PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

What limits the evolutionary emergence of pathogens?

S. Gandon, M. E. Hochberg, R. D. Holt and T. Day

Phil. Trans. R. Soc. B 2013 **368**, 20120086, published online 3 December 2012

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

rstb.royalsocietypublishing.org

Research

Cite this article: Gandon S, Hochberg ME, Holt RD, Day T. 2012 What limits the evolutionary emergence of pathogens? Phil Trans R Soc B 368: 20120086.
http://dx.doi.org/10.1098/rstb.2012.0086

One contribution of 15 to a Theme Issue 'Evolutionary rescue in changing environments'.

Subject Areas:
evolution, ecology, health and disease and epidemiology, computational biology

Keywords:
evolution, extinction, stochasticity, disease, mutation

Author for correspondence:
S. Gandon
e-mail: sylvain.gandon@cefe.cnrs.fr

What limits the evolutionary emergence of pathogens?

S. Gandon[1], M. E. Hochberg[2,3], R. D. Holt[4] and T. Day[5]

[1]CEFE, CNRS, 1919 route de Mende, Montpellier, 34293, France
[2]ISEM UMR 5554, CNRS, Montpellier, France
[3]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
[4]Department of Biology, University of Florida, Gainesville, FL, USA
[5]Departments of Mathematics and Biology, Queen's University, Kingston, Ontario, Canada

The ability of a pathogen to cause an epidemic when introduced in a new host population often relies on its ability to adapt to this new environment. Here, we give a brief overview of recent theoretical and empirical studies of such evolutionary emergence of pathogens. We discuss the effects of several ecological and genetic factors that may affect the likelihood of emergence: migration, life history of the infectious agent, host heterogeneity, and the rate and effects of mutations. We contrast different modelling approaches and indicate how details in the way we model each step of a life cycle can have important consequences on the predicted probability of evolutionary emergence. These different theoretical perspectives yield important insights into optimal surveillance and intervention strategies, which should aim for a reduction in the emergence (and re-emergence) of infectious diseases.

1. Introduction

Evolutionary rescue occurs when a population in a given environment is expected to go extinct, but nonetheless persists because evolution by natural selection increases fitness rapidly enough to prevent extinction (see [1]). This is a process that is likely to be a recurrent and widespread feature of the coevolutionary dynamics of hosts and pathogens, defining both realized host ranges for pathogens and the responses by each to environmental change. Here, we examine the interplay of infectious disease emergence, evolutionary rescue and responses of coupled host–pathogen systems to environmental change, emphasizing largely a conceptual framework that has itself 'emerged' recently, but also touching on very important public-health problems.

Imagine that a host individual acquires a new infection, and that it is placed in an isolated uninfected host population. When will this primary case lead to an epidemic? Early models in mathematical epidemiology predict that whether or not an epidemic emerges depends on the basic reproductive ratio of the pathogen (R_0), which is the expected number of secondary cases per primary case in an otherwise uninfected population (reviewed in [2,3]). In the classical deterministic description of disease transmission, the pathogen will spread if $R_0 > 1$ and will go extinct otherwise. This simple description of pathogen invasion relies on the underlying assumptions of the deterministic process. The early stages of an invasion are, however, typically characterized by a small number, n, of infected hosts. In such cases, it is necessary to take into account demographic stochasticity in processes such as transmission, recovery and mortality. Using a probabilistic approach leads to a more refined answer to the above question. For example, it can be shown using a branching process that the probability of emergence (i.e. the probability that, after the introduction of n-infected hosts, a non-evolving pathogen avoids initial extinction and leads to an epidemic) is zero when $R_0 < 1$ and is equal to $P = 1 - (1/R_0)^n$ when $R_0 > 1$ (this result holds in classical epidemiological models assuming that the duration of infection is exponentially distributed and contacts follow a Poisson process [3]). As the initial number n of introduced individuals

Royal Society Publishing
Informing the science of the future

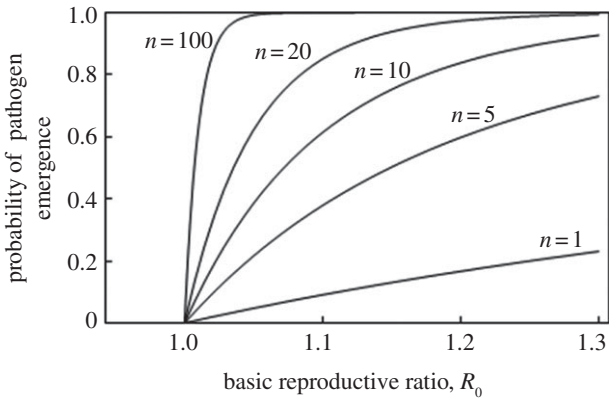

Figure 1. Effect of the basic reproductive ratio, R_0, and the inoculum size, n, on the probability of emergence of a pathogen.

becomes large, the probability of emergence approaches the all-or-nothing deterministic description (figure 1). Even at low n, a large R_0 implies a high probability of establishment [4]. When R_0 is not much greater than unity, interesting complexities arise in characterizing the probability of emergence, for instance, owing to heterogeneity in the host population [5]. But even in such cases, if there are recurrent introduction events, eventually the disease will emerge. By contrast, if $R_0 < 1$, without evolution the pathogen will never emerge, no matter how many spillover events occur onto the novel host population. For such host–pathogen combinations, disease emergence requires evolutionary rescue.

If the pathogen can evolve, then an epidemic might occur even if $R_0 < 1$ initially, because mutations could arise that make $R_0 > 1$ before extinction occurs. In such cases, there is a race between the process of extinction and the process of adaptation to the new host [6,7]. In this study, we will focus on such situations. If a pathogen can evolve to its new host, then this can dramatically increase the range of situations leading to epidemics. Our aim is to identify the main factors that govern the probability of such evolutionary emergence (i.e. the probability that, after the introduction of a maladapted form of the pathogen, the pathogen evolves thereby avoiding initial extinction and in so doing, generates an epidemic). Because we are effectively dealing with the question of 'evolutionary rescue', many of our conclusions have analogies with other studies in this special issue.

We first present a derivation of the probability of evolutionary emergence in a simple, but quite general, ecological scenario. This permits the evaluation of how several factors affect the risk of evolutionary emergence. We will then relax some of the assumptions behind this ecological scenario, which lead to more complex, yet realistic and relevant, situations. Finally, we discuss the available empirical evidence. Our aim is to use these different evolutionary scenarios to better understand what limits the adaptation of pathogens, which is key to managing the risks of infectious disease emergence.

2. Probability of evolutionary emergence

We begin by considering the following ecological scenario. A novel pathogen with clonal, asexual reproduction is introduced into a large host population of size N, which is closed to immigration and emigration. We assume direct transmission. The transmission rate of the pathogen to susceptible hosts per infected host is β, and the constant *per capita* mortality induced by infection (i.e. the virulence) is α. If we assume that the *per capita* natural host mortality rate and pathogen clearance (recovery) rate to be constants δ and γ, respectively, this yields the following expression for the basic reproduction ratio:

$$R_0 = \frac{\beta N}{\delta + \alpha + \gamma}. \tag{2.1}$$

Because we focus on evolutionary emergence, we are interested in situations in which the novel pathogen is maladapted and thus doomed to extinction in the absence of adaptation (i.e. $R_0 < 1$ initially). Adaptation permitting persistence could occur by the acquisition of mutations that will affect various pathogen life-history traits. In principle, adaptation could occur through an increase in transmission or a decrease in virulence or recovery (i.e. clearance). Ultimately, such adaptation will lead to a 'new' pathogen with a basic reproductive ratio R_0^* exceeding unity. Under the assumptions that: (i) a single mutational step is required to reach the adapted genotype, (ii) mutation is directional (no back-mutation towards the maladapted wild-type) and (iii) the mutation rate is small, the probability of evolutionary emergence from a single initially infected individual is [6–8]

$$P_e \approx \frac{1}{1 - R_0}[uR_0 + \mu L]P^*, \tag{2.2}$$

where u is the probability of adaptive mutation during a transmission event to a new host, μ is the rate of fixation of adaptive mutations within a host during the infection, $L = 1/(\delta + \alpha + \gamma)$ is the average duration of an infection and $P^* = 1 - 1/R_0^*$ (see §1 and figure 1 for $n = 1$) is the probability of emergence of the 'new' pathogen with a basic reproductive ratio R_0^*.

The above expression isolates three distinct quantities driving evolutionary emergence. First, the quantity $1/(1 - R_0)$ measures the expected cumulative number of cases induced after the introduction of a single infected host with the maladapted pathogen, before it goes extinct. This is equal to $\sum_{i=0}^{\infty} R_0^i$, where i refers to the position in the epidemic chain that derives from the first case (i.e. $i = 1$ refers to secondary cases derived from the first infectious case, $i = 2$ refers to the infections deriving from the secondary cases...). In other words, the probability of emergence is proportional to the expected size of the epidemic induced by the maladapted mutant. Second, the above expression depends linearly on two different mutation processes. Mutation may occur conditional on a transmission event, and this will occur on average uR_0 times per host infected with the maladapted genotype. The fixation of adaptive mutations may also take place within the host during the course of the infection, and because the average infection is expected to last L units of time, this will produce μL new adaptive mutations per host initially infected with the maladapted genotype. Third, once the adapted genotype is present in the local pathogen population, it must 'escape' initial extinction and persist, and this occurs with probability $P^* = 1 - 1/R_0^*$.

This analytical expression is useful to gain an understanding of the factors governing evolutionary emergence. In particular, the above three terms clearly show the impact of (i) demography (the chain of infections before the appearance of an adapted genotype), (ii) the mutation process and

(iii) the level of adaptation of the emerging pathogen. In the following, we will use the above description as a framework for discussing these three effects in light of other theoretical developments, specifically various complexities associated with pathogen epidemiology.

3. Migration and reintroductions

Migration is classically viewed as a force that counteracts natural selection through the flow of maladapted genes and, as such, a force that limits adaptation [9]. Yet, this classical result for clonal or major gene models relies on the underlying assumption of pronounced density dependence. Source–sink metapopulation models have shown that the effect of migration is strongly contingent on the amount of density dependence in the sink. In the absence of density dependence, recurrent migration can enhance adaptation to the sink by infusing more variation sampled from the source, and by sustaining a higher sink population size, which results in more mutational input in the local environment [10]. The occurrence of some density dependence in the sink may lead to non-monotonic effects of migration [11, fig. 6].

In the context of infectious disease emergence and the transmission from animals to humans (i.e. a zoonosis), the animal reservoir can be viewed as the source and the human population as the (initial) sink. Migration in this case refers to the recurrent introduction of pathogens into the human population. Pathogen progeny are likely to have access to a large number of uninfected, naive human hosts, which means that in general they are unlikely to experience significant competition during the initial phases of the epidemic. In this situation, the more immigration events, the more likely it is that the pathogen can adapt to the novel host. Indeed, one can use the above criteria to study the effect of the number (n) of independent introduction events on the probability of evolutionary emergence in the human population, which is $1 - (1 - P_e)^n$, where P_e is the probability for a single introduction. (This assumes that introductions are separated enough in time or space that a given introduction either goes extinct or adapts before overlapping with any other colonizing attempt). This confirms that the higher the propagule pressure (both in terms of number of infected individuals per infection episode, and the number of distinct infection episodes), the higher the probability of emergence [12].

If a human population is in contact with an animal reservoir, emergence may also be facilitated by indirect transmission routes, where the emergent strain could also circulate via back-dispersal into the reservoir. Reluga *et al.* [13] modelled this process and confirmed that more contact with the reservoir host can facilitate pathogen emergence. They assumed that a mutation increasing transmission in the novel host has a neutral effect in the reservoir, and that back-transmission does not reduce potential transmission rates within the novel host. Modifying these assumptions would make emergence less likely.

What is less obvious is whether the probability of evolutionary emergence is higher or lower when initial pathogen introductions are clustered rather than being spread out in space or in time. In a spatially and temporally homogeneous environment, it often does not matter. When there are temporal [14] or spatial [15] heterogeneities, however, in purely ecological models, clustered introductions always lead to a lower probability of emergence. Although it has not yet been formally analysed, this is likely true for evolutionary rescue scenarios as well. The reason is that clustered introductions do not benefit from an 'exploration' of environmental heterogeneity, which leads to an increase in the chance of an introduction occurring at the right point in space and time.

However, there may be some situations for which clustering of infections is advantageous and promotes disease emergence via evolutionary rescue. This can arise if there are analogues of Allee effects in transmission, or in host demography, at low numbers of infected individuals. For example, the reason mortality rate may be elevated in infected hosts is that they become vulnerable to predation. If predators can be readily satiated, an increase in the local abundance of infected hosts can reduce the mortality rate per infected host. This may increase the duration of the infection, L, and the likelihood that appropriate mutations permitting persistence will arise and fuel emergence. Holt *et al.* [16] considered source–sink models with an Allee effect, and showed for several different scenarios that an increase in the number of individuals introduced per colonizing episode (immigration rate) could enhance adaptation to the sink. The models were not explicitly about host–pathogen interactions, but can be readily interpreted to encompass them.

On a related issue, if the introduction consists of n different genotypes, as n increases, this will increase the chance of a 'better' genotype being present. A situation with initial genetic variation is discussed elsewhere [17, eqn 2.1a, p. 2947]. If the pathogen is sexual, an additional effect of recurrent migration can arise: matings between immigrants and better-adapted residents impose a migrational load on the latter. This can lead to alternative evolutionary states in a sink population: a low density one, permanently maladapted owing to gene flow constraining local adaptation, and one at high density, for which immigration is quantitatively small relative to local carrying capacity [18,19]. Disease emergence then can be influenced by transient temporal variation in the host population, boosting transmission or inhibiting recovery (as is shown for a more general case, not specific to host–pathogen systems, in [18]).

4. Maladaptation and life history of the ancestral pathogen

As pointed out above, R_0 of the ancestral pathogen in the novel host governs the length of the epidemic chain before extinction. This directly affects the opportunities for mutating away from the ancestral type. But beyond this effect, the details of the life cycle of the ancestral pathogen may strongly influence the likelihood of evolutionary disease emergence.

It is important to note that different pathogens with the same R_0 may have different probabilities of emergence if they have different values of L, the average duration of the infection. Indeed, mutation and evolution are likely to operate during the course of the infection in each individual host. Hence, the longer the duration of infection, the more opportunities for the emergence of an adapted mutant. André & Day [6] show that this result is robust to alterations in the life cycle assumed in our baseline model. In particular, the expression for P_e (equation (2.1)) holds even when transmission, death and clearance rates vary with the age of the

infection (these modifications of the life cycle may, however, affect the detailed calculation of R_0 and L). These authors also explore a situation in which the rate of mutation μ (which refers to the process of within-host pathogen adaptation) may vary with the age of the infection. Again, the expression for the probability of evolutionary emergence still holds, provided that μ is replaced by $\bar{\mu}(L)$, the average rate of within-host adaptation for an infection of duration L. This generalization shows the robustness of the previous conclusion but also opens avenues for further developments. Allowing for variation in the rate of within-host adaptation is a first step towards a more explicit description of the process of selection occurring among pathogens competing within each host. In many situations, a more efficient within-host exploitation strategy leads to more transmission. Yet selection within and between hosts may be very different. There are cases of short-sighted evolution [20], where within-host evolution can lead to lower transmission ability (and lower R_0), because the factors underlying within-host competitiveness are not necessarily those that maximize between-host transmission. This effect could be formalized by allowing the rate of within-host evolution to the new adapted strain to decrease with time. In this case, one can show that an increase in L may not necessarily lead to an increase in P_e. Further investigations are required to study scenarios with a more detailed description of within-host processes that would relax the unrealistic assumption that the sweep of mutations is instantaneous [21–23].

5. Host heterogeneity and contact networks

The above reasoning assumes that the host population is homogeneous and that death, transmission and recovery do not differ among hosts. Infected hosts, however, may differ greatly in age, sex, behaviour, spatial aggregation, genetic background and so forth, and each of these factors may affect pathogen life-history traits and the potential for disease emergence. In particular, the occurrence of superspreading, in which a few individuals infect an unusually large number of secondary cases, has been observed in many infectious diseases [24,25]. Taking into account this heterogeneity presents a major theoretical challenge. Several earlier studies have investigated the impact of different forms of heterogeneity on the probability of emergence in the absence of evolution [4,25]. Holding the expected value of R_0 constant, heterogeneity among hosts may also affect the probability of emergence. Indeed, using a phenomenological approach that enables the use of various distributions of the expected number of secondary cases caused by a particular infected host (individual reproductive number), Lloyd-Smith *et al.* [25] have shown that for the same average value of R_0, an increase in the variation of the individual reproductive number has two main effects. More variation reduces the probability of emergence, but when an outbreak does occur the epidemic size is increased. These two effects can be illustrated with a simple example (see appendix A). Imagine two pathogens with the same expected value of $R_0 = 2$. In the first pathogen, there is no variation in the individual reproductive number, but in the second, the individual value varies and is either 0 or 4, with equal probability. In the first pathogen, $P = 0.5$ (see §1). For the second pathogen, the probability of emergence

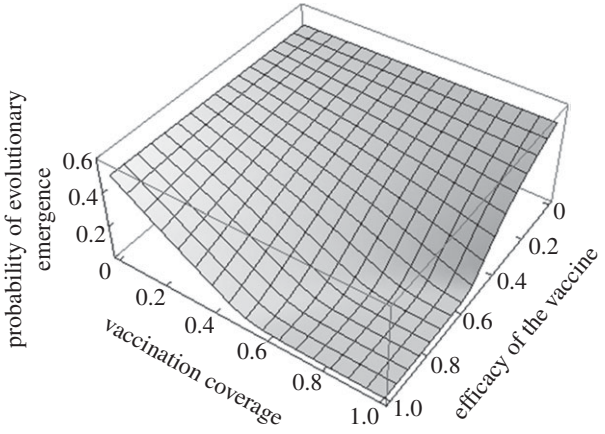

Figure 2. Effect of vaccination coverage, $1 - f$, and vaccine efficacy, r, on the probability of evolutionary emergence of a pathogen with homogeneous mixing, $\pi = 0.5$. Vaccine efficacy only affects the infectiousness of the vaccinated host, $b_2 = (1 - r)b_1$, and does not influence its susceptibility, $\sigma = 1$. Here, we consider the situation in which the vaccine has an effect only on the maladapted strain. The adapted strain is considered so fit (on both naive and vaccinated hosts) that we can neglect its risk of early extinction (see appendix B for more details on this scenario). Parameter values: $b_1 = 2$, $d_1 = 1$, $d_2 = 1$, $u = \mu = 10^{-3}$.

is equal to $P = 0.25$ (see appendix A). In this case, the risk of early extinction is increased when the pathogen encounters a poor-quality host, and the higher probability of emergence when the pathogen gets lucky and infects a good-quality host does not compensate for this effect. Yet when the epidemic takes off, the presence of good-quality hosts results in a larger epidemic size. Note that, interestingly, Yates *et al.* [5] found that the type of heterogeneity matters as well, and in particular that variations in some host properties (e.g. susceptibility to infection) have no impact on emergence. In appendix A, we study a very similar version of the model in Yates *et al.* [5], but for a continuous time birth–death model.

The above situations did not allow the pathogen to adapt to the new host. The impact of host heterogeneity on evolutionary emergence has been explored in only a handful of studies. Yates *et al.* [5] showed that host heterogeneity has a very weak effect on the probability of evolutionary emergence. An approach based on an explicit description of the contact process between hosts and the network of mutations has been used to study evolutionary emergence [26]. Studying the effect of host heterogeneity is perhaps more complex in this situation, since a modification of the contact network has direct effects on the variance as well as on the expected value of R_0 (see [4]) and this latter effect has a well-known direct effect on emergence (see above expression for P_e). Yet this approach is very promising because it is based on a more detailed presentation of the environment in which infection chains play out. As such, it paves the way for several important directions of future research, such as the study of evolutionary disease emergence in more realistic spatially structured models. In appendix B, we study evolutionary emergence assuming a fraction $1 - f$ of the population is vaccinated against the pathogen. We derive some analytical results in the very special case in which the vaccine is perfect; the effects of the efficacy of imperfect vaccines and of vaccination coverage on the probability of evolutionary emergence are shown in figure 2. This analysis confirms that vaccination may be an efficient measure to limit evolutionary emergence.

6. Mutations

Mutation is the ultimate fuel of evolution and it plays a key role in the process of evolutionary emergence. We have already discussed in §4 the importance of whether mutations are conditional on transmission or not, but other details of mutation matter as well.

(a) Distribution of fitness effects

How many mutations confer a benefit in the novel environment? When multiple mutations are simultaneously present, are their effects on fitness simply additive, or more complex? The answers to these questions require an underlying description of the fitness landscape. The above calculation for the probability of evolutionary emergence relies on a very simple fitness landscape, where the adaptive mutation can be reached in one step. Some generalization can be obtained in other relatively more complex scenarios. For example, if m individually neutral mutations (i.e. each single mutation does not change the traits of the maladapted pathogen on the novel host) are required before reaching a significant increase in fitness, then the probability of evolutionary emergence becomes [6,7]:

$$P_e \approx \frac{1}{(1-R_0)^m}[uR_0 + \mu L]^m P^*. \tag{6.1}$$

More realistic fitness landscapes would help refine these predictions. For example, Fisher's geometric model of adaptation provides a framework to incorporate very important feedbacks between the level of maladaptation to the novel host and the fraction of beneficial mutations. One interpretation of this geometric model might be that when the pathogen is initially less adapted to the host (i.e. lower R_0), the fraction of mutations that are beneficial in the novel host may be larger, which may reduce the impact of initial maladaptation on the probability of evolutionary emergence. This requires further theoretical and experimental developments. Orr & Unckless [27] integrated Fisher's approach with the evolutionary rescue scenario considered in [26], and the former's models could be modified to describe disease emergence. On the theoretical side, the approach of Alexander & Day [26] with an explicit description of the network of mutations may provide a useful framework to study this.

(b) Effects of mutation on life-history traits

By definition, we assume that the adapted genotype has $R_0^* > 1$ and thus has a probability of emergence of $P^* = 1 - 1/R_0^*$. In other words, not surprisingly, when comparing alternative mutations that can potentially lead to the evolutionary emergence of a persistent infectious disease, the strain with the highest R_0 has the highest probability of emergence. Once established, in the early phase of an epidemic, however, the strain with the highest instantaneous *per capita* growth rate (i.e. Malthusian fitness $r_0 = \beta N - (\delta + \alpha + \gamma)$) increases faster, and may thus be viewed as the most competitive one. Indeed, the strain with the highest r_0 will increase faster in a fully susceptible host population. Strains with high r_0 generally have high R_0, but this is not always the case. So a strain with a higher R_0 may actually be less competitive because a different strain could have a higher r_0 [21,28,29]. Yet, it is the one with the higher R_0 that will be better at initially avoiding extinction. As an attempt to better understand this result, consider it in the light of a classical diffusion approximation that shows the importance of the distribution of offspring number on the probability of early extinction. In particular, to escape early extinction, a strain benefits from an increase in its expected Malthusian growth rate r_0, but also a decrease in the variance in its growth rate (see appendix C). In our simple epidemiological model, there is a link between the mean and the variance in the growth rate, and maximizing R_0 strikes the appropriate balance between the two (it maximizes the mean to variance ratio). Figure 3 (modified from [33] and [34]) presents the potential implications of this result for the evolutionary dynamics in the very early phase of an emergence. Figure 3 shows the basic elements of the life history of the infection: birth (transmission) on the vertical axis, and death (actual death whatever the cause, plus pathogen clearance) on the horizontal axis. A line of slope unity from the origin refers to the condition $R_0 = 1$, or equivalently $r_0 = 0$. The ancestral maladapted strain is described by a point (black dot) below this line, where transmission cannot match losses. But mutants arise in some neighbourhood (the circular region) of this ancestral state, and some mutants can have $R_0 > 1$ (black area). We note that in this space, the family of dashed black lines emanating from the origin describe different equivalence sets in terms of R_0. By contrast, the family of grey lines parallel to the $r_0 = 0$ line are contour lines with the same value of r_0. Note that any prediction of the evolutionary trajectories will require some knowledge about the effects of mutations on the various pathogen life-history traits. In particular, in this heuristic figure, we allowed each cloud of feasible mutations to have the same size, in effect assuming that mutations have constant additive effects on these demographic parameters, regardless of initial conditions. Whether or not this is a reasonable assumption will depend upon the biological mechanisms underlying each of these demographic parameters.

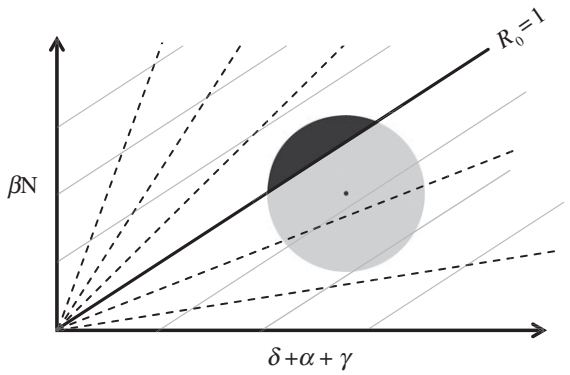

Figure 3. Schematic of the effect of mutations on various life-history traits of the pathogen (virulence, α, and transmission, β) on the probability of evolutionary emergence. The ancestral pathogen (black dot) is producing a cloud of mutants (in grey). Among those mutants, those with $R_0 > 1$ will have a chance to emerge (in black). The probability of emergence is not governed by the *per capita* growth rate of those mutants (the grey lines indicate contour lines with the same value of r_0) but by their basic reproduction ratio (the dashed lines are contour lines with the same value of R_0).

Although the results discussed above are based on a rather general birth–death process, deviations from such processes are also common in many infectious diseases. The relationship between R_0 and r_0 is very sensitive to the details of the pathogen life cycle and the distribution of generation time [35]. For example, some pathogens exhibit a lytic life

cycle in which the propagules are formed and stored during an infection and then all simultaneously released upon killing the host. Models for such life cycles have been referred to as 'burst–death' processes [34] and allow for evolutionary adaptation in burst size, time to burst or clearance rate. As with the above considerations, these models also reveal that pathogen life history can play an important role in evolutionary emergence, with mutations affecting some traits being more likely to lead to adaptation than others [34].

(c) Mutation rates

The above approximation for the probability of evolutionary emergence shows that it increases linearly with the input of beneficial mutations. An underlying assumption behind this calculation is that the mutation rate is relatively low. Although this is reasonable for the majority of pathogens, some viruses, and in particular RNA viruses, may have very high mutation rates [36]. This may violate the above model and could alter the effect of mutation rate on disease emergence. Indeed, because the vast majority of mutations are deleterious, a large increase in the mutation rate could load the genome and prevent potentially beneficial mutations from rescuing a maladapted population. This is the idea of an 'error threshold' that may ultimately lead pathogen populations to extinction [37,38]. Some interesting scenarios are explored in [26]. Two main situations may arise. If the original strain is very maladapted to the new host (i.e. $R_0 \ll 1$), a large fraction of new mutations will be beneficial, and increasing mutation will always favour evolutionary emergence. By contrast, if the original strain is only weakly maladapted (i.e. $R_0 \approx 1$), then increasing mutation rate can have the opposite effect on evolutionary emergence. This effect, however, requires further theoretical investigation to determine when, in general, increased mutation rate is expected to favour evolutionary emergence.

7. Discussion

In the above sections, we derive the probability of evolutionary emergence under various scenarios. This theoretical approach helps identify a diverse range of factors that play key roles in evolutionary emergence. Before discussing the implications of these theoretical predictions, we want to review briefly the empirical evidence that bears on these questions.

(a) Empirical evidence supporting the above theoretical predictions

Regarding the impact of migration and reintroduction, a direct prediction from the above models is that species jumps are more likely to occur between species that regularly share the same environment simply because there are more opportunities for frequent reintroductions between sympatric species. At a broad scale, geography has been found to be a major determinant of species jumps among pathogens of wild primates and humans [39]. Similarly, transmission of rabies virus across different North American bat species appears to be limited by geographical range overlap of the different bat species [40]. At more local scales, one would predict greater likelihood of disease emergence for species that have similar habitat requirements and phenologies, which would increase overlap in space and time and permit multiple attempts at cross-host colonization. We are unaware of direct assessments of this prediction.

Concerning the effect of initial maladaptation, a direct prediction is that jumps are more likely to occur between species that are phylogenetically more similar (because the 'gap' between pathogen fitness in the two host species should be smaller). In the above two empirical studies [39,40], there was a strong negative effect of the phylogenetic distance between host species. Davies & Pedersen [39] found, however, that among viral pathogens, cross-species transmission was more limited by geography than by divergence time between hosts. They argue that this could be due to the higher evolutionary potential of viruses compared with other pathogens. A well-documented example of a species jump (between related species) leading to virus emergence is the outbreaks of Chikungunya virus in the Indian Ocean. Sequencing of viral isolates revealed that emergence was linked to a few mutations allowing the virus to adapt to a new vector species, *Aedes albopictus* (the virus is usually transmitted by *A. aegypti*) [41,42]. Here, the adaptation to a new vector species led to a massive increase in transmission and to the re-emergence of the disease in human populations.

The evolutionary potential for emergence is mainly governed by the pathogen mutation rate. The above theory shows that an increase in the mutation rate is generally expected to increase the probability of evolutionary emergence. There are several studies [43] showing indeed that emergence is more likely in RNA viruses, which are characterized by large mutation rates. Beyond this simple qualitative prediction of the effect of mutation rates, there is very little empirical evidence to use to investigate the importance of the distribution of mutation effects on fitness. There are an increasing number of studies measuring the distribution of fitness effects (DFE) of mutations (especially in viruses, [44]). These studies, however, are limited to a measure of fitness in a single environment, and typically in the original host where the parasite is already well adapted. A recent study [45] provides a measure of the DFE of a plant virus on eight different host species. This study confirms the prediction of Fisher's geometric model of adaptation that more beneficial mutations are observed in novel host species. As pointed out already, it is important to determine the life-history traits of adaptive mutations to quantify their probability of emergence. Further experimental work is required in this area to obtain the effects of mutations on transmission, virulence and recovery rates.

Although there is some empirical evidence supporting some qualitative predictions of the theory, there are still very few experimental attempts to test predictions on emergence and evolutionary emergence. The fact that these are stochastic processes means multiple replicate populations are required, and this experimental effort is simply impossible in some biological systems (e.g. pathogens of vertebrate species). Pathogens of microbes may, however, provide a good model system to study emergence [46,47].

(b) Which particular pathogens should we watch out for?

The above theoretical framework indicates that if we want to limit the risk of emergence, we should focus on pathogens with the following three properties: (i) pathogens that have many opportunities to enter into contact with human populations, (ii) pathogens with large mutation rates,

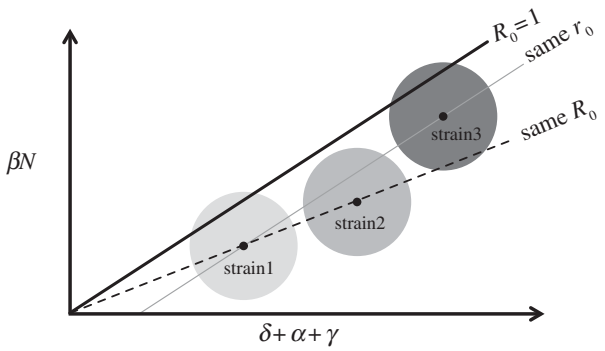

Figure 4. Contrasting the risk of evolutionary emergence for three pathogens. The first and third pathogens (strains 1 and 3) have the same *per capita* growth rate, r_0. The first and the second pathogen (strains 1 and 2) have the same basic reproduction ratio, R_0.

(iii) pathogens that are already reasonably adapted to the host, or that infect related hosts (e.g. primates). Also of importance are the life-history characteristics of pathogens that have a particularly high chance of emergence. Common sense would predict the evolutionary emergence of more deadly pathogens (such as strains 2 and 3 in figure 4, assuming they are further to the right due to a high pathogen-induced death rate). By contrast, the pathogen with the highest probability of evolutionary emergence may be that with the lowest virulence and transmission (strain 1 in figure 3). Under the assumption that the effects of mutations on the traits are the same, mutations will have a higher effect on R_0 for an avirulent pathogen than for a virulent one, all else being equal, because the avirulent pathogen is closer to the origin. Hence, if two pathogens have the same R_0 (e.g. strains 1 and 2), it is the one with the lower virulence (strain 1) that has the higher probability of evolutionary emergence. If the two pathogens have the same r_0 (e.g. strains 1 and 3), it is not so clear which one is more likely to emerge. Strain 3 has an advantage because it has a higher R_0, which means it will create a higher number of cases initially. Yet the effect on R_0 of each of the adaptive mutations on strain 3 is lower than on strain 1. Besides, because strain 1 will generate longer infections (because of lower virulence), the rate of production of new adapted genotypes will be higher for strain 1 (see §4). All this can be formalized using the equation for P_e and additional assumptions regarding the distribution of the effects of mutations on each trait.

(c) How do we limit pathogen emergence?

The theoretical studies we reviewed above point towards general rules of thumb to limit evolutionary emergence (see also [17] and [26]). First, a reduction in the rate of effective contact (through hygiene, vaccination and other measures) with novel pathogens is always expected to limit emergence (appendix A) and evolutionary emergence (see figure 2 and appendix B). Second, one may also try to limit the duration of the infectious period (through treatment and quarantine). It may also be possible to use more complex strategies targeting super-spreaders to limit transmission more effectively [25]. It is difficult, however, to go beyond these classical public-health control interventions. In principle, the above theoretical framework may provide ways to quantify the risk of emergence for each pathogen. But to do this we would need a better knowledge of the distribution of mutational effects on the life-history traits of these pathogens. So far, there are very few data available on this and we hope the present study motivates future experimental study in this direction.

One important area of inquiry both for our fundamental understanding of pathogen emergence, and for applications to areas as diverse as species conservation, agriculture and human health, is to understand how pathogen virulence may be associated with the probability of emergence and how virulence may evolve during the course of an evolutionary rescue. André & Hochberg [48], using a model with density-dependent disease transmission, found that the size of the host population into which the pathogen invades is not only crucial for emergence, but also for the evolved virulence. Specifically, only low virulence pathogens can emerge in small host populations, whereas a range of virulences can succeed in large host populations (see also [49,50]). But it may be worth noting that after pathogen emergence the evolutionary dynamics can be described within a very different framework that neglects the impact of stochasticity. Indeed, as soon as the pathogen population becomes large enough to escape the risk of extinction, the evolutionary trajectories need to be tracked together with the epidemiological dynamics because the two dynamics will feed back on one another [28,29,33,51–54].

Our theoretical treatment has specifically focused on infectious diseases with direct transmission. Parasites that are transmitted by vectors, or have complex life histories with multiple host species, will have different expressions than equation (2.1) for R_0, and expression (2.2) may not directly describe the probability of emergence for such infectious diseases. It would be valuable to develop comparable theories for such parasites, in order to make more detailed comparative statements about the relative likelihood of disease emergence for different classes of infectious diseases. But our general conclusions that to minimize the risk of evolutionary emergence, one should lower the initial R_0 of the infection as much as possible, and likewise reduce whenever feasible the frequency of contacts between ancestral hosts of the pathogen and potential novel hosts, are likely robust across a wide spectrum of host–parasite scenarios. We believe that understanding the evolutionary dimensions of emerging diseases is a topic of vital concern for human well-being, and for species conservation. Theoretical studies such as those we have presented here can help clarify the rationales for particular mitigation and intervention strategies.

We thank the organizers of the Montpellier conference on evolutionary rescue for their invitation, Mike Barfield, Amaury Lambert and Guillaume Martin for comments and discussions. R.D.H. thanks the University of Florida Foundation and NIH GM083192 for research support. M.E.H. thanks the ANR 'EvolStress' (09-BLAN-099-01) and the McDonnell Foundation (JSMF 220020294/SCS-Research Award) for research funding. S.G. acknowledges financial support from CNRS and European Research Council Starting grant 243054 EVOLEPID. This work was also supported by the French Agropolis Fondation (RTRA—Montpellier, BIOFIS project no. 1001–001).

Appendix A. Probability of emergence in a heterogeneous host population with no evolution

Consider an infection caused by a pathogen in a heterogeneous population that consists of two different types of

hosts: good-quality hosts in proportion f (where the birth rate and the death rates of the infection are b_1 and d_1, respectively), and bad-quality hosts in proportion $(1 - f)$ (where the birth rate and the death rates of the infection are b_2 and d_2, respectively). The 'birth' of an infection means infection of an additional host, while the 'death' means the termination of an infection through host death or recovery. We further assume that bad-quality hosts may have an equal or lower probability σ (i.e. $\sigma \leq 1$) of becoming infected upon contact with the pathogen. In order to calculate the probability of emergence, we first derive the probability $Q(t)$ that an introduced pathogen, present in the population at time t in a single host, ultimately goes extinct. This probability is equal to $Q(t) = fQ_1(t) + (1 - f)Q_2(t)$, where $Q_1(t)$ and $Q_2(t)$ refer to the probability of ultimate extinction when the pathogen is initially introduced into a good or a bad-quality host, respectively. These two quantities can be derived by considering all the events that might occur during an infinitesimal period $\mathrm{d}t$ (so that either birth or death is possible, but not both; the three possible states of the world at time $t + \mathrm{d}t$ are thus that the initial infected host has survived and infected another host, or it has died, or nothing has been changed):

$$Q_1(t) = b_1\mathrm{d}t\, Q_1(t+\mathrm{d}t)\, Q'(t+\mathrm{d}t) + d_1\mathrm{d}t + Q_1(t+\mathrm{d}t) \times (1 - b_1\mathrm{d}t - d_1\mathrm{d}t)$$

and

$$Q_2(t) = b_2\mathrm{d}t\, Q_2(t+\mathrm{d}t)\, Q''(t+\mathrm{d}t) + d_2\mathrm{d}t + Q_2(t+\mathrm{d}t) \times (1 - b_2\mathrm{d}t - d_2\mathrm{d}t),$$

where $Q'(t+\mathrm{d}t) = AQ_1(t+\mathrm{d}t) + (1-A)Q_2(t+\mathrm{d}t)$ and

$$Q''(t+\mathrm{d}t) = BQ_1(t+\mathrm{d}t) + (1-B)Q_2(t+\mathrm{d}t),$$

with

$$A = \frac{f\pi}{f\pi + \sigma(1-f)(1-\pi)}$$

$$B = \frac{f(1-\pi)}{f(1-\pi) + \sigma(1-f)\pi}.$$

The parameter π refers to the contact structure: $\pi = 1/2$ refers to homogeneous mixing, $\pi > 1/2$ to assortative mixing and $\pi < 1/2$ to dissortative mixing. Playing with π, σ, b_2 and/or d_2 allows one to study the effects of different types of variability in the host population on emergence (see [5] for a similar life cycle but with a discrete time branching model).

The above equations assume that the density of susceptible hosts and the relative frequency of the two host classes are constant during the whole stochastic process. The probabilities of ultimate loss are thus independent of time and the probability of emergence is $P = 1 - Q$, where Q is obtained from the resolution of

$$b_1Q_1Q' + d_1 - Q_1(b_1 + d_1) = 0$$

and

$$b_2Q_2Q'' + d_2 - Q_2(b_2 + d_2) = 0.$$

In the main text, we discuss a situation in which $f = 1/2$, $\sigma = 1$, $b_1 = 4$, $b_2 = 0$ and $d_1 = d_2 = 1$ which yields $P = 0.25$. More generally the probability of emergence is (when $b_2 = 0$): $P = f - (d_1/b_1)$ (when $f > d_1/b_1$) and zero, otherwise. This expression can also be rewritten as $P = f(1 - 1/R_{0,1})$ which is perhaps simpler to interpret. The first f is the probability that the initial infected host is a good-quality one, while the second term is simply the probability of emergence given in the main text, after replacing R_0 by $fR_{0,1}$, where $R_{0,1} = b_1/d_1$ (i.e. the basic reproduction ratio of the pathogen in a good-quality host population). For an even more general case where the bad-quality host is infectious (i.e. $b_2 \neq 0$) we get

$$P = \frac{-b_2d_1 + b_1(b_2 - d_2) + \sqrt{(b_2(b_1+d_1) - b_1d_2)^2 + 4b_1b_2(-b_2d_1 + b_1d_2)f}}{2b_1b_2}.$$

For cases where $\pi \neq 1/2$ and $\sigma < 1$, the above conditions can be solved numerically. These results are similar to those reported by Yates *et al.* [5].

Appendix B. Probability of evolutionary emergence in a heterogeneous host population

What is the effect of host heterogeneity on the probability of evolutionary emergence? In particular, what is the effect of vaccinating a fraction $1 - f$ of the population against a pathogen to limit its probability of evolutionary emergence? To answer this question, we can use a very similar approach to account for the additional effect of different mutation pathways (see main text) towards adaptive mutations. As in appendix A, we can derive recurrence equations for the probabilities of ultimate extinction of the maladapted pathogen (in both naive and vaccinated hosts, Q_1 and Q_2, respectively), and the probabilities of ultimate extinction of the adapted pathogen (in both naive and vaccinated hosts, $Q_{a,1}$ and $Q_{a,2}$, respectively):

$$Q_1(t) = b_1(1-u)\mathrm{d}tQ_1(t+\mathrm{d}t)Q'(t+\mathrm{d}t) + b_1u\mathrm{d}tQ_1(t+\mathrm{d}t) \times Q'_a(t+\mathrm{d}t) + d_1\mathrm{d}t + \mu\mathrm{d}tQ_{a,1}(t+\mathrm{d}t) + Q_1(t+\mathrm{d}t) \times (1 - b_1\mathrm{d}t - \mu\mathrm{d}t - d_1\mathrm{d}t)$$

$$Q_2(t) = b_2(1-u)\mathrm{d}tQ_2(t+\mathrm{d}t)Q''(t+\mathrm{d}t) + b_2u\mathrm{d}tQ_2(t+\mathrm{d}t)Q'_a(t+\mathrm{d}t) + d_2\mathrm{d}t + \mu\mathrm{d}tQ_{a,2}(t+\mathrm{d}t) + Q_2(t+\mathrm{d}t)(1 - b_2\mathrm{d}t - \mu\mathrm{d}t - d_2\mathrm{d}t)$$

$$Q_{a,1}(t) = b_{a,1}\mathrm{d}tQ_{a,1}(t+\mathrm{d}t)Q'_a(t+\mathrm{d}t) + d_{a,1}\mathrm{d}t + Q_{a,1}(t+\mathrm{d}t) \times (1 - b_{a,1}\mathrm{d}t - d_{a,1}\mathrm{d}t)$$

and

$$Q_{a,2}(t) = b_{a,2}\mathrm{d}tQ_{a,2}(t+\mathrm{d}t)Q''_a(t+\mathrm{d}t) + d_{a,2}\mathrm{d}t + Q_{a,2}(t+\mathrm{d}t) \times (1 - b_{a,2}\mathrm{d}t - d_{a,2}\mathrm{d}t),$$

where the Q' and Q'' terms are defined as in appendix A and

$$Q'_a(t+\mathrm{d}t) = AQ_{a,1}(t+\mathrm{d}t) + (1-A)Q_{a,2}(t+\mathrm{d}t)$$

and

$$Q''_a(t+\mathrm{d}t) = BQ_{a,1}(t+\mathrm{d}t) + (1-B)Q_{a,2}(t+\mathrm{d}t),$$

are the analogous terms for the adapted pathogen, where the parameters A and B are also defined in appendix A (σ is assumed to be the same for both pathogens).

The above equations assume that the density of susceptible hosts and the vaccination coverage f is constant during the whole stochastic process. The probabilities of ultimate loss are thus independent of time and the probability of evolutionary emergence is $P_e = 1 - Q$, with $Q = fQ_1 + (1 - f)Q_2$. The above system of equations can thus be used to study the effect of vaccination in a broad

range of situations on the probability of evolutionary emergence. We will consider two extreme cases below.

First, let us assume that vaccine is perfect and prevents infection of the vaccinated hosts from both the maladapted and the adapted strain (i.e. $b_2 = b_{a,2} = 0$). In this case, we find that when the mutation rates are assumed to be low a good approximation for the probability of evolutionary emergence is

$$P_e \approx \frac{1}{1 - fR_{0,1}}[ufR_{0,1} + \mu L]P^*,$$

with $P^* = 1 - (fQ_{a,1} + (1-f)Q_{a,2}) = f - 1/R^*_{0,1}$ when $R^*_{0,1} = b_{a,1}/d_{a,1} > 1/f$, and $P^* = 0$, otherwise. This is a generalization of the expression given in the main text, which corresponds to a situation with no vaccination (i.e. $f = 1$). The above expression clearly shows that vaccination with such a perfect vaccine may be an efficient way to reduce the probability of emergence.

Second, to focus on the effect of heterogeneity on the first step of evolutionary emergence (that is, the mutation towards the adaptive mutation), one may also assume that the adaptive mutation has a very high basic reproduction ratio on both the naive and vaccinated hosts (see [5]). In other words, the adaptive mutation, as soon as it arises, can no longer go extinct (i.e. $Q_{a,1} = Q_{a,2} = 0$ in the above equations and $P^* = 1$). The probability of evolutionary emergence can thus be obtained from the following condition:

$$0 = b_1(1-u)Q_1Q' + d_1 - Q_1(b_1 + \mu + d_1)$$

and

$$0 = b_2(1-u)Q_2Q'' + d_2 - Q_2(b_2 + \mu + d_2).$$

Note that this condition is very similar to the one given above without evolution. We plot on figure 2 the effect of vaccination coverage $(1 - f)$ and vaccine efficacy (we assume the efficacy of the vaccine only affects the infectiousness of the vaccinated host, see legend of figure 2) on the probability of evolutionary emergence with homogeneous mixing. When the vaccine is perfect against the maladapted strain, one obtains the following expression for the probability of evolutionary emergence:

$$P_e \approx \frac{1}{1 - fR_{0,1}}[uR_{0,1} + \mu L]$$

In this case, again, vaccination will limit the risk of evolutionary emergence through a reduction of the epidemic size of the maladapted strain. More complex scenarios can be studied with this approach to look, for instance, at the impact of different types of heterogeneities on evolutionary emergence, as in Yates *et al.* [5].

Appendix C. Importance of the mean and variance of the distribution of offspring

The diffusion approximation [30,31] is an alternative way to obtain the probability of escaping extinction when n individuals are initially present (see also the paper of Martin *et al.* [32]):

$$P \approx 1 - e^{-2nr/v},$$

where r and v are the mean and the variance of the offspring number, respectively. This expression is an approximation but holds under a broad range of scenarios. It formalizes the idea that the extinction is sensitive to the whole offspring distribution and in particular the mean and the variance of the growth rate of the population. Population persistence is increased by increases in the mean and decreases in the variance. In many situations, these two quantities covary, and in particular in the simple epidemiological model we study here.

Let us assume a classical birth–death model with *per capita* parameters b (birth) and d (death). In a short interval dt, three things can happen to an individual

— giving birth (+1 individual) with probability $b\,dt$,
— death (−1) with probability $b\,dt$, and
— nothing (0) with probability $1 - b\,dt - d\,dt$.

The expected change in population size in a small interval of time dt owing to a focal individual is thus equal to $r dt = b dt(+1) + b dt(-1) + (1 - b dt - d dt)(0) = (b - d)dt$. The variance in the change in population in a small interval of time dt due to a focal individual is thus equal to: $v\,dt = b\,dt(+1 - r)^2 + d\,dt(-1 - r)^2 + (1 - b\,dt - d\,dt)(0 - r)^2$. After neglecting the higher-order terms in dt (i.e. dt^2, dt^3) we obtain: $v\,dt = (b + d)dt$. The ratio of the mean to the variance in growth rate is $r/v = (b - d)/(b + d) = (R_0 - 1)/(R_0 + 1)$, where $R_0 = b/d$ is the number of births over the average lifespan $1/d$. The diffusion approximation given above (with $n = 1$) thus yields

$$P \approx 1 - e^{-2(R_0-1)/(R_0+1)}.$$

This expression is indeed a good approximation (for R_0 not too high [31]) of the exact probability of escaping extinction given in the main text (i.e. $P = 1 - 1/R_0$). The point we want to make here is that the probability of escaping extinction in our simple epidemiological model is governed by a single parameter, R_0. Maximizing the basic reproduction ratio always strikes a balance between increasing r and decreasing v. This may help us understand the seemingly counterintuitive result that, although the Malthusian growth rate, r, does provide a relevant measure of the competitivity of a strain, it is its basic reproduction ratio that governs the probability of escaping early extinction.

References

1. Gonzalez A, Ronce O, Ferriere R, Hochberg ME. 2012 Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Phil. Trans. R. Soc. B* **368**, 20120404. (doi:10.1098/rstb.2012.0404)
2. Anderson RM, May RM. 1991 *Infectious diseases of humans: dynamics and control*. Oxford, UK: Oxford University Press.
3. Diekmann O, Heesterbeek JAP. 2000 *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Chichester, UK: Wiley.
4. May RM, Gupta S, McLean AR. 2001 Infectious disease dynamics: what characterizes a successful invader? *Phil. Trans. R. Soc. Lond. B* **356**, 901–910. (doi:10.1098/rstb.2001.0866)
5. Yates A, Antia R, Regoes RR. 2006 How do pathogen evolution and host heterogeneity interact in disease emergence? *Proc. R. Soc. B* **273**, 3075–3083. (doi:10.1098/rspb.2006.3681)
6. André JB, Day T. 2005 The effect of disease life history on the evolutionary emergence of novel

pathogens. *Proc. R. Soc. B* **272**, 1949–1956. (doi:10.1098/rspb.2005.3170)

7. Antia R, Regoes RR, Koella JC, Bergstrom CT. 2003 The role of evolution in the emergence of infectious diseases. *Nature* **426**, 658–661. (doi:10.1038/nature02104)
8. Iwasa Y, Michor F, Nowak M. 2003 Evolutionary dynamics of escape from biomedical intervention. *Proc. R. Soc. Lond. B* **270**, 2573–2578. (doi:10.1098/rspb.2003.2539)
9. Lenormand T. 2002 Gene flow and the limits to natural selection. *Trends Ecol. Evol.* **17**, 183–189. (doi:10.1016/S0169-5347(02)02497-7)
10. Holt RD, Gomulkiewicz R. 1997 How does immigration influence local adaptation? A re-examination of a familiar paradigm. *Am. Nat.* **149**, 563–572. (doi:10.1086/286005)
11. Gomulkiewicz R, Holt RD, Barfield M. 1999 The effects of density dependence and immigration on adaptation and niche evolution in a black-hole sink environment. *Theor. Popul. Biol.* **55**, 283–296. (doi:10.1006/tpbi.1998.1405)
12. Holt RD, Barfield M, Gomulkiewicz R. 2005 Theories of niche conservatism and evolution: could exotic species be potential tests? In *Species invasions: insights into ecology, evolution, and biogeography* (eds D Sax, J Stachowicz, SD Gaines), pp. 259–290. Sunderland, MA: Sinauer Associates.
13. Reluga T, Meza R, Walton DB, Galvani AP. 2007 Reservoir interactions and disease emergence. *Theor. Popul. Biol.* **72**, 400–408. (doi:10.1016/j.tpb.2007.07.001)
14. Haccou P, Iwasa Y. 1996 Establishment probability in fluctuating environments: a branching process model. *Theor. Popul. Biol.* **50**, 254–280. (doi:10.1006/tpbi.1996.0031)
15. Schreiber SJ, Lloyd-Smith JO. 2009 Invasion dynamics in spatially heterogeneous environments. *Am. Nat.* **174**, 490–505. (doi:10.1086/605405)
16. Holt RD, Knight TM, Barfield M. 2004 Allee effects, immigration, and the evolution of species' niches. *Am. Nat.* **163**, 253–262. (doi:10.1086/381408)
17. Day T, Andre J-B, Park A. 2006 The evolutionary emergence of pandemic influenza. *Proc. R. Soc. B* **273**, 2945–2953. (doi:10.1098/rspb.2006.3638)
18. Holt RD, Gomulkiewicz R, Barfield M. 2004 Temporal variation can facilitate niche evolution in harsh sink environments. *Am. Nat.* **164**, 187–200. (doi:10.1086/422343)
19. Ronce O, Kirkpatrick M. 2001 When sources become sinks: migrational meltdown in heterogeneous habitats. *Evolution* **55**, 1520–1531. (doi:10.1554/0014-3820(2001)055[1520:WSBSMM]2.0.CO;2)
20. Levin BR, Bull JJ. 1994 Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* **2**, 76–81. (doi:10.1016/0966-842X(94)90538-X)
21. Holt RD, Barfield M. 2006 Within-host pathogen dynamics: some ecological and evolutionary consequence of transients, dispersal mode, and within-host spatial heterogeneity. In *Disease evolution: models, concepts, and data analyses* (eds Z Feng, U Dieckmann, S Levin), pp. 45–66. Providence, RI: American Mathematical Society.
22. Ganusov VV, Bergstrom CT, Antia R. 2002 Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. *Evolution* **56**, 213–223.
23. Mideo N, Alizon S, Day T. 2008 Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol. Evol.* **23**, 511–517. (doi:10.1016/j.tree.2008.05.009)
24. Woolhouse ME *et al*. 1997 Heterogeneities in the transmission of infectious agents: implications for the design of control programs. *Proc. Natl Acad. Sci. USA* **94**, 338–342. (doi:10.1073/pnas.94.1.338)
25. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. 2005 Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359. (doi:10.1038/nature04153)
26. Alexander HK, Day T. 2010 Risk factors for the evolutionary emergence of pathogens. *J. R. Soc. Interface* **7**, 1455–1474. (doi:10.1098/rsif.2010.0123)
27. Orr HA, Unckless RL. 2008 Population extinction and the genetics of adaptation. *Am. Nat.* **172**, 160–169. (doi:10.1086/589460)
28. Gandon S, Day T. 2007 The evolutionary epidemiology of vaccination. *J. R. Soc. Interface* **4**, 803–817. (doi:10.1098/rsif.2006.0207)
29. Gandon S, Day T. 2009 Evolutionary epidemiology and the dynamics of adaptation. *Evolution* **6**, 826–838. (doi:10.1111/j.1558-5646.2009.00609.x)
30. Feller W. 1951 Diffusion processes in genetics. In *Proc. of the Second Berkeley Symp. on Mathematical Statistics and Probability*, pp. 227–246. Berkeley, CA: University of California Press.
31. Lambert A. 2006 Probability of fixation under weak selection: a branching process unifying approach. *Theor. Popul. Biol.* **69**, 419–441. (doi:10.1016/j.tpb.2006.01.002)
32. Martin G, Aguilée R, Ramsayer J, Kaltz O, Ronce O. 2012 The probability of evolutionary rescue: towards a quantitative comparison between theory and evolution experiments. *Phil. Trans. R. Soc. B* **368**, 20120088. (doi:10.1098/rstb.2012.0088)
33. Holt RD, Hochberg ME. 2002 Virulence on the edge: a source–sink perspective. In *Adaptive dynamics of infectious diseases: in pursuit of virulence management* (eds U Dieckmann, JAJ Metz, MW Sabelis, K Sigmund), pp. 104–120. Cambridge, UK: Cambridge University Press.
34. Alexander HK, Wahl LM. 2008 Fixation probabilities depend on life history: fecundity, generation time and survival in a burst-death model. *Evolution* **62**, 1600–1609. (doi:10.1111/j.1558-5646.2008.00396.x)
35. Wallinga J, Lipsitch M. 2007 How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc. R. Soc. B* **274**, 599–604. (doi:10.1098/rspb.2006.3754)
36. Holmes EC. 2009 *The evolution and emergence of RNA viruses*. Oxford Series in Ecology and Evolution (OSEE). Oxford, UK: Oxford University Press.
37. Bull JJ, Sanjuán R, Wilke CO. 2007 Theory of lethal mutagenesis for viruses. *J. Virol.* **81**, 2930–2939. (doi:10.1128/JVI.01624-06)
38. Martin G, Gandon S. 2010 Lethal mutagenesis and evolutionary epidemiology. *Phil. Trans. R. Soc. B* **365**, 1953–1963. (doi:10.1098/rstb.2010.0058)
39. Davies TJ, Pedersen AB. 2008 Phylogeny and geography predict pathogen community similarity in wild primates and humans. *Proc. R. Soc. B* **275**, 1695–1701. (doi:10.1098/rspb.2008.0284)
40. Streicker DG, Turmelle AS, Vonhof MJ, Kuzmin IV, McCracken GF, Rupprecht CE. 2010 Host phylogeny constrains cross-species emergence and establishment of rabies virus in bats. *Science* **329**, 676–679. (doi:10.1126/science.1188836)
41. de Lamballerie X, Leroy E, Charrel RN, Ttsetsarkin K, Higgs S, Gould EA. 2008 Chikungunya virus adapts to tiger mosquito via evolutionary convergence: a sign of things to come? *Virol. J.* **5**, 33. (doi:10.1186/1743-422X-5-33)
42. Schuffenecker I *et al*. 2006 Genome microevolution of Chikungunya viruses causing the Indian Ocean outbreak. *PLoS Med.* **3**, e263. (doi:10.1371/journal.pmed.0030263)
43. Taylor LH, Latham SM, Woolhouse MEJ. 2001 Risk factors for human disease emergence. *Phil. Trans. R. Soc. Lond. B* **356**, 983–989. (doi:10.1098/rstb.2001.0888)
44. Sanjuán R. 2010 Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. *Phil. Trans. R. Soc. B* **365**, 1975–1982. (doi:10.1098/rstb.2010.0063)
45. Lalić J, Cuevas JM, Elena SF. 2011 Effect of host species on the distribution of 549 mutational fitness effects for an RNA virus. *PLoS Genet.* **7**, e1002378. (doi:10.1371/journal.pgen.1002378)
46. Dennehy JJ. 2009 Bacteriophages as model organisms for virus emergence research. *Trends Microbiol.* **17**, 450–457. (doi:10.1016/j.tim.2009.07.006)
47. Dennehy JJ, Friedenberg NA, McBride RC, Holt RD, Turner PE. 2010 Experimental evidence that source genetic variation drives pathogen emergence. *Proc. R. Soc. B* **277**, 3113–3121. (doi:10.1098/rspb.2010.0342)
48. André JB, Hochberg ME. 2005 Virulence evolution in emerging infectious diseases. *Evolution* **59**, 1406–1412. (doi:10.1554/05-111)
49. Bonhoeffer S, Lenski RE, Ebert D. 1996 The curse of the pharaoh: the evolution of virulence in pathogens with long living propagules. *Proc. R. Soc. Lond. B* **263**, 715–721. (doi:10.1098/rspb.1996.0107)
50. Day T. 2003 Virulence evolution and the timing of disease life-history events. *Trends Ecol. Evol.* **18**, 113–118. (doi:10.1016/S0169-5347(02)00049-6)
51. Day T, Proulx SR. 2004 A general theory for the evolutionary dynamics of virulence. *Am. Nat.* **163**, E40–E63. (doi:10.1086/382548)
52. Day T, Gandon S. 2006 Insights from Price's equation into evolutionary epidemiology. In *Disease evolution: models, concepts and data analyses* (eds Z Feng, U Dieckmann, S Levin), pp. 23–43. Providence, RI: American Mathematical Society.
53. Day T, Gandon S. 2007 Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* **10**, 876–888. (doi:10.1111/j.1461)
54. Bull JJ, Ebert D. 2008 Invasion thresholds and the evolution of non-equilibrium virulence. *Evol. Appl.* **1**, 172–182. (doi:10.1111/j.1752-4571.2007.00003.x)