulation. However, the reduction in drug A prescription means that for some serotypes, drug A resistance will no longer be favoured, $s_A^X < 0$. Because drug A resistance has reached fixation for many serotypes, drug B resistance is often more likely to occur in an infection with a genetic background resistant to drug A. This will cause both doubly-resistant infections (which benefit from resistance to drug B) and sensitive infections (which have lost resistance to drug A but have yet to gain resistance to drug B) to rise to high frequencies. As more time elapses, in the serotypes for which drug A sensitivity and drug B resistance is favoured, doubly-resistant and sensitive infections will be replaced by infections singly-resistant to drug B. Depending upon the mutation/recombination rates, this process can take enormous amounts of time, generating long periods of apparent stasis in which the population appears to be in equilibrium (see the first row of Fig. 4).

Although this process will generate significant transient total LD due to the covariance in resistance frequency across serotypes, in equation (4) there is nothing generating serotype LD since migration (i.e., antigenic recombination) is infrequent. This leaves epistasis as the remaining (deterministic) force capable of generating serotype LD. Indeed, the addition of epistasis can generate significant, transient LD within serotypes as the transient selective sweeps are ongoing (Fig. 4). Thus transient dynamics coupled with epistasis can explain the significant within-serotype LD observed in *Streptococcus pneumoniae*. Critically, in all cases in Figure 4,

the LD at both the metapopulation and serotype level is transient, and the final state is linkage equilibrium. How population structure maintains LD at equilibrium is the focus of the next example.

LD perspective explains equilibrium patterns of MDR

The paper by [19] focused upon MDR evolution in a metapopulation consisting of independent host populations (so migration is restricted, $m^{X\to X'} \approx 0$). They found that at equilibrium, population differences could lead to MDR overrepresentation ($D_{\text{tot}} > 0$), and that populations with a longer duration of pathogen carriage were more likely to exhibit MDR, a result they attributed to an increased likelihood of antibiotic exposure. Here we show how employing the LD perspective: (i) reveals the evolutionary logic behind what populations differences maintain LD at equilibrium, and (ii) using these insights allows us to generalize the results to a broader range of scenarios, beyond variation in duration of carriage. For simplicity, we will assume there is no epistasis.

There are two required conditions to maintain total LD at equilibrium. First, some mechanism needs to maintain resistance diversity (variation in p_A^X and p_B^X) in the metapopulation. There are variety of ways in which this could occur [19, 35–38], but [19, 35] assume it is due to population differences in the conditions favouring resistance evolution. Since there is no mechanism maintaining within-population diversity, this implies that at equilibrium $D^X = 0$, and so from equation (5) it follows that $D_{\text{tot}} = \text{cov}(p_A, p_B)$. Thus the second condition required for total LD is that p_A and p_B covary. Specifically, whenever p_A^X and p_B^X (or their dynamical equations, (3)), are uncorrelated, the metapopulation will be in linkage equilibrium. From equation (3) we see that if the additive selection coefficients, s_A^X and s_B^X, are uncorrelated, then so too are the dynamics of p_A^X and p_B^X, and so $\text{cov}(p_A, p_B) = 0$. Hence only when population differences generate correlations between the selection coefficients will they generate LD.

Using this insight, why are populations with a longer duration of carriage associated with MDR? And should we expect associations between MDR and any other population attributes? Our primary focus is whether (and how) the selection coefficients are correlated. It is straightforward to compute (see Sup. Mat. 1.4.2),

$$s_A^X = -(\beta_{ab}^X - \beta_{Ab}^X)S^X - (\alpha_{Ab}^X - \alpha_{ab}^X) + \tau_A^X, \tag{8}$$

where we have used slightly different notation from [19]. Now, consider a scenario in which both the treatment rates and the parameters controlling the (additive) costs of resistance (e.g., $\beta_{ab}^X - \beta_{Ab}^X$ and $\alpha_{Ab}^X - \alpha_{ab}^X$) are uncorrelated (this is one of the scenarios presented in Figure 4 of [19]). From equation (8), the only remaining source of correlation is susceptible density, S^X, which plays a role whenever there are explicit transmission costs, $\beta_{Ab}^X < \beta_{ab}^X$. At equilibrium, S^X is determined by pathogen traits such as transmission and duration of carriage, such that 'fitter' populations (i.e., those in which pathogens are more transmissible or have longer duration of carriage) will more substantially deplete susceptibles. By reducing S^X, 'fitter' populations lower the transmission costs for resistance to either drug, and so double-resistance is more likely to be selectively advantageous, even when treatment rates are uncorrelated. In turn, this overrepresentation of doubly-resistant infections will generate total LD.

Thus although variation in duration of carriage can lead to MDR evolution and LD through its effect upon susceptible density (Fig. 5a), it is neither necessary (the same pattern can be produced by variation in transmissibility; Fig. 5b) nor sufficient (variation in duration of carriage

has no effect without explicit transmission costs, Fig. 5c). More broadly, if there are more than two drugs, then provided that there are explicit transmission costs for resistance to each drug, susceptible density will generate a correlation between all the selection coefficients, which in turn will yield the pattern of 'nestedness' observed by [19]. What is critical for this effect to be prominent, however, is that there is clear differentiation in population susceptible density, and that the parameters controlling cost of resistance (i.e., $\beta^X_{ab} - \beta^X_{Ab}$), are large enough so as to ensure a strong correlation amongst selection coefficients.

LD perspective helps identify drug prescription strategies limiting the evolution of MDR

Owing to its practical relevance for public health, often we are interested in the consequences different antibiotic deployment strategies across/between populations can have. The populations of interest could correspond to physically distinct groups such as a hospital and its broader community, or different geographical regions (e.g., countries). From a public health perspective, when considering different antibiotic deployment strategies, a variety of factors must be considered, but in general the goal is to successfully treat as many people as possible, thereby reducing the total burden [3, 39]. In this circumstance, the LD in the metapopulation and/or populations can provide important information about the likelihood of treatment success. In particular, for a given population frequency of drug A and drug B resistance, negative LD (MDR underrepresentation) increases the likelihood that if treatment with one drug fails (due to resistance), treatment with the other drug will succeed. On the other hand, positive LD (MDR overrepresentation) increases the likelihood of treatment failure, since a greater proportion of resistant infections are doubly-resistant and so cannot be successfully treated with either drug.

Equations (4) and (7) show that to generate negative LD, drugs should be deployed in a population specific fashion, that is, drug A should be restricted to some populations and drug B restricted to the remaining populations [see also 18, 19]. Doing so will create a negative covariance in selection, such that resistance to drug A (resp. drug B) will be favoured in some populations and disfavoured in the others. This negative covariance in selection will give rise to negative LD and MDR underrepresentation. This outcome can occur even when drugs have to be prescribed at a higher rate in some populations (e.g., some populations are higher risk groups). If instead drugs are deployed indiscriminately across populations, and in addition, some populations require more frequent antibiotic prescription, this will yield a positive covariance of selection and so generate positive LD and MDR (Fig. 2).

As an application of this principle, suppose there are two populations corresponding to a 'hospital' and a 'community'. In this scenario, the three most commonly debated antibiotic deployment strategies are: cycling, in which drugs are temporally rotated in the hospital; mixing, in which hospital patients are randomly assigned different antibiotics; and combination therapy in which drugs are prescribed to patients in combination [e.g., 3–5, 17, 39–41]. There are two relevant points that hold irrespective of which antibiotic deployment strategy is used. First, antibiotics are prescribed at significantly higher rates in hospitals than in the community. Second, because the focal bacteria are commensal all that differs between strategies is what drug(s) people are prescribed and not the treatment rate. In this scenario, we immediately see the problems that can arise (Fig. 6). Because the treatment rate is higher in the hospital than in the community, both mixing and combination therapy will generate a greater relative selective

advantage for both types of resistance in the hospital. In turn, this will generate a positive covariance of selection, leading to positive LD and MDR overrepresentation. On the other hand, if we cycle the drugs between the hospital and the community, such that if drug A is deployed in the hospital, drug B is deployed in the community, this will generate a negative covariance in selection, leading to MDR underrepresentation. Note that this logic could equally be applied if we were considering a network of hospitals; since in that case if we have (say) N hospitals, the LD of the metapopulation is still the average population LD plus the covariance [26]. Thus although cycling can be either the best or worst option for single drug resistance [17] (see also Fig. 6), by generating negative LD it can lead to MDR underrepresentation and improved clinical outcomes.

Conclusions

The evolution of multidrug resistant pathogens is a pressing health concern, and is a topic which is increasingly gaining attention from evolutionary biologists and mathematical modellers alike. However, the typical process in studying the problem of MDR is to introduce a model of the form of (2), and then either proceed to a numerical analysis of these equations or simplify the model further by neglecting the dynamics of double resistant infections [3, 5, 17]. This is because models of MDR evolution rapidly become intractable, a problem which is particularly acute when incorporating aspects of population structure. Here we have argued that a more insightful and simplifying approach is the 'linkage disequilibrium perspective': after specifying the model of interest, as in (2), it is desirable to transform the model into the form of equations (3), (4), (6), and (7), which brings to the forefront the role played by linkage disequilibrium for MDR evolution in structured populations. Using the linkage disequilibrium perspective leaves us better equipped to determine what factors are responsible for generating MDR, and their generality. Moreover, taking such an approach leads to a more straightforward comparison with existing models and results.

References

[1] O'Neill, J., 2015 Tackling a global health crisis: initial steps. *The Review on Antimicrobial Resistance Chaired by Jim O'Neill* .

[2] Blanquart, F., 2019 Evolutionary epidemiology models to predict the dynamics of antibiotic resistance. *Evol. Appl.* **12**, 365–383.

[3] Bonhoeffer, S., Lipsitch, M. & Levin, B. R., 1997 Evaluating treatment protocols to prevent antibiotic resistance. *Proc. Natl. Acad. Sci.* **94**, 12106–12111.

[4] Lipsitch, M., Bergstrom, C. T. & Levin, B. R., 2000 The epidemiology of antibiotic resistance in hospitals: paradoxes and prescriptions. *PNAS* **97**, 1938–1943.

[5] Bergstrom, C. T., Lo, M. & Lipsitch, M., 2004 Ecological theory suggests that antimicrobial cycling will not reduce antimicrobial resistance in hospitals. *PNAS* **101**, 13285–13290.

[6] Austin, D. J. & Anderson, R. M., 1999 Studies of antibiotic resistance within the patient, hospitals and the community using simple mathematical models. *Phil. Trans. R. Soc. B* **354**, 721–738.

[7] Lewontin, R. C., 1964 The interaction of selection and linkage I: General conditions; heterotic models. *Genetics* **49**, 49–67.

[8] Felsenstein, J., 1965 The effect of linkage on directional selection. *Genetics* **52**, 349–363.

[9] Ohta, T., 1982 Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proc. Natl. Acad. Sci.* **79**, 1940–1944.

[10] Barton, N. H., 1995 Linkage and the limits to natural selection. *Genetics* **140**, 821–841.

[11] Rice, S. H., 2004 *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sunderland, MA, USA: Sinauer Associates.

[12] Slatkin, M., 2008 Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. *Nat. Rev. Genetics* **9**, 477–485.

[13] Ohta, T., 1982 Linkage disequilibrium with the island model. *Genetics* **101**, 139–155.

[14] Slatkin, M., 1975 Gene flow and selection in a two-locus system. *Genetics* **81**, 787–802.

[15] Li, W. & Nei, M., 1974 Stable linkage disequilibrium without epistasis in subdivided populations. *Theor. Pop. Biol.* **6**, 173–183.

[16] Nei, M. & Li, W., 1973 Linkage disequilibrium in subdivided populations. *Genetics* **75**, 213–219.

[17] Beardmore, R. E., Peña-Miller, R., Gori, F. & Iredell, J., 2017 Antibiotic cycling and antibiotic mixing: which one best mitigates antibiotic resistance? *Mol. Biol. Evol.* **34**, 802–817.

[18] Day, T. & Gandon, S., 2012 The evolutionary epidemiology of multilocus drug resistance. *Evolution* **66**, 1582–1597.

[19] Lehtinen, S., Blanquart, F., Lipsitch, M. & Fraser, C., 2019 On the evolutionary ecology of multidrug resistance in bacteria. *PLoS Pathogens* .

[20] Lewontin, R. C. & Kojima, K., 1960 The evolutionary dynamics of complex polymorphisms. *Evolution* **14**, 458–472.

[21] Lewontin, R. C., 1964 The interaction of selection and linkage. II: Optimum models. *Genetics* **50**, 757–782.

[22] Trindade, S., Sousa, A., Xavier, K. B., Dionisio, F., Ferreira, M. G. & Gordo, I., 2009 Positive epistasis drives the acquisition of multidrug resistance. *PLoS Genetics* **5**.

[23] MacLean, R. C., Hall, A. R., Perron, G. G. & Buckling, A., 2010 The population genetics of antibiotic resistance: integrating molecular mechanisms and treatment contexts. *Nature Reviews Genetics* **11**, 405–414.

[24] Hall, A. R. & MacLean, R. C., 2011 Epistasis buffers the fitness effects of rifampicin-resistance mutations in *Pseudomonas Aeruginosa*. *Evolution* **65**, 2370–2379.

[25] Bretscher, M. T., Althaus, C. L., Muller, V. & Bonhoeffer, S., 2004 Recombination in HIV and the evolution of drug resistance: for better or for worse? *BioEssays* **26**, 180–188.

[26] Feldman, M. W. & Christiansen, F. B., 1975 The effect of population subdivision on two loci without selection. *Genet. Res. Camb.* **24**, 151–162.

[27] Neher, R. A. & Shraiman, B. I., 2011 Statistical genetics and evolution of quantitative traits. *Rev. Mod. Phys.* **83**, 1283–1300.

[28] Rouzine, I. M., Rodrigo, A. & Coffin, J. M., 2001 Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiology and Molecular Biology Reviews* **65**, 151–185.

[29] Hastings, A., Abbott, K. C., Cuddington, K., Francis, T., Gellner, G., Lai, Y., Morozov, A., Petrovskii, S., Scranton, K. & Zeeman, M., 2018 Transient phenomena in ecology. *Science* **361**, 1–9.

[30] Kimura, M., 1965 Attainment of quasi-linkage equilibrium when gene frequencies are changing by natural selection. *Genetics* **52**, 875–890.

[31] Otto, S. P. & Day, T., 2007 *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press.

[32] Turner, P., Turner, C., Jankhot, A., Helen, N., Lee, S. J., Day, N. P., White, N. J., Nosten, F. & Goldblatt, D., 2012 A longitudinal study of *Streptococcus pneumoniae* carriage in a cohort of infants and their mothers on the Thailand-Myanmar border. *PLoS ONE* **7**.

[33] Henriques-Normark, B. & Tuomanen, E. I., 2013 The pneumococcus: epidemiology, microbiology, and pathogenesis. *Cold Spring Harb. Perspect. Med.* pp. 1–15.

[34] Cobey, S. & Lipsitch, M., 2012 Niche and neutral effects of acquired immunity permit coexistence of pneumococcal serotypes. *Science* **335**, 1376–1380.

[35] Lehtinen, S., Blanquart, F., Croucher, N. J., Turner, P., Lipsitch, M. & Fraser, C., 2017 Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proc. Natl. Acad. Sci.* **114**, 1075–1080.

[36] Lipsitch, M., Colijn, C., Cohen, T., Hanage, W. P. & Fraser, C., 2009 No coexistence for free: neutral null models for multistrain pathogens. *Epidemics* **1**, 2–13.

[37] Colijn, C., Cohen, T., Fraser, C., Hanage, W., Goldstein, E., Givon-Lavi, N., Dagan, R. & Lipsitch, M., 2009 What is the mechanism for persistent coexistence of drug-susceptible and drug-resistant strains of *Streptococcus pneumoniae*? *J. R. Soc. Interface* **7**, 905–919.

[38] Davies, N. G., Flasche, S., Jit, M. & Atkins, K. E., 2019 Within-host dynamics shape antibiotic resistance in commensal bacteria. *Nat. Ecol. Evol.* **3**, 440–449.

[39] zur Wiesch, P. A., Kouyos, R., Abel, S., Viechtbauer, W. & Bonhoeffer, S., 2014 Cycling empirical antibiotic therapy in hosptials: meta-analysis and models. *PLoS Pathogens* **10**.

[40] Beardmore, R. E. & Peña-Miller, R., 2010 Antibiotic cycling versus mixing: the difficulty of using mathematical models to definitively quantify their relative merits. *Math. Biosci. Eng.* **7**, 923–933.

[41] Tepekule, B., Uecker, H., Derungs, I., Frenoy, A. & Bonhoeffer, S., 2017 Modeling antibiotic treatment in hospitals: A systematic approach shows benefits of combination therapy over cycling, mixing, and mono-drug therapies. *PLoS Comp. Biol.* **13**.

[42] Nowak, M. A. & May, R. M., 1994 Superinfection and the evolution of parasite virulence. *Proc. R. Soc. B* **255**, 81–89.

[43] Alizon, S., 2013 Co-infection and super-infection models in evolutionary epidemiology. *Interface Focus* **3**.

[44] Weinberger, D. M., Trzcinski, K., Lu, Y., Bogaert, D., Brandes, A., Galagan, J., Anderson, P. W., Malley, R. & Lipsitch, M., 2009 Pneumococcal capsular polysaccharide structure predicts serotype prevalence. *PLoS Pathogens* **5**.

[45] Zafar, M. A., Hamaguchi, S., Zangari, T., Cammer, M. & Weiser, J. N., 2017 Capsule type and amount affect shedding and transmission of *Streptococcus pneumoniae*. *mBio* **8**.

1 Supplementary Material

Here we provide more comprehensive details on the different equations, variables, and definitions used in the main text.

Our focus is on an asymptomatically carried bacteria species in a metapopulation consisting of N populations. Focus upon an arbitrarily chosen population X. Let S^X and I^X_{ij} denote the density of susceptible hosts and ij-infections, respectively, at time t, where i indicates if the infection is resistant ($i = A$) or not ($i = a$) to drug A and j indicates if the infection is resistant ($j = B$) or not ($j = b$) to drug B. Susceptible hosts contract ij infections at a per-capita rate $\beta^X_{ij} I^X_{ij}$, where β^X_{ij} is a rate constant, while ij-infections are naturally cleared at a per-capita rate α^X_{ij}. Hosts in population X are treated with antibiotics A, B, or both in combination, at per-capita rates τ^X_A, τ^X_B, and τ^X_{AB}, respectively. Hosts move from population X to population X' at a per-capita rate $m^{X\to X'}$.

The resistance profile of an infection changes through two processes. First, there may be de novo mutation, and so let μ^X_ℓ be the per-capita rate at which an infection in population X acquires allele ℓ through mutation. Second, a ij-infection may be superinfected by a $k\ell$-strain [42, 43]; in this circumstance recombination may occur. Specifically, $k\ell$-strains are transmitted to ij-infections at rate $\beta^X_{k\ell} I^X_{k\ell} I^X_{ij}$, whereupon with probability σ superinfection occurs. In the event of superinfection, with probability $1-\rho$, recombination does not occur, in which case with equal probability the ij-infection either remains unchanged or becomes a $k\ell$-infection. With probability ρ, recombination does occur, in which case with equal probability the ij-infection becomes either an $i\ell$- or kj-infection. Because our focus is upon the role of population structure, we do not allow for coinfection or within-host competitive differences based upon resistance profiles [e.g., 38] but these are straightforward extensions. Moreover, at this stage we do not make any further specification of the dynamics of uninfected hosts, be they susceptible or recovered, as doing so is not essential for a qualitative understanding of MDR evolution.

Rather than immediately writing down the set of differential equations corresponding to these epidemiological assumptions, we instead group the terms based upon the four biological processes that are occurring. In particular, the change in I^X_{ij} can be written as the sum of:

(1) The net change due to *mutation*, denoted $\Delta\mu^X_{ij}$. As an example, focus upon the change in Ab-infections in population X due to mutation, $\Delta\mu^X_{Ab}$. These infections can increase through mutation in one of two ways: (i) ab-infections acquiring allele A at rate $\mu^X_A I^X_{ab}$ or (ii) AB-infections acquiring allele b at rate $\mu^X_b I^X_{AB}$. On the other hand, I^X_{Ab} infections are lost due to mutation whenever they (i) acquire allele a at a per-capita rate μ^X_a, or (ii) acquire allele B at a per-capita rate μ^X_B. Combining this information gives the change in Ab-infections in population X as

$$\Delta\mu^X_{Ab} = \mu^X_A I^X_{ab} + \mu^X_b I^X_{AB} - (\mu^X_a + \mu^X_B) I^X_{Ab}, \tag{9}$$

which is mathematically equivalent to

$$\Delta\mu^X_{Ab} = \mu^X_A (I^X_{ab} + I^X_{Ab}) + \mu^X_b (I^X_{Ab} + I^X_{AB}) - \mu^X I^X_{Ab}, \tag{10}$$

where $\mu^X \equiv \mu^X_a + \mu^X_A + \mu^X_b + \mu^X_B$ is the per-capita mutation rate in population X. The only difference between the two formulations is interpretation: equation (9) shows only mutations

which lead to a change in state, whereas equation (10) shows all possible mutations, even those which do not. This is why the per-capita loss term, μ^X, in (10) can be considered the total per-capita mutation rate in population X. More generally, we can write $\Delta\mu^X_{ij}$ as

$$\Delta\mu^X_{ij} \equiv \mu^X_i (I^X_{aj} + I^X_{Aj}) + \mu^X_j (I^X_{ib} + I^X_{iB}) - \mu^X I^X_{ij}. \tag{11}$$

(2) The net change due to *recombination*, denoted $\Delta\rho^X_{ij}$. Let ρ^X_ℓ be the per-capita rate at which infections gain allele ℓ through recombination. For example, consider ρ^X_A. In particular, ij-infections are challenged by strains carrying allele A at rate $(\beta^X_{Ab} I^X_{Ab} + \beta^X_{AB} I^X_{AB}) I^X_{ij}$. With probability σ, a superinfection event occurs. Given an superinfection event, with probability ρ recombination happens, in which case with probability $1/2$ the recombinant strain Aj will replace the ij infection. Thus

$$\rho^X_A = \rho \frac{\sigma}{2} (\beta^X_{Ab} I^X_{Ab} + \beta^X_{AB} I^X_{AB}), \tag{12}$$

and ij-infections acquire allele A in population X at rate $\rho^X_A I^W_{ij}$. Therefore the change in ij-infections in population X due to recombination is

$$\Delta\rho^X_{ij} \equiv \rho^X_i (I^X_{aj} + I^X_{Aj}) + \rho^X_j (I^X_{ib} + I^X_{iB}) - \rho^X I^X_{ij} \tag{13}$$

where ρ^X is the per-capita rate of recombination in population X, that is,

$$\rho^X \equiv \rho\sigma \sum_{k\ell} \beta^X_{k\ell} I^X_{k\ell} = \rho^X_a + \rho^X_A + \rho^X_b + \rho^X_B.$$

(3) The net change due to host *migration* between populations,

$$-\sum_{k=1}^{N} m^{X\to k} I^X_{ij} + \sum_{k=1}^{N} m^{k\to X} I^k_{ij}. \tag{14}$$

(4) The net change due to *per-capita growth*,

$$r^X_{ij} \equiv \beta^X_{ij} S^X - \alpha^X_{ij} - \mathbf{1}_a(i)\tau^X_A - \mathbf{1}_b(j)\tau^X_B - (1 - \mathbf{1}_{AB}(ij))\tau^X_{AB} - (1-\rho)\frac{\sigma}{2}\sum_{k\ell}(\beta^X_{k\ell} - \beta^X_{ij}) I^X_{k\ell},$$

where $\mathbf{1}_\ell(k)$ is an indicator variable and is equal to 1 if $\ell = k$ and 0 otherwise.

With these four processes in hand, the dynamics of infection densities are given by the system of $4N$ differential equations

$$\frac{\mathrm{d}I^X_{ij}}{\mathrm{dt}} = \Delta\mu^X_{ij} + \Delta\rho^X_{ij} - \sum_{k=1}^{N} (m^{X\to k} I^X_{ij} - m^{k\to X} I^k_{ij}) + r^X_{ij} I^X_{ij},\ X = 1,2,\ldots,N,\ i \in \{a, A\},\ j \in \{b, B\}. \tag{15}$$

1.1 Population LD and MDR

In what follows, we provide more details for the calculations of population LD and MDR. First, the frequency of infections with allele(s) ℓ or $k\ell$ in population X are

$$p_A^X = \frac{\sum_\ell I_{A\ell}^X}{I^X}, \quad p_B^X = \frac{\sum_\ell I_{\ell B}^X}{I^X}, \quad \text{and} \quad p_{k\ell}^X = \frac{I_{k\ell}^X}{I^X}, \tag{16}$$

where $I^X = \sum_{ij} I_{ij}^X$ is the total density of infections in population X. Using these definitions, the standard measure of linkage equilibrium in population X is

$$D^X = p_{AB}^X - p_A^X p_B^X, \tag{17}$$

which is mathematically equivalent to

$$D^X = p_{AB}^X p_{ab}^X - p_{Ab}^X p_{aB}^X. \tag{18}$$

The three dynamical equations of interest for studying MDR in population X are

$$\begin{aligned}
\frac{\mathrm{d}p_A^X}{\mathrm{d}t} &= s_A^X p_A^X(1-p_A^X) + s_B^X D^X + s_{AB}^X p_A^X(1-p_A^X)\frac{p_{AB}^X}{p_A^X} + (\mu_A^X + \rho_A^X)(1-p_A^X) - (\mu_a^X + \rho_a^X)p_A^X \\
&\quad - \sum_{k=1}^N m^{k\to X}\frac{I^k}{I^X}(p_A^X - p_A^k), \\
\frac{\mathrm{d}p_B^X}{\mathrm{d}t} &= s_B^X p_B^X(1-p_B^X) + s_A^X D^X + s_{AB}^X p_B^X(1-p_B^X)\frac{p_{AB}^X}{p_B^X} + (\mu_B^X + \rho_B^X)(1-p_B^X) - (\mu_b^X + \rho_b^X)p_B^X \\
&\quad - \sum_{k=1}^N m^{k\to X}\frac{I^k}{I^X}(p_B^X - p_B^k), \\
\frac{\mathrm{d}D^X}{\mathrm{dt}} &= (s_A^X - s^X + s_B^X - s^X)D^X - (\mu^X + \rho^X)D^X + s_{AB}^X p_{AB}^X p_{ab}^X \\
&\quad - \sum_{k=1}^N m^{k\to X}\frac{I^k}{I^X}\left(D^X - D^k - (p_A^X - p_A^k)(p_B^X - p_B^k)\right).
\end{aligned} \tag{19}$$

System (19) contains a number of quantities that we now define in more detail. First, the (additive) selection coefficient for resistance to drugs A and B in population X are defined as

$$s_A^X = r_{Ab}^X - r_{ab}^X \quad \text{and} \quad s_B^X = r_{aB}^X - r_{ab}^X, \tag{20}$$

respectively, while epistasis in population X is $s_{AB}^X = r_{AB}^X + r_{ab}^X - r_{Ab}^X - r_{aB}^X$. It follows that we can write each of the per-capita growth rates, r_{ij}^X, as

$$r_{ij}^X = r^X + \mathbf{1}_A(i)s_A^X + \mathbf{1}_B(j)s_B^X + \mathbf{1}_{AB}(ij)s_{AB}^X. \tag{21}$$

This is why $r_{ab}^X = r^X$ can be thought of as 'baseline' per-capita growth. We define the average selection for resistance in population X as

$$s^X = s_A^X p_A^X + s_B^X p_B^X + s_{AB}^X p_{AB}^X. \tag{22}$$

Note that the average per-capita growth rate in population X is therefore $r^X + s^X$, that is, average per-capita growth rate is the sum of the 'baseline' per-capita growth rate and the average selection for resistance.

1.2 Metapopulation LD and MDR

Next, consider metapopulation (or total) LD and MDR. First, the metapopulation level equivalents of equations (16) are

$$p_A = \sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} p_A^k, \quad p_B = \sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} p_B^k, \quad \text{and} \quad p_{ij} = \sum_{k=1}^{N} \frac{I^k}{\sum_{\ell=1}^{N} I^\ell} p_{ij}^k. \tag{23}$$

The standard measure of linkage disequilibrium at the level of the total population is

$$D_{\text{tot}} = p_{AB} - p_A p_B. \tag{24}$$

which in terms of the population level variables is

$$D_{\text{tot}} \equiv \frac{\sum_{k=1}^{N} I^k D^k}{\sum_{j=1}^{N} I^j} + \frac{\sum_{k=1}^{N} I^k p_A^k p_B^k}{\sum_{j=1}^{N} I^j} - \frac{\sum_{k=1}^{N} I^k p_A^k}{\sum_{j=1}^{N} I^j} \frac{\sum_{\ell=1}^{N} I^\ell p_B^\ell}{\sum_{j=1}^{N} I^j} = D + \text{cov}(p_A, p_B) \tag{25}$$

where D is the average population LD and $\text{cov}(p_A, p_B)$ is the covariance between resistance to drug A and resistance to drug B.

Using these variables, the three dynamical equations for studying metapopulation MDR are

$$\begin{aligned}
\frac{\mathrm{d}p_A}{\mathrm{d}t} &= s_A p_A(1-p_A) + s_B D_{\text{tot}} + s_{AB} p_A (1-p_A) \frac{p_{AB}}{p_A} \\
&\quad + (\mu_A + \rho_A)(1-p_A) - (\mu_a + \rho_a) p_A + \text{cov}(r, p_A) + p_B \text{cov}\left(s_B, \frac{p_{AB}}{p_B}\right), \\
\frac{\mathrm{d}p_B}{\mathrm{d}t} &= s_B p_B(1-p_B) + s_A D_{\text{tot}} + s_{AB} p_B (1-p_B) \frac{p_{AB}}{p_B} \\
&\quad + (\mu_B + \rho_B)(1-p_B) - (\mu_b + \rho_b) p_B + \text{cov}(r, p_B) + p_A \text{cov}\left(s_A, \frac{p_{AB}}{p_A}\right), \\
\frac{\mathrm{d}D_{\text{tot}}}{\mathrm{d}t} &= (s_A - s + s_B - s) D_{\text{tot}} - (\mu + \rho) D_{\text{tot}} + s_{AB} p_{ab} p_{AB} + \text{cov}(r, D_{\text{tot}}) + \text{coskew}(r, p_A, p_B) \\
&\quad + \sum_{\ell \in \{A,B\}} (1-p_\ell) p_\ell \text{cov}\left(s_\ell, \frac{p_{AB}}{p_\ell}\right) + (1-p_A)\Lambda_{Aa} - p_A \Lambda_{aA} + (1-p_B)\Lambda_{Bb} - p_B \Lambda_{bB}.
\end{aligned} \tag{26}$$

Note that in the equation $\mathrm{d}D_{\text{tot}}/\mathrm{d}t$, there are terms involving Λ_{ij} which do not appear in the main text. These terms are

$$\Lambda_{Aa} = \text{cov}\left(\mu_A + \rho_A, \frac{p_{aB}}{1-p_A}\right) \quad \text{and} \quad \Lambda_{aA} = \text{cov}\left(\mu_a + \rho_a, \frac{p_{AB}}{p_A}\right), \tag{27}$$

while

$$\Lambda_{Bb} = \text{cov}\left(\mu_B + \rho_B, \frac{p_{aB}}{1-p_B}\right) \quad \text{and} \quad \Lambda_{bB} = \text{cov}\left(\mu_b + \rho_b, \frac{p_{AB}}{p_B}\right). \tag{28}$$

Thus the expression

$$(1-p_A)\Lambda_{Aa} - p_A \Lambda_{aA} + (1-p_B)\Lambda_{Bb} - p_B \Lambda_{bB} \tag{29}$$

in the equation $\mathrm{d}D_{\mathrm{tot}}/\mathrm{d}t$ is the effect upon D_{tot} of spatial heterogeneity in mutation and recombination rates ($\mu_\ell^X \neq \mu_\ell^{X'}$ and/or $\rho_\ell^X \neq \rho_\ell^{X'}$) coupled with differences in the proportion of infections with allele i (e.g., $i = A$ or $i = a$) that are resistant to the other drug ($j = B$). In particular, populations in which infections are more likely to acquire resistance through mutation/recombination disproportionately effect total LD through an increase in doubly-resistant infections. However, these terms are likely to be quite small because they require that substantial differences in mutation/recombination rates exist between populations. Since these terms are unlikely to be a significant contributor to the dynamics of D_{tot}, we ignore them in the main text.

There remains a number of other quantities in system (26) that we now define in more detail. First, the probability that an infection resistant to drug ℓ is found in population X is

$$\frac{I^X}{\sum_{j=1}^N I^j}\frac{p_\ell^X}{p_\ell}. \tag{30}$$

For example, if we apply our variable definitions, it is straightforward to show that

$$\frac{I^X}{\sum_{j=1}^N I^j}\frac{p_A^X}{p_A} = \frac{I_{Ab}^X + I_{AB}^X}{\sum_{k=1}^N (I_{Ab}^k + I_{AB}^k)}. \tag{31}$$

Next, to compute the metapopulation-level selection coefficients, and mutation/recombination rates, we need to compute the weighted average of the population quantities, where the weights are the probability that an infection of a particular type is in population X (calculated above). Applying this logic, the metapopulation-level selection coefficients and epistasis are

$$s_\ell = \sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j}\frac{p_\ell^k}{p_\ell}s_\ell^k \quad \text{and} \quad s_{AB} = \sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j}\frac{p_{AB}^k}{p_{AB}}s_{AB}^k. \tag{32}$$

The average selection for resistance in the metapopulation is

$$s = s_A p_A + s_B p_B + s_{AB} p_{AB}. \tag{33}$$

The per-capita mutation and recombination rates follow similarly. Recall that μ_ℓ and ρ_ℓ are the per-capita rates at which infections gain allele ℓ. Thus, for example,

$$\mu_A = \sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j}\frac{1-p_A^k}{1-p_A}\mu_A^k \quad \text{and} \quad \mu_a = \sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j}\frac{p_A^k}{p_A}\mu_a^k. \tag{34}$$

Similar calculations can be made to arrive at μ_B, μ_b, and the various ρ_ℓ. The total per-capita mutation and recombination rates are

$$\mu = \mu_a + \mu_A + \mu_b + \mu_B \quad \text{and} \quad \rho = \rho_a + \rho_A + \rho_b + \rho_B. \tag{35}$$

1.3 Covariance and coskewness

Finally, we also use a number of covariance terms and a coskewness terms. Let $\mathbb{E}[c]$ denote the expectation of the quantity c. Then applying the definition of covariance, we have

$$\begin{aligned}\mathrm{cov}(p_A, p_B) &= \mathbb{E}[p_A p_B] - \mathbb{E}[p_A]\mathbb{E}[p_B] \\ &= \sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j} p_A^k p_B^k - \left(\sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j} p_A^k\right)\left(\sum_{k=1}^N \frac{I^k}{\sum_{j=1}^N I^j} p_B^k\right)\end{aligned}$$

Following the same procedure, we can calculate $\mathrm{cov}(r, p_A)$ and $\mathrm{cov}(r, D_{\mathrm{tot}})$. When the covariance involves quantities that also specifically depend upon particular allele(s), the only difference is that when computing the expectation the probability used is the probability that an allele ℓ is in population X. For example,

$$\begin{aligned}
\mathrm{cov}\left(s_A, \frac{p_{AB}}{p_A}\right) &= \mathbb{E}\left[s_A \frac{p_{AB}}{p_A}\right] - \mathbb{E}[s_A]\mathbb{E}\left[\frac{p_{AB}}{p_A}\right] \\
&= \sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{p_A^k}{p_A} s_A^k \frac{p_{AB}^k}{p_A^k} - \left(\sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{p_A^k}{p_A} s_A^k\right)\left(\sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{p_A^k}{p_A} \frac{p_{AB}^k}{p_A^k}\right) \\
&= \sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{s_A^k p_{AB}^k}{p_A} - \left(\sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{p_A^k}{p_A} s_A^k\right)\left(\sum_{k=1}^{N} \frac{I^k}{\sum_{j=1}^{N} I^j} \frac{p_{AB}^k}{p_A}\right) \\
&= \sum_{k=1}^{N} \frac{I^k p_{AB}^k}{\sum_{j=1}^{N} I^j} (s_A^k - s_A).
\end{aligned}$$

The covariance terms involving the recombination and mutation rates follow similarly, with the appropriate exchanges of variables. Finally, we have the coskewness term, which can be calculated as

$$\begin{aligned}
\mathrm{coskew}(r_{ab}, p_A, p_B) &= \mathbb{E}\left[(r_{ab} - \mathbb{E}[r_{ab}])(p_A - \mathbb{E}[p_A])(p_B - \mathbb{E}[p_B])\right] \\
&= \mathrm{cov}(r, p_A p_B) - p_B \mathrm{cov}(r, p_A) - p_A \mathrm{cov}(r, p_B).
\end{aligned}$$

1.4 Specific examples

1.4.1 Transient dynamics and MDR in *Streptococcus pneumoniae*

Here we use a variant of the model originally proposed by [19, 35] in which the populations represent different serotypes. 'Migration' between serotypes occurs via antigenic recombination with probability m, given transmission between hosts infected with different serotypes has occurred. Resistance is gained and lost through unbiased mutation at a per-capita rate μ and there is no recombination of resistance loci.

Applying these assumptions and using the notation presented with our model from the main text, this yields

$$\begin{aligned}
\frac{\mathrm{d}I_{ij}^X}{\mathrm{d}t} = &\left(\beta_{ij}^X \nu(I, X) S - \alpha_{ij}^X - \mathbf{1}_a(i)\tau_A - \mathbf{1}_b(j)\tau_B - (1 - \mathbf{1}_{AB}(ij))\tau_{AB}\right) I_{ij}^X \\
&+ \mu\left(\sum_{\ell} (I_{\ell j}^X + I_{i\ell}^X) - 4I_{ij}^X\right) + m \sum_{k=1}^{N} (\beta_{ij}^X - \beta_{ij}^k) I_{ij}^X I_{ij}^k \qquad (36)
\end{aligned}$$

where

$$\nu(I, X) = \left(1 - \left[\frac{\sum_{ij} I_{ij}^X}{\sum_{k=1}^{N} \sum_{ij} I_{ij}^k} - \frac{1}{N}\right]\right)^{\omega} \qquad (37)$$

is a balancing function intended to mimic the stabilizing effect adaptive host immunity has upon serotype diversity (ω controls the strength of this effect; see [35]). Note that the treatment rates are assumed to be independent of serotype.

If we let r_{ij}^X denote the per-capita growth term of an ij-infection belonging to serotype X (the first term in brackets in equation (36)), we can partition this as

$$r_{ij}^X = r^X + \mathbf{1}_A(i)s_A^X + \mathbf{1}_B(j)s_B^X + \mathbf{1}_{AB}(ij)s_{AB}^X \tag{38}$$

where

$$\begin{aligned} r^X &= \beta_{ab}^X \nu(I,X)S - \alpha_{ab}^X - \tau_A - \tau_B - \tau_{AB} \\ s_A^X &= -(\beta_{ab}^X - \beta_{Ab}^X)\nu(I,X)S - (\alpha_{Ab}^X - \alpha_{ab}^X) + \tau_A \\ s_B^X &= -(\beta_{ab}^X - \beta_{aB}^X)\nu(I,X)S - (\alpha_{aB}^X - \alpha_{ab}^X) + \tau_B \\ s_{AB}^X &= \left(\beta_{ab}^X + \beta_{AB}^X - \beta_{Ab}^X - \beta_{aB}^X\right)\nu(I,X)S - (\alpha_{ab}^X + \alpha_{AB}^X - \alpha_{Ab}^X - \alpha_{aB}^X) + \tau_{AB} \end{aligned} \tag{39}$$

For simplicity we keep total population size constant, and so set $S = 1 - \sum_{k=1}^N \sum_{ij} I_{ij}^k$.

The simulations in Figure 4 assume the metapopulation is initially treated at per-capita rates $(\tau_A, \tau_B, \tau_{AB}) = (0.12, 0, 0)$, until $t = 500$ when these rates switch to $(\tau_A, \tau_B, \tau_{AB}) = (0.07, 0.1, 0)$. Other parameters values used are $n = 15$, $\omega = 4$, $\beta_{ab}^X - \beta_{Ab}^X = \beta_{ab}^X - \beta_{aB}^X = 0.2$, $\alpha_{Ab}^X - \alpha_{ab}^X = \alpha_{aB}^X - \alpha_{ab}^X = 0.05$, $\mu = 10^{-8}$, and $m = 10^{-8}$. Finally, because *Streptococcus* serotypes differ based upon duration of carriage and transmissibility, and there is evidence of a positive correlation between the two [44, 45], α_{ab}^X was chosen to assume evenly spaced parameter values from $\alpha_{ab}^X = 0.2$ to $\alpha_{ab}^X = 1$, while β_{ab}^X was chosen to assume evenly spaced parameter values from $\beta_{ab}^X = 3$ to $\beta_{ab}^X = 3.5$. Cost epistasis is assumed to solely effect transmissibility. When there is positive epistasis, $\beta_{AB}^X + \beta_{ab}^X - \beta_{Ab}^X - \beta_{aB}^X = 0.05$, whereas for negative epistasis, $\beta_{AB}^X + \beta_{ab}^X - \beta_{Ab}^X - \beta_{aB}^X = -0.05$.

1.4.2 Equilibrium analysis of metapopulation consisting of independent populations

This is a version of one of the models presented in [19]. The metapopulation consists of N populations. The populations are independent (i.e, there is no migration between populations), and each population is assumed to be of fixed size 1 so $S^X = 1 - \sum_{ij} I_{ij}^X$. Resistance is gained and lost through unbiased mutation occurring at rate μ and there is no recombination. Therefore

$$\frac{\mathrm{d}I_{ij}^X}{\mathrm{d}t} = \left(\beta_{ij}^X S^X - \alpha_{ij}^X - \mathbf{1}_a(i)\tau_A^X - \mathbf{1}_b(j)\tau_B^X - (1 - \mathbf{1}_{AB}(ij))\tau_{AB}^X\right) I_{ij}^X + \mu\left(\sum_\ell (I_{\ell j}^X + I_{i\ell}^X) - 4I_{ij}^X\right). \tag{40}$$

If we let r_{ij}^X denote the per-capita growth term of an ij-infection in subpopulation X (the first term in brackets in equation (40)), we can partition this as

$$r_{ij}^X = r^X + \mathbf{1}_A(i)s_A^X + \mathbf{1}_B(j)s_B^X + \mathbf{1}_{AB}(ij)s_{AB}^X \tag{41}$$

where

$$\begin{aligned} r^X &= \beta_{ab}^X S^X - \alpha_{ab}^X - \tau_A^X - \tau_B^X - \tau_{AB}^X \\ s_A^X &= -(\beta_{ab}^X - \beta_{Ab}^X)S^X - (\alpha_{Ab}^X - \alpha_{ab}^X) + \tau_A^X \\ s_B^X &= -(\beta_{ab}^X - \beta_{aB}^X)S^X - (\alpha_{aB}^X - \alpha_{ab}^X) + \tau_B^X \\ s_{AB}^X &= \left(\beta_{ab}^X + \beta_{AB}^X - \beta_{Ab}^X - \beta_{aB}^X\right)S^X - (\alpha_{ab}^X + \alpha_{AB}^X - \alpha_{Ab}^X - \alpha_{aB}^X) + \tau_{AB}^X \end{aligned} \tag{42}$$

This notation and formulation differs from that of [19, 35] in that they assumed costs were multiplicative, that is,

$$\beta_{ab}^X = \beta^X, \quad \beta_{Ab}^X = \beta^X c_{\beta_A}^X, \quad \beta_{aB}^X = \beta^X c_{\beta_B}^X, \quad \beta_{AB}^X = \beta^X c_{\beta_A}^X c_{\beta_B}^X \tag{43}$$

and

$$\alpha^X_{ab} = \alpha^X, \quad \alpha^X_{Ab} = \frac{\alpha^X}{c^X_{\alpha_A}}, \quad \alpha^X_{aB} = \frac{\alpha^X}{c^X_{\alpha_B}}, \quad \alpha^X_{AB} = \frac{\alpha^X}{c^X_{\alpha_A} c^X_{\alpha_B}} \tag{44}$$

where $0 \le c^X_{\beta_\ell} \le 1$ and $0 \le c^X_{\alpha_\ell} \le 1$. The problem with multiplicative costs is apparent when we compute epistasis,

$$s^X_{AB} = \beta^X (1 - c^X_{\beta_A})(1 - c^X_{\beta_B}) S^X - \alpha^X \frac{(1 - c^X_{\alpha_A})(1 - c^X_{\alpha_B})}{c^X_{\alpha_A} c^X_{\alpha_B}} + \tau^X_{AB}. \tag{45}$$

Here we can see that for the model of [19], only when there are no costs of resistance and no combination treatment will there be no epistasis. Thus transmission costs will produce positive epistasis and duration of carriage costs will produce negative epistasis in the model of [19].

In Figure 5 we consider three scenarios; whenever possible we choose parameter values to agree with those of Figure 4 in [19]. In each scenario we assume there are 20 independent populations, that the per-capita mutation rate is $\mu = 10^{-10}$, and there is no epistasis, $s^X_{AB} = 0$. In subplot 5**a**, we set $\beta^X_{ab} = 2$, while duration of carriage varies by population from $\alpha^X_{ab} = 0.25$ to $\alpha^X_{ab} = 1.75$. In subplot 5**b** we set $\alpha^X_{ab} = 0.5$, while transmission varies by population from $\beta^X_{ab} = 1$ to $\beta^X_{ab} = 3$. In both subplots 5**a** and 5**b**, $\alpha^X_{Ab} = \alpha^X_{aB} = \alpha^X_{ab}$, while $\beta^X_{ab} - \beta^X_{Ab} = \beta^X_{ab} - \beta^X_{aB} = 0.1$. Finally in subplot 5**c**, $\beta^X_{ab} = \beta^X_{Ab} = \beta^X_{aB} = 2$, while duration of carriage varies by population from $\alpha^X_{ab} = 0.25$ to $\alpha^X_{ab} = 1.75$, with $\alpha^X_{Ab} - \alpha^X_{ab} = \alpha^X_{aB} - \alpha^X_{ab} = 0.05$.

1.4.3 Constrasting drug prescription strategies in a hospital-community setting

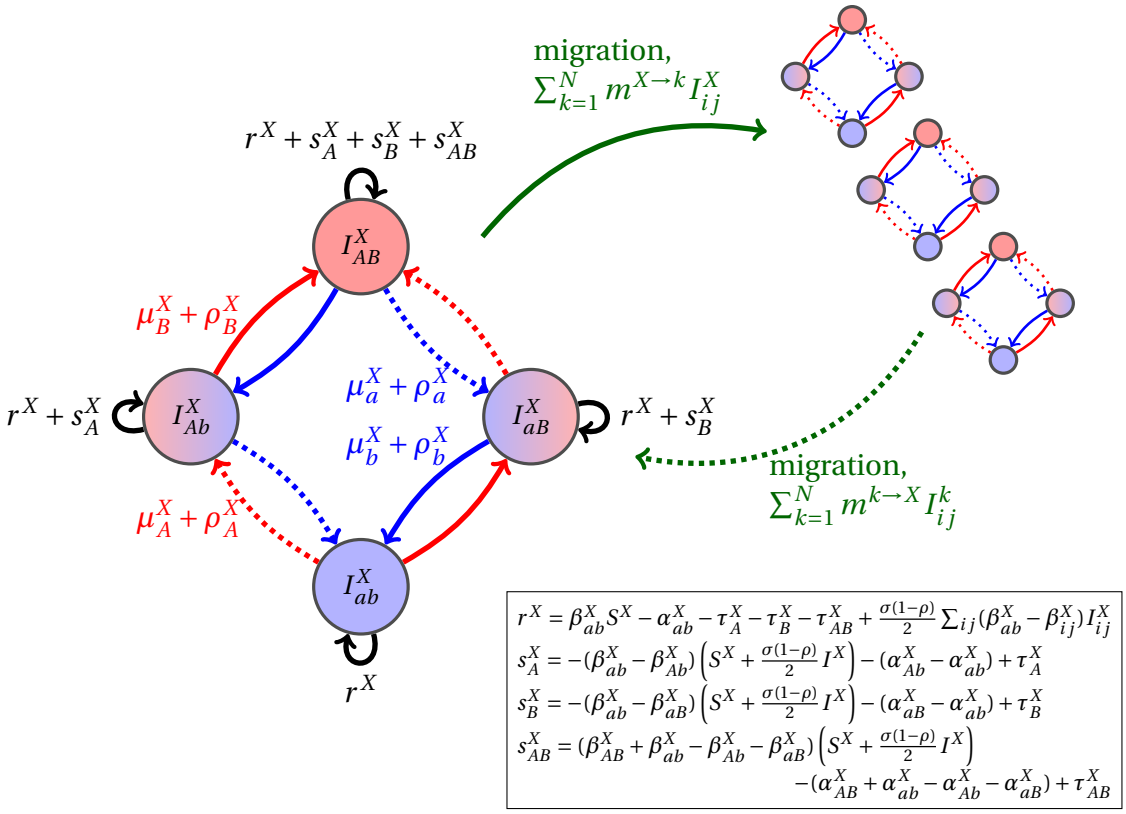
When we model the hospital and community, we use equation (2) and assume the susceptible host density is controlled by

$$\begin{aligned}\frac{\mathrm{d}S^X}{\mathrm{d}t} = \theta^X - dS^X - m^{X\to X'} S^X + m^{X'\to X} S^{X'} - \sum_{ij} \beta^X_{ij} I^X_{ij} S^X \\ + \sum_{ij} (\alpha^X_{ij} - d) I^X_{ij} + \sum_{ij} \left(\tau^X_A \mathbf{1}_a(i) + \tau^X_B \mathbf{1}_b(j) + \tau^X_{AB}(1 - \mathbf{1}_{AB}(ij))\right) I^X_{ij}\end{aligned} \tag{46}$$

where θ^X is the influx of new hosts and d is the background mortality rate.

In the hospital/community model, we assume population C is the 'community' and population H is the 'hospital'. Therefore we let $\theta^H = 0$, and $m^{C\to H} = m/\sum_{ij} I^C_{ij}$ be the rate at which individuals are admitted to the hospital, which is independent of population size. Individuals exit the hospital at a constant rate $m^{H\to C}$, so they spend on average $1/m^{H\to C}$ time units in hospital (assuming background mortality is low). The specification of the migration rates in this way allows us to ensure the 'community' is always much larger than the 'hospital'.

Parameters used in Figure 6 are $\beta^X_{ab} = 2$, $\beta^X_{ab} - \beta^X_{Ab} = \beta^X_{ab} - \beta^X_{aB} = 0.4$, $\alpha^X_{ab} = 0.1$, $\alpha^X_{Ab} - \alpha^X_{ab} = \alpha^X_{aB} - \alpha^X_{ab} = 0.02$, $d = 0.01$, $\theta^C = 0.2$, $\theta^H = 0$, $m^{H\to C} = 0.5$, $m = 0.2$, $\mu = 10^{-7}$, $\sigma = 0$.



$$r^X = \beta_{ab}^X S^X - \alpha_{ab}^X - \tau_A^X - \tau_B^X - \tau_{AB}^X + \frac{\sigma(1-\rho)}{2} \sum_{ij} (\beta_{ab}^X - \beta_{ij}^X) I_{ij}^X$$
$$s_A^X = -(\beta_{ab}^X - \beta_{Ab}^X)\left(S^X + \frac{\sigma(1-\rho)}{2} I^X\right) - (\alpha_{Ab}^X - \alpha_{ab}^X) + \tau_A^X$$
$$s_B^X = -(\beta_{ab}^X - \beta_{aB}^X)\left(S^X + \frac{\sigma(1-\rho)}{2} I^X\right) - (\alpha_{aB}^X - \alpha_{ab}^X) + \tau_B^X$$
$$s_{AB}^X = (\beta_{AB}^X + \beta_{ab}^X - \beta_{Ab}^X - \beta_{aB}^X)\left(S^X + \frac{\sigma(1-\rho)}{2} I^X\right) - (\alpha_{AB}^X + \alpha_{ab}^X - \alpha_{Ab}^X - \alpha_{aB}^X) + \tau_{AB}^X$$

Figure 1: **Schematic of the dynamics of system** (2). The metapopulation consists of N connected populations. Each population has four possible types of infections, linked by one-step mutation or recombination (blue and red arrows), whose per-capita rates are independent of genetic background. The 'baseline' per-capita growth rate of sensitive infections is r^X, the additive selection coefficients for drug A and B resistance are s_A^X and s_B^X, respectively, while s_{AB}^X denotes any epistatic interactions. In the inset, we compute these quantities for the specific model introduced in the main text.

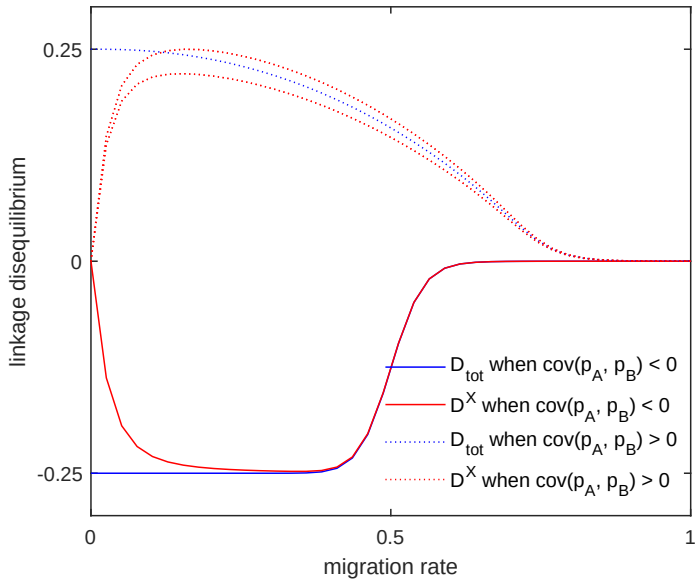


Figure 2: **The effect of migration upon LD at equilibrium depends upon the scale at which LD is measured.** Here we show equilibrium LD in a metapopulation consisting of four populations. Two scenarios are shown. In the first scenario, drug A is prescribed in two populations and drug B is prescribed in the other two populations at the same rate; this yields $\text{cov}(p_A, p_B) < 0$. Because we assume costs of resistance to either drug are identical, all the populations have identical LD. In the second scenario, drug A and drug B are prescribed in the same two populations while the other two populations receive no drugs; this yields $\text{cov}(p_A, p_B) > 0$. Since the drugs are prescribed unequally across populations, the LD observed in each of the two pairs of populations diverge.

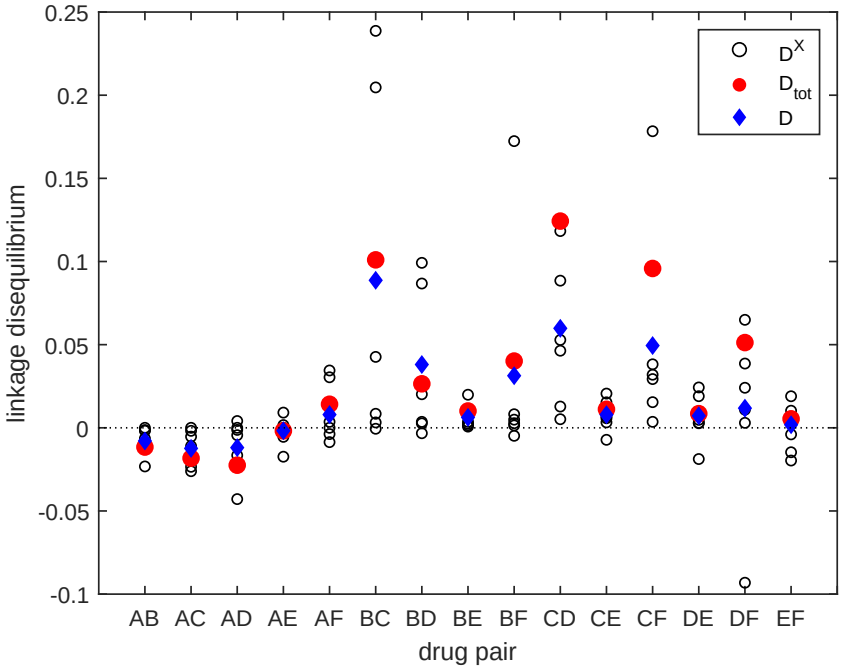


Figure 3: **Linkage disequilibrium for different drug pairs in *Streptococcus pneumoniae*.** Data is from the Maela data set of [19, 32]. The red circles are the observed population LD, D_{tot}, the blue diamonds are the average LD across serotypes, D, and the black circles are the LD of each serotype, D^X. We have restricted the data to serotypes involving 100 or more samples (serotypes 14, 6A/C, 6B, 15B/C, 19F, 23F). The drugs considered are: A = chloramphenicol, B = clindamycin, C = erythromycin, D = penicillin, E = sulphatrimethoprim, and F = tetracycline.

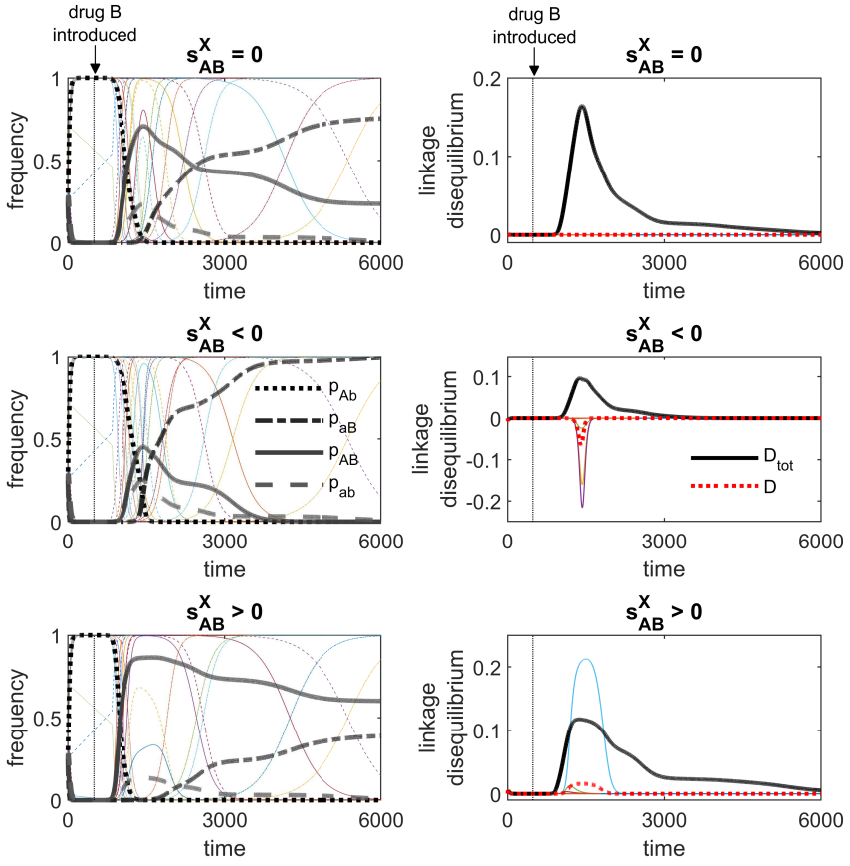


Figure 4: **Transient dynamics and epistasis can explain patterns of LD in *Streptococcus pneumoniae*.** In all simulations, serotypes differ based upon duration of carriage and transmissibility. Hosts are initially treated with drug A at a rate of $\tau_A = 0.12$ per month. At $t = 500$ (months), drug B is introduced, and drug A prescription reduced, $(\tau_A, \tau_B) = (0.07, 0.1)$. In the first row, there is no epistasis, while in the second row there is negative epistasis and in the third row, there is positive epistasis. The thin multicoloured lines denote the within-serotype dynamics. In all cases, at equilibrium both the serotypes and the metapopulation will be in linkage equilibrium, however, transient LD occurs on sufficiently long timescales so as to appear permanent.

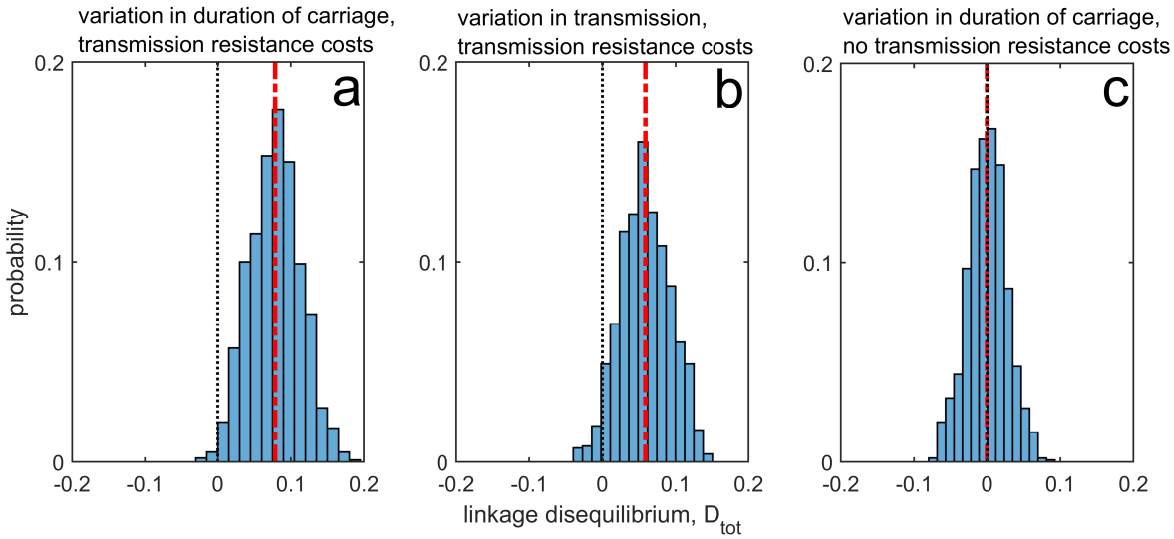


Figure 5: **Duration of carriage is one of many potential explanations for MDR overrepresentation at equilibrium.** Variation in duration of carriage across independent populations can lead to linkage disequilibrium (subplot **a**), but it is neither necessary (**b**), nor sufficient (**c**). We simulate 1000 populations (blue bars), each consisting of 20 independent populations in which treatment rates for each population are randomly chosen to be either $\tau_{\max} = 0.075$ or $\tau_{\min} = 0.025$ with equal probability while simultaneously satisfying $\mathrm{cov}(\tau_A, \tau_B) = 0$. The dashed red line is the mean LD across the simulations for each scenario. In subplot **a**, duration of carriage varies across populations and there are transmission resistance costs; in subplot **b**, transmission varies and there are transmission resistance costs; while in subplot **c**, duration of carriage varies and there are no transmission costs. These simulations diverge slightly from those of [19] in that their model always includes epistasis (see Sup. Info.), whereas here we only consider nonepistatic scenarios.

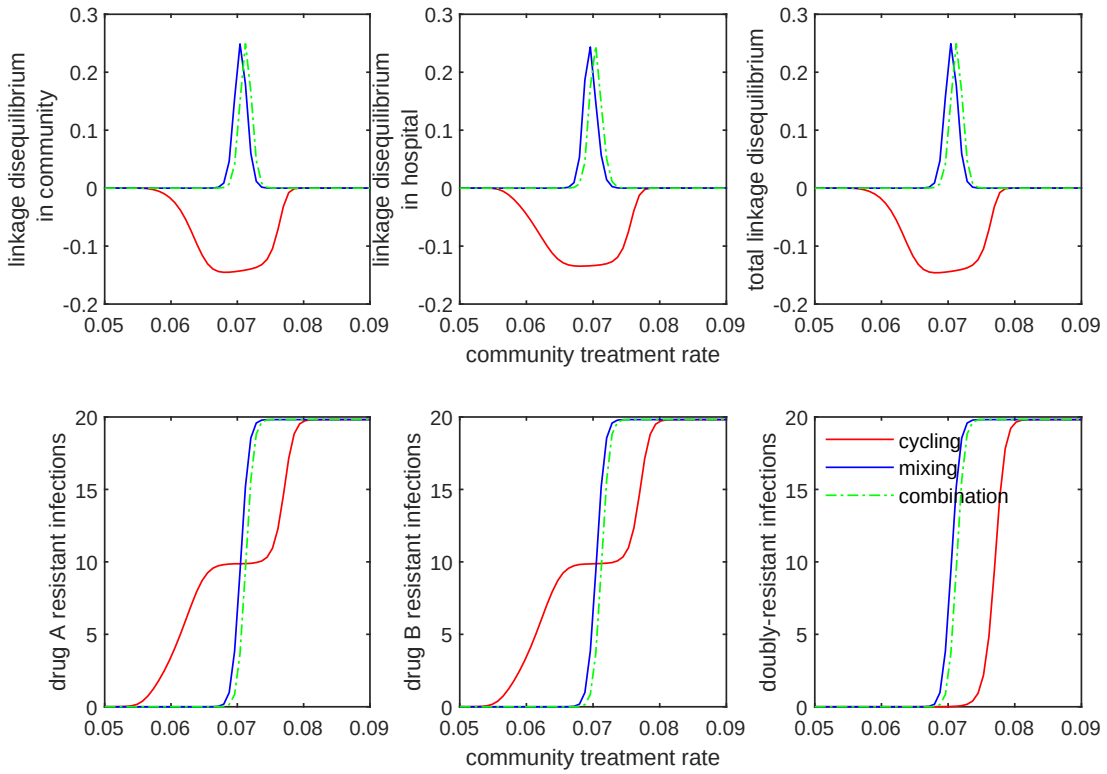


Figure 6: **Different antibiotic prescription strategies generate different patterns of LD at equilibrium**. Here we focus upon a population divided into a community and a hospital. Individuals enter the hospital at a fixed rate and spend a fifth of the time in the hospital that it takes to naturally clear a sensitive infection. The hospital/community size split corresponds to 20 beds per 1000 people, while individuals in the hospital receive antibiotics at 15x the rate they do in the community. We integrate system (2) until equilibrium is reached.

Symbol	Description
I_{ij}^X	Density of ij-infections in population X, where $i = A$ (resp. $i = a$) if infection is resistant (resp. sensitive) to drug A and $j = B$ (resp. $j = b$) if infection is resistant (resp. sensitive) to drug B.
I^X	Density of total infections in population X.
p_ℓ^X, p_ℓ	Frequency of infections resistant to drug(s) ℓ in population X and the metapopulation, respectively.
D^X, D_{tot}, D	Linkage disequilibrium in population X, the metapopulation, and the average across populations, respectively.
$m^{X\to X'}$	Per-capita rate at which hosts migrate from population X to X'.
r^X, r	Per-capita growth rate of sensitive infections in population X and metapopulation (or 'baseline' per-capita growth rate).
s_ℓ^X, s_ℓ	Additive selection coefficient for drug ℓ in population X and the metapopulation, respectively.
s_{AB}^X, s_{AB}	Epistatic effect of being doubly-resistant in population X and the metapopulation, respectively.
$\Delta\mu_{ij}^X$, $\Delta\rho_{ij}^X$	Net change in ij-infections in population X due to mutation or recombination, respectively.
μ_ℓ^X, μ_ℓ	Per-capita rate at which mutations generate allele ℓ.
ρ_ℓ^X, ρ_ℓ	Per-capita rate at which recombination leads to gain of allele ℓ.
s^X, s	Average selection for resistance of any type.
$\text{cov}(F,G)$	Covariance between the quantities F and G, that is, $\text{cov}(F,G) = \mathbb{E}[FG] - \mathbb{E}[F]\mathbb{E}[G]$, where $\mathbb{E}[c]$ denotes the expectation of quantity c.
$\text{coskew}(F,G,H)$	Coskewness between the quantities F, G, H, that is, $\text{coskew}(F,G,H) = \mathbb{E}[(F-\mathbb{E}[F])(G-\mathbb{E}[G])(H-\mathbb{E}[H])]$.

Table 1: Notation used in main text. In all cases, a quantity indexed with a superscript X is the population X quantity, whereas the absence of a superscript X implies the quantity is for the metapopulation.