The interplay between migration and selection on the dynamics of pathogen variants

Wakinyan Benhamou[iD,1,2,⋆], Rémi Choquet[iD,2], and Sylvain Gandon[iD,2]

[1] *High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA*
[2] *CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France*

[⋆] *E-mail: wakinyan@princeton.edu*

June 22, 2025

Abstract

The fitness advantage of an emerging pathogen variant is typically estimated from its change in frequency over time. This approach relies on the assumption that the pathogen spreads in a well-mixed population, where frequency changes are solely driven by selection. Yet, spatial structure can have major consequences on the spread of the new variant. Here, we model the change in frequency across time and across space of a new pathogen variant spreading in a two-patch host metapopulation. Crucially, we show that even small rates of migration can interfere with the effects of selection and may bias the estimation of fitness. We illustrate this effect with the evolution of SARS-CoV-2 by contrasting the spread of the Alpha and the Delta variants in England. We contend that the observed heterogeneity of fitness estimates across space could result from the variation of host mobility which is expected to affect pathogen migration. This work highlights the need of a comprehensive theoretical framework accounting for the interplay between selection and migration on the spread of new pathogen variants.

Keywords: Evolutionary epidemiology theory; selection; migration; SARS-CoV-2

1 Introduction

Estimating the fitness advantage of new pathogen variants is essential for evaluating epidemic risk and optimizing control strategies. This fitness advantage, relative to co-circulating lineages, is usually estimated from the increase in variant frequency on the logit scale (aka the selection coefficient) [1, 2]. This approach has been widely used during the COVID-19 pandemic, in the context of the emergence and replacement of successive variants of concern [3–5]. For instance, **Fig. 1** shows the dynamics of the sweep of the Alpha variant (lineage B.1.1.7) and the Delta variant (lineage B.1.617.2) across regions in England. The analysis of the variant dynamics across space is particularly insightful. First, for both variants, the time

of introduction of the new variant differs substantially among regions, which generates major differences in variant frequency across space throughout their invasion. Second, the fitness estimated from the change in frequency reveals that the fitness of the Alpha variant (relative to the ancestral strain in October 2020) is smaller than the fitness of the Delta variant (relative to Alpha): 0.065 day^{-1} (95% CI [0.052, 0.077]) vs. 0.097 day^{-1} (95% CI [0.093, 0.101]), respectively. Third, the fitness estimations for Alpha remain relatively constant across regions, but is much more variable for Delta. In particular, the fitness seems to increase in regions where the variant was introduced later (e.g, about 60% higher in the Yorkshire and Humber region than in the London or South West regions). How can we explain this spatial variation in fitness among different regions? This pattern could emerge if there were systematic sampling biases as pointed out in [4] and [6]. Yet, we want to explore the alternative mechanistic possibility that this spatio-temporal pattern is driven by other evolutionary forces acting on the change in frequency of a new variant. First, demographic stochasticity and superspreading events may have massive effects on the dynamics of a new variant at the start of its invasion, when its population size is relatively small [7–9]. These stochastic events are likely to explain the difference in introduction time among different regions (random founder effect) but should have relatively minor effects on the subsequent dynamics of the variant. Second, variations of selection among different regions could also lead to different changes in frequency over time, e.g., heterogeneity in non-pharmaceutical interventions (NPIs) [10–12], vaccination [13–15]. Third, host mobility among regions (allowing pathogen migration) may also alter the dynamics of variant frequency. In this study, we will focus on the interplay between selection and migration, and we thus need to develop a model accounting for spatial structure, epidemiological dynamics and pathogen evolution.

Classical epidemiological models, such as the Susceptible-Infected-Recovered (SIR) model popularized by [16], assume a homogeneous-mixing population: each individual has the same probability of meeting any other individual [17, 18], so that there is no spatial heterogeneity that may impact the spread of the disease. Yet, hosts typically interact more with a reduced set of neighbors [19, 20]. A common way to account for spatial structure is to use metapopulation models, that is, a network of spatially separated patches (e.g., households, cities, regions, countries), interconnected by explicit flows of migration [21, 22]. Such models have been extensively used in epidemiology (e.g., [23–31]), and the recent COVID-19 pandemic was no exception (e.g., [32, 33]). Modeling spatial heterogeneity and migration have also been a major focus in the field of evolutionary biology, including in the context of host-pathogen systems – e.g., the phenotypic evolution of pathogens in spatially heterogeneous environments, focusing on life-history traits such as drug resistance [34, 35] or virulence and transmission [15, 36–45]. Yet, most studies focus on the long-term evolutionary outcome and the analysis of the interplay between migration and selection on the transient evolutionary dynamics of

pathogens remains unclear.

Evolutionary epidemiology theory is well suited to study the short-term dynamics of host-pathogen systems, when epidemiological and evolutionary time scales overlap [46], but has been mostly used in well-mixed environments (e.g., [47–51]). In this study, we extend evolutionary epidemiology theory to follow the evolutionary dynamics of a pathogen population in a two-patch host metapopulation. This fully deterministic model tracks simultaneously the epidemiological dynamics and the change in frequency of a new pathogen variant. We focus on the dynamics of the variant frequency across time and on the differentiation across populations, two quantities that are increasingly becoming available from sequencing data. Our model thus provides a useful theoretical framework to build a more comprehensive understanding of the evolutionary forces that shape the dynamics of emerging pathogen variants.

2 Evolutionary epidemiology across time and space

2.1 The model

We model the spread of a horizontally transmitted pathogen in a two-patch host metapopulation. Let β refer to the *per capita* transmission rate, and γ to the *per capita* recovery rate. We track the dynamics of two strains of the pathogen: the wildtype (w) and the mutant strain (m, or variant). The variant may differ phenotypically from the wildtype in terms of transmission rate ($\beta_m = \beta_w + \Delta\beta$) and/or recovery rate ($\gamma_m = \gamma_w + \Delta\gamma$). Besides, hosts are distributed in two populations, labeled A and B, and coupled by migration. Throughout, we use superscripts to distinguish the two populations. We denote N^A and N^B, the total densities of populations A and B, respectively, such that $N^A + N^B = N$. Define μ^i, the probability for a host from the focal population $i \in \{A, B\}$ to visit the non-focal population. For the sake of simplicity, we assume that μ^i is independent of both the host's epidemiological status (S, I or R) and the pathogen's strain (w, m). Visits are assumed to be only instantaneous to-and-fro host movements, such as commuting [18, 25]. The force of infection is thus the sum of four types of intra- or inter-community interactions. Last, due to potential heterogeneities between the two populations – encapsulating e.g., local NPIs, behaviors –, we also allow phenotypic traits β and γ to vary between populations, so that the variant phenotypes in population i are given by $\beta_m^i = \beta_w^i + \Delta\beta^i$ and $\gamma_m^i = \gamma_w^i + \Delta\gamma^i$. **Fig. 2** provides an illustration of the model and notations are summarized in **Table 1** (more details are available in the **SI Appendix §S1-S2**).

2.2 Dynamics of the variant frequency across time

To follow pathogen evolution, we track $q^i = I^i_m/I^i$, the local frequency of the variant in each population $i \in \{A, B\}$ (we drop the time dependence in our notations for compactness and readability). Next, we assume that probabilities μ^A and μ^B are relatively small because contact rates are likely to be much higher within than between populations. Under this assumption, we show in **SI Appendix §S4** that the temporal dynamics of the logit-frequency of the variant logit $\left(q^i\right)$ can be approximated at the first order by

$$\frac{\mathrm{d}\,\mathrm{logit}\left(q^i\right)}{\mathrm{d}t} \approx \mathcal{S}^i + \frac{\mathcal{H}^i}{\underbrace{q^i\left(1-q^i\right)}_{\text{Genetic variance}}} \tag{1}$$

with

$$\begin{aligned}
\mathcal{S}^i &= \overbrace{\Delta\beta^i \frac{S^i}{N^i} - \Delta\gamma^i}^{\text{Local selection}} + \overbrace{\left[\mu^i\left(\Delta\beta^j \frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - \Delta\beta^i\right) + \frac{N^j}{N^i}\mu^j\Delta\beta^i\left(\frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - 1\right)\right]\frac{S^i}{N^i}}^{\text{Interaction between selection and migration}} \\
\mathcal{H}^i &= \underbrace{-\left(q^i - q^j\right)\left(\mu^i\beta^j_w + \frac{N^j}{N^i}\mu^j\beta^i_w\right)\frac{I^j/N^j}{I^i/N^i}\frac{S^i}{N^i}}_{\text{Between-pop. homogenization}}
\end{aligned} \tag{2}$$

Equations (1)-(2) are useful to better understand how selection and migration shape the evolution of the local frequency of the variant. First, $\mathcal{S}^i$ collects all the terms that depend on the phenotypic differences between the two strains and refer to the effects of local selection on the focal population as well as the interaction between selection and migration between populations. Second, the component $\mathcal{H}^i$ affects the local dynamics of the variant as soon as the variant frequency varies between the two populations (i.e., $q^i \neq q^j$).

As expected, in the absence of host mobility ($\mu^i = \mu^j = 0$), we recover the single-population scenario: $\mathcal{S}^i = \Delta\beta^i S^i/N^i - \Delta\gamma^i$ and $\mathcal{H}^i = 0$, where the dynamics of the variant logit-frequency is solely governed by the local selection and thus depends only on the phenotypic differences between the two strains and on the local availability of susceptibles [10, 12, 48, 49]. More generally, if the two strains differ in terms of transmission, $\mathcal{S}^i$ depends on potential asymmetries in population size and prevalence between the two populations because migration will give access to susceptible hosts from other populations and thus to new opportunities of transmission. Note that the interaction between selection and migration may arise even when the variant frequency is identical across populations (see also [48, 52, 53]).

When the local variant frequency varies across populations, the magnitude of $\mathcal{H}^i$ is governed by the intensity of migration, which depends on host mobility (governed by μ^i and μ^j) but also on the density of susceptibles in the focal population S^i. Indeed, what matters ultimately is the migration of the disease itself (i.e, inter-community transmission), which can

be reduced through a more limited availability of susceptibles. Interestingly, $\mathcal{H}^i$ is scaled by the inverse of the local genetic variance $q^i\left(1-q^i\right)$. Therefore, even with very low rates of migration, $\mathcal{H}^i$ cannot be neglected when the genetic variance is very small. This is in particular what we expect upon the introduction of a new variant in the focal population, as the variant is still rare at the early stages of the invasion (**Fig. 3** and **S1**) – this can be generalized to an arbitrary number of populations (see **Fig. S2** and **SI Appendix §S5**). Besides, these effects are magnified when the non-focal population is bigger (**Fig. S3-A**). Hence, even in the absence of selection ($\mathcal{S}^i = 0$), migration can still lead to a transient increase of the variant frequency until it reaches an equilibrium where $q^i = q^j$ (**Fig. S3-B**).

2.3 Dynamics of the differentiation across space

2.3.1 A measure of spatial differentiation

The genetic differentiation across space – hereafter referred to as spatial differentiation – can intuitively be measured as the difference $q^A - q^B$. Yet, the dynamics of this measure of spatial differentiation is difficult to interpret because it depends on the dynamics of genetic variances. In the following we focus on an alternative measure of differentiation [12, 51]

$$\mathcal{Q} = \frac{q^A}{1-q^A}\frac{1-q^B}{q^B},$$

such that $\ln\left(\mathcal{Q}\right) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$. Hence, when the variant frequency is the same in both populations (no differentiation), $\mathcal{Q} = 1$ and $\ln\left(\mathcal{Q}\right) = 0$. Under the assumption of weak migration, the temporal dynamics of $\ln\left(\mathcal{Q}\right)$ is approximated by (see **SI Appendix §S4**)

$$\frac{\mathrm{d}\ln\left(\mathcal{Q}\right)}{\mathrm{d}t} \approx \left(\mathcal{S}^A - \mathcal{S}^B\right) - \left(\mathcal{Q}-1\right)\left(\frac{\beta_w^A}{N^A}\mu^B + \frac{\beta_w^B}{N^B}\mu^A\right)\left(\frac{q^B}{q^A}\frac{I^B}{I^A}S^A + \frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right) \tag{3}$$

The above dynamical equation shows that the spatial differentiation is driven by two evolutionary forces. The first, $\mathcal{S}^A - \mathcal{S}^B$, is the difference in selection (both local selection and interaction term with migration) between population A and B and thus only matters if selection is heterogeneous across populations. The second tends to homogenize the frequency of the variant across populations and can be partitioned into two components: (i) the current level of differentiation $(\mathcal{Q}-1)$ and (ii) a term correlating with the intensity of migration. In the neutral case (i.e., no selection, $\mathcal{S}^A = \mathcal{S}^B = 0$), $\mathcal{Q}$ converges towards unity.

2.3.2 Homogeneous phenotypic differences across populations

First, let's assume for simplicity that the phenotypic traits of each strain are the same in both population A and B, i.e., $\Delta\beta^A = \Delta\beta^B = \Delta\beta$ and $\Delta\gamma^A = \Delta\gamma^B = \Delta\gamma$. As in **Fig. 1**

and **Fig. 3**, an initial spatial differentiation can be due to different times of introduction of the variant in each population. This can happen for instance if the two populations were previously isolated (e.g., lockdown, travel ban), and/or because of the inherent stochasticity of the introduction and emergence of a rare variant. This differentiation can be maintained over time in the absence of migration between the two populations (**Fig. S4-A**) but, otherwise, is always transient and vanishes in the longer term (**Fig. S4-C** and **E**). It is important to note that this scenario does not necessarily imply that selection is also homogeneous across populations. Interestingly, when the two strains differ in terms of transmission, a differential in the availability of susceptible hosts between the tow populations can generate a transient dynamics of $\mathcal{Q}$, even without pre-existing spatial differentiation (**Fig. S5**).

2.3.3 Heterogeneous phenotypic differences across populations

Second, let's assume that the phenotypic differences $\Delta\beta^i$ and $\Delta\gamma^i$ differ in each population $i \in \{A, B\}$. In this case, the difference $\left(\mathcal{S}^{\mathcal{A}} - \mathcal{S}^{\mathcal{B}}\right)$ can generate and maintain $\mathcal{Q}$, whose dynamics depends on the balance between selection and migration (**Fig. S4**, second column).

In the SIR model, the epidemic eventually dies out and both the wildtype and the variant go extinct. We can easily relax this assumption by using an SIRS model where infection does not provide lifelong immunity but recovered individuals return to the susceptible compartment (waning immunity). The mathematical derivation of our model is not sensitive to any extensions involving a recruitment of new susceptibles and their effects would be inherently captured by the density of susceptibles in the previous equations. In these conditions, $\mathcal{Q}$ might persist, converging towards approximately

$$\mathcal{Q} \approx 1 + \frac{\mathcal{S}^A - \mathcal{S}^B}{\left(\frac{\beta_w^A}{N^A}\mu^B + \frac{\beta_w^B}{N^B}\mu^A\right)\left(\frac{q^B}{q^A}\frac{I^B}{I^A}S^A + \frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right)}, \tag{4}$$

which can be viewed as a balance between the effects of selection (numerator) and the effects of migration (denominator). In particular, if the variant is adapted in one population but maladapted in the other, $\mathcal{Q}$ can be maintained in the long term, resulting in polymorphism maintenance (**Fig. S6**). Migration should not be too strong though, since it is known to erode local adaptation [54–56].

3 Discussion

Gene flow is a potent evolutionary force that tends to homogenize genetic variability across space via migration. It usually slows down adaptation because it counteracts the effect of local selection, but it may also promote adaptation when it introduces adaptive mutations [57]. Unraveling the tangled contributions of different evolutionary forces to the changes in

frequency can be carried out using genomic time series [58]. Here, we study the transient evolutionary dynamics of a pathogen population in a host metapopulation to understand the interplay between selection and migration on the dynamics of different strains of pathogens across time and space. We show that migration can bias or mimic the effect of selection, thereby complicating the estimation of the fitness of the new variant. In particular, the slope of the logit-frequency of a new variant (used as an estimation of its selection coefficient) can become locally very steep during the emergence of a variant because of the effect of migration from another population. To detect the influence of migration it is important to track the dynamics of the spatial differentiation among populations. We identify three main mechanisms that can generate spatial differentiation: (i) different times of introduction of the variant, (ii) a differential availability of susceptible hosts (if the variant differs in terms of transmissibility), and (iii) heterogeneous phenotypic differences across populations. While the first mechanism depends on the timing of introduction (and thus indirectly on the amount of migration), the latter two mechanisms are driven by spatial heterogeneity in selection. Selection and migration jointly shape the dynamics of spatial differentiation: if phenotypic differences are the same across populations, the spatial differentiation is always transient unless migration is absent. In contrast, when phenotypic differences vary across populations, the spatial differentiation can persist in the long term.

Understanding the interplay between selection and migration may yield some insights on the spread of SARS-CoV-2 variants. In particular, the analysis of the dynamics of the spatial differentiation can help explain the difference between the Alpha and the Delta variants in England (**Fig. 1**). First, the computation of pairwise measures of differentiation across the nine regions of England yields higher initial values for the Alpha variant (**Fig. 4**). As discussed above, higher levels of differentiation could be due to longer delays between successive introductions for Alpha, induced by a lower degree of connectivity between regions, i.e., fewer migration events (see also [59]). Second, the dynamics of the spatial differentiation may also reveal the amount of connectivity among regions. We show in **Fig. S7-A** and **B** that the rate at which the differentiation drops depends both on the initial amount of differentiation and the amount of migration. Interestingly, the patterns we obtain for SARS-CoV-2 are consistent with the hypothesis that migration was higher during the sweep of Delta because the drop in differentiation is steeper than for the Alpha variant (**Fig. S7-C**). The hypothesis that what may drive the difference between the two variants is driven by contrasted amounts of migration is further supported by the increase in human mobility during the sweep of Delta compared to Alpha according to Google mobility reports (**Fig. S8**) – with a relative increase of about +87%. Although Google mobility reports do not explicitly captured inter-region mobility, they reflect the progressive relaxation of NPIs throughout the pandemic. Additionally, the analysis of time series of neutral allele frequency have been used very recently to

estimate inter-community transmission rates, confirming the higher mixing rates during the spread of the Delta variant than during the spread of the Alpha variant [60]. This analysis further shows that the reproductive values (i.e. the relative contribution to the future of the pathogen population) vary among regions and across time. From the pathogen point of view, these reproductive values measure the relative "quality" of the different regions [53, 61]. Weighting the change in variant frequency by the reproductive values of different regions may thus provide an alternative way to model the heterogeneous and time-varying contribution of the different regions (i.e., source-sink dynamics) to the evolution of the pathogen [53, 62].

Our general model relies on numerous simplifying assumptions. For example, we assume that all hosts from the same population can commute with the same probability, regardless of whether they are infected or not. We thus neglect the potential influence the pathogen, or each strain, may have on the mobility of its host – e.g., if hosts tend to limit their movements when they present symptoms or are aware of being infected. In **SI Appendix §S6** and **Fig. S9**, we relax the assumption that infected hosts can commute with the same probability as non-infected hosts and show that it has little impact on the evolutionary dynamics of the variant. In our model, transmission rates depend on the population where the infection occurs, not on the population to which the infected host belongs. This is suitable to model scenarios with, e.g., differences in control strategies or behavioral differences where hosts align with local habits, but not for scenarios where transmission is an intrinsic property of the host, e.g., transmission-blocking vaccine. In contrast, population differences in recovery rates can be due to, e.g., differences in treatment/vaccination which affect the infection clearance. Note that spatial heterogeneity in host immunity profiles (natural or vaccine-induced) can also induce a heterogeneous selection on other phenotypic traits and favor immune-escape variants [14, 15, 52, 63, 64]. Though beyond the scope of this paper, demographic stochasticity can also affect the dynamics of pathogen variants. For instance, there was no clear consensus early in the COVID-19 pandemic on whether the global increase in the frequency of SARS-CoV-2 substitution D614G was due to positive selection or to a random founder effect [7, 8, 65–67], underscoring the complexity of disentangling the effects of selection from genetic drift. Last, while our model is a first step to take spatial structure into account, it remains an overly simplified representation of real-world spatial and social networks, which are characterized by complex and time-varying interaction processes [68]. More sophisticated models usually offer a better depiction of the reality, but general conclusions are also more difficult to draw, as they are also inherently less tractable [28]. Nonetheless, our parsimonious framework can be extended to an arbitrary number of populations (see **SI Appendix §S5** and **Fig. S2**), and migration probabilities can be calibrated to capture more complex scenarios – such as in, e.g., gravity models (from transportation theory [69]), which assume transmissions to be correlated positively with population sizes and negatively with separating distances [70–72].

To conclude, the dynamics of variant frequency and spatial differentiation contain valuable information about pathogen evolution. Our work underscores the importance of the availability of the distribution of variant frequency among different "classes" of hosts that may belong to different regions. We could extend our analysis to track the dynamics of differentiation across other types of class-structure (e.g. age or social group, vaccination status, ...). It would be interesting to explore further how the availability of such stratified data may yield a more quantitative approach to disentangle the effect of selection from the effects of migration among different classes of hosts, and thus refine the estimation of variant fitness. In our highly interconnected and interdependent world [73], human traveling behavior can indeed be quantified with increasingly accuracy at different scales – e.g., travel history data [74, 75], air-traffic data [30, 73], cell-phone mobility data [33, 76] –, yielding independent measures of the intensity of host mobility and contact among different types of hosts.

Supplementary material

Supplementary Figures (PDF)

Supplementary Information (SI Appendix) (PDF)

Data and code accessibility

Data that were used in this study, along with the scripts for the analyses (in R), are available in this GitHub repository. Numerical simulations were performed using R [77] version 4.4.2 (2024-10-31) and ordinary differential equations were numerically solved with the function `lsoda` from the package `deSolve` [78].

Acknowledgments

We thank Sébastien Lion for numerous inspiring discussions on the analysis of class-structured models. We also thank Erik Volz for sharing the data of the SARS-CoV-2 Delta variant.

Funding

We acknowledge the French Ministry of Higher Education and Research and the École Normale Supérieure Paris-Saclay for the PhD scholarship of WB. We also acknowledge support from the High Meadows Environmental Institute at Princeton University for the postdoctoral position of WB.

Conflict of interest

The authors declare no conflict of interest.

Table 1: **Notations.** The subscript $k \in \{w,\ m\}$ refers to the strain of the pathogen: wildtype strain w or the mutant strain m (or variant). The superscript $i \in \{A, B\}$ refer to the focal population i.

Term	**Definition**
$N,\ N^i$	Total population densities (global and local, respectively)
$S,\ S^i$	Susceptible hosts (global and local, respectively)
$I,\ I^i$	Infected/infectious hosts (global and local, respectively)
$I_k,\ I_k^i$	Hosts infected by strain k (global and local, respectively)
$R,\ R^i$	Recovered hosts (global and local, respectively)
$q,\ q^i$	Frequency of the variant m among infected hosts: $q = I_m/I$ (global) and $q^i = I_m^i/I^i$ (local)
β_k^i	*Per capita* transmission rate of strain k in population i
γ_k^i	*Per capita* recovery rate of strain k in population i
μ^i	Probability for hosts from population i to commute to the other population
$\Delta\beta^i,\ \Delta\gamma^i$	Phenotypic differences between the variant m and the wildtype w in population i: $\Delta\beta^i = \beta_m^i - \beta_w^i$ and $\Delta\gamma^i = \gamma_m^i - \gamma_w^i$
$\mathcal{S}^i$	Effects of local selection and interaction between selection and migration on the dynamics of the variant frequency in population i
$\mathcal{H}^i$	Effect of between-population homogenization on the dynamics of the variant frequency in population i
$\mathcal{Q}$	Spatial differentiation (between population A and B): $\mathcal{Q} = q^A/\left(1-q^A\right) \times \left(1-q^B\right)/q^B$

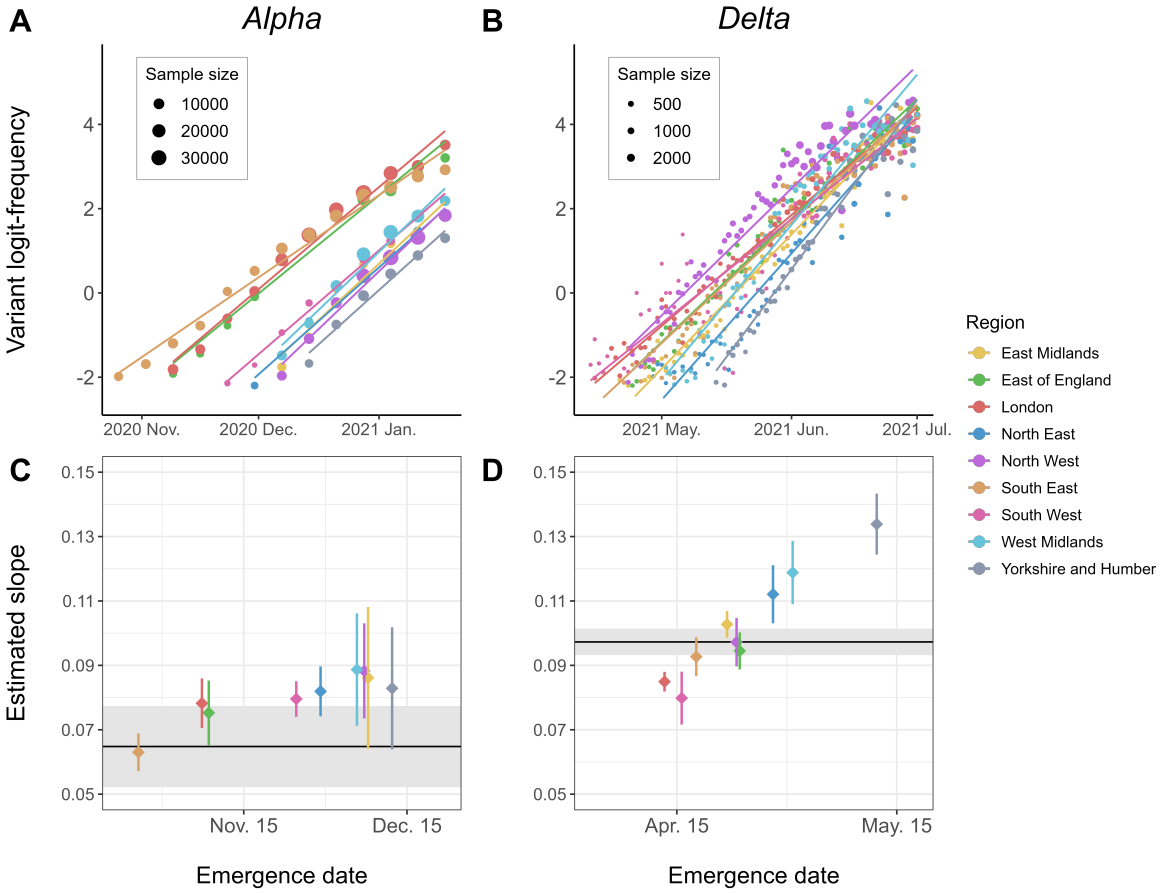

Figure 1: **Growth of the SARS-CoV-2 Alpha and Delta variants of concern in England.** Time series data (points) and linear regressions (lines) of the logit-frequency of two SARS-CoV-2 variants: (A) Alpha (lineage B.1.1.7) and (B) Delta (lineage B.1.617.2). Data for Alpha correspond to time series of regional frequencies of S Gene Target Failure (qPCR performed after swab sampling in the wider population) in England from [79]; data for Delta were shared by Erik Volz who previously used them in [4]. To minimize the influence of demographic stochasticity, we remove all frequencies lower than 0.1 or greater than 0.99. The second row (C-D) displays the estimated values of the slope of each linear regression (mean (diamonds) $\pm$ 95% confidence interval (segments), expressed per day) vs. emergence date (first date the frequency has reached 0.1). These slopes are typically used to estimate the selection coefficient of a variant. The horizontal line and the gray envelope represent the same estimation but for data aggregated at the national scale (no spatial structure).

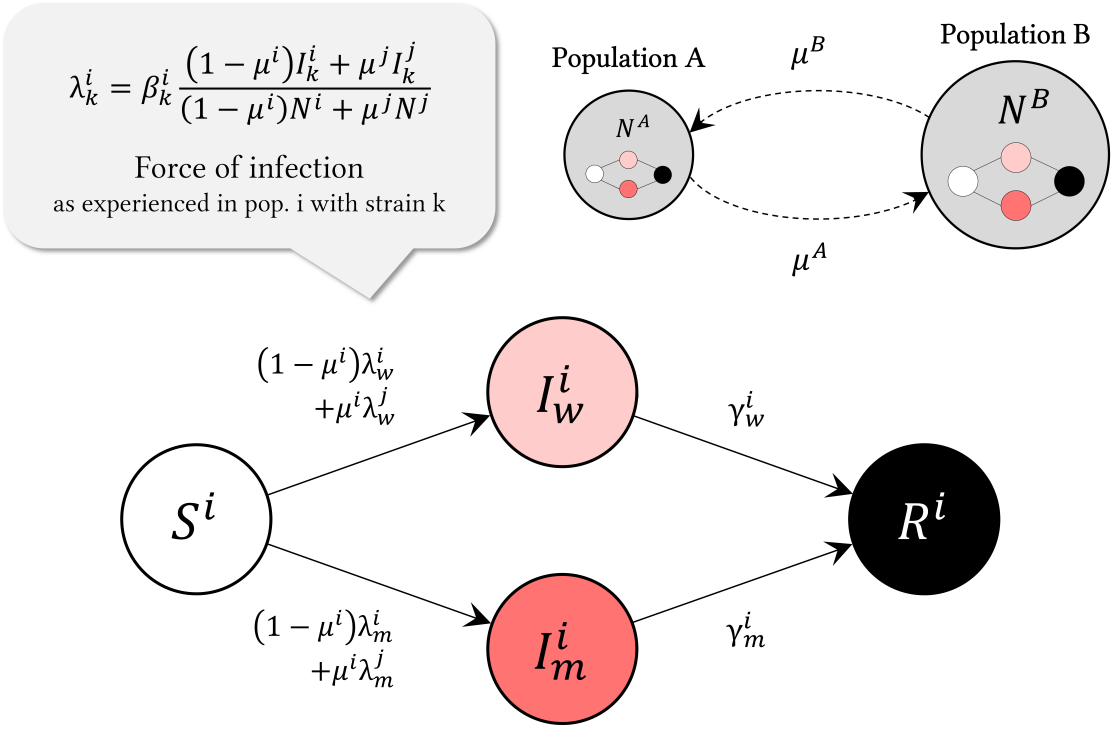

Figure 2: **Schematic of the compartmental model.** (Top right) We model a two-patch host metapopulation where individuals from population A (resp. B) may instantaneously visit population B (resp. A) with probability μ^A (resp. μ^B), or remain in their home population with complementary probabilities; circle sizes represent population densities, N^A and N^B (here for example, $N^B > N^A$). (Bottom) SIR model for the focal population $i \in \{A, B\}$ (accordingly, the non-focal population is denoted j); subscripts w and m refer to the wildtype and mutant strain (or variant) of the pathogen. (Top left) The force of infection λ_k^i represents the rate at which susceptible hosts are infected in population $i \in \{A, B\}$ with strain $k \in \{w, m\}$. See **Table 1** for notation and **SI Appendix §S1–S2** for more details.

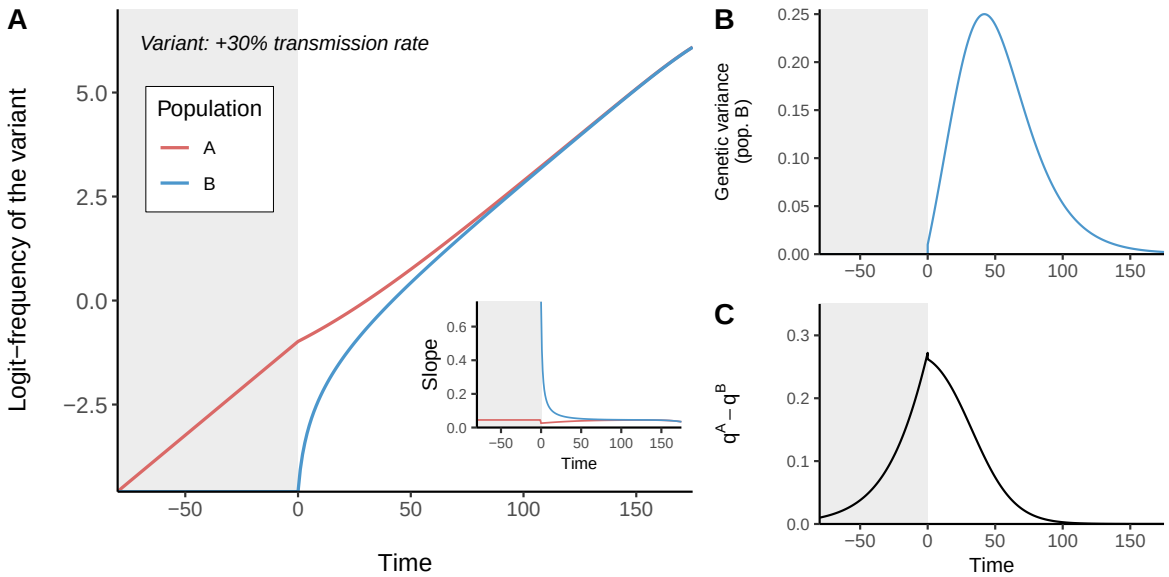

Figure 3: **Evolutionary dynamics of an emerging variant in a two-patch metapopulation.** The host metapopulation is divided in two populations, A and B. The variant has a +30% transmission rate selective advantage over the wildtype. We simulate the model with identical total population densities ($N^A = N^B = 100$), and parameter values (same for both populations): $\beta_w = 0.15$, $\beta_m = 0.195$ and $\gamma_w = \gamma_m = 0.1$. The wildtype is initially introduced at very low density (10^{-7}) in both populations and the variant is only introduced in population A such that it represents 1% of the infected hosts I^A – both populations are otherwise fully susceptible. Before $t = 0$, the two populations are isolated ($\mu^A = \mu^B = 0$, gray background). At $t = 0$, the variant is introduced in population B (1% of I^B) and, from that time point onward, hosts from population B may visit population A ($\mu^B = 0.1$). We plot over time the dynamics of: (A) the variant logit-frequency in population A (red) and B (blue) – the inset plot shows the dynamics of the slope using the approximation (1)-(2) (see also **Fig. S1**) –, (B) the genetic variance in population B, $q^B\left(1 - q^B\right)$ and (C) the frequency difference $q^A - q^B$. Simulations for which the variant has a longer duration of infectiousness exhibits similar trends (not shown).

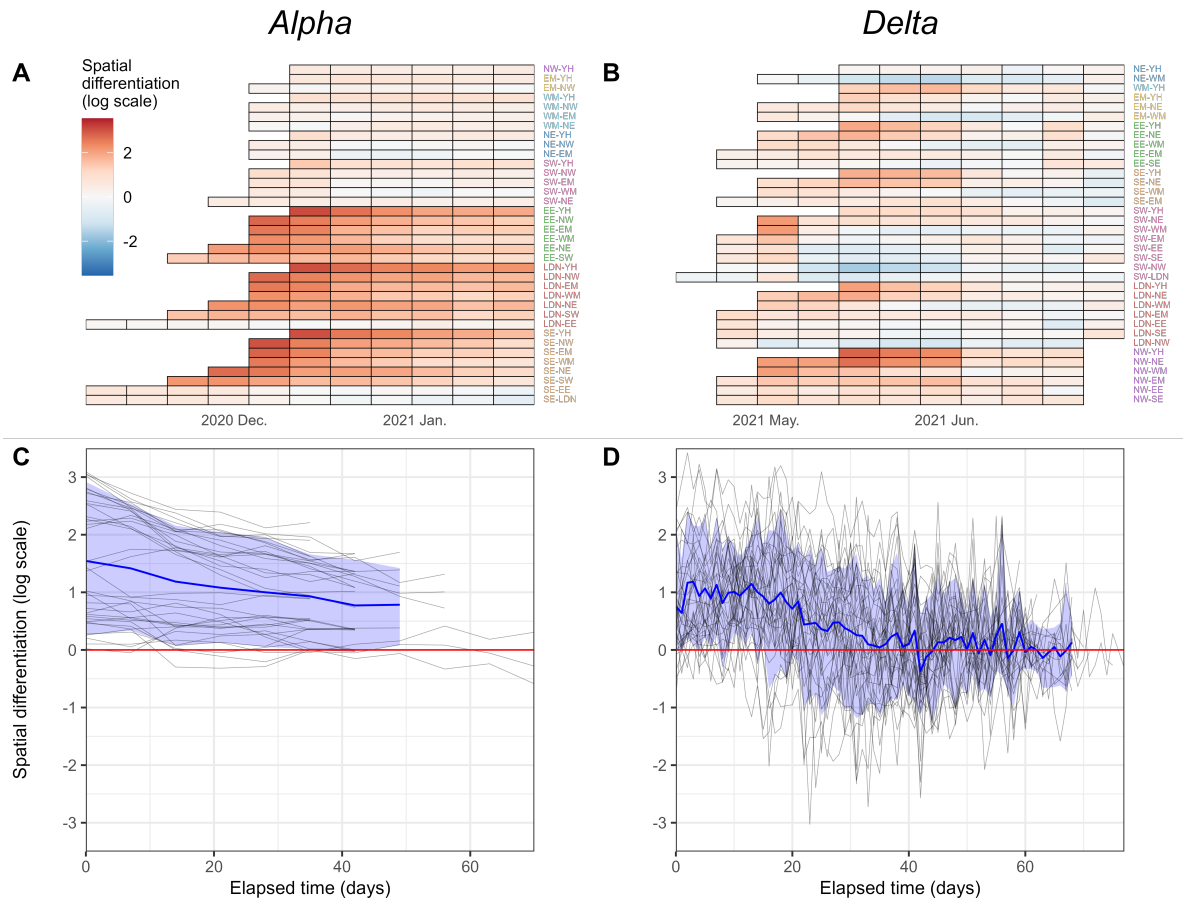

Figure 4: **Dynamics of the spatial differentiation during the sweep of the SARS-CoV-2 Alpha and Delta variants of concern.** Using the regional time series of the Alpha and Delta variant logit-frequency in the 9 regions of England (**Fig. 1**), we compute $\binom{9}{2} = 36$ pairwise regional differences for each available time point (log-differentiation $\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$). We choose the pairs so that the early values of $\ln(\mathcal{Q})$ are positive (see **Fig. S7**). Computed values are plotted (A-B) by date and region pair (colored according to the first region) and (C-D) against the time elapsed since the first value of each region pair. The thick black line represents the mean and the envelopes the empirical interval between the 10% and 90% quantiles (only computed with at least 10 values). The horizontal red line indicates the absence of differentiation ($\ln(\mathcal{Q}) = 0$). For better visualization, we use weekly data for the Delta variant in A. Abbreviations: East Midlands (EM), East of England (EE), London (LDN), North East (NE), North West (NW), South East (SE), South West (SW), West Midlands (WM) and Yorkshire and Humber (YH).

References

1. Fisher, R. A. *The genetical theory of natural selection: a complete variorum edition* (Oxford University Press, 1999).
2. Chevin, L.-M. On measuring selection in experimental evolution. *Biology letters* **7,** 210–213 (2011).
3. Boyle, L. *et al.* Selective sweeps in SARS-CoV-2 variant competition. *Proceedings of the National Academy of Sciences* **119,** e2213879119 (2022).
4. Volz, E. Fitness, growth and transmissibility of SARS-CoV-2 genetic variants. *Nature Reviews Genetics* **24,** 724–734 (2023).
5. Donker, T. *et al.* Estimation of SARS-CoV-2 fitness gains from genomic surveillance data without prior lineage classification. *Proceedings of the National Academy of Sciences* **121,** e2314262121 (2024).
6. Lythgoe, K. A. *et al.* Lineage replacement and evolution captured by 3 years of the United Kingdom Coronavirus (COVID-19) Infection Survey. *Proceedings of the Royal Society B* **290,** 20231284 (2023).
7. Grubaugh, N. D., Hanage, W. P. & Rasmussen, A. L. Making sense of mutation: what D614G means for the COVID-19 pandemic remains unclear. *Cell* **182,** 794–795 (2020).
8. Volz, E. *et al.* Evaluating the effects of SARS-CoV-2 spike mutation D614G on transmissibility and pathogenicity. *Cell* **184,** 64–75 (2021).
9. Yu, Q. *et al.* Lineage frequency time series reveal elevated levels of genetic drift in SARS-CoV-2 transmission in England. *Plos Pathogens* **20,** e1012090 (2024).
10. Day, T., Gandon, S., Lion, S. & Otto, S. P. On the evolutionary epidemiology of SARS-CoV-2. *Current Biology* **30,** R849–R857 (2020).
11. Otto, S. P. *et al.* The origins and potential future of SARS-CoV-2 variants of concern in the evolving COVID-19 pandemic. *Current Biology* **31,** R918–R929 (2021).
12. Benhamou, W., Lion, S., Choquet, R. & Gandon, S. Phenotypic evolution of SARS-CoV-2: a statistical inference approach. *Evolution* **77,** 2213–2223 (2023).
13. Walter, A. & Lion, S. Epidemiological and evolutionary consequences of periodicity in treatment coverage. *Proceedings of the Royal Society B* **288,** 20203007 (2021).
14. Day, T., Kennedy, D. A., Read, A. F. & Gandon, S. Pathogen evolution during vaccination campaigns. *PLoS biology* **20,** e3001804 (2022).
15. Walter, A., Gandon, S. & Lion, S. Effect of unequal vaccination coverage and migration on long-term pathogen evolution in a metapopulation. *Journal of Evolutionary Biology* **37,** 189–200 (2024).

16. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* **115,** 700–721 (1927).

17. Anderson, R. M. & May, R. M. *Infectious diseases of humans: dynamics and control* (Oxford University Press, 1991).

18. Lipshtat, A., Alimi, R. & Ben-Horin, Y. Commuting in metapopulation epidemic modeling. *Scientific reports* **11,** 15198 (2021).

19. Ball, F., Mollison, D. & Scalia-Tomba, G. Epidemics with two levels of mixing. *The Annals of Applied Probability,* 46–89 (1997).

20. Ferguson, N. M., Galvani, A. P. & Bush, R. M. Ecological and immunological determinants of influenza evolution. *Nature* **422,** 428–433 (2003).

21. Grenfell, B. & Harwood, J. (Meta)population dynamics of infectious diseases. *Trends in ecology & evolution* **12,** 395–399 (1997).

22. Arino, J. in *Modeling and dynamics of infectious diseases* 64–122 (World Scientific, 2009).

23. Lajmanovich, A. & Yorke, J. A. A deterministic model for gonorrhea in a nonhomogeneous population. *Mathematical Biosciences* **28,** 221–236 (1976).

24. Hethcote, H. W. An immunization model for a heterogeneous population. *Theoretical population biology* **14,** 338–349 (1978).

25. Post, W., DeAngelis, D. & Travis, C. Endemic disease in environments with spatially heterogeneous host populations. *Mathematical Biosciences* **63,** 289–302 (1983).

26. May, R. M. & Anderson, R. M. Spatial heterogeneity and the design of immunization programs. *Mathematical Biosciences* **72,** 83–111 (1984).

27. Bolker, B. & Grenfell, B. T. Space, persistence and dynamics of measles epidemics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **348,** 309–320 (1995).

28. Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proceedings of the National Academy of Sciences* **102,** 11157–11162 (2005).

29. Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *Journal of theoretical biology* **251,** 450–467 (2008).

30. Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *science* **342,** 1337–1342 (2013).

31. Yuksel, M. K., Remien, C. H., Karki, B., Bull, J. J. & Krone, S. M. Vector dynamics influence spatially imperfect genetic interventions against disease. *Evolution, Medicine, and Public Health* **9,** 1–10 (2021).

32. Roques, L., Bonnefon, O., Baudrot, V., Soubeyrand, S. & Berestycki, H. A parsimonious approach for spatial transmission and heterogeneity in the COVID-19 propagation. *Royal Society Open Science* **7,** 201382 (2020).

33. Le Treut, G. *et al.* A high-resolution flux-matrix model describes the spread of diseases in a spatial network and the effect of mitigation strategies. *Scientific Reports* **12,** 15946 (2022).

34. Débarre, F., Bonhoeffer, S. & Regoes, R. R. The effect of population structure on the emergence of drug resistance during influenza pandemics. *Journal of the Royal Society Interface* **4,** 893–906 (2007).

35. Freire, T. F. A., Hu, Z., Wood, K. B. & Gjini, E. Modeling spatial evolution of multidrug resistance under drug environmental gradients. *PLOS Computational Biology* **20,** e1012098 (2024).

36. Boots, M. & Sasaki, A. 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **266,** 1933–1938 (1999).

37. Boots, M., Hudson, P. J. & Sasaki, A. Large shifts in pathogen virulence relate to host population structure. *Science* **303,** 842–844 (2004).

38. Lion, S. & Boots, M. Are parasites "prudent"in space? *Ecology Letters* **13,** 1245–1255 (2010).

39. Osnas, E. E., Hurtado, P. J. & Dobson, A. P. Evolution of pathogen virulence across space during an epidemic. *The American Naturalist* **185,** 332–342 (2015).

40. Berngruber, T. W., Lion, S. & Gandon, S. Spatial structure, transmission modes and the evolution of viral exploitation strategies. *PLoS pathogens* **11,** e1004810 (2015).

41. Griette, Q., Raoul, G. & Gandon, S. Virulence evolution at the front line of spreading epidemics. *Evolution* **69,** 2810–2819 (2015).

42. Zurita-Gutiérrez, Y. H. & Lion, S. Spatial structure, host heterogeneity and parasite virulence: implications for vaccine-driven evolution. *Ecology letters* **18,** 779–789 (2015).

43. Lion, S. & Gandon, S. Evolution of spatially structured host–parasite interactions. *Journal of Evolutionary Biology* **28,** 10–28 (2015).

44. Lion, S. & Gandon, S. Spatial evolutionary epidemiology of spreading epidemics. *Proceedings of the Royal Society B: Biological Sciences* **283,** 20161170 (2016).

45. Griette, Q., Alfaro, M., Raoul, G. & Gandon, S. Evolution and spread of multiadapted pathogens in a spatially heterogeneous environment. *Evolution Letters* **8,** 427–436 (2024).

46. Lion, S., Sasaki, A. & Boots, M. Extending eco-evolutionary theory with oligomorphic dynamics. *Ecology Letters* **26,** S22–S46 (2023).

47. Day, T. & Proulx, S. R. A general theory for the evolutionary dynamics of virulence. *The American Naturalist* **163,** E40–E63 (2004).

48. Day, T. & Gandon, S. Insights from Price's equation into evolutionary epidemiology. *Disease evolution: models, concepts, and data analyses* **71,** 23–44 (2006).

49. Day, T. & Gandon, S. Applying population-genetic models in theoretical evolutionary epidemiology. *Ecology Letters* **10,** 876–888 (2007).

50. Berngruber, T. W., Froissart, R., Choisy, M. & Gandon, S. Evolution of virulence in emerging epidemics. *PLoS pathogens* **9,** e1003209 (2013).

51. Benhamou, W. *et al.* Evolution of virulence in emerging epidemics: from theory to experimental evolution and back. *Virus Evolution* **10,** veae069 (2024).

52. Gandon, S. & Day, T. The evolutionary epidemiology of vaccination. *Journal of the Royal Society Interface* **4,** 803–817 (2007).

53. Lion, S. & Gandon, S. Evolution of class-structured populations in periodic environments. *Evolution* **76,** 1674–1688 (2022).

54. Blanquart, F., Gandon, S. & Nuismer, S. The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of evolutionary biology* **25,** 1351–1363 (2012).

55. Blanquart, F., Kaltz, O., Nuismer, S. L. & Gandon, S. A practical guide to measuring local adaptation. *Ecology letters* **16,** 1195–1205 (2013).

56. Mirrahimi, S. & Gandon, S. Evolution of specialization in heterogeneous environments: equilibrium between selection, mutation and migration. *Genetics* **214,** 479–491 (2020).

57. Slatkin, M. Gene flow and the geographic structure of natural populations. *Science* **236,** 787–792 (1987).

58. Simon, A. & Coop, G. The contribution of gene flow, selection, and genetic drift to five thousand years of human allele frequency change. *Proceedings of the National Academy of Sciences* **121,** e2312377121 (2024).

59. Tegally, H. *et al.* Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern. *Cell* **186,** 3277–3290 (2023).

60. Okada, T., Isacchini, G., Yu, Q. & Hallatschek, O. Uncovering heterogeneous intercommunity disease transmission from neutral allele frequency time series. *medRxiv,* 2024–12 (2024).

61. Gandon, S. Evolution of multihost parasites. *Evolution* **58,** 455–469 (2004).

62. Lion, S. Class structure, demography, and selection: reproductive-value weighting in nonequilibrium, polymorphic populations. *The American Naturalist* **191,** 620–637 (2018).

63. McLeod, D. V. & Gandon, S. Effects of epistasis and recombination between vaccine-escape and virulence alleles on the dynamics of pathogen adaptation. *Nature ecology & evolution* **6,** 786–793 (2022).

64. Gog, J. R., Hill, E. M., Danon, L. & Thompson, R. N. Vaccine escape in a heterogeneous population: insights for SARS-CoV-2 from a simple model. *Royal Society Open Science* **8,** 210530 (2021).

65. Hodcroft, E. B. *et al.* Spread of a SARS-CoV-2 variant through Europe in the summer of 2020. *Nature* **595,** 707–712 (2021).

66. Korber, B. *et al.* Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182,** 812–827 (2020).

67. Plante, J. A. *et al.* Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592,** 116–121 (2021).

68. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Reviews of modern physics* **87,** 925–979 (2015).

69. Erlander, S. & Stewart, N. F. *The gravity model in transportation analysis: theory and extensions* (Vsp, 1990).

70. Murray, G. D. & Cliff, A. D. A stochastic model for measles epidemics in a multi-region setting. *Transactions of the Institute of British Geographers,* 158–174 (1977).

71. Xia, Y., Bjørnstad, O. N. & Grenfell, B. T. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *The American Naturalist* **164,** 267–281 (2004).

72. Balcan, D. *et al.* Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the national academy of sciences* **106,** 21484–21489 (2009).

73. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proceedings of the national academy of sciences* **101,** 15124–15129 (2004).

74. Lemey, P. *et al.* Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nature communications* **11,** 5110 (2020).

75. Butera, Y. *et al.* Genomic sequencing of SARS-CoV-2 in Rwanda reveals the importance of incoming travelers on lineage diversity. *Nature communications* **12,** 5705 (2021).

76. Kraemer, M. U. *et al.* Spatiotemporal invasion dynamics of SARS-CoV-2 lineage B.1.1.7 emergence. *Science* **373,** 889–895 (2021).

77. R Core Team. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing (Vienna, Austria, 2024). https://www.R-project.org/.

78. Soetaert, K., Petzoldt, T. & Setzer, R. W. Solving differential equations in R: package deSolve. *Journal of statistical software* **33,** 1–25 (2010).

79. Public Health England. *Investigation of novel SARS-COV-2 variant 202012/01: technical briefing 5* tech. rep. (2020). https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/959426/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5.pdf.

Supplementary figures

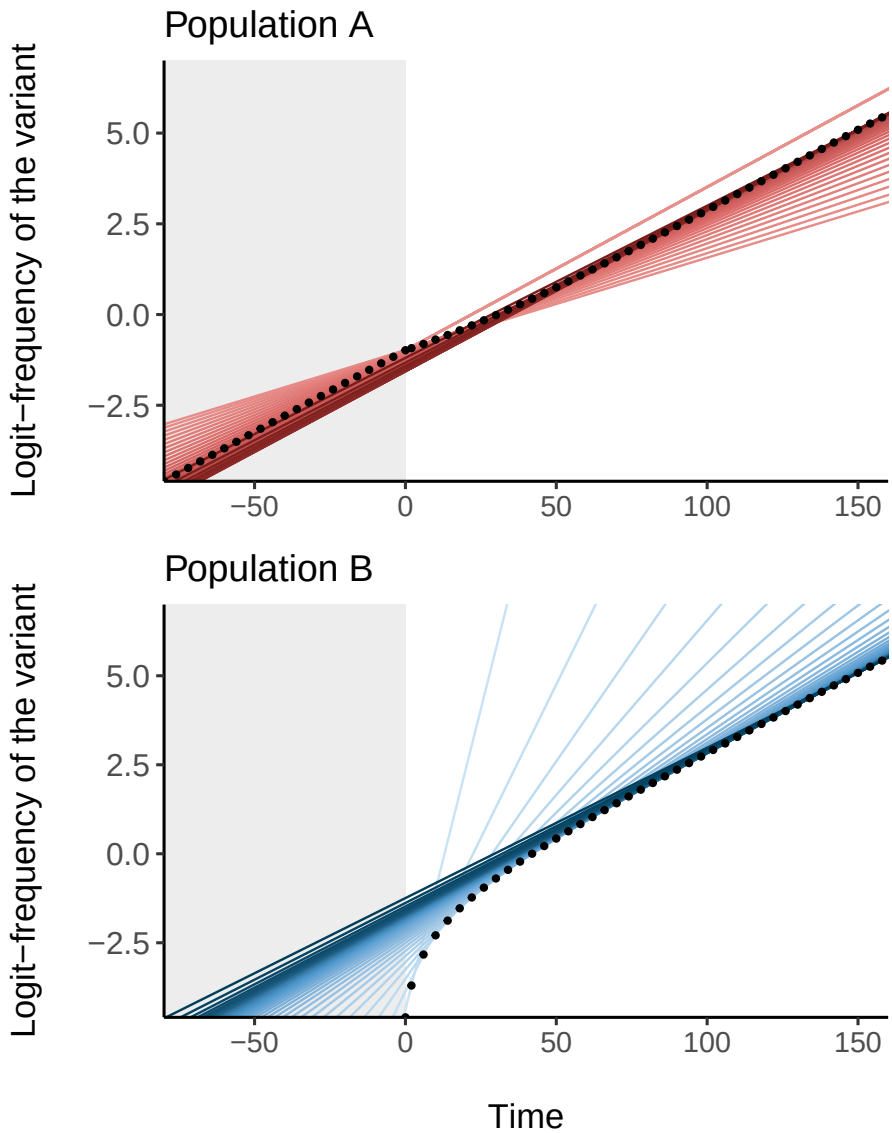

Figure S1: **Dynamics of the slope of the variant logit-frequency in a two-patch metapopulation.** This figure is built upon the simulation presented in the main text in **Fig. 1**. The host metapopulation is divided in two populations, A (top) and B (bottom). The variant has a +30% transmission rate selective advantage over the wildtype. We simulate the model with identical total population densities ($N^A = N^B = 100$), and parameter values (same for both populations): $\beta_w = 0.15$, $\beta_m = 0.195$ and $\gamma_w = \gamma_m = 0.1$. The wildtype is initially introduced at very low density (10^{-7}) in both populations and the variant is only introduced in population A such that it represents 1% of the infected hosts I^A – both populations are otherwise fully susceptible. Before $t = 0$, the two populations are isolated ($\mu^A = \mu^B = 0$, gray background). At $t = 0$, the variant is introduced in population B (1% of I^B) and, from that time point onward, hosts from population B may visit population A ($\mu^B = 0.1$). Tangents (red and blue lines) to each logit-frequency values (black points) indicate the corresponding current rate of change. Tangent slopes are approximated using equations (1)-(2) from the main text; color shades (from lighter to darker) represent the course of time.

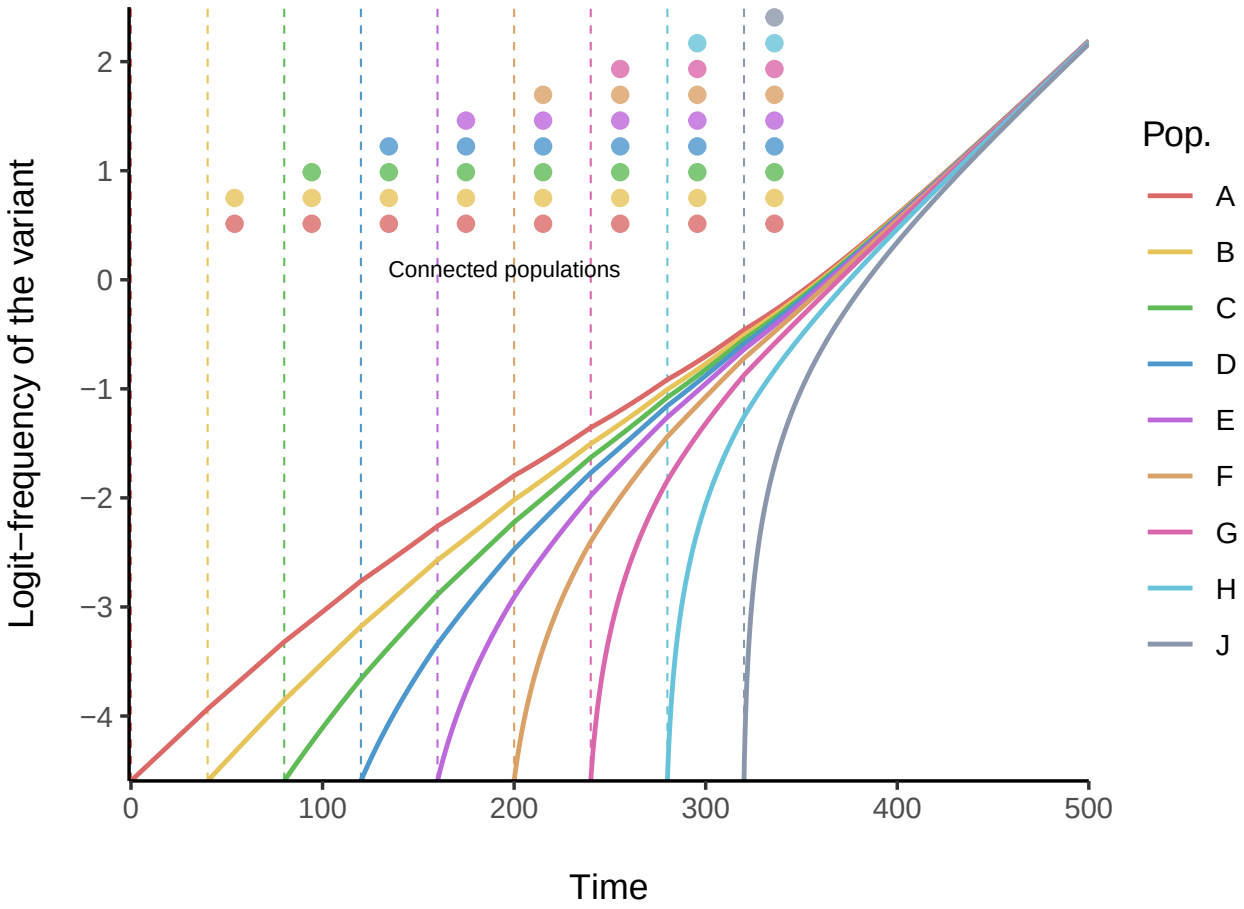

Figure S2: **Evolutionary dynamics of an emerging variant in a nine-patch host metapopulation.** We extend in **SI Appendix §S5** the model presented in the main text to any number of populations. Here, the host metapopulation is divided in nine populations (such as the nine regions of England). For the sake of simplicity, the variant has only a transmission advantage and the phenotypes of both strains do not vary across populations: $\beta_w = 0.11$, $\beta_m = 0.1265$ (+15% transmission advantage) and $\gamma_w = \gamma_m = 0.1$. We also assume same total density for each population (100) and same probability of migration (0.01). At $t = 0$, the wildtype is introduced at low density (10^{-7}) in all populations, but the variant is only introduced in population A at a lower density (1% of infected hosts I^A) – all populations are otherwise fully susceptible. All populations are first isolated from each other. Populations are then sequentially connected by migration, starting from population A and B. Every time a population is connected (as indicated by the vertical dashed lines), we introduce the variant such that it represents 1% of the infected hosts in that population.

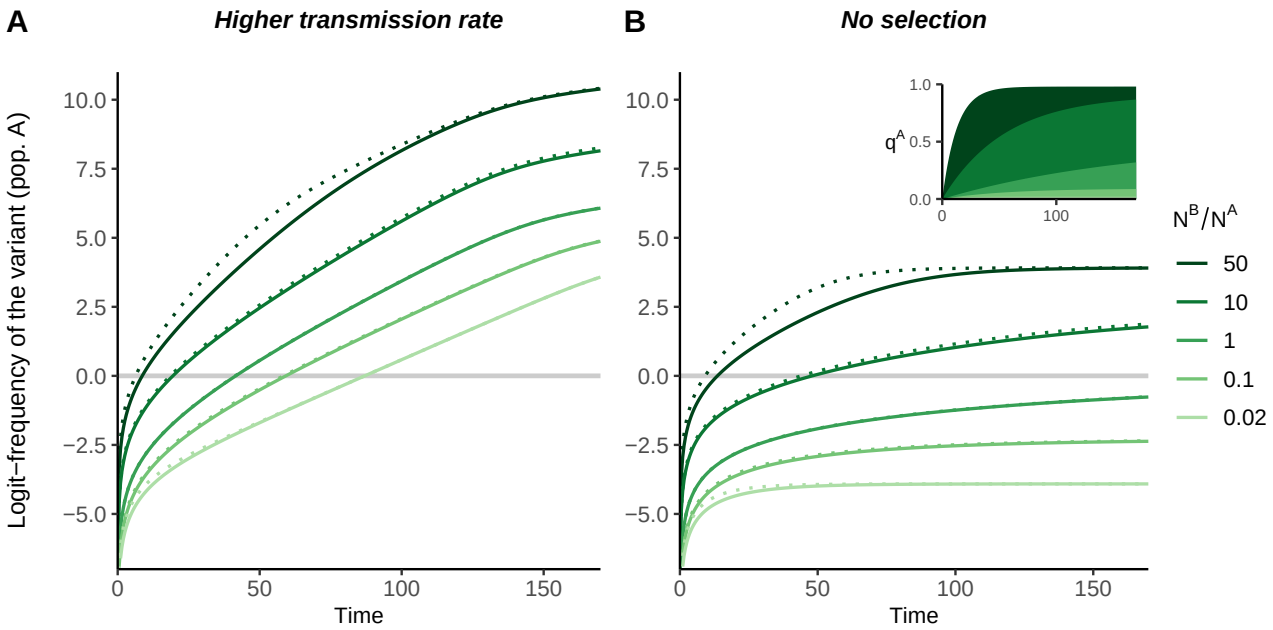

Figure S3: **The effect of population sizes on the dynamics of the variant frequency in the focal population.** The host metapopulation is divided in two populations, A and B. Varying the total population density ratio N^B/N^A (in all cases, $N^A = 100$), we simulate our full model (**SI Appendix §S2.2**, system (S4), solid lines) along with our first-order approximation under weak migration (dotted lines). Trajectories of the logit-frequency of a variant under: (A) positive selection (+30% transmission rate, $\beta_m > \beta_w \;\Rightarrow\; \Delta\beta > 0$), (B) neutral evolution ($\beta_m = \beta_w \;\Rightarrow\; \Delta\beta = 0$). In all simulations, we take (same for both populations): $\beta_w = 0.15$, $\gamma_w = \gamma_m = 0.1$ and $\mu^A = \mu^B = 10^{-2}$. At $t = 0$, the wildtype is introduced at very low prevalence in population A ($I_w^A/N^A = 10^{-6}$ and $I_m^A = 0$) and the variant is introduced at the same prevalence in population B ($I_m^B/N^B = 10^{-6}$ and $I_w^B = 0$) – both populations are otherwise fully susceptible.

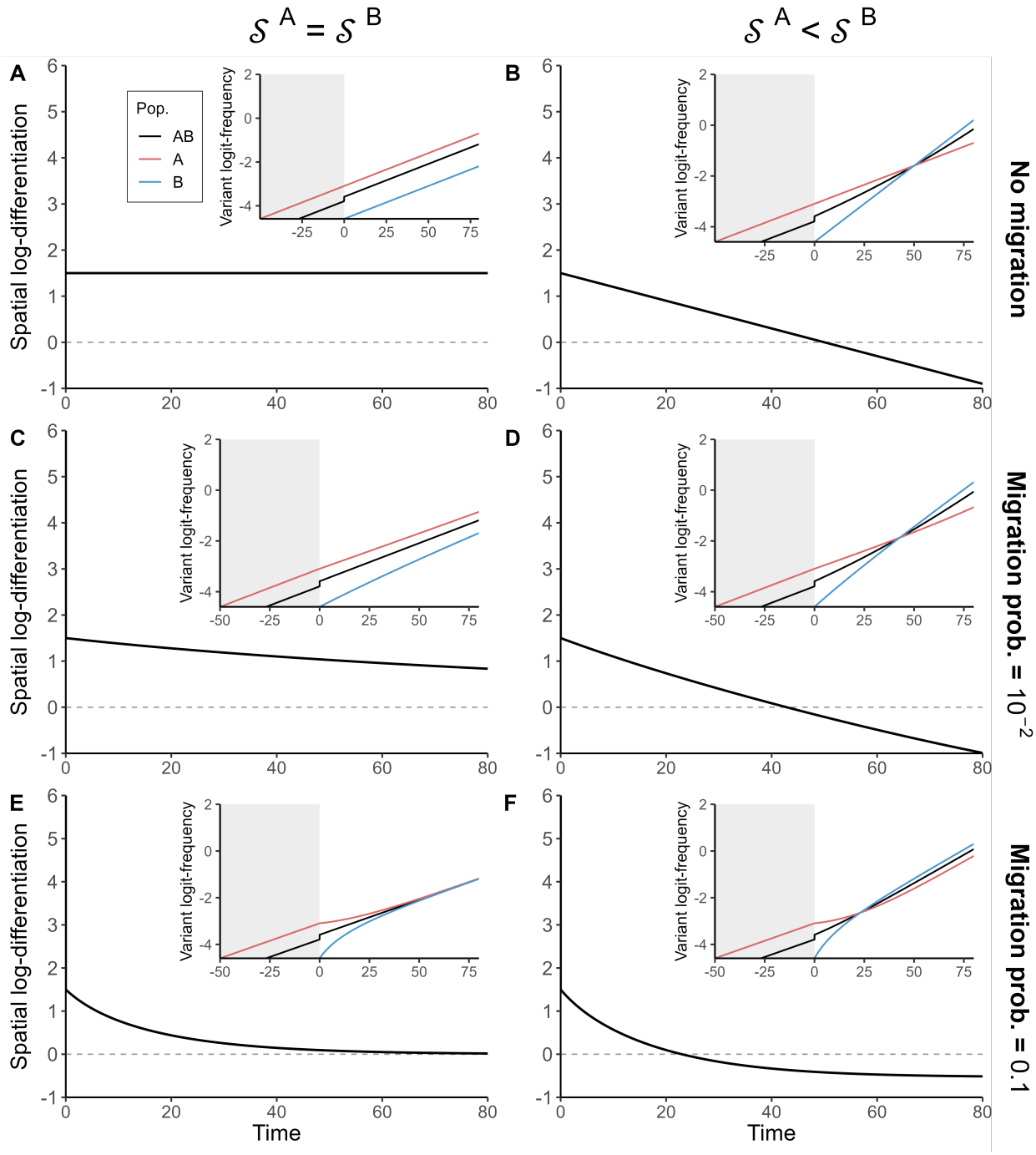

Figure S4: **The transient dynamics of spatial differentiation and the migration–selection balance.** For the sake of simplicity, we assume that both strains share the same transmissibility. This allows us to set aside the effects due to the availability of susceptible hosts. Fixed parameter values are: $\forall i \in \{A, B\}$, $N^i = 100$, $\beta_w^i = \beta_m^i = 0.15$ and $\gamma_w^i = 0.1$. (A-E) The recovery rate of the variant do not vary across populations (homogeneous selection, $\mathcal{S}^A = \mathcal{S}^B$): $\forall i \in \{A, B\}$, $\gamma_m^i = 0.07$ ($\mathcal{S}^i = -\Delta\gamma^i = 0.03$). (B-F) The variant is more adapted in population B (heterogeneous selection, $\mathcal{S}^A < \mathcal{S}^B$): $\gamma_m^A = 0.07$ ($\mathcal{S}^A = -\Delta\gamma^A = 0.03$) and $\gamma_m^B = 0.04$ ($\mathcal{S}^B = -\Delta\gamma^B = 0.06$). The wildtype is initially introduced in both populations at the same density (10^{-4}) and the variant is only introduced in population A such that it represents 1% of infected hosts I^A – both populations are otherwise fully susceptible. Until $t = 0$, the two populations are isolated ($\mu^A = \mu^B = 0$) and the variant is only present in population A (gray background). At $t = 0$, the variant is introduced in population B at low density (1% of I^B) and the two populations (A-B) stay isolated, or are coupled with (C-D) low or (E-F) higher levels of migration.

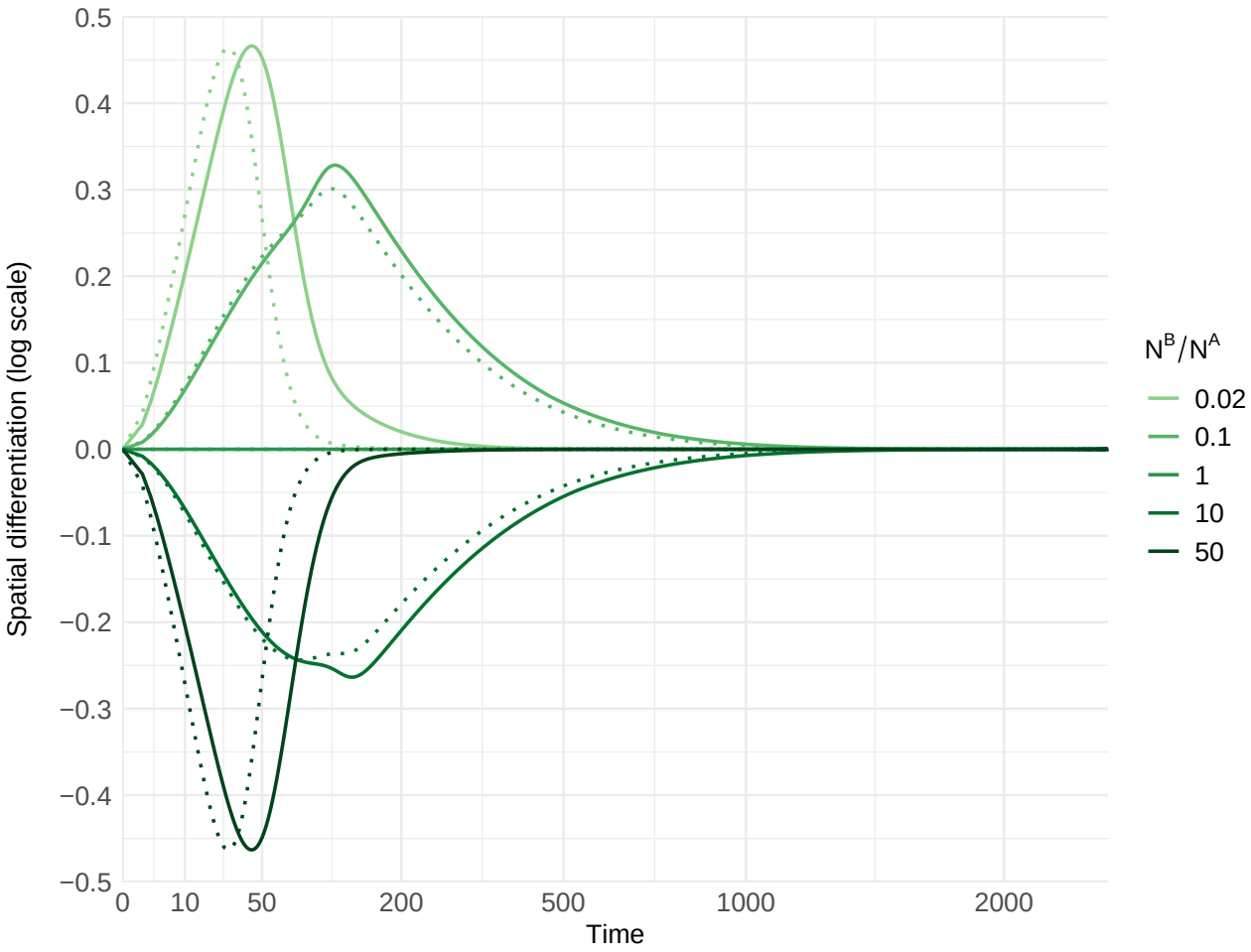

Figure S5: **The transient disruption of spatial differentiation.** The host metapopulation is divided in two populations, A and B. Varying the total population density ratio N^B/N^A (in all cases, $N^A = 100$), we simulate our full model (**SI Appendix §S2.2**, system (S4), solid lines) with parameter values (same for both populations): $\beta_w = 0.15$, $\beta_m = 0.195$ (+30% transmission rate, $\Delta\beta = 0.045$), $\gamma_w = \gamma_m = 0.1$ ($\Delta\gamma = 0$) and $\mu^A = \mu^B = 10^{-2}$. We also run these simulations using our first-order approximation under weak migration (dotted lines). At $t = 0$, the wildtype and the variant are introduced at the same low *density* in both populations ($\forall i \in \{A, B\},\ I_w^i = I_m^i = 10^{-3}$) – both populations are otherwise fully susceptible. The log-differentiation $\ln(\mathcal{Q})$ is then calculated as $\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$. Although there is no spatial differentiation initially ($\ln(\mathcal{Q}) = 0$), the spatial differentiation is then transiently disrupted as soon as $N^A \neq N^B$ because the initial differential in the proportions of susceptible hosts results in a differential in selection when $\Delta\beta \neq 0$.

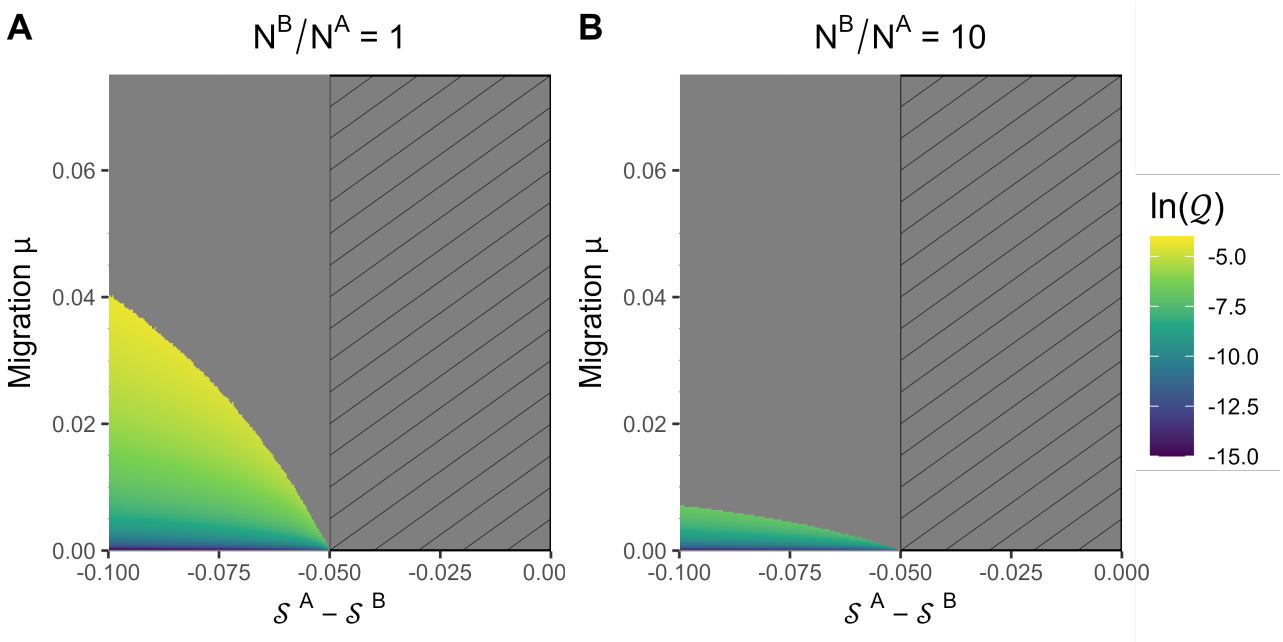

Figure S6: **Long-term spatial differentiation and local adaptation.** We extend the previous SIR model presented in the main text to an SIRS model where recovered individuals return to the susceptible compartment at rate $\omega = 0.01$ (waning immunity). This enables the pathogen to avoid extinction and the system to eventually reach an endemic equilibrium. For the sake of simplicity, we assume that both strains share the same transmissibility and only differ in terms of recovery rates. This allows us to set aside the selective effects due to the availability of susceptible hosts such that the difference $\mathcal{S}^A - \mathcal{S}^B$ reduces to $-\Delta\gamma^A + \Delta\gamma^B$. Parameter values shared by both populations are: $\beta_w = \beta_m = 0.15$, $\gamma_w = 0.1$ and $\gamma_m^B = 0.05$; the variant is thus selected for in population B ($\mathcal{S}^B = -\Delta\gamma^B > 0$). We compute the long-term spatial differentiation by varying the parameter value of γ_m^A from 0.05 to 0.15 and of the migration probability $\mu = \mu^A = \mu^B$ from 0 to 0.075. In population A, depending on the value of γ_m^A, the variant is either selected for (hatching background, $\mathcal{S}^A = -\Delta\gamma^A > 0$) or against ($\mathcal{S}^A = -\Delta\gamma^A < 0$). Simulations were performed with (A) $N^A = N^B = 100$ or (B) $N^A = 100$ and $N^B = 1000$. The gray color corresponds to cases where a strain has numerically reached fixation in a population.

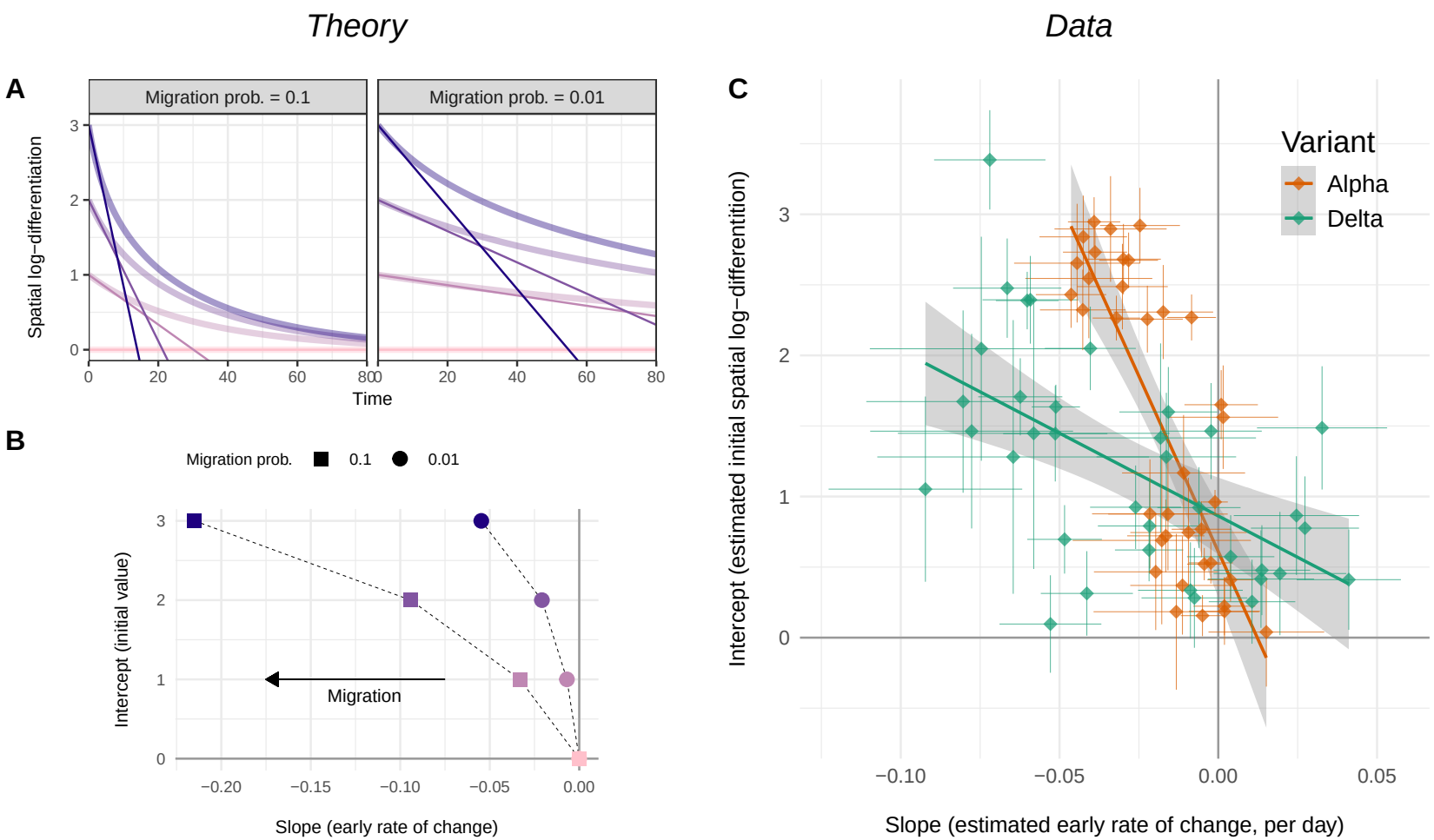

Figure S7: **Approximations of the early dynamics of the spatial differentiation.** (A-B) Theory. According to equation (3), the dynamics of the spatial differentiation $\mathcal{Q}$ depends on the differential in selection between the two populations, the current level of differentiation and the intensity of migration. In the following we assume that selection is homogeneous across populations and we only focus on the last two. (A) Dynamics of the spatial log-differentiation $\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$ starting from different initial conditions and with two level of migration ($\mu^A = \mu^B = 0.01$ or 0.1) – parameter values (same for all populations): $\beta_w = \beta_m = 0.15$, $\gamma_w = 0.1$ and $\gamma_m = 0.05$. Straight lines represent linear approximations of the early dynamics of $\ln(\mathcal{Q})$ and we reported the corresponding slopes and intercepts in panel B. (C) Data. We estimated the early dynamics of the spatial log-differentiation (slope and intercept) of the Alpha and Delta SARS-CoV-2 variants of concern using simple linear regressions. We used the data and the pairs of regions presented in **Fig. 4-C** and **D** for which elapsed time did not exceed 35 days. We oriented each pair so that the point estimate of the intercept was always positive. Interestingly, the patterns we obtain for SARS-CoV-2 are consistent with the hypothesis that migration was higher during the sweep of Delta because the drop in differentiation is steeper than for the Alpha variant (straight lines and gray envelopes show a linear regression on point estimates).

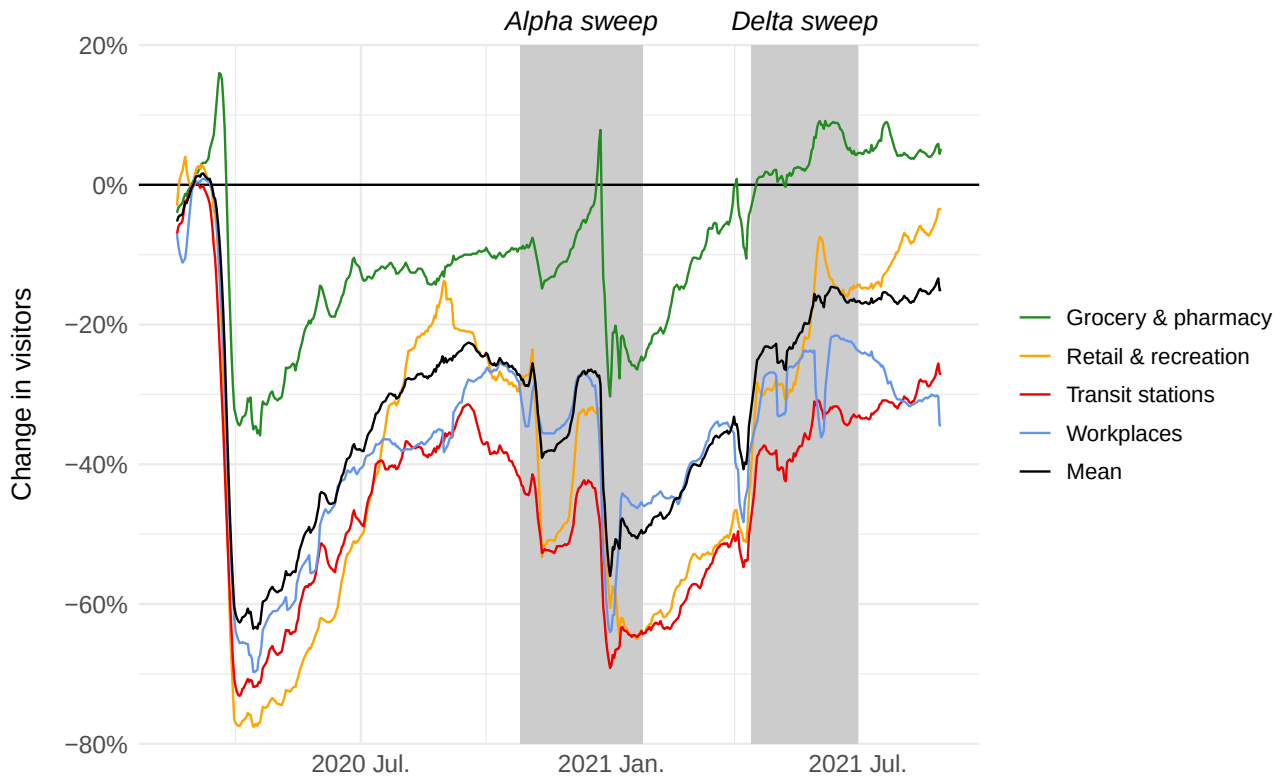

Figure S8: **Change in visitors in the United Kingdom (Google mobility reports).** Changes in the number of visitors compared to baseline days (median value from the 5-week period from Jan. 3 to Feb. 6, 2020). Mobility data were downloaded from the website *Our World in Data* (https://ourworldindata.org/covid-google-mobility-trends) and we only focused on the following locations: grocery & pharmacy, retail & recreation, transit stations and workplaces (the mean across these four place categories is shown in black). The gray rectangles indicate the period of the sweep of the Alpha and Delta variants.

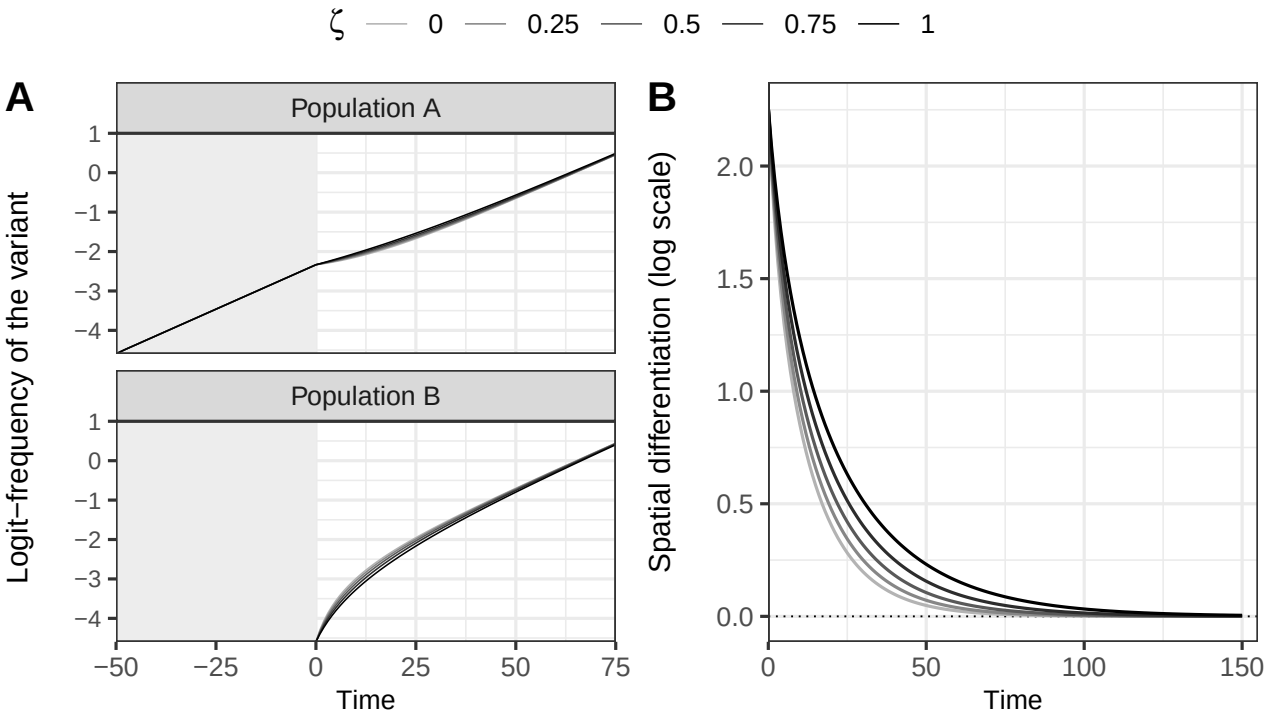

Figure S9: **The effect of reduced mobility in infected hosts on the evolutionary dynamics.** The host metapopulation is divided in two populations, A and B. The variant has a +30% transmission rate selective advantage over the wildtype. We simulate the model with identical total population densities ($N^A = N^B = 100$), and parameter values (same for both populations): $\beta_w = 0.15$, $\beta_m = 0.195$ and $\gamma_w = \gamma_m = 0.1$. The wildtype is initially introduced at very low density (10^{-7}) in both populations and the variant is only introduced in population A such that it represents 1% of the infected hosts I^A – both populations are otherwise fully susceptible. Before $t = 0$, the two populations are isolated ($\mu^A = \mu^B = 0$, gray background). At $t = 0$, the variant is introduced in population B (1% of I^B) and, from that time point onward, hosts may visit the other population ($\mu^A = \mu^B = 0.1$). In **SI Appendix §S6**, we relax the assumption that infected hosts (I) commute with the same probability that non-infected hosts (S and R), such that ζ represents the reduction in host mobility due to the disease, ranging from 0 (no impact, same as in the main text) to 1 (no mobility of infected hosts). Varying the value of ζ, we plot the temporal dynamics of (A) the variant logit-frequency and of (B) the spatial log-differentiation $\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$.

Supplementary Information (SI Appendix)

S1 A two-patch host metapopulation coupling epidemiology and evolution

We model the dynamics of a directly and horizontally transmitted infectious disease. To do so, we rely on a classical susceptible-infected-recovered (SIR) models. Hosts are either susceptible (S), infected/infectious (I) or recovered (R). For each compartment "X", we denote X its density at the current time t (we drop the time dependence in our notations for compactness and readability). For the sake of simplicity, we neglect host birth and death processes and assume that the total host population density $N = S + I + R$ remains constant over time. Let β be the *per capita* transmission rate – which captures both the host contact rate and the probability of transmission per contact with an infected host –, and γ, the *per capita* recovery rate (infection clearance). We assume that there is no pathogen virulence (i.e., disease-induced mortality).

We focus on a scenario where two strains (genotypes) of the pathogen are spreading in the host population. Without loss of generality, let us call them the wildtype (w) and the mutant strain (m), or variant. Throughout, we use the subscripts w and m to refer to each strain. We neglect the appearance of new mutations. In addition, we assume that a host can only be infected once, i.e., we assume no super-infection (including co-infections by both strains) and that, after recovery, infection provides full, lifelong immunity to both strains (full cross-immunity). The total density of infected hosts I is thus $I = I_w + I_m$. Crucially, the phenotype of the variant may differ from the wildtype in terms of transmission rate ($\beta_m = \beta_w + \Delta\beta$) and/or recovery rate ($\gamma_m = \gamma_w + \Delta\gamma$).

We also consider a two-patch metapopulation, that is, two host populations, labeled A and B, coupled by migration (we extend this model to $n \geq 2$ populations in the section **§S5**). Throughout, we use superscripts to distinguish the two populations. We denote N^A and N^B, the (constant) total densities of populations A and B, respectively, such that $N^A + N^B = N$. Now define μ^i, the probability for a host from the focal population $i \in \{A, B\}$ to visit the non-focal population. For the sake of simplicity, we assume that μ^i is independent of both the host's epidemiological status (S, I or R), and the pathogen's strain (w, m). We thus neglect the potential influence the pathogen may have on the mobility of its host – e.g., if hosts tend to limit their movements when they present symptoms or are aware of being infected – (we relax this assumption in the last section **§S6**), and we assume that migration is not itself under selection. Visits are considered to be only (instantaneous) to-and-fro host movements, such as commuting (Lipshtat et al., 2021), neglecting the impact of longer or permanent stays. Strictly speaking, we thus model the migration of the disease, rather than that of hosts (Post et al., 1983). Finally, we assume that, due to some spatial heterogeneities between the two populations (e.g., non-pharmaceutical interventions, behaviors, treatments) phenotypic traits β and γ can vary between populations, so that the phenotypic traits of the variant in population i is given by

$$\begin{aligned} \beta_m^i &= \beta_w^i + \Delta\beta^i \\ \gamma_m^i &= \gamma_w^i + \Delta\gamma^i \end{aligned} \tag{S1}$$

For example, if the efficacy of control measures implemented to reduce transmission in the focal population i is denoted c^i, then we would have $\beta_k^i = \left(1 - c^i\right)\beta_k$, with β_k, the baseline transmissibility of strain k.

S2 Epidemiology

S2.1 At the metapopulation level

Epidemiological dynamics at the metapopulation level are given by the following system of ordinary differential equations (ODEs)

$$\begin{aligned}
\frac{\mathrm{d}S}{\mathrm{d}t} &= -\sum_i \lambda^i \Big(\left(1-\mu^i\right) S^i + \mu^j S^j \Big) \\
\frac{\mathrm{d}I}{\mathrm{d}t} &= \sum_i \lambda^i \Big(\left(1-\mu^i\right) S^i + \mu^j S^j \Big) - \sum_i \overline{\gamma}^i I^i \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= \sum_i \overline{\gamma}^i I^i
\end{aligned} \tag{S2}$$

with $(i,j) \in \{A,B\}^2,\ j \neq i$; $\overline{\gamma}^i$ is the mean recovery rate in population i (averaged across pathogen genotypes)

$$\overline{\gamma}^i = \sum_k \frac{I_k^i}{I^i} \gamma_k^i, \tag{S3}$$

with $k \in \{w,m\}$, and λ^i represents the force of infection (i.e., the *per capita* infection rate) experienced in population i. Assuming transmission to be frequency-dependent (McCallum et al., 2001), we have

$$\lambda^i = \sum_k \beta_k^i \frac{\left(1-\mu^i\right) I_k^i + \mu^j I_k^j}{\left(1-\mu^i\right) N^i + \mu^j N^j}. \tag{S4}$$

The total force of infection is therefore the sum of the contributions of four types of intra or inter-community interactions (Le Treut et al., 2022). The intra-community contributions correspond to transmissions between a susceptible and an infected host who belong to the same population, which can occur either (i) in the local population or (ii) in the visited population (visitor-to-visitor infection). In contrast, the inter-community contributions correspond to transmissions between a susceptible and an infected host who do not belong to the same population, which can occur (i) when a susceptible visitor gets infected by a native or (ii) when a susceptible native gets infected by a visitor.

S2.2 At the level of the focal population

Epidemiological dynamics for the focal population $i \in \{A,B\}$ (accordingly, the non-focal population is denoted by j) are given by the following system of ODEs

$$\begin{aligned}
\frac{\mathrm{d}S^i}{\mathrm{d}t} &= -\Big(\left(1-\mu^i\right) \lambda^i + \mu^i \lambda^j \Big) S^i \\
\frac{\mathrm{d}I^i}{\mathrm{d}t} &= \Big(\left(1-\mu^i\right) \lambda^i + \mu^i \lambda^j \Big) S^i - \overline{\gamma}^i I^i \\
\frac{\mathrm{d}R^i}{\mathrm{d}t} &= \overline{\gamma}^i I^i
\end{aligned} \tag{S5}$$

with $k \in \{w, m\}$ and with λ^i and λ^j, the forces of infection experienced in the focal and non-focal population, respectively. Focusing only on the epidemiological dynamics of hosts from the focal population i who are infected by the variant, we have

$$\frac{\mathrm{d}I_m^i}{\mathrm{d}t} = \left(\left(1-\mu^i\right) \underbrace{\beta_m^i \frac{\left(1-\mu^i\right) I_m^i + \mu^j I_m^j}{\left(1-\mu^i\right) N^i + \mu^j N^j}}_{\lambda_m^i} + \mu^i \underbrace{\beta_m^j \frac{\left(1-\mu^j\right) I_m^j + \mu^i I_m^i}{\left(1-\mu^j\right) N^j + \mu^i N^i}}_{\lambda_m^j} \right) S^i - \gamma_m^i I_m^i, \tag{S6}$$

with λ_m^i and λ_m^j, the force of infection of the variant in population i and j, respectively.

S3 Evolution

Alongside epidemiological dynamics, we use a population genetics approach and track the evolution of the system over time (strain frequency and spatial differentiation).

S3.1 Frequencies of the variant

First, at the level of the focal population $i \in \{A, B\}$, define

$$q^i \;=\; \frac{I_m^i}{I^i} \;=\; \frac{I_m^i}{\sum_k I_k^i}, \tag{S7}$$

with $k \in \{w, m\}$, the frequency of the variant in the focal population i at the current time t. The temporal dynamics of the variant in the focal population i is given by the derivative of q^i with respect to time, that is,

$$\begin{aligned} \frac{\mathrm{d}q^i}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{I_m^i}{I^i}\right) \\ &= \left(\frac{\mathrm{d}I_m^i}{\mathrm{d}t} - q^i \frac{\mathrm{d}I^i}{\mathrm{d}t}\right) \Big/ I^i \end{aligned} \tag{S8}$$

Second, at the metapopulation level, let q be the overall frequency of the variant

$$q \;=\; \frac{I_m}{I} \;=\; \frac{\sum_i I_m^i}{\sum_i \sum_k I_k^i} \;=\; f q^A + (1-f)\, q^B, \tag{S9}$$

where $f = I^A/I$ is the frequency of infected hosts that belong to population A. Likewise, the temporal dynamics of q is given by

$$\begin{aligned} \frac{\mathrm{d}q}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{I_m}{I}\right) \\ &= \left(\frac{\mathrm{d}I_m}{\mathrm{d}t} - q \frac{\mathrm{d}I}{\mathrm{d}t}\right) \Big/ I \end{aligned} \tag{S10}$$

S3.2 Spatial differentiation

We now introduce the following measure of the genetic differentiation between population A and B (hereafter referred to as spatial differentiation)

$$\mathcal{Q} = \frac{q^A}{1-q^A}\frac{1-q^B}{q^B}, \tag{S11}$$

such that

$$\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right). \tag{S12}$$

When the variant is more frequent in population A than in population B, $\mathcal{Q} > 1$ and $\ln(\mathcal{Q}) > 0$. When the variant is found in equal frequency in both populations, $\mathcal{Q} = 1$ and $\ln(\mathcal{Q}) = 0$ (no differentiation). We used similar quantities in previous work to track the differentiation across pathogen life stages (Benhamou et al., 2023, 2024), rather than across space, but the principle remains exactly the same. Though the spatial differentiation would be more classically expressed on the natural scale as $q^A - q^B$, the corresponding dynamics is affected by the genetic variances. We thus focus more conveniently on $\mathcal{Q}$. Using equation (S12), the temporal dynamics of the log-differentiation $\ln(\mathcal{Q})$ is given by

$$\frac{\mathrm{d}\ln(\mathcal{Q})}{\mathrm{d}t} = \frac{\mathrm{d}\,\text{logit}\left(q^A\right)}{\mathrm{d}t} - \frac{\mathrm{d}\,\text{logit}\left(q^B\right)}{\mathrm{d}t} = \frac{1}{\mathcal{Q}}\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}t}, \tag{S13}$$

S3.3 Relation between local and global variables

Using either the pair of variables q^A and q^B ("local" variables) or the pair q and $\mathcal{Q}$ ("global" variables) are sufficient to fully describe the evolutionary dynamics of the system. The definitions of q and $\mathcal{Q}$ are based on q^A and q^B and, the other way around, q^A and q^B can be expressed as functions of q and $\mathcal{Q}$

$$\begin{aligned} q^A &= \frac{1+(\mathcal{Q}-1)(f+q) - \sqrt{\left(1+(f+q)(\mathcal{Q}-1)\right)^2 - 4fqQ(\mathcal{Q}-1)}}{2f(\mathcal{Q}-1)} \\ q^B &= \frac{-1-(\mathcal{Q}-1)(f-q) + \sqrt{\left(1+(f+q)(\mathcal{Q}-1)\right)^2 - 4fqQ(\mathcal{Q}-1)}}{2(1-f)(\mathcal{Q}-1)} \end{aligned} \tag{S14}$$

The case $\mathcal{Q} = 1$ is a removable discontinuity for which $q^A = q^B = q$. When the differentiation $\mathcal{Q}$ is not too far from 1 (no differentiation), a Taylor expansion about the case $\mathcal{Q} = 1$ yields the approximations

$$\begin{aligned} q^A &\approx q + q(1-q)(1-f)(\mathcal{Q}-1) \\ q^B &\approx q - q(1-q)f(\mathcal{Q}-1) \end{aligned} \tag{S15}$$

S4 Weak migration approximation

We now assume that the amount of host mobility is small, i.e., hosts interact much more locally rather than with other populations. In the following, we treat μ^i as $\epsilon_{\mathcal{M}} M^i$ and μ^j as $\epsilon_{\mathcal{M}} M^j$ to emphasize that the migration probabilities of the hosts are small, of order $\epsilon_{\mathcal{M}} \ll 1$. We referred to this approximation in the main text as the weak migration assumption.

S4.1 Epidemiological dynamics

A Taylor expansion of the force of infection λ^i (S4) about $\epsilon_{\mathcal{M}} = 0$ yields

$$\lambda^i = \sum_k \beta_k^i \frac{\left(1 - \frac{N^j}{N^i}\mu^j\right) I_k^i + \mu^j I_k^j}{N^i} + O\left(\epsilon_{\mathcal{M}}^2\right). \tag{S16}$$

Note that the term $O(\epsilon_{\mathcal{M}})$ depends on μ^j but not on μ^i anymore. Likewise, a Taylor expansion of the ODE system (S5) about $\epsilon_{\mathcal{M}} = 0$ yields the following epidemiological dynamics for the infected hosts from population i

$$\frac{\mathrm{d}I^i}{\mathrm{d}t} = \sum_k \left(\left[\overbrace{\underbrace{\left(1 - \mu^i - \frac{N^j}{N^i}\mu^j\right)\beta_k^i \frac{I_k^i}{N^i}}_{\text{Endogenous}}}^{\text{Intra-community}} + \overbrace{\underbrace{\mu^j \beta_k^i \frac{I_k^j}{N^i}}_{\text{Endogenous}} + \underbrace{\mu^i \beta_k^j \frac{I_k^j}{N^j}}_{\text{Exogenous}}}^{\text{Inter-community}} \right] S^i - \gamma_k^i I_k^i \right) + O\left(\epsilon_{\mathcal{M}}^2\right). \tag{S17}$$

Here, "endogenous" (resp. "exogenous") refers to infections that occur within (resp. outside) the focal population, regardless of the population to which the infector belongs. Note that all the visitor-to-visitor contribution to the force of infection (intra-community but exogenous) is now included in the term $O\left(\epsilon_{\mathcal{M}}^2\right)$.

S4.2 Evolutionary dynamics

S4.2.1 Dynamics of the variant frequency at the population level

Using equations (S8) and (S29), the dynamics of the variant frequency q^i is given after some rearrangements by

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \underbrace{q^i\left(1 - q^i\right)}_{\substack{\text{Genetic}\\ \text{variance}}} \mathcal{S}^i + \mathcal{H}^i + O\left(\epsilon_{\mathcal{M}}^2\right). \tag{S18}$$

with

$$\begin{aligned} \mathcal{S}^i &= \overbrace{\Delta\beta^i \frac{S^i}{N^i} - \Delta\gamma^i}^{\text{Local selection}} + \overbrace{\left[\mu^i\left(\Delta\beta^j \frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - \Delta\beta^i\right) + \frac{N^j}{N^i}\mu^j \Delta\beta^i\left(\frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - 1\right)\right]\frac{S^i}{N^i}}^{\text{Interaction between selection and migration}} \\ \mathcal{H}^i &= \underbrace{-\left(q^i - q^j\right)\left(\mu^i \beta_w^j + \frac{N^j}{N^i}\mu^j \beta_w^i\right)\frac{I^j/N^j}{I^i/N^i}\frac{S^i}{N^i}}_{\text{Between-pop. homogenization}} \end{aligned} \tag{S19}$$

Both $\mathcal{S}^i$ and $\mathcal{H}^i$ are first-order approximations in $\epsilon_{\mathcal{M}}$. The component $\mathcal{S}^i$ collects all the terms that depend on the phenotypic differences between the two strains and refer to the effects of local selection on the focal population as well as the interaction between selection and migration between populations. The component $\mathcal{H}^i$ is a homogenization force that affects the local dynamics of the variant as soon as the variant frequency varies between the two connected populations (i.e., $q^i \neq q^j$).

The selection coefficient of the variant is typically quantified by estimating the regression slope of its frequency changes on the logit scale (Chevin, 2011; Fisher, 1999). We thus focus instead on the dynamics of

the logit-frequency of the variant which is given by

$$\frac{\mathrm{d}\,\mathrm{logit}\left(q^i\right)}{\mathrm{d}t} = \frac{\mathrm{d}q^i/\mathrm{d}t}{q^i\left(1-q^i\right)} = \mathcal{S}^i + \frac{\mathcal{H}^i}{q^i\left(1-q^i\right)} + O\left(\epsilon_{\mathcal{M}}^2\right). \tag{S20}$$

S4.2.2 Dynamics of the variant frequency at the metapopulation level

Likewise, the temporal dynamics of the variant logit-frequency at the metapopulation level (S10) is given under the weak migration assumption by

$$\begin{aligned}\frac{\mathrm{d}\,\mathrm{logit}\left(q\right)}{\mathrm{d}t} =& \frac{fq^A\mathcal{S}^A + \left(1-f\right)q^B\mathcal{S}^B}{q} + \\ & \frac{q^A-q^B}{q\left(1-q\right)}f\left(1-f\right)\left[\left(\frac{\beta_w^A}{N^A}\left(1-\mu^A-\frac{N}{N^A}\mu^B\right)-\frac{\beta_w^B}{N^B}\mu^A\right)S^A-\gamma_w^A-\right. \\ & \left.\left(\left(\frac{\beta_w^B}{N^B}\left(1-\mu^B-\frac{N}{N^B}\mu^A\right)-\frac{\beta_w^A}{N^A}\mu^B\right)S^B-\gamma_w^B\right)\right]+O\left(\epsilon_{\mathcal{M}}^2\right),\end{aligned} \tag{S21}$$

In the absence of host mobility (i.e., $\mu^A=\mu^B=0$), the previous equation reduces to:

$$\begin{aligned}\frac{\mathrm{d}\,\mathrm{logit}\left(q\right)}{\mathrm{d}t} =& \frac{fq^A\left(\Delta\beta^A\frac{S^A}{N^A}-\Delta\gamma^A\right)+\left(1-f\right)q^B\left(\Delta\beta^B\frac{S^B}{N^B}-\Delta\gamma^B\right)}{q} + \\ & \frac{q^A-q^B}{q\left(1-q\right)}f\left(1-f\right)\left[\frac{\beta_w^A}{N^A}S^A-\gamma_w^A-\left(\frac{\beta_w^B}{N^B}S^B-\gamma_w^B\right)\right]+O\left(\epsilon_{\mathcal{M}}^2\right),\end{aligned} \tag{S22}$$

S4.2.3 Dynamics of the spatial differentiation

Plugging equation (S20) into the temporal dynamics of the log-differentiation $\ln\left(\mathcal{Q}\right)$ (S13) yields

$$\frac{\mathrm{d}\ln\left(\mathcal{Q}\right)}{\mathrm{d}t} = \left(\mathcal{S}^A-\mathcal{S}^B\right)+\frac{\mathcal{H}^{\mathcal{A}}}{q^A\left(1-q^A\right)}-\frac{\mathcal{H}^{\mathcal{B}}}{q^B\left(1-q^B\right)}+O\left(\epsilon_{\mathcal{M}}^2\right), \tag{S23}$$

where, after some rearrangements

$$\frac{\mathcal{H}^{\mathcal{A}}}{q^A\left(1-q^A\right)}-\frac{\mathcal{H}^{\mathcal{B}}}{q^B\left(1-q^B\right)} = -\left(\mathcal{Q}-1\right)\left(\frac{\beta_w^A}{N^A}\mu^B+\frac{\beta_w^B}{N^B}\mu^A\right)\left(\frac{q^B}{q^A}\frac{I^B}{I^A}S^A+\frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right).$$

We can then look at the long term spatial differentiation. Intuitively, setting the right-hand side of (S23) to 0 and solving for $\mathcal{Q}$ gives

$$\mathcal{Q}\approx 1+\frac{\mathcal{S}^A-\mathcal{S}^B}{\left(\frac{\beta_w^A}{N^A}\mu^B+\frac{\beta_w^B}{N^B}\mu^A\right)\left(\frac{q^B}{q^A}\frac{I^B}{I^A}S^A+\frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right)}. \tag{S24}$$

However, in the current model, the pathogen always goes extinct because the host population eventually reaches the herd immunity threshold, so that there are no longer enough susceptible hosts to sustain the epidemic. This assumption can, however, be easily relaxed. Our derivation is not sensitive to extending the model to account for waning immunity or other recruitment of new susceptibles (e.g., birth, immigration),

as their effects would be inherently captured through the density of susceptibles.

S5 Extension to an arbitrary number of populations

More generally, we can extend the previous model to $n \geq 2$ populations.

S5.1 Epidemiological dynamics

At the level of the focal population i, the dynamics of the density of infected hosts in model (S5) becomes

$$\frac{\mathrm{d}I^i}{\mathrm{d}t} = \left[\left(1 - \sum_{j\neq i} \mu^{i\to j}\right)\lambda^i + \sum_{j\neq i} \mu^{i\to j}\lambda^j\right] S^i - \overline{\gamma}^i I^i \tag{S25}$$

with $k \in \{w, m\}$ and where j refers to the index of the jth population and $\mu^{i\to j}$ (resp. $\mu^{j\to i}$), to the probability of migration from the focal population i (resp. from the non-focal population j) to population j (resp. i). As before, λ^i and λ^j are the forces of infection experienced in population i and j, respectively, such that

$$\lambda^i = \sum_k \beta_k^i \frac{\left(1 - \sum_{j\neq i} \mu^{i\to j}\right) I_k^i + \sum_{j\neq i} \mu^{j\to i} I_k^j}{\left(1 - \sum_{j\neq i} \mu^{i\to j}\right) N^i + \sum_{j\neq i} \mu^{j\to i} N^j}. \tag{S26}$$

Under the weak migration assumption, a Taylor expansion of the previous equation about $\epsilon_{\mathcal{M}} = 0$ (see details in **§S4**) yields

$$\lambda^i = \sum_k \beta_k^i \frac{\left(1 - \sum_{j\neq i} \frac{N^j}{N^i}\mu^{j\to i}\right) I_k^i + \sum_{j\neq i} \mu^{j\to i} I_k^j}{N^i} + O\left(\epsilon_{\mathcal{M}}^2\right),$$

and, similarly, the dynamics of the density of infected hosts (S25) is now given by

$$\frac{\mathrm{d}I^i}{\mathrm{d}t} = \sum_k \left[\underbrace{\overbrace{\left(1 - \sum_{j\neq i} \mu^{i\to j} - \sum_{j\neq i} \frac{N^j}{N^i}\mu^{j\to i}\right)\beta_k^i \frac{I_k^i}{N^i}}^{\text{Intra-community}}}_{\text{Endogenous}} + \overbrace{\underbrace{\sum_{j\neq i} \mu^{j\to i}\beta_k^i \frac{I_k^j}{N^i}}_{\text{Endogenous}} + \underbrace{\sum_{i\neq j} \mu^{i\to j}\beta_k^j \frac{I_k^j}{N^j}}_{\text{Exogenous}}}^{\text{Inter-community}}\right] S^i - \sum_k \gamma_k^i I_k^i + O\left(\epsilon_{\mathcal{M}}^2\right).$$

S5.2 Evolutionary dynamics

At the level of the focal population i, the dynamics of the variant logit-frequency $\mathrm{logit}\left(q^i\right) = \mathrm{logit}\left(I_m^i / I^i\right)$ can still be expressed as

$$\frac{\mathrm{d}\,\mathrm{logit}\left(q^i\right)}{\mathrm{d}t} = \mathcal{S}^i + \frac{\mathcal{H}^i}{q^i\left(1 - q^i\right)} + O\left(\epsilon_{\mathcal{M}}^2\right),$$

but now with

$$\begin{aligned}
\mathcal{S}^i &= \Delta\beta^i \frac{S^i}{N^i} - \Delta\gamma^i + \left[\sum_{j\neq i} \mu^{i\to j}\left(\Delta\beta^j \frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - \Delta\beta^i\right) + \Delta\beta^i \sum_{j\neq i}\frac{N^j}{N^i}\mu^{j\to i}\left(\frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - 1\right)\right]\frac{S^i}{N^i} \\
\mathcal{H}^i &= -\sum_{j\neq i}\left(q^i - q^j\right)\left(\beta_w^j \mu^{i\to j} + \frac{N^j}{N^i}\beta_w^i \mu^{j\to i}\right)\frac{I^j/N^j}{I^i/N^i}\frac{S^i}{N^i}
\end{aligned} \tag{S27}$$

We illustrate the dynamics of the logit-frequency of an emerging variant in a nine-patch host metapopulation (such as the nine regions of England) in **Fig. S2**.

We can also define the genetic differentiation between any pair of populations, say i and $l \neq i$, as

$$\mathcal{Q}^{il} = \frac{q^i}{1-q^i}\frac{1-q^l}{q^l},$$

such that $\ln\left(\mathcal{Q}^{il}\right) = \text{logit}\left(q^i\right) - \text{logit}\left(q^l\right)$. Again, under the assumption of weak migration, the dynamics of the spatial log-differentiation $\ln\left(\mathcal{Q}^{il}\right)$ is given by

$$\frac{\mathrm{d}\ln\left(\mathcal{Q}^{il}\right)}{\mathrm{d}t} = \left(\mathcal{S}^i - \mathcal{S}^l\right) + \frac{\mathcal{H}^i}{q^i\left(1-q^i\right)} - \frac{\mathcal{H}^l}{q^l\left(1-q^l\right)} + O\left(\epsilon_{\mathcal{M}}^2\right),$$

which, after some rearrangements, yields

$$\begin{aligned}
\frac{\mathrm{d}\ln\left(\mathcal{Q}^{il}\right)}{\mathrm{d}t} = \left(\mathcal{S}^i - \mathcal{S}^l\right) &- \sum_{j\neq i}\left(\mathcal{Q}^{ij} - 1\right)\left(\mu^{j\to i}\frac{\beta_w^i}{N^i} + \mu^{i\to j}\frac{\beta_w^j}{N^j}\right)\frac{q^j}{q^i}\frac{I^j}{I^i}S^i \\
&- \sum_{j\neq l}\left(\mathcal{Q}^{jl} - 1\right)\left(\mu^{j\to l}\frac{\beta_w^l}{N^l} + \mu^{l\to j}\frac{\beta_w^j}{N^j}\right)\frac{1-q^j}{1-q^l}\frac{I^j}{I^l}S^l + O\left(\epsilon_{\mathcal{M}}^2\right).
\end{aligned}$$

This expression can be difficult to use in practice as the dynamics of the spatial differentiation between two populations now depends not only on itself, but on the whole spatial differentiation structure in which populations i and l are embedded, including all the corresponding pairwise differentiation values. The weight of each corresponding pair of populations can either speed up or slow down the overall dynamics of the pairwise spatial differentiation.

S6 Extension to reduced mobility in infected hosts

In this section, we relax the assumption that infected hosts (I) commute with the same probability that non-infected hosts (S and R). To do so, we assume that infected individuals from the focal population i can commute to the non-focal population with probability $(1-\zeta)\mu^i$, where $\zeta \in [0, 1]$ represents the reduction in host mobility due to the disease. For the sake of simplicity, ζ is assumed to be the same for both populations and for both strains. Note that the case $\zeta = 0$ corresponds thus to the scenario presented in the main text, in which the infection does not affect host mobility.

S6.1 Epidemiological dynamics

In these conditions, the ODE systems (S2) and (S5) are unchanged but the force of infection in the focal population i is now given by

$$\begin{aligned}\lambda^i &= \sum_k \beta_k^i \frac{\left(1-(1-\zeta)\mu^i\right) I_k^i + (1-\zeta)\mu^j I_k^j}{N^i - \mu^i\left(S^i + (1-\zeta)I^i + R^i\right) + \mu^j\left(S^j + (1-\zeta)I^j + R^j\right)} \\ &= \sum_k \beta_k^i \frac{\left[1 - \frac{N^j}{N^i}\mu^j\left(1-\zeta\frac{I^j}{N^j}\right) + \zeta\mu^i\left(1-\frac{I^i}{N^i}\right)\right] I_k^i + (1-\zeta)\mu^j I_k^j}{N^i} + O\left(\epsilon_{\mathcal{M}}^2\right),\end{aligned} \tag{S28}$$

where the second line is a Taylor expansion of the first line under the assumption of the weak migration (see details in **§S4**). Plugging equation (S28) into the ODE system (S5), the dynamics of the density of infected hosts thus becomes

$$\begin{aligned}\frac{\mathrm{d}I^i}{\mathrm{d}t} =& \sum_k \Bigg[\overbrace{\underbrace{\left(1-\mu^i\left(1-\zeta\left(1-\frac{I^i}{N^i}\right)\right) - \frac{N^j}{N^i}\mu^j\left(1-\zeta\frac{I^j}{N^j}\right)\right)\beta_k^i\frac{I_k^i}{N^i}}_{\text{Endogenous}}}^{\text{Intra-community}} + \overbrace{\underbrace{(1-\zeta)\mu^j\beta_k^i\frac{I_k^j}{N^i}}_{\text{Endogenous}} + \underbrace{\mu^i\beta_k^j\frac{I_k^j}{N^j}}_{\text{Exogenous}}}^{\text{Inter-community}}\Bigg] S^i \\ &- \sum_k \gamma_k^i I_k^i + O\left(\epsilon_{\mathcal{M}}^2\right).\end{aligned} \tag{S29}$$

Note that, if $\zeta = 1$ – i.e., infection prevents infected hosts from commuting between populations –, the endogenous inter-community term disappears because inter-community transmissions would solely occur through the commuting of susceptible hosts (infection in the non-focal population and pathogen importation in the focal population).

S6.2 Evolutionary dynamics

At the level of the focal population i, the dynamics of the variant logit-frequency can still be expressed as

$$\frac{\mathrm{d}\,\mathrm{logit}\left(q^i\right)}{\mathrm{d}t} = \mathcal{S}^i + \frac{\mathcal{H}^i}{q^i\left(1-q^i\right)} + O\left(\epsilon_{\mathcal{M}}^2\right),$$

but now with

$$\begin{aligned}\mathcal{S}^i =& \overbrace{\Delta\beta^i\frac{S^i}{N^i} - \Delta\gamma^i}^{\text{Local selection}} + \\ &\overbrace{\left[\mu^i\left(\Delta\beta^j\frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} - \Delta\beta^i\left(1-\zeta\left(1-\frac{I^i}{N^i}\right)\right)\right) + \frac{N^j}{N^i}\mu^j\Delta\beta^i\left((1-\zeta)\frac{q^j}{q^i}\frac{I^j/N^j}{I^i/N^i} + \zeta\frac{I^j}{N^j} - 1\right)\right]\frac{S^i}{N^i}}^{\text{Interaction between selection and migration}}\end{aligned} \tag{S30}$$

$$\mathcal{H}^i = \underbrace{-\left(q^i - q^j\right)\left(\mu^i\beta_w^j + \frac{N^j}{N^i}(1-\zeta)\mu^j\beta_w^i\right)\frac{I^j/N^j}{I^i/N^i}\frac{S^i}{N^i}}_{\text{Between-pop. homogenization}}$$

When $\zeta = 0$ (no reduced mobility in infected hosts), we recover equations (S19), as presented in the main text. At the other end of the spectrum, when $\zeta = 1$, equation (S30) reduces to

$$\begin{aligned}
\mathcal{S}^i &= \Delta\beta^i \frac{S^i}{N^i} - \Delta\gamma^i + \left[\mu^i \left(\Delta\beta^j \frac{q^j}{q^i} \frac{I^j/N^j}{I^i/N^i} - \Delta\beta^i \frac{I^i}{N^i}\right) + \frac{N^j}{N^i}\mu^j \Delta\beta^i \left(\frac{I^j}{N^j} - 1\right)\right] \frac{S^i}{N^i} \\
\mathcal{H}^i &= -\left(q^i - q^j\right) \mu^i \beta_w^j \frac{I^j/N^j}{I^i/N^i} \frac{S^i}{N^i}
\end{aligned}$$

Again, the homogenization component $\mathcal{H}^i$ now only depends on μ^i because inter-community transmissions can only be achieved through the mobility of susceptible hosts. In **Fig. S9-A**, we plot some examples for the dynamics of the variant logit-frequency by varying the value of ζ, from 0 to 1. Simulated trajectories are extremely similar, in particular because the proportion of infected hosts remains really small compared to the proportion of susceptibles.

Besides, the dynamics of the spatial log-differentiation $\ln(\mathcal{Q}) = \text{logit}\left(q^A\right) - \text{logit}\left(q^B\right)$ is now given by

$$\begin{aligned}
\frac{\mathrm{d}\ln(\mathcal{Q})}{\mathrm{d}t} = \left(\mathcal{S}^A - \mathcal{S}^B\right) - (\mathcal{Q} - 1) \Bigg[&\mu^B \frac{\beta_w^A}{N^A} \left((1-\zeta)\frac{q^B}{q^A}\frac{I^B}{I^A}S^A + \frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right) + \\
&\mu^A \frac{\beta_w^B}{N^B} \left(\frac{q^B}{q^A}\frac{I^B}{I^A}S^A + (1-\zeta)\frac{1-q^A}{1-q^B}\frac{I^A}{I^B}S^B\right)\Bigg] + O\left(\epsilon_{\mathcal{M}}^2\right),
\end{aligned} \tag{S31}$$

Again, when $\zeta = 0$, we recover equations (S23). We show in **Fig. S9-B** some simulated dynamics for the spatial log-differentiation by varying the value of ζ, from 0 to 1. As expected, lower amounts of migration, due to reduced mobility rates in infected hosts ($\zeta \to 1$), slow down the homogenization process between the two populations; note however that, again, this effect is small because the proportion of infected hosts is small.

References

Benhamou, W., Blanquart, F., Choisy, M., Berngruber, T. W., Choquet, R., & Gandon, S. (2024). Evolution of virulence in emerging epidemics: From theory to experimental evolution and back. *Virus Evolution*, *10*(1), veae069. https://doi.org/10.1093/ve/veae069

Benhamou, W., Lion, S., Choquet, R., & Gandon, S. (2023). Phenotypic evolution of SARS-CoV-2: A statistical inference approach. *Evolution*, *77*(10), 2213–2223. https://doi.org/10.1093/evolut/qpad133

Chevin, L.-M. (2011). On measuring selection in experimental evolution. *Biology letters*, *7*(2), 210–213. https://doi.org/10.1098/rsbl.2010.0580

Fisher, R. A. (1999). *The genetical theory of natural selection: A complete variorum edition*. Oxford University Press.

Le Treut, G., Huber, G., Kamb, M., Kawagoe, K., McGeever, A., Miller, J., Pnini, R., Veytsman, B., & Yllanes, D. (2022). A high-resolution flux-matrix model describes the spread of diseases in a spatial network and the effect of mitigation strategies. *Scientific Reports*, *12*(1), 15946. https://doi.org/10.1038/s41598-022-19931-w

Lipshtat, A., Alimi, R., & Ben-Horin, Y. (2021). Commuting in metapopulation epidemic modeling. *Scientific reports*, *11*(1), 15198. https://doi.org/10.1038/s41598-021-94672-w

McCallum, H., Barlow, N., & Hone, J. (2001). How should pathogen transmission be modelled? *Trends in ecology & evolution*, *16*(6), 295–300. https://doi.org/10.1016/S0169-5347(01)02144-9

Post, W., DeAngelis, D., & Travis, C. (1983). Endemic disease in environments with spatially heterogeneous host populations. *Mathematical Biosciences*, *63*(2), 289–302. https://doi.org/10.1016/0025-5564(82)90044-X