

RESEARCH ARTICLE

# Evolutionary emergence of infectious diseases in heterogeneous host populations

Hélène Chabas<sup>1</sup>, Sébastien Lion<sup>1</sup>, Antoine Nicot<sup>1</sup>, Sean Meaden<sup>2</sup>, Stineke van Houte<sup>2</sup>, Sylvain Moineau<sup>3,4</sup>, Lindi M. Wahl<sup>5</sup>, Edze R. Westra<sup>2</sup>, Sylvain Gandon<sup>1\*</sup>

**1** CEFE UMR 5175, CNRS - Université de Montpellier - Université Paul-Valéry Montpellier – EPHE, Montpellier, France, **2** ESI and CEC, Biosciences, University of Exeter, Cornwall Campus, Penryn, United Kingdom, **3** Département de biochimie, microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, Québec City, Canada, **4** Félix d'Hérelle Reference Center for Bacterial Viruses, Faculté de médecine dentaire, Université Laval, Québec City, Canada, **5** Applied Mathematics, Western University, London, Ontario, Canada

\* [sylvain.gandon@cefe.cnrs.fr](mailto:sylvain.gandon@cefe.cnrs.fr)



**OPEN ACCESS**

**Citation:** Chabas H, Lion S, Nicot A, Meaden S, van Houte S, Moineau S, et al. (2018) Evolutionary emergence of infectious diseases in heterogeneous host populations. *PLoS Biol* 16(9): e2006738. <https://doi.org/10.1371/journal.pbio.2006738>

**Academic Editor:** John Drake, University of Georgia, United States of America

**Received:** May 21, 2018

**Accepted:** September 5, 2018

**Published:** September 24, 2018

**Copyright:** © 2018 Chabas et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All the data are available in the Supporting Information file, [S1 Data](#).

**Funding:** CNRS (grant number PEPS MPI grant). received by SG. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. EMBO Short Term Fellowship program (grant number 7191). received by HC. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. BBSRC (grant number BB/N017412/1). received by

## Abstract

The emergence and re-emergence of pathogens remains a major public health concern. Unfortunately, when and where pathogens will (re-)emerge is notoriously difficult to predict, as the erratic nature of those events is reinforced by the stochastic nature of pathogen evolution during the early phase of an epidemic. For instance, mutations allowing pathogens to escape host resistance may boost pathogen spread and promote emergence. Yet, the ecological factors that govern such evolutionary emergence remain elusive because of the lack of ecological realism of current theoretical frameworks and the difficulty of experimentally testing their predictions. Here, we develop a theoretical model to explore the effects of the heterogeneity of the host population on the probability of pathogen emergence, with or without pathogen evolution. We show that evolutionary emergence and the spread of escape mutations in the pathogen population is more likely to occur when the host population contains an intermediate proportion of resistant hosts. We also show that the probability of pathogen emergence rapidly declines with the diversity of resistance in the host population. Experimental tests using lytic bacteriophages infecting their bacterial hosts containing Clustered Regularly Interspaced Short Palindromic Repeat and CRISPR-associated (CRISPR-Cas) immune defenses confirm these theoretical predictions. These results suggest effective strategies for cross-species spillover and for the management of emerging infectious diseases.

## Author summary

The probability that an epidemic will break out is highly dependent on the ability of the pathogen to acquire new adaptive mutations and to induce evolutionary emergence. Forecasting pathogen emergence thus requires a good understanding of the interplay between the epidemiology and evolution taking place at the onset of an outbreak. Here, we provide a comprehensive theoretical framework to analyze the impact of host

ERW. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. NERC (grant number NE/M018350/1). received by ERW. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Wellcome Trust (grant number 109776/Z/15/Z). received by ERW. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. European Research Council (grant number ERC-STG-2016-714478-EVOIMMECH). received by ERW. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Marie Curie Actions (grant number REA grant agreement 660039). received by SvH. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. BBSRC (grant number BB/R010781/1). received by SvH. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Natural Sciences and Engineering Research Council of Canada (grant number). supports SM and LMW. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Leverhulme Trust (grant number Visiting Professorships grant). received by SG. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Canada Research Chair (Tier 1) (grant number). received by SM. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** BIM, bacteriophage-insensitive mutant; Cas, CRISPR-associated genes; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeat.

population heterogeneity on the probability of pathogen evolutionary emergence. We use this model to predict the impact of the fraction of susceptible hosts, the inoculum size of the pathogen, and the diversity of host resistance on pathogen emergence. Our experiments using lytic bacteriophages and CRISPR-resistant bacteria support our theoretical predictions and demonstrate that manipulating the diversity of resistance alleles in a host population may be an effective way to limit the emergence of new pathogens.

## Introduction

Understanding the factors that govern the ability of pathogens to invade a new host population is of paramount importance to design better surveillance systems and control policies. Mathematical epidemiology can provide key insights into these dynamics [1–4]. For instance, simple deterministic models identified critical vaccination thresholds, above which pathogens are driven extinct, which informed policy guidelines for vaccination campaigns [1]. However, chance events and rapid pathogen evolution can also play a critical role in determining the outcome of disease dynamics [2,4–6]. For example, recent experimental studies indicated that the dramatic size of the 2013–2016 Ebola epidemic can at least be partially explained by the acquisition of genetic mutations that increased transmissibility to humans [7,8].

Stochastic models of epidemiology can help to understand the emergence of evolving pathogen populations [5,9–13]. These models, however, often make the unrealistic assumption that the pathogen is spreading in a well-mixed and homogeneous host population, in which all hosts are equally susceptible. Although a handful of theoretical studies have shown that host heterogeneity could have an important impact on pathogen emergence, these models either relied on phenomenological or numerical approaches [12,13], or assumed that the hosts only differ in their number of contacts but not their susceptibility to pathogens [11]. Here, we extend this line of inquiry by (i) building a mechanistic model of pathogen emergence in a diverse host population, in which only some hosts are resistant to the pathogen, (ii) deriving analytical expressions for the probability of evolutionary emergence of the pathogen, and (iii) providing the first experimental test of theoretical predictions on pathogen evolutionary emergence using a bacteria–phage interaction. We demonstrate that realistic increases in the diversity of host resistance alleles strongly reduce the probability of evolutionary emergence of novel pathogens, hence suggesting new strategies to manage the emergence of diseases. Crucially, using bacteria with distinct Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) immunity and their lytic viruses (bacteriophages) [14–17], we experimentally explore the effect of host population heterogeneity on the emergence and evolution of pathogens. The experimental validation of our theoretical predictions with this microbial system confirms the ability of our mathematical model to capture the complexity of the interplay between the epidemiology and evolution of emerging pathogens in this model system.

## Results

### Theory: Understanding the probability of pathogen emergence

In order to predict how the composition of host populations impacts the probability of pathogen emergence, we developed a branching process model [5,9–13]. We aimed to capture host–pathogen interactions in which different groups of individuals within a host population each carry unique resistance alleles that recognize different pathogen epitopes, and pathogens can evade recognition by acquiring “escape” mutations in the corresponding epitopes. In this

model, we assume that the host population contains a fraction  $(1 - f_R)$  of individuals that are fully susceptible to the pathogen, while the remaining fraction  $f_R$  of the population is resistant and composed of a mixture of  $n$  host types in equal frequencies, each of which has a different resistance allele. The efficacy of resistance is assumed to be perfect (we relax this assumption in section S1.2 of S1 Text). Therefore, a pathogen with  $i$  escape mutations ( $i$  between 0 and  $n$ ) can infect a fraction  $(1 - f_R) + f_R i/n$  of the total host population.

We further assume that a host infected with a pathogen that does not carry escape mutations transmits at rate  $b$  and dies at rate  $d$ . Host resistance prevents infection without affecting  $b$  or  $d$ . Whereas escape mutations allow the pathogen to infect a larger fraction of the host population, they also carry a fitness cost,  $c$  which causes pathogens with  $i$  escape mutations to reproduce at rate  $b_i = b(1 - c)^i$ . The probability of acquiring an escape mutation is a function of  $n$ , the number of resistance alleles in the population, as well as  $i$ , the number of escape mutations already encoded by the pathogen. The probability that a pathogen with  $i$  escape mutations will acquire an additional one equals  $u_{i,n} = 1 - (1 - \mu)^{n-i}$ , where  $\mu$  is the pathogen mutation rate per target site (a target site is a region of the pathogen genome where a point mutation or a deletion may allow escape from recognition by host immunity). This simplifies to  $u_{i,n} \approx \mu(n - i)$  when the pathogen mutation rate is assumed to be small (note how the rate of escape mutations increases with  $(n - i)$ ). For the sake of simplicity, we assume that escape mutations cannot revert to the ancestral types. These reversions are expected to have a negligible effect on the probability of evolutionary emergence when the target site mutation rate remains small [11]. To account for the effect of spatial structure, we assume that when a pathogen is released from an infected host, it will land with probability  $\phi$  on the same type of host (i.e., a host susceptible to this pathogen) and with probability  $(1 - \phi)$  on a random host from the population, which may or not be of the same type.

The expected number of secondary infections caused by a pathogen with  $i$  escape mutations in an uninfected host population is given by its basic reproduction ratio:

$$R_{i,n} = \frac{b_i}{d} F_{i,n} \tag{1}$$

where  $F_{i,n} = (\phi + (1 - \phi)(f_R i/n + (1 - f_R)))$  is the effective fraction of hosts that can be infected by the focal pathogen. A pathogen with  $n$  escape mutations has a basic reproduction ratio equal to  $R_0(1 - c)^n$ , where  $R_0 = b/d$  refers to the basic reproduction ratio of the pathogen with 0 escape mutations in a fully susceptible host population. Note, however, that a pathogen with 0 escape mutations introduced in a diverse host population has a basic reproduction ratio equal to  $R_{0,n} \leq R_0$ .

The key question we wish to address with this model is how the composition and structure of the host population determines the ultimate fate of a pathogen (i.e., extinction versus emergence, see S1 and S2 Figs). We detailed in the Materials and methods section the calculation of the probability of emergence,  $P_{i,m}$  which is the probability that an inoculum of  $V_0$  pathogens with  $i$  escape mutations will not go extinct when introduced in a host population with  $n$  different resistance alleles. To understand the role of pathogen evolution in this process, we also derive the probability of evolutionary emergence, which quantifies the importance of escape mutations to pathogen emergence.

**Evolutionary emergence is maximized for intermediate proportions of resistant hosts.** To understand how the composition of the host population determines the probability that a pathogen emerges, we consider the simple scenario in which the population consists exclusively of sensitive hosts and one type of resistant host with the same resistance allele (i.e.,  $n = 1$ ). In this scenario, the probability of emergence of a pathogen with no escape mutation is

(see [Materials and methods](#))

$$P_{0,1} = 1 - \left( f_R + (1 - f_R) \frac{C - \sqrt{-4dA + C^2}}{2A} \right)^{V_0} \quad (2)$$

where  $A = b(1 - \mu)(1 - f_R(1 - \phi))$ ,  $B = b\mu$  and  $C = A + B(1 - 1/(R_0(1 - c))) + d$ .

In the absence of mutation, the probability of pathogen emergence is thus

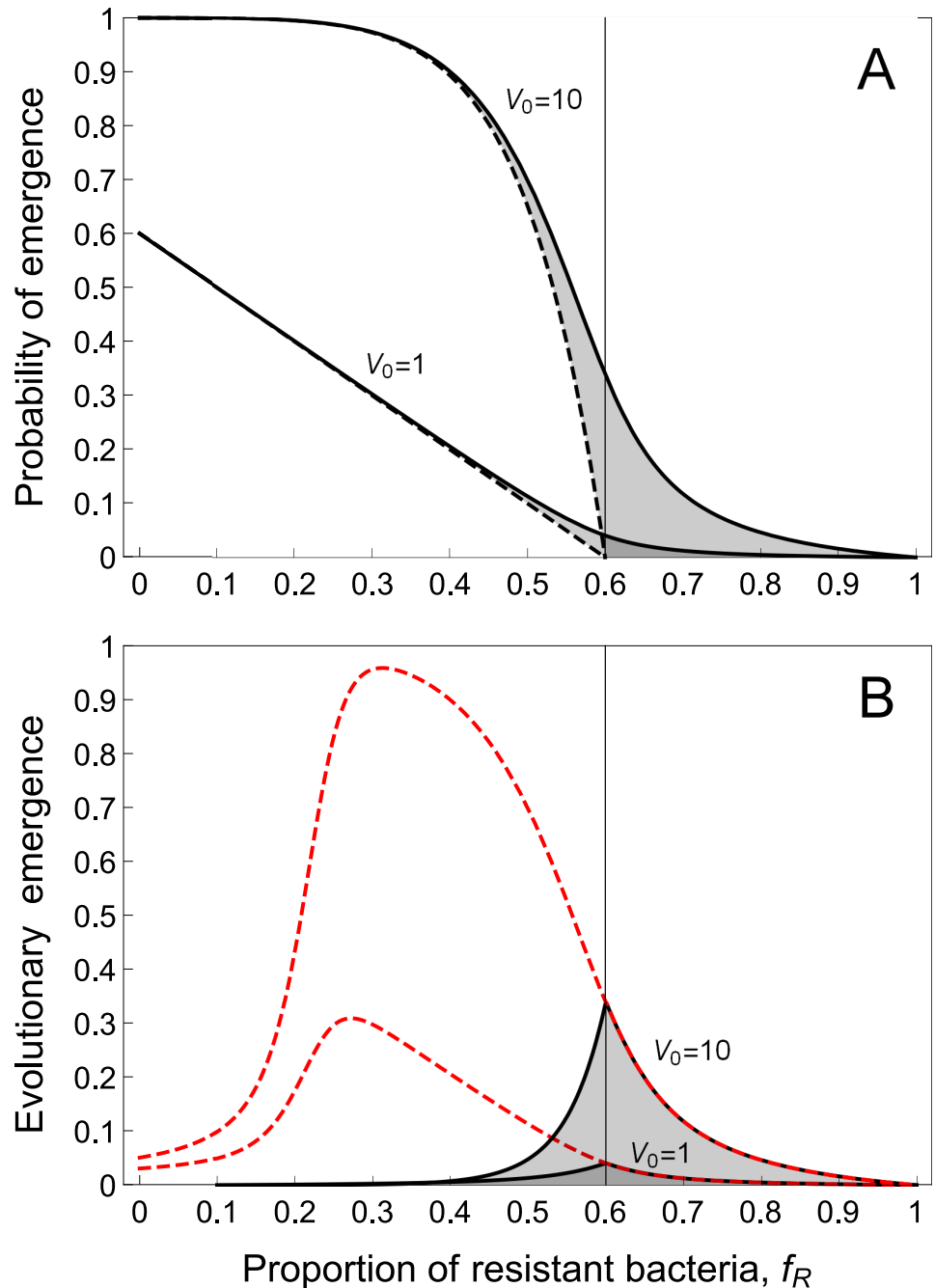
$$P_{0,1}^{\mu=0} = 1 - \left( f_R + \frac{1 - f_R}{R_0(1 - f_R(1 - \phi))} \right)^{V_0} \quad (3)$$

Varying the fraction  $f_R$  of resistant hosts shows that the probability of emergence decreases linearly with the fraction of resistant hosts in well-mixed populations (when  $\phi = 0$  and  $V_0 = 1$ , [Fig 1](#)). As expected, the probability of emergence increases with  $V_0$ , the size of the pathogen inoculum ([Fig 1](#) and [S3 Fig](#)). Population structure has also a positive effect on pathogen emergence, because it helps the pathogen to persist in the susceptible subpopulation. Interestingly, there is a threshold value  $f_T$  for the fraction of host resistance, at which the probability of emergence vanishes ([Fig 1](#)):

$$f_T = \frac{R_0 - 1}{(1 - \phi)R_0} \quad (4)$$

Next, we wanted to understand how pathogen evolution can help pathogens emerge. As expected, pathogen mutation generally increases the probability of emergence (unless there is a significant fitness cost associated with escape mutations) because those mutations allow escape from host resistance. The gray area in [Fig 1A](#) measures the amount of evolutionary emergence. Interestingly, evolutionary emergence is maximized for intermediate values of the frequency of resistance (usually when  $f_R = f_T$ , see [Fig 1B](#)). Indeed, when the frequency of resistance is low, there is no selection for escape mutations, and those mutations get rapidly lost if they are associated with fitness costs. In contrast, when the frequency of resistance is high, the chains of transmission driven by the wild-type strain are very short and there is a lower rate of appearance of escape mutations. Hence, in spite of the strong selection for escape mutations, the rate of evolutionary emergence is reduced. Intermediate resistance frequency therefore maximizes evolutionary emergence because this is where both the influx and the selection for escape mutations are high. When the host population is spatially structured, the role of pathogen evolution is smaller, simply because the probability of emergence without evolution is higher under these conditions (see [S4 Fig](#)).

So far, the model considered only the role of pathogen mutation during the very early stages of the epidemic, when stochastic effects play an important role. However, whenever we observe the emergence of a novel pathogen, these initial stages will usually have already passed, and pathogen evolution may therefore deviate from the patterns that are predicted by the above model. For example, even when escape mutations are not necessary for emergence, those mutations may increase in frequency after the onset of an epidemic. This is particularly true when the proportion of resistant hosts is large, relative to the cost of mutation. After emergence, the size of the pathogen population increases rapidly and one can start to neglect the effect of demographic stochasticity on the change in escape mutation frequencies. To understand how pathogens evolve over longer timescales during an epidemic (i.e., beyond the initial emergence), we combined our stochastic description of pathogen emergence with a deterministic model for the change in the frequency of escape mutations (see section S1.3 of [S1 Text](#)). This model predicts that the probability of observing an escape mutation evolving in a



**Fig 1. Effect of the proportion of resistant hosts ( $f_R$ ) on pathogen emergence when there is a single type of resistant host ( $n = 1$ ), and for two values of the pathogen inoculum size ( $V_0 = 1$  and  $10$ ).** (A) Probability of pathogen emergence with (full curve,  $u_{0,1} = 0.01$ ) or without mutations (dashed curve,  $u_{0,1} = 0$ ). The shaded area indicates the effect of pathogen adaptation on emergence. The threshold value  $f_R = f_T$  that prevents pathogen emergence in the absence of pathogen adaptation is indicated with a vertical dashed line. (B) Evolutionary emergence of pathogens (the dashed red curve in A) is maximized for an intermediate value of the fraction of resistant hosts. The dashed red curve shows the theoretical prediction when we account for the change in the escape mutation frequency after the pathogen emergence took place (see section S1.3 of S1 Text). Other parameter values:  $b = 2.5$ ,  $d = 1$ ,  $\phi = 0$ ,  $c = 0.2$ ,  $T = 24$ .

<https://doi.org/10.1371/journal.pbio.2006738.g001>

pathogen population is maximized for a low frequency of host resistance (red dashed curve in Fig 1B and in S4 and S5 Figs).

**Diversity of host resistance decreases pathogen emergence.** Natural host populations usually consist of multiple host genotypes, each carrying their own resistance allele. To understand how this will impact the predicted patterns of pathogen emergence, we analyzed the situation in which the resistant host population consists of  $n$  types in equal frequencies, each carrying a unique resistance allele. In the absence of pathogen mutations, emergence is governed by Eq 3 and does not depend on host diversity because all hosts are equally resistant to a pathogen with no escape mutations. Each extra escape mutation, however, allows the pathogen to infect a fraction  $1/n$  of the resistant host population, and the pathogen needs  $n$  escape mutations to exploit the whole host population. Yet, the probability  $u_{i,n}$  to acquire an escape mutation increases with  $n$  because there are more genetic loci (i.e., target sites) involved in the interaction with the host. In other words, both the number of mutations required to reach the top of the fitness landscape and the rate of acquisition of escape mutations increase with the diversity of host resistance. Because these two processes have opposite effects on the rate of pathogen adaptation, it is not immediately obvious how host diversity affects pathogen emergence.

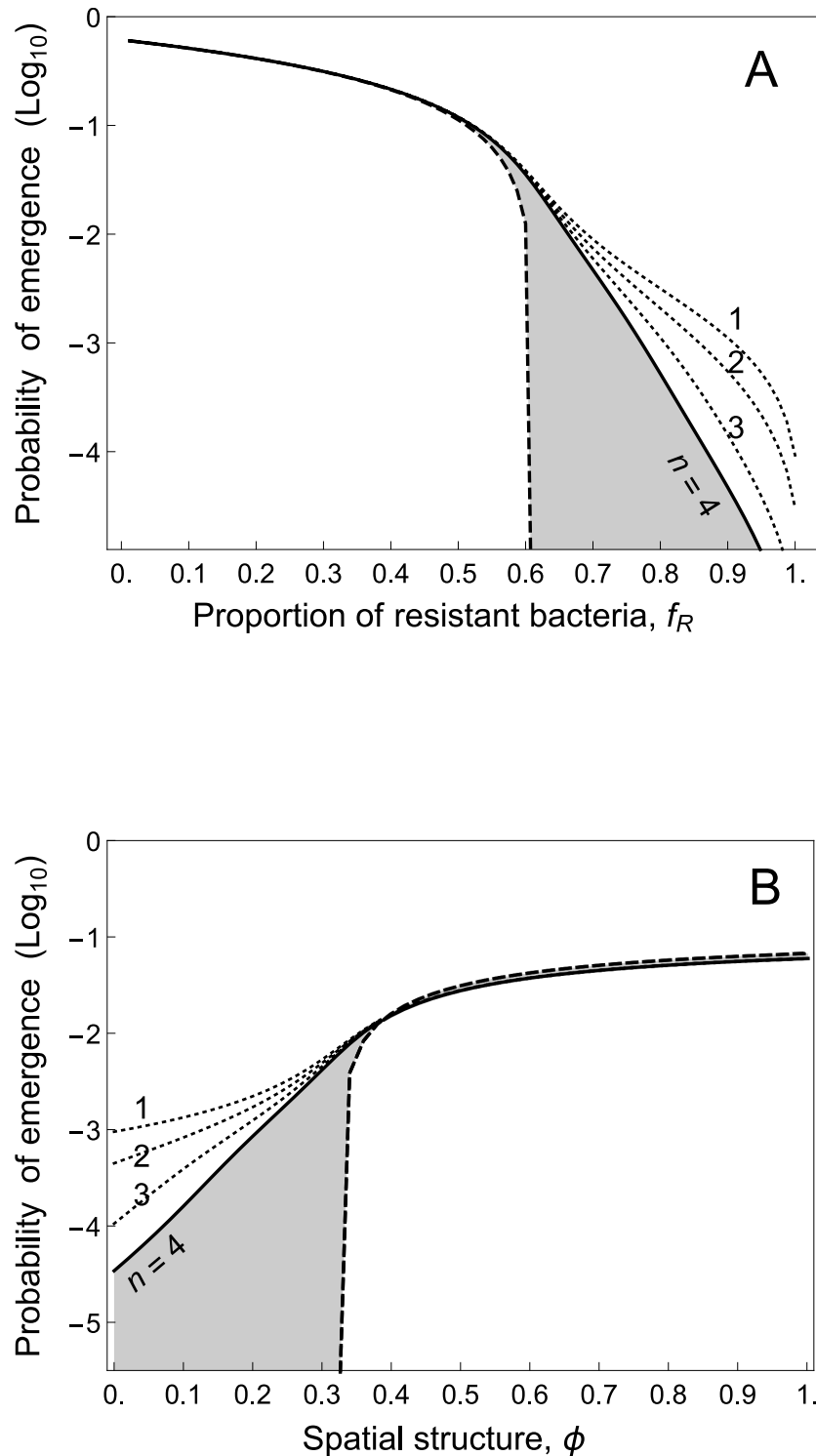
We derive numerically the probability of pathogen emergence under a broad range of scenarios (see Materials and methods). In Fig 2A, we show the joint effects of the frequency of resistance and the diversity of resistance. Increasing host diversity always decreases the probability of emergence of the pathogen. We also find a very strong interaction between host diversity and spatial structure. Because spatial structure induces a clustering of the different types of resistant hosts, it reduces the effective diversity of host resistance and promotes pathogen emergence (Fig 2B).

The above analysis relies on (i) the assumption that the pathogen life cycle can be approximated by the birth–death model and (ii) the assumption that the frequency of host resistance does not vary through time. We relaxed both these assumptions with individual-based simulations in section S1.4 of S1 Text and obtained very similar results (see S6, S7 and S8 Figs).

## Experiments on phage emergence confirm theoretical predictions

Next, we wanted to experimentally explore the validity of the above predictions. While this is challenging given the paucity of suitable empirical systems that are amenable to experimental manipulations in a timely fashion, we explored whether this could be achieved by studying the evolutionary emergence of “escape” phages against bacteria with a CRISPR–CRISPR-associated (Cas) system. This immune defense provides full protection against a phage infection by adding phage-derived sequences (known as “spacers”) in a CRISPR locus carried by the bacterial host chromosome [14]. This empirical system allowed us to overcome three important technical challenges (see details of the experimental protocols in the Materials and methods section). First, the stochastic nature of extinction requires a large number of replicate populations to measure a probability of emergence, which is possible using bacteria and phages in 96-well plates. Second, by mixing bacteria with different and unique CRISPR resistance alleles, we could manipulate the fraction of resistant hosts and the diversity in resistance alleles without affecting other traits of the host [18]. Third, unlike most other empirical systems, the mechanism of phage adaptation to CRISPR-based immunity is well known: lytic phages “escape” CRISPR resistance through mutation of their target sequence (the “protospacer”) [13,15,18,19,20].

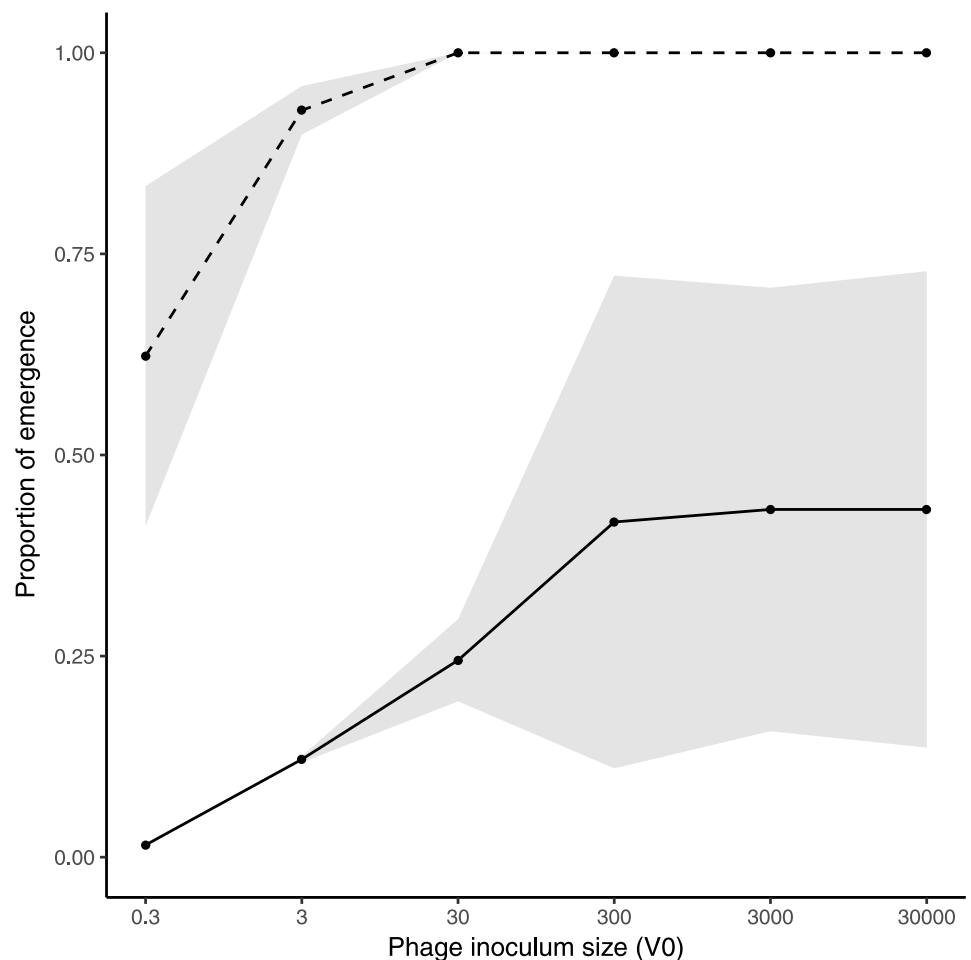
In order to validate the model using this empirical system, we used eight CRISPR-resistant clones (also referred as bacteriophage-insensitive mutants [BIMs]) of the gram-negative



**Fig 2. Effect of the fraction of resistant hosts, host resistance diversity ( $n$ ), and spatial structure ( $\phi$ ) on pathogen emergence.** (A) Probability of pathogen emergence without ( $u_{0,n} = 0$ , dashed curve) or with ( $u_{0,n} = 5 \times 10^{-3}$ , full curve) mutations (note the logarithmic scale) when  $n = 4$  and for increasing values of the fraction of resistant hosts in the absence of spatial structure (i.e.,  $\phi = 0$ ). The shaded area illustrates the fraction of pathogen emergence caused by pathogen adaptation. The dotted curves indicate the probability of emergence for lower levels of host diversity. In (B), the probability of pathogen emergence is shown for increasing values of  $\phi$  (i.e., spatial structure) when  $f_R = 0.88$ . Other parameter values:  $b = 2.5$ ,  $d = 1$ ,  $c = 0.05$ .

<https://doi.org/10.1371/journal.pbio.2006738.g002>

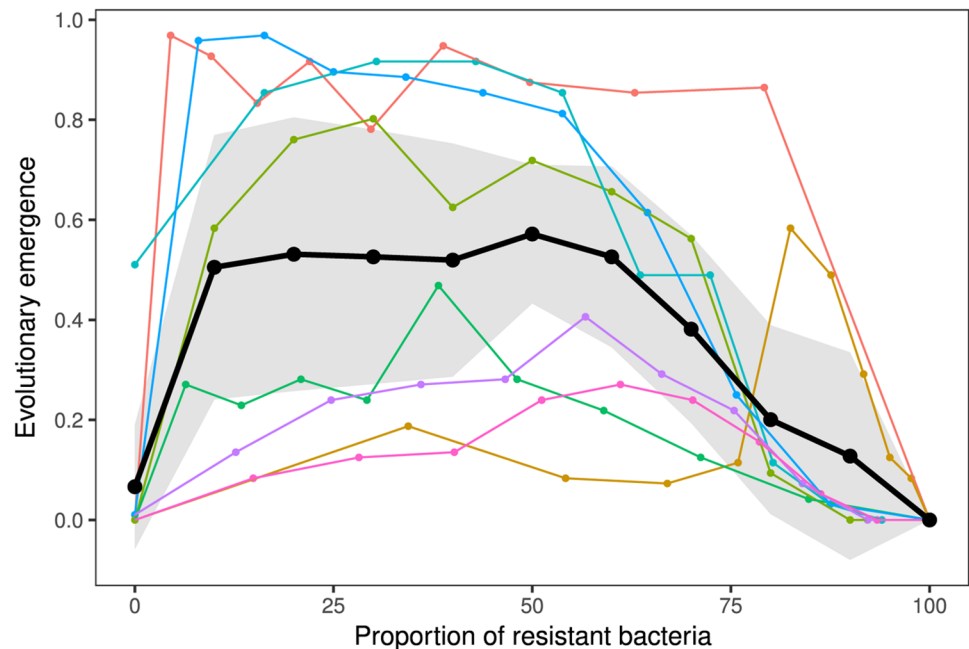
*Pseudomonas aeruginosa* strain UCBPP-PA14, each of which carried a single and distinct spacer targeting the lytic phage DMS3vir. Each of these spacers provides full resistance to infection. For each of these eight CRISPR-resistant clones, the rate at which the phage acquires escape mutations was found to be approximately equal to  $2.8 \times 10^{-7}$  mutations/locus/replication, as determined using Luria-Delbrück experiments (see section S2.1.6 in S1 Text and S9 Fig). Using one of these BIMs, we first tested the theoretical prediction that the probability of emergence increases with the size of the virus inoculum ( $V_0$ ). To this end, 96 replicate populations, each composed of an equal mix of sensitive bacteria and a CRISPR-resistant clone, were exposed with five different inoculum sizes of the phage (corresponding to a mean  $V_0$  of approximately 0.3, 3, 30, 300, and 3,000 phages). After 24 hours, we measured the fraction of phage-infected bacterial populations in which emergence had occurred. Consistent with the model predictions, we observed that the larger the phage inoculum size, the higher the probability of pathogen emergence (Fig 3, dashed line). In addition, we measured the fraction of



**Fig 3. Probability of evolutionary emergence increases with the size of phage inoculum ( $V_0$ ).** The proportion of replicate populations in which emergence (i.e., in which the amplification of the phages is detected, dashed line) or evolutionary emergence (i.e., in which the amplification of an escape phage is detected, solid line) was observed following inoculation with  $V_0$  unevolved phages in 96 independent replicate populations, each consisting of 50% sensitive bacteria and 50% BIMs ( $f_R = 0.5$ ). The different values of phage inoculum were obtained by serial dilution and correspond approximately equal to  $V_0 = 0.3, 3, 30, 300, \text{ or } 3,000$ , where  $V_0$  refers to the mean of a Poisson distributed number of viruses. Shaded areas represent 95% confidence intervals of the mean of two experiments. Data are available in S1 Data. BIM, bacteriophage-insensitive mutant.

<https://doi.org/10.1371/journal.pbio.2006738.g003>





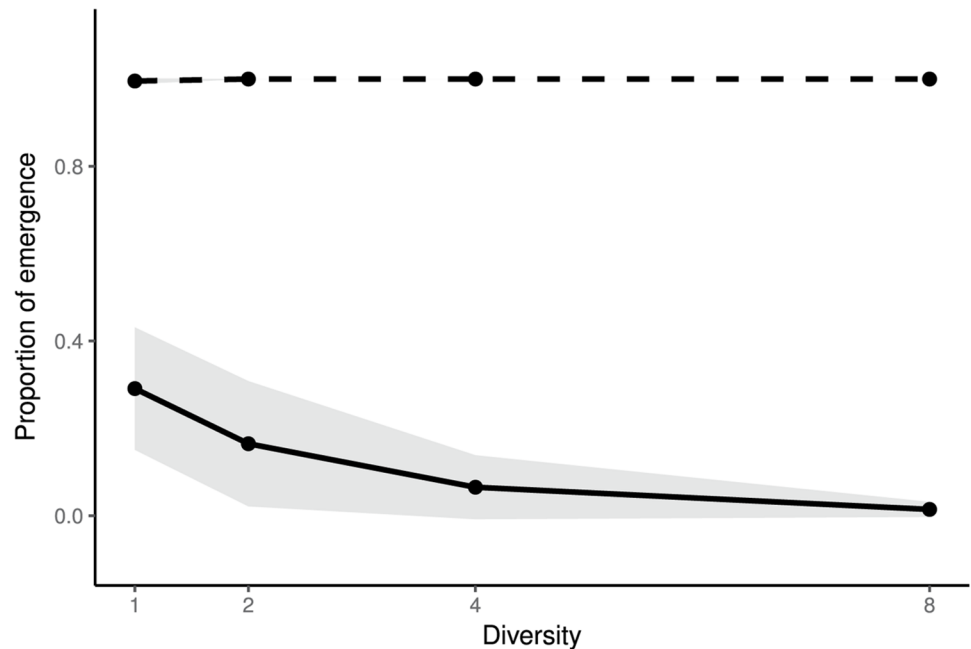
**Fig 4. Intermediate proportion of resistant hosts maximizes the probability of evolutionary emergence.** Proportion of replicate populations with evolutionary emergence (i.e., in which the amplification of an escape phage is detected) for increasing values of the proportion of resistant bacteria ( $f_R$ ). The different colors correspond to replicate experiments performed using eight different BIMs (see S2 Table and S9 Fig). For each treatment, each of the 96 replicate host populations was inoculated with an initial quantity approximately equal to  $V_0 = 300$  unevolved phages. Black lines indicate the mean across the eight BIMs; gray shaded areas represent 95% confidence intervals of the mean. Data are available in S1 Data. BIM, bacteriophage-insensitive mutant.

<https://doi.org/10.1371/journal.pbio.2006738.g004>

viral populations in which the phages had evolved to escape CRISPR resistance. Again, in accordance with the theory, we found that larger phage inocula were associated with an increased evolution of phage escape mutations (Fig 3, full line, Kendall,  $z = 3.416$ ,  $\tau = 0.784$ ,  $p < 0.001$ ). Furthermore, we obtained very similar results using a different empirical system consisting of the lytic phage 2972 and its gram-positive bacterial host *Streptococcus thermophilus* DGCC7710. In this experiment, 96 populations composed of sensitive bacteria and a CRISPR-resistant clone were infected with three different inoculum sizes of the phage. As above, we found that a larger phage inoculum led to both a higher probability of emergence and a higher probability of evolutionary emergence (S10 Fig).

Next, we tested the theoretical prediction that the probability of pathogen evolutionary emergence is highest in populations with an intermediate fraction of resistant hosts (Fig 4). For each of the eight BIMs, we generated populations composed of sensitive bacteria and a variable proportion of CRISPR-resistant bacteria, ranging from 0% to 100% in 10% increments. These populations were subsequently infected with  $V_0 = 300$  phages, and the fractions of emergence and evolutionary emergence were measured. As expected, pathogen/phage emergence dropped when the proportion of host/bacteria resistance reached a threshold level (S11 Fig). Interestingly, examination of phage evolution among emerging phage populations also confirmed that the probability of observing escape mutations is maximized for intermediate proportions of host resistance (Fig 4). Again, we obtained very consistent results with phage 2972 and *S. thermophilus* (S10 Fig).

We noticed substantial variation among CRISPR-resistant hosts in the observed frequencies of escape phage evolution (Fig 4). Variations in phage mutation rates are unlikely to explain



**Fig 5. Increasing the diversity of host resistance decreases the probability of evolutionary emergence.** Proportion of replicate populations with phage emergence (i.e., in which the amplification of the phages is detected, dashed line) or evolutionary emergence (i.e., in which the amplification of an escape phage is detected, solid line) for increasing values of the diversity of host resistance ( $n = 1, 2, 4,$  or  $8$  BIMs) when the proportion of host resistance is  $f_R = 0.5$ . For each treatment, each of the 96 replicate host populations was inoculated with an initial quantity approximately equal to  $V_0 = 300$  unevolved phages. Shaded areas represent 95% confidence intervals of the mean of two experiments. Data are available in [S1 Data](#). BIM, bacteriophage-insensitive mutant.

<https://doi.org/10.1371/journal.pbio.2006738.g005>

this variability because, as pointed out above, we failed to detect significant variations in the rate of escape mutations to the different CRISPR-resistant hosts (see [S9 Fig](#)). Variations in the fitness cost associated with these mutations could, however, explain the observed variations in the final frequency of escape mutations (see [S5 Fig](#)).

Finally, we experimentally explored the effect of resistance allele diversity on evolutionary emergence for a fixed proportion of host resistance ( $f_R = 0.5$ ). To this end, we generated bacterial populations that were composed of sensitive bacteria and an equal mix of one, two, four, or eight CRISPR-resistant clones. In this case, as expected, an inoculum size of 300 phages always led to pathogen emergence, but increasing host diversity had a strong negative effect on the ability of the phage to evolve to escape host resistance ([Fig 5](#)). We also found higher probabilities of observing multiple escape mutations in the low diversity treatment (Kendall,  $z = -4.8771$ ,  $\text{Tau} = -0.3259$ ,  $p = 1.07 \times 10^{-6}$ ), which further supports the prediction that host diversity hampers the evolution of the phage population.

## Discussion

The emergence and re-emergence of pathogens has far-reaching negative impacts on wildlife, agriculture, and public health. Unfortunately, pathogen emergence events are notoriously difficult to predict and we need good biological models to experimentally explore the interplay between epidemiology and evolution taking place at the early stages of an epidemic. Here, we used a combination of diverse theoretical and experimental analyses to examine how the composition of a host population impacts the probability of pathogen emergence and evolution. Our theory is tailored to the biology of CRISPR-phage interactions, and subsequent validation

using this experimental system demonstrates the predictive power of this theoretical framework. However, we suggest that this framework may be suitable for predicting pathogen emergence whenever hosts recognize specific pathogen epitopes and resistance can be overcome by epitope mutations. For instance, the specificity of the host–parasite interaction driven by CRISPR immunity (S12 Fig) is akin to the classical gene-for-gene system described in plant pathosystems [21]. However, host immunity may not always be perfect, which will impact both the dynamics and the evolution of the pathogen population [22–24]. To further generalize our findings, we derived the probability of pathogen emergence when immunity is imperfect (see section S1.2 in S1 Text). Note, however, that this should be considered separately from the more complex epidemiological dynamics that occur when the probability of a successful infection depends on the pathogen dose or when the pathogen causes immunosuppression, both of which can cause emergence to become dependent on the pathogen population density [25,26].

Our framework provides several insights on emergence and re-emergence in both the presence and absence of pathogen evolution. For instance, this model captures how the composition and diversity of the host population impacts the emergence of a nonevolving pathogen. In this context, a larger proportion of resistant hosts decreases pathogen emergence, but this effect is weaker in spatially structured populations in which transmission is more likely to occur between the same host types, which allows for pathogen persistence in sensitive subpopulations. This effect is akin to the effect of the spatial distribution of suitable habitats on extinction thresholds [27–30] and consistent with earlier work that shows that host composition and spatial structure impact the growth rate of bacteriophage  $\phi 6$  [31]. In the context of an evolving pathogen, our theory helps to explain the general observation that evolutionary emergence and the spread of escape mutations is maximal for an intermediate proportion of resistant hosts in the population [32]. Specifically, this is because increasing host resistance in the population has two opposite effects: (i) the influx of new mutations decreases because the ancestral pathogen cannot replicate on resistant hosts, and (ii) selection for escape mutations increases.

Second, our model predicts that diversity in host resistance alleles decreases the probability of evolutionary emergence. Even though larger host diversity increases the number of adaptive mutations for the pathogen (i.e., a larger number of targets of selection), each mutation is associated with a smaller fitness advantage (i.e., a smaller increase in the fraction of the host population that can be infected). The theory presented here therefore helps to explain previous empirical data on the impact of host CRISPR diversity on the evolution of escape phages [18]. The link between host biodiversity and infectious diseases has attracted substantial attention recently [33–43]. Several studies support the “dilution effect” hypothesis, which postulates that host diversity limits disease spread [39,40, 43]. For example, host diversity may limit the spread of a pathogen by increasing the fraction of bad-quality hosts in the population [43]. Indeed, increasing the fraction of resistant hosts (but not the diversity of resistance alleles) decreases the basic reproduction ratio of the wild-type pathogen [44,45]. In addition, host diversity per se may also limit disease spread, and several studies have shown the negative effect of host diversity on the deterministic growth rate of the pathogen under specific patterns of host–parasite specificity [35,46,47].

Notwithstanding these important insights, what sets our theoretical model apart is its ability to understand the factors that impact the initial pathogen emergence, rather than the downstream spread of a pathogen once it has already emerged. Studying this requires stochastic models, which are critical to model the probability of rare events, for example, pathogen spillover across species, including at the human–animal interface [48,49,450], the emergence of drug resistance [51,52], the evolution of vaccine resistance [53], and the reversion of live vaccines [54–58]. In all these public health issues, understanding pathogen emergence requires models accounting for the stochastic nature of epidemiological and evolutionary dynamics.

The present study focuses on the effect of the diversity of host resistance when each resistant host carries a single resistance allele (i.e., a single spacer in CRISPR). Our joint theoretical and experimental approach could be readily extended to evaluate the impact of the accumulation of multiple resistance alleles in a single host genotype rather than mixing multiple genotypes with a single resistance allele in the host population. The impact of such alternative strategies on the durability of resistance and on disease spread is particularly relevant in agriculture [59,60]. Our work provides a theoretical framework to study these different issues, and our experimental model system can be used to evaluate the ability of different control strategies to limit pathogen adaptation and emergence.

## Materials and methods

### Theory

We detail the derivation of the probability of pathogen emergence presented in the main text (the main parameters of the model are listed in S1 Table). We are interested in the ultimate fate (extinction or not) of a single pathogen with  $i$  escape mutations dropped into a very large host population with a proportion  $f_R$  of resistant hosts. This resistant population is composed of an equal frequency of  $n$  different resistance genotypes. This free infectious particle first has to infect a host to avoid extinction, and the probability of ultimate extinction of this pathogen is

$$Q_{i,n} = (1 - f_R)q_{i,n} + f_R \left( \frac{i}{n}q_{i,n} + \frac{n-i}{n} \right) \tag{5}$$

where  $q_{i,n}$  is the probability of ultimate extinction of the pathogen when it is currently infecting a host.

Next, we focus on the probability  $q_{i,n}(t)$  at time  $t$  that a pathogen with  $i$  mutations in an infected host will ultimately go extinct. In a small interval of time,  $dt$ , four different events may take place. First, the pathogen may transmit to a new host without additional escape mutations. Second, after a mutation event, the pathogen may transmit a pathogen with  $i + 1$  escape mutations to a new host. Third, the infected host (and the pathogen in the host) may die. Fourth, nothing may happen during this interval of time  $dt$ . Collecting these different terms allows us to write down recursions for the probability  $q_{i,n}(t)$ , at time  $t$ , as a function of the probability  $q_{i,n}(t + dt)$  and  $q_{i+1,n}(t + dt)$ , at time  $t + dt$ :

$$q_{i,n}(t) = \underbrace{A_{i,n} dt q_{i,n}(t + dt) q_{i,n}(t + dt)}_{\text{reproduction without mutation}} + \underbrace{B_{i,n} dt q_{i,n}(t + dt) q_{i+1,n}(t + dt)}_{\text{reproduction with mutation}} + \underbrace{d dt}_{\text{death}} + \underbrace{q_{i,n}(t + dt)(1 - A_{i,n} dt - B_{i,n} dt - d dt)}_{\text{no event}} \tag{6}$$

with:

$$A_{i,n} = b_i(1 - u_{i,n})F_{i,n}$$

$$B_{i,n} = b_i u_{i,n} F_{i+1,n}$$

$$F_{i,n} = (\phi + (1 - \phi)(f_R i/n + (1 - f_R))).$$

The above calculation is based on the assumption that the pathogen never reaches a high prevalence and that the composition of the host population remains constant (i.e.,  $F_{i,n}$  is assumed to remain constant). In other words, the probabilities  $q_{i,n}(t)$  are assumed to be invariant with time. We can thus set  $q_{i,n}(t) = q_{i,n}(t + dt)$  to obtain a recursion equation that allows us to derive  $q_{i,n}$  from  $q_{i+1,n}$ .

The first term of this recursion gives the probability of extinction,  $q_{n,n}$  that a pathogen with  $n$  escape mutations (a pathogen fully adapted to the novel host population) will go extinct. The heterogeneity of the environment has no impact on a fully adapted pathogen, and its probability of extinction is simply the extinction probability of the birth–death process:

$$q_{n,n} = 1/(R_0(1 - c)^n) \tag{7}$$

Next, to derive  $q_{n-1,n}$  from  $q_{n,n}$  we need the recursion equation for  $q_{i,n}$ . However, we have to distinguish two different scenarios. First, if  $A_{i,n} = 0$ , for example, the case of a cell infected by a fully maladapted pathogen (i.e.,  $i = 0$ ) in a well-mixed population with no susceptible hosts (i.e.,  $\phi = 0, f_R = 1$ ), we find:

$$q_{i,n} = \frac{d}{d + B_{i,n}(1 - q_{i+1,n})} \tag{8}$$

Second, in the more general scenario, in which  $A_{i,n} > 0$ , we have

$$q_{i,n} = \frac{C_{i,n} - \sqrt{-4dA_{i,n} + C_{i,n}^2}}{2A_{i,n}} \tag{9}$$

with:  $C_{i,n} = A_{i,n} + B_{i,n}(1 - q_{i+1,n}) + d$ .

Knowing  $q_{n,n}$  and the above recursion equations, we can derive  $q_{n-1,n}$  and next  $q_{n-2,n} \dots$  until we get  $q_{0,n}$ . We are particularly interested in  $q_{0,n}$  and  $Q_{0,n}$  because these quantities measure the probability of extinction of a pathogen with no escape mutations (in an infected host or as an infectious particle, respectively). Ultimately, we obtain the probability of emergence of an inoculum of  $V_0$  propagules of pathogen with no escape mutations (when  $n = 1$ , this yields Eq 2 in the main text):

$$P_{0,n} = 1 - (Q_{0,n})^{V_0} \tag{10}$$

We show in Fig 2 how the diversity of host resistance affects the probability of pathogen emergence through a reduction of evolutionary emergence. In S4 Fig, we illustrate the interaction between host diversity and spatial structure in pathogen emergence. We show that more spatial structure decreases the impact of host diversity on evolutionary emergence and increases the overall probability of pathogen emergence.

## Experiments

To study the impact of the host population composition on the probability of evolutionary emergence, we used two different microbial systems: (i) the gram-negative *P. aeruginosa* and its lytic phage DMS3vir, and (ii) the gram-positive *S. thermophilus* and its lytic phage 2972. All the resistant bacteria (i.e., BIMs) derived from the phage-sensitive wild-type strains *P. aeruginosa* UCBPP PA14 and *S. thermophilus* DGCC7710 rely on CRISPR-Cas immunity for complete resistance against the corresponding phage [14,61].

For all treatments, we performed 96 replicate infections of the corresponding host populations. We manipulated the composition of the host populations by mixing overnight cultures of sensitive bacteria and BIMs in the proportions indicated in the text, figures, and figure legends. Each replicate population was inoculated 1:100 into fresh growth media and infected with a quantity  $V_0$  of phages (the inoculum size), as indicated in the text, figures, and figure legends. After 23 hours, we monitored within each population (i) the occurrence of phage epidemics (i.e., an emergence) and (ii) the presence of escape mutants (i.e., an evolutionary emergence). A detailed description of these experiments is provided in section S2 of S1 Text.

## Supporting information

**S1 Text. Supplementary information.** In this supplementary file, we provide (1) an explicit derivation of the probability of evolutionary emergence and additional theoretical results, and (2) more details on the experimental part of the project.  
(PDF)

**S1 Fig. Summary figure.** In this study, we explore the effects of three main components of the composition of the host population on the evolutionary emergence of pathogens.  
(TIF)

**S2 Fig. The three alternative outcomes following the inoculation of the host population by a pathogen.** The figure (A) is a schematic representation of the host population before and after the emergence. The dots represent uninfected hosts (pink), hosts infected with the wild-type pathogen (blue), and hosts infected with the escape mutant (red) that can infect resistant hosts (indicated with a black contour line). The figure (B) is a schematic representation of the continuous time branching process that accounts for pathogen transmission to a new host (a vertical line), host recovery/death (a cross), and pathogen mutation (a vertical line connecting an infection by a wild type to an infection by an escape mutant).  
(TIF)

**S3 Fig. Effect of the pathogen inoculum size ( $V_0$ ) on pathogen emergence when there is a single type of resistant host ( $n = 1$ ) for different values of the fraction of resistant hosts ( $f_R$ ).** In this figure, we plot  $P_{0,1}$  against  $V_0$  (dashed lines) and  $\sum_{V_0=0}^{\infty} \frac{e^{-E[V_0]} E[V_0]^{V_0}}{V_0!} P_{0,1}$  against  $E[V_0]$  (full line) under the assumption that the number of phages inoculated follows a Poisson distribution with mean  $E[V_0]$ . Other parameter values:  $b = 2.5$ ,  $d = 1$ ,  $u_{0,1} = 0.01$ ,  $c = 0.01$ .  
(TIF)

**S4 Fig. Effect of the proportion of resistant hosts ( $f_R$ ) on pathogen emergence when there is a single type of resistant host ( $n = 1$ ) and for two values of the phage inoculum size ( $V_0 = 1$  and  $10$ ) when  $\phi = 0.3$ .** (A) Probability of pathogen emergence without ( $u_{0,1} = 0$ , dashed curve) or with ( $u_{0,1} = 0.01$ , full curve) mutations. The shaded area refers to the fraction of pathogen emergence caused by pathogen adaptation. The threshold value  $f_T$  of the fraction of resistant hosts preventing pathogen emergence in the absence of pathogen adaptation is indicated with a vertical dashed line. (B) Evolutionary emergence of pathogens (the shaded area in A) is maximized for an intermediate value of the fraction of resistant hosts. The dashed red curve represents the theoretical prediction when we track the change in the frequency of escape mutations after emergence (see section S1.3 in S1 Text). Other parameter values (same as in Fig 1, except for  $\phi$ ):  $b = 2.5$ ,  $d = 1$ ,  $\rho = 1$ ,  $c = 0.2$ ,  $T = 24$ .  
(TIF)

**S5 Fig. Effect of the cost of escape mutations ( $c$ ) on evolutionary emergence.** The shaded area refers to the fraction of pathogen emergence caused by pathogen adaptation (when  $c = 0.01$ ) and is maximized for an intermediate value of the fraction of resistant hosts. The dashed curves represent the theoretical prediction when we track the change of the frequency of escape mutations after pathogen emergence (see section 1.3 in S1 Text) for three different values of the cost of escape mutations:  $c = 0.01$  (red),  $c = 0.2$  (orange), and  $c = 0.4$  (green). Other parameter values:  $b = 2.5$ ,  $d = 1$ ,  $u_{0,n} = 10^{-3}$ ,  $\phi = 0$ ,  $\rho = 1$ ,  $T = 20$ .  
(TIF)

**S6 Fig. Results from individual-based simulations (burst-death life cycle with fixed proportions of resistant hosts): Probability of emergence versus the proportion of resistant bacteria when a single propagule ( $V_0 = 1$ ) is introduced in a well-mixed population ( $\phi = 0$ ) with a single type of resistance ( $n = 1$ ).** Results are shown in the absence ( $u_{0,1} = 0$ , crosses) or presence ( $u_{0,1} = 0.1$ , circles) of mutation for burst sizes of 2, 10, and 50. Other parameter values:  $\hat{b} = 2.5$ ,  $d = 1$ ,  $c = 0$ . Results are shown for 10,000 simulation runs. The solid black line in panel A illustrates results for the birth–death process, demonstrating that the two processes are similar but not precisely equivalent when  $B = 2$ .

(TIF)

**S7 Fig. Results from individual-based simulations (burst-death life cycle with fixed proportions of resistant hosts): Probability of emergence versus the proportion of resistant bacteria when a single propagule ( $V_0 = 1$ ) is introduced in a well-mixed population ( $\phi = 0$ ).** Results are shown on a log scale in the absence of mutation ( $u_{0,1} = 0$ , crosses, dotted line) or in the presence of mutation ( $u_{0,1} = 0.005$ ), with one to four types of resistance ( $n = 1, 2, 3$ , dashed lines;  $n = 4$ , circles, solid line) for burst sizes  $B = 2, 10$ , and  $20$ . Other parameter values:  $\hat{b} = 2.5$ ,  $d = 1$ ,  $c = 0.05$ . Results are shown for 100,000 simulation runs.

(TIF)

**S8 Fig. Results from individual-based simulations (birth–death life cycle with evolving proportions of resistant hosts).** Upper panels (A and B): The probability of emergence versus the initial proportion of resistant bacteria in a well-mixed population ( $\phi = 0$ ) with a single type of resistance ( $n = 1$ ). Results are shown in the absence ( $u_{0,1} = 0$ , crosses) or presence ( $u_{0,1} = 0.1$ , circles) of mutation, for initial inoculum sizes  $V_0 = 1$  (A) or  $V_0 = 10$  (B). For comparison, the predictions of the branching process approximation are shown (red lines). Lower panels (C and D): the fraction of resistant bacteria among uninfected bacteria at the end of the simulation for  $V_0 = 1$  (C) or  $V_0 = 10$  (D). Note the deviation from the  $y = x$  (dotted) line showing that the frequency of resistance is not constant. Other parameter values:  $b = 2.5$ ,  $d = 1$ ,  $c = 0.2$ . Results are shown for 100 simulation runs. Each simulation was stopped either when the parasite went extinct or when the maximal simulation time was reached ( $t = 200$ ), depending on which event happened first.

(TIF)

**S9 Fig. Proportion ( $1 - P_0$ ) of the 96 replicates in which phage escape mutants were generated in the absence of selection (Luria-Delbrück experiment) after plating the phage against 40 different BIM of *Pseudomonas aeruginosa*.** In Figs 4 and 5, we used eight different BIMs (indicated in the gray rectangle, see S2 Table), on which our estimation of the phage mutation rate was based (see section S2.1.6 of S1 Text). Data are available in S1 Data. BIM, bacteriophage-insensitive mutant.

(TIF)

**S10 Fig. Increased phage inoculum size ( $V_0$ ) and intermediate proportions of resistant hosts ( $f_R$ ) maximize the probability of evolutionary emergence in the *Streptococcus thermophilus* DGCC 7710–phage 2972 system.** The proportion of populations (96 total) infected with either  $V_0 \approx 2$  (red),  $V_0 \approx 20$  (green), or  $V_0 \approx 200$  (blue) phages, in which emergence (A, solid lines) or evolutionary emergence (B, dashed lines) was observed. Shaded areas represent estimated 95% confidence intervals. Data are available in S1 Data.

(TIF)

**S11 Fig. High proportion of resistant hosts minimizes the probability of emergence in the *Pseudomonas aeruginosa* UCBPP-PA14–phage DMS3vir system.** Probability of emergence

(i.e., when the amplification of the phage is detected) for increasing values of the proportion of a single resistant bacterium ( $f_R$ ). In contrast, Fig 4 presents the probability of evolutionary emergence (i.e., when the amplification of an escape phage is detected). The different colors correspond to replicate experiments performed using eight different BIMs. For each treatment, each of the 96 replicate host populations were inoculated with an initial quantity of  $V_0 \approx 300$  unevolved phages. Black lines indicate the mean across the eight BIMs; gray shaded areas represent 95% confidence intervals of the mean. Data are available in S1 Data. BIM, bacteriophage-insensitive mutant.

(TIF)

**S12 Fig. Fitness landscape of the phage in a bacterial population with  $n = 4$  different types of resistance.** The figure (A) presents the network of 16 different strains with an increasing number of escape mutations. Arrows indicate possible mutation steps. The figure (B) presents the infectivity of five of these 16 different strains against each of the five different bacterial types (where a cross indicates bacteria resistance).

(TIF)

**S1 Table. Definitions of the main parameters of the model.**

(DOCX)

**S2 Table. Spacer sequences of the eight BIMs of *Pseudomonas aeruginosa* used in our study (see S9 Fig).** BIM, bacteriophage-insensitive mutant.

(DOCX)

**S1 Data. Raw data of the experiments presented in Figs 3, 4 and 5, S9, S10, and S11 Figs.**

The file contains five tabs and the data in each tab have been used to generate one (or two) figure(s). (1) **Tab Fig 3.** The table is composed of four columns: “Phages” indicates the number of phages initially inoculated, “Fraction” indicates the proportion of replicates (between 0 and 1) in which an epidemic occurred, “Type” indicates whether this epidemic is an emergence or an evolutionary emergence, and “Replicate” indicates in which replicated experiments the data were obtained. (2) **Tab Fig 4 and S11 Fig.** The table is composed of four columns: “Fraction sensitive” indicates the proportion of sensitive bacteria added at the beginning of the experiment, “Epidemics (on 96 replicates)” indicates the number of replicates (between 0 and 96) in which an epidemic occurred, “Type of epidemics” indicates whether this is an emergence or an evolutionary emergence, and “BIM” indicates the name of the BIM. When “Type of epidemic” is “Evolutionary emergence,” the data were used to draw Fig 4, and when it is “Emergence,” the data were used to draw S11 Fig. (3) **Tab Fig 5.** The table contains four columns: “Strain” contains the name of the strain tested, “Diversity” contains the level of diversity tested, “Epidemics on strain” contains the number of replicates (between 0 and 96) on which an epidemic occurred, and “Rep” indicates in which replicated experiments the data were obtained. (4) **Tab S9 Fig.** The column named “BIM” indicates the name of the tested strain, and the column named “Mutants” indicates the number of replicates (between 0 and 96) in which a mutant was detected. The eight BIMs used in the following steps of this study are the BIMs 25, 38, 19, 39, 7, 9, 13, and 16. These eight bacterial strains are referred to as BIMs 1–8, respectively, in S1 Text. (5) **Tab S10 Fig.** The table contains four columns. The column called “Fraction Sensitive” indicates the fraction of sensitive bacteria at the beginning of the experiment, the column called “Epidemics” indicates the number of replicates (between 0 and 96) in which an epidemic occurred, the column called “Phages per replicate” indicates the number of phages inoculated at the beginning of the experiment, and the column “Type of epidemics” indicates whether the observed epidemics correspond to an emergence or an evolutionary emergence.



BIM, bacteriophage-insensitive mutant.  
(XLSX)

## Acknowledgments

The authors thank Jack Common for providing the different *P. aeruginosa* BIMs and Denise Tremblay for many useful pieces of advice to handle *S. thermophilus* and its lytic phage.

## Author Contributions

**Conceptualization:** Edze R. Westra, Sylvain Gandon.

**Data curation:** H  l  ne Chabas, Antoine Nicot, Sean Meaden, Edze R. Westra, Sylvain Gandon.

**Formal analysis:** S  bastien Lion, Lindi M. Wahl, Sylvain Gandon.

**Funding acquisition:** H  l  ne Chabas, Edze R. Westra, Sylvain Gandon.

**Investigation:** H  l  ne Chabas, S  bastien Lion, Sylvain Gandon.

**Methodology:** H  l  ne Chabas, Antoine Nicot, Sean Meaden, Stineke van Houte, Sylvain Moineau, Edze R. Westra, Sylvain Gandon.

**Project administration:** Sylvain Gandon.

**Software:** S  bastien Lion, Lindi M. Wahl, Sylvain Gandon.

**Supervision:** Edze R. Westra, Sylvain Gandon.

**Visualization:** H  l  ne Chabas, S  bastien Lion, Sean Meaden, Lindi M. Wahl, Sylvain Gandon.

**Writing – original draft:** Sylvain Gandon.

**Writing – review & editing:** H  l  ne Chabas, S  bastien Lion, Stineke van Houte, Sylvain Moineau, Lindi M. Wahl, Edze R. Westra, Sylvain Gandon.

## References

1. Anderson RM, May RM. Infectious diseases of humans: dynamics and control. Wiley Online Library; 1992.
2. Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious disease dynamics. Princeton University Press; 2013.
3. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton: Princeton University Press; 2008.
4. Heesterbeek H, Anderson RM, Andreasen V, Bansal S, De Angelis D, Dye C, et al. Modeling infectious disease dynamics in the complex landscape of global health. *Science*. 2015; 347: aaa4339. <https://doi.org/10.1126/science.aaa4339> PMID: 25766240
5. Antia R, Regoes RR, Koella JC, Bergstrom CT. The role of evolution in the emergence of infectious diseases. *Nature*. 2003; 426: 658–661. <https://doi.org/10.1038/nature02104> PMID: 14668863
6. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature*. 2005; 438: 355–359. <https://doi.org/10.1038/nature04153> PMID: 16292310
7. Urbanowicz RA, McClure CP, Sakuntabhai A, Sall AA, Kobinger G, M  ller MA, et al. Human adaptation of ebola virus during the West African outbreak. *Cell*. 2016; 167: 1079–1087.e5. <https://doi.org/10.1016/j.cell.2016.10.013> PMID: 27814505
8. Diehl WE, Lin AE, Grubaugh ND, Carvalho LM, Kim K, Kyaw PP, et al. Ebola virus glycoprotein with increased infectivity dominated the 2013–2016 epidemic. *Cell*. 2016; 167: 1088–1098.e6. <https://doi.org/10.1016/j.cell.2016.10.014> PMID: 27814506

9. Woolhouse MEJ, Haydon DT, Antia R. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol Evol (Amst)*. 2005; 20: 238–244. <https://doi.org/10.1016/j.tree.2005.02.009> PMID: [16701375](https://pubmed.ncbi.nlm.nih.gov/16701375/)
10. André J-B, Day T. The effect of disease life history on the evolutionary emergence of novel pathogens. *Proc Biol Sci*. 2005; 272: 1949–1956. <https://doi.org/10.1098/rspb.2005.3170> PMID: [16191602](https://pubmed.ncbi.nlm.nih.gov/16191602/)
11. Alexander HK, Day T. Risk factors for the evolutionary emergence of pathogens. *J R Soc Interface*. 2010; 7: 1455–1474. <https://doi.org/10.1098/rsif.2010.0123> PMID: [20410190](https://pubmed.ncbi.nlm.nih.gov/20410190/)
12. Yates A, Antia R, Regoes RR. How do pathogen evolution and host heterogeneity interact in disease emergence? *Proc Biol Sci*. 2006; 273: 3075–3083. <https://doi.org/10.1098/rspb.2006.3681> PMID: [17015347](https://pubmed.ncbi.nlm.nih.gov/17015347/)
13. Gandon S, Hochberg ME, Holt RD, Day T. What limits the evolutionary emergence of pathogens? *Philos Trans R Soc Lond, B, Biol Sci*. 2013; 368: 20120086. <https://doi.org/10.1098/rstb.2012.0086> PMID: [23209168](https://pubmed.ncbi.nlm.nih.gov/23209168/)
14. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007; 315: 1709–1712. <https://doi.org/10.1126/science.1138140> PMID: [17379808](https://pubmed.ncbi.nlm.nih.gov/17379808/)
15. Garneau JE, Dupuis M-È, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010; 468: 67–71. <https://doi.org/10.1038/nature09523> PMID: [21048762](https://pubmed.ncbi.nlm.nih.gov/21048762/)
16. Deveau H, Garneau JE, Moineau S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu Rev Microbiol*. 2010; 64: 475–493. <https://doi.org/10.1146/annurev.micro.112408.134123> PMID: [20528693](https://pubmed.ncbi.nlm.nih.gov/20528693/)
17. Westra ER, Swarts DC, Staals RHJ, Jore MM, Brouns SJJ, van der Oost J. The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity. *Annu Rev Genet*. 2012; 46: 311–339. <https://doi.org/10.1146/annurev-genet-110711-155447> PMID: [23145983](https://pubmed.ncbi.nlm.nih.gov/23145983/)
18. van Houte S, Ekroth AKE, Broniewski JM, Chabas H, Ashby B, Bondy-Denomy J, et al. The diversity-generating benefits of a prokaryotic adaptive immune system. *Nature*. 2016; 532: 385–388. <https://doi.org/10.1038/nature17436> PMID: [27074511](https://pubmed.ncbi.nlm.nih.gov/27074511/)
19. Deveau H, Barrangou R, Garneau JE, Labonté J, Fremaux C, Boyaval P, et al. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol*. 2008; 190: 1390–1400. <https://doi.org/10.1128/JB.01412-07> PMID: [18065545](https://pubmed.ncbi.nlm.nih.gov/18065545/)
20. Semenova E, Jore MM, Datsenko KA, Semenova A, Westra ER, Wanner B, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci USA*. 2011; 108: 10098–10103. <https://doi.org/10.1073/pnas.1104144108> PMID: [21646539](https://pubmed.ncbi.nlm.nih.gov/21646539/)
21. Flor HH. Current Status of the Gene-For-Gene Concept. *Annual Review of Phytopathology*. 1971; 9: 275–296. <https://doi.org/10.1146/annurev.py.09.090171.001423>
22. Gandon S, Michalakis Y. Evolution of parasite virulence against qualitative or quantitative host resistance. *Proceedings of the Royal Society of London B: Biological Sciences*. 2000; 267: 985–990. <https://doi.org/10.1098/rspb.2000.1100> PMID: [10874747](https://pubmed.ncbi.nlm.nih.gov/10874747/)
23. Gandon S, Mackinnon MJ, Nee S, Read AF. Imperfect vaccines and the evolution of pathogen virulence. *Nature*. 2001; 414: 751–756. <https://doi.org/10.1038/414751a> PMID: [11742400](https://pubmed.ncbi.nlm.nih.gov/11742400/)
24. Gandon S, Day T. The evolutionary epidemiology of vaccination. *J R Soc Interface*. 2007; 4: 803–817. <https://doi.org/10.1098/rsif.2006.0207> PMID: [17264055](https://pubmed.ncbi.nlm.nih.gov/17264055/)
25. Regoes RR, Ebert D, Bonhoeffer S. Dose-dependent infection rates of parasites produce the Allee effect in epidemiology. *Proceedings of the Royal Society of London B: Biological Sciences*. 2002; 269: 271–279. <https://doi.org/10.1098/rspb.2001.1816> PMID: [11839196](https://pubmed.ncbi.nlm.nih.gov/11839196/)
26. Landsberger M, Gandon S, Meaden S, Chabas H, Buckling A, Westra ER, et al. Anti-CRISPR phages cooperate to overcome CRISPR-Cas immunity. *bioRxiv*. 2018; 279026. <https://doi.org/10.1101/279026>
27. Bascompte J, Sole RV. Habitat fragmentation and extinction thresholds in spatially explicit models. *Journal of Animal Ecology*. 1996; 65: 465–473. <https://doi.org/10.2307/5781>
28. Hill M f., Caswell H. Habitat fragmentation and extinction thresholds on fractal landscapes. *Ecology Letters*. 1999; 2: 121–127. <https://doi.org/10.1046/j.1461-0248.1999.22061.x>
29. Fahrig L. Effect of habitat fragmentation on the extinction threshold: a synthesis. *Ecological Applications*. 2002; 12: 346–353. [https://doi.org/10.1890/1051-0761\(2002\)012\[0346:EOHFOT\]2.0.CO;2](https://doi.org/10.1890/1051-0761(2002)012[0346:EOHFOT]2.0.CO;2)
30. Suzuki SU, Sasaki A. How does the resistance threshold in spatially explicit epidemic dynamics depend on the basic reproductive ratio and spatial correlation of crop genotypes? *J Theor Biol*. 2011; 276: 117–125. <https://doi.org/10.1016/j.jtbi.2011.02.002> PMID: [21300071](https://pubmed.ncbi.nlm.nih.gov/21300071/)

31. Dennehy JJ, Friedenber NA, Yang YW, Turner PE. Virus population extinction via ecological traps. *Ecol Lett.* 2007; 10: 230–240. <https://doi.org/10.1111/j.1461-0248.2006.01013.x> PMID: 17305806
32. Benmayor R, Hodgson DJ, Perron GG, Buckling A. Host mixing and disease emergence. *Current Biology.* 2009; 19: 764–767. <https://doi.org/10.1016/j.cub.2009.03.023> PMID: 19375316
33. Keesing F, Holt RD, Ostfeld RS. Effects of species diversity on disease risk. *Ecol Lett.* 2006; 9: 485–498. <https://doi.org/10.1111/j.1461-0248.2006.00885.x> PMID: 16623733
34. Keesing F, Belden LK, Daszak P, Dobson A, Harvell CD, Holt RD, et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature.* 2010; 468: 647–652. <https://doi.org/10.1038/nature09575> PMID: 21124449
35. Lively CM. The effect of host genetic diversity on disease spread. *Am Nat.* 2010; 175: E149–152. <https://doi.org/10.1086/652430> PMID: 20388005
36. King KC, Lively CM. Does genetic diversity limit disease spread in natural host populations? *Heredity (Edinb).* 2012; 109: 199–203. <https://doi.org/10.1038/hdy.2012.33> PMID: 22713998
37. Lively CM, de Roode JC, Duffy MA, Graham AL, Koskella B. Interesting open questions in disease ecology and evolution. *The American Naturalist.* 2014; 184: S1–S8. <https://doi.org/10.1086/677032> PMID: 25061674
38. Wood CL, Lafferty KD, DeLeo G, Young HS, Hudson PJ, Kuris AM. Does biodiversity protect humans against infectious disease? *Ecology.* 2014; 95: 817–832. <https://doi.org/10.1890/13-1041.1> PMID: 24933803
39. Civitello DJ, Cohen J, Fatima H, Halstead NT, Liriano J, McMahon TA, et al. Biodiversity inhibits parasites: Broad evidence for the dilution effect. *Proc Natl Acad Sci U S A.* 2015; 112: 8667–8671. <https://doi.org/10.1073/pnas.1506279112> PMID: 26069208
40. Keesing F, Ostfeld RS. Is biodiversity good for your health? *Science.* 2015; 349: 235–236. <https://doi.org/10.1126/science.aac7892> PMID: 26185230
41. Ostfeld RS, Keesing F. Effects of host diversity on infectious disease. *Annual Review of Ecology, Evolution, and Systematics.* 2012; 43: 157–182. <https://doi.org/10.1146/annurev-ecolsys-102710-145022>
42. Westra ER, S nderhauf D, Landsberger M, Buckling A. Mechanisms and consequences of diversity-generating immune strategies. *Nat Rev Immunol.* 2017; 17: 719–728. <https://doi.org/10.1038/nri.2017.78> PMID: 28787398
43. Ostfeld RS, Keesing F. Is biodiversity bad for your health? *Ecosphere.* 2017; 8: e01676. <https://doi.org/10.1002/ecs2.1676>
44. Dobson A. Population dynamics of pathogens with multiple host species. *Am Nat.* 2004; 164 Suppl 5: S64–78. <https://doi.org/10.1086/424681> PMID: 15540143
45. Gandon S. Evolution of multihost parasites. *Evolution.* 2004; 58: 455–469. PMID: 15119430
46. Sasaki A. Host-parasite coevolution in a multilocus gene-for-gene system. *Proc Biol Sci.* 2000; 267: 2183–2188. <https://doi.org/10.1098/rspb.2000.1267> PMID: 11413631
47. Ohtsuki A, Sasaki A. Epidemiology and disease-control under gene-for-gene plant-pathogen interaction. *J Theor Biol.* 2006; 238: 780–794. <https://doi.org/10.1016/j.jtbi.2005.06.030> PMID: 16085107
48. Lloyd-Smith JO, George D, Pepin KM, Pitzer VE, Pulliam JRC, Dobson AP, et al. Epidemic dynamics at the human-animal interface. *Science.* 2009; 326: 1362–1367. <https://doi.org/10.1126/science.1177345> PMID: 19965751
49. Pepin KM, Lass S, Pulliam JRC, Read AF, Lloyd-Smith JO. Identifying genetic markers of adaptation for surveillance of viral host jumps. *Nat Rev Microbiol.* 2010; 8: 802–813. <https://doi.org/10.1038/nrmicro2440> PMID: 20938453
50. Olival KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL, Daszak P. Host and viral traits predict zoonotic spillover from mammals. *Nature.* 2017; 546: 646. <https://doi.org/10.1038/nature22975> PMID: 28636590
51. Day T, Read AF. Does high-dose antimicrobial chemotherapy prevent the evolution of resistance? *PLoS Comput Biol.* 2016; 12: e1004689. <https://doi.org/10.1371/journal.pcbi.1004689> PMID: 26820986
52. Bell G, MacLean C. The search for ‘evolution-proof’ antibiotics. *Trends in Microbiology.* 2017; <https://doi.org/10.1016/j.tim.2017.11.005> PMID: 29191398
53. Kennedy DA, Read AF. Why does drug resistance readily evolve but vaccine resistance does not? *Proc Biol Sci.* 2017; 284. <https://doi.org/10.1098/rspb.2016.2562> PMID: 28356449
54. Kew OM, Sutter RW, de Gourville EM, Dowdle WR, Pallansch MA. Vaccine-derived polioviruses and the endgame strategy for global polio eradication. *Annual Review of Microbiology.* 2005; 59: 587–635. <https://doi.org/10.1146/annurev.micro.58.030603.123625> PMID: 16153180

55. Hanley KA. The double-edged sword: how evolution can make or break a live-attenuated virus vaccine. *Evolution: Education and Outreach*. 2011; 4: 635. <https://doi.org/10.1007/s12052-011-0365-y> PMID: [22468165](https://pubmed.ncbi.nlm.nih.gov/22468165/)
56. Sasaki A, Haraguchi Y, Yoshida H. Estimating the risk of re-emergence after stopping polio vaccination. *Front Microbiol*. 2012; 3: 178. <https://doi.org/10.3389/fmicb.2012.00178> PMID: [22783231](https://pubmed.ncbi.nlm.nih.gov/22783231/)
57. Pons-Salort M, Burns CC, Lyons H, Blake IM, Jafari H, Oberste MS, et al. Preventing vaccine-derived poliovirus emergence during the polio endgame. *PLoS Pathog*. 2016; 12: e1005728. <https://doi.org/10.1371/journal.ppat.1005728> PMID: [27384947](https://pubmed.ncbi.nlm.nih.gov/27384947/)
58. Bull JJ, Smithson MW, Nuismer SL. Transmissible viral vaccines. *Trends in Microbiology*. 2018; 26: 6–15. <https://doi.org/10.1016/j.tim.2017.09.007> PMID: [29033339](https://pubmed.ncbi.nlm.nih.gov/29033339/)
59. Delmotte F, Bourguet D, Franck P, Guillemaud T, Reboud X, Vacher C, et al. Combining Selective Pressures to Enhance the Durability of Disease Resistance Genes. *Front Plant Sci*. 2016; 7. <https://doi.org/10.3389/fpls.2016.01916> PMID: [28066472](https://pubmed.ncbi.nlm.nih.gov/28066472/)
60. Mundt CC. Pyramiding for resistance durability: theory and practice. *Phytopathology*. 2018; 108: 792–802. <https://doi.org/10.1094/PHYTO-12-17-0426-RVW> PMID: [29648947](https://pubmed.ncbi.nlm.nih.gov/29648947/)
61. Cady KC, Bondy-Denomy J, Heussler GE, Davidson AR, O'Toole GA. The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *J Bacteriol*. 2012; 194: 5728–5738. <https://doi.org/10.1128/JB.01184-12> PMID: [22885297](https://pubmed.ncbi.nlm.nih.gov/22885297/)