Evolutionary dynamics of phenotype-structured populations: from individual-level mechanisms to population-level consequences

Rebecca H. Chisholm, Tommaso Lorenzi, Laurent Desvillettes and Barry D. Hughes

Abstract. Epigenetic mechanisms are increasingly recognised as integral to the adaptation of species that face environmental changes. In particular, empirical work has provided important insights into the contribution of epigenetic mechanisms to the persistence of clonal species, from which a number of verbal explanations have emerged that are suited to logical testing by proof-of-concept mathematical models. Here, we present a stochastic agent-based model and a related deterministic integrodifferential equation model for the evolution of a phenotype-structured population composed of asexually-reproducing and competing organisms which are exposed to novel environmental conditions. This setting has relevance to the study of biological systems where colonising asexual populations must survive and rapidly adapt to hostile environments, like pathogenesis, invasion and tumour metastasis. We explore how evolution might proceed when epigenetic variation in gene expression can change the reproductive capacity of individuals within the population in the new environment. Simulations and analyses of our models clarify the conditions under which certain evolutionary paths are possible, and illustrate that whilst epigenetic mechanisms may facilitate adaptation in asexual species faced with environmental change, they can also lead to a type of "epigenetic load" and contribute to extinction. Moreover, our results offer a formal basis for the claim that constant environments favour individuals with low rates of stochastic phenotypic variation. Finally, our model provides a "proof of concept" of the verbal hypothesis that phenotypic stability is a key driver in rescuing the adaptive potential of an asexual lineage, and supports the notion that intense selection pressure can, to an extent, offset the deleterious effects of high phenotypic instability and biased epimutations, and steer an asexual population back from the brink of an evolutionary dead end.

1. Introduction

In the physical sciences, mathematical modelling can frequently be performed for systems for which the prevailing physical laws are uncontroversial, their mathematical structure is known, and relevant physical constants can be determined by processes that are not part of the modelling exercise. Precise quantitative predictions can often be made, and compared against observations of the real system that can also be made with decent accuracy. The example *par excellence* is the study of the energy levels of the hydrogen atom, but there is no shortage of other fine exemplars.

In the biological and medical sciences, the situation is seldom so satisfactory, and models can sometimes play a rather different role [1]. Mathematical models can be used to explore whether qualitative verbal explanations are actually consistent with simple model implementations and lead to conclusions with broad structural stability under parameter changes, or a suite of mathematical models capturing in simple ways different plausible views of the system of interest can be explored to see which postulated qualitative attributes of a system are consistent with the observed behaviour.

Models of evolution typically focus on allelic variation and Mendelian inheritance as the source of heritable novel phenotypic variants. However, accumulating evidence suggests that there are other nongenetic forms of inheritance, including the transmission of behaviours [2], a modified niche [3] and epigenetic mechanisms such as differential DNA-methylation, histone modifications, and microRNAs. Epigenetic mechanisms are increasingly implicated in the successful adaptation of species which face environmental change. For instance, the invasion of cancer cells, which occurs in a progressively malignant tumour niche [4], has been linked to epigenetic disruptions [5]. Moreover, the silencing of tumour suppressor gene expression by promoter hypermethylation at CpG-rich islands is common among several human malignancies [6, 7]. Epigenetic mechanisms have been implicated in acquired drug-tolerance of cancer cells during *in vitro* experiments [8–10]. They are also recognised as key drivers in the success of invasive clonal plant species [11]. In fact, a recent empirical study of the methylomes (DNA methylation patterns) of invasive plant populations characterised by low genotypic diversity, and spread over a large geographic area, reported correlations between epigenetic differentiation and habitat [12], suggesting an epigenetic role in adaptation in this system. So it may be the case in a variety of contexts that epigenetic mechanisms increase a population's potential for adaptation, particularly in the face of rapid environmental change. A corollary of these ideas is that the stability of epialleles in asexual populations is crucial for successful epigenetic adaptation [13]. However, it remains unclear what level of epimutation stability is required for selection-based effects to occur in organisms faced with novel selective pressures, and when these selection-based effects are sufficient to avoid extinction.

In this paper, we present a stochastic agent-based model and a related deterministic integrodifferential equation model for the evolution of a phenotype-structured population composed of asexually-reproducing and competing organisms which are exposed to novel environmental conditions. At this stage, our model does not contain spatial structure, although it could be extended to include spatially distributed phenotypic attributes and migration of individuals at the cost of significantly increased complexity. Despite this simplification, analogous models have already been proven to be capable of shedding some light on the mechanisms which underlie patterns of evolution and adaptation in asexual populations [14–19]. From the mathematical point of view, our work follows earlier papers on the derivation of deterministic mesoscopic models from stochastic agent-based models [20–25] and the analysis of integrodifferential equations that arise in models of evolutionary dynamics within phenotype-structured populations [26–33].

In Section 2 we develop the agent-based model and explore its behaviour through simulations. Initially we assume that the agents (individual organisms) are genetically identical, but of variable phenotype (Section 2.1), before probing the more complex case in which there are several competing subpopulations ("strains", here, for convenience of terminology), each with distributed phenotypes (Section 2.2). Important attributes of the agent-based model are (i) the dependence of the reproduction rate of agents on their phenotype; (ii) stochastic variation in phenotype, which may include both an unbiased random component and a systematic drift component; and (iii) death by competition. We discuss in the light of representative simulations how the organisms can maladapt, in the sense that the most prevalent phenotype is not the one which reproduces most readily, and how under certain conditions the organisms can be driven to extinction. We also demonstrate through simulations that in the multiple-strain version of the model, a single strain (while retaining its own phenotypic spread) can drive all other strains to extinction, but cases also arise in which all strains are doomed to extinction.

In Section 3 we show how we can associate with the discrete stochastic agent-based model a deterministic integrodifferential equation model, and we identify all equilibrium solutions of the model. The correspondence between the discrete stochastic and continuous deterministic models discussed in Section 3.1 comes through a formal limit as the discrete stochastic model's time increment and its phenotype-space length scale approach zero in a constrained manner, together with a mean-field treatment of death under competition. The integrodifferential equation model we obtain in this way is one that we have explored partially in an earlier paper [17], but its derivation from the agent-based model, and all of the rigorous analysis of the mathematical consequences of the integrodifferential

equation model, are new. The only possible equilibrium solutions, which are identified in Section 3.2, are related to the eigenfunctions of a second-order linear differential operator.

In Section 4 we establish a number of results concerning the time-evolution of solutions of the integrodifferential equation model. For the special case of Gaussian initial phenotype distributions, we establish in Section 4.1 the existence of corresponding time-dependent Gaussian solutions, for which the long-term asymptotic behaviour of the mean and variance can be explicitly determined, together with a number of results concerning the long-term fate of total strain populations. Results from our study of Gaussian solutions enable us to prove in Section 4.2 the existence for all subsequent times of solutions for arbitrary compact support initial phenotype distributions. In Section 4.3 we use an entirely different approach, based on the Laplace transform, to confirm asymptotic results we obtained earlier in the paper by other means, while also providing some stronger results.

The parameter space of the integrodifferential equation model can be explored completely, and we are able to draw some striking rigorous conclusions in Sections 3 and 4. As we demonstrate in Section 5, representative computations show a close relation between the quantitative predictions of the integrodifferential equation model and those of the agent-based model, except under circumstances where the population falls very low and stochastic effects are more evident. It would be interesting to pursue further in a fully rigorous manner the limiting behaviour of the stochastic agent-based model, but that somewhat ambitious enterprise is beyond the scope of the present paper.

We emphasise that we have not bench-marked our models against particular organisms, or identified our phenotypic coordinate with any particular epigenetic trait, although degree of DNA methylation might be proposed as an exemplar. Our work is presented with the perspective that mathematical modelling can complement more traditional methods of evolutionary biology research by capturing in abstract terms the implicit assumptions of a verbal hypothesis (and any hidden underlying assumptions) and clarifying the conditions under which certain evolutionary paths are possible [1]. The modelling framework potentially covers a wide range of ecological scenarios including the colonisation of a new niche by tumour cells following metastasis, the invasion of new host tissues by a pathogen (*e.g.*, the invasion of organs outside of the lungs by *Mycobacterium tuberculosis*), and the colonisation of new habitats by invasive and (primarily) asexually reproducing plant species. Initial conditions that are far from the long-time limiting behaviour embody an organism encountering a new environment to which it is not currently phenotypically well-adapted, a ubiquitous scenario in biology and medicine.

Simulations and analyses of our model clarify the conditions under which successful adaptation of the colonising asexual species is possible, and illustrate that whilst epigenetic mechanisms may facilitate adaptation in asexual species faced with environmental change, they can also lead to a type of "epigenetic load" and contribute to extinction. Moreover, our results offer a formal basis for the claim that constant environments favour individuals with low rates of stochastic phenotypic variation. Finally, our model provides a "proof of concept" of the verbal hypothesis that phenotypic stability is a key driver in rescuing the adaptive potential of an asexual lineage, and supports the notion that intense selection pressure can, to an extent, offset the deleterious effects of high phenotypic instability and biased epimutations, and steer an asexual population back from the brink of an evolutionary dead end.

Also, as we show, the agent-based and integrodifferential equation models raise interesting mathematical questions that are worthy of investigation, most of which we are able to answer rigorously for the integrodifferential equation model.

2. An agent-based model of epigenetic evolution

In the agent-based modelling framework, agents represent individual asexual organisms in the colonising population, and are characterised by a phenotype which we associate with the level of expression of the fitness-regulating gene. Agents undergo stochastic changes in phenotype, and react to the environment according to their phenotype and a set of rules which we describe below. In Section 2.1 we introduce our agent-based model in the simplest case when there is only a single strain present

with a distributed phenotype, and report some representative simulations, from which interesting observations with biological implications emerge. We extend our model to the case of several strains in Section 2.2. For brevity, $\mathbb{N}_0$ denotes the natural numbers with 0 included, that is, the non-negative integers, while as usual $\mathbb{Z}$ denotes the set of integers.

2.1. One genetic strain with phenotypic diversity

Stochastic epigenetic variation is modelled by allowing agents in the population to undergo a discrete-time branching random walk in a discretised one-dimensional phenotype space. In this setting, each agent occupies a position on a one-dimensional lattice with lattice spacing Δ. If $i \in \mathbb{Z}$ denotes a phenotype lattice site, the level of normalised gene expression of an agent at site i is given by $x_i = i\Delta \in [0, 1]$. A generic time-step is indexed by $n \in \mathbb{N}_0$, and the corresponding time is $t = \tau n$, where τ is the duration of each time step. We define the random variables $N_n(i) \in \mathbb{N}_0$ and $N_n \in \mathbb{N}_0$ to be, respectively, the number of agents at phenotype lattice site i, and the total number of agents alive in the system at time step n, so that $N_n = \sum_i N_n(i)$. The mean phenotypic variant μ_n (the average level of expression of the fitness-regulating gene) of the population at the nth time step is defined by

$$\mu_n := \frac{1}{N_n} \sum_i x_i N_n(i). \tag{1}$$

Accordingly, μ_0 is the mean phenotypic variant of the initial population (the average level of expression of the fitness-regulating gene of the colonising population), and we define the equilibrium mean phenotypic variant by

$$\overline{\mu} := \lim_{n\to\infty} \mu_n. \tag{2}$$

In the modern two-step view of evolution [34, 35], individuals with novel phenotypes are first introduced into a population, which then undergoes natural selection—individuals survive, reproduce and die according to their fitness. Accordingly, we model evolution as a two-step process. During each time step of the agent-based model, agents simultaneously undergo two operations to simulate first the introduction of novel phenotypic variants into the population due to stochastic epimutations, followed by the increase in frequency of variants characterised by a greater fitness, due to natural selection (see the algorithm pictured in Fig. 1).

More specifically, stochastic epimutations are modelled during the first step by allowing each agent with phenotype given by $x_i = i\Delta$ to update its phenotype according to the random walk. To simulate a stochastic change in phenotype or level of gene expression, agents simultaneously move a distance Δ in the phenotype space according to movement probabilities. In terms of movement on the lattice, we consider agents as independent of each other, so that the number of agents occupying any particular lattice site in phenotype space is unconstrained. For an agent at site i, let P_{L} and P_{R} denote the respective left and right movement probabilities into the nearest neighbour sites ($i - 1$ or $i + 1$). These movements correspond, respectively, to an incremental decrease or increase in the level of fitness-regulating gene expression. Hence, $P = P_{\mathrm{L}} + P_{\mathrm{R}}$ is the probability that an agent will move a distance Δ within time τ, given that the agent is at site i at time step n. Furthermore, $1 - P_{\mathrm{L}} - P_{\mathrm{R}}$ is the probability that an agent's gene expression will remain unchanged during this time. We identify P/τ as the average rate of epimutations.

To account for the fact that epimutations may be biased towards either increased or decreased gene expression, we introduce a bias parameter $h \in [0, 1]$ and write

$$P_{\mathrm{L}} = \frac{P}{1+h} \quad \text{and} \quad P_{\mathrm{R}} = \frac{Ph}{1+h}. \tag{3}$$

The extreme cases $h = 0$ and $h = 1$ correspond, respectively, to a deterministic system in which all steps in phenotype space are in a specified direction (here, to the left, without loss of generality), and the most random system where steps to the left and right are equally probable.

Finally, we constrain the phenotype of each agent to the unit interval (that is, $x_i = \Delta i \in [0, 1]$ for all i) by implementing reflecting boundary conditions. In more detail, during a time-step, if an agent with phenotype $x_i = 1$ chooses to move right, then this move is aborted and the agent retains

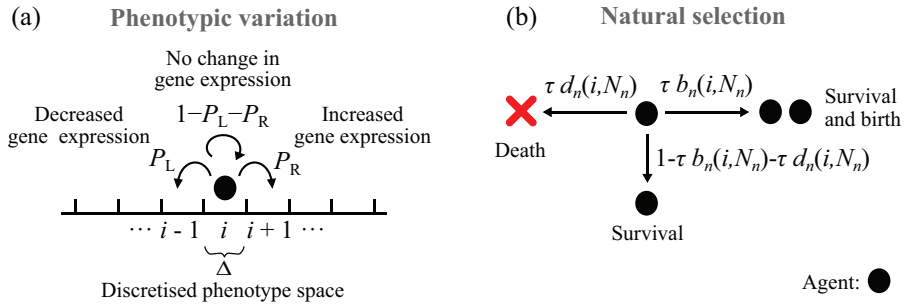

FIGURE 1. **Two-step algorithm of the agent-based model of stochastic epigenetic variation and natural selection in colonising asexual populations.** (a) Stochastic variation in phenotype due to epigenetic mechanisms is modelled as a random walk in the discretised phenotype space. (b) Natural selection is modelled as a stochastic birth and death process, where the phenotype of an agent determines the likelihood of proliferation and death. Phenotypic variation occurs before natural selection in the algorithm. Here, P_{L} and P_{R} denote the respective left and right movement probabilities into the nearest neighbour phenotype sites $i-1$ and $i+1$ (which correspond to a respective decrease and increase in the degree of expression of the fitness-regulating gene), $\tau b_n(i,N_n)$ is the probability an agent will proliferate during a time-step of length τ and $\tau d_n(i,N_n)$ is the probability an agent will die during a time-step of length τ, where N_n is the total number of agents alive in the system at the nth time step.

its phenotype $x_i = 1$ for the duration of the time-step. Analogously, if an agent with phenotype $x_i = 0$ chooses to move left, then this move is also aborted and the agent retains its phenotype $x_i = 0$ for the duration of the time-step.

After the phenotype-update step, each agent will simultaneously undergo the effects of selection. With probability $\tau b_n(i,N_n)$ the agent at site i will proliferate and produce two identical agents with the same phenotype $i\Delta$ (we assume that phenotype is heritable), while with probability $\tau d_n(i,N_n)$ the agent will die, and with probability $1-\tau[b_n(i,N_n)+d_n(i,N_n)]$ the agent will remain quiescent. Notice that these probabilities may depend not only on the level of gene expression of each agent, but also on the total number of agents in the system N_n. Therefore, the agents are not strictly independent of each other—we consider the colonising asexual population to be a system of quasi-interacting random walkers.

We focus on a population that has a smooth fitness landscape, and where there is only one possible phenotypic variant (degree of gene expression) that maximises fitness. In particular, we consider the case where the death probability $d_n(i,N_n)$ is an increasing function of N_n and invariant with respect to i, and where the proliferation function $b_n(i,N_n)$ is concave with respect to i, and invariant with respect to N_n, such that

$$d_n(i,N_n) := \kappa N_n \quad \text{and} \quad b_n(i) := \gamma - \epsilon\Delta^2(i-i^*)^2. \tag{4}$$

These assumptions replicate a scenario where there exists an optimal degree of gene expression, and hence one possible phenotypic variant with $x^* = i^*\Delta$, that maximises an organism's fitness through maximising its probability of proliferation, and where competition for space and resources amongst organisms in the population acts to increase the probability of death, irrespective of phenotype. Here, the choice of κ, γ and ϵ determine the carrying capacity of the system (*i.e.*, the equilibrium size of the population). We note that a consequence of the simple functional form used is that, depending on the values of the parameters γ and ϵ, sufficiently large deviations from the fittest phenotype render the birth rate negative, corresponding to failure to reproduce and effectively adding an additional component to the death rate.

Figure 2: $\gamma = 100$, $\epsilon = 100$, $\kappa = 10^{-2}$, $N_0 = 500$, $P = 0.1$;
(A),(B),(D) $\mu_0 = 0.5$; (C) $\mu_0 = j/10$, $j \in \{0, 1, 2, ..., 10\}$;
(A),(C),(D) $h = 0.6$; (B) $h = 1$.

Figure 3: $\mu_0 = 0.5$, $\gamma = 100$,
(A) $h = j/10$, $j \in \{0, 1, 2, ..., 10\}$, $\epsilon = 100$, $\kappa = 10^{-2}$, $N_0 = 500$, $P = 0.1$;
(B) $P \in \{0.01, 0.05, 0.1, 0.5, 1\}$, $\epsilon = 100$, $\kappa = 10^{-2}$, $h = 0.6$, $N_0 = 500$;
(C) $\epsilon \in \{5, 10, 20, 50, 100, 150, 250\}$, $P = 0.1$, $h = 0.6$, $N_0 = 500$, $\kappa = 10^{-2}$;
(D) $\epsilon = 100$, $P = 0.1$, $h = 0.6$, $\kappa \in \{0.002, 0.01, 0.02, 0.1, 0.2, 1, 2\}$,
$N_0 = 500$ for $\kappa < 1$ and $N_0 = 50$ for $\kappa \geq 1$.

TABLE 1. Simulation parameters for Figures 2 and 3. Parameters common to both figures are given in the main text above Observation 1.

We present in Figures 2 and 3 representative results obtained from MATLAB simulations of the model with the death and proliferation probabilities (4). We also considered the case where the death probability is constant, and where the proliferation function $b_n(i, N_n)$ is concave with respect to i, and a decreasing function of N_n. In this alternative setup, we assume that competition for space and resources amongst agents in the population decreases the probability of proliferation rather than increases the probability of death. However, we achieved qualitatively similar results to those from the aforementioned case (results not presented here).

Before we address the biologically relevant observations that may be made from Figures 2 and 3, it is convenient for brevity to summarize here simulation parameters shared in Figures 2 and 3. For all simulations, we set the fittest phenotypic variant to be $x^* = 0.9$, the time step $\tau = 10^{-3}$ and the lattice spacing $\Delta = 10^{-2}$. The initial phenotype distribution was uniformly distributed across the interval $x \in [\mu_0 - 0.05, \mu_0 + 0.05]$ for the homogeneous case, and $x \in [0, 1]$ for the heterogeneous case. Additional parameter details can be found in Table 1. To facilitate comparisons between the agent-based model and the related integrodifferential equation model discussed later, we introduce the parameters

$$\alpha = \frac{\Delta(P_{\mathrm{R}} - P_{\mathrm{L}})}{\tau} = \frac{\Delta P(h-1)}{\tau(h+1)}, \quad \beta = \frac{\Delta^2(P_{\mathrm{L}} + P_{\mathrm{R}})}{2\tau} = \frac{\Delta^2 P}{2\tau}, \quad M = \frac{1}{\kappa}\Big[\gamma - (\epsilon\beta)^{1/2} - \frac{\alpha^2}{4\beta}\Big]. \tag{5}$$

The role of $\max(M, 0)$ as a carrying capacity for the model will emerge from the analysis of the integrodifferential equation model.

Observation 1—Epigenetic variation facilitates adaptation in asexuals and a bias in the direction of epimutations can influence the degree of adaptation. To replicate the colonisation of an asexual species in a new habitat with novel selective forces, we simulate the agent-based model so that the optimal level of gene expression in the new environment ($x^* = 0.9$) differs from the average level of gene expression of the colonising population ($\mu_0 = 0.5$), which we assume was well-adapted to its previous environment. A comparison between simulations when there is no bias in the direction of epimutations (so that the bias parameter $h = 1$), and when there is a bias towards epimutations (so that $h < 1$) is presented in Fig. 2(a)–(b).

When there is a bias in the direction of epimutations [Panel (a)], the population fails to adapt to the reproductively fittest phenotypic trait variant $x^* = 0.9$ (indicated by the dashed-white line). However, epigenetic variation does allow the colonising population to establish an equilibrium phenotype distribution with mean equilibrium phenotypic trait variant $\overline{\mu} \approx 0.7$. On the other hand, for the case when there are unbiased epimutations [Panel (b)], the population quickly adapts to the phenotypic trait variant that corresponds to the highest fitness (so that $\overline{\mu} = x^*$).

These results are robust with respect to the value of the mean phenotypic variant μ_0 of the initial homogeneous phenotypic distribution [see Fig. 2(c)], suggesting that epigenetic variation can facilitate adaptation in asexuals faced with novel selective pressures.

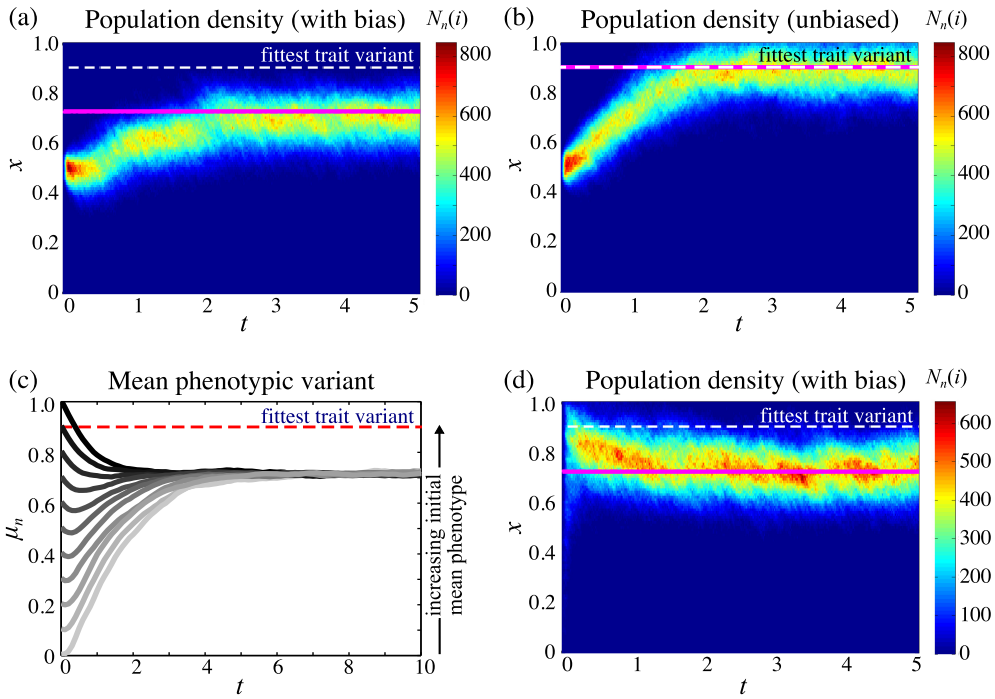

FIGURE 2. **Bias in the direction of epimutations can influence the outcome of evolution.** We present simulation results from the agent-based model showing (a),(b),(d) the population density $N_n(i)$ evolving in time in a single simulation and (c) the average over 10 simulations of the mean phenotypic variant as it evolves in time. (a) For the case when there is a bias in the direction of epimutations (with bias parameter $h = 0.6$), the population fails to adapt to the fittest phenotypic trait variant ($x^* = 0.9$) from an initially homogenous phenotypic distribution centred at $x = 0.5$. Instead, the equilibrium phenotypic trait distribution centres around $x = 0.7$. (b) For the case when there are unbiased epimutations ($h = 1$), the population quickly adapts to the phenotypic trait variant that has the highest fitness. Now, the equilibrium phenotypic trait distribution is centred around $x = 0.9$. (c) When the bias parameter and the rate of epimutations are fixed ($h = 0.6$ and $P = 0.1$), changing the value of the mean phenotypic variant of the initial homogeneous population (*i.e.*, μ_0) does not affect the equilibrium phenotypic distribution. (d) When the initial population is heterogeneous with respect to phenotype, and when there is a bias in the direction of epimutations (with $h = 0.6$), again the population fails to adapt to the fittest phenotypic trait variant ($x^* = 0.9$). Instead, the phenotypic trait distribution centres around $x = 0.7$. The magenta lines correspond to the estimate for the equilibrium mean phenotypic variant from the analogous PDE model [equation (21)], and the dashed lines [white in (a), (b) and (d); red in (c)] correspond to the fittest phenotypic trait variant x^* in the fitness landscape.

Observation 2—Epigenetic diversity in the colonising population does not neutralise the effect of biased epimutations. In a previous work [35] it was suggested that if an initial population is heterogeneous, so that all possible phenotypic variants are present, then a bias in the introduction of phenotypic variation will have no effect on the evolutionary outcome. To test whether this is the case

in our model, instead of starting from a relatively homogeneous initial phenotypic distribution (as we did above), now we begin from a heterogeneous population, and track the evolution of the phenotypic distribution through time when there are biased epimutations (with $h < 1$). Our results, displayed in Fig. 2(d), show that, analogous to the more homogenous case presented in Fig. 2(a), the colonising population fails to adapt to the fittest phenotypic trait variant $x^* = 0.9$ (indicated by the dashed white line). Instead, the equilibrium phenotypic trait distribution centres around $\overline{\mu} \approx 0.7$. Therefore, our simulations suggest that the initial presence of phenotype heterogeneity does not neutralise the adaptive effect of biased epimutations.

Observation 3—The adaptive potential of colonising asexual populations increases with the stability of epimutations. Next we compare averages of the evolution of the mean phenotypic variant μ_n of the population over ten simulations of the agent-based model while holding all parameters constant except for the rate of epimutations P. Predictably, we find that increasing the instability of epimutations (by increasing P) acts to shift the mean equilibrium value $\overline{\mu}$ of the phenotype of the population farther away from the fittest phenotypic trait variant x^* [see Fig. 3(b)]. This is also the case when we increase the strength of bias h (by decreasing h) and fix all other parameters [see Fig. 3(a)]. Therefore, our model predicts that the adaptive potential of colonising asexual populations decreases as epimutations become increasingly unstable and biased.

Observation 4—Strong natural selection can overcome the effects of epimutation instability and bias. To investigate whether natural selection can overcome the effects of epimutaion bias and instability, we compare averages of the evolution of the mean phenotypic variant μ_n of the population over twenty simulations of the agent-based model while holding all parameters constant except for ϵ, which controls the effect that an agent's phenotype has on its reproduction rate [see definitions (4)]. We find that increasing the strength of natural selection (by increasing ϵ) acts to shift the mean equilibrium phenotype $\overline{\mu}$ closer to the fittest phenotypic trait variant x^* [see Fig. 3(c)]. Therefore, our model predicts that natural selection can, to an extent, neutralise the effects of epimutation instability and bias when it is sufficiently strong.

Observation 5—Epimutation bias has the greatest effect on small populations. Next we investigate the effect of the population size on evolutionary outcome in the presence of biased epimutations. We compare averages of the evolution of the mean phenotypic variant of the population over twenty simulations of the agent-based model, while holding all parameters constant except for the carrying capacity of the population [which we vary by changing the parameter κ, see definitions (4)]. We find that, for large populations ($N_n \gtrsim 10^4$), the population size has no effect on evolutionary outcome. However, when the population size is small ($N_n \lesssim 10^3$), decreasing the population size shifts the equilibrium value of the mean phenotypic variant of the population further away from the fittest phenotypic trait variant [see Fig. 3(d)].

Therefore, our model predicts that for smaller populations, epimutation bias has an increasing effect on evolutionary outcome as the population size decreases, while for larger populations, varying the population size does not alter the effect that bias has on evolutionary outcome. An analogy can be drawn between this result and the concept of genetic load [36], where the accumulation of deleterious genetic mutations occurs in small populations, despite the action of purifying selection.

2.2. Several genetic strains, each with phenotypic diversity

To extend the agent-based model of Section 2.1 in a natural way to a model with a finite number N of strains, we associate with each strain an integer j, and agents in the model represent individual members of the competing strains and are characterised by a phenotypic character $x_j \in [0, 1]$ which again represents the level of expression of a particular fitness-regulating gene that varies due to stochastic epimutations. As before, evolution is simulated in discrete time (with time-step length τ) as a two-step process with the first part of each time step (epimutation) implemented as a random walk in a discretised one-dimensional phenotype space (with lattice spacing Δ such that $x_j = i\Delta$).

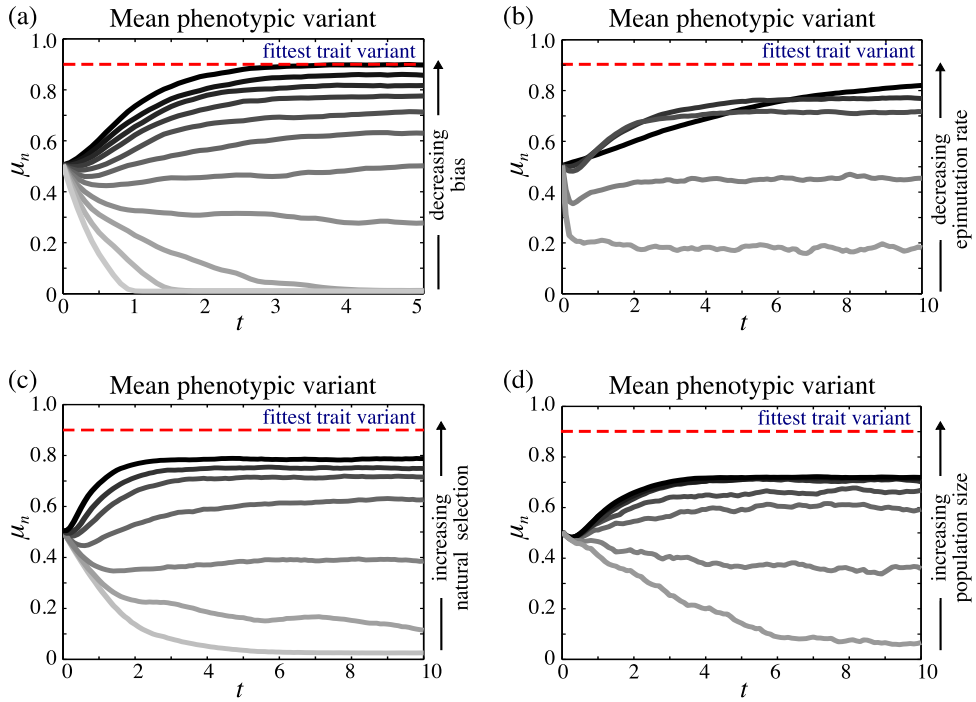

FIGURE 3. **The adaptive potential of colonising asexual populations is reduced as the stability of epimutations declines, as natural selective pressures become weaker, and as the strength of epimutation bias increases. These effects are exacerbated in small populations.** Simulation data of the agent-based model from an initially homogeneous population (with respect to phenotype), showing the average mean phenotypic variant as it evolves in time. When all other parameters remain fixed, (a) increasing the strength of the bias in epimutations (*i.e.*, decreasing the parameter h) pushes the equilibrium phenotypic distribution further away from the phenotypic variant that maximises fitness; (b) increasing the rate of epimutations (*i.e.*, increasing the parameter P) pushes the equilibrium phenotypic distribution further away from the phenotypic variant that maximises fitness; (c) increasing the strength of natural selection (by increasing ϵ) shifts the equilibrium value of the mean phenotypic variant of the population closer to the fittest phenotypic trait variant; (d) decreasing the population size (by decreasing the carrying capacity of the population) shifts the equilibrium value of the mean phenotypic variant of the population further away from the fittest phenotypic trait variant. Here, averages are calculated from either (a) 10 or (b)–(d) 20 simulations of the agent-based model. The red dashed lines correspond to the fittest phenotypic trait variant x^* in the fitness landscape.

The probabilities of stepping left or right during this element of the procedure are, respectively,

$$P_{\mathrm{L}}^j = \frac{P_j}{1+h_j}, \qquad P_{\mathrm{R}}^j = \frac{P_j h_j}{1+h_j}, \tag{6}$$

so an agent of strain j actually moves in phenotype space with probability P_j, pauses instead with probability $1-P_j$, and has a bias in moving measured by the parameter $h_j \in [0,1]$. To complete the update procedure, agents then simultaneously proliferate, die or remain quiescent with the respective probabilities

$$\tau\left[\gamma_j - \epsilon_j(x-x^*)^2\right], \quad \tau\kappa_j N_n \quad \text{and} \quad 1-\tau\left[\gamma_j - \epsilon_j(x-x^*)^2\right] - \tau\kappa_j N_n.$$

parameter set		Parameter Set 1		Parameter Set 2	
strain name		strain 1	strain 2	strain 1	strain 1
epimutation rate descriptor		slower	faster	slower	faster
phenotype lattice spacing	Δ	0.01	0.01	0.01	0.01
time increment	τ	0.001	0.001	0.0001	0.0001
initial number of agents	N	500	500	500	500
epimutation bias parameter	h_j	0.6	0.6	0.85	0.85
related epimutation bias parameter	α_j	-0.25	-0.50	-2.43	-4.05
epimutation rate	P	0.1	0.2	0.3	0.5
related epimutation rate	β_j	0.005	0.010	0.15	0.25
maximum proliferation rate	γ_j	100	100	10	10
fitness falloff parameter	ϵ_j	100	100	10	10
death rate	κ_j	0.01	0.01	9	9
mean phenotype at equilibrium	$\overline{x}_j$	0.323	0.250	-0.493	-0.782
carrying capacity	M_j	9617	9275	-0.1207	-0.8907

TABLE 2. Parameters used for two agent-based simulation studies of a two-strain system. Italicised entries in the left column refer to parameters for the related continuum model, matched to the agent-based system parameters using Eqs (7) or computed using Eqs (8) and (22) and rounded for display; all other entries are exact values.

The different strains compete for the same resources, so that the death rate depends on the total population N_n summed over all strains. We assume that new agents inherit the phenotypic character and strain of its parent.

As for the single-strain model, we introduce parameters that will facilitate later comparisons between the agent-based model and related integrodifferential equation:

$$\alpha_j = \frac{\Delta(P_{\mathrm{R}}^j - P_{\mathrm{L}}^j)}{\tau} = \frac{\Delta P_j(h_j - 1)}{\tau(h_j + 1)}, \qquad \beta_j = \frac{\Delta^2(P_{\mathrm{L}}^j + P_{\mathrm{R}}^j)}{2\tau} = \frac{\Delta^2 P_j}{2\tau}, \tag{7}$$

and

$$M_j = \frac{1}{\kappa_j}\Big[\gamma_j - \frac{\alpha_j^2}{4\beta_j} - (\beta_j\epsilon_j)^{1/2}\Big]. \tag{8}$$

The significance of $M^* = \max_{1\le j\le N} M_j$ will be elucidated in Section 3.

We have performed some simulations to discover typical properties of the multistrain agent-based model, with the results shown in Figures 4 and 5. The phenotypic characters for each strain are constrained to the interval $0 \le x_j = i\Delta \le 1$, with reflecting boundary conditions for the random walk element of the procedure and fittest phenotype coordinate $x_j^* = i^*\Delta = 0.5$ for each strain. The initial condition used in each case for each strain was a uniform distribution over all discrete phenotype coordinates. Values of the parameters for the two parameter sets considered are given in Table 2.

Observation 6—In multistrain problems the typical outcomes are either indefinite survival of a single strain, or extinction of all strains. In Figure 4 we show the result of a simulation for Parameter Set 1, in which two strains with the same initial population (500 of each strain) interact and evolve from the uniform initial condition. As the initial populations are low, there is significant proliferation of both strains early in the simulation, with the strain that has a higher epimutation rate (strain 2, in panel (b) of Figure 4) briefly flourishing, before it is rapidly displaced by the strain with a lower epimutation rate (strain 1, in panel (a) of Figure 4), which evolves to a steady state with the most probable phenotype close to that predicted by the continuum model in the case in which only strain 1 were present (we establish this in Section 3 below). In terms of the parameters M_j, defined by Eq. (8), for the simulation in Figure 4, the victorious strain has $M_1 = 9617$, while the strain that goes extinct has $M_2 = 9275$. The simulation suggests to us that one strain will eventually attain total dominance, and as observed in the single-strain problem (Section 2.1), if there is drift present ($P_{\mathrm{L}} \neq P_{\mathrm{R}}$) then

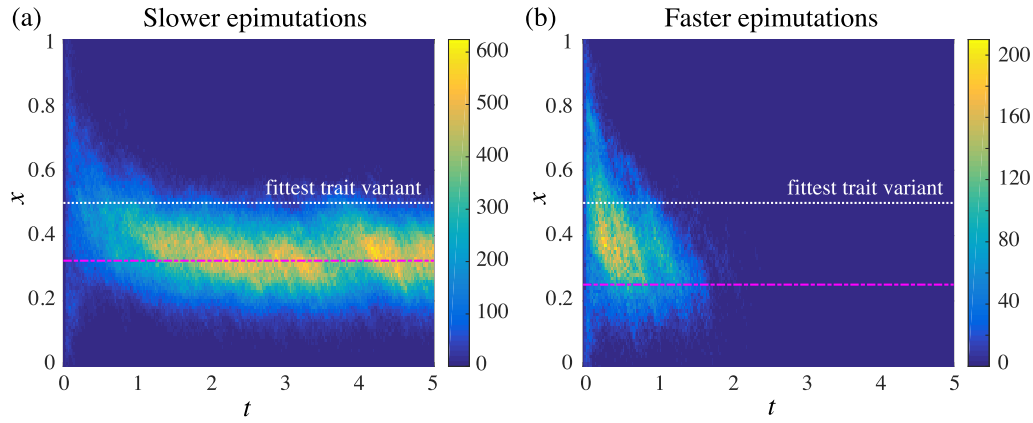

FIGURE 4. An agent-based simulation of two competing strains for Parameter Set 1 from Table 2. Initially there are 500 individuals of each species present, uniformly distributed over the discretized phenotype domain. The white broken line shows the fittest phenotype for reproduction ($x_j^* = 0.5$). The magenta broken line corresponds to the mean phenotype $\overline{x}_j$ predicted by the continuum model if only strain j were present initially. (a) (strain 1—"slower epimutations"): $\beta_1 = 0.005$, $\overline{x}_1 = 0.323$ and $M_1 = 9617$. (b) (strain 2—"faster epimutations"): $\beta_2 = 0.01$, $\overline{x}_2 = 0.250$ and $M_2 = 9275$.

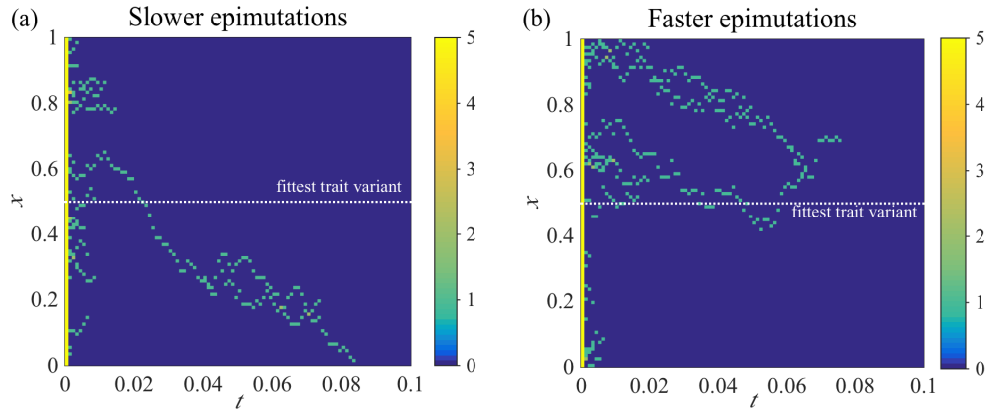

FIGURE 5. An agent-based simulation of two competing strains for Parameter Set 2 from Table 2. Initially there are 500 individuals of each species present, uniformly distributed over the discretized phenotype domain. The white broken line shows the fittest phenotype for reproduction ($x_j^* = 0.5$). (a) (strain 1—"slower epimutations"): $\beta_1 = 0.15$, $\overline{x}_1 = -0.493$ and $M_1 = -0.1207$. (b) (strain 2—"faster epimutations"): $\beta_2 = 0.25$, $\overline{x}_2 = -0.782$ and $M_2 = -0.8907$.

the most prevalent phenotype of the surviving strain is not the fittest phenotype for reproductive purposes.

In Figure 5 we show the results of simulations for Parameter Set 2. Again, both strains have initial population 500, uniformly distributed over phenotype. Neither the strain with the lower epimutation rate ($\beta_1 = 0.15$, $M_1 = -0.1207$) nor the strain with the higher epimutation rate ($\beta_2 = 0.25$, $M_2 = -0.8907$) survive, suggesting that if no strain has $M_j > 0$, then all strains are doomed to extinction. We explore the robustness of these conclusions using the continuum model in Section 3 (where all possible equilibrium states are determined) and Section 4, where we discover which equilibrium states

are accessible from given initial conditions and determining the time-dependence of the approach to equilibrium.

3. A continuum model of epigenetic evolution

Experience in a variety of contexts has demonstrated the value of relating discrete-time, discrete-space, agent-based stochastic models to deterministic continuum models [22–25]. Apart from a few special fortuitous examples, such as the unbiased simple exclusion process on a regular lattice, the correspondence is approximate rather than exact, and is exhibited using some form of "mean-field" argument, in which correlations between the locations of distinct agents are either ignored or are treated in some truncated or other approximate way [37]. Often numerical solutions of the mean-field continuum models reproduce well results obtained by averaging over many realizations of the agent-based stochastic model [24, 25] and there is the prospect of obtaining analytic results for the continuum model, enabling parameter space to be explored fully and asymptotic behaviour extracted. As we shall demonstrate, this is so for the agent-based mode of epigenetic evolution that we have discussed in Section 2.

3.1. Obtaining the integrodifferential equation model

We begin by considering the case in which there is a single strain present, which has a distributed phenotype. In the analysis that follows, we continue with the notation established in Section 2, so that $i \in \mathbb{Z}$ is a phenotype position coordinate and $n \in \mathbb{N}_0$ is a discrete time point, and the random variable $N_n(i) \in \mathbb{N}_0$ is the number of agents of phenotype i at discrete time n. We denote by $I(\omega, p)$ an indicator random variable associated with an event ω involving a single agent, which takes the value 1 with probability p and is zero otherwise. The event ω may be subscripted. In a sum over k involving ω_k, the events for distinct values of k are assumed independent. Even if bearing the same value of the subscript, events appearing in summations with different dummy indices of summation are also assumed independent (an abuse of notation that avoids ugly symbolism). We denote by $N_n^{\ddagger}(i)$ the number of agents with phenotype i when the agents have just updated their phenotype, before any reproduction or death has occurred. Then

$$N_n^{\ddagger}(i) = N_n(i) + \sum_{\ell=1}^{N_n(i+1)} I(\omega_\ell, P_{\mathrm{L}}) + \sum_{r=1}^{N_n(i-1)} I(\omega_r, P_{\mathrm{R}}) - \sum_{k=1}^{N_n(i)} I(\omega_k, P_{\mathrm{L}} + P_{\mathrm{R}}),$$

where the sum over ℓ accounts for the movement choices of agents at phenotype $i+1$ (who must move left to become of phenotype i), the sum over r accounts for the movement choices of agents at phenotype $i-1$ (who must move right to become of phenotype i) and the sum over k accounts for those agents currently of phenotype i that mutate. Taking expectation (denoted by angle brackets) we find that

$$\langle N_n^{\ddagger}(i)\rangle = \langle N_n(i)\rangle + P_{\mathrm{L}}\langle N_n(i+1)\rangle + P_{\mathrm{R}}\langle N_n(i-1)\rangle - (P_{\mathrm{L}} + P_{\mathrm{R}})\langle N_n(i)\rangle. \tag{9}$$

We now consider the reproduction and death events for agents of current phenotype i. For each such agent we assume the agent either:

divides into two agents with probability $\tau b_n(i)$;
dies with probability $\tau\kappa F(\sum_j N_n^{\ddagger}(j))$;
does nothing (and so remains in place) with probability $1 - \tau b_n(i) - \tau\kappa F(\sum_j N_n^{\ddagger}(j))$.

The time step τ is assumed to be small enough that the three probabilities lie, as required, in the interval $[0, 1]$.

We remark that $\sum_j N_n^{\ddagger}(j)$ is a random variable for which the expected value is taken as $\varrho(t)$ in the continuum limit, and our primary interest is in the case $F(\varrho) = \varrho$. If $N_n^{\ddagger}(i) = \nu$ and $\sum_j N_n^{\ddagger}(j) = \theta$,

then the expected number of agents of phenotype i that replace a given agent after the reproduction and death stage at time n is

$$2\tau b_n(i) + 1 - \tau b_n(i) - \tau\kappa F(\theta) + 0 = 1 + \tau b_n(i) - \tau\kappa F(\theta).$$

Hence we find that

$$\langle N_{n+1}(i)\rangle = \sum_\mu \sum_\nu \nu\big[1 + \tau b_n(i) - \tau\kappa F(\theta)\big]\,\Pr\{N_n^\ddagger(i) = \nu,\ \sum_j N_n^\ddagger(j) = \theta\}. \tag{10}$$

We now invoke a single probabilistic approximation, which is of the mean-field type frequently used for the analysis of agent-based systems (see, e.g., [37]): in the summand we replace $F(\theta) = F(\sum_j N_n^\ddagger(j))$ by $F(\langle\sum_j N_n^\ddagger(j)\rangle)$, and we note that as

$$\langle\sum_j N_n^\ddagger(j)\rangle = \langle\sum_j N_n(j)\rangle = \varrho \tag{11}$$

(since the transition from $N_n(j)$ to $N_n^\ddagger(j)$ in the simulation does not change total agent numbers) we have

$$\langle N_{n+1}(i)\rangle \approx \big[1 + \tau b_n(i) - \tau\kappa F(\varrho)\big]\langle N_n^\ddagger(i)\rangle. \tag{12}$$

If we eliminate $\langle N_n^\ddagger(i)\rangle$ from Eqs (9) and (12) and take $t = n\tau$, $x = \Delta i$ and $\Delta c(x,t) = \langle N_n(i)\rangle$ we have

$$\begin{aligned} c(x,t+\tau) &= \big[1 + \tau b_n(i) - \tau\kappa F(\varrho)\big]\big[c(x,t) + P_{\mathrm{L}}c(x+\Delta,t) + P_{\mathrm{R}}c(x-\Delta,t) - (P_{\mathrm{L}}+P_{\mathrm{R}})c(x,t)\big] \\ &= \big[1 + \tau b_n(i) - \tau\kappa F(\varrho)\big]\Big[c(x,t) + \Delta(P_{\mathrm{L}} - P_{\mathrm{R}})\frac{\partial c}{\partial x} + \frac{\Delta^2(P_{\mathrm{L}}+P_{\mathrm{R}})}{2}\frac{\partial^2 c}{\partial x^2} + o(\Delta^2)\Big]. \end{aligned}$$

Hence, writing $b_n(i) = b(x,t)$,

$$\begin{aligned} \frac{c(x,t+\tau) - c(x,t)}{\tau} &= \frac{\Delta(P_{\mathrm{L}} - P_{\mathrm{R}})}{\tau}\frac{\partial c}{\partial x} + \frac{\Delta^2(P_{\mathrm{L}}+P_{\mathrm{R}})}{2\tau}\frac{\partial^2 c}{\partial x^2} + o\Big(\frac{\Delta^2}{\tau}\Big) \\ &\quad + \frac{\tau b(x,t) - \tau\kappa F(\varrho)}{\tau}\Big[c(x,t) + \Delta(P_{\mathrm{L}} - P_{\mathrm{R}})\frac{\partial c}{\partial x} + \frac{\Delta^2(P_{\mathrm{L}}+P_{\mathrm{R}})}{2}\frac{\partial^2 c}{\partial x^2} + o(\Delta^2)\Big]. \end{aligned}$$

To facilitate the emergence of a sensible continuum limit we assume that

$$\lim_{\Delta,\tau\to 0}\frac{\Delta^2(P_{\mathrm{L}}+P_{\mathrm{R}})}{2\tau} = \beta, \quad \lim_{\Delta,\tau\to 0}\frac{\Delta(P_{\mathrm{R}} - P_{\mathrm{L}})}{\tau} = \alpha \tag{13}$$

and so

$$\frac{\partial c}{\partial t} = \beta\frac{\partial^2 c}{\partial x^2} - \alpha\frac{\partial c}{\partial x} + \Big\{\lim_{\Delta,\tau\to 0}\big[b(x) - \kappa F(\varrho)\big]\Big\}c, \tag{14}$$

where

$$\varrho(t) = \int_{-\infty}^{\infty} c(x,t)dx. \tag{15}$$

For our simulations $F(Z) = Z$ and $b_n(i) = \gamma - \epsilon\Delta^2(i - i^*)^2$, so $b(x,t) = \gamma - \epsilon(x - x^*)^2$, and we arrive at the integrodifferential equation

$$\frac{\partial c}{\partial t} = \beta\frac{\partial^2 c}{\partial x^2} - \alpha\frac{\partial c}{\partial x} + \Big[\gamma - \epsilon(x - x^*)^2 - \kappa\varrho\Big]c. \tag{16}$$

The preceding analysis should be performed on a finite interval initially, with τ chosen small enough that quantities interpreted as probabilities fall in the interval $(0,1)$. However, if we extend the interval to $-\infty < x < \infty$, then the birth contribution term $[\gamma - \epsilon(x - x^*)^2]c$ becomes a strong death term for $|x - x^*| \gg (\gamma/\epsilon)^{1/2}$, mimicking absorbing boundaries on a finite interval.

Lorenzi *et al.* [17] have studied the continuum model (15)–(16) we have just derived (and taking $-\infty < x < \infty$) in the case where the fittest phenotype x^* is a periodic function of time (corresponding, for example, to seasonal variation). They have demonstrated the existence of nontrivial periodic solutions under certain parameter conditions. These periodic solutions oscillate with the same frequency as the fittest phenotype, except under very special circumstances requiring temporary symmetry and

the absence of systematic phenotype drift, in which case the total population oscillates twice as fast as the fittest phenotype. We shall not address time-dependence of x^* here.

The same arguments can be used to associate a deterministic continuum model with the multistrain agent-based model of Section 2.2. If we index the parameters and functions associated with a given strain using a subscript $j \in \{1, 2, \ldots, N\}$, then the evolution equation for strain j is

$$\frac{\partial c_j}{\partial t} + \alpha_j \frac{\partial c_j}{\partial x} = \beta_j \frac{\partial^2 c_j}{\partial x^2} + [\gamma_j - \epsilon_j (x - x_j^*)^2 - \kappa_j \varrho] c_j, \tag{17}$$

where

$$\varrho(t) = \sum_j \varrho_j(t), \qquad \varrho_j(t) = \int_{-\infty}^{\infty} c_j(x,t)dx. \tag{18}$$

When working in the infinite interval, which we shall always do for our integrodifferential equation models, there is no loss of generality in translating coordinates so that $x^* = 0$, so we do this in the subsequent analysis of the models, whether for a single strain system, or for multiple strains. Only solutions of the problems (15)–(16) or (17)–(18) for which

$$c(x,t) \geq 0, \quad c(x,t) \to 0 \text{ as } x \to \pm\infty, \quad \frac{\partial^n}{\partial x^n} c(x,t) \to 0 \text{ as } x \to \pm\infty \text{ for all } n \in \mathbb{N}, \tag{19}$$

or

$$c_j(x,t) \geq 0, \quad c_j(x,t) \to 0 \text{ as } x \to \pm\infty, \quad \frac{\partial^n}{\partial x^n} c_j(x,t) \to 0 \text{ as } x \to \pm\infty \text{ for all } n \in \mathbb{N}, \tag{20}$$

respectively, are relevant in the context of the model.

3.2. Preliminary observations on the model

As we shall discuss briefly at the end of this section and more fully in Section 5, we have found a close quantitative correspondence between simulations of the agent-based model on a finite phenotype interval and numerical solutions of the corresponding integrodifferential equation model with $x \in (-\infty, \infty)$. Although the link between the agent-based and continuum models requires $\Delta \to 0^+$, $\tau \to 0^+$ and $h \to 1^-$, we have found that $\Delta = 10^{-2}$ and $\tau = 10^{-3}$ are usually small enough, even with significant bias per step, e.g., $h = 0.6$. A major advantage of the continuum model over the agent-based model is that the former enables conclusions to be drawn over the entire parameter space.

We postpone to Section 4 the detailed consideration of time-dependent solutions of the model: the issue of sufficient conditions for existence and uniqueness of solutions, and the long-term behaviour of any solutions, are matters of evident interest. Two major issues, however, can be dealt with immediately: the possibility of explosive population growth, and the possibility of nontrivial equilibrium solutions (that is, equilibrium solutions with a nonzero total population). We present the answers in the form of two lemmas. Both are stated for the multistrain case, but apply for the single-strain case when we take $N = 1$ with the redundant strain subscript discarded.

If solutions exist for all times $t > 0$, Lemma 3.1 rules out unbounded population growth.

Lemma 3.1. *If $\varrho(0) < \infty$ then, for all solutions of the problem (17)–(20), $\varrho(t)$ is bounded as $t \to \infty$.*

Proof of Lemma 3.1. Integrating Eq. (17) over the interval $-\infty < x < \infty$ and we find that

$$\begin{aligned} \varrho_j'(t) &= [\gamma_j - \kappa_j \varrho(t)]\varrho_j(t) - \epsilon_j \int_{-\infty}^{\infty} (x - x_j^*)^2 c_j(x,t)dx \\ &\leq [\gamma_j - \kappa_j \varrho(t)]\varrho_j(t) \leq [\max_\ell \gamma_\ell - \min_\ell \kappa_\ell \varrho(t)]\varrho_j(t). \end{aligned}$$

Summing over j gives $\varrho'(t) \leq [\max_\ell \gamma_\ell - \min_\ell \kappa_\ell \varrho(t)]\varrho(t)$ and so $\varrho'(t) \leq 0$ when $\varrho(t) \geq \max_\ell \gamma_l / \min_\ell \kappa_\ell$. It follows that if $\varrho(0) \geq \max_\ell \gamma_l / \min_\ell \kappa_\ell$, then $\varrho(t) \leq \varrho(0)$ for all $t > 0$. If $0 < \varrho(0) < \max_\ell \gamma_l / \min_\ell \kappa_\ell$, although $\varrho(t)$ may increase initially, it cannot increase beyond $\max_\ell \gamma_l / \min_\ell \kappa_\ell$. □

There is always a trivial equilibrium solution $c_j(x,t) \equiv 0$ $(1 \leq j \leq N)$. The question of whether nontrivial equilibrium solutions (for which at least one strain has nonzero population) exist is answered by the following lemma, in which the significance of the parameters M_j (or M in the single-strain case) is revealed. Loosely speaking, these parameters are potential carrying capacities for the corresponding

strains. Lemma 3.2 shows that any strain j with $M_j < 0$ cannot reach a nontrivial equilibrium, but when several strains have positive potential carrying capacities, only those strains with the largest potential carrying capacity can endure. The implications of the Lemma are discussed further after the proof.

Lemma 3.2. *Let* $M_j = \dfrac{1}{\kappa_j}\Big[\gamma_j - \dfrac{\alpha_j^2}{4\beta_j} - (\beta_j\epsilon_j)^{1/2}\Big]$ *for* $1 \le j \le N$. *Then:*

(a) *if* $M_j \le 0$ *for all* j, *then the only equilibrium solution is* $c_j(x,t) \equiv 0$ $(1 \le j \le N)$;
(b) *if* $M^* = \max_{1\le j\le N} M_j > 0$, *then*
  (i) *there exist equilibrium solutions in which at least one strain survives* (*that is, at least one strain* j *has nonzero population* $\varrho_j(\infty)$);
  (ii) *if strain* j *survives, then its phenotype distribution has the Gaussian form*

$$c_j(x,t) = \frac{\varrho_j(\infty)(\epsilon_j/\beta_j)^{1/4}}{(2\pi)^{1/2}} \exp\Big\{-\frac{1}{2}\Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2}\Big[x - x_j^* - \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}}\Big]^2\Big\};$$

  and the total population (*summed over all surviving strains*) *is* $\varrho(\infty) = M_j$;
  (iii) *if the values of* M_1, M_2, $\ldots M_N$ *are distinct, there are no equilibrium solutions in which more than one strain survives;*
  (iv) *if strains* j *and* ℓ *are both to survive it is necessary that* $M_j = M_\ell > 0$.

Proof of Lemma 3.2. By a change of phenotype coordinate,we can always centre the problem for each strain about the fittest phenotype. It therefore suffices to prove the stated results for the case $x_j^* = 0$ for all j.

Let $C_j(x)$ denote the equilibrium phenotype distribution of strain j and, with an innocent abuse of notation, write $\varrho_j(\infty)$ and $\varrho(\infty) = \sum_j \varrho_j(\infty)$ for the corresponding population of strain j and the total population of all strains, respectively. Write

$$C_j(x) = \exp\Big(\frac{\alpha_j x}{2\beta_j}\Big)Y_j(z), \quad z = \Big(\frac{4\epsilon_j}{\beta_j}\Big)^{1/4}x.$$

Then we find that $Y_j(z)$ satisfies the differential equation

$$Y''(z) - \Big(\frac{z^2}{4} + a_j\Big)Y(z) = 0, \quad a_j = \frac{\kappa_j}{2(\beta_j\epsilon_j)^{1/2}}\Big(\varrho(\infty) + \frac{\alpha_j^2}{4\beta_j\kappa_j} - \frac{\gamma_j}{\kappa_j}\Big).$$

Since Y satisfies Weber's equation, familiar from quantum mechanics and other contexts, we know [40,41] that it has solutions which are bounded for all z if and only if $a_j = -n - 1/2$, where n is a non-negative integer. These bounded solutions are the Gaussians $\exp(-z^2/4)$ multiplied by polynomials of degree n, which form an orthogonal set of functions, and so are everywhere non-negative if and only if $n = 0$. The existence of a nontrivial solution thus requires $a_j = -1/2$ and recalling the definition of M_j we find that for strain j to have a nontrivial equilibrium solution we require

$$\varrho(\infty) = M_j$$

and if this condition is met, then for some positive K_j (independent of x and t)

$$C_j(x) = K_j \exp\Big\{-\frac{1}{2}\Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2}\Big[x - \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}}\Big]^2\Big\}.$$

The constant K_j can be evaluated in terms of the strain population $\varrho_j(\infty)$ by integrating this equation. We find that

$$C_j(x) = \frac{\varrho_j(\infty)(\epsilon_j/\beta_j)^{1/4}}{(2\pi)^{1/2}} \exp\Big\{-\frac{1}{2}\Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2}\Big[x - \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}}\Big]^2\Big\}.$$

Since we need $\varrho(\infty) > 0$ if at least one strain is to survive, we see that if $M_j \le 0$ for all j, no strain can survive at equilibrium. This establishes the claims (a) and (b)–(ii).

We first note that if $M_\ell > 0$, then taking $\varrho_j(\infty) = 0$ for all $j \neq \ell$ and $\varrho_\ell(\infty) = M_\ell$, we have the equilibrium solution

$$c_j(x,t) = \begin{cases} 0, & j \neq \ell, \\ \dfrac{M_\ell(\epsilon_\ell/\beta_\ell)^{1/4}}{(2\pi)^{1/2}} \exp\Big\{-\dfrac{1}{2}\Big(\dfrac{\epsilon_\ell}{\beta_\ell}\Big)^{1/2}\Big[x - \dfrac{\alpha_\ell}{2(\beta_\ell\epsilon_\ell)^{1/2}}\Big]^2\Big\}, & j = \ell, \end{cases}$$

establishing claim (b)–(i). Next we observe that the necessary condition $M_j = \varrho(\infty)$ for survival of strain j shows that the survival of two strains is precluded if those strains have different values of the composite parameter M_j, and this completes the proof of claims (b)–(iii) and (b)–(iv). □

The analysis of equilibria by itself does not tell us which equilibrium solution, if any, is obtained in the long-time limit, starting from a given initial condition and we address this matter in Section 4. However, the equilibrium analysis already sheds light on the agent-based model simulations. For the single strain case, the only possible equilibrium solution for $M \leq 0$ has $\varrho(\infty) = 0$ (extinction). For $M > 0$, there is a unique nontrivial equilibrium solution, in which $\varrho(\infty) = M$, and this solution is Gaussian with mean

$$\overline{x} = x^* + \frac{\alpha}{2(\beta\epsilon)^{1/2}}. \tag{21}$$

The magenta lines in Figures 2 and 3, which match well with the simulation mean phenotype distributions $\overline{\mu}$ after many time steps, are computed from Eq. (21). This close match is our first empirical validation of the practical relevance of the integrodifferential equation model to the stochastic agent-based model, though we emphasise that as yet we have not discussed the issue of convergence to equilibrium from a given initial condition (we address this point in Section 4).

For the multistrain case, if strain j survives in the limit $t \to \infty$ and its phenotype distribution converges to an equilibrium solution, then we know that the limit distribution is Gaussian with mean

$$\overline{x}_j = x_j^* + \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}}. \tag{22}$$

In Figure 4 we show a two-species simulation, with M_1 and M_2 positive, so that if only one of the strains were initially present, a nontrivial equilibrium for that strain would be possible. The magenta line is the mean of the Gaussian equilibrium solution associated with the strain whose fate is displayed in the figure. Since $M_1 \neq M_2$, the integrodifferential equation model tells us that though both strains are initially present, if the system converges to equilibrium only one strain can survive. In this case we observe that the survivor is the strain with the higher M_j value (the one that would have the higher equilibrium population if it alone were present). Moreover, the mean phenotype for the surviving strain obtained from the simulation is close to that predicted by the integrodifferential equation model. The integrodifferential equation model predicts extinction of all strains if $M_j \leq 0$ for all strains. Figure 5 illustrates the analogous phenomenon for the agent-based model.

4. Time evolution in the multistrain continuum model

We now address time-evolving solutions of the integrodifferential equation model. Where the multistrain case is as easy to analyse as the single-strain case, we give the proof for the multistrain case, with accompanying remarks about any specific implications for the single-strain case that merit special mention.

4.1. Gaussian initial conditions

We shall prove that when $M^* > 0$, some of the equilibrium solutions have nontrivial basins of attraction, by establishing via explicit solutions the fate of the system for all Gaussian initial phenotype distributions. Importantly, in the multistrain case, some of the equilibrium solutions can never be attained from Gaussian initial conditions. More strongly, in the single strain case, if $M > 0$ all Gaussian

initial conditions initiate time-dependent solutions that converge to the unique nontrivial equilibrium solution, while if $M \le 0$ all Gaussian initial conditions initiate time-dependent solutions with asymptotic extinction.

Theorem 4.1. *Let $M^* = \max_{1\le j\le N} M_j$. For all positive values of $\varrho_j(0)$ $[1 \le j \le N]$, if each strain has a Gaussian phenotype distribution at time $t = 0$, then*

(a) *the phenotype distributions remain Gaussian for all $t > 0$ and*

$$\lim_{t\to\infty} \frac{c_j(x,t)}{\varrho_j(t)} = \frac{(\epsilon_j/\beta_j)^{1/4}}{(2\pi)^{1/2}} \exp\Big\{-\frac{1}{2}\Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2}\Big[x - \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}}\Big]^2\Big\};$$

(b) *if $M_j < 0$ then $\lim_{t\to\infty} \varrho_j(t) = 0$;*
(c) *if $M_j = 0$ then $\liminf_{t\to\infty} \varrho_j(t) = 0$;*
(d) *$\lim_{t\to\infty} \varrho_j(t) = 0$ for all strains j for which $M_j < M^*$;*
(e) *if $M^* > 0$ then $\limsup_{t\to\infty} \varrho(t) \ge M^*$ and $\liminf_{t\to\infty} \varrho_j(t) > 0$ for every strain j for which $M_j = M^*$.*

Proof of Theorem 4.1. As in the proof of Lemma 3.2 we need only consider the case $x_j^* = 0$. We seek a solution of the form

$$c_j(x,t) = \frac{\varrho_j(t)v_j(t)^{1/2}}{\sqrt{2\pi}} \exp\big\{-\frac{1}{2}v_j(t)[x-\mu_j(t)]^2\big\} \qquad (1 \le j \le N), \tag{23}$$

corresponding to a Gaussian distribution over phenotype for strain j with mean $\mu_j(t)$, variance $1/v_j(t)$ and population $\varrho_j(t)$. As usual the total population over all strains is $\varrho(t) = \sum_{1\le j\le N} \varrho_j(t)$. Construction of the unique solutions for the functions $\varrho_j(t)$, $\mu_j(t)$ and $v_j(t)$ for all $t \ge 0$—given arbitrary positive values of all of the constants $\varrho_j(0)$ and $v_j(0)$ and arbitrary real values of the initial means $\mu_j(0)$—and verifying some $t \to \infty$ asymptotic properties of $\varrho_j(t)$, $\mu_j(t)$ and $v_j(t)$ will establish the theorem.

(a) Inserting our Gaussian trial solution into the $x_j^* = 0$ version of Eq. (17) we find that Eq. (17) is satisfied if and only if

$$\begin{aligned}\frac{\varrho_j'(t)}{\varrho_j(t)} + \frac{v_j'(t)}{2v_j(t)} - \frac{v_j'(t)}{2}[x-\mu_j(t)]^2 + \mu_j'(t)v_j(t)[x-\mu_j(t)] - \alpha_j v_j(t)[x-\mu_j(t)] \\ = \beta_j\big\{-v_j(t) + v_j(t)^2[x-\mu_j(t)]^2\big\} + \gamma_j - \epsilon_j x^2 - \kappa_j\varrho(t)\end{aligned} \tag{24}$$

for $-\infty < x < \infty$ and $t \ge 0$. In particular, taking $x = \mu_j(t)$ we deduce from Eq. (24) that

$$\frac{\varrho_j'(t)}{\varrho_j(t)} + \frac{v_j'(t)}{2v_j(t)} = -\beta_j v_j(t) + \gamma_j - \epsilon_j\mu_j(t)^2 - \kappa_j\varrho(t). \tag{25}$$

Having imposed this requirement, we can eliminate the terms involving $\varrho_j(t)$ and $\varrho(t)$ from Eq. (24) using Eq. (25) and we find after a little algebra that we require

$$[x-\mu_j(t)]\Big\{\frac{v_j'(t)}{2}[x-\mu_j(t)] - \mu_j'(t)v_j(t) + \alpha_j v_j(t) + \beta_j v_j(t)^2[x-\mu_j(t)] - \epsilon_j[x+\mu_j(t)]\Big\} = 0$$

for $-\infty < x < \infty$ and $t \ge 0$. It is easy to deduce from this that the necessary and sufficient conditions for the Gaussian trial solution to give a valid solution for $-\infty < x < \infty$ and $t \ge 0$ are that the differential equation (25) and the differential equations

$$v_j'(t) + 2\beta_j v_j(t)^2 = 2\epsilon_j \tag{26}$$

$$\mu_j'(t) + \frac{2\epsilon_j\mu_j(t)}{v_j(t)} = \alpha_j \tag{27}$$

have solutions for all $t \geq 0$ which preserve the requirements that $v_j(t) > 0$ and $\varrho_j(t) > 0$, given arbitrary $\mu_j(0)$, $v_j(0) > 0$ and $\varrho_j(0) > 0$. Since Eq. (26) is a separable first-order differential equation, its solution is easily constructed in the standard way and found to be

$$v_j(t) = \Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2} \frac{\big\{(\epsilon_j/\beta_j)^{1/2} + v_j(0) - [(\epsilon_j/\beta_j)^{1/2} - v_j(0)]\exp[-4(\beta_j\epsilon_j)^{1/2}t]\big\}}{\big\{(\epsilon_j/\beta_j)^{1/2} + v_j(0) + [(\epsilon_j/\beta_j)^{1/2} - v_j(0)]\exp[-4(\beta_j\epsilon_j)^{1/2}t]\big\}}. \tag{28}$$

This remains positive for all $t \geq 0$ and $\lim_{t\to\infty} v_j(t) = (\epsilon_j/\beta_j)^{1/2}$. Moreover, the convergence to the limit is exponentially rapid.

Having determined $v_j(t)$, we may now solve the linear differential equation (27) by the integrating factor method, and we find that

$$\mu_j(t) = \Big\{\mu_j(0) + \alpha_j \int_0^t \exp\Big[2\epsilon_j \int_0^\xi \frac{d\tau}{v_j(\tau)}\Big] d\xi\Big\} \exp\Big[-2\epsilon_j \int_0^t \frac{d\tau}{v_j(\tau)}\Big]. \tag{29}$$

Evaluating the integrals and writing for brevity

$$\lambda_j = \frac{(\epsilon_j/\beta_j)^{1/2} - v_j(0)}{(\epsilon_j/\beta_j)^{1/2} + v_j(0)} \tag{30}$$

(so that $-1 < \lambda_j < 1$) yields

$$\mu_j(t) = \frac{(1-\lambda_j)\mu_j(0)}{\exp[2(\beta_j\epsilon_j)^{1/2}t] - \lambda_j \exp[-2(\beta_j\epsilon_j)^{1/2}t]} + \frac{\alpha_j}{2(\beta_j\epsilon_j)^{1/2}} \frac{\exp[2(\beta_j\epsilon_j)^{1/2}t] + \lambda_j \exp[-2(\beta_j\epsilon_j)^{1/2}t] - (1+\lambda_j)}{\exp[2(\beta_j\epsilon_j)^{1/2}t] - \lambda_j \exp[-2(\beta_j\epsilon_j)^{1/2}t]}.$$

It is easily seen that $\mu_j(t)$ exists for all $t \geq 0$ and that $\lim_{t\to\infty} \mu_j(t) = \alpha_j/[2(\beta_j\epsilon_j)^{1/2}]$.

Where $1 \leq j \leq n$, the differential equations (25) have the form

$$\varrho_j'(t) = \Big[f_j(t) - \kappa_j \sum_{\ell=1}^{N} \varrho_\ell(t)\Big]\varrho_j(t), \tag{31}$$

where, using Eq. (26) to effect some further simplification, we find that

$$f_j(t) = \gamma_j - \frac{\epsilon_j}{v_j(t)} - \epsilon_j \mu_j(t)^2 \tag{32}$$

is continuous on $[0,\infty)$. Picard's Theorem on the unique solvability of the initial value problem for first-order systems then guarantees the existence of solutions for all $t \geq 0$. Having settled global existence of $\varrho_j(t)$, we turn to the issue of strict positivity of $\varrho_j(t)$. We deduce from Eq. (31) that

$$\varrho_j(t) = \varrho_j(0) \exp\Big\{\int_0^t \big[f_j(\tau) - \kappa_j \varrho(\tau)\big] d\tau\Big\}. \tag{33}$$

The right-hand side is strictly positive and this precludes any of the strain populations $\varrho_j(t)$ vanishing in finite time. We have now established the asserted existence of the Gaussian phenotype solutions for all time and the stated Gaussian value of $\lim_{t\to\infty} c_j(x,t)/\varrho_j(t)$, completing the proof of property (a) in the statement of the theorem. It remains only to address the long-term fates of the strain populations that we have listed as (b)–(d).

(b) From Lemma 3.1 we know that for some constant m we have $0 \leq \varrho(t) < m < \infty$ for all $t \geq 0$, while from Eq. (32)

$$\lim_{t\to\infty} f_j(t) = \gamma_j - (\beta_j\epsilon_j)^{1/2} - \frac{\alpha_j^2}{4\beta_j} = \kappa_j M_j \tag{34}$$

with the convergence being exponentially rapid, so that

$$\int_0^t \big[f_j(\tau) - \kappa_j \varrho(\tau)\big] d\tau = \kappa_j \int_0^t \big[M_j - \varrho(\tau)\big] d\tau + O(1) \quad \text{as } t \to \infty. \tag{35}$$

If we denote lower and upper bounds on the $O(1)$ term by $\kappa_j A_j$ and $\kappa_j B_j$ respectively, we know that there are real constants A_j and B_j such that

$$\exp\Big\{\kappa_j \int_0^t \big[M_j - \varrho(\tau)\big]d\tau + \kappa_j A_j\Big\} \le \frac{\varrho_j(t)}{\varrho_j(0)} \le \exp\Big\{\kappa_j \int_0^t \big[M_j - \varrho(\tau)\big]d\tau + \kappa_j B_j\Big\}. \tag{36}$$

It follows immediately that $\lim_{t\to\infty} \varrho_j(t) = 0$ if $M_j < 0$.

(c) In the case $M_j = 0$, if $\liminf_{t\to\infty} \varrho(t) = 0$, then because $\varrho_j(t) \le \varrho(t)$, we find that $\liminf_{t\to\infty} \varrho_j(t) = 0$, so we need only address the case in which $\xi = \liminf_{t\to\infty} \varrho(t) > 0$, so that for all sufficiently large time, $\varrho(t) > \xi/2$. It follows then from the inequalities (36) that $\varrho_j(t) \to 0$ as $t \to \infty$, which is a stronger result than the result we set out to prove that $\liminf_{t\to\infty} \varrho_j(t) = 0$

(d) It also follows from the inequalities (36) that

$$\exp[(M_j - M_\ell)t + A_j - B_\ell] \le \frac{[\varrho_j(t)/\varrho_j(0)]^{1/\kappa_j}}{[\varrho_\ell(t)/\varrho_\ell(0)]^{1/\kappa_\ell}} \le \exp[(M_j - M_\ell)t + B_j - A_\ell]. \tag{37}$$

We observe now that if $M_j < M_\ell$, then $\varrho_j(t) \to 0$ as $t \to \infty$. The central expression in the inequalities (37) cannot vanish from a divergence to infinity of the denominator, since $\varrho_\ell(t) \le \varrho(t) \le m$, so the numerator must decay to zero, and indeed will do so exponentially rapidly.

(e) Consider a strain j for which $M_j = M^*$. If it were the case that $\limsup_{t\to\infty} \varrho(t) = \zeta < M^*$, then for sufficiently large t we would have $M_j - \varrho(t) > (M_j - \zeta)/2$ and so

$$\frac{\varrho_j(t)}{\varrho_j(0)} \ge C_j \exp[(M_j - \zeta)t/2] \tag{38}$$

for some positive constant C_j, violating the boundedness of the total population (which requires boundedness for every strain). Hence $\limsup_{t\to\infty} \varrho(t) \ge M^*$. Since all strains M_j that have $M_j < M^*$ die out as $t \to \infty$ we see that only strains for which $M_j = M^*$ can contribute to $\limsup_{t\to\infty} \varrho(t)$. If there is more than one strain for which $M_j = M^*$, then applying the inequality (37) for all pairs of such strains, we see that no such strain can have $\liminf_{t\to\infty} \varrho_j(t) = 0$. □

For the single strain model, there is a pleasingly complete characterization of the long-term fate for a Gaussian initial phenotype distribution.

Corollary 4.1. *In the case when only a single strain is present, starting from a Gaussian initial phenotype distribution the total population is given by*

$$\varrho(t) = \frac{\varrho(0)\exp\Big[\int_0^t f(\tau)d\tau\Big]}{1 + \kappa\varrho(0)\displaystyle\int_0^t \exp\Big[\int_0^\zeta f(\tau)d\tau\Big]d\zeta}. \tag{39}$$

As $t \to \infty$, we have $\varrho(t) \to 0$ if $M < 0$, $\varrho(t) \to M$ if $M > 0$, and $\varrho(t) \sim \varrho(0)/(\kappa t)$ if $M = 0$.

Proof of Corollary 4.1. When only one strain is present, Eq. (33) reduces to

$$\varrho(t) = \varrho(0)\exp\Big\{\int_0^t \big[f(\tau) - \kappa\varrho(\tau)\big]d\tau\Big\}, \tag{40}$$

where $f(t) \to \kappa M$ as $t \to \infty$. We can rewrite this equation as

$$\frac{d}{dt}\exp\Big[\kappa\int_0^t \varrho(\tau)d\tau\Big] = \kappa\varrho(0)\exp\Big[\int_0^t f(\tau)d\tau\Big]. \tag{41}$$

Integrating this equation from time 0 to an arbitrary positive time we find that

$$\exp\Big[\kappa\int_0^t \varrho(\tau)d\tau\Big] = 1 + \kappa\varrho(0)\int_0^t \exp\Big[\int_0^\zeta f(\tau)d\tau\Big]d\zeta. \tag{42}$$

Taking logarithms and differentiating produces the explicit solution (39) for the total population. Since we know that $f(t) \sim \kappa M$ as $t \to \infty$ (with $f(t) - \kappa M$ exponentially small), we know that for some constant c, as $t \to \infty$ we have

$$\exp\Big[\int_0^t f(\tau)d\tau\Big] \sim ce^{\kappa M t} \quad \text{and} \quad \int_0^t \exp\Big[\int_0^\zeta f(\tau)d\tau\Big] \sim \begin{cases} \text{constant} & \text{if } M < 0 \\ ct, & \text{if } M = 0 \\ \dfrac{c}{\kappa M} e^{\kappa M t} & \text{if } M > 0 \end{cases}$$

and the asserted asymptotic behaviour of $\varrho(t)$ follows. □

4.2. General initial conditions

We now address the existence of solutions for arbitrary initial phenotype distributions. Our results are rigorous for initial conditions that are of compact support, and we denote by Ω a closed real interval that contains as proper subsets the support of each initial phenotype distribution. The analysis also goes through for initial phenotype distributions supported by the full real line, subject to appropriate restrictions on the spatial decay of the initial conditions, but we do not discuss this here. The proof is constructive: we use a trial solution involving several functions and show how these functions can be determined.

Theorem 4.2. *If all initial strain distributions have compact support then the system of coupled integrodifferential equations for $c_j(x,t)$, $\varrho_j(t)$ and $\varrho(t)$ has a solution for all $t > 0$ with the properties that*

(a) *for all strains j with $\varrho_j(0) > 0$ we have $\varrho_j(t) > 0$ and $c_j(x,t) > 0$ for all $t > 0$;*
(b) *if $M_j < 0$ then $\varrho_j(t) \to 0$ as $t \to \infty$.*

Proof of Theorem 4.2. A special limiting case of the Gaussian solutions exhibited in Theorem 4.1 is obtained by taking the limiting initial conditions $v_j(0) \to \infty$, corresponding to a delta-function initial condition for each strain, with the initial population of strain j all having phenotype $\mu_j(0)$. The explicit solutions for the functions $v_j(t)$ and $\mu_j(t)$ introduced in Eq. (23) become

$$v_j(t) = \Big(\frac{\epsilon_j}{\beta_j}\Big)^{1/2} \coth[2(\beta_j\epsilon_j)^{1/2}t], \qquad \mu_j(t) = \frac{\mu_j(0)}{\cosh[2(\beta_j\epsilon_j)^{1/2}t]} + \frac{\alpha_j \tanh[2(\beta_j\epsilon_j)^{1/2}t]}{2(\beta_j\epsilon_j)^{1/2}}, \tag{43}$$

and for these special $v_j(t)$ and $\mu_j(t)$, we write $\mu_j(0) = \xi$ and we seek a solution of the form

$$c_j(x,t) = \int_\Omega \frac{R_j(\xi,t)v_j(t)^{1/2}}{\sqrt{2\pi}} \exp\big\{-\frac{1}{2}v_j(t)[x - \mu_j(t)]^2\big\}d\xi, \tag{44}$$

so that $R_j(x,0) = c_j(x,0)$ for $x \in \Omega$.

Inserting our trial solution (44) into Eq. (17), interchanging orders of integration and differentiation, and recalling the ordinary differential equations satisfied by $v_j(t)$ and $\mu_j(t)$, we find after some straightforward algebra that the trial solution works provided that

$$\frac{1}{R_j(\xi,t)}\frac{\partial R_j}{\partial t} = \gamma_j - \frac{\epsilon_j}{v_j(t)} - \epsilon_j\mu_j(t)^2 - \kappa_j\varrho(t). \tag{45}$$

Recalling the explicit solutions (43) for $v_j(t)$ and $\mu_j(t)$ and the notation $\mu_j(0) = \xi$ we can integrate the ordinary differential equation (45) and deduce that

$$\log\Big[\frac{R_j(\xi,t)}{R_j(\xi,0)}\Big] = \mathcal{M}_j(\xi,t) - \kappa_j\int_0^t \varrho(\tau)d\tau, \tag{46}$$

where

$$\begin{aligned}\mathcal{M}_j(\xi,t) &= \int_0^t \Big[\gamma_j - \frac{\epsilon_j}{v_j(\tau)} - \epsilon_j \mu_j(\tau)^2\Big] d\tau \\ &= \gamma_j t - \frac{1}{2}\log\big\{\cosh\big[2(\beta_j\epsilon_j)^{1/2}t\big]\big\} - \frac{\epsilon_j\xi^2}{2(\beta_j\epsilon_j)^{1/2}}\tanh[2(\beta_j\epsilon_j)^{1/2}t] \\ &\quad + \frac{\alpha_j\xi}{2\beta_j}\Big\{\frac{1}{\cosh[2(\beta_j\epsilon_j)^{1/2}t]} - 1\Big\} - \frac{\alpha_j^2}{4\beta_j}\Big\{t - \frac{\tanh[2(\beta_j\epsilon_j)^{1/2}t]}{2(\beta_j\epsilon_j)^{1/2}}\Big\}. \end{aligned} \tag{47}$$

Recalling that $R_j(x,0) = c_j(x,0)$ and eliminating $R_j(\xi,t)$ from Eq. (44) using Eq. (46) we now have an integral solution for $c_j(x,t)$, though it involves the as yet undetermined total population $\varrho(t)$:

$$c_j(x,t) = \int_\Omega \frac{c_j(\xi,0)v_j(t)^{1/2}}{\sqrt{2\pi}} \exp\Big\{\mathcal{M}_j(\xi,t) - \kappa_j \int_0^t \varrho(\tau)d\tau - \frac{v_j(t)}{2}[x-\mu_j(t)]^2\Big\} d\xi. \tag{48}$$

If we now integrate with respect to x over $(-\infty,\infty)$ and interchange orders of integration we find that

$$\varrho_j(t) = \int_\Omega c_j(\xi,0)\exp\big[\mathcal{M}_j(\xi,t)\big]d\xi \exp\Big(-\kappa_j\int_0^t \varrho(\tau)d\tau\Big). \tag{49}$$

It remains only to show how

$$r(t) = \int_0^t \varrho(\tau)d\tau \tag{50}$$

can be determined in order to complete our constructive proof of the existence of everywhere non-negative solutions $c_j(x,t)$ for all $t>0$ for the broad class of initial conditions that we have allowed in this section. We have

$$r'(t) = \varrho(t) = \sum_j \varrho_j(t) = \sum_j \int_\Omega c_j(\xi,0)\exp\big[\mathcal{M}_j(\xi,t)\big]d\xi \exp[-\kappa_j r(t)] \tag{51}$$

and Picard's Theorem on the existence and uniqueness of solutions of the initial value problem ensures the existence of a solution $r(t)$ for all $t>0$. Since $r'(t)>0$, we find that $r(t)$ is strictly increasing, and so $\varrho(t)$ cannot vanish for any finite value of t and it follows that $\varrho_j(t)>0$ and $c_j(x,t)>0$ for all $t>0$.

Finally, to demonstrate that the solution we have constructed has the property that $\varrho_j(t)\to 0$ as $t\to\infty$ for any strain j with $M_j<0$ we need only note that

$$\mathcal{M}_j(\xi,t) = \kappa_j M_j t + \frac{1}{2}\ln(2) - \frac{\epsilon_j\xi^2}{2(\beta_j\epsilon_j)^{1/2}} + \frac{\alpha_j\xi}{2\beta_j} + \frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}} + O(\exp[-2(\beta_j\epsilon_j)^{1/2}t]) \tag{52}$$

and that

$$0 < \varrho_j(t) \le \int_\Omega c_j(\xi,0)\exp\big[\mathcal{M}_j(\xi,t)\big]d\xi \tag{53}$$

from Eq. (49). □

Although we could probe further the asymptotic behaviour of the general solution that we have constructed, our proof of existence has not resolved the question of uniqueness, so there is a small gap in the logic if we wish to draw general conclusions about asymptotic behaviour. The Laplace transform approach of Section 4.3 avoids this problem: it gives more complete conclusions on the long-time behaviour of the system and also establishes uniqueness of bounded strictly positive solutions $c_j(x,t)$.

4.3. Laplace transform analysis

Having established the existence of bounded strictly positive solutions to the system of coupled integrodifferential equations (17)–(18) for all $t > 0$, we show how informative results for such solutions can be obtained using the Laplace transform. We write

$$c_j(x,t) = u_j(x,t)\exp\Big[-\kappa_j\int_0^t \varrho(\tau)d\tau + \frac{\alpha_j x}{2\beta_j}\Big], \tag{54}$$

where the new unknown functions $u_j(x,t)$ are to be strictly positive, remain bounded for $x \in (-\infty,\infty)$, and satisfy the partial differential equations

$$\frac{\partial u_j}{\partial t} = \beta_j\frac{\partial^2 u_j}{\partial x^2} - \Big[\frac{\alpha_j^2}{4\beta_j} - \gamma_j + \epsilon_j x^2\Big]u_j. \tag{55}$$

Using the Laplace transform defined in the usual way by [38]

$$\widehat{f}(s) = \mathcal{L}\{f(t); t\to s\} = \int_0^\infty e^{-st}f(t)dt, \tag{56}$$

one obtains the differential equation

$$\frac{\partial^2}{\partial x^2}\widehat{u}_j(x,s) - \Big[\frac{s}{\beta_j} + \frac{\alpha_j^2}{4\beta_j^2} - \frac{\gamma_j}{\beta_j} + \frac{\epsilon_j x^2}{\beta_j}\Big]\widehat{u}_j(x,s) = -\frac{u_j(x,0)}{\beta_j}, \tag{57}$$

for which a Green function solution can be constructed in the standard way [39], leading to an explicit solution in terms of integrals for the Laplace transform of $u_j(x,t)$, namely

$$\widehat{u}_j(x,s) = \frac{1}{\beta_j}\int_{-\infty}^{\infty}\widehat{G}_j(x-\zeta,s)u_j(\zeta,0)d\zeta, \tag{58}$$

where the Green function is the unique solution that decays as $|x|\to\infty$ of the differential equation

$$\frac{\partial^2}{\partial x^2}\widehat{G}_j(x,s) - \Big[\frac{s}{\beta_j} + \frac{\alpha_j^2}{4\beta_j^2} - \frac{\gamma_j}{\beta_j} + \frac{\epsilon_j x^2}{\beta_j}\Big]\widehat{G}_j(x,s) = -\delta(x), \tag{59}$$

which is closely related to Weber's equation, which arose in the proof of Lemma 3.2.

We find that

$$\widehat{G}_j(x,s) = \frac{1}{2[-U'(a,0)]}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4}U\Big(a,\Big(\frac{4\epsilon_j}{\beta_j}\Big)^{1/4}|x|\Big). \tag{60}$$

In Eq. (60), the s-dependence all comes through

$$a = \Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/2}\Big(\frac{s}{\beta_j} + \frac{\alpha_j^2}{4\beta_j^2} - \frac{\gamma_j}{\beta_j}\Big) = \frac{s-\kappa_j M_j}{2(\beta_j\epsilon_j)^{1/2}} - \frac{1}{2}. \tag{61}$$

The parameter a depends on j, of course, but it will be notationally convenient to suppress this dependence for the moment. The function $U(a,z)$ is the parabolic cylinder function [40, 41] and $U'(a,z)$ denotes its derivative with respect to its second argument. From Eqs (19.3.5) and (19.5.3) of Miller [40] and the duplication formula for the gamma function, it is easily shown that

$$\frac{U(a,z)}{2[-U'(a,0)]} = \frac{\exp(-z^2/4)}{2^{a/2+3/4}\Gamma(a/2+3/4)}\int_0^\infty \exp\Big(-zq - \frac{q^2}{2}\Big)q^{a-1/2}dq. \tag{62}$$

If we introduce the inverse Laplace transform

$$G_j(x,t) = \mathcal{L}^{-1}\{\widehat{G}_j(x,s); s\to t\}, \tag{63}$$

of the Green function (60) then Eqs (54) and (58) give

$$u_j(x,t) = \frac{1}{\beta_j}\int_{-\infty}^{\infty}G_j(x-\zeta,t)u_j(\zeta,0)d\zeta, \tag{64}$$

and

$$c_j(x,t) = \frac{1}{\beta_j}\exp\Big[-\kappa_j\int_0^t\varrho(\tau)d\tau\Big]\int_{-\infty}^{\infty}G_j(x-\zeta,t)\exp\Big[\frac{\alpha_j}{2\beta_j}(x-\zeta)\Big]c_j(\zeta,0)d\zeta. \tag{65}$$

If we integrate over x from $-\infty$ to ∞, interchange orders of integration on the right and let $\eta = x - \zeta$ in the inner integral, we find that

$$\varrho_j(t) = \frac{\varrho_j(0)}{\beta_j} \exp\Big[-\kappa_j \int_0^t \varrho(\tau)d\tau\Big] g_j(t), \tag{66}$$

where

$$g_j(t) = \int_{-\infty}^{\infty} G_j(\eta, t) \exp\Big[\frac{\alpha_j \eta}{2\beta_j}\Big] d\eta. \tag{67}$$

The functions $g_j(t)$ do not depend on the initial conditions used, and from the solutions we have obtained in Theorem 4.1 for Gaussian initial conditions we can deduce that $g_j(t) > 0$ for all $t \geq 0$.

By logarithmic differentiation of Eq. (66) we find that

$$\frac{\varrho_j'(t)}{\varrho_j(t)} = -\kappa_j \varrho(t) + \frac{g_j'(t)}{g_j(t)}. \tag{68}$$

We can rewrite this as

$$-\frac{\varrho_j'(t) g_j(t)}{\varrho_j(t)^2} + \frac{g_j'(t)}{\varrho_j(t)} = \frac{\kappa_j g_j(t) \varrho(t)}{\varrho_j(t)} \geq \kappa_j g_j(t). \tag{69}$$

Integrating, we find that

$$\frac{g_j(t)}{\varrho_j(t)} - \frac{g_j(0)}{\varrho_j(0)} \geq \kappa_j \int_0^t g_j(\tau)d\tau, \tag{70}$$

and it follows that

$$\varrho_j(t) \leq \frac{g_j(t)\varrho_j(0)}{g_j(0) + \kappa_j \varrho_j(0) \int_0^t g_j(\tau)d\tau}. \tag{71}$$

We now see that if $g_j(t)$ decays exponentially rapidly as $t \to \infty$, then so does $\varrho_j(t)$, while if $g_j(t)$ has a positive limit as $t \to \infty$, then $\varrho_j(t) = O(t^{-1})$.

Remark 4.1. In the degenerate case in which there is only one strain initially present (so we may drop the subscript j), the inequality (69) becomes an equality, and we have

$$\varrho(t) = \frac{g(t)\varrho(0)}{g(0) + \kappa\varrho(0) \int_0^t g(\tau)d\tau}. \tag{72}$$

It is clear from Eq. (72) that for $\varrho(t)$ to have a strictly positive limit as $t \to \infty$ we need $g(t) \sim A\exp(\lambda t)$ as $t \to \infty$ with $\lambda > 0$ and if that is the case, then $\lim_{t\to\infty} \varrho(t) = \lambda/\kappa$, independent of the value of the initial population $\varrho(0)$. On the other hand, if $g(t)$ has slower than exponential growth as $t \to \infty$, or if it decays, then $\lim_{t\to\infty} \varrho(t) = 0$.

Lemma 4.1. *Where the constant Δ_j is defined by $\Delta_j = \sqrt{2}\,\beta_j \exp\{\alpha_j^2/[8\beta_j(\beta_j\epsilon_j)^{1/2}]\}$,*

$$g_j(t) = \Delta_j e^{\kappa_j M_j t}\Big\{1 - \Big[\frac{1}{2} + \frac{\alpha_j^2}{4\beta_j(\beta_j\epsilon_j)^{1/2}}\Big] e^{-4(\beta_j\epsilon_j)^{1/2}t} + O(e^{-8(\beta_j\epsilon_j)^{1/2}t})\Big\} \text{ as } t \to \infty.$$

Proof of Lemma 4.1. To determine the large-t asymptotic behaviour of $g(t)$, we need only determine the location and nature of the relevant singularities of the Laplace transform of $g(t)$. We take the Laplace transform of Eq. (67), using Eqs (60) and (62), make the change of integration variable $z = (4\epsilon_j/\beta_j)^{1/4}\eta$, and interchange orders of integration, finding that

$$\begin{aligned}\widehat{g}_j(s) = {} & \frac{1}{2^{a/2+3/4}\Gamma(a/2+3/4)} \Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/2} \\ & \times \int_0^\infty \exp\Big(-\frac{q^2}{2}\Big) q^{a-1/2} \int_{-\infty}^{\infty} \exp\Big[-\frac{z^2}{4} + \frac{\alpha_j}{2\beta_j}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4} z - |z|q\Big] dz\, dq.\end{aligned} \tag{73}$$

If we can obtain an analytic continuation of $\widehat{g}(s)$ in the form

$$\widehat{g}_j(s) = \sum_{n=0}^{N} \frac{c_n}{s - \lambda_n} + \text{function holomorphic in } \mathrm{Re}(s) \geq \lambda_N \tag{74}$$

where $\lambda_0 > \lambda_1 > \lambda_2 \cdots$, then it will follow that as $t \to \infty$

$$g_j(t) = \sum_{n=0}^{N} c_n e^{\lambda_n t} + o(e^{\lambda_N t}). \tag{75}$$

In view of the relation (61) between a and s, we simply need to expand the singular part of $\widehat{g}_j(s)$ in terms of the parameter a to determine the asymptotic expansion for $g(t)$. The inner integral in Eq. (73) is a holomorphic function of q at the origin and its Taylor series is

$$\int_{-\infty}^{\infty} \exp\Big[-\frac{z^2}{4} + \frac{\alpha_j}{2\beta_j}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4} z\Big] dz - q \int_{-\infty}^{\infty} |z| \exp\Big[-\frac{z^2}{4} + \frac{\alpha_j}{2\beta_j}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4} z\Big] + \frac{q^2}{2} \int_{-\infty}^{\infty} z^2 \exp\Big[-\frac{z^2}{4} + \frac{\alpha_j}{2\beta_j}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4} z\Big] + O(q^3).$$

The first and third integrals over z can be evaluated by completing the square in the exponential and making an appropriate linear change of variables, to produce relatively elementary integrals associated with the Gaussian distribution (or alternatively, easily converted to gamma functions). The second integral can be evaluated in terms of the error function, though we shall never need its actual value and we introduce a constant B to account for it in the analysis that follows. We find that the inner integral in Eq. (73) is

$$2\sqrt{\pi} \exp\Big[\frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]\Big\{1 - Bq + \frac{1}{4}\Big[4 + \frac{\alpha_j^2}{\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]q^2 + O(q^3)\Big\}.$$

All singularities of the double integral in Eq. (73) at finite points in the complex a plane are shared in location and residue with those of

$$\begin{aligned}
&\int_0^1 \exp\Big(-\frac{q^2}{2}\Big) q^{a-1/2} \int_{-\infty}^{\infty} \exp\Big[-\frac{z^2}{4} + \frac{\alpha_j}{2\beta_j}\Big(\frac{\beta_j}{4\epsilon_j}\Big)^{1/4} z - |z|q\Big] dz\, dq \\
&= 2\sqrt{\pi} \exp\Big[\frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}}\Big] \int_0^1 \Big[1 - \frac{q^2}{2} + O(q^4)\Big] q^{a-1/2} \Big\{1 - Bq + \frac{1}{4}\Big[4 + \frac{\alpha_j^2}{\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]q^2 + O(q^3)\Big\} dq \\
&= 2\sqrt{\pi} \exp\Big[\frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]\Big\{\frac{1}{a+1/2} - \frac{B}{a+3/2} + \frac{1}{(a+5/2)}\Big[\frac{1}{2} + \frac{\alpha_j^2}{4\beta_j(\beta_j\epsilon_j)^{1/2}}\Big] + F(a)\Big\},
\end{aligned}$$

where $F(a)$ is holomorphic in $\mathrm{Re}(a) > -7/2$. Now the prefactor

$$\frac{1}{2^{a/2+3/4}\Gamma(a/2+3/4)}$$

in Eq. (73) is everywhere holomorphic, with simple zeros at $a = -3/2 - 2m$ ($m = 0, 1, 2, \ldots$) due to the gamma function, and it takes the values

$$\frac{1}{2^{1/2}\Gamma(1/2)} = \frac{1}{\sqrt{2\pi}} \quad \text{and} \quad \frac{1}{2^{-1/2}\Gamma(-1/2)} = -\frac{1}{\sqrt{2\pi}}$$

at $a = -1/2$ and $a = -5/2$, respectively. If follows that

$$\widehat{g}_j(s) = \frac{1}{\sqrt{2}}\Big(\frac{\beta_j}{\epsilon_j}\Big)^{1/2} \exp\Big[\frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]\Big\{\frac{1}{a+1/2} - \frac{1}{(a+5/2)}\Big[\frac{1}{2} + \frac{\alpha_j^2}{4\beta_j(\beta_j\epsilon_j)^{1/2}}\Big] + \cdots\Big\},$$

with the rightmost pole not exhibited located at $a = -9/2$. Using Eq. (61) we find that

$$\widehat{g}_j(s) = \sqrt{2}\beta_j \exp\Big[\frac{\alpha_j^2}{8\beta_j(\beta_j\epsilon_j)^{1/2}}\Big]\Big\{\frac{1}{s - \kappa_j M_j} - \frac{1}{s - \kappa_j M_j + 4(\beta_j\epsilon_j)^{1/2}}\Big[\frac{1}{2} + \frac{\alpha_j^2}{4\beta_j(\beta_j\epsilon_j)^{1/2}}\Big] + \cdots\Big\},$$

with the rightmost pole not exhibited located at $s = \kappa_j M_j - 8(\beta_j\epsilon_j)^{1/2}$. The asymptotic expansion for $g(t)$ asserted in the Lemma now follows immediately from the complex inversion formula for the Laplace transform in the usual manner. □

Corollary 4.2. *For the multistrain problem,*

(a) *if $M_j < 0$, then for all initial phenotype distributions, strain j goes extinct asymptotically, with $\varrho_j(t) = O(e^{\kappa_j M_j t})$ as $t \to \infty$;*
(b) *if $M_j = 0$ then for all initial phenotype distributions, strain j goes extinct asymptotically, with $\varrho_j(t) \le (\kappa_j t)^{-1}[1 + o(1)]$;*
(c) *if $M_j > 0$ then for all initial phenotype distributions, $\limsup_{t\to\infty} \varrho_j(t) \le M_j$.*

Corollary 4.2 follows immediately from Lemma 4.1, the inequality (71) and the observation that as $t \to \infty$

$$\int_0^t e^{\kappa_j M_j \tau}[1 + O(e^{-4(\beta_j \epsilon_j)^{1/2}\tau}]d\tau = \begin{cases} O(1) & \text{if } M_j < 0, \\ t + O(1) & \text{if } M_j = 0, \\ (\kappa_j M_j)^{-1} e^{\kappa_j M_j t}[1 + o(1)] & \text{if } M_j > 0. \end{cases} \tag{76}$$

Corollary 4.3. *For the single strain problem,*

(a) *if $M < 0$, then for all initial phenotype distributions, $\varrho(t) = O(e^{\kappa M t})$ as $t \to \infty$;*
(b) *if $M = 0$ then for all initial phenotype distributions, $\varrho(t) \le (\kappa t)^{-1}[1 + o(1)]$ as $t \to \infty$;*
(c) *if $M > 0$ then for all initial phenotype distributions, $\lim_{t\to\infty} \varrho(t) = M$.*

Proof of Corollary 4.3. This follows immediately from Lemma 4.1, Remark 4.1 and Eq. (76). □

Remark 4.2. By treating the corrections to the dominant asymptotic behaviour of $g(t)$ in the single-strain problem more carefully, we can address the rate of convergence of the total population to equilibrium when $M > 0$. Integrating the asymptotic expansion we find that

$$\int_0^t g(\tau)d\tau = \sqrt{2}\,\beta \exp\Big[\frac{\alpha^2}{8\beta(\beta\epsilon)^{1/2}}\Big]\Big\{\frac{\exp(\kappa M t)}{\kappa M} + \Delta(t)\Big\}, \tag{77}$$

where

$$\Delta(t) = \begin{cases} O(\exp\{[\kappa M - 4(\beta\epsilon)^{1/2}]t\}), & \kappa M > 4(\beta\epsilon)^{1/2}, \\ O(t), & \kappa M = 4(\beta\epsilon)^{1/2}, \\ O(1), & \kappa M < 4(\beta\epsilon)^{1/2}. \end{cases} \tag{78}$$

After a little algebra, Lemma 4.1, the expansion (77) and the exact relation (72) lead to the conclusion that

$$\varrho(t) = M + O(e^{-4(\beta\epsilon)^{1/2}t}) + O(\Delta(t)e^{-\kappa M t}) = \begin{cases} M + O(e^{-4(\beta\epsilon)^{1/2}t}) & \kappa M > 4(\beta\epsilon)^{1/2}, \\ M + O(te^{-4(\beta\epsilon)^{1/2}t}), & \kappa M = 4(\beta\epsilon)^{1/2}, \\ M + O(e^{-\kappa M t}), & \kappa M < 4(\beta\epsilon)^{1/2}. \end{cases} \tag{79}$$

Lemma 4.2. *Suppose that $M^* = \max_{1\le j\le N} M_j > 0$. Then*

(a) $\limsup_{t\to\infty} \varrho(t) \ge M^*$;
(b) *if $\liminf_{t\to\infty} \varrho_j(t) > 0$ then $\lim_{t\to\infty} \frac{1}{t}\int_0^t \varrho(\tau)d\tau = M_j$;*
(c) *if $\liminf_{t\to\infty} \varrho_j(t) > 0$ and $\liminf_{t\to\infty} \varrho_k(t) > 0$ then $M_j = M_k$.*

Proof of Lemma 4.2. The proof of part (a) is by contradiction. Suppose that $\varrho(t) \le M^* - \delta$ for all sufficiently large values of t (for $t \ge T$, say). Then for any strain k for which $M_k = M^*$ we have [from Eq. (66)]

$$\log[\varrho_k(t)] = \log[g_k(t)] - \kappa_k \int_0^t \varrho(\tau)d\tau + \text{constant} \ge \kappa_k M^* t - \kappa_k(M^* - \delta)t + o(t), \tag{80}$$

so we see that $\log[\varrho_k(t)] \to \infty$ as $t \to \infty$, so $\varrho_k(t) \to \infty$, contradicting the result proved in Corollary 4.2 that $\limsup_{t\to\infty} \varrho_k(t) \le M_k$.

For part (b) we note that

$$\log[\varrho_j(t)] = \kappa_j \Big[M_j t - \int_0^t \varrho(\tau) d\tau \Big] + O(1). \tag{81}$$

To ensure that the right-hand side is bounded below (required to keep $\liminf_{t\to\infty} \varrho_j(t) > 0$) and above (required by Corollary 4.2), it must follow that

$$M_j t - \int_0^t \varrho(\tau) d\tau = O(1) \text{ as } t \to \infty \tag{82}$$

and the claim made in part (b) is proved. Part (c) follows immediately from part (b). □

With the required preliminary calculations out of the way, we can now deduce a complete characterisation of the behaviour of a system in which there are N strains, all of which are initially present [that is, $\varrho_j(0) > 0$ for $1 \le j \le N$].

Theorem 4.3. *Let* $M^* = \max_{1\le j\le N} M_j$.

(a) *If* $M^* \le 0$, *then* $\varrho(t) \to 0$ *as* $t \to \infty$. *Moreover, if* $M_j < 0$, *we have* $\varrho_j(t) = O(e^{\kappa_j M_j t})$ *as* $t \to \infty$, *while if* $M_j = 0$, *we have* $\varrho_j(t) \le (\kappa_j t)^{-1}[1 + o(1)]$ *as* $t \to \infty$.

(b) *For* $M^* > 0$,

(i) *if* $M_j < M^*$, *then* $\varrho_j(t) \to 0$ *as* $t \to \infty$, *with* $\varrho_j(t) = O(e^{-\kappa_j (M^* - M_j) t})$;

(ii) *if* $M_j = M^*$, *then* $0 < \liminf_{t\to\infty} \varrho_j(t) \le \limsup_{t\to\infty} \varrho_j(t) \le M^*$;

(iii) $\lim_{t\to\infty} \frac{1}{t} \int_0^t \varrho(\tau) d\tau = M^*$;

(iv) *if there is a unique strain* j *for which* $M_j = M^*$, *then* $\lim_{t\to\infty} \varrho_j(t) = \lim_{t\to\infty} \varrho(t) = M^*$.

Proof of Theorem 4.3. Part (a) is a restatement of results from Corollary 4.2, so we need only address the three results listed under part (b).

(i) From the proof of Lemma 4.2 we know that where j is any strain and ℓ is a surviving strain, as $t \to \infty$ we must have

$$\log[\varrho_j(t)] = \kappa_j \Big[M_j t - \int_0^t \varrho(\tau) d\tau \Big] + O(1) \quad \text{and} \quad M_\ell t - \int_0^t \varrho(\tau) d\tau = O(1)$$

and so

$$\log[\varrho_j(t)] = \kappa_j (M_j - M_\ell) t + O(1).$$

We see that if $M_\ell < M_j$, then $\varrho_j(t) \to \infty$ as $t \to \infty$, contradicting the known boundedness of $\varrho_j(t)$. Hence $M_\ell = M^*$: only the strain or strains with the highest carrying capacity can survive, and the result (i) follows.

(ii) For all strains for which $M_j = M^*$, we have from equations (81) and (82) that

$$\log[\varrho_j(t)] = \kappa_j \Big[M^* t - \int_0^t \varrho(\tau) d\tau \Big] + O(1) \quad \text{and} \quad M^* t - \int_0^t \varrho(\tau) d\tau = O(1)$$

and so $\log[\varrho_j(t)] = O(1)$ as $t \to \infty$, which rules out the possibility that $\liminf_{t\to\infty} \log[\varrho_j(t)] = -\infty$ or $\limsup_{t\to\infty} \log[\varrho_j(t)] = \infty$, and correspondingly rules out the possibility that $\liminf_{t\to\infty} \varrho_j(t) = 0$ or that $\limsup_{t\to\infty} \varrho_j(t) = 0$. So if several strains with the maximal positive carrying capacity are initially present, none of them can go extinct in the $t \to \infty$ limit.

(iii) Since we now know that at least one strain does not go extinct in the case $M^* > 0$, part (iii) follows from part (b) of Lemma 4.2.

(iv) When we have a unique strain with maximal carrying capacity, we can use Corollary 4.2(c) and Lemma 4.2(a) to deduce that

$$\limsup_{t\to\infty} \varrho(t) = \limsup_{t\to\infty} \varrho_j(t) = M^*.$$

strain name		strain 1	strain 2	strain 3
epimutation rate descriptor		slow	intermediate	fast
phenotype lattice spacing	Δ	0.01	0.01	0.01
time increment	τ	0.001	0.001	0.001
initial number of agents	N	250	250	250
epimutation bias parameter	h_j	0.6	0.6	0.6
related epimutation bias parameter	α_j	-0.8108	-4.0541	-2.4324
epimutation rate	P	0.1	0.2	0.3
related epimutation rate	β_j	0.005	0.01	0.015
maximum proliferation rate	γ_j	100	100	100
fitness falloff parameter	ϵ_j	100	100	100
death rate	κ_j	9	9	9
mean phenotype at equilibrium	$\overline{x}_j$	0.323	0.250	0.193
carrying capacity	M_j	9617	9275	8940

TABLE 3. Parameters used for one agent-based simulation study of a three-strain system. Italicised entries in the left column refer to parameters for the related continuum model, matched to the agent-based system parameters using Eqs (7) or computed using Eqs (8) and (22) and rounded for display; all other entries are exact values.

The differentiability of $\varrho(t)$ and the limit proved as part (iii) exclude the possibility that

$$\liminf_{t\to\infty} \varrho(t) < M^*$$

and so part (iv) follows. □

We have arrived at the important conclusion that although Gaussian equilibrium solutions in the multistrain problem are possible with any one strain j for which $M_j > 0$ being the only one present, if we start with all strains initially present, then of those strains for which $M_j > 0$, only strains with $M_j = M^* = \max_\ell M_\ell$ have Gaussian equilibria that are attainable as the $t \to \infty$ limits. If there is a single strain j with $M_j = M^* > 0$, then only that strain survives in the $t \to \infty$ limit, and we have $\lim_{t\to\infty} \varrho(t) = \lim_{t\to\infty} \varrho_j(t) = M^*$. That is, *the strain that would have the highest equilibrium population if it were the only one present drives all other strains to extinction.*

When there are several strains for which $M_j = M^* > 0$, then we still have $\lim_{t\to\infty} \varrho(t) = M^*$, but it does not appear to be a simple matter to work out how the equilibrium population is partitioned between these strains. However, indefinite coexistence of two or more strains can occur only under very fortuitous circumstances that constitute a set of measure zero in the parameter space of the problem.

We observe also that if β_j is the only parameter of the model that is strain dependent and there is no phenotypic drift for any strain (that is, $\alpha_j = 0$, $\gamma_j = \gamma$, $\kappa_j = \kappa$, $\epsilon_j = \epsilon$) then decreasing β_j increases M_j, so that *the strain with the slowest phenotype diffusion is the only survivor.* Thus the continuum model exhibits for a broad parameter space the striking phenomenon for the agent-based model illustrated in Figures 4 and 5.

5. The empirical correspondence of the stochastic and deterministic models

We have been able to present a rigorous discussion of the limiting behaviour of the integrodifferential equation model. The connection between the stochastic agent-based model and the deterministic integrodifferential equation model was exhibited by a plausible approximate analysis, rather than being rigorously derived, but the integrodifferential equation model has shed light on the observed behaviour of the agent-based model. For example, for the single-strain problem, the predictions of long-term survival when $M > 0$ and extinction if $M \leq 0$ are replicated reliably in simulations provided

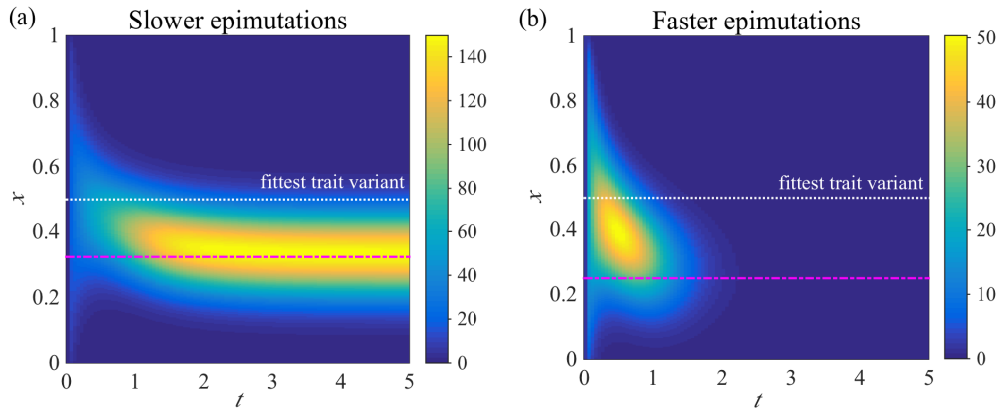

FIGURE 6. Numerical solutions of the continuum model for two competing strains with equal uniform initial distributions and parameters corresponding to the agent-based simulations for Parameter Set 1 in Table 2. For each strain, the fittest phenotype corresponds to $x^* = 0.5$ (white broken line), and the magenta broken line corresponds to the predicted mean phenotype $\overline{x} = x^* + (\alpha_i/2)(\epsilon_i\beta_i)^{-1/2}$ if only the strain depicted were present. (a): (strain 1—"slower epimutations"): $\beta_1 = 0.005$ and $\alpha_1 = -0.25$, corresponding to $\overline{x} = 0.323$ and $M_1 = 9617$. (b): (strain 2—"faster epimutations"): $\beta_2 = 0.01$ and $\alpha_1 = -0.5$, corresponding to $\overline{x} = 0.250$ and $M_2 = 9275$.

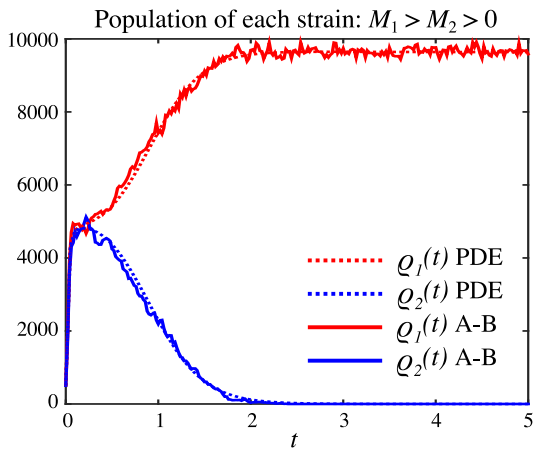

FIGURE 7. Comparison of the total populations for each strain as a function of time for a single agent-based simulation and the numerical solution of the corresponding continuum model (using Parameter Set 1 in Table 2).

the initial agent population is not too small. In the multistrain case, simulations show either the extinction of all strains, or the survival only of the strain with the highest carrying capacity, and the integrodifferential equation model predicts such behaviour. The correspondence is not merely qualitative, as we have demonstrated empirically that the long-time mean phenotype distributions of the agent-based model agree with the the mean for the integrodifferential equation model's equilibrium solutions.

We have obtained numerical solutions of the integrodifferential equations for the two-strain and the three-strain systems taking the relevant parameters from Parameter Set 1 in Table 2 and Table 3, respectively. We describe here the procedure for constructing numerical solutions of the mathematical

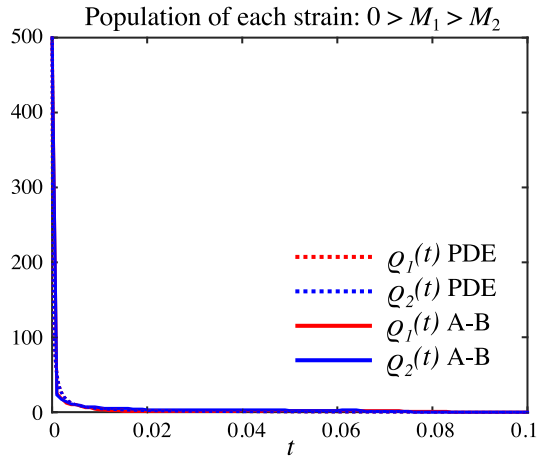

FIGURE 8. Total populations for each strain as a function of time for a single agent-based simulation and parameters chosen so that $0 > M_1 > M_2$ (using Parameter Set 2 in Table 2).

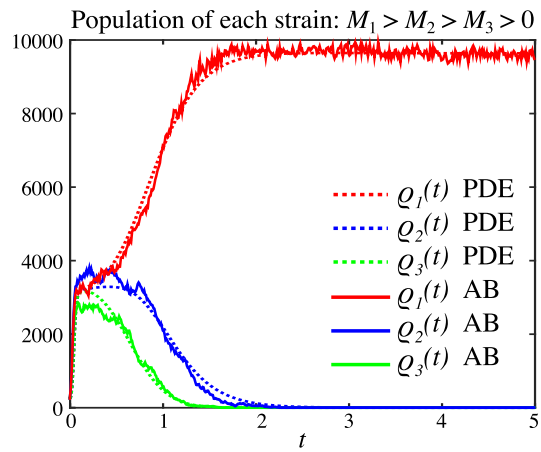

FIGURE 9. Total populations for each strain as a function of time for three competing strains, with the same uniform initial condition for each strain and parameters chosen so that $M_1 > M_2 > M_3$ (using parameters in Table 3).

problem defined by endowing (17)–(18) and (20) with the initial conditions

$$c_j(x,0) = C_0\, \mathbb{1}_{(-L;L)}(x), \qquad C_0 \in \mathbb{R}_+. \tag{83}$$

We fix a time step Δt and set $t_k = k\Delta t$. The method is based on a time splitting scheme between the conservative part and the reaction term, that is, the approximation c_j^{k+1} of $c_j(t_{k+1})$ is computed from the approximation c_j^k of $c_j(t_k)$ in two steps:

$$c_j^{k+1/2} = c_j^k - \Delta t \left(\alpha_j \frac{\partial c_j^k}{\partial x} - \beta_j \frac{\partial^2 c_j^k}{\partial x^2}\right) \tag{84}$$

and

$$c_j^{k+1} = c_j^{k+1/2} + \Delta t\, c_j^{k+1/2}\, R_j\big(x, \varrho^k\big), \tag{85}$$

where

$$R_j\big(x,\varrho^k\big) := \gamma_j - \epsilon_j (x - x_j^*)^2 - \kappa_j \varrho^k, \qquad \varrho^k = \sum_j \varrho_j^k, \tag{86}$$

and ϱ_j^k is the integral of c_j^k.

We next turn to the space discretisation and we use a uniform grid with N points on the interval $[-L, L]$, with $\Delta x = 2L/N$ the space step.

We approximate $c_j^k(x_i)$ and $R\big(x_i, \varphi^k, \varrho^k\big)$ by discrete values $c_j^{k,i}$ and $R_j^i\big(\varrho^k\big)$, and we recover ϱ_j^k through numerical integration. To be consistent with assumptions (20), we set

$$c_j^k(-L) = c_j^k(L) = 0 \quad \text{and} \quad \frac{\partial}{\partial x} c_j^k(-L) = \frac{\partial}{\partial x} c_j^k(L) = 0 \tag{87}$$

for all values of k and j.

We solve Eq. (84) by using a second-order upwind scheme for the advection term and a three-point explicit scheme for the diffusion term. For the reaction term, we use an implicit-explicit finite difference scheme [17, 19], that is, we compute $c_j^{k+1,i}$ as

$$c_j^{k+1,i} = c_j^{k+1/2,i} \, \frac{1 + \Delta t \; R_j^i\big(\varrho^k\big)_+}{1 + \Delta t \; R_j^i\big(\varrho^k\big)_-} \, . \tag{88}$$

Numerical computations were performed in MATLAB. We selected a uniform discretisation consisting of 900 points on the interval $[-L, L]$ with $L = 1.5$ as the spatial domain, and the interval $[0, 5]$ as the time domain (time step $\Delta t = 10^{-5}$).

Our numerical solutions are presented in Figure 6 and they reproduce very closely the behaviour that we found in the agent-based model and showed in Figure 4. Thus the integrodifferential equation model is informative not merely about mean values of the phenotype coordinate, but also gives useful information about distributions.

Experience with the relation between agent-based models and related continuum models [24, 25] might have led us to hope for reasonable agreement between an average over many simulations and the solution of the continuum equations. In this case, we have a good match between a single simulation and the continuum solution. As a more precise test of the match between the agent-based simulations and the continuum model, we show in Figures 7–9 the time evolution of the total population for the agent-based simulations and the continuum model solutions. We have chosen representative examples of a two-strain system in which one strain survives (Figure 7), a two-strain system in which both strains become extinct in the long-time limit (Figure 8), and a three-strain system (Figure 9). The match is excellent in each case.

For very small initial populations, or when any positive values of composite parameters M_j are sufficiently small that the equilibrium carrying capacity of the system is small, an agent-based system is to be preferred to the continuum model, so that the inevitable stochastic effects in small populations can be modelled effectively. Both the discrete stochastic and deterministic continuum models may shed light on the response of phenotypically dispersed populations introduced into a new environment and on the difficulty of maintaining phenotypic diversity (which in our case corresponds to having several long-surviving strains) under differential reproductive fitness, epigenetic drift and epigenetic evolution.

We conclude with a few remarks about some related work in other contexts and possible extensions of the present work. The analysis of nonlinear reaction–diffusion systems by maximal exploitation of relations to the ordinary diffusion equation has parallels in recent work of Alfaro and Carles [42] on non-local reaction-diffusion equations, while the notion of interacting, evolving phenotype distributions (in a genetic, rather than epigentic context, and with a somewhat different mathematical formulation) was pursued by May and Nowak two decades ago [43–45]. The inclusion of variation across physical space as well as across phenotype space or the incorporation of non-local couplings with more structure than in our simple coupling via the total population could produce a richer class of behaviours [46–55].

We note in particular that in our modelling framework the long-term coexistence of different strains can occur only for a very limited range of parameter values, namely those cases in which $M_j = M_\ell$ for some $j \neq \ell$. This is due to the fact that the effects of competition between individuals in our model are described by a logistic term involving only the total number of individuals of all strains, that is, the function ϱ). Such a modelling choice is known to lead to competitive exclusion in

the case of ODEs, and the same is true in our setting. It seems likely that if ϱ were replaced by a whole matrix of competition terms modelling the competition effect of each strain on another strain, then (provided that those terms satisfy suitable assumptions, involving stronger intra-strain competition than inter-strain competition) different strains could survive in the long run, so that coexistence would arise, as in competitive Lotka–Volterra systems of ODEs.

In our agent-based modelling, movements in the phenotype space are restricted to nearest neighbour sites only, which corresponds to the assumption that epimutations may only cause an incremental change in gene expression. A possible extension of the model would be to allow movements to sites farther away than nearest neighbours, or even to all sites, to incorporate the possibility that single epimutations may lead to larger changes in fitness, as was considered recently in the context of genetic instability in heterogeneous tumors by Asatryan and Komarova [56]. In passing to the continuum limit, any extension of our lattice-based model that accommodates steps of more than one lattice spacing would still lead to a classical diffusion mechanism so long as the mean-square displacement per step is finite, as is well known in the theory of random walks [57]. Although it would not be biologically plausible, a model on an infinite phenotype lattice with the probability of a jump of ℓ lattice sites proportional to $|\ell|^{-\mu-1}$ with $\mu \in (0,2)$ would lead to the diffusion operator $\beta\nabla^2$ being replaced by the infinitesimal generator for a stable Lévy distribution of order μ [57].

Acknowledgements

This research was supported in part by the Australian Research Council (DP140100339) and by the French National Research Agency through the ANR blanche project Kibord [ANR-13-BS01-0004] and the "ANR JC" project Modevol [ANR-13-JS01-0009]. TL was also supported in part by the Hadamard Mathematics Labex, backed by the Fondation Mathématique Jacques Hadamard, through a grant overseen by the French National Research Agency [ANR-11-LABX-0056-LMH]. LD was also supported in part by Université Sorbonne Paris Cité "Investissements d'Avenir"[ANR-11-IDEX-0005].

References

[1] Servedio, M.R., Brandvain, Y., Dhole, S., Fitzpatrick, C.L., Goldberg, E.E., Stern, C.A., Van Cleve, J., Yeh, D.J.: Not just a theory—the utility of mathematical models in evolutionary biology. PLoS Biol. **12**, e1002017 (2014)

[2] Bonduriansky, R., Crean, A.J., Day, T.: The implications of nongenetic inheritance for evolution in changing environments. Evol Applic **5**, 192–201 (2012)

[3] Odling-Smee, F.J., Laland, K.N., Feldman, M.W.: Niche construction: the neglected process in evolution, Issue 37 of Monographs in Population Biology, Princeton University Press (2003)

[4] Barcellos-Hoff M.H., Lyden D, Wang T.C.: The evolution of the cancer niche during multistage carcinogenesis. Nat Rev Cancer **13**, 511–518 (2013)

[5] Lujambio A, Esteller M.: How epigenetics can explain human metastasis: a new role for microRNAs. Cell Cycle **8**, 377–382 (2009)

[6] Esteller M, Corn PG, Baylin SB, Herman J.G.: A gene hypermethylation profile of human cancer. Cancer Res **61**, 3225–3229 (2001)

[7] House M.G., Herman J.G., Guo M.Z., Hooker C.M., Schulick R.D., Lillemoe K.D., et al.: Aberrant hypermethylation of tumor suppressor genes in pancreatic endocrine neoplasms. Ann Surg **238**, 423 (2003)

[8] Sharma, S.V., Lee D.Y., Li B., Quinlan M.P., Takahashi F., Maheswaran S., et al.: A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. Cell **141**, 69–80 (2010)

[9] Pisco, A.O., Brock, A., Zhou, J., Moor, A., Mojtahedi, M., Jackson, D., Huang, S.: Non-Darwinian dynamics in therapy-induced cancer drug resistance. Nat Commun **4**, 2467 (2013)

[10] Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., Lander, E.S.: Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. Cell **146**, 633–644 (2011)

[11] Verhoeven KJ, Preite V.: Epigenetic variation in asexually reproducing organisms. Evolution **68**, 644–655 (2014)

[12] Richards, C.L., Schrey, A.W., Pigliucci, M.: Invasion of diverse habitats by few 385 Japanese knotweed genotypes is correlated with epigenetic differentiation. Ecol Lett **15**, 1016–1025 (2012)

[13] Shea N., Pen I., Uller T.: Three epigenetic information channels and their different roles in evolution. J Evol Biol **24**, 1178–1187 (2011)

[14] Chisholm, R.H., Lorenzi, T., Lorz, A., Larsen, A.K., Neves de Almeida, L., Escargueil, A. et al.: Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, non-genetic instability and stress-induced adaptation. Cancer Res **75**, 930–939 (2015)

[15] Gudelj I., Coman C.D., Beardmore R.E.: Classifying the role of trade-offs in the evolutionary diversity of pathogens. Proc R Soc A **462**, 97–116 (2006)

[16] Lavi, O., Greene, J., Levy, D., Gottesman, M.: Simplifying the complexity of resistance heterogeneity in metastatic cancer. Trends Mol Med **20**, 129–136 (2014)

[17] Lorenzi, T., Chisholm, R.H., Desvillettes, L., Hughes, B.D.: Dissecting the dynamics of epigenetic changes in phenotype-structured populations exposed to fluctuating environments. J Theor Biol **7**, 166–176 (2015)

[18] Lorenzi, T., Chisholm, R.H., Melensi, M., Lorz, A., Delitala, M.: Mathematical model reveals how regulating the three phases of T-cell response could counteract immune evasion. Immunology **146**, 271–280 (2015)

[19] Lorz, A., Lorenzi, T., Hochberg, M.E., Clairambault, J., Perthame, B.: Populational adaptive evolution, chemotherapeutic resistance and multiple anticancer therapies. ESAIM-Math Model Num **47**, 377–399 (2013)

[20] Champagnat, N., Ferrière, R., Méléard, S.: Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. Theor Popul Biol **69**, 297–321 (2006)

[21] Champagnat, N., Ferrière, R., Ben Arous, G.: The canonical equation of adaptive dynamics: A mathematical view. Selection **2**, 73–83 (2001)

[22] Deroulers, C., Aubert, M., Badoual, M, Grammaticos, B.: Modeling tumor cell migration: from microscopic to macroscopic models. Phys Rev E **79**, 031917 (2009)

[23] Sengers, B.G., Please, C.P., Oreffo, R.O.C.: Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. J Roy Soc Interface **4**, 1107–1117 (2007)

[24] Simpson, M.J., Landman, K.A., Hughes, B.D., Fernando, A.E.: A model for mesoscale patterns in motile populations. Physica A **389**, 1412–1424 (2010)

[25] Simpson, M.J., Landman, K.A., Hughes, B.D.: Multi-species simple exclusion processes. Physica A **388**, 399–406 (2009)

[26] Calsina, À., Cuadrado, S., Desvillettes, L., Raoul G.: Asymptotics of steady states of a selection-mutation equation for small mutation rate. P Roy Soc Edinb A **143**, 1123–1146 (2013)

[27] Chisholm, R.H., Lorenzi, T., Lorz, A.: Effects of an advection term in nonlocal Lotka-Volterra equations. Commun Math Sci, in press, (2015)

[28] Desvillettes L., Jabin, P.-E., Mischler, S., Raoul, G.: On selection dynamics for continuous structured populations. Commun Math Sci **6**, 729–747 (2008)

[29] Lorenzi, T., Lorz, A., Restori, G.: Asymptotic dynamics in populations structured by sensitivity to global warming and habitat shrinking. Acta Appl Math **131**, 49–67 (2013)

[30] Lorz, A., Mirrahimi, S., Perthame, B.: Dirac mass dynamics in a multidimensional nonlocal parabolic equation. Commun Part Differ Equ **36**, 1071–1098 (2011)

[31] Mirrahimi, S., Perthame, B., Souganidis, P.E.: Time fluctuations in a population model of adaptative dynamics. Ann I H Poincaré-AN **32**, 41–58 (2015)

[32] Perthame, B., Barles, G.: Dirac concentrations in Lotka-Volterra parabolic PDEs. Indiana Univ Math J **57**, 3275–3301 (2008)

[33] Raoul, G.: Long time evolution of populations under selection and vanishing mutations. Acta Appl Math **114**, 1–14 (2011)

[34] Kimura M.: The neutral theory of molecular evolution. Cambridge: Cambridge University Press (1984)

[35] Stoltzfus, A., Yampolsky, LY.: Climbing mount probable: mutation as a cause of nonrandomness in evolution. J Hered **100**, 637–647 (2009)

[36] Haldane, J.: The association of characters as a result of inbreeding and linkage. Ann Eugen **15**, 15–23 (1949)

[37] Penington, C.J., Hughes, B.D., Landman, K.A.: Building macroscale models from microscale probabilistic models: a general probabilistic approach for nonlinear diffusion and multispecies phenomena. Phys Rev E **84**, 041120 (2011)

[38] Doetsch, G.: *Handbuch der Laplace-Transformation*, Band II. Birkhäuser, Basel (1955)

[39] Sagan, H.: *Boundary and Eigenvalue Problems in Mathematical Physics*, Ch. 9. Wiley, New York (1961)

[40] Miller, J.C.P.: Parabolic cylinder functions. In: *Handbook of Mathematical Functions*, ed. Abramowitz, M., Stegun, I.A, Ch. 19. Dover, New York (1965)

[41] Temme, N.M.: Parabolic cylinder functions. In *NIST Handbook of Mathematical Functions*, ed. Olver, F.W.J, Lozier, D.W., Boisvert, R.F., Clark, C.W., Ch. 12. Cambridge University Press (2010)

[42] Alfaro, M., Carles, R.: Explicit solutions for replicator-mutator equations: extinction versus acceleration. SIAM J Appl Math **74**, 1919–1934 (2014)

[43] Nowak, M.A., May, R.M.: Superinfection and the evolution of parasite virulence. Proc R Soc Lond B **255**, 81–89 (1994)

[44] May, R.M., Nowak, M.A.: Coinfection and the evolution of parasite virulence. Proc R Soc Lond B **261**, 209–215 (1995)

[45] May, R.M., Nowak, M.A.: Superinfection, metapopulation dynamics, and the evolution of diversity. J Theor Biol **170**, 95–114 (1994)

[46] Alfaro, M., Coville, J., Raoul, G.: Travelling waves in a nonlocal reaction-diffusion equation as a model for a population structured by a space variable and a phenotypic trait. Comm Partial Diff Eq **38**, 2126–2154 (2013)

[47] Arnold, A., Desvillettes, L., Prévost, C.: Existence of nontrivial steady states for populations structured with respect to space and a continuous trait. Comm on Pure and Applied Analysis **11**, 83–96 (2012)

[48] Bouin, E., Calvez, V.: Travelling waves for the cane toads equation with bounded traits. Nonlinearity **27**, 2233 (2014)

[49] Bouin, E., Calvez, V., Meunier, N., Mirrahimi, S., Perthame, B., Raoul, G., et al.: Invasion fronts with variable motility: phenotype selection, spatial sorting and wave acceleration. Comptes Rendus Math **350**, 761–766 (2012)

[50] Delitala, M., Lorenzi, T.: Asymptotic dynamics in continuous structured populations with mutations, competition and mutualism. J Math Anal Appl **389**, 439–451 (2012)

[51] Hughes, B.D., Fellner, K.: Continuum models of cohesive stochastic swarms: the effect of motility on aggregation patterns. Physica D **260**, 26–48 (2013)

[52] Jabin, P.E., Raoul, G.: On selection dynamics for competitive interactions. J Math Biol **63**, 493–517 (2011)

[53] Mirrahimi, S., Perthame, B.: Asymptotic analysis of a selection model with space. J Math Pures Appl **104**, 1108–1118 (2015)

[54] Mirrahimi, S., Raoul G.: Population structured by a space variable and a phenotypical trait. Theor Popul Biol **84**, 87–103 (2013)

[55] Perthame, B., Souganidis, P.E.: Rare mutations limit of a steady state dispersion trait model, preprint, (2015)

[56] Asatryan, A.D., Komarova, N.L.: Evolution of genetic instability in heterogeneous tumors. J Theor Biol **396**, 1–12 (2016)

[57] Hughes, B.D.: Random Walks and Random Environments, Vol. 1. Oxford University Press (1995)

Rebecca H. Chisholm
School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney NSW 2052, Australia
e-mail: `rebecca.chisholm@unsw.edu.au`

Tommaso Lorenzi
School of Mathematics and Statistics, University of St Andrews, St Andrews KY16 9SS, United Kingdom & Centre de mathématiques et de leurs applications, UMR 8536, CNRS, École normale supérieure de Cachan, F-94235, Cachan Cedex, France & Laboratoire Jacques-Louis Lions, Équipe MAMBA, UMR 7598, CNRS, INRIA, Sorbonne Universités, UPMC Univ. Paris 06, F-75005, Paris, France
e-mail: `tl47@st-andrews.ac.uk`

Laurent Desvillettes
Université Paris Diderot, Sorbonne Paris Cité, Institut de Mathématiques de Jussieu-Paris Rive Gauche, UMR 7586, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, F-75013, Paris, France
e-mail: desvillettes@math.univ-paris-diderot.fr

Barry D. Hughes
School of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia
e-mail: barrydh@unimelb.edu.au