Evolution and spread of multi-adapted pathogens in a spatially heterogeneous environment

Quentin Griette[1], Matthieu Alfaro[2], Gaël Raoul[3*], and Sylvain Gandon[4*]

[1]Normandie Univ, UNIHAVRE, LMAH, FR-CNRS-3335, ISCN, Le Havre 76600, France.

[2]Univ. Rouen Normandie, CNRS, LMRS UMR 6085, F-76000 Rouen, France.

[3]CMAP, CNRS, Ecole polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France.

[4]CEFE, CNRS, Univ Montpellier, EPHE, IRD, Montpellier, France.

[*]Contributed equally

QG - quentin.griette@univ-lehavre.fr

MA - matthieu.alfaro@univ-rouen.fr

GR - gael.raoul@polytechnique.edu

SG - sylvain.gandon@cefe.cnrs.fr

Type of article: Letter

Running title: The spread of multi-adapted pathogens

Keywords: epidemic spread, spatial heterogeneity, host types, single-adapted pathogen, multi-adapted pathogen

Abstract word count: 216

Total word count: 4255

Number of references: 45

Number of figures: 6

Corresponding author: Sylvain Gandon, CEFE, 1919 route de Mende, 34293 Montpellier Cedex 5, France

Abstract

Pathogen adaptation to multiple selective pressures challenges our ability to control their spread. Here we analyse the evolutionary dynamics of pathogens spreading in a heterogeneous host population where selection varies periodically in space. We study both the transient dynamics taking place at the front of the epidemic and the long-term evolution far behind the front. We identify five types of epidemic profiles arising for different levels of spatial heterogeneity and different costs of adaptation. In particular, we identify the conditions where a generalist pathogen carrying multiple adaptations can outrace a coalition of specialist pathogens. We also show that finite host populations promote the spread of generalist pathogens because demographic stochasticity enhances the extinction of locally maladapted pathogens. But higher mutation rates between genotypes can rescue the coalition of specialists and speed up the spread of epidemics for intermediate levels of spatial heterogenity. Our work provides a comprehensive analysis of the interplay between migration, local selection, mutation and genetic drift on the spread and on the evolution of pathogens in heterogeneous environments. This work extends our fundamental understanding of the outcome of the competition between two specialists and a generalist strategy (single- versus multi-adapted pathogens). These results have practical implications for the design of more durable control strategies against multi-adapted pathogens in agriculture and in public health.

Impact summary: Pathogen adaptation is constantly eroding the efficacy of prophylactic and therapeutic measures against the spread of infectious diseases. A promising way to limit the spread of multi-adapted pathogens is to distribute different control measures across space (e.g., different vaccines, different resistant varieties of crop in agriculture). Yet, the influence of the spatial deployment of these interventions on the genetic composition of spreading epidemics remains unclear. Is it possible to identify optimal deployment strategies that reduce the spread and the speed of adaptation of resistant pathogens? We analyse the evolution of pathogen adaptations throughout an epidemic spreading in a heterogeneous host population where selection varies periodically in space. We show how lower spatial heterogeneity can speed up the epidemic spread and disfavour multi-adapted pathogens. But this effect can be altered qualitatively by the demographic stochasticity taking place at the edge of the front and by higher rates of mutation between different pathogen genotypes. We predict the composition of the pathogen population both *far behind* and *at* the front of the epidemic. This analysis allows us to elucidate the consequences of the effects of spatial heterogeneity on the coexistence between specialist (single-adaptated) and generalist (multi-adapted) pathogen strategies.

1 Introduction

Pathogen epidemics can have devastating consequences for animal and plant species and it is particularly important to understand which factors govern the speed of epidemics to predict and potentially prevent their spread. Determining the speed of biological invasions has attracted a lot of attention from theoretical biologists (Fisher 1937; Kolmogorov, Petrovsky, and Piskunov 1937; Skellam 1951). Under the simplifying assumption that the invasion takes place in a homogeneous environment (e.g. an epidemic spreading in a fully susceptible host population), diffusion models can be used to predict the asymptotic speed of the epidemic (Fisher 1937; Kot, Lewis, and Driessche 1996; Shigesada and Kawasaki 1997). In this case the population is expected to spread as a travelling wave with a constant speed equal to $2\sqrt{\sigma r}$, where r is the growth rate of the population at low density and σ is the diffusion coefficient that measures how quickly the organisms disperse. Spatial heterogeneity in the environment, however, may dramatically affect the spread of the invading organism (Shigesada and Kawasaki 1997). If the spatial variation is periodic, the natural extension of the travelling front is the so-called *pulsating front* characterized by its average speed (Berestycki, Hamel, and Roques 2005a,b; Shigesada and Kawasaki 1997). Earlier studies have mostly focused on the spatial dynamics of invasions under the assumption that evolutionary dynamics could be neglected. Yet, evolution can be very rapid during invasions and this evolution can affect the speed of the spread in homogeneous environments (Griette, Raoul, and Gandon 2015; Osnas, Hurtado, and Dobson 2015; Perkins et al. 2013; Wei and Krone 2005).

Here we study how the pathogen evolution can affect the spread of an epidemic taking place in a spatially heterogeneous host population. Host variation is assumed to affect resistance to infection and pathogen transmission. Many different situations could generate this type of spatial heterogeneity. For instance, in agriculture the use of different resistant varieties in crops could be a way to manipulate the spatial distribution of host resistance to a specific pathogen (Gilligan 2008; Mikaberidze, McDonald, and Bonhoeffer 2015; Mundt 2002). In animal species, the use of different vaccines at different locations could also generate a spatial mosaic of immunity (McLeod, Wahl, and Mideo 2021). Crucially, we allow the pathogen to adapt to this diversity of host resistance and we consider different types of adaptations. First, the pathogen may evolve a specialist strategy allowing the optimal exploitation of a single resistant host. Second, the pathogen may evolve a generalist strategy allowing the pathogen to exploit distinct resistant hosts. But this ability to infect multiple host may carry intrinsic fitness costs (e.g. a lower transmission rate). The analysis of the competition between specialist and generalist strategies is a classical evolutionary question which has been explored by theoretical studies under different biological scenarios (Levins 1968; Parvinen and Egas 2004; Wilson and Yoshimura

1994). These studies have shown that the long-term evolutionary outcome and the potential coexistence between multiple strategies depend on the balance between the amount of spatial heterogeneity and the homogenizing effect of migration. Yet, it is unclear if the same principle holds away from the equilibrium, at the front of a population that is spreading in a heterogeneous environment. In particular, it is unclear if one expects the generalist strategy to be more frequent at the edge or far behind the front, and how this evolution can affect the speed of the spread. Besides, a better understanding of the consequences of the heterogeneity of host resistance on pathogen dynamics could have practical implications for disease control. For instance, we could optimize the composition of the host population to reduce epidemic spread and limit the evolution of multi-adapted pathogens which are expected to erode dramatically the efficacy of control efforts.

In the following, we take advantage of the theoretical framework of pulsating fronts to examine the spatial dynamics of different pathogens spreading in a one-dimensional environment. First, we study the effect of the spatial heterogeneity on the speed of a monomorphic pathogen population. In a second step, we allow mutations between different pathogen genotypes and we analyse the evolution of a coalition of different pathogen genotypes. We contrast the composition of the pathogen population at the edge and behind the front and we identify five different types of epidemic profiles. Finally, we examine the effect of demographic stochasticity on the speed of spreading epidemics when the host population is assumed to be of finite size.

2 Methods

We model the dynamics of a directly transmitted pathogen in a one-dimensional habitat. At time t and position x, the host population is divided into uninfected individuals, $S(t,x)$, and infected individuals, $I(t,x)$. We assume that dead hosts are immediately replaced by new susceptible hosts (because host fecundity is assumed to be large and not limiting) so that the total density of hosts is assumed to remain constant over space and time: $K = S(t,x) + I(t,x)$. We focus on a scenario where the environment is divided into two different habitats where the hosts are either of type A or of type B. For instance, this scenario could result from the use of two different vaccines at different locations or, if we consider the spread of a phytopathogen in crop, by the use of distinct host resistant varieties in different fields. We consider a simple spatial pattern where host composition varies periodically and we use L to denote the period of the spatial fluctuation of host composition. Because all the hosts are resistant to some pathogen genotype we expect that the pathogens fully sensitive to both types of host resistance will be rapidly outcompeted by single- or multi-adapted genotypes. We thus focus our analysis on the dynamics of three adapted pathogen genotypes circulating in the host population:

(i) the density of hosts infected with the genotype only able to infect host of type A is noted $I_a(t,x)$ (single-adapted genotype a to host type A), (ii) the density of hosts infected with the genotype only able to infect host of type B is noted $I_b(t,x)$ (single-adapted genotype b to host type B) and (iii) the density of hosts infected with the genotype able to infect both types of hosts is noted $I_m(t,x)$ (m for multi-adaptation to both types of hosts). Coinfection by different genotypes is not allowed and each genotype i is characterized by $\beta_i(x)$, the rate at which transmission occurs between infected and susceptible hosts after a contact at position x. The rate of transmission of the multi-adapted genotype β_m is independent of space because multi-adaptation implies that the rate of transmission is not affected by the treatment. In contrast, the rates of transmission $\beta_a(x)$ and $\beta_b(x)$ vary in space because we assume that host resistance reduces transmission (without affecting the other life history traits). All the infections are assumed to end (because of clearance and/or increased mortality due to pathogen virulence) at a rate α. More precisely, we assume that β_a (resp. β_b) takes values $\alpha + r$ in populations of host A only (resp. B only), and value $\alpha - r$ in populations of host B only (resp. A only), see **Fig. 1**. This symmetry between the two specialists simplifies the following analysis of the model. Note, however, that we also examine a scenario when we introduce some asymmetry in the maximal growth rates of the two specialists in the **Supplementary Information** (section 1.2.1 and Fig. S3). Mutations may occur between these three genotypes and μ_{ij} stands for the rate of mutation from genotype i to genotype j.

The transmission of the pathogen is assumed to be local (infected hosts can only infect susceptible hosts at the same spatial location) but both susceptible and infected hosts are allowed to diffuse in one dimension with a fixed rate σ. In other words, we neglect the influence the pathogen may have on the mobility of its host. Our model can thus be written as the following set of reaction–diffusion equations (for readability, we drop the time and space dependence notation on host densities):

$$\begin{cases} \dfrac{\partial I_a}{\partial t} = I_a\left[r_a(x) - \beta_a(x)\dfrac{I}{K}\right] + \sigma\dfrac{\partial^2 I_a}{\partial x^2} + \mu_{ba}I_b + \mu_{ma}I_m - (\mu_{ab}+\mu_{am})I_a \\ \dfrac{\partial I_b}{\partial t} = I_b\left[r_b(x) - \beta_b(x)\dfrac{I}{K}\right] + \sigma\dfrac{\partial^2 I_b}{\partial x^2} + \mu_{ab}I_a + \mu_{mb}I_m - (\mu_{ba}+\mu_{bm})I_b \\ \dfrac{\partial I_m}{\partial t} = I_m\left[r_m - \beta_m\dfrac{I}{K}\right] + \sigma\dfrac{\partial^2 I_m}{\partial x^2} + \mu_{am}I_a + \mu_{bm}I_b - (\mu_{ma}+\mu_{mb})I_m \end{cases} \tag{1}$$

where $I = I_a + I_b + I_m$. Note that $r_i(x) = \beta_i(x) - \alpha$ is the malthusian growth rate of the single-adapted genotype i (with $i \in \{a,b\}$) and $r_m = \beta_m - \alpha$ is the malthusian growth rate of the multi-adapted genotype m, when most of the hosts are uninfected (i.e. at the edge of the epidemic). Yet, when the pathogen population starts to increase locally the density of uninfected hosts drops and decreases the transmission opportunities as in classical epidemiological models with direct-transmission (see

also (Débarre, Lenormand, and Gandon 2009; Griette, Raoul, and Gandon 2015). This drop in host density would be even stronger if host fecundity was not able to compensate host mortality (the total density of the host population would drop due to the spread of the pathogen). For simplicity, however, we restrict our analysis to the case where $S(t,x) + I(t,x)$ remains constant and equal to K.

In the following we study the speed of spreading epidemics in a spatially heterogeneous environment as a function of (i) the period of the spatial fluctuation in the composition of the host population, (ii) the transmission rates of the different genotypes in the different habitats. We first consider the spread of single genotypes before analysing the effect of mutations among genotypes on the speed of a polymorphic pathogen population. Finally, we explore the effect of demographic stochasticity on the speed of monomorphic and polymorphic epidemics spreading in heterogeneous environments.

3 Results

3.1 The speed of a monomorphic pathogen population

The multi-adapted genotype m does not "feel" the spatial heterogeneity of host population. When such a genotype is introduced in the host population and if we assume no mutation ($\mu_{ma} = \mu_{mb} = 0$) the above system reduces to the spread of a single pathogen in a uniform environment. The pathogen population spreads as a travelling wave with a speed equal to (Griette, Raoul, and Gandon 2015; Osnas, Hurtado, and Dobson 2015; Shigesada and Kawasaki 1997):

$$c_m = 2\sqrt{\sigma r_m}. \tag{2}$$

The analysis of the speed of a single-adapted genotype $i \in \{a, b\}$ is more challenging because the growth rate of the pathogen varies periodically in space between $r_i(x) = r$ (when the genotype is adapted to the host in x) and $r_i(x) = -r$ (when the genotype is not adapted to the host in x). It is possible to derive good approximations for the speed of the epidemic in two limit cases (Hamel, Fayard, and Roques 2010; Hamel, Nadin, and Roques 2011), namely when L is small and when L is large. When the period of the fluctuation of the environment is very small (i.e. $L \to 0$) the *grain* of the environment is so small that the growth rate of the pathogen is equal to the average growth rate in the two habitats: $\overline{r} = \frac{r+(-r)}{2} = 0$. In contrast, when the period of the fluctuation is large the pathogen will move very fast when it is adapted to the host and it will slow down when the host resistance reduces its transmission rate. In the limit when $L \to \infty$ the speed reaches an asymptote

that can be described explicitly. We then get, for $i \in \{a, b\}$,

$$c_i \sim 0 \text{ when } 0 < L \ll 1, \quad c_i \sim \left(\frac{2}{\sqrt{3}}\right)^{3/2} \sqrt{r} \text{ when } 1 \ll L. \tag{3}$$

Moreover the speed of the single-adapted genotype epidemic increases with L, the period of the spatial fluctuation of the environment (**Fig. 2**).

3.2 The speed of a polymorphic pathogen population

Before considering the full system (with the 3 pathogen genotypes: a, b and m) we examine the dynamics of a coalition of two single-adapted genotypes (a and b) each adapted to distinct types of hosts. When the mutation rates are very low (i.e. $\mu_{aj} = \mu_{bj} \approx 0$) we recover the result of a monomorphic population (red line in **Fig. 2**). However, numerical simulations with a fixed mutation rate μ between single-adapted genotypes indicate that increasing the mutation rate has a complex effect on the speed of the polymorphic population (**Fig. 2**). When L is small, increasing the mutation rate has only a weak effect on epidemic speed because the environment changes so fast that both specialist genotypes are almost equifrequent. For intermediate values of L, the size of the area populated by a single host type allows the adapted genotype to outcompete the other genotype and to take up some speed. Hence, the composition of the epidemic fluctuates between the two specialist genotypes and a higher mutation rate speeds up the emergence of this locally adapted genotype and increases the propagation speed. For larger values of L, however, this effect is dominated by the detrimental emergence of ill-adapted mutants (*mutation load*) that slows down the propagation within an area populated by a single host type. Hence, the composition of the pathogen population at the front of the epidemic depends on the balance between local selection, mutation and L which measures the amount of spatial heterogeneity. We show in the **Supplementary Information** (section 1.2.1) that there is a threshold value L_c below which the whole epidemic can be driven by a single specialist:

$$L_c \sim \frac{2\sqrt{2}}{3^{3/4} - \sqrt{2}} \sqrt{\frac{\sigma}{r}} \ln\left(\frac{\sqrt{\sigma r}}{\mu}\right). \tag{4}$$

When $L < L_c$ the propagation of each specialist is independent because they can move through the "bad habitat" by diffusion. In contrast, when $L > L_c$ the bad habitat slows down the spread of the maladapted specialist and the coalition of two specialists is faster than a single specialist because they "pass the baton" when they move to a different habitat. The composition of the pathogen population at the front of the epidemic fluctuates between the two specialist genotypes. Higher mutation rates speed up the epidemic because mutation speeds up the switch between the two specialists at the tip

of the front. Note, however, that high mutation rates generate a *mutation load* when $L \gg L_c$ via the recurrent introduction of a single-adapted genotype unable to infect the local host type. This is why the maximal speed of the coalition of single-adapted genotypes can never reach the speed of a universally adapted pathogen ($c_{a+b} < 2\sqrt{\sigma r}$ in **Fig. 2**).

When we assume a fixed mutation rate μ among the three pathogen genotypes, the epidemic spreads faster than epidemics where only the coalition of two specialists is present, provided the period of the fluctuation is small (**Fig. 3**). Indeed, when L is small the multi-adapted genotype m outpaces the single-adapted genotypes at the front of the epidemic (**Fig. 3**). In contrast, when L is large, the multi-adapted genotype is outcompeted by the coalition of the two specialists because we assume the maximal growth rate r of the specialits is higher than the growth rate r_m of the generalist (in particular when the mutation rate between single-adapted genotypes is large enough). Increasing the mutation rate tends to lower the speed of the epidemic when L is small or very large, because mutations reintroduce maladapted genotypes and build up the mutation load (**Fig. 4**). For intermediate values of L, however, increasing the mutation rate can increase the speed of the pathogen spread, by speeding up the propagation of a the coalition of specialists a and b (**Fig. 4**). This is due to the beneficial effects of mutations on the speed of the coalition of two single-adapted genotypes that we discussed above (**Fig. 2**).

3.3 The speed of stochastic epidemics

The above results rely on the assumption that the deterministic model we are using provides a good description of the spread of a pathogen epidemics. Yet, the front of the epidemic is driven by a small number of infections. The finite nature of the pathogen population at the edge of the epidemics yields demographic stochasticity and is expected to slow down its spread (Brunet and Derrida 1997; Griette, Raoul, and Gandon 2015; Mueller, Mytnik, and Quastel 2011; Snyder 2003). In the following we explore the effect of stochasticity using an individual-based model that takes into account the finite number N of hosts at each spatial location. The individual transitions between the different states of the hosts are described by a list of random events (transmission, mutation, death; see the **Supplementary Information** section 2.1 for a detailed description of the individual-based model). As expected, this stochastic model converges to the above deterministic model when N is assumed to be very large. To study the effect of demographic stochasticity on epidemic spread we performed simulations with our individual-based model and measured the average speed on a long time interval after the influence of the initial condition is lost.

First, we discuss the speed of monomorphic epidemics in the absence of mutations. The speed of

the multi-adapted genotype is decreased by the effect of stochasticity but remains very close to the deterministic approximation (see Brunet and Derrida 1997; Griette, Raoul, and Gandon 2015). The magnitude of this drop is expected to be proportional to $\left(\ln\left(\frac{N}{\delta x}\right)\right)^{-2}$, where $\frac{N}{\delta x}$ represents the number of hosts per unit of space. In contrast, the speed of the single-adapted specialist is dramatically altered by stochasticity (**Fig. 3**). This speed is always lower than the speed of the deterministic approximation but, when L is large the speed can drop abruptly to zero which indicates that the pathogen cannot spread any more. Indeed, when the period of the fluctuation of the environment reaches a threshold value $L_e \sim \frac{4}{3}\sqrt{\frac{\sigma}{r}}\ln\left(\frac{N}{\delta x}\right)$ the pathogen cannot cross the unfavourable habitat (see **Supplementary Information**, section 2.2.2). In particular, the pathogen is very likely to go extinct in the unfavourable habitat when the population size is small, the diffusion rate is limited and its growth rate is very negative (remember that we assume the growth rate to be $-r$ in the unfavourable habitat). Note that this critical period L_e only increases logarithmically with the population size N, so that this "blocking effect" can be observed even with relatively large population sizes. This explains why the propagation speed of a single-adapted genotype is maximised for intermediate values of L. In the deterministic approximation, in contrast, the pathogen can always cross unfavourable habitats because extinctions do not occur and the speed of epidemic spread increases monotonically with L.

Second, if we allow some mutation between the two single-adapted genotypes, the epidemic can cross those unfavorable environments because mutations will rescue pathogen populations when $L > L_e$. Consequently, increasing mutation rates can have a dramatic impact on the speed of epidemics when L is large (**Fig. 4**). Finally, when we allow the mutation between the three different genotypes, the speed of the epidemics is close to (but lower than) the deterministic approximation, and this speed can decrease when $L > L_e$ and the mutation rates are small enough (**Fig. 4**). As pointed out above, the magnitude of this effect on the reduction of the epidemic speed is of the order $(\ln(N))^{-2}$ when N is large enough.

3.4 Pathogen diversity far behind the epidemic front

In the previous sections we focused on the speed and the composition of the pathogen population at the edge of the epidemic. Next, we characterise the composition of the pathogen population far behind the front, when it reaches an endemic equilibrium. Note that the composition of the pathogen population behind the front is much less sensitive to the effect of demographic stochasticity because at the endemic equilibrium, the number of pathogens present is much larger than at the front of the epidemics, diminishing greatly the risk of genotype extinctions. Hence, we do not need to distinguish the deterministic and stochastic models in this section. Three cases can be observed (**Fig. 5**):

(i) The multi-adapted genotype dominates: If both the cost of being multi-adapted (i.e. $r - r_m$) and L are low, the generalist strategy outcompetes the specialists and goes to fixation.

(ii) The coalition of specialist genotypes dominates: When both the cost of being multi-adapted (i.e. $r - r_m$) and L are large, the coalition of specialists outcompetes the generalist strategy.

(iii) The three genotypes coexist: The coexistence between the three different genotypes is also possible for a range of parameter values when both r_m and L are relatively large. Indeed, as pointed by (Débarre and Lenormand 2011), a generalist strategy can outcompete specialists at the interface between habitats.

3.5 Five epidemic profiles

The above analysis shows how the composition of the pathogen population is dominated by different genotypes at the edge and behind the front of the epidemic. Indeed, even if all genotypes are reintroduced locally by mutation, the spatial variability of the environement and the spread of the population affects the relative competitive abilities of the different genotypes at different locations. In particular, when we vary both the period of host heterogeneity L and the growth rate r_m of the multi-adapted genotype, we can distinguish five different profiles of epidemics (**Fig. 5**). Interestingly, we identify an epidemic type (marked by **III** in **Fig. 5**, see also **Fig. 6**) where the multi-adapted genotype m drives the spread of the epidemic but is outcompeted later on by the coalition of the two specialists (single-adapted genotypes a and b). In other words, the analysis of the transitory dynamics reveals conditions where the multi-adapted genotype is able to emerge, taking advantage of the presence of numerous uninfected host populations, even though specialized strategies are better competitors once the epidemics has developed and many hosts have been infected.

We recover the same five epidemic profiles with finite host population sizes (**Fig. 5**) but demographic stochasticity affects the genetic diversity at the front of the epidemic where the size of the pathogen population is reduced. Single-adapted genotypes are most sensitive to the influence of stochasticity because these specialized genotypes can reach very low density in unfavourable habitats. The multi-adapted genotype m benefits from the influence of this demographic stochasticity (compare the size of epidemic type marked by **III** in the deterministic and stochastic cases illustrated by **Fig. 5**).

4 Discussion

Our study provides a comprehensive analysis of the evolution of pathogen specialization in a spreading epidemics. Our model allows us to examine both the long-term evolutionary outcome far behind the front of the epidemic, and the transient evolution taking place at the front of the epidemic. We

recover the classical result of previous evolutionary analyses showing that the long-term evolutionary outcome depends on the balance between spatial heterogeneity and the amount of migration among habitats. Larger patches of homogeneous habitats favor the coalition of locally adapted specialists in each habitat, but migration tends to favor generalist strategies able to cope with a diversity of habitats (Christiansen 1975; Day 2000; Débarre and Gandon 2010; Débarre, Ronce, and Gandon 2013; Mirrahimi and Gandon 2020). We also recover the possibility to maintain the coexistence of specialists and generalist strategies when the generalist can be stably maintained at the interface between habitats (Débarre and Lenormand 2011). Interestingly, our analysis of the transient evolutionary dynamics of the pathogen in a spreading epidemic reveals that the composition of the pathogen population can be very different at the front of the epidemic. Indeed, even if the local composition of the host population does not change in time, the pathogen present at the front of the epidemic experiences temporal fluctuations of the environment. Frequent temporal fluctuations favor the generalist strategy because, in spite of its constitutive fitness cost (i.e. $r_m < r$ in our model), the generalist strategy does not feel the heterogeneity of the environment. Consequently, we show that multi-adapted pathogens are expected to drive the spread of epidemics in finely grained environments. In contrast, when the spatial fluctuations are larger, the coalition of specialists is expected to drive the epidemics. Indeed, even if the transition between the two habitats can slow down the average speed of a coalition of specialists, the speed of each specialist is maximized when they are locally adapted. Contrasting the composition of the pathogen population at the edge and at the back of the epidemic allowed us to identify five different types of epidemic profiles in **Fig. 5**. This figure shows that the coexistence of specialists and generalists strategies is promoted by a lower fitness of the multi-adapted genotype and a larger period of host heterogeneity. In general we find that the speed of the epidemic is increased with larger period of host heterogeneity but, as discussed below, these results are modulated by the pathogen mutation rates and by the amount of demographic stochasticity.

We found that mutation among pathogen genotypes is a double edged sword: (i) it allows the pathogen to acquire adaptive mutations but (ii) it can also produce a mutation load with the recurrent introduction of locally maladapted genotypes. The balance between these two effects depends on the heterogeneity of the environment which, in turn, depends on the ratio between the period L of the fluctuation of the environment and the diffusion coefficient σ. The beneficial effect of a higher mutation rate is maximal for intermediate levels of this ratio. Indeed, it is not profitable for the pathogen population to mutate often when the environment keeps changing (i.e., $L \sim 0$) or when the environment changes very slowly (i.e., $L \to \infty$). Several earlier studies obtained similar conclusions in non-spatial models where it is possible to show that there is an optimal stochastic switching rate

between specialized phenotypes that maximizes the growth rate of a population in a fluctuating environment (Kussell and Leibler 2005; Lachmann and Jablonka 1996). In all these different scenarios, the introduction of genetic variation provides a way to "pass the baton" between different specialist genotypes and allows the population to exploit more efficiently a fluctuating environment.

As expected from earlier theoretical studies (Brunet and Derrida 1997; Griette, Raoul, and Gandon 2015; Mueller, Mytnik, and Quastel 2011; Snyder 2003), demographic stochasticity lowers the speed of the epidemic spread. Most of the results of the deterministic model hold in finite host populations. The only notable exception occurs when large values of L can prevent the spread of single-resistance genotypes. The input of new mutations may then provide a way to adapt to the new host type. Hence the speed of pathogen epidemics may be constrained by both the stochastic nature of the demographic process and the stochastic nature of the mutation events occurring at the edge of the epidemic. Several earlier studies have shown how the increased intensity of genetic drift in expanding populations could result in an "expansion load" due to the accumulation of deleterious mutations (Hallatschek and Nelson 2010; Peischl, Kirkpatrick, and Excoffier 2015). In our model, however, deleterious mutations at some location (e.g. genotype a in host type B) are adaptive at other locations (e.g. in host type A). It would be interesting to study the effects of finite population size in a more realistic model allowing for the accumulation of unconditionally deleterious mutations.

Our models can be used to make practical recommendations regarding the manipulation of the spatial structure of the host population to limit the speed of pathogen epidemics. The spatial structure of the host population can be manipulated by mixing hosts with different levels of resistance to the pathogen. This variation in host resistance can either be due to genetic heterogeneity ~~among~~ (e.g. resistant crop varieties), to immunological heterogeneity (e.g. vaccination) or other therapeutic interventions (e.g. the use of drugs against the pathogen). Earlier studies have analysed the impact of the local manipulation of the heterogeneity of the environment on the adaptation of pests and pathogens (Comins 1977; Débarre, Bonhoeffer, and Regoes 2007; Lenormand and Raymond 1998; Park et al. 2015; Raymond 2019). In particular these models have determined the critical area size of host resistance above which adaptation to the host does not occur because local selection is swamped by the influence of migration. The present study expands these earlier studies that focused on the migration-selection equilibrium and examines transient dynamics of adaptation in the presence of two types of host resistance. Hence, our analysis may be particularly relevant in agriculture where multiple resistance varieties may be used to limit pathogen spread (Djidjou-Demasse, Moury, and Fabre 2017; Rimbaud et al. 2018a,b, 2021). If the objective is to limit the speed of the epidemic spread, a lower value of L should be recommended. Lower L values imply that a spreading epidemics

is exposed to a more variable environment. This prevents the pathogen to specialize to a specific environment and, consequently, to speed up in a favourable environment. Interestingly, fine-scale environmental heterogeneity (low L values) are also expected to reduce the probability of pathogen emergence (Chabas et al. 2018). This fine-scale heterogeneity, however, may promote the spread of generalist and multi-adapted pathogens. Those generalist pathogens are likely to spread more slowly because of the potential fitness cost associated with the acquisition of additional mutations. But additional compensatory mutations (not considered in our model) may restore the competitivity of generalist pathogens against specialist pathogens. In other words, the optimal deployment of control measures in space varies with the forecast horizon. Our model helps clarify the consequences of these interventions on the short term epidemiological dynamics (the speed of the spreading epidemic) as well as the evolutionary dynamics of the pathogen population.

Several experimental studies have monitored and quantified the spread and the evolution of a bacteria in laboratory conditions (Baym et al. 2016; Deforet et al. 2019). In particular, the MEGA-plate experiment of Baym et al followed the spread of *Escherichia coli* in a spatially heterogeneous environment characterised by increasing concentrations of antibiotics. This fascinating experiment allowed to visualize pathogen spread and evolution in real time. This experimental procedure could be used to test some of our predictions. For instance, we could monitor the influence of the scale of spatial heterogeneity with a manipulation of the parameter L in the MEGA-plate. We hope that the present theoretical framework may stimulate an experimental validation of our theoretical predictions using experimental evolution of microbes in spatially heterogeneous environments.

Acknowledgements

We thank the CNRS MITI for funding the project VIRADAPT&SPREAD. This work has received funding from the French ANR DEEV (ANR-20-CE40-0011-01) projects, and from the région Normandie BIOMA-NORMAN (21E04343) project. GR was partially funded by the ERC SINGER ADG 101054787.

Author contributions

QG carried out the simulations and made the figures, MA contributed to the analysis of the model, GR derived the approximations of the threshold values of habitat size, SG wrote the first draft of the manuscript, and all authors contributed to the revisions of the manuscript. The project was initiated by GR and SG, and all authors contributed to the development of the final version of the model.

Data accessibility

This is a theoretical study and no data were used.

References

Baym, M. et al. (2016). Spatiotemporal microbial evolution on antibiotic landscapes. *Science* **353.6304**, pp. 1147–1151.

Berestycki, H., F. Hamel, and L. Roques (2005a). Analysis of the periodically fragmented environment model. I. Species persistence. *J. Math. Biol.* **51.1**, pp. 75–113.

— (2005b). Analysis of the periodically fragmented environment model. II. Biological invasions and pulsating travelling fronts. *J. Math. Pures Appl. (9)* **84.8**, pp. 1101–1146.

Brunet, E. and B. Derrida (1997). Shift in the velocity of a front due to a cutoff. *Physical Review E* **56.3**, p. 2597.

Chabas, H. et al. (2018). Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS biology* **16.9**, e2006738.

Christiansen, F. B. (1975). Hard and soft selection in a subdivided population. *The American Naturalist* **109.965**, pp. 11–16.

Comins, H. N. (1977). The management of pesticide resistance. *Journal of Theoretical Biology* **65.3**, pp. 399–420.

Day, T. (2000). Competition and the effect of spatial resource heterogeneity on evolutionary diversification. *The American Naturalist* **155.6**, pp. 790–803.

Débarre, F. and S. Gandon (2010). Evolution of specialization in a spatially continuous environment. *Journal of evolutionary biology* **23.5**, pp. 1090–1099.

Débarre, F. and T. Lenormand (2011). Distance-limited dispersal promotes coexistence at habitat boundaries: reconsidering the competitive exclusion principle. *Ecology Letters* **14.3**, pp. 260–266.

Débarre, F., T. Lenormand, and S. Gandon (2009). Evolutionary epidemiology of drug-resistance in space. *PLoS Computational Biology* **5.4**, e1000337.

Débarre, F., O. Ronce, and S. Gandon (2013). Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology* **26.6**, pp. 1185–1202.

Deforet, M., C. Carmona-Fontaine, K. S. Korolev, and J. B. Xavier (2019). Evolution at the edge of expanding populations. *The American Naturalist* **194.3**, pp. 291–305.

Djidjou-Demasse, R., B. Moury, and F. Fabre (2017). Mosaics often outperform pyramids: insights from a model comparing strategies for the deployment of plant resistance genes against viruses in agricultural landscapes. *New Phytologist* **216.1**, pp. 239–253.

Débarre, F., S. Bonhoeffer, and R. R. Regoes (2007). The effect of population structure on the emergence of drug resistance during influenza pandemics. *Journal of The Royal Society Interface* **4.16**, pp. 893–906.

Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of eugenics* **7.4**, pp. 355–369.

Gilligan, C. A. (2008). Sustainable agriculture and plant diseases: an epidemiological perspective. *Philosophical Transactions of the Royal Society B: Biological Sciences* **363.1492**, pp. 741–759.

Griette, Q., G. Raoul, and S. Gandon (2015). Virulence evolution at the front line of spreading epidemics. *Evolution* **69.11**, pp. 2810–2819.

Hallatschek, O. and D. R. Nelson (2010). Life at the front of an expanding population. *Evolution* **64.1**, pp. 193–206.

Hamel, F., J. Fayard, and L. Roques (2010). Spreading speeds in slowly oscillating environments. *Bull. Math. Biol.* **72.5**, pp. 1166–1191.

Hamel, F., G. Nadin, and L. Roques (2011). A viscosity solution method for the spreading speed formula in slowly varying media. *Indiana Univ. Math. J.* **60.4**, pp. 1229–1247.

Kolmogorov, A. N., I. G. Petrovsky, and N. S. Piskunov (1937). Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. Etat Moscou* **Sér. Inter. A 1**, pp. 1–26.

Kot, M., M. Lewis, and P. van den Driessche (1996). Dispersal Data and the Spread of Invading Organisms. *Ecology* **77.7**, pp. 2027–2042.

Kussell, E. and S. Leibler (2005). Phenotypic diversity, population growth, and information in fluctuating environments. *Science* **309.5743**, pp. 2075–2078.

Lachmann, M. and E. Jablonka (1996). The inheritance of phenotypes: an adaptation to fluctuating environments. *Journal of theoretical biology* **181.1**, pp. 1–9.

Lenormand, T. and M. Raymond (1998). Resistance management: the stable zone strategy. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **265.1409**, pp. 1985–1990.

Levins, R. (1968). *Evolution in changing environments: some theoretical explorations.* 2. Princeton University Press.

McLeod, D. V., L. M. Wahl, and N. Mideo (2021). Mosaic vaccination: how distributing different vaccines across a population could improve epidemic control. *Evolution Letters* **5.5**, pp. 458–471.

Mikaberidze, A., B. A. McDonald, and S. Bonhoeffer (2015). Developing smarter host mixtures to control plant disease. *Plant Pathology* **64.4**, pp. 996–1004.

Mirrahimi, S. and S. Gandon (2020). Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. *Genetics* **214.2**, pp. 479–491.

Mueller, C., L. Mytnik, and J. Quastel (2011). Effect of noise on front propagation in reaction-diffusion equations of KPP type. *Invent. Math.* **184.2**, pp. 405–453.

Mundt, C. C. (2002). Use of multiline cultivars and cultivar mixtures for disease management. *Annual review of phytopathology* **40.1**, pp. 381–410.

Osnas, E. E., P. J. Hurtado, and A. P. Dobson (2015). Evolution of pathogen virulence across space during an epidemic. *The American Naturalist* **185.3**, pp. 332–342.

Park, A. W., J. Haven, R. Kaplan, and S. Gandon (2015). Refugia and the evolutionary epidemiology of drug resistance. *Biology Letters* **11.11**, p. 20150783.

Parvinen, K. and M. Egas (2004). Dispersal and the evolution of specialisation in a two-habitat type metapopulation. *Theoretical Population Biology* **66.3**, pp. 233–248.

Peischl, S., M. Kirkpatrick, and L. Excoffier (2015). Expansion load and the evolutionary dynamics of a species range. *The American Naturalist* **185.4**, E81–E93.

Perkins, T., B. Phillips, M. Baskett, and A. Hastings (2013). Evolution of dispersal and life history interact to drive accelerating spread of an invasive species. *Ecol Lett* **16**, pp. 1079–1087.

Raymond, B. (2019). Five rules for resistance management in the antibiotic apocalypse, a road map for integrated microbial management. *Evolutionary applications* **12.6**, pp. 1079–1091.

Rimbaud, L., J. Papaix, L. G. Barrett, J. J. Burdon, and P. H. Thrall (2018a). Mosaics, mixtures, rotations or pyramiding: What is the optimal strategy to deploy major gene resistance? *Evolutionary Applications* **11.10**, pp. 1791–1810.

Rimbaud, L., J. Papaix, J.-F. Rey, L. G. Barrett, and P. H. Thrall (2018b). Assessing the durability and efficiency of landscape-based strategies to deploy plant resistance to pathogens. *PLoS computational biology* **14.4**, e1006067.

Rimbaud, L. et al. (2021). Models of plant resistance deployment. *Annual Review of Phytopathology* **59**, pp. 125–152.

Shigesada, N. and K. Kawasaki (1997). *Biological invasions: theory and practice.* Oxford University Press, UK.

Skellam, J. (1951). Random dispersal in theoretical populations. *Biometrika* **38**, pp. 196–268.

Snyder, R. E. (2003). How demographic stochasticity can slow biological invasions. *Ecology* **84.5**, pp. 1333–1339.

Wei, W. and S. M. Krone (2005). Spatial invasion by a mutant pathogen. *Journal of theoretical biology* **236.3**, pp. 335–348.

Wilson, D. S. and J. Yoshimura (1994). On the coexistence of specialists and generalists. *The American Naturalist* **144.4**, pp. 692–707.

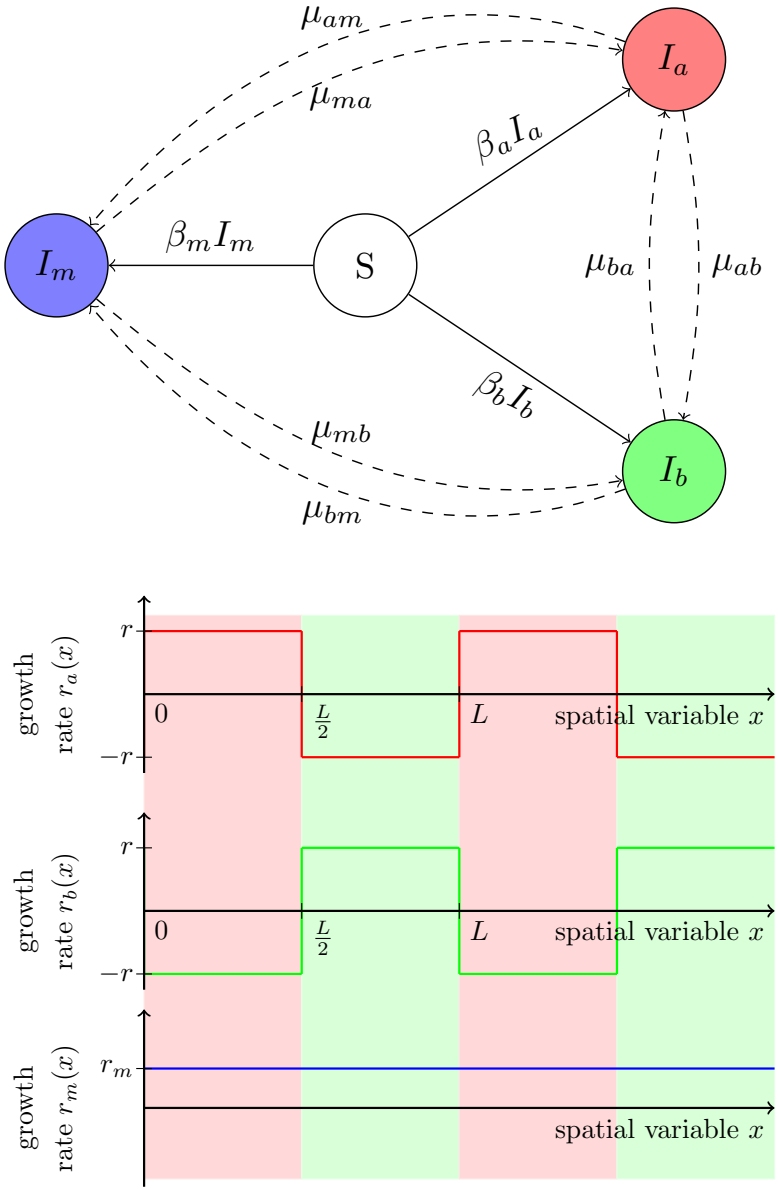

Figure 1: Schematic presentation of the evolutionary epidemiology model and the spatial heterogeneity of the environment. Top figure: diagram of the compartimental model. S represents susceptible hosts, I_a (resp. I_b, I_m) represents hosts infected by single-adapted genotype a, which is only able to infect host type A (resp. single-adapted genotype b only able to infect host type B, and the multi-adapted pathogen able to infect both types of hosts). In dashed we have represented mutations that typically happen at a much lower rate than transmissions. Bottom figure: Values of the intrinsic growth rates $x \mapsto r_a(x) = \beta_a(x) - \alpha$, $x \mapsto r_b(x) = \beta_b(x) - \alpha$, $x \mapsto r_m = \beta_m - \alpha$ as a function of the spatial variable $x \in \mathbb{R}$, where, for $x \in (0, L)$, $\beta_a(x) = 2r\mathbb{1}_{(0,\frac{L}{2})}(x)$ while $\beta_b(x) = 2r\mathbb{1}_{(\frac{L}{2},L)}(x)$, and $\alpha = r$. The maximal growth rate of the single-adapted genotypes is assumed to be higher than the growth rate of the multi-adapted genotype: $r \geq r_m$. The red (resp. green) area represents the locations $x \in \mathbb{R}$ where hosts of type A (resp. B) are present.

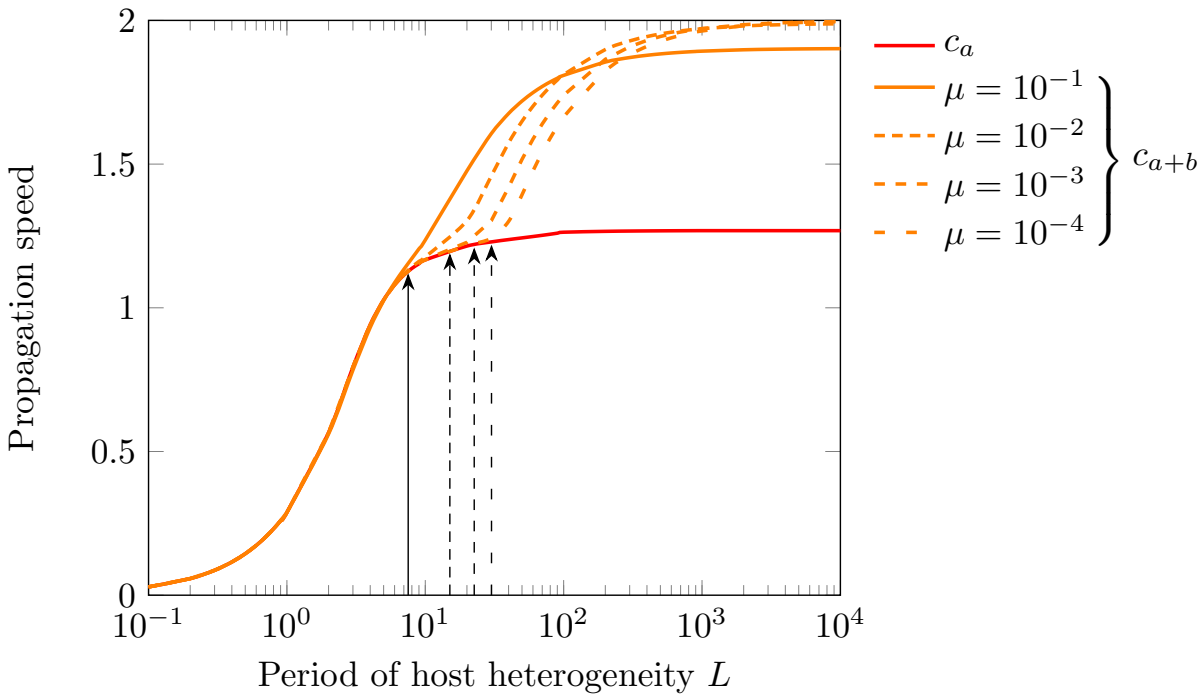

Figure 2: Impact of the mutation rate μ on the propagation speed of a coalition of the two specialist pathogen types, for the determinist model. We plot the speed c_a of a single specialist genotype (red line) and the speed c_{a+b} of a coalition of both specialist genotypes propagating together (orange lines) when $\mu_{ab} = \mu_{ba} = \mu$ (with $\mu_{am} = \mu_{bm} = 0$). The final values for c_a are extrapolated (from $L = 2000$ inclusive). The black arrows indicate the values of L_c for the different rates of mutation (see equation (4)). Parameters: $\sigma = 1$, $r = 1$, and the functions $\beta_a(x)$, $\beta_b(x)$, $r_a(x)$ and $r_b(x)$ are as in Fig. 1.

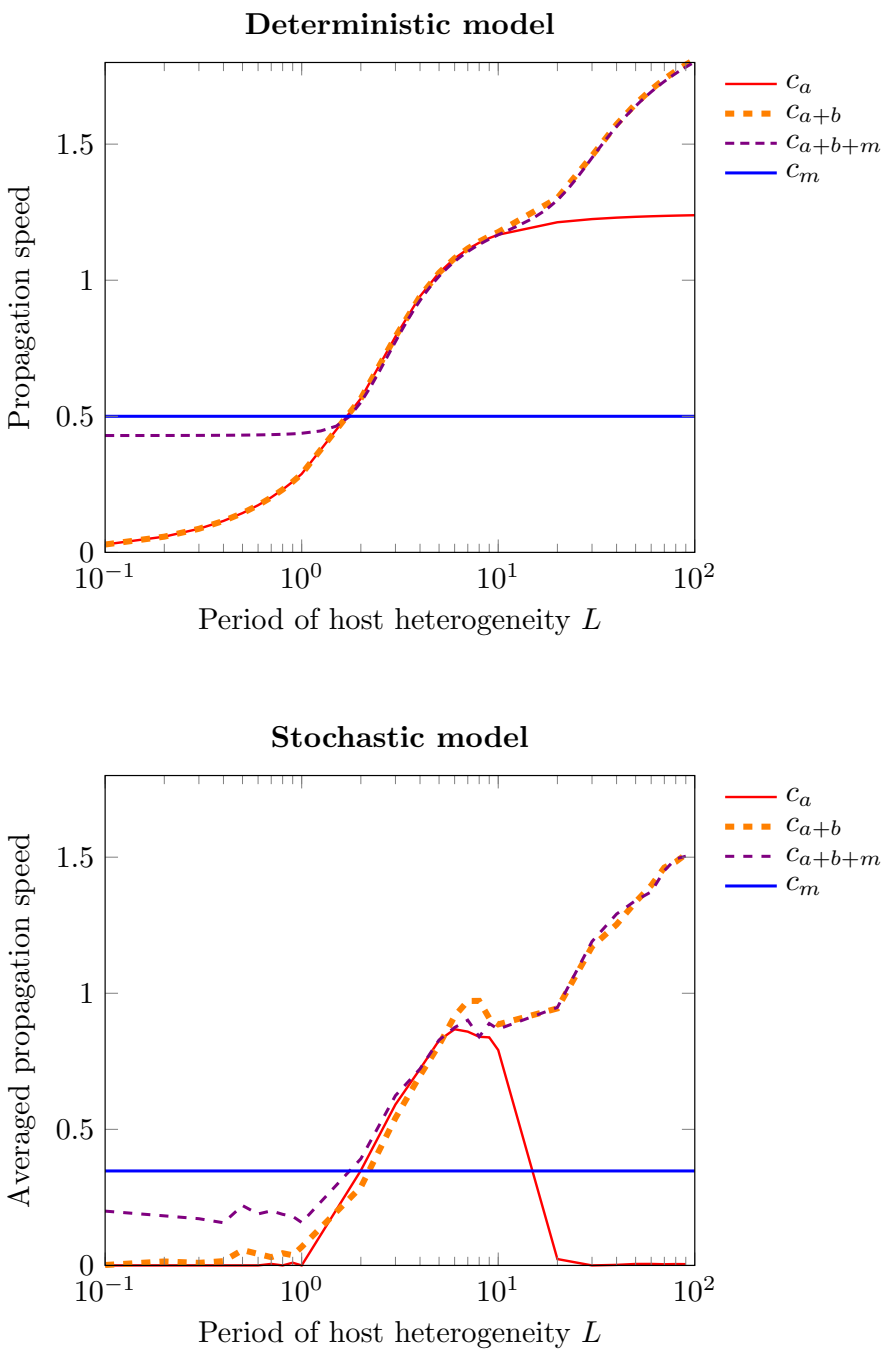

Figure 3: Propagation speed when only one specialist genotype is present (c_a), when both specialist genotypes are present (c_{a+b} with $\mu_{ab} = \mu_{ba} = \mu$) and when all the three pathogen genotypes are present (c_{a+b+m} with $\mu_{ij} = \mu, \forall\, i, j \in \{a, b, m\}$). Top figure: speed of the epidemic in the *deterministic model* (1) against the period L for the coalition of specialist genotypes (orange line: c_{a+b} with $\mu_{ab} = \mu_{ba} = \mu$), the multi-adapted genotype alone (blue line: c_m) and the full model with both the specialist genotypes and the multi-adapted genotype (purple line: c_{a+b+m} with $\mu_{ij} = \mu, \forall\, i, j \in \{a, b, m\}$). Bottom figure: speed of the epidemic in the *stochastic model* with $N = 100$ and $\delta x = 0.1$. Parameters: $r = 1$, $r_m = \frac{1}{16}$, $\sigma = 1$, $\mu = 0.01$, $\beta_m = 1 + \frac{1}{16}$, and the functions $\beta_a(x)$, $\beta_b(x)$, $r_a(x)$ and $r_b(x)$ are as in Fig. 1.

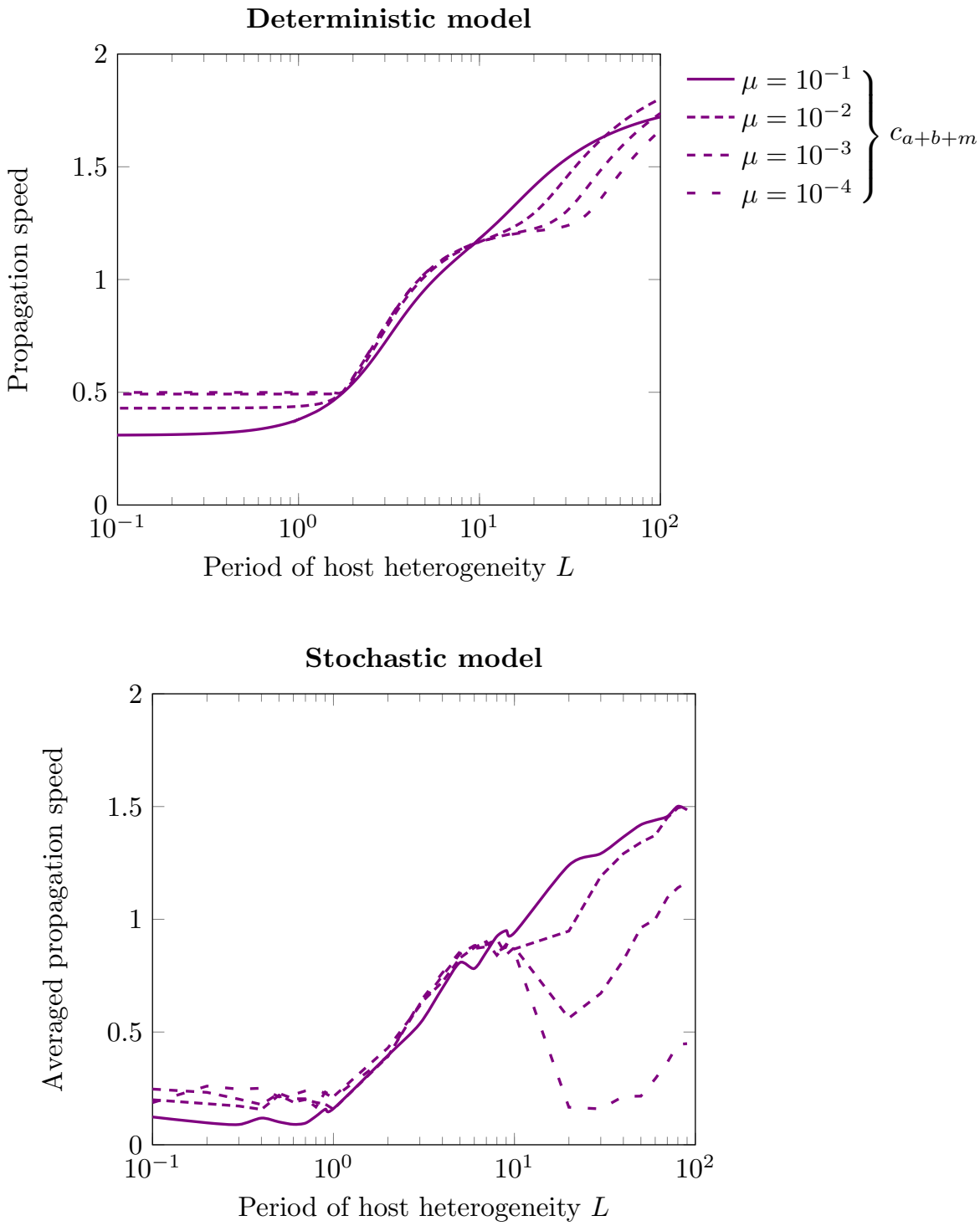

Figure 4: Effect of the mutation rate μ on the propagation speed of the epidemics when all three pathogen types are present (c_{a+b+m} with $\mu_{ij} = \mu, \forall\, i, j \in \{a, b, m\}$). Top figure: *deterministic model*. Bottom figure: *stochastic model* with $N = 100$ and $\delta x = 0.1$. Parameters: $\sigma = 1$, $r_m = \frac{1}{16}$, $r = 1$, $\beta_m = 1 + \frac{1}{16}$, and the functions $\beta_a(x)$, $\beta_b(x)$, $r_a(x)$ and $r_b(x)$ are as in Fig. 1.

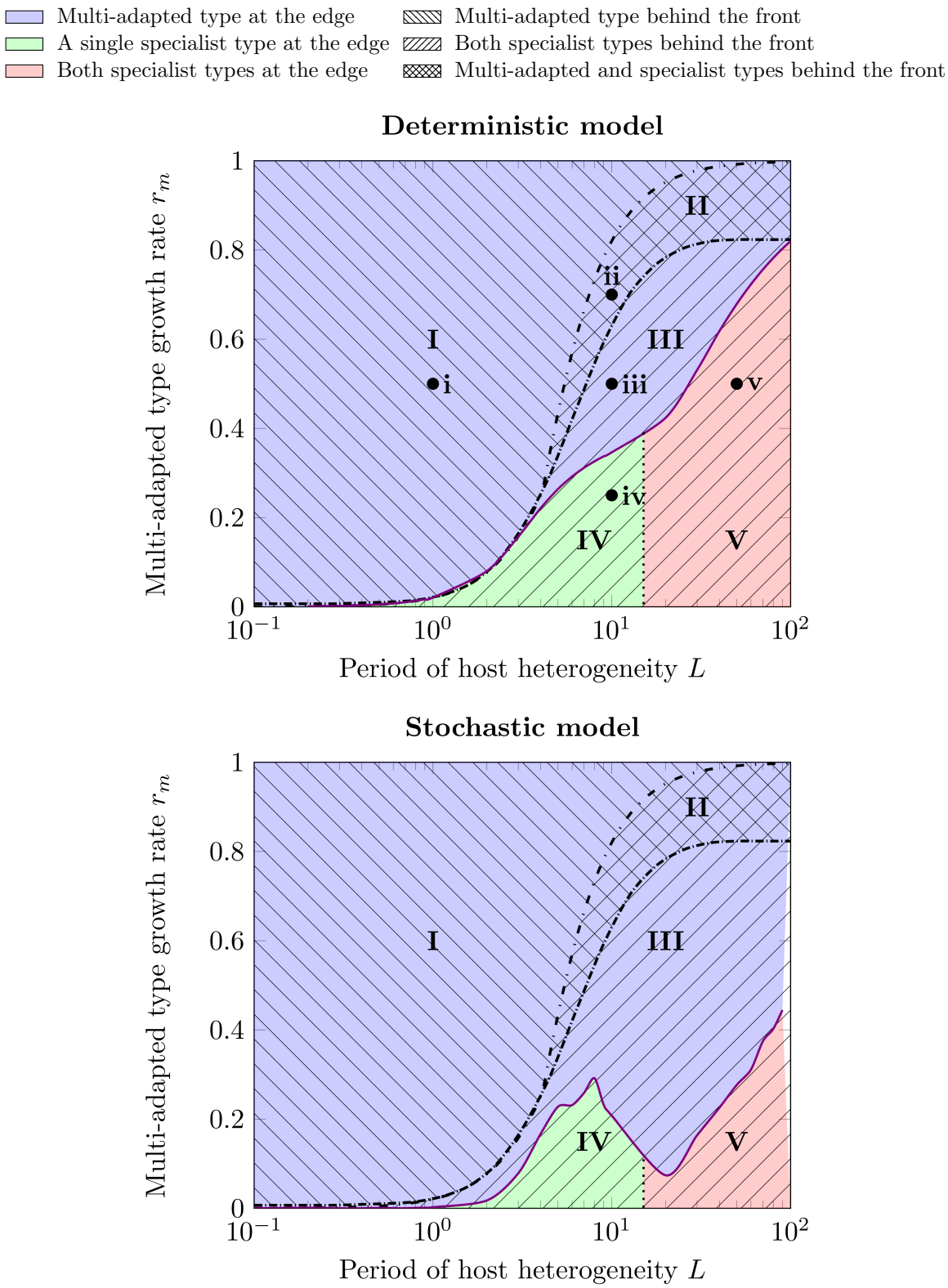

Figure 5: The five epidemic profiles. Composition of the population at the edge of the front (colors), and behind the front (hatches), as a function of r_m and L with $\mu_{ij} = \mu, \forall\, i, j \in \{a, b, m\}$. See also Fig. 6 for the description of these different epidemic profiles obtained with the parameters noted **i** to **v** in the top figure . Top figure: *deterministic model*. Bottom figure: *stochastic model* with $N = 100$ and $\delta x = 0.1$. Parameters: $\sigma = 1$, $\mu = 0.01$, $r = 1$, $\beta_m = 1 + r_m$, and the functions $\beta_a(x)$, $\beta_b(x)$, $r_a(x)$ and $r_b(x)$ are as in Fig. 1.

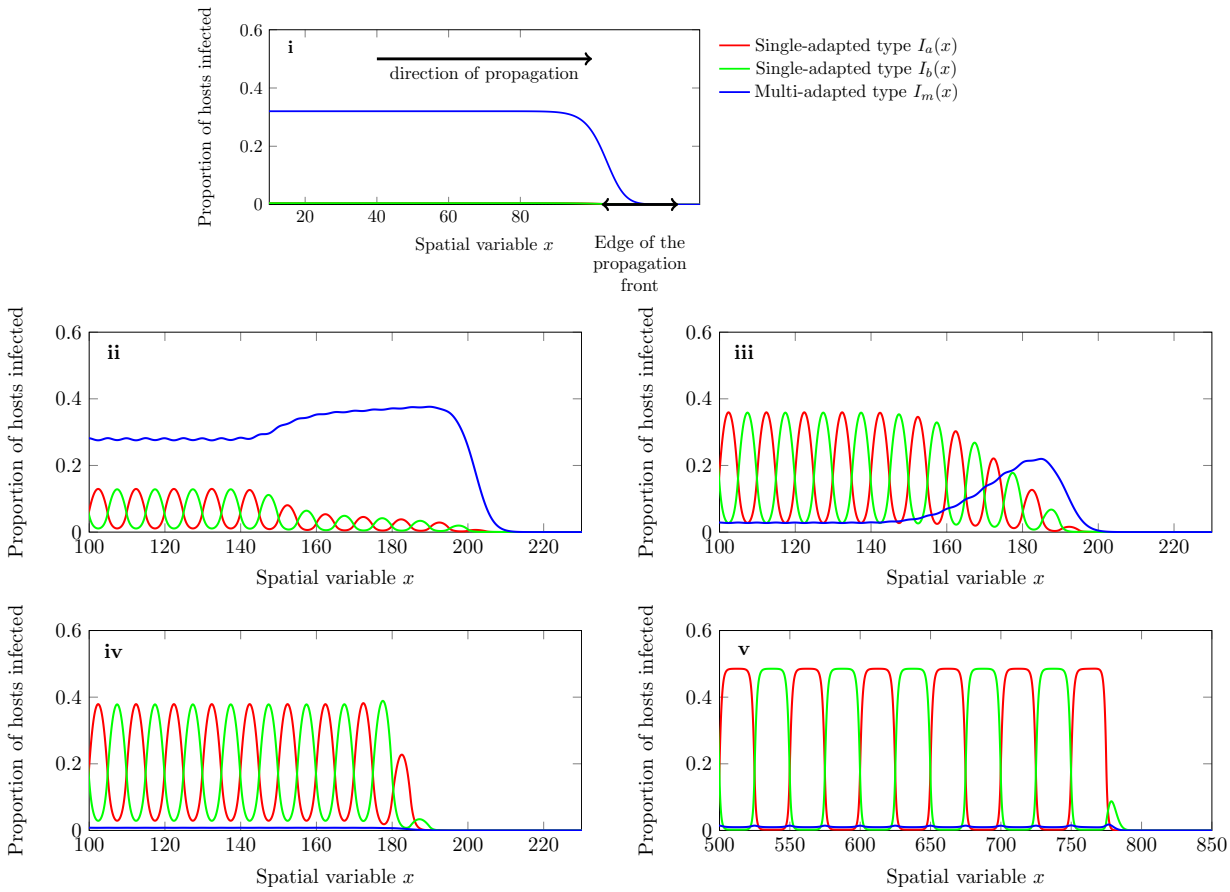

Figure 6: Composition of the pathogen population. For the parameters (L, r_m) noted **i** to **v** in Fig. 5: $(1, 0.5)$, $(10, 0.7)$, $(10, 0.5)$, $(10, 0.25)$, $(50, 0.5)$. Other parameters $r = 1$, $\mu = 0.001$, $\beta_m = 0.5$, and the functions $\beta_a(x)$, $\beta_b(x)$, $r_a(x)$ and $r_b(x)$ are as in Fig. 1.