

Research



Cite this article: Saad-Roy CM, Grenfell BT, Levin SA, van den Driessche P, Wingreen NS. 2021 Evolution of an asymptomatic first stage of infection in a heterogeneous population.

J. R. Soc. Interface **18**: 20210175.

<https://doi.org/10.1098/rsif.2021.0175>

Received: 2 March 2021

Accepted: 24 May 2021

Subject Category:

Life Sciences—Mathematics interface

Subject Areas:

biomathematics, evolution

Keywords:

asymptomatic infection stage, population heterogeneity, pathogen evolution, disease control

Author for correspondence:

Chadi M. Saad-Roy

e-mail: csaadroy@princeton.edu

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5448618>.

Evolution of an asymptomatic first stage of infection in a heterogeneous population

Chadi M. Saad-Roy¹, Bryan T. Grenfell^{2,3}, Simon A. Levin², P. van den Driessche⁵ and Ned S. Wingreen⁴

¹Lewis-Sigler Institute for Integrative Genomics, ²Department of Ecology and Evolutionary Biology, ³Princeton School of Public and International Affairs, and ⁴Department of Molecular Biology, Princeton University, Princeton, NJ, USA

⁵Department of Mathematics and Statistics, University of Victoria, Victoria, British Columbia, Canada

CMS-R, 0000-0002-2217-3071

Pathogens evolve different life-history strategies, which depend in part on differences in their host populations. A central feature of hosts is their population structure (e.g. spatial). Additionally, hosts themselves can exhibit different degrees of symptoms when newly infected; this latency is a key life-history property of pathogens. With an evolutionary-epidemiological model, we examine the role of population structure on the evolutionary dynamics of latency. We focus on specific power-law-like formulations for transmission and progression from the first infectious stage as a function of latency, assuming that the across-group to within-group transmission ratio increases if hosts are less symptomatic. We find that simple population heterogeneity can lead to local evolutionarily stable strategies (ESSs) at zero and infinite latency in situations where a unique ESS exists in the corresponding homogeneous case. Furthermore, there can exist more than one interior evolutionarily singular strategy. We find that this diversity of outcomes is due to the (possibly slight) advantage of across-group transmission for pathogens that produce fewer symptoms in a first infectious stage. Thus, our work reveals that allowing individuals without symptoms to travel can have important unintended evolutionary effects and is thus fundamentally problematic in view of the evolutionary dynamics of latency.

1. Introduction

The traits of pathogens are shaped by numerous trade-offs, and these competing forces mould their life-history strategies. In this context, perhaps the best studied characteristic is pathogen virulence, with the underlying trade-off that increased pathogen virulence can lead to host death before transmission. This idea was first introduced by Anderson & May [1] (for reviews of recent literature on evolution of virulence, see [2,3] and references therein). In addition to virulence, other pathogen characteristics that have been examined include transmission and contact rates in a model with parallel severe and mild infections [4], persistence and invasion constraints [5], and transmission versus recovery [6].

A key feature of certain pathogens is their ability to transmit early in an infection, while the host is either fully asymptomatic or only mildly symptomatic before progressing to a fully symptomatic infectious stage. It is crucial to characterize this stage for pathogen control [7]. Indeed, if the number of asymptomatic infections is high, then control is substantially more challenging. It is thus imperative to understand why certain infections can transmit early on before the appearance of symptoms, e.g. HIV, whereas others cannot, e.g. measles or smallpox. The potential importance of ‘silent’ transmission in the

current COVID-19 outbreak further underlines the importance of understanding the evolutionary dynamics of a latent first stage of infection [8].

An initial latent infectious stage can be studied from a life-history perspective. Saad-Roy *et al.* [9] formulated an evolutionary-epidemiological model for a homogeneous population with two infectious stages. Here, latency is defined as the degree to which a host does not exhibit symptoms in the first infectious stage, and ranges from zero to infinite (fully asymptomatic). Due to smaller pathogen loads and thus reduced host immune responses, a host with larger latency may progress more slowly to the second infectious stage (i.e. remain in the first infectious stage for longer), at the cost of reduced transmission. For general dependences of progression and transmission on latency, the authors of [9] investigated the existence of local ESSs, and found that bistability, i.e. two strategies that each locally maximize fitness, can occur. Furthermore, for specific formulations of simple power-law-like functional dependences of transmission and progression on latency, they proved there are exactly four distinct evolutionary scenarios: (i) the unique evolutionarily stable strategy (ESS) is at zero latency, i.e. the first infectious stage is fully symptomatic, (ii) the unique ESS is at infinite latency, i.e. the first infectious stage is fully asymptomatic, (iii) the unique ESS is at intermediate latency, i.e. the first infectious stage is mildly symptomatic ($0 < \lambda < \infty$) or (iv) there is a unique evolutionarily singular strategy that is *unstable*, yielding bistability with either zero or infinite latency. The main parameters that determine evolutionary outcomes are the exponents that govern the decrease in transmission and progression as functions of latency, and in addition the transmission rate of an individual that is completely asymptomatic in the first infectious stage. Thus, the shape of the transmission–progression trade-off (figure 2b) can dramatically influence pathogen evolutionary dynamics, as has been previously noted more generally [10,11].

However, populations are rarely homogeneous. To address and incorporate heterogeneity in mathematical models of epidemiological dynamics, numerous approaches have been developed. In a nutshell, population structure can be considered either through discrete groups (e.g. [12]), or more explicitly as continuous variables (e.g. [13]). For example, Lajmanovich & Yorke [12] introduced a mathematical model for disease transmission in a heterogeneous population with an arbitrary number of discrete groups. These groups could represent physical separation, such as different patches on a landscape [14–16], or represent another grouping with assortative mixing, such as one based on age, social, or economic factors.

In addition to affecting epidemiological dynamics, host population structure can also affect the evolution of pathogen traits. For example, a series of studies has examined the evolution of virulence under spatial heterogeneity of host populations from both a theoretical [17–21] and experimental standpoint [22]. These findings illustrate that space can play an important role in virulence evolution. In particular, due to relatedness [23,24], lower virulence can arise from rapid depletion of susceptibles at a local scale by more virulent pathogens, a process described in the literature as ‘self-shading’ [3]. In related work on population heterogeneity, Gandon [25] found that the existence of multiple possible hosts may have important evolutionary effects on pathogen virulence; van Baalen & Sabelis [26] and Lion [27] showed that multiple infections (a form of population structure) can substantially affect virulence evolution.

In an otherwise homogeneous population, it is also possible that structure stems from efforts to mitigate disease transmission. Imperfect regional quarantine partially restricts contacts between hosts from different regions, and so gives rise to heterogeneity. Another example of disease control generating population structure is unequal degrees of vaccination leading to differential transmission [28]. Thus, human interventions can lead to host population heterogeneity, which could itself alter the evolutionary dynamics of pathogens and their life-history strategies. It is therefore important to investigate the relationship between population heterogeneity and initial infection dynamics from a life-history perspective, especially due to the general public health relevance of initial asymptomatic transmission.

In this paper, we aim to understand how the interplay of population structure and transmission shapes the evolutionary dynamics of the initial infectious stage. Whether population structure is inherent due to biological constraints, e.g. restricted movement in space, or imposed by social processes, such heterogeneity invariably affects the nature of the trade-offs between transmission and progression in the initial infectious stage. Here, we use a simple, analytical model with identical groups in order to distill the effect of heterogeneity on the evolutionary dynamics of latency. However, in order to focus on this effect, we omit many other important biological and social factors (which themselves lead to very complex systems). Understanding the interplay of these factors with heterogeneity and latency evolution are salient areas for future work, especially to inform disease mitigation in human populations. We first formulate an epidemiological model with interacting groups (separated in space). Then, we examine the resulting evolutionary dynamics for general formulations, in addition to power-law-like relations, for transmission and progression as functions of latency. Additionally, we first consider a general formulation for the ratio of across-group to within-group transmission rates. Then, when we utilize specific power-law-like relations, we assume that this ratio increases with increased latency. Thus, hosts that are more symptomatic have lower relative across-group to within-group transmission in contrast to hosts that exhibit fewer symptoms and can thus mix more readily between groups. We conclude by examining the implications of our findings for disease control strategies.

2. Epidemiological model

We extend the SIIRS model of [9] to include two groups of individuals, with a fraction N_1 and N_2 of individuals in group 1 and group 2, respectively. The model is depicted in figure 1 and, for group $i \neq k$, is formulated as

$$\left. \begin{aligned} \frac{dS_i}{dt} &= \delta N_i - \alpha_{1,i} I_{1,i} S_i - \alpha_{2,i} I_{2,i} S_i - p_k \sigma_k \alpha_{1,k} I_{1,k} S_i \\ &\quad - \sigma_k \alpha_{2,k} I_{2,k} S_i - \delta S_i + \mu R_i, \\ \frac{dI_{1,i}}{dt} &= \alpha_{1,i} I_{1,i} S_i + \alpha_{2,i} I_{2,i} S_i + p_k \sigma_k \alpha_{1,k} I_{1,k} S_i \\ &\quad + \sigma_k \alpha_{2,k} I_{2,k} S_i - v_{1,i} I_{1,i} - \delta I_{1,i}, \\ \frac{dI_{2,i}}{dt} &= v_{1,i} I_{1,i} - (v_{2,i} + \delta) I_{2,i} \\ \text{and } \frac{dR_i}{dt} &= v_{2,i} I_{2,i} - \delta R_i - \mu R_i. \end{aligned} \right\} \quad (2.1)$$

For $i, q = 1, 2$, S_i is the fraction of individuals that are in group i and are susceptible, $I_{q,i}$ is the fraction of individuals that are in group i and in the q th infectious stage, and R_i is the fraction of individuals that are in group i and are recovered.

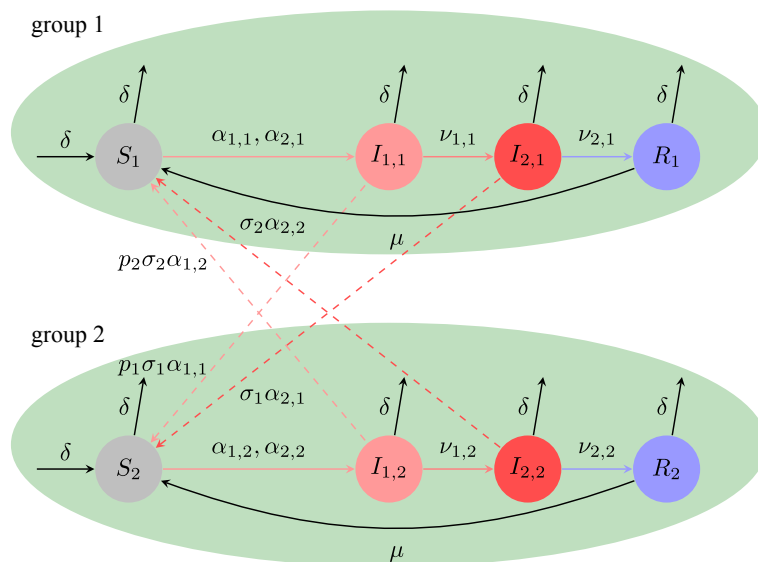


Figure 1. Flowchart of the two-group model for disease transmission. Here, the fraction of individuals that are susceptible (S_i), infectious in the first stage ($I_{1,i}$), infectious in the second stage ($I_{2,i}$), and recovered (R_i) in group i are depicted within each group. The within-group transmission rate for group i and infectious stage j is $\alpha_{j,i}$. The transmission rates from group 1 to group 2 are multiplied by $p_1\sigma_1$ and σ_1 for infectious individuals in the first and second stages, respectively, and those from group 2 to group 1 are analogously multiplied by $p_2\sigma_2$ and σ_2 . For each group, these are depicted as flowing from S_i to $I_{1,i}$ because individuals do not transfer between groups. The progression rates from infectious stage j in group i is $\nu_{j,i}$. The rate of loss of immunity (μ) and the birth (death) rate (δ) are the same in both groups. This figure is extended from [9].

Furthermore, $\alpha_{q,i}$ is the transmission rate of infectious individuals in stage q and group i , $\nu_{q,i}$ is the rate of progression for stage q in group i , $p_i\sigma_i$ is the across-group decrease in initial (stage 1) transmission for group i , and σ_i is the decrease in across-group stage 2 transmission for group i . Note that we assume $p_i \in [1, 1/\sigma_i]$, i.e. the ratio of the across-group to within-group transmission rate in the first stage $p_i\sigma_i$ can vary from equivalency with stage 2 ($p_i\sigma_i = \sigma_i$) to across-group and within-group transmission being equivalent ($p_i\sigma_i = 1$) (although p_i could also be a constant that is decoupled from these equivalencies). Lastly, we assume that the birth/death rate (δ) and rate of loss of immunity (μ) are the same in both groups.

Adding these equations for each group gives that N_1 and N_2 are constant. In this model, there is a unique disease-free equilibrium P_0 , where $S_1 = N_1$, $S_2 = N_2$, and $I_{1,1} = I_{2,1} = R_1 = I_{1,2} = I_{2,2} = R_2 = 0$. Therefore, for $i = 1, 2$, we define the within-group basic reproduction number, i.e. the number of infections in a fully susceptible group caused by one infection of the same group, as

$$\mathcal{R}_0^{(i)} = \frac{\alpha_{1,i}N_i}{\nu_{1,i} + \delta} + \frac{\nu_{1,i}}{(\nu_{1,i} + \delta)(\nu_{2,i} + \delta)} \alpha_{2,i}N_i, \quad (2.2)$$

and the across-group basic reproduction number, i.e. the average number of infected individuals in one group, that is completely susceptible, arising from an infected individual in the other group, as

$$\mathcal{R}_{C,i} = \frac{p_i\sigma_i\alpha_{1,i}N_j}{\nu_{1,i} + \delta} + \frac{\nu_{1,i}}{(\nu_{1,i} + \delta)(\nu_{2,i} + \delta)} \sigma_i\alpha_{2,i}N_j. \quad (2.3)$$

Here, for group i , $\alpha_{1,i}N_i/(\nu_{1,i} + \delta)$ and $\alpha_{2,i}N_i/(\nu_{2,i} + \delta)$ are the numbers of infections arising within this same group while the initial infectious host is in the first and second stage, respectively, and $\nu_{1,i}/(\nu_{1,i} + \delta)$ is the probability of surviving the first stage. Similarly, $p_i\sigma_i\alpha_{1,i}N_j/(\nu_{1,i} + \delta)$ and $\sigma_i\alpha_{2,i}N_j/(\nu_{2,i} + \delta)$ are the corresponding numbers of infections in the other group. It follows from the next-generation matrix

method [29,30] that the basic reproduction number $\mathcal{R}_0$ is the largest eigenvalue of the matrix of new infections that result among susceptibles in each group from a single infectious individual. Thus, for our two-group model,

$$\mathcal{R}_0 = \sqrt{\mathcal{R}_{C,1}\mathcal{R}_{C,2} + \left(\frac{\mathcal{R}_0^{(1)} - \mathcal{R}_0^{(2)}}{2}\right)^2} + \frac{\mathcal{R}_0^{(1)} + \mathcal{R}_0^{(2)}}{2}. \quad (2.4)$$

Since N_i is fixed and $R_i = N_i - S_i - I_{1,i} - I_{2,i}$, we consider the reduced six-dimensional model. In this space, the disease-free equilibrium is globally asymptotically stable if $\mathcal{R}_0 < 1$ (theorem 1, section 2, electronic supplementary material).

3. Evolutionary dynamics

Here, λ is a measure of ‘latency’, i.e. the degree to which the first infectious stage is an asymptomatic infectious stage. Thus, $\lambda = 0$ corresponds to a fully symptomatic first stage of infection, whereas $\lambda \rightarrow \infty$ denotes a fully asymptomatic first stage of infection. We assume that λ governs a trade-off between transmission ($\alpha_1[\lambda]$), progression ($\nu_1[\lambda]$), and across- to within-group first-stage relative to second-stage transmission rates ($p[\lambda]$), all of which relate to the first infectious stage. Modulated by host immunity, a decrease in latency may increase transmissibility during the first infectious stage but also accelerate progression to the second infectious stage (for further rationale of this trade-off, see [9] and references therein including Fraser *et al.* [31] for transmission and progression of HIV, and other references that tie symptoms to transmission). Since symptoms can increase the probability of transmission per contact (through coughs, sneezes, etc.), then even for pathogens where viral load is independent of latency, the transmission rate may decrease as latency increases. Furthermore, hosts with more symptoms (and thus a more severe infection) often have a reduced ability to disperse (e.g. due to screening for symptoms in airports [32,33]), and this corresponds to $p[\lambda]$ increasing as λ increases.

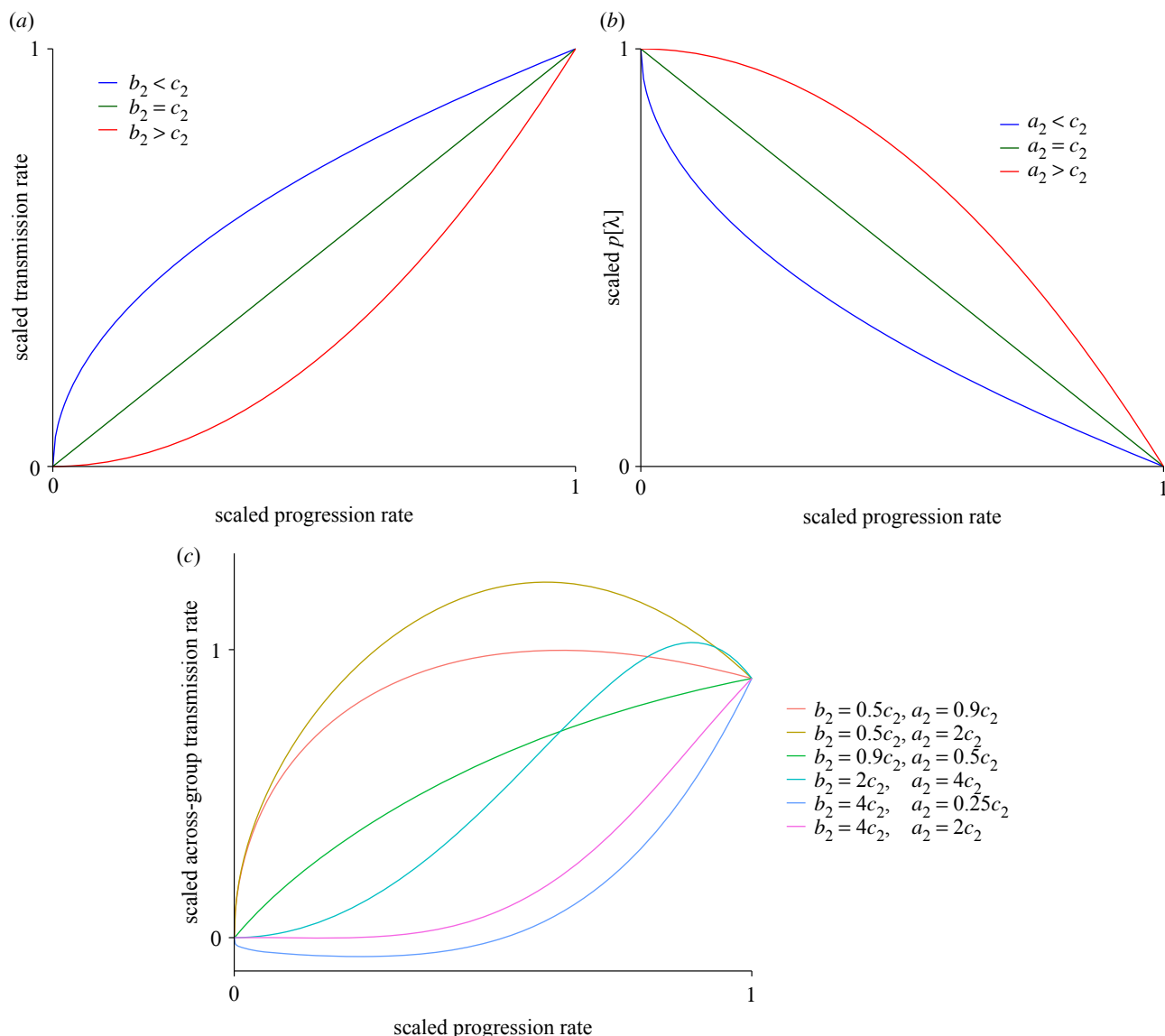


Figure 2. Relationships between key parameters. (a) Scaled transmission rate, $\frac{\alpha_1 - \alpha_{1,\infty}}{b_1}$, as a function of the scaled progression rate. This panel is as in fig. 1D in Saad-Roy *et al.* [9]. (b) Scaled within-group to across-group first-stage transmission ratio $(p - 1)/a_1$ as a function of the scaled progression rate, $(v_1 - v_{1,\infty})/c_1$. (c) Examples of scaled across-group transmission rate, $(p\alpha_1 - p[\infty]\alpha_{1,\infty})/a_1b_1$, as a function of the scaled progression rate. Letting y denote the scaled across-group transmission rate and x denote the scaled progression rate, the relationship between these variables is $y = -x^{(a_2+b_2)/c_2} + (1 + (1/a_1))x^{b_2/c_2} - (\alpha_{1,\infty}/b_1)x^{a_2/c_2}$. For the schematics in this panel, $(1 + (1/a_1)) = 2$ and $\alpha_{1,\infty}/b_1 = 0.1$.

On the other hand, in the biologically less realistic case that symptoms and severity do not correlate with reduced dispersal, then $p[\lambda]$ is constant.

The theory of adaptive dynamics [34,35] assumes ecological and evolutionary dynamics occur on separate timescales, and is used to study the fate of a rare, nearby mutant that arises once the ecological equilibrium has been reached in the environment established by the resident (i.e. weak selection). Ultimately, over evolutionary time, a strategy that maximizes the mutant's invasion fitness is an evolutionarily stable strategy (ESS). Note that this approach is used to study evolutionary outcomes in systems where equilibria have been reached, and the disease is endemic. For newly emerging diseases, such an equilibrium has not been attained yet, and numerous other factors are also important for persistence (e.g. depths of transient troughs [36]).

Here, elucidating epidemiological dynamics for the full two-group SI₁I₂RS model is elaborate, and is currently an open problem. However, if an endemic equilibrium $\hat{E} = (\hat{S}_1, \hat{I}_{1,1}, \hat{I}_{2,1}, \hat{R}_1, \hat{S}_2, \hat{I}_{1,2}, \hat{I}_{2,2}, \hat{R}_2)$ is reached for a phenotype

$\hat{\lambda}$, then, following the next-generation matrix [29,30], the basic reproduction number for a mutant with phenotype λ is

$$\mathcal{R}_m[\lambda, \hat{\lambda}] = \frac{1}{2} \sqrt{4 \frac{\hat{S}_1 \mathcal{R}_{C,1}}{N_1} \frac{\hat{S}_2 \mathcal{R}_{C,2}}{N_2} + \left(\frac{\hat{S}_1 \mathcal{R}_0^{(1)}}{N_1} - \frac{\hat{S}_2 \mathcal{R}_0^{(2)}}{N_2} \right)^2} + \frac{1}{2} \left(\frac{\hat{S}_1 \mathcal{R}_0^{(1)}}{N_1} + \frac{\hat{S}_2 \mathcal{R}_0^{(2)}}{N_2} \right), \quad (3.1)$$

where $\mathcal{R}_0^{(1)} = \mathcal{R}_0^{(1)}[\lambda]$, $\mathcal{R}_0^{(2)} = \mathcal{R}_0^{(2)}[\lambda]$, $\mathcal{R}_{C,1} = \mathcal{R}_{C,1}[\lambda]$, $\mathcal{R}_{C,2} = \mathcal{R}_{C,2}[\lambda]$ are functions of the mutant phenotype λ , and $\hat{S}_1 = \hat{S}_1[\hat{\lambda}]$, $\hat{S}_2 = \hat{S}_2[\hat{\lambda}]$ are functions of the resident phenotype $\hat{\lambda}$. Note that, by definition, $\mathcal{R}_m[\hat{\lambda}, \hat{\lambda}] = 1$.

3.1. Two identical groups

Inspired by Lloyd & May [37], we first focus on the case of two identical groups, i.e. $\alpha_{1,1} = \alpha_{1,2} = \alpha_1$, $\alpha_{2,1} = \alpha_{2,2} = \alpha_2$, $v_{1,1} = v_{1,2} = v_1$, $v_{2,1} = v_{2,2} = v_2$, $\sigma_1 = \sigma_2 = \sigma$, $p_1 = p_2 = p$, $N_1 = N_2 = N = 1/2$ and so $\mathcal{R}_0^{(1)} = \mathcal{R}_0^{(2)} = \mathcal{R}$ and $\mathcal{R}_{C,1} = \mathcal{R}_{C,2} = \mathcal{R}_C$ (for a summary of parameter definitions, see table S1A, electronic

supplementary material). Note that in this case, the basic reproduction number is

$$\mathcal{R}_0[\lambda] = (1 + p\sigma) \frac{\alpha_1[\lambda]N}{v_1[\lambda] + \delta} + (1 + \sigma) \frac{v_1[\lambda]}{v_1[\lambda] + \delta} \frac{\alpha_2 N}{(v_2 + \delta)}. \quad (3.2)$$

If we let $R = (\alpha_1/(v_1 + \delta)) + (v_1/(v_1 + \delta))(\alpha_2/(v_2 + \delta))$, i.e. the homogeneous basic reproduction number (setting $p = \sigma = 1$, since $N = 1/2$), $\mathcal{R}_0$ can be rearranged to give $\mathcal{R}_0[\lambda] = N(\sigma + 1)R[\lambda] + N\sigma(p[\lambda] - 1)(\alpha_1[\lambda]/(v_1[\lambda] + \delta))$. (It can also be seen that $\mathcal{R}_0 \leq R$). Here, for two identical groups, there is a unique endemic equilibrium $\hat{E}$ if and only if $\mathcal{R}_0 > 1$, with $\hat{S} = \frac{N}{\mathcal{R}_0}$ (theorem 2, section 2, electronic supplementary material). If $\mathcal{R}_0 > 1$, the equilibrium $\hat{E}$ is locally asymptotically stable with respect to symmetric perturbations (theorem 3, section 2, electronic supplementary material).

Once the epidemiological dynamics have reached the endemic equilibrium, the reproduction number for an invading mutant (i.e. the invasion fitness) with phenotype λ in a population with resident phenotype $\hat{\lambda}$ is

$$\begin{aligned} \mathcal{R}_m[\lambda, \hat{\lambda}] &= \hat{S}[\hat{\lambda}] \left(\frac{\mathcal{R}_c[\lambda] + \mathcal{R}[\lambda]}{N} \right) \\ &= \hat{S}[\hat{\lambda}] \left((\sigma + 1)R[\lambda] + \sigma(p[\lambda] - 1) \frac{\alpha_1[\lambda]}{v_1[\lambda] + \delta} \right). \end{aligned} \quad (3.3)$$

In this setting with two identical groups, we aim to maximize the mutant's fitness, i.e. $\mathcal{R}_m[\hat{\lambda}, \lambda]$ in equation (3.3), over all possible mutant strategies λ . Here, $\mathcal{R}_m[\hat{\lambda}, \lambda]$ is simply the product of the resident's equilibrium susceptible fraction (i.e. $2\hat{S}[\lambda]$ since $N = 1/2$) multiplied by the basic reproduction number $\mathcal{R}_0[\hat{\lambda}]$ (in a totally naive population) for the mutant. Therefore, this maximization is equivalent to finding the strategy that minimizes the resource in a pure exploitation model [38]. For this minimization, the 'resource' is the fraction of susceptible hosts (cf. [9]). Thus, population structure decreases the basic reproduction number (if $\sigma < 1$, and since $p\sigma < 1$, then $\mathcal{R}_0 < R$), and evolution then maximizes it.

Noting that v_1 , α_1 and p are functions of λ , we seek to find extremal values of $2\hat{S}[\lambda]$, which are therefore extremal values of $\hat{S}[\lambda]$. Since $\hat{S}[\lambda] = \frac{N}{\mathcal{R}_0[\lambda]}$ (and recalling $N = 1/2$ is constant), a value of λ that maximizes $\hat{S}$ minimizes $\mathcal{R}_0$, and a value of λ that minimizes $\hat{S}$ maximizes $\mathcal{R}_0$. Thus, since $\mathcal{R}_0[\lambda] = N(\sigma + 1)R[\lambda] + N\sigma(p[\lambda] - 1)(\alpha_1/(v_1 + \delta))$, evolutionarily singular strategies are values of λ such that

$$\begin{aligned} \frac{d\mathcal{R}_0}{d\lambda} &= N(\sigma + 1) \frac{dR}{d\lambda} + N\sigma \frac{\alpha_1}{v_1 + \delta} \frac{dp}{d\lambda} \\ &+ N \frac{(p - 1)\sigma}{v_1 + \delta} \left[\frac{d\alpha_1}{d\lambda} - \frac{\alpha_1}{v_1 + \delta} \frac{dv_1}{d\lambda} \right] = 0. \end{aligned} \quad (3.4)$$

The first term alone of equation (3.4) gives qualitatively equivalent evolutionary dynamics to the homogeneous model. Thus, the second and third terms are responsible for any emergent phenomena. In the limiting case as $v_1 \rightarrow \infty$, i.e. when the first stage is negligible, these terms vanish. Additionally, if the ratio of the first to second stage relative across-group to within-group transmission rates (p) is not a function of latency, then the second term is zero. Lastly, if $p = 1$, the third term disappears. These observations indicate an important interplay between infectious stages and heterogeneity. The second term denotes the change in pathogen fitness as a function of latency due to a change in the ratio of the first-stage transmission rates as latency increases. The third term can be rearranged as $N(p - 1)\sigma(d/d\lambda)(\alpha_1/$

$(v_1 + \delta)$). Thus, this term denotes the change in pathogen fitness due to differential across-group transmission in the first relative to the second infectious stages.

Additionally, since $\mathcal{R}_0$ does not depend on the rate of waning immunity μ , the evolutionary dynamics are independent of this parameter. Thus, our evolutionary analyses hold for any duration of immunity (including both SI_1I_2S and SI_1I_2R extremes). Furthermore, the following evolutionary analyses also hold for a model that includes death due to infection (see section 3, electronic supplementary material).

In the homogeneous case [9], i.e. $p = \sigma = 1$, and with a general progression–transmission trade-off, there are relations that guarantee the existence of an evolutionarily singular strategy that is either stable or unstable. With power-law-like formulations of transmission and progression as functions of latency, there are four distinct evolutionary outcomes: a unique non-zero ESS $\bar{\lambda} > 0$, bistability with local ESSs at $\bar{\lambda} = 0$ and $\bar{\lambda} \rightarrow \infty$, or a unique ESS at either $\bar{\lambda} = 0$ or $\bar{\lambda} \rightarrow \infty$. These ESSs are all trivially convergence stable, since they emerge from susceptible minimization. (Note that nearby mutants cannot invade an ESS at a finite value of λ . By contrast, we say $\lambda \rightarrow \infty$ is an ESS when $\mathcal{R}_0[\lambda]$ is eventually an increasing function of λ , i.e. nearby mutants with progressively larger latency can invade.)

How does population structure affect these outcomes? In this model with heterogeneity and general formulations for $\alpha_1[\lambda]$, $v_1[\lambda]$ and $p[\lambda]$, modified conditions (see equations (S1)–(S3), electronic supplementary material) guarantee that there is an ESS at zero latency, positive latency, or at least two alternative stable states (section 1.1, electronic supplementary material). The subsequent focus was then on power-law-like formulations of these trade-offs (figure 2a, table S1B–C, electronic supplementary material), such as

$$\alpha_1[\lambda] = b_1(F[\lambda])^{-b_2} + \alpha_{1,\infty} \quad (3.5)$$

and

$$v_1[\lambda] = c_1(F[\lambda])^{-c_2} + v_{1,\infty}, \quad (3.6)$$

where $F[0] = 1$, $F[\infty] \rightarrow \infty$, and $F'[\lambda] > 0$. Importantly in [9], the authors noted that the key relation is the dependence between α_1 and v_1 , and so the choice of $F[\lambda]$ does not affect the biological interpretation. In particular, rearranging equations (3.5) and (3.6), we obtain α_1 as a function of v_1 , and we see that this function is concave up if $b_2 > c_2$ and concave down if $b_2 < c_2$ (figure 2a). Overall, the power-law-like formulation represents a situation where additional symptoms (i.e. a decrease in λ) lead to progressively larger increases in transmission (see [9] for specific examples). In related literature on the evolution of virulence, Gandon *et al.* [39] modelled transmission and recovery as power-law functions of virulence. In what follows, we present our results and the evolutionary behaviour we obtain for the two-group model.

3.1.1. General across-group to within-group first-stage transmission ratio

Here, we assume these formulations for $\alpha_1[\lambda]$ and $v_1[\lambda]$, but consider any non-decreasing function $p[\lambda] \geq 1$ not identical to 1. We examine the possible evolutionary outcomes in comparison to the homogeneous case. We find that, if there exists an interior ESS $\bar{\lambda}$ in the homogeneous setting, then there is an analogous ESS in the heterogeneous two-group model, i.e. $0 < \sigma < 1$, with positive latency λ^* with $\lambda^* > \bar{\lambda}$ (theorem 5,

Table 1. Summary of the evolutionary dynamics classifications in model (1) with two identical groups. These results are explained in detail in Expanded Results, electronic supplementary material, and the relevant theorems are proved in the electronic supplementary material. Additionally, for certain special cases, we proved the uniqueness of the interior evolutionarily singular strategy or the boundary ESS (see sections 1.2 and 1.3 electronic supplementary material for more details).

parameter inequalities			$\hat{S}[0]$	location of local ESSs (possibly other interior ESSs)	relevant theorems
$c_2 > b_2$	$a_2 < b_2$	$\alpha_{1,\infty} > 0$	pos.	$\lambda = 0, \lambda \rightarrow \infty$	9
			neg.	$\lambda \rightarrow \infty$	9
		$\alpha_{1,\infty} = 0$	pos.	$\lambda = 0$	10
			neg.	interior	10
	$a_2 > b_2$	n.a.	pos.	$\lambda = 0$	12, 13
			neg.	interior	12, 13
	$a_2 < c_2$	$\frac{\alpha_{1,\infty}}{\delta} < \frac{\alpha_2}{v_2 + \delta} \frac{\sigma + 1}{a_1 \sigma + \sigma + 1}$	pos.	$\lambda = 0$	8
			neg.	interior	8
		$\frac{\alpha_{1,\infty}}{\delta} > \frac{\alpha_2}{v_2 + \delta} \frac{\sigma + 1}{a_1 \sigma + \sigma + 1}$	pos.	$\lambda = 0, \lambda \rightarrow \infty$	8
			neg.	$\lambda \rightarrow \infty$	8
		$a_2 < c_2$	pos.	$\lambda = 0, \lambda \rightarrow \infty$	9
			neg.	$\lambda \rightarrow \infty$	9
		$\alpha_{1,\infty} = 0$	pos.	$\lambda = 0$	11
			neg.	interior	11

section 2, electronic supplementary material). Thus, even with two identical groups, heterogeneity increases latency. Furthermore, if the unique ESS is at maximal latency, i.e. $\bar{\lambda} \rightarrow \infty$, then the corresponding unique ESS is also at $\lambda^* \rightarrow \infty$ (theorem 4, section 2, electronic supplementary material). Lastly, if there is bistability with zero and maximal latency in the homogeneous setting, then there is at least a local ESS at $\lambda^* \rightarrow \infty$ in the corresponding heterogeneous model with two identical groups (theorem 6, section 2, electronic supplementary material). These results are summarized in electronic supplementary material, figure S1. If $p[\lambda] = p > 1$ is constant, then $p[\lambda]$ is non-decreasing and greater than 1, and so the above results hold. If $p[\lambda] = 1$, then $\mathcal{R}_0[\lambda] = N(\sigma + 1)R[\lambda]$, and the evolutionary dynamics are equivalent to the homogeneous model.

3.1.2. Power-law or exponential across-group to within-group first-stage transmission ratio

To obtain more specific results on the evolutionary dynamics of latency with heterogeneity, we now assume a specific form for the non-decreasing function $p[\lambda]$, namely

$$p[\lambda] = a_1(1 - F[\lambda]^{-a_2}) + 1, \quad (3.7)$$

where the function $F[\lambda]$ is the same as in equations (3.5) and (3.6). For example, $F[\lambda] = (1 + \lambda)$ gives trade-offs formulated as power-laws, whereas $F[\lambda] = e^\lambda$ gives an exponential formulation. Here, a_2 can be interpreted biologically as the exponent that governs the increase of the across-group to within-group transmission ratio as a function of latency. Therefore, we refer to a_2 as the ‘relative heterogeneous transmission growth exponent’. Since $F[0] = 1$ and $F[\infty] \rightarrow \infty$, it follows that $p[0] = 1$ and $p[\infty] \rightarrow a_1 + 1$. As a special case, it is possible that the fully asymptomatic first-stage transmission between groups is the same as within groups, giving $a_1 = \frac{1}{\sigma} + 1$.

Figure 2*b,c* presents schematics of the key relationships between transmission and progression rates that are parameterized by latency (λ), focusing on the across-group to within-group first-stage transmission governed by $p[\lambda]$, and the across-group first-stage transmission rate governed by $p[\lambda]\alpha_1[\lambda]$ as functions of the first-stage progression rate $v_1[\lambda]$.

Rearranging $p[\lambda]$ and $v_1[\lambda]$ gives p as a function of the progression rate v_1 . This function is concave up if $a_2 < c_2$ and concave down if $a_2 > c_2$ (figure 2*b*). As a function of the progression rate, the across-group transmission rate is more complicated, and depends on the exact values of multiple parameters (figure 2*c* and see the caption of figure 2 for the exact equation). The shapes of these curves depict the interactions between progression, transmission and heterogeneity: it is apparent that complicated trade-offs may emerge.

In our model, evolutionarily stable strategies (ESSs) are those strategies λ^* that minimize the fraction of susceptibles $\hat{S}$. We separate our analyses depending upon the relative magnitudes of the progression decay exponent (c_2) and the transmission decay exponent (b_2). We present a summary of the classification of evolutionary outcomes according to the relative magnitudes of these parameters in table 1 (see sections 1.2 and 1.3, in electronic supplementary material for the mathematical details). In what follows in figures 3–6, we illustrate these results with well-chosen examples that capture the important contrasts in evolutionary dynamics between our model and prior work that assumed homogeneity [9].

Figure 3 provides numerical examples to summarize the effect of heterogeneity on the evolutionary dynamics of latency when the progression rate decreases faster than the transmission rate, i.e. $c_2 > b_2$, for different values of $\hat{S}[0]$. In figure 3*a*, $\hat{S}[\lambda]$ is decreasing at zero in all three cases examined. If the relative heterogeneous transmission growth exponent is smaller than the transmission decay exponent, i.e. $a_2 < b_2$, then there is a unique ESS at infinity. By contrast, if $a_2 > b_2$ or

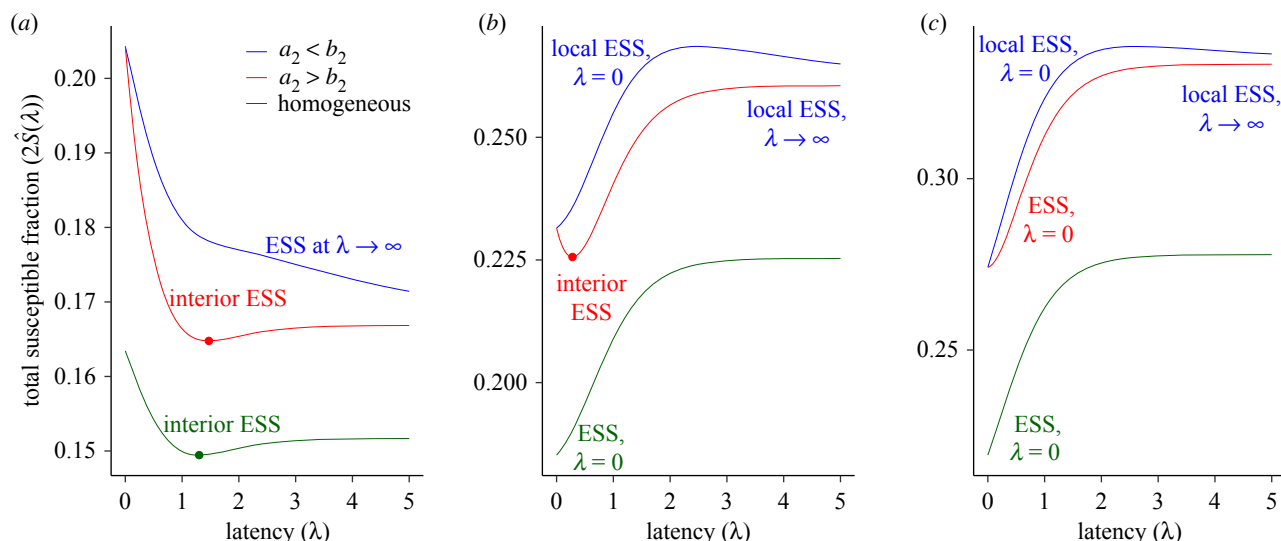


Figure 3. Examples of possible evolutionary dynamics in the identical two-group model when progression decreases faster than transmission, i.e. $c_2 > b_2$ for different values of $\hat{S}'[0]$. (a) $\hat{S}'[0] < 0$ for all three examples, (b) the sign of $\hat{S}'[0]$ is affected by the value of the relative heterogeneous transmission ratio growth exponent a_2 , and (c) $\hat{S}'[0] > 0$ for all three examples. Within each panel, parameters are the same ($b_1 = 0.6$, $c_1 = 0.2$, $b_2 = 2$, $c_2 = 2.25$, $k = \alpha_2/(v_2 + \delta) = 3$, $\alpha_{1,\infty} = 0.18$, $\delta = 1/(365(50))$) except for a_2 , and $p = \sigma = 1$ for the homogeneous case (green line). The heterogeneity parameters are $\sigma = 0.6$, $a_1 = (1/\sigma) - 1$, and $a_2 = 0.3$ and $a_2 = 2.5$ for the blue and red curves, respectively. Across the three panels, these fixed parameters are also the same, except for $v_{1,\infty}$, which then determines whether $\hat{S}'[0]$ is positive. (a) $v_{1,\infty} = 0.05$, (b) $v_{1,\infty} = 0.125$ and (c) $v_{1,\infty} = 0.3$. Note that the vertical axis denotes the total fraction of susceptibles in both groups, i.e. twice the susceptible fraction in a single group.

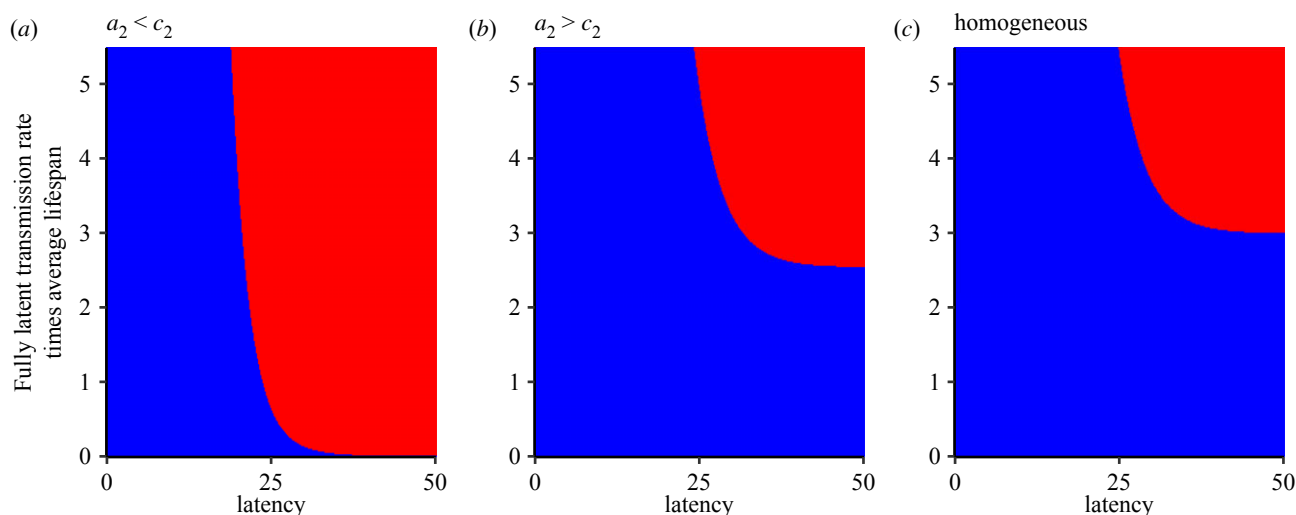


Figure 4. Examples of evolutionary effects of heterogeneity when progression decreases slower than transmission ($c_2 < b_2$) and the emergence of alternative stable states. In all three panels, we plot the sign of $\hat{S}'[\lambda]$ as a function of the fully latent transmission rate times the average lifespan ($\alpha_{1,\infty}/\delta$) and latency. Here, red and blue denote that $\hat{S}[\lambda]$ is a decreasing and increasing function, respectively. The heterogeneous transmission ratio growth exponent is (a) smaller than the progression decay exponent and (b) larger than the progression decay exponent, and in both $\sigma = 0.7$ and $a_1 = \frac{1}{\sigma} - 1$. In (c), $p = \sigma = 1$. Across all panels, $b_1 = 0.6$, $c_1 = 0.1$, $b_2 = 2.25$, $c_2 = 2$, $v_{1,\infty} = 0.01$, $\delta = 1/(365(50))$ and $k = \alpha_2/(v_2 + \delta) = 3$ are fixed. In (a), $a_2 = 1.75$, and in (b), $a_2 = 2.5$.

in the homogeneous model, there is a unique interior ESS. In figure 3b, whether $\hat{S}[\lambda]$ is increasing or decreasing at zero depends on the magnitude of a_2 . Thus, this provides an example where the evolutionary dynamics in all three scenarios are different. If $a_2 < b_2$, there is bistability with zero and infinite latency. If $a_2 > b_2$, there is a unique interior ESS. If there is no heterogeneity, then the unique ESS is at zero latency. In figure 3c, $\hat{S}[\lambda]$ is increasing at zero in all three situations, and the homogeneous model and the example with $a_2 > b_2$ both have a unique ESS at zero latency. By contrast, if $a_2 < b_2$, then zero and infinite latency are bistable.

When the progression rate decreases more slowly than the transmission rate, i.e. $c_2 < b_2$, figure 4 illustrates how the inclusion of heterogeneity favours alternative stable states

of latency. Across a range of values of the fully latent transmission rate times the average lifespan ($\alpha_{1,\infty}/\delta$) and latency, $\hat{S}[\lambda]$ is a decreasing function of λ in the red regions, whereas $\hat{S}[\lambda]$ is an increasing function in the blue regions. Thus, for any horizontal transect, an evolutionarily singular strategy occurs at the intersection of a red and blue region: if a blue region is followed by red, then this is an unstable evolutionarily singular strategy. In figure 4a, we illustrate the important qualitative difference that heterogeneity can introduce in contrast with the homogeneous model (figure 4c). As long as fully latent transmission can occur (i.e. $\alpha_{1,\infty} > 0$), there is an unstable evolutionarily singular strategy and thus alternative states that are evolutionarily stable. The evolutionary outcomes depicted in figure 4b (i.e. $a_2 > c_2$) are qualitatively

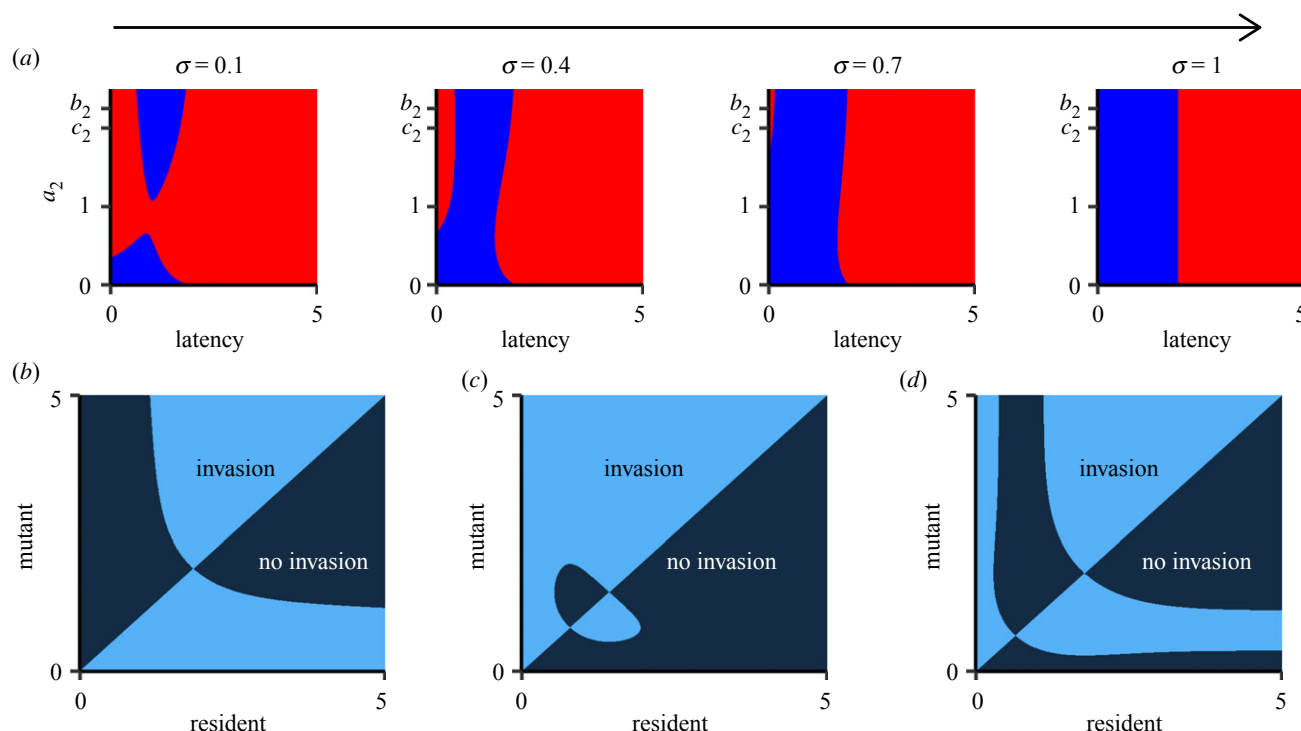


Figure 5. (a) Schematics of possible transitions in evolutionary dynamics due to changes in heterogeneity parameters. An interior evolutionarily singular strategy is where $\hat{S}[\lambda]$ is maximized or minimized. Here, red indicates that $\hat{S}[\lambda]$ is decreasing, whereas blue indicates $\hat{S}[\lambda]$ is increasing, and $a_1 = (1/\sigma) - 1$. The relative across-group to within-group transmission σ is increased from left to right panels, with $\sigma = 0.1$, $\sigma = 0.4$, $\sigma = 0.7$ and $\sigma = 1$, respectively, and $c_2 = 2$, $b_2 = 2.25$. (b) Schematic of pairwise invasibility plot (PIP) with $b_2 > c_2$, $\sigma = 0.1$, and $a_2 = 0.01$ selected so that there is a unique interior evolutionarily singular strategy which is unstable. (c) Schematic of PIP for $b_2 > c_2$, with $\sigma = 0.1$ and $a_2 = 1.5$ chosen so that the system has two interior evolutionarily singular strategies, one stable and one unstable. (d) As in c, but with a bigger a_2 value ($a_2 = 2.3$). In (b–d), light blue denotes a region where the mutant phenotype can successfully invade, whereas dark blue denotes a region where the mutant phenotype dies out. Other parameter values across all panels are $\delta = 1/(365(50))$, $k = \alpha_2/(\nu_2 + \delta) = 3$, $b_1 = 0.56$, $c_1 = 0.1$, $\alpha_{1,\infty} = 0.04$, $\nu_{1,\infty} = 0.01$.

similar to the homogeneous model (figure 4c). However, the quantitative difference in the threshold of $\alpha_{1,\infty}/\delta$ required for bistability is observed: this phenomenon occurs for a larger range in the presence of heterogeneity.

3.2. Other interior evolutionarily singular strategies

Analytically, we have proven the existence of at least one interior ESS or of local ESSs at either zero or maximal latency, and we have proven that there is a unique interior ESS in three special cases (see theorems 10, 13 and 14; section 2, electronic supplementary material). However, in many other cases, there could be other interior evolutionarily singular strategies.

Numerically, we investigated the evolutionary dynamics of latency in the first infectious stage as a function of parameters that govern heterogeneity, and we present these results in figure 5. (Note that numerically we cannot distinguish a local ESS at $\lambda \rightarrow \infty$ from one at large λ .) The panels in figure 5a are as in figure 4, except that now there are horizontal transects which, for increasing latency, are red followed by blue, i.e. $\hat{S}[\lambda]$ decreases and then increases. The intersections of these regions are therefore ESSs. For small relative across-group to within-group transmission (i.e. $\sigma \ll 1$, figure 5a, leftmost panel), a small relative heterogeneous transmission ratio growth exponent a_2 implies that there is a single unstable evolutionarily singular strategy. As a_2 increases, the boundary ESS at $\lambda = 0$ loses its stability, and an interior ESS emerges and is bistable with $\lambda \rightarrow \infty$. As σ increases, the local ESS $\lambda = 0$ loses its stability at progressively higher values of a_2 . At $p = \sigma = 1$, the two groups behave homogeneously.

Schematics of possible pairwise invasibility plots (PIPs) for the evolutionary dynamics of latency in the identical two-group model are presented in figure 5b–d. In the dark blue regions, the mutant phenotype cannot invade the resident phenotype, whereas in the light blue region, the mutant phenotype can successfully invade and becomes the new resident. Figure 5b is a PIP schematic for the case with a unique interior evolutionarily singular strategy that is unstable, leading to bistability with no and maximal latency. By contrast, a schematic PIP with two interior evolutionarily singular strategies, one stable and one unstable, is presented in figure 5c,d, illustrating the possibility of bistability with an interior strategy and $\lambda^* \rightarrow \infty$, a major difference with the homogeneous setting where this was not possible unless more complicated trade-offs, formulated as logistic-like functions of latency, were considered [9].

The main difference between figure 5c,d is that a_2 is smaller in figure 5c, giving rise to a major difference in the range of values for which mutants can invade the interior ESS. In particular, the interior ESS in figure 5c is only locally stable, whereas the ESS at infinite latency is globally stable. Conversely, with larger a_2 , the interior ESS in figure 5d is globally stable, whereas the ESS at maximal latency is only locally stable in this case. It is sensible that a larger a_2 favours the interior ESS at the expense of the ESS at maximal latency, since, for larger values of a_2 , there is a corresponding smaller increase in across-group transmission for large values of latency.

Figure 6 illustrates another major difference between the identical two-group model and the homogeneous case. Here, for the transmission decay exponent greater than the

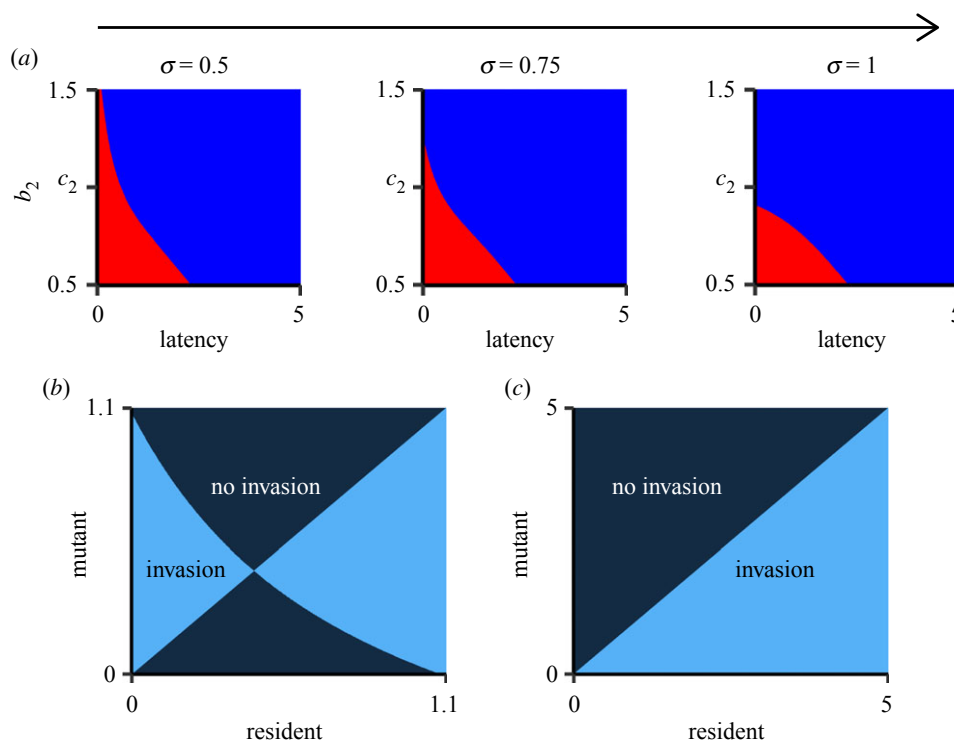


Figure 6. Schematics illustrating that the inclusion of heterogeneity allows for an interior ESS when $b_2 > c_2$, in contrast to the homogeneous case (cf. theorem 15, section 2, electronic supplementary material). (a) Sign of $d\hat{S}/d\lambda$ for different values of latency and b_2 , for $\sigma = 0.5, 0.75$ and 1 . Here, red and blue are as in figure 5. Thus, across a horizontal transect for a fixed value of b_2 , a change in colour denotes an interior evolutionarily singular strategy, and from red to blue denotes an ESS. In all panels, $a_1 = (1/\sigma) - 1$, giving $p[\infty] = 1$, and a_2 is chosen so that it is larger than the maximal value of $b_2 + c_2$ (i.e. $c_2 = 1$, b_2 values from 0.5 to 1.5 and $a_2 = 2.55$) (cf. theorem 14, section 2, electronic supplementary material). (b) PIP schematic for $b_2 = 1.1 > c_2$ and $\sigma = 0.5$, giving rise to an interior ESS. (c) Schematic for $\sigma = 0.75$ and $b_2 = 1.5$ large enough so that there is no interior ESS. Other parameter values across all panels are $\delta = 1/(50(365))$, $k = \alpha_2/(v_2 + \delta) = 3$, $b_1 = 0.6$, $c_1 = 0.1$, $\alpha_{1,\infty} = 0.00005$, $v_{1,\infty} = 0.01$.

progression decay exponent, i.e. $b_2 > c_2$, and for a very small latent transmission rate, there can exist a unique interior ESS. Thus, heterogeneity introduces a qualitative difference to the evolutionary outcomes of latency: the homogeneous model in the analogous setting exhibits only a unique ESS at zero latency. Figure 6a illustrates how the ESS landscape changes for different ratios of across-group and within-group fully symptomatic transmission rates (σ), with $\sigma = 1$ reducing the two-group model to one homogeneous group. Figure 6b presents a PIP schematic for $b_2 > c_2$ and an interior ESS, whereas figure 6c illustrates the case of no interior ESS.

4. Discussion and conclusion

We have formulated an evolutionary-epidemiological model with two interacting groups and two infectious stages to examine evolutionary outcomes of latency for structured populations. We have shown that under the biologically plausible assumption where, normalized by the respective within-group transmission rates, progressively more latent individuals can transmit progressively better to other groups relative to more symptomatic individuals (e.g. increased dispersal due to fewer symptoms and less severe infection), the incorporation of heterogeneity (through identical groups) has important qualitative and quantitative effects on the evolutionary dynamics of latency. Overall, heterogeneity can induce (i) a shift to alternative stable states, (ii) an increase in latency (i.e. the ESSs move to larger values of latency), (iii) a transition from zero latency to partial latency and (iv) the appearance of more than one local ESS with partial latency. In the

appendix, we show that our conclusions are qualitatively the same for an arbitrary number of identical groups with identical across-group interactions, and, as a crude model for transmission around the globe, also for a ring of groups. Below, we explain the biological mechanisms that underlie our results, and follow with their detailed implications.

If the ratio between the across-group and within-group first-stage transmission rates relative to that of the second stage is constant but greater than one, i.e. $dp/d\lambda = 0$ and $p > 1$, then the third term of equation (3.4) remains. In this setting, the evolutionary outcomes of latency are qualitatively similar to the homogeneous model. However, there are quantitative differences that arise. In this setting, the fully latent transmission rate is replaced by the 'effective' fully latent transmission rate $\hat{\alpha}_{1,\infty} = ((1 + p\sigma)/(1 + \sigma))\alpha_{1,\infty}$ and b_1 is replaced by $\hat{b}_1 = ((1 + p\sigma)/(1 + \sigma))b_1$ to determine evolutionary dynamics. Since $\hat{\alpha}_{1,\infty} > \alpha_{1,\infty}$, outcomes with an ESS at positive (non-zero) latency (including bistability) are favoured (remark 6, section 2, electronic supplementary material).

Thus, the changes in qualitative evolutionary dynamics that are induced by heterogeneity are due to the presence of the second term of equation (3.4). In other words, they are due to an increase, which may be small, in the ratio of the first-stage transmission rates as latency increases. To intuitively understand the basis for the richer evolutionary dynamics of latency in this model compared to the homogeneous one and confirm that this mechanism is responsible for the new evolutionary behaviour, we examine a highly simplified version of our two-group model (remark 5, section 2, electronic supplementary material). This means that any slight advantage that less symptomatic individuals have in

heterogeneous transmission can lead to very different evolutionary behaviour. For example, such a situation arises if less symptomatic individuals are more likely to travel between regions (e.g. by evading detection [32,33]). While it is known that the movement of asymptomatic individuals contributes to 'hidden' transmission that is difficult to trace, we have shown, through our model with heterogeneity, that this movement can have very important evolutionary implications that need to be considered for proper public health planning, especially for endemic diseases.

The inclusion of simple heterogeneity increases the range of parameters for which there are alternative stable states of latency. Thus, the ensuing biological implications are important to examine. For example, consider the case with alternative stable states at zero and infinite latency, and suppose a pathogen exhibits zero latency. The parameters governing the trade-offs can change and give rise to an interior local ESS. If the parameters change back, it is possible that the pathogen's latency is now in the basin of attraction of the local ESS at infinite latency. Thus, this reversal would not restore latency to its previous value. It should also be noted that while our model with two identical groups assumes density-dependent transmission, an analogous formulation with frequency-dependence [40], i.e. where transmission is normalized by the total population of the transmitting group, does not alter the evolutionary dynamics (remark 3, section 2, electronic supplementary material, and see remark 4 for the extension to multiple groups).

Through expansions of their homogeneous model, Saad-Roy *et al.* [9] illustrated that their conclusions were robust to other biological additions. First, these authors showed that the introduction of other infectious stages that do not interact with latency in their model gives rise to qualitatively similar behaviour, and we have shown that this continues to hold with simple population structure (see section 4, electronic supplementary material). Moreover, further constraints on the underlying biology of the system can result in a maximal possible latency, and these authors showed intuitively how these constraints can change evolutionary dynamics depending on where interior evolutionarily singular strategies are located. Similar analyses also directly translate to our new framework. Additionally, we have shown in section 3, electronic supplementary material that an epidemiological model with disease-induced death yields qualitatively equivalent evolutionary dynamics of latency. Lastly, with logistic-like trade-offs for transmission and progression, heterogeneity in transmission can qualitatively and quantitatively affect evolutionary dynamics of latency (figure S4, electronic supplementary material). For example, we find that including such heterogeneity can result in three interior evolutionarily singular strategies (figure S4B, electronic supplementary material).

4.1. Interplay with strategies for control

With their model of disease transmission in a homogeneous population, Saad-Roy *et al.* [9] explored the effect of disease control strategies on transmission rate and progression rate trade-offs, and additionally highlighted the importance and implications of multiple local ESSs of latency. Our theoretical analyses further illustrate how control strategies could influence pathogen evolution under different, more broad, settings.

In particular, certain disease control measures could result in host population structure through the separation of a

homogeneous population into groups, in an effort to reduce transmission. An archetypal example of such a measure is imperfect regional quarantine, where less symptomatic individuals are able to mix more readily between regions (e.g. if detection of infection is based on symptoms, such as travel screening [32,33]). For these individuals in the first infectious stage, the across-group to within-group transmission rate increases as a function of latency, and so our framework applies. Thus, these control measures could lead to substantially different evolutionary dynamics, with possible consequences including higher latency, more circumstances leading to alternative stable states at zero and maximal latency, or more than one interior evolutionarily singular strategy. These outcomes are especially negative since they can substantially increase the difficulty of control measures that rely upon the identification and isolation of symptomatic individuals, e.g. through contact tracing. Consequently, the overall duration of an intervention relative to evolutionary timescales is important, and it is therefore imperative to quantify this latter timescale.

Beyond the increase in diversity of possible evolutionary outcomes due to population structure, control strategies can influence trade-offs and thus can affect the qualitative evolutionary dynamics. Notably, the transmission decay exponent, the fully latent transmission rate, and the across-group to within-group transmission ratio growth exponent can all be altered by control strategies. Saad-Roy *et al.* [9] extensively discussed the effect of control measures on evolutionary dynamics through changes in the transmission decay exponent and in the fully latent transmission rate. In addition, the exponent governing the across-group to within-group transmission ratio as a function of latency can decrease if stronger population structure is imposed, or increase if population structure is relaxed (e.g. more symptomatic individuals are able to successfully travel during imperfect regional quarantine).

4.2. Future work

It would be interesting to incorporate additional infectious stages that are affected by the evolutionary outcome of latency and other strain interactions mechanisms such as superinfection of the first infectious stage (which can lead to branching in a homogeneous model [41]). Furthermore, connecting our model to the kinetics that describe the biological processes of infection within a host would also be valuable. Such analyses could reveal the interplay between these kinetic processes and population structure, and the roles they have in latency evolution.

Additionally, we have only proved uniqueness of interior evolutionarily singular strategies for certain special cases, and have shown numerically that two interior evolutionarily singular strategies can exist for specific parameter values. Thus, future work could also extend our theoretical analyses to determine the exact number of interior evolutionarily singular strategies for all conditions.

Numerous further questions arise that are specific to our model, ranging from epidemiological to evolutionary dynamics. The long-term epidemiological dynamics for a model with groups that are not identical remains an open question. If a unique endemic equilibrium also exists in this situation, similar evolutionary analyses would give insight into the effect of certain parameters that influence the trade-offs in only one group. For example, changes in latency ESSs for one group could result from different control strategies in another group,

and investigating changes in evolutionary dynamics that result from these indirect influences would be interesting, and possibly further inform strategies to mitigate pathogen evolution towards a more latent first infectious stage. On the other hand, if there can exist more than one endemic equilibrium, then these equilibria may have central implications for evolutionary dynamics [42]. Thus, characterizing the long-term epidemiological dynamics is an important area for future work. Extending epidemiological-evolutionary analyses to an arbitrary number of non-identical groups each with potentially distinct within-group and across-group transmission rates is also a promising area of future research.

In particular, a specific example of differential mixing is due to age. With an appropriate epidemiological model that reflects the flow of individuals as they age, it would be particularly interesting to understand the changes in evolutionary dynamics of latency that arise due to heterogeneous interactions among children and adults. In principle, an extension of our model with different group characteristics could crudely represent age-structure (including different immune durations, which we have so far ignored). Thus, it would be valuable to determine the long-term epidemiological dynamics (in particular, if there is a unique or multiple endemic equilibria) in such an extension, with which the evolutionary consequences of age-structure on latency could be examined.

Data accessibility. This article has no additional data.

Authors' contributions. Designed research (C.M.S.-R.). Performed research (C.M.S.-R., B.T.G., S.A.L., P.v.d.D. and N.S.W.). Wrote the paper (C.M.S.-R., B.T.G., S.A.L., P.v.d.D. and N.S.W.).

Competing interests. We declare we have no competing interests.

Funding. This work was supported in part by the National Science Foundation, through grant no. CNS-2027908, through the Center for the Physics of Biological Function (PHY-1734030) and through Expeditions in Computing (CCF-1917819), and was also supported by the RAPIDD programme of the Science and Technology Directorate Department of Homeland Security and the Fogarty International Center, NIH, the US Centers for Disease Control and Prevention, the James S. McDonnell Foundation 21st Century Science Initiative Collaborative Award in Understanding Dynamic and Multi-scale Systems, the C3.ai Digital Transformation Institute and Microsoft Corporation, Gift from Google, LLC, the Natural Sciences and Engineering Research Council (NSERC) of Canada through a Discovery grant (P.v.d.D.) and a Postgraduate-Doctoral Scholarship (C.M.S.-R.).

Acknowledgements. We thank anonymous reviewers for suggestions that have improved the manuscript.

Appendix A. Numerous identical groups

Instead of two identical groups, our model can be extended to consider an arbitrary number of identical groups. The within-group transmission in each stage is the same and has force of infection α_1 for stage 1 and α_2 for stage 2. Here, we assume that the across-group transmission is the same for any pair of groups, as in §3.1 of Lloyd & May [37]. Therefore, for the first and second stages, the across-group transmission rates are $p\sigma\alpha_1$ and $\sigma\alpha_2$, respectively, and the fraction of the population in each group is N . Using the next-generation approach [29,30] (section 1.4, electronic supplementary material), it follows that

$$\mathcal{R}_0 = (1 + p\sigma(n-1)) \frac{\alpha_1 N}{v_1 + \delta} + (1 + \sigma(n-1)) \frac{v_1}{(v_1 + \delta)(v_2 + \delta)} \frac{\alpha_2 N}{(v_1 + \delta)(v_2 + \delta)}. \quad (\text{A1})$$

Furthermore, solving the analogous model equations to equation (2.1) shows that there is a unique endemic equilibrium if $\mathcal{R}_0 > 1$ with

$$\hat{S}_i = \hat{S} = \frac{N}{\mathcal{R}_0} = \frac{1}{(\sigma(n-1) + 1)R + (p-1)\sigma(n-1) \frac{\alpha_1}{v_1 + \delta}}. \quad (\text{A2})$$

By a similar argument as for the two identical groups model (cf. theorem 3, section 2, electronic supplementary material), if $\mathcal{R}_0 > 1$ then this endemic equilibrium is also locally asymptotically stable with respect to symmetric perturbations.

To determine the possible evolutionary outcomes for this model, we seek to maximize $\mathcal{R}_0$ or equivalently minimize $\hat{S}[\lambda]$, where $\alpha_1[\lambda]$, $v_1[\lambda]$ and $p[\lambda]$ are functions of λ as in equations (3.5)–(3.7). These calculations for n identical groups at the equilibrium $\hat{S}$ are the same as for the model with two identical groups, except that σ is now replaced by $\sigma(n-1)$ (remark 1, section 2, electronic supplementary material).

In what follows, we summarize our results on the evolutionary outcomes in this model with n groups. Since the calculations are equivalent, all the previous results translate directly to the case with n identical groups. Furthermore, the limit of infinite groups, i.e. $n \rightarrow \infty$, can be taken (section 1.5, electronic supplementary material). For $n \rightarrow \infty$, the thresholds that determine evolutionary dynamics do not directly depend on σ , the relative strength of the across-group symptomatic transmission compared to within-group symptomatic transmission. Rather, they only depend on the fraction $\frac{1}{a_1+1} = \frac{p[0]}{p[\infty]}$, which is the ratio of the fully symptomatic to the fully latent across-group transmission rates.

A.1. Extension to radially symmetric case

The previous model with an arbitrary number of identical groups assumes that the transmission rate between any two different groups is the same. However, transmission around the globe is not symmetrical. To crudely model global transmission, we assume instead that the groups are radially oriented and that the transmission rate between groups i and $j \neq i$ depends only on the 'distance' between these groups, defined as the modular distance between i and j , i.e. with periodic boundaries (e.g. Rohani *et al.* [43]). Thus, instead of a single across-group transmission rate, the transmission rate depends upon which two groups are considered. It therefore follows that the basic reproduction number is now

$$\mathcal{R}_0 = (1 + p\sigma\chi_n) \frac{\alpha_1 N}{v_1 + \delta} + (1 + \sigma\chi_n) \frac{v_1}{(v_1 + \delta)(v_2 + \delta)} \frac{\alpha_2 N}{(v_1 + \delta)(v_2 + \delta)}, \quad (\text{A3})$$

where χ_n is the sum of the relative strengths of the pairwise interactions, i.e.

$$\chi_n = \begin{cases} 2 \sum_{k=1}^{(n/2)-1} f[k] + f\left[\frac{n}{2}\right] & \text{if } n \text{ is even,} \\ 2 \sum_{k=1}^{(n-1)/2} f[k] & \text{if } n \text{ is odd,} \end{cases} \quad (\text{A4})$$

and $f[k]$ is a value that depends on the 'distance' k between two groups. For instance, to recover the previous model, $f[k] = 1$ for all k , so that across-group transmission rates are all $p\sigma\alpha_1$ and $\sigma\alpha_2$ in the first and second stages, respectively.

Below, we briefly highlight the implications of our previous results for this extended model. By radial

symmetry, the n groups are identical, and thus analogues of the main results (section 1.6, electronic supplementary material) hold for this model also, and the unique endemic equilibrium is

$$\hat{S}_i = \hat{S} = \frac{N}{\mathcal{R}_0} = \frac{1}{(\sigma\chi_n + 1)R + (p - 1)\sigma\chi_n(\alpha_1/(v_1 + \delta))}. \quad (\text{A } 5)$$

References

- Anderson RM, May RM. 1982 Coevolution of hosts and parasites. *Parasitology* **85**, 411–426. (doi:10.1017/S003118200055360)
- Alizon S, Hurford A, Mideo N, Van Baalen M. 2009 Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* **22**, 245–259. (doi:10.1111/j.1420-9101.2008.01658.x)
- Cressler CE, McLeod DV, Rozins C, Van Den Hoogen J, Day T. 2016 The adaptive evolution of virulence: a review of theoretical predictions and empirical tests. *Parasitology* **143**, 915–930. (doi:10.1017/S003118201500092X)
- Lin CJ, Deger KA, Tien JH. 2016 Modeling the trade-off between transmissibility and contact in infectious disease dynamics. *Math. Biosci.* **277**, 15–24. (doi:10.1016/j.mbs.2016.03.010)
- King A, Shrestha S, Harvill E, Bjørnstad O. 2009 Evolution of acute infections and the invasion–persistence trade-off. *Am. Nat.* **173**, 446–455. (doi:10.1086/597217)
- Alizon S. 2008 Transmission–recovery trade-offs to study parasite evolution. *Am. Nat.* **172**, E113–E121. (doi:10.1086/589892)
- Fraser C, Riley S, Anderson RM, Ferguson NM. 2004 Factors that make an infectious disease outbreak controllable. *Proc. Natl Acad. Sci. USA* **101**, 6146–6151. (doi:10.1073/pnas.0307506101)
- Tindale L *et al.* 2020 Transmission interval estimates suggest pre-symptomatic spread of COVID-19. *eLife* **9**, e57149. (doi:10.7554/eLife.57149)
- Saad-Roy CM, Wingreen NS, Levin SA, Grenfell BT. 2020 Dynamics in a simple evolutionary–epidemiological model for the evolution of an initial asymptomatic infection stage. *Proc. Natl Acad. Sci. USA* **117**, 11 541–11 550. (doi:10.1073/pnas.1920761117)
- Mazancourt C de, Dieckmann U. 2004 Trade-off geometries and frequency-dependent selection. *Am. Nat.* **164**, 765–778. (doi:10.1086/424762)
- Bowers RG, Hoyle A, White A, Boots M. 2005 The geometric theory of adaptive evolution: trade-off and invasion plots. *J. Theor. Biol.* **233**, 363–377. (doi:10.1016/j.jtbi.2004.10.017)
- Lajmanovich A, Yorke JA. 1976 A deterministic model for gonorrhea in a nonhomogeneous population. *Math. Biosci.* **28**, 221–236. (doi:10.1016/0025-5564(76)90125-5)
- Wu J. 2008 Spatial structure: partial differential equations models. In *Mathematical epidemiology* (eds F Brauer, P van den Driessche, J Wu), pp. 191–203. Berlin, Germany: Springer.
- Grenfell BT, Bolker BM, Kleczkowski A. 1995 Seasonality and extinction in chaotic metapopulations. *Proc. R. Soc. Lond. B* **259**, 97–103. (doi:10.1098/rspb.1995.0015)
- Grenfell B, Kleczkowski A, Gilligan C, Bolker B. 1995 Spatial heterogeneity, nonlinear dynamics and chaos in infectious diseases. *Stat. Methods Med. Res.* **4**, 160–183. (doi:10.1177/096228029 500400205)
- Bolker B, Grenfell BT. 1995 Space, persistence and dynamics of measles epidemics. *Phil. Trans. R. Soc. Lond. B* **348**, 309–320. (doi:10.1098/rstb.1995.0070)
- Boots M, Sasaki A. 1999 Small worlds and the evolution of virulence: infection occurs locally and at a distance. *Phil. Trans. R. Soc. Lond. B* **266**, 1933–1938. (doi:10.1098/rspb.1999.0869)
- Boots M, Sasaki A. 2000 The evolutionary dynamics of local infection and global reproduction in host–parasite interactions. *Ecol. Lett.* **3**, 181–185. (doi:10.1046/j.1461-0248.2000.00139.x)
- Boots M, Hudson PJ, Sasaki A. 2004 Large shifts in pathogen virulence relate to host population structure. *Science* **303**, 842–844. (doi:10.1126/science.1088542)
- Haraguchi Y, Sasaki A. 2000 The evolution of parasite virulence and transmission rate in a spatially structured population. *J. Theor. Biol.* **203**, 85–96. (doi:10.1006/jtbi.1999.1065)
- Lipsitch M, Herre EA, Nowak MA. 1995 Host population structure and the evolution of virulence: a ‘law of diminishing returns’. *Evolution* **49**, 743–748. (doi:10.1111/j.1558-5646.1995.tb02310.x)
- Boots M, Meador M. 2007 Local interactions select for lower pathogen infectivity. *Science* **315**, 1284–1286. (doi:10.1126/science.1137126)
- Wild G, Gardner A, West SA. 2009 Adaptation and the evolution of parasite virulence in a connected world. *Nature* **459**, 983–986. (doi:10.1038/nature08071)
- Lion S, Boots M. 2010 Are parasites ‘prudent’ in space? *Ecol. Lett.* **13**, 1245–1255. (doi:10.1111/j.1461-0248.2010.01516.x)
- Gandon S. 2004 Evolution of multihost parasites. *Evolution* **58**, 455–469. (doi:10.1111/j.0014-3820.2004.tb01669.x)
- van Baalen M, Sabelis MW. 1995 The dynamics of multiple infection and the evolution of virulence. *Am. Nat.* **146**, 881–910. (doi:10.1086/285830)
- Lion S. 2013 Multiple infections, kin selection and the evolutionary epidemiology of parasite traits. *J. Evol. Biol.* **26**, 2107–2122. (doi:10.1111/jeb.12207)
- Glass K, Kappey J, Grenfell BT. 2004 The effect of heterogeneity in measles vaccination on population immunity. *Epidemiol. Infect.* **132**, 675–683. (doi:10.1017/S0950268804002080)
- van den Driessche P, Watmough J. 2002 Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* **180**, 29–48. (doi:10.1016/S0025-5564(02)00108-6)
- Diekmann O, Heesterbeek JAP, Metz JAJ. 1990 On the definition and the computation of the basic reproduction ratio $\mathcal{R}_0$ in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382. (doi:10.1007/BF00178324)
- Fraser C, Hollingsworth TD, Chapman R, Wolf F de, Hanage WP. 2007 Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl Acad. Sci. USA* **104**, 17 441–17 446. (doi:10.1073/pnas.0708559104)
- Gostic KM, Kucharski AJ, Lloyd-Smith JO. 2015 Effectiveness of traveller screening for emerging pathogens is shaped by epidemiology and natural history of infection. *eLife* **4**, e05564. (doi:10.7554/eLife.05564)
- Gostic K, Gomez ACR, Mummah RO, Kucharski AJ, Lloyd-Smith JO. 2020 Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife* **9**, e55570. (doi:10.7554/eLife.55570)
- Geritz SAH, Metz JAJ, Kisdi E, Meszéna G. 1997 Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **78**, 2024–2027. (doi:10.1103/PhysRevLett.78.2024)
- Brannstrom A, Johansson J, Von Festenberg N. 2013 The hitchhiker’s guide to adaptive dynamics. *Games* **4**, 304–328. (doi:10.3390/g4030304)
- Restif O, Grenfell BT. 2006 Integrating life history and cross-immunity into the evolutionary dynamics of pathogens. *Proc. R. Soc. B* **273**, 409–416. (doi:10.1098/rspb.2005.3335)
- Lloyd AL, May RM. 1996 Spatial heterogeneity in epidemic models. *J. Theor. Biol.* **179**, 1–11. (doi:10.1006/jtbi.1996.0042)
- Tilman D. 1982 *Resource competition and community structure*. Princeton, NJ: Princeton University Press.
- Gandon S, Mackinnon MJ, Nee S, Read AF. 2001 Imperfect vaccines and the evolution of pathogen virulence. *Nature* **414**, 751–756. (doi:10.1038/414751a)

40. McCallum H, Barlow N, Hone J. 2001 How should pathogen transmission be modelled?. *Trends Ecol. Evol.* **16**, 295–300. (doi:10.1016/S0169-5347(01)02144-9)
41. Saad-Roy CM, Grenfell BT, Levin SA, Pellis L, Stage HB, van den Driessche P, Wingreen NS. 2021 Superinfection and the evolution of an initial asymptomatic stage. *R. Soc. Open Sci.* **8**, 202212. (doi:10.1098/rsos.202212)
42. Dercole F, Ferriere R, Rinaldi S. 2002 Ecological bistability and evolutionary reversals under asymmetrical competition. *Evolution* **56**, 1081–1090. (doi:10.1111/j.0014-3820.2002.tb01422.x)
43. Rohani P, May RM, Hassell MP. 1996 Metapopulation and equilibrium stability: the effects of spatial structure. *J. Theor. Biol.* **181**, 97–109. (doi:10.1006/jtbi.1996.0118)