PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Within-Host Stochastic Emergence Dynamics of Immune-Escape Mutants

Matthew Hartfield*¤, Samuel Alizon

Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2), 911 avenue Agropolis, Montpellier, France

¤ Current address: Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada.
* matt.hartfield@gmail.com

OPEN ACCESS

Citation: Hartfield M, Alizon S (2015) Within-Host Stochastic Emergence Dynamics of Immune-Escape Mutants. PLoS Comput Biol 11(3): e1004149. doi:10.1371/journal.pcbi.1004149

Editor: Becca Asquith, Imperial College London, UNITED KINGDOM

Received: August 18, 2014

Accepted: January 22, 2015

Published: March 18, 2015

Data Availability Statement: Supplementary Mathematica files are included as Supporting Information files. Simulation code and results are available from the Dryad data depository (doi: 10.5061/dryad.df1vk).

Funding: MH was funded by an ATIP-Avenir grant from CNRS and INSERM to SA; SA and MH acknowledge additional support from the CNRS and the IRD. The funders 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.

Abstract

Predicting the emergence of new pathogenic strains is a key goal of evolutionary epidemiology. However, the majority of existing studies have focussed on emergence at the population level, and not within a host. In particular, the coexistence of pre-existing and mutated strains triggers a heightened immune response due to the larger total pathogen population; this feedback can smother mutated strains before they reach an ample size and establish. Here, we extend previous work for measuring emergence probabilities in non-equilibrium populations, to within-host models of acute infections. We create a mathematical model to investigate the emergence probability of a fitter strain if it mutates from a self-limiting strain that is guaranteed to go extinct in the long-term. We show that ongoing immune cell proliferation during the initial stages of infection causes a drastic reduction in the probability of emergence of mutated strains; we further outline how this effect can be accurately measured. Further analysis of the model shows that, in the short-term, mutant strains that enlarge their replication rate due to evolving an increased growth rate are more favoured than strains that suffer a lower immune-mediated death rate ('immune tolerance'), as the latter does not completely evade ongoing immune proliferation due to inter-parasitic competition. We end by discussing the model in relation to within-host evolution of human pathogens (including HIV, hepatitis C virus, and cancer), and how ongoing immune growth can affect their evolutionary dynamics.

Author Summary

The ongoing evolution of infectious diseases provides a constant health threat. This evolution can either result in the production of new pathogens, or new strains of existing pathogens that escape prevailing drug treatments or immune responses. The latter process, also known as immune escape, is a predominant reason for the persistence of several viruses, including HIV and hepatitis C virus (HCV), in their human host. As a consequence, the within-host emergence of new strains has been the intense focus of modelling studies. However, existing models have neglected important feedbacks that affects this emergence probability. Specifically, once a mutated pathogen arises that spreads more quickly than

the initial (resident) strain, it potentially triggers a heightened immune response that can eliminate the mutated strain before it spreads. Our study outlines novel mathematical modelling techniques that accurately quantify how ongoing immune growth reduces the emergence probability of mutated pathogenic strains over the course of an infection. Analysis of this model suggests that, in order to enlarge its emergence probability, it is evolutionary beneficial for a mutated strain to increase its growth rate rather than tolerate immunity by having a lower immune-mediated death-rate. Our model can be readily applied to existing within-host data, as demonstrated with application to HIV, HCV, and cancer dynamics.

Introduction

Parasites and pathogens pose a continuous threat to human, livestock, and plant health since new strains can readily emerge, via mutation or recombination, from pre-existing strains. Generally, the focus has been on detection of emerging diseases at the population level, in order to track and control their spread [1, 2]. Modelling approaches to predicting emergence have therefore primarily concentrated on detecting infections arising between individual hosts [3, 4], and the contribution of within-host processes to pathogen emergence has often been overlooked. It is now well known that within-host evolution has strong effects on the epidemiology of many pathogens (reviewed in [5]), and can substantially affect the course of an infection, as illustrated by the cases of HIV [6] and hepatits C virus (HCV) [7].

Each step of the within-host evolutionary dynamics consists of the emergence of a rare mutant that takes over the pathogen population. Since mutated infections always initially appear as a few copies, they are prone to extinction so their emergence is best captured using stochastic dynamics (as opposed to deterministic approaches). Only a handful of previous models have investigated this stochastic within-host process. A widespread use of within-host models in static populations is to calculate the probability that infections will evolve drug resistance, and what regime is needed to avoid treatment failure [8–11]. One exception [12] studied viral emergence within a host, if most mutations were deleterious, in order to determine how different viral replication mechanisms affected the establishment of beneficial alleles.

Parasite and pathogen evolution can radically affect the course of infections in hosts able to mount immune responses. For instance, HIV is known to successfully fix mutations within individuals which enable it to evade immune pressures [6]. This is in line with evidence that target cell limitation cannot account for HIV dynamics, and that immune limitation also needs to be present [13]. Arguably, this continual evasion of immunity prompts the chronic nature of HIV infections (see [14, 15] for an illustration of these 'Red Queen' dynamics). In the case of HCV, it has also been found that chronic infections were associated with higher rates of within-host evolution [16]. Concerning a different chronic disease, it is now well-known that the ability of malignant cells to rapidly expand in size, ignoring biological signals to arrest growth (e.g. through mutating the *p*53 tumour suppression gene), and escaping the patient's immune system are key steps in cancer development [17]. All these scenarios can be analysed in the larger framework of evolutionary rescue [18, 19], where a change in the environment (in this case, the activation of an immune response) will cause the population to go extinct unless it evolves (develops an increased replicative ability, then subsequently rise to a large enough size to avoid stochastic loss). Although chronic infections are those most likely to evolve strains that evade immune pressures, evidence exists that such immune escape also plays a role in acute infections, like influenza [20]. Note that the latter evidence arises from serial passage

experiments; although these tend to maximise selection at the within-host level, they still include a slight selective pressure for increased transmission.

Modelling mutant emergence in an immune evasion context raises a technical challenge, as it is a non-equilibrium process. In most of the within-host models presented above [8–10, 12], it is sufficient to know the initial state of the system to calculate emergence probability of immune or treatment escape. A non-equilibrium model was studied by Alexander and Bonhoeffer [11], who accounted for the reduction in available target cells when determining the emergence of drug resistance using an evolutionary rescue framework. With immune-escape however, the problem is that the immune state when the mutant infection appears is only temporary. In this case, there will be an initial strain present within the host, which triggers immunity. This strain can then mutate into a faster-replicating form, but if the mutated strain arises at a low frequency, immunity can destroy it before it has a chance to spread. This effect can be exacerbated by the fact that the mutated strain also prompts an increased immune response, so the emerging infection has a stronger defence to initially compete with (assuming immune growth is proportional to the total size of the pathogen population). This feedback, where increased immune growth prevents emergence of mutated strains with higher replication rates, can strongly affect the appearance of mutated strains within-host, and needs to be accounted for. Existing emergence models have not yet accounted for such within-host population feedbacks, especially those arising from the immune system.

Recently, Hartfield and Alizon [21] tackled a related problem, regarding how epidemiological feedbacks affect disease emergence at the host population level. In their model, a faster-replicating strain emerged via mutation from a pre-existing infection; however, the continuing outbreak caused by the initial strain removed susceptible individuals from the population, which limited the initial spread of the mutated strain. It was shown that the ongoing depletion of susceptible individuals due to the initial strain spreading has a stronger effect on reducing pathogen emergence, than assumed by just scaling down the reproductive ratio by the frequency of susceptible individuals present when the mutated infection appears. That is, the feedback produced by the first strain in removing susceptible individuals caused a drastic decrease in the emergence probability of the mutated strain.

Building on this previous study, we derive here an analytical approximation for the probability that an immune-escape mutant will emerge and maintain itself within a host. We use 'immune-escape' in the sense that while the mutated strain can be killed by immunity, it can ultimately outgrow immune growth and chronically persist. Furthermore, we use an acute infection setting, which are commonly used to study the within-host dynamics of 'flu-like' diseases [22–28]. Analysis of the model demonstrates how the ongoing proliferation of immune cells acts to decrease the emergence probability of mutated strains. Acute infection models are also useful in studying the first stages of chronic infections, where one observes exponential growth of the virus population, followed by a decline in the first weeks of infection. In addition, there are two questions that are worth investigating with this model from a biological standpoint. First, what is the fittest evolutionary strategy for an escape mutant: is it to overgrow the immune response (that is, increase its inherent replication rate to enable its persistence, even when immune cells are at capacity), or to tolerate it (prevent immunity from killing as many pathogens per immune cell)? Both these actions will lead to an increase in the pathogen's net reproduction rate and can thus be described as immune evasion, but it is unclear if one process is favoured over the other. In addition, note that disabling the immune response is not possible, since the wild-type infection activates it. Second, to what extent do we need to account for the ongoing proliferation of the immune response? In other words, if we calculate the emergence probability based on the system state when the mutation occurs, how inaccurate would this

estimate be? We end by discussing our results in the light of what is known about within-host evolution for several human infections.

Materials and Methods

Model outline

In order to find an analytical solution for the within-host emergence of a mutated strain, we follow the approach of Hartfield and Alizon [21], which investigated the appearance of a new infectious pathogen from a pre-existing strain at the population level. To consider the within-host case, we construct and subsequently analyse a specific scenario of acute, immunising infections. Here, the first strain will go extinct because it does not replicate at a high enough rate. However, before vanishing, it can mutate into a form that grows unboundedly; we are interested in calculating the probability of this event occurring. We focus on this case to cover a general range of within-host evolution scenarios, which is important since there is yet no consensus on how to best model within-host infections (reviewed in [29]). Although it is feasible that the mutated strain has a maximum population size, mathematical analysis of pathogen emergence only needs to consider the dynamics of mutants when they are present in a few copies, rather than the long-term behaviour once they have already established. Therefore, the general results outlined in this paper are also broadly applicable to cases where the infection population sizes are bounded.

Our analytical approach involves using a set of deterministic differential equations to ascertain pathogen spread in a stochastic birth-death process, where an infection (or immune cell) can only either die or produce 1 offspring. This process is one of the most common ways of investigating stochastic disease spread [30]. A list of nomenclature used in the model is outlined in Table 1. Assume there exists an initial pathogen, or infected cell-line, the size of which at time t is denoted $x_1(t)$. This line grows in size over time according to the following equation, which is well-used for within-host infection models [29]:

$$\frac{dx_1}{dt} = x_1(\varphi_1 - \sigma_1 y) \tag{1}$$

Table 1. Glossary of notation.

Symbol	Usage
t	Time (measured as number of generation since start of process)
φ_1, φ_2	Growth rate of initial, mutated infection
σ_1, σ_2	Immune-mediated death-rate of initial, mutated infection
x_1, x_2	Size of initial, mutated infection
x_{init}	Size of x_1 at time $t = 0$ (or $y = y_{init}$)
y	Size of immune response
y_{init}	Initial size of immune response at $t = 0$
K	Maximum size of immune response
r	Unscaled growth rate of immune response
R_1, R_2	Reproductive ratio of original or mutated infection, φ_i/σ_i
R^*	'Effective' initial reproductive ratio in the presence of immunity, $R - y_0$
ρ	Scaled immune growth rate, r/σ_1
μ	Individual mutation rate from first to second infection
Π	Emergence probability of second strain in an non-equilibrium population
P	Overall emergence probability of second strain
P^*_{ext}	Extinction probability, given an 'effective' reproductive ratio

doi:10.1371/journal.pcbi.1004149.t001

Here, φ_1 is the growth rate of the infection, σ_1 is the rate of destruction of the pathogen per immune cell, and $y(t)$ is the number of immune cells (i.e. lymphocytes). For simplicity, we assume that there is complete cross-immunity between the various pathogen strains, so it is not necessary to model immune cell diversity.

The growth of the immune cell population is modelled using a logistic-growth curve:

$$\frac{dy}{dt} = rx_1 y\left(1 - \frac{y}{K}\right) \tag{2}$$

Here, r is the proliferation rate of immune cells, and K is the maximum population size they can achieve. This formulation is an extension of the model developed by Gilchrist and Sasaki [24] to study acute infections. The main difference in that in their model, the density of immune cells is allowed to reach any value in order to clear the infection. Here, we impose that immune density does not go above a maximum threshold K, which correspond to an intrinsic limitation in the host resources allocated to immunity.

Profile plots of typical infection responses are shown in Section 1 of S1 Text. If $\varphi_1/\sigma_1 < K$, then the first strain will increase in size until the immune-cells reach a maximum. After this point, the infection will decrease towards eventual extinction, while the immune response will be maintained at a non-zero size. However, if $\varphi_1/\sigma_1 \geq K$ then the first strain will continue to expand. A formal rational for this behaviour will be shown below.

To proceed with finding an analytical solution for the emergence probability, we proceed as in previous analyses [21, 24, 31], and note that since y is monotonically increasing, we can use the immune cell population size as a surrogate measure of time. By dividing Equation 1 by Equation 2, we obtain a differential equation for x_1 as a function of y:

$$\frac{dx_1}{dy} = \frac{(\varphi_1 - \sigma_1 y)x_1}{rx_1 y\left(1 - \frac{y}{K}\right)} \tag{3}$$

To simplify subsequent analyses, we make the following substitutions. We define the reproductive rate of the infection, where there is a single immune cell ($y = 1$) equals $R_1 = \varphi_1/\sigma_1$. We also set $\rho = r/\sigma_1$ (this can be formally shown by rescaling time by $\tau = \sigma_1 t$). We use the notation R_1 to draw parallels between the scaled pathogenic replication rate, and the reproductive ratio R_0 in population-level, epidemiological models [32]. After making the required substitutions, Equation 3 can be rewritten as:

$$\frac{dx_1}{dy} = \frac{K(R_1 - y)}{\rho y(K - y)} \tag{4}$$

Equation 4 formally shows that the infected cell line increases in size ($dx_1/dy > 0$) if $y < R_1 = \varphi_1/\sigma_1$, and decreases if $y > R_1$. A corollary of this result is that if R_1 of a infection exceeds K, then it cannot go extinct in the long term.

This differential equation is straightforward to solve (Section 1 of S1 Text), and yields the following function for $x_1(y)$:

$$x_1(y) = x_{init} + \frac{1}{\rho}\left(\log\left[\left(\frac{y}{y_{init}}\right)^{R_1}\left(\frac{K - y}{K - y_{init}}\right)^{(K - R_1)}\right]\right) \tag{5}$$

where x_{init} and y_{init} are, respectively, the number of pathogen and immune cells at the start of the process (time $t = 0$), and log is the natural logarithm. It is clear from Equation 4 that the maximum value of x_1 occurs for $y = R_1$. By substituting this value into Equation 5, we obtain

the maximum value of x_1 (denoted x_M) as:

$$x_M = x_{init} + \frac{1}{\rho}\left(\log\left[\left(\frac{R_1}{y_{init}}\right)^{R_1}\left(\frac{K-R_1}{K-y_{init}}\right)^{(K-R_1)}\right]\right) \quad (6)$$

Note that ρ has no effect on the position of the peak (that is, the value of y leading to the maximal infection level). Since it affects the growth rate of the immune response (and therefore also of the pathogen), it does determine the maximum value itself. As this maximum is inversely proportional to ρ, smaller immune growth rates lead to larger peaks. Finally, the maximum infection time (as a function of y) needed for the first infection to go extinct can be determined by solving $x_1(y) = 0$ numerically (Section 1 of S1 Text).

Formulating emergence probability

Our goal is to calculate the probability of 'evolutionary rescue'. That is, the initial strain has $R_1 < K$ so is guaranteed to go extinct in the long-term. However, a mutated form (with reproductive ratio R_2) could arise with $R_2 > K$, and if it does not go extinct when rare, it can outgrow immune proliferation. Previous theory on the emergence of novel pathogenic strains [21] showed that if mutated strains arise at rate μ per time step, the overall emergence probability P is given by:

$$P = 1 - \exp\left(-\mu \int_{y_{init}}^{y_M} x_1(y)\Pi(y)\mathrm{d}y\right) \quad (7)$$

for y_M the maximum immune size for which it is possible for immune escape to arise, and $\Pi(y)$ the emergence probability of an escape mutation were it to appear. The formulation of each of these will be discussed in turn.

Emergence probability, $\Pi(y)$. The emergence probability of the new strain has to account for both the increase in immune response before the mutation occurs, which reduces the 'effective' growth-rate of a mutated strain when it appears due to an heightened immune size (greater y in the model); and the continuing growth of immunity once it has appeared. Both these points factor in how the overall immune strength affects emergence probabilities, but this strength can increase over time. In theory, this emergence probability can be calculated from first principles using a time-inhomogeneous branching-process equation, which accounts for the birth and death of pathogens as the lymphocyte population size is changing (see [4, 21] for examples of branching-process formulations in epidemiology). However, due to the complexity of the model, specifically the non-linear form of the change in immune response (Equation 2), it is not possible to derive an exact analytical solution for Π (Section 1 of S1 Text).

We proceed by comparing a previously-found solution to Π in a changing susceptible population to our model, and adapting the result accordingly. When investigating a second strain emerging from a pre-existing epidemic, Hartfield and Alizon [21] found that a good approximation for Π is a function of the form:

$$\Pi = \left(\frac{R^*}{\rho(x_1+x_2)y(K-y)+R^*}\right)\left(1-P^*_{ext}\right) \quad (8)$$

where R^* is the 'effective' reproductive ratio of the mutated strain at the start of the process, which is dependent on the size of a secondary resource (e.g. susceptible individuals, immune cells); and P^*_{ext} is the extinction probability in a birth-death model, given the 'effective' reproductive ratio is reduced depending on the frequency of the same secondary resource. Equation 8 arose during analysis of pathogen emergence in an epidemiological SIR framework [21], by

noting how for a single strain, the ratio of the rate of change of infected and susceptible individuals over time reflected the emergence probability. This logic was then carried over to a two-strain scenario, so it details how the ongoing spread of the first pathogen affected emergence of mutated strains. To give an example, when applied to a model of pathogen emergence in an SIR setting, the effective reproductive ratio R^* would equal the standard reproductive ratio R_0, reduced by a factor S_0/N, if there initially existed S_0 susceptible individuals out of a total population of size N. There would also be a term in the denominator that is proportional to the rate of spread of the initial, unmutated strain.

However, in the current model, mutated pathogen lines have a reduced emergence probability due to ongoing production of immunity triggered by the first strain. The rational behind Equation 8 is that the $\rho(x_1 + x_2)y(K - y)$ accounts for the reduction in emergence probability due to the continuing onset of immunity, given a total infection size $(x_1 + x_2)$. If it reaches a steady-state [that is, $\rho(x_1 + x_2)y(K - y) \to 0$] then Π collapses mathematically to $(1 - P^*_{ext})$. Hence Equation 8 accounts for how the presence of infection triggers immune responses that restrict the emergence probability of mutated infections; this is the feedback not yet considered in previous within-host evolution models.

We use Equation 8 in our model by setting $R^* = R_2 - y_{init}$, which is the rescaled growth rate of the mutated strain, corrected for the fact that the baseline immunity rate will reduce its initial selective advantage. Note we use y_{init} instead of $y(t)$ in this argument, since the R^* term should describe the baseline replication rate if the second strain is only present. For the emergence probability, we first note that as with Equations 3 and 4, a single mutated cell has an effective birth rate of R_2 and death rate of y. Standard results from birth-death models states that the mean growth rate is equal to $R_2 - y$, with variance equal to $R_2 + y$ [33]. We therefore determine the baseline emergence probability by substituting these terms into the classic result by Feller [34]:

$$(1 - P^*_{ext}) = 1 - \exp\left(-2\frac{(R_2 - y)}{(R_2 + y)}\right) \tag{9}$$

Note that Equation 9 is not constant over time since it is a function of y. Overall, we obtain the following approximation for Π:

$$\Pi(y) = \left(\frac{R_2 - y_{init}}{\rho(x_1(y) + 1)y(K - y) + R_2 - y_{init}}\right)\left(1 - \exp\left(-2\frac{(R_2 - y)}{(R_2 + y)}\right)\right) \tag{10}$$

Also note that we use $x_1(y)$+1 in the denominator, since the total pathogen population is accounted for by the initial pathogen size when the mutant appears, $x_1(y)$, plus a single mutated individual.

Maximum value of y, y_M. The emergence of mutated infections can be arrested in one of two ways. Either the immune response becomes high enough so as to drive the first infection to extinction ($x_1(y) = 0$), or so that any further emergence is impossible ($\Pi(y) = 0$). Therefore, y_M can be easily found by finding the smallest solution out of $x_1(y) = 0$ or $\Pi(y) = 0$. These equations have to be solved numerically (Section 1 of S1 Text).

Example plots of x_1 and Π. Fig. 1 give example behaviour of both x_1 and Π, as determined by Equations 5 and 10 respectively. x_1 exhibits an inverted-U shape, with the maximum size decreasing if ρ increases, as expected from the form of Equation 5. From a biological standpoint, it illustrates the fact that the parasite population grows until the immune response is activated, at which point it decreases. Π, on the other hand, demonstrates a U shape. Emergence is initially high due to fewer immune cells being present, and again as the first strain dies out, since the immune population is not proliferating as fast and is less likely to remove mutated

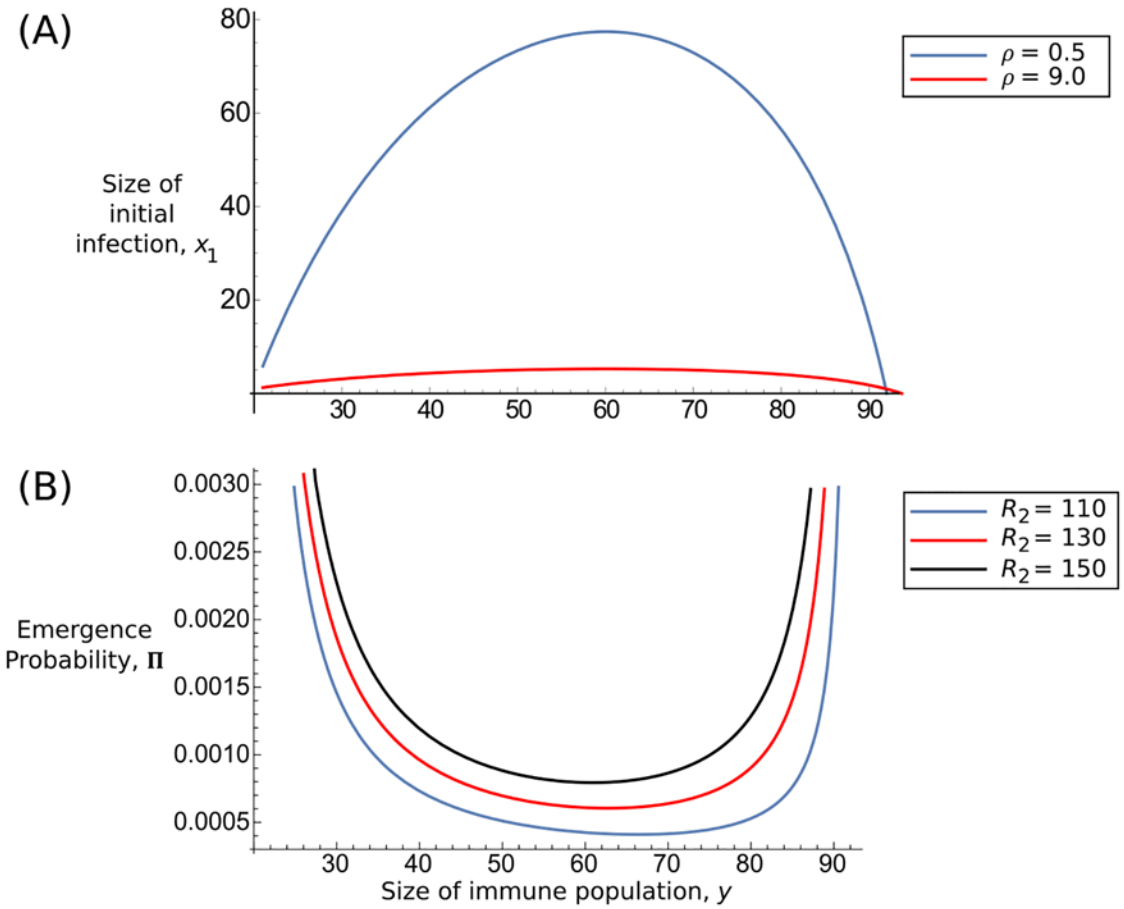

Fig 1. Sample behaviour of x_1 and Π. Example plots of Equations 5 (A) and 10 (B), as a function of the immune-cell population y. Different line colours reflect either different values of ρ in (A), or R_2 in (B). Other parameters are $K = 100$, $R_1 = 60$, $x_{init} = 1$, $y_{init} = 20$.

doi:10.1371/journal.pcbi.1004149.g001

strains as they appear. Results are qualitatively similar for different parameter values, with Π further decreasing with higher values of ρ (Section 1 of S1 Text).

Simulation methods

We verified our analytical solution by comparing it to simulation data. Simulations were written in C and based on the Gillespie Algorithm with tau-leaping [35, 36]; source code has been deposited online in the Dryad data depository (doi:10.5061/dryad.df1vk). The time step was set to be very low: $\Delta\tau = 0.00005$. This is because the tau-leaping algorithm is accurate only if the expected number of events per time step is small [37]; since the growth rates of the pathogen strains and the lymphocytes are both large, a small time step is needed to make the simulation valid.

The growth of both the original and mutated strains are simulated using scaled parameter rates. That is, the birth rate per time step for each pathogen is Poisson-distributed with mean $(R_i \cdot x_i \cdot \Delta\tau)$, and death rate with mean $(x_i \cdot y \cdot \Delta\tau)$ for $i = 1, 2$. This is done to reduce the number of parameters in the model, and also enables accurate comparison with the scaled results.

The change in size of the immune response is determined by standard logistic-growth dynamics for the Gillespie algorithm. That is, the Poisson mean number of births per time step equaled $\rho_b(x_1 + x_2)y$, for ρ_b is the birth rate parameter, and the mean number of deaths equalled

$y(x_1 + x_2)(\rho_d + \rho(y/K))$, where ρ_d is the death rate and $\rho = \rho_b - \rho_d$. Since there are no distinct ρ_b and ρ_d terms in the model (just ρ), then we set $\rho_d = 1$ and varied ρ_b, so $\rho = \rho_b - 1$. The net growth rate ρ was varied, generally between 0.5 and 19 depending on the size of the other parameters.

Note that in the analytical solution, it is assumed that the immune response does not die off. In order to maintain this assumption, we set $y_{init} = 20$ in the simulations. This also makes intuitive sense, because it is unlikely that the initial immune response is limited to just one cell. If the immune response goes extinct before both infected strains go extinct, or the mutated strain emerges, then that run is discarded, the simulation is reset and restarted.

In simulations, R_1 is either set to 60 if $K = 100$, $R_1 = 100$ if $K = 250$ or 1000, or $R_1 = 2{,}000$ for $K = 10{,}000$. R_2 was varied, ensuring that $R_2 > K$ so the infected strain can outgrow the immune response (see Equation 4). The mutation rate was also varied over several orders of magnitude. The first strain is reintroduced (from an initial frequency of 1 cell) 10,000,000 times as separate replication runs. Since the emergence probabilities were predicted to be low, a large number of runs were needed in order to produced a meaningful estimate of emergence probability. The second strain is said to have emerged once it exceeded 20 cells; since meaningful values of R_2 were large, only a handful of cells would have been needed to guarantee emergence, hence a low outbreak threshold could be used [38].

Results

Comparison of analytical results with simulations data

The first results we checked were individual profiles of the stochastic simulations, since these can be used to demonstrate the behaviour of within-host emergence. In the majority of cases, the first strain increase in size, but once immune proliferation also reaches its maximum then the first strain goes extinct soon after (Fig. 2(a)). However, emergence of the mutated strain can occur even if the immune cells are at their maximum size (Fig. 2(b)). This is to be expected, given the form of Π (Equation 10), which demonstrates that emergence is likeliest once the immune cells reach a steady-state, so ongoing proliferation does not restrict their establishment. Conversely, emergence probability is reduced if the immune response is spreading; this is because when the mutated infection is rare, it is less likely to reproduce as immune cells increase in number. Since the mutated infection population is only present in a few copies, this negative effect this immune growth has on reproductive ability will be drastic (see also equivalent population genetics results by Otto and Whitlock [39]).

Fig. 3 compares the full analytical solution (Equation 7, with Π given by Equation 10) against simulation data. If $K = 100$, $R_1 = 60$, we see that there is an accurate overlap between the two for a variety of mutation rates that span several orders of magnitude (Fig. 3(a) and (b)). This match demonstrates how our analytical solutions can be used to accurately predict emergence probabilities in the face of different scenarios, including antibiotic resistance and tumour formation, as both these processes are characterised by high mutation rates.

We also tested a parameter set where the carrying capacity and initial growth rate was much higher ($K = 1{,}000$ and $R_1 = 100$, or $K = 10{,}000$ and $R_1 = 2{,}000$). Fig. 3(c)-(f) demonstrates that the analytical results slightly underestimate the simulation results to a small degree, especially if the mutated strain's growth rate is high and R_2 is close to K, but becomes more accurate as R_2 increases and generally provides a good approximation. These inaccuracies probably arise due to our analytical solution not fully accounting for the increased variance in both infection and immune growth rates that can arise if parameters are large, as in this model [30]. However, the error does not appear to be great, so the model can still be used to provide accurate estimates of emergence probability for this parameter set.

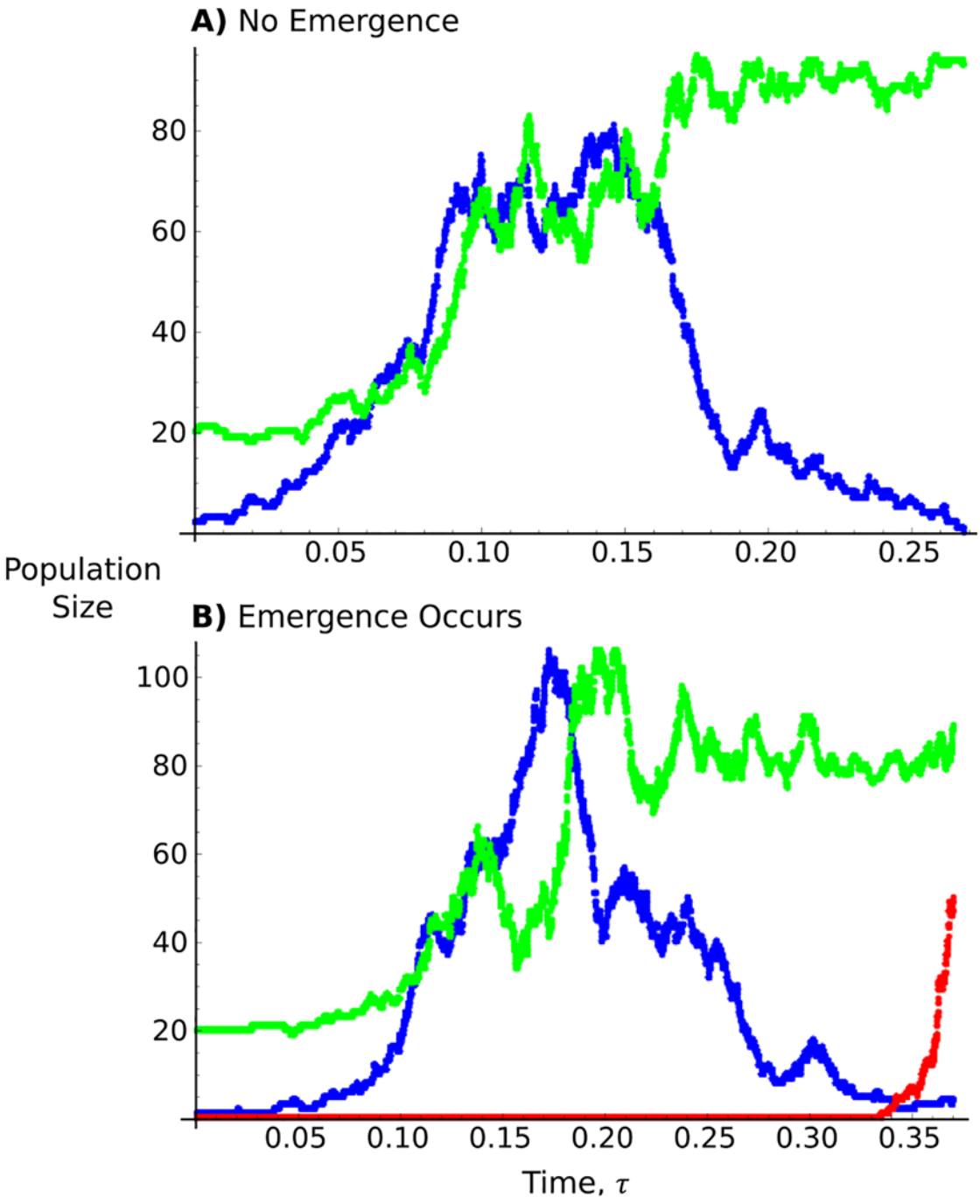

Fig 2. Example of simulation runs. Profiles of example simulation runs over time. In (A), the first strain goes extinct before a mutated pathogen arises, while in (B) emergence occurs. Blue dots represent the initial pathogen, red dots in (B) represent the mutated strain, while green dots show immune cell proliferation. The constant stream of red dots in (B) indicate the mutated strain at zero copies. Parameters are $K = 100$, $R_1 = 60$, $R_2 = 200$, $\rho = 0.5$, and $\mu = 0.01$.

doi:10.1371/journal.pcbi.1004149.g002

Finally, we also tested how well analytical solutions work for cases where $R_1 < R_2 < K$. Although such a mutated strain replicates more quickly, Equation 4 shows that it will die out in the long-term. Hence our analytical solutions might not correctly reflect the emergence probability of these infections. Nevertheless, even in this case, Equation 7 accurately matches up with

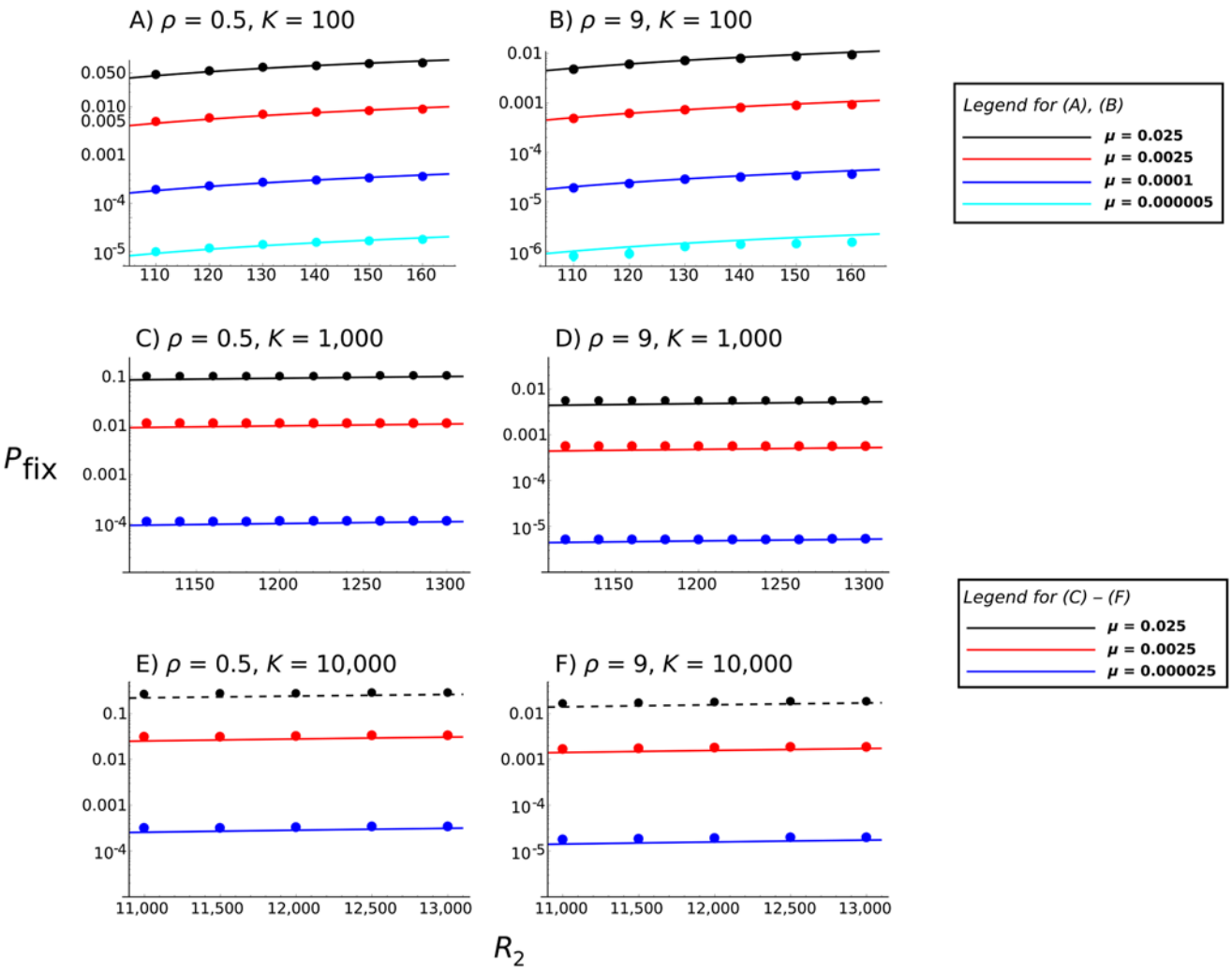

Fig 3. Comparison of simulations with analytical solutions. Comparisons of the full analytical solution (Equation 7, with Π given by Equation 10) with simulation results. Solid lines represent analytical solutions; points are simulation calculations. Graphs are plotted as a function of the reproductive ratio of the second strain, R_2. Note that the y axis is plotted on a log scale. Different colours denote different mutation rates, as shown in the accompanying legend. Other parameters are (A and B) $K = 100$, $R_1 = 60$; (C and D) $K = 1000$, $R_1 = 100$; or (E and F) $K = 10{,}000$, $R_1 = 2000$. In all panels, $x_{init} = 1$ and $y_{init} = 20$. ρ equals either 0.5 (A, C, and E) or 9 (B, D, and F). All error bars, as calculated using binomial confidence intervals, lie within the points. Further results are shown in Section 2 of S1 Text.

doi:10.1371/journal.pcbi.1004149.g003

simulation results in this parameter range, although some inaccuracies arise for $K = 1{,}000$ (S1 Fig).

Effect of ongoing immunity on emergence

To exemplify why it is important to account for ongoing immune growth, we compared our full solution for the emergence probability (Equation 7, with Π equal to Equation 10), to a 'naive' estimate that does not assume ongoing immune proliferation ($\Pi = 1 - 1/R_2$). We see that our result leads to a greatly reduced emergence probability, with values from our model being $\sim$1,000 times lower than the naive estimate (Fig. 4, and Section 3 of S1 Text). Similar results apply if feedbacks only affect the mutant after it has arose (that is, $\Pi = (1 - P^*_{ext})$ as given by Equation 9). Hence, as in previous non-equilibrium models [21], one needs to account for

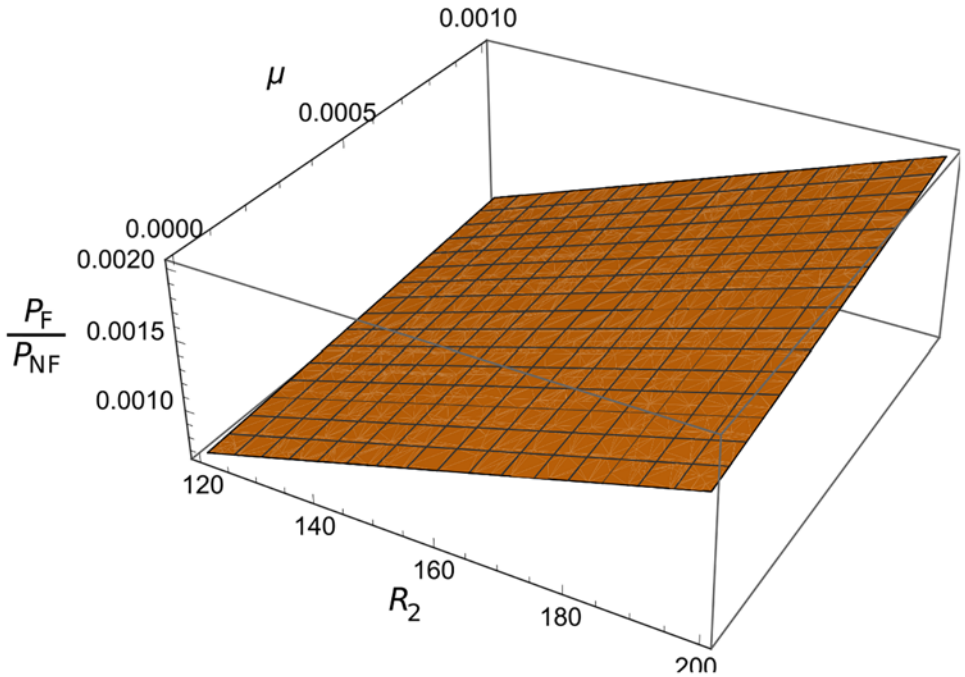

Fig 4. Quantifying effect of population-level feedbacks. 3D plots comparing the total emergence probability of a mutated strain with feedbacks (denoted P_F), to one where feedbacks are not considered (denoted P_{NF}), as a function of mutation rate μ and R_2. Other parameters are $K = 100$, $R_1 = 60$, $\rho = 5$, $x_{init} = 1$, $y_{init} = 20$.

doi:10.1371/journal.pcbi.1004149.g004

dynamical feedbacks, both before and after the mutated strain appears, in order to account for the reduced emergence probability in these scenarios.

Comparing pathogen growth against death rate

We next studied what process has a larger effect on pathogen emergence. Immune escape can either be achieved by overgrowing the immune response (increasing the intrinsic pathogen growth rate φ), or by tolerating it (reducing the immune-mediated death rate σ). One might expect that the two processes would lead to similar increase in escape probability, since both affect the effective reproductive rate R^*. However, this intuition need not hold in the face of immune-mediated feedbacks. We commence with a heuristic analysis based on the deterministic model to predict general behaviour for a newly-emerging strain, then use numerical analyses to check this reasoning.

If there are two strains spreading concurrently, the deterministic rate of change of immunity, y, and the second strain x_2, is given by the following set of differential equations:

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = x_2(\varphi_2 - \sigma_2 y) \tag{11a}$$

$$\frac{\mathrm{d}y}{\mathrm{d}t} = r(x_1 + x_2)y\left(1 - \frac{y}{K}\right) \tag{11b}$$

Further recall that that the basic pathogen growth rate with only one immune cell is $R_2 = \varphi_2/\sigma_2$. In order to augment its spread, the infection can either increase its intrinsic growth rate by a certain factor (i.e. change $\varphi_2 \to \varphi_2\, c$, where $c > 1$ is a numerical constant), or instead

tolerate the immune response (mathematically equivalent to the transformation $\sigma_2 \rightarrow \sigma_2/c$). We can investigate the effect of both these rescaled variables on the rate of change of pathogen spread by substituting them into Equation 4. After making the substitution $\varphi_2 \rightarrow \varphi_2 c$, the pathogen rate of change becomes:

$$\frac{dx_2}{dy} = \frac{K(cR_2 - y)}{\rho y(K - y)} \left(\frac{x_2}{x_1 + x_2}\right) \left(\frac{\sigma_2}{\sigma_1}\right) \tag{12}$$

The $x_2/(x_1 + x_2)$ term arises due to the presence of the pre-existing initial strain x_1, and the death-rate ratio σ_2/σ_1 appears since ρ was initially scaled by σ_1; this term disappears if we assume equal death rates. Apart from these terms, Equation 12 is conceptually the same as Equation 4, but with R scaled by a factor c, as expected. However, if we instead scale $\sigma_2 \rightarrow \sigma_2 / c$, a different result emerges:

$$\frac{dx_2}{dy} = \frac{K(cR_2 - y)}{c\rho y(K - y)} \left(\frac{x_2}{x_1 + x_2}\right) \left(\frac{\sigma_2}{\sigma_1}\right) \tag{13}$$

Here, not only is the reproductive ratio R_2 scaled, but the immune growth term ρ is also altered. Mathematically, this is a simple consequence of the fact that the growth rate is scaled by $1/\sigma_1$, so any rescaling of σ_2 also affects ρ through its effect on the σ_2/σ_1 ratio. Biologically, this outcome reflects the fact that a rescaling of immune-mediated death would not have the same effect on the growth rate than changing φ_2 by the same ratio, since pathogen death is also a function of the immune population, y (see Equation 1). So although tolerating immunity might increase the infection growth rate, it comes at a cost of increasing the effective growth rate of immunity, as each immune cell would have a larger average impact on pathogen death.

Therefore, while a rise in growth rate (φ_2) would only affect R_2, reducing the death rate (σ_2) comes at a cost of increasing the effective proliferation rate of immune cells. Hence, one expects that it is more advantageous for an infection to increase φ_2 and outgrow the immune response, instead of reducing σ_2 and tolerating it. Furthermore, note that the emergence probability Π in the presence of epidemiological feedbacks contains a term of order $1/\rho$ (as with dx_2/dy), so the effective rise in immunity will further lead to a negative overall effect on emergence probability.

We verified this intuition by comparing the analytical results for cases where φ_2 is increased by a set factor (so that only R_2 is changed by this rescaling), to outcomes where σ_2 is scaled (which will not just affect R_2 but will also increase ρ by the same factor, as outlined above). Fig. 5(a) demonstrates how, for large parameter values, increasing φ_2 only produces a higher overall emergence probability. Qualitatively similar results arise if using different parameter values.

Comparison with earlier results. These results appear to contradict a previous analysis by Alizon [27]. Using nested models, it was shown that in order to maximise the epidemiological (between-host) reproductive ratio R_0, increasing the growth rate might not always be the best option for a pathogen because it shortens the duration of the infection.

There are several changes between the models that could explain the different outcomes. Principally, [27] aimed to determine the long-term R_0 across a population, while our model concerns the emergence of new strains intra-host. In addition, [27] used deterministic equations, and did not investigate the stochastic emergence of new strains (that is, whether those with a smaller R_0 are likely to emerge if they appear at a low frequency). However, one major difference that is worthy of further investigation concerns the different immune-response functions used in each model.

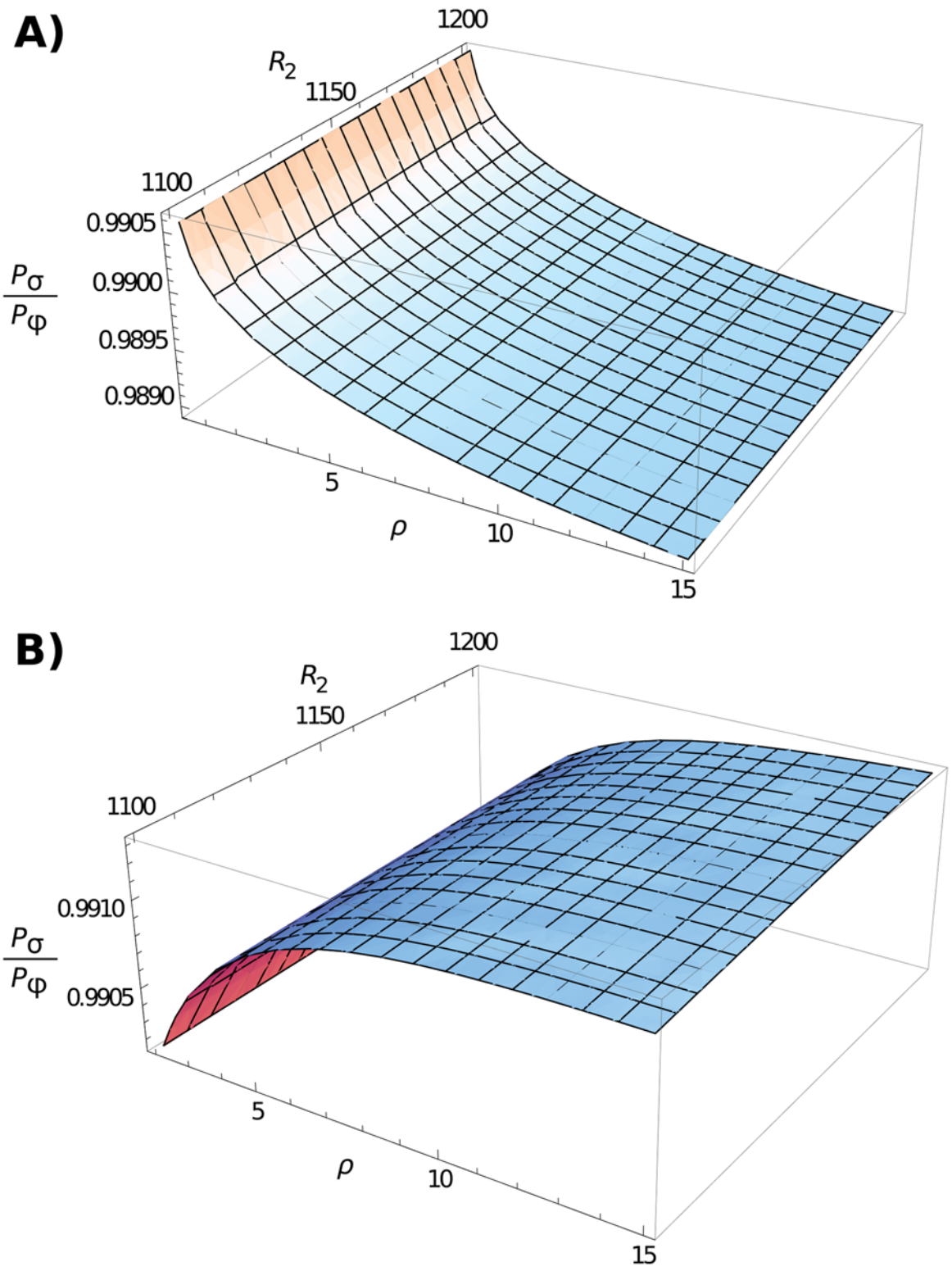

Fig 5. Comparing pathogen growth against death rate. 3D plot comparing the emergence probability if the death rate σ is reduced by a factor $1/c = 1.01$ (here denoted P_σ), to the case where the growth rate φ is increased by $c = 1.01$ (denoted P_φ). Plot (A) is for the case where immune response $dy / dx = \rho xy(1 - y/K)$, as in our model, and (B) assumes $dy / dx = \rho\varphi xy(1 - y/K)$ to enable comparison of our model with [27]. Other parameters are $K = 1{,}000$, $R = 100$, $x_{init} = 1$, $y_{init} = 20$, $\mu = 0.025$ and $\sigma_1 = 1$ for (B).

doi:10.1371/journal.pcbi.1004149.g005

In [27], the model assumed that the dynamics of the immune cell population took the form $dy / dt = r\varphi^b xy$ (Equation 3 in that paper). Hence, the immune response was not just stimulated by the presence of an infection, but is also dependent on pathogen replication. The reason for this assumption is that the expression of pathogen peptides on the infected cell surface is strongly dependent on the replication activity [40]. This is why latent viruses, such as herpes virus, can go unnoticed from the immune system and persist for long periods of time. Interestingly, if $b = 1$ then, according to this form of dy / dt, overgrowing and tolerating the immune response will have equal effects on pathogen emergence. This can be seen by forming dx_2 / dy as before, and after substituting either $\varphi_2 \rightarrow \varphi_2 c$ or $\sigma_2 \rightarrow \sigma_2/c$, one sees that each transformation results in the same rescaled differential equation (Section 3 of S1 Text):

$$\frac{dx_2}{dy} = \frac{K(cR_2 - y)}{cR_2\sigma_1\rho y(K - y)}\left(\frac{x_2}{x_1 + x_2}\right) \tag{14}$$

Hence, if immunity is also linearly dependent on pathogenic growth, then increasing growth and reducing death will have equal effects on long-term emergence in a static population. However, due to the ongoing spread of immunity in our model, increasing immune tolerance might still cause the greatest negative impact on the emergence of mutated strains. This is due to its ensuing effect of increasing the 'effective' immune growth, which can restrict new strains from emerging (as reflected in the $1/\rho$ term in Π). This is indeed what is observed; an example of this behaviour is given in Fig. 5(b). Note that increasing ρ in this case leads to opposing behaviour than to what is seen for our model, in the sense that higher ρ reduces the advantage of increased φ_2. Presumably, this is due to the φ_2 term affecting the immune growth-rate, which can offset the reduction in emergence probability caused by the ongoing immune growth.

Discussion

Since infectious diseases can exhibit high mutation rates, infectious diseases always pose a strong risk that they can evolve into new strains, especially ones that circumvent host defences. While this question has been extensively studied at the epidemiological level, the question of evolutionary emergence within hosts has received less scrutiny. Furthermore, few existing models have accounted for population dynamics feedbacks that can impact and restrict the emergence of mutated strains.

Here, we fill a key gap in existing modelling knowledge by deriving analytical methods for determining the emergence of new strains via within-host evolution. To this end, we tailor a broadly-applicable population dynamics model to the interaction between pathogens and immune cells, a system which share similarities with predator-prey models [29]. As with previous work studying feedbacks at the epidemiological level [21], we highlight how ongoing population spread (in this case, of immune cells) can strongly limit the emergence of new strains. This result is reflected in the fact that the emergence probability is greatly reduced due to ongoing immune proliferation (Equation 10); accounting for this feedback shows how it strongly restricts emergence of mutated strains (Section 3 of S1 Text).

After verifying the accuracy of the model, we subsequently determined whether it is more advantageous for a new infection to increase its growth rate (φ), or instead increase tolerance to the immune response (decrease σ), both of which will increase a pathogen's reproductive ratio R and hence qualifies as immune escape phenomena. We demonstrate that, rather surprisingly, it always pays in the short-term for a pathogen to increase its replication rate, and to try and outgrow the immune response. The rationale behind this behaviour is that if immune tolerance rises by a certain factor, it will not cause an equal reduction on pathogen death since immunity is still present at a high frequency. Intuitively, the presence of multiple strains will

cause indirect competition between them, since the presence of both causes a higher spread of immunity which reduces the mutated strain's growth rate when rare. This phenomenon manifests itself mathematically as an increase in the scaled immune growth rate, ρ. Since increasing effective immunity would also reduce the emergence probability of new strains (Equation 10), it is evolutionary advantageous for pathogens to evolve a higher growth rate instead (Fig. 5).

One elegant outcome of the modelling framework is that these general outcomes should be robust to the function used to model immune-cell proliferation. It is a concern in epidemiological modelling that key results are heavily dependent on the specific mathematical functions used [41], and several models have been previously used to reflect immune growth [29]. However, most of these functions approximate to exponential growth when immunity initially appears; furthermore, the form of the emergence probability (Equation 8) makes clear that any immune growth will feedback onto the emergence of the mutated strain, limiting its appearance. For this reason, our conclusion that infectious diseases would benefit more from higher growth rates should also be robust to how immunity spreads over time.

There exists evidence, in line with these findings, that HIV increases its growth rate over the course of an infection [42]. However, the observed increase is small. A slight increase might only be necessary if too large a growth rate would trigger a larger immune response, or kill off the host too quickly and thus limit the epidemiological spread of the pathogen (both effects have been previously investigated by [27]). It has been proposed that immune escape for HIV becomes less efficient over time, which might be the cause of a low increase in replication rate [15]. Hence it might be mechanistically more feasible for a pathogen to instead evade the immune response, by mutating in or close to a virus epitope, so as to avoid recognition by lymphocytes that would otherwise neutralise it (a variant on the 'immune tolerance' scenario outlined above). These type of escape mutations have been widely documented in HIV infections, which is one reason why the virus is able to persist for long periods of time, and also why vaccination is so difficult [43, 44]. Similarly, although not a chronic illness, evidence exists demonstrating that influenza A/H1N1 evades immunity via evolution causing increased virus affinity to cell receptors, which enlarges its replication rate [20]. To further investigate this issue, future models need to be created that combine short-term effects of pathogen emergence, as used here, with long-term forecasts of the pathogenic steady-state (since our model only deals with the initial, emergence stage).

More generally, the huge challenge exerted by the immune system on pathogen populations *de facto* generates apparent competition between different infection strains (or even species) co-infecting the same host [45]. In this case, it has been postulated that the faster-replicating strain could persist in the long-term by sheer force of numbers [46], as predicted with our model. Several empirical studies exist that support this intuition, as observed with malaria [47, 48] and schistosomes [49] (although competition for another resource, such as red-blood cells for malaria infections, could have partly explained these results).

Implications of the model for different cases of immune escape

Our model can also be used to shed light on different processes of within-host emergence that have been observed in clinical studies. Different diseases show varying outcomes with regards to the production of immune-escape mutations. Two extensively-studied human diseases, HIV and HCV, are both characterised by a successive emergence of new strains over time. In particular, HIV is well-known to produce 'escape mutations' that evade T cell-mediated immunity [43, 44]. On a within-host phylogeny, this behaviour is characterised by creation of new subclades [5–7]. This behaviour can be intuitively explained by noting that HIV has extremely high mutation rates, estimated at ~ 0.2 errors made per replication cycle and the ability to

produce 10^{10} – 10^{12} virions per day [6] (although mutation creation might be limited by a lower N_e, which has been estimated to lie between 1,000 [50] and 100,000 [51]). Furthermore, a sizeable proportion of mutations are beneficial, with estimates of adaptive substitutions in the *env* gene placed at ∼55% [52] (although the actual proportion of spontaneous beneficial mutations will be less than this, as observed from mutagenesis studies with viruses [53, 54]). What may seem surprising is that given this evolutionary potential, we do not see a more pronounced increase in HIV replication rate throughout the course of an infection. Yet, conversely, immune escape is very strongly selected for [14]. One possibility could be that given the constrains on RNA virus genomes [55], evading the immune response might be easier to achieve than increasing the replication rate. Furthermore, if the maximum immunity population size is high, it could be too complicated mechanistically to outgrow it, rather than than simply evading it.

HCV also shows a similar propensity to produce immune-escape mutations, with an estimated substitution rate of 1.2×10^{-4} per replication cycle and 10^{12} virions generated each day [7] (although, as with HIV, the effective population size is greatly lower than this value [7]). Acute hepatitis C infections are characterised by little diversity accumulating over time, except in one specific genetic region (NS5B; [56]). Chronic infections accumulate much more diversity, in line with the hypothesis that they continuously evolve to evade the immune system [16]. These infections are also characterised by different immune profiles: acute infections lead to a high, sustained immune response, which tends to be greatly lowered in chronic infections [57]. Our model suggests that if the immune response naturally increased rapidly (i.e. a higher intrinsic ρ in the model), it can greatly limit the emergence of new pathogens as it rises, preventing immune escape and a chronic illness. There exists evidence that mutations in hosts correlate with infection outcome, mainly due to SNPs at the *IL28B* loci, which could cause this effect [58–60]. However, it remains controversial as to what determines virus clearance, especially since there exist evidence for virus control over infection outcome [61]. Therefore, the effect on immune response on the creation of escape mutations requires further study.

Other diseases are characterised by low probability of emergence despite frequent mutation, for which immune feedbacks could be the cause. Cancer growth is characterised by evading pre-programmed cell death and extremely rapid cell replication [17]. Furthermore, genomes present in cancer cells usually carry 'mutator' alleles, causing additional cell instability [62]. Therefore, cell mutation can be extremely common, but can generally be stopped by the immune response. Yet it is known that tumours can mutate immune-checkpoint networks to generate protection from the immune system. One prominent example is the up-regulation of ligands for the programmed cell death protein 1 (PD1) pathway, which can block antitumour immune responses [63]. Our model suggests that due to the rapid replication of tumour cells, they could also strongly trigger a heightened immune rate, the increased spread of which will greatly prevent cancer emergence. This mechanism could explain why tumour emergence is rare relative to the potential mutation rate.

Overall summary

By accounting for the ongoing spread of immunity, we have quantified how this particular effect can feedback onto emergence of mutated pathogens intra-host, and inhibit the appearance of mutated strains. Further analysis of the model demonstrates how it is beneficial for a pathogen to increase its replication rate and attempt to outgrow immunity, as opposed to tolerate it. This model can therefore shed light on expected within-host evolutionary dynamics of infections, as well as determine why extremely rapidly replicating pathogens do not emerge as often as expected.

Supporting Information

S1 Fig. Simulation comparisons where $R_1 < R_2 < K$. Comparisons of the full analytical solution (Equation 7, with Π given by Equation 10) with simulation results. Solid lines represent analytical solutions; points are simulation calculations. Graphs are plotted as a function of the reproductive ratio of the second strain, R_2. Note that the y axis is plotted on a log scale. Different colours denote different mutation rates, as shown in the accompanying legend. Other parameters are (A and B) $K = 100$, $R_1 = 60$; (C and D) $K = 1{,}000$, $R_1 = 100$; or (E and F) $K = 10{,}000$, $R_1 = 2000$. In all panels, $x_{init} = 1$ and $y_{init} = 20$. ρ equals either 0.5 (A, C, and E) or 9 (B, D, and F). Most error bars, as calculated using binomial confidence intervals, lie within the points if they cannot be seen.
(EPS)

S1 Text. Contains additional information on derivations, and further comparisons against simulation results (*Mathematica* NB format). The file is separated into three sections:

Section 1 Setting up the mathematical model. In-depth mathematical analyses of the differential equations used, and how to derive the emergence probability if affected by immune growth (Equation 10 in the main text).

Section 2 Testing against stochastic simulations. Full list of analytical comparisons with simulation outcomes, which contributed to Fig. 3.

Section 3 Mathematical analysis of analytical solution. Further outline of mathematical analysis and numerical computation, demonstrating (i) how immune feedbacks affect population growth, and (ii) how different φ and σ values affect mutated pathogen emergence.
(PDF)

S2 Text. Same as Text S1, but in PDF format.
(PDF)

Acknowledgments

We would like to thank Helen Alexander for helpful comments on the manuscript.

Author Contributions

Conceived and designed the experiments: MH SA. Performed the experiments: MH SA. Analyzed the data: MH SA. Contributed reagents/materials/analysis tools: MH SA. Wrote the paper: MH SA.

References

1. Morens DM, Folkers GK, Fauci AS (2004) The challenge of emerging and re-emerging infectious diseases. Nature 430: 242–249. doi: 10.1038/nature02759 PMID: 15241422
2. Woolhouse M, Scott F, Hudson Z, Howey R, Chase-Topping M (2012) Human viruses: discovery and emergence. Philos Trans R Soc Lond B Biol Sci 367: 2864–2871. doi: 10.1098/rstb.2011.0354 PMID: 22966141
3. 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 PMID: 14668863
4. Gandon S, Hochberg ME, Holt RD, Day T (2013) What limits the evolutionary emergence of pathogens? Philos Trans R Soc Lond B Biol Sci 368. doi: 10.1098/rstb.2012.0086 PMID: 23209168
5. Alizon S, Luciani F, Regoes RR (2011) Epidemiological and clinical consequences of within-host evolution. Trends Microbiol 19: 24–32. doi: 10.1016/j.tim.2010.09.005 PMID: 21055948
6. Rambaut A, Posada D, Crandall KA, Holmes EC (2004) The causes and consequences of HIV evolution. Nat Rev Genet 5: 52–61. doi: 10.1038/nrg1246 PMID: 14708016

7. Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, et al. (2011) Sequential Bottlenecks Drive Viral Evolution in Early Acute Hepatitis C Virus Infection. PLoS Pathog 7: e1002243. doi: 10.1371/journal.ppat.1002243 PMID: 21912520

8. Goldie J, Coldman A (1979) A mathematic model for relating the drug sensitivity of tumors to their spontaneous mutation rate. Cancer Treat Rep 63. PMID: 526911

9. Iwasa Y, Michor F, Nowak MA (2003) Evolutionary dynamics of escape from biomedical intervention. Proc R Soc B 270: 2573–2578. doi: 10.1098/rspb.2003.2539 PMID: 14728779

10. Komarova NL, Wodarz D (2005) Drug resistance in cancer: Principles of emergence and prevention. Proc Natl Acad Sci USA 102: 9714–9719. doi: 10.1073/pnas.0501870102 PMID: 15980154

11. Alexander HK, Bonhoeffer S (2012) Pre-existence and emergence of drug resistance in a generalized model of intra-host viral dynamics. Epidemics 4: 187–202. doi: 10.1016/j.epidem.2012.10.001 PMID: 23351371

12. Loverdo C, Park M, Schreiber SJ, Lloyd-Smith JO (2012) Influence of viral replication mechanisms on within-host evolutionary dynamics. Evolution 66: 3462–3471. doi: 10.1111/j.1558-5646.2012.01687.x PMID: 23106710

13. Bonhoeffer S, May RM, Shaw GM, Nowak MA (1997) Virus dynamics and drug therapy. Proc Natl Acad Sci USA 94: 6971–6976. doi: 10.1073/pnas.94.13.6971 PMID: 9192676

14. Richman DD, Wrin T, Little SJ, Petropoulos CJ (2003) Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc Natl Acad Sci USA 100: 4144–4149. doi: 10.1073/pnas.0630530100 PMID: 12644702

15. Blanquart F, Gandon S (2013) Time-shift experiments and patterns of adaptation across time and space. Ecol Lett 16: 31–38. doi: 10.1111/ele.12007 PMID: 23034076

16. Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, et al. (2000) The Outcome of Acute Hepatitis C Predicted by the Evolution of the Viral Quasispecies. Science 288: 339–344. doi: 10.1126/science.288.5464.339 PMID: 10764648

17. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100: 57–70. doi: 10.1016/S0092-8674(00)81683-9 PMID: 10647931

18. Martin G, Aguilée R, Ramsayer J, Kaltz O, Ronce O (2013) The probability of evolutionary rescue: towards a quantitative comparison between theory and evolution experiments. Philos Trans R Soc Lond B Biol Sci 368.

19. Alexander HK, Martin G, Martin OY, Bonhoeffer S (2014) Evolutionary rescue: linking theory for conservation and medicine. Evol Appl. doi: 10.1111/eva.12221 PMID: 25558278

20. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, et al. (2009) Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift. Science 326: 734–736. doi: 10.1126/science.1178258 PMID: 19900932

21. Hartfield M, Alizon S (2014) Epidemiological Feedbacks Affect Evolutionary Emergence of Pathogens. Am Nat 183: E105–E117. doi: 10.1086/674795 PMID: 24642501

22. Antia R, Levin BR, May RM (1994) Within-host population dynamics and the evolution and maintenance of microparasite virulence. Am Nat 144: 457–472. doi: 10.1086/285686

23. Sasaki A (1994) Evolution of antigen drift/switching: Continuously evading pathogens. J Theor Biol 168: 291–308. doi: 10.1006/jtbi.1994.1110 PMID: 7520957

24. Gilchrist MA, Sasaki A (2002) Modeling Host–Parasite Coevolution: A Nested Approach Based on Mechanistic Models. J Theor Biol 218: 289–308. doi: 10.1006/jtbi.2002.3076 PMID: 12381431

25. Ganusov VV, Bergstrom CT, Antia R (2002) Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. Evolution 52: 213–223. doi: 10.1111/j.0014-3820.2002.tb01332.x

26. André JB, Ferdy JB, Godelle B (2003) Within-host parasite dynamics, emerging trade-off, and evolution of virulence with immune system. Evolution 57: 1489–1497. doi: 10.1111/j.0014-3820.2003.tb00357.x PMID: 12940354

27. Alizon S (2008) Transmission–recovery trade–offs to study parasite evolution. Am Nat 172: E113–E121. doi: 10.1086/589892 PMID: 18710338

28. Pepin KM, Volkov I, Banavar JR, Wilke CO, Grenfell BT (2010) Phenotypic differences in viral immune escape explained by linking within-host dynamics to host-population immunity. J Theor Biol 265: 501–510. doi: 10.1016/j.jtbi.2010.05.036 PMID: 20570681

29. Alizon S, van Baalen M (2008) Acute or Chronic? Within–Host Models with Immune Dynamics, Infection Outcome, and Parasite Evolution. Am Nat 172: E244–E256. doi: 10.1086/592404 PMID: 18999939

30. Allen L (2008) An introduction to stochastic epidemic models. In: Brauer F, van den Driessche P, Wu J, editors, Mathematical Epidemiology, Berlin / Heidelberg: Springer, volume 1945 of *Lecture Notes in Mathematics*, chapter 3. pp. 81–130.

31. Hansen E, Day T (2011) Optimal antiviral treatment strategies and the effects of resistance. Proc R Soc B 278: 1082–1089. doi: 10.1098/rspb.2010.1469 PMID: 20880883

32. Anderson RM, May RM (1991) Infectious Diseases of Humans. Dynamics and Control. Oxford: Oxford University Press.

33. Otto SP, Day T (2007) A Biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton: Princeton University Press.

34. Feller W (1951) Diffusion processes in genetics. In: Proc. Second Berkeley Symp. on Math. Statist. and Prob., Berkeley and Los Angeles: Univ. of Calif. Press. pp. 227–246.

35. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81: 2340–2361. doi: 10.1021/j100540a008

36. Gillespie DT (2001) Approximate accelerated stochastic simulation of chemically reacting systems. J Chem Phys 115: 1716–1733. doi: 10.1063/1.1378322

37. Keeling MJ, Rohani P (2007) Modeling Infectious Diseases in Humans and Animals. Princeton University Press.

38. Hartfield M, Alizon S (2013) Introducing the outbreak threshold in epidemiology. PLoS Pathog 9: e1003277. doi: 10.1371/journal.ppat.1003277 PMID: 23785276

39. Otto SP, Whitlock MC (1997) The probability of fixation in populations of changing size. Genetics 146: 723–733. PMID: 9178020

40. Yewdell JW (2007) Plumbing the sources of endogenous MHC class I peptide ligands. Curr Opin Immunol 19: 79–86. doi: 10.1016/j.coi.2006.11.010 PMID: 17140786

41. De Boer RJ (2012) Which of our modeling predictions are robust? PLoS Comput Biol 8: e1002593 EP. doi: 10.1371/journal.pcbi.1002593 PMID: 22844235

42. Kouyos RD, von Wyl V, Hinkley T, Petropoulos CJ, Haddad M, et al. (2011) Assessing Predicted HIV-1 Replicative Capacity in a Clinical Setting. PLoS Pathog 7: e1002321. doi: 10.1371/journal.ppat.1002321 PMID: 22072960

43. Goulder PJR, Watkins DI (2004) HIV and SIV CTL escape: implications for vaccine design. Nat Rev Immunol 4: 630–640. doi: 10.1038/nri1417 PMID: 15286729

44. Palmer D, Frater J, Phillips R, McLean AR, McVean G (2013) Integrating genealogical and dynamical modelling to infer escape and reversion rates in HIV epitopes. Proc R Soc B 280: 20130696. doi: 10.1098/rspb.2013.0696 PMID: 23677344

45. Read AF, Taylor LH (2001) The ecology of genetically diverse infections. Science 292: 1099–1102. doi: 10.1126/science.1059410 PMID: 11352063

46. Mideo N (2009) Parasite adaptations to within-host competition. Trends Parasitol 25: 261–268. doi: 10.1016/j.pt.2009.03.001 PMID: 19409846

47. de Roode JC, Helinski MEH, Anwar MA, Read AF (2005) Dynamics of multiple infection and within-host competition in genetically diverse malaria infections. Am Nat 166: 531–542. doi: 10.1086/491659 PMID: 16224719

48. Vardo AM, Kaufhold KD, Schall J (2007) Experimental test for premunition in a lizard malaria parasite. J Parasitol 93: 280–282. doi: 10.1645/GE-1005R.1 PMID: 17539410

49. Dumont M, Moné H, Mouahid G, Idris MA, Shaban M, et al. (2007) Influence of pattern of exposure, parasite genetic diversity and sex on the degree of protection against reinfection with *Schistosoma mansoni*. Parasitol Res 101: 247–252. doi: 10.1007/s00436-007-0476-0 PMID: 17310396

50. Leigh Brown AJ (1997) Analysis of HIV-1 env gene sequences reveals evidence for a low effective number in the viral population. Proc Natl Acad Sci USA 94: 1862–1865. doi: 10.1073/pnas.94.5.1862

51. Pennings PS, Kryazhimskiy S, Wakeley J (2014) Loss and Recovery of Genetic Diversity in Adapting Populations of HIV. PLoS Genet 10: e1004000. doi: 10.1371/journal.pgen.1004000 PMID: 24465214

52. Williamson S (2003) Adaptation in the *env* Gene of HIV-1 and Evolutionary Theories of Disease Progression. Mol Biol Evol 20: 1318–1325. doi: 10.1093/molbev/msg144 PMID: 12777505

53. Eyre-Walker A, Keightley PD (2007) The distribution of fitness effects of new mutations. Nat Rev Genet 8: 610–618. doi: 10.1038/nrg2146 PMID: 17637733

54. Sanjuán R (2010) Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philos Trans R Soc Lond B Biol Sci 365: 1975–1982. doi: 10.1098/rstb.2010.0063 PMID: 20478892

55. Belshaw R, Gardner A, Rambaut A, Pybus OG (2008) Pacing a small cage: mutation and RNA viruses. Trends Ecol Evol 23: 188–193. doi: 10.1016/j.tree.2007.11.010 PMID: 18295930

56. Liu L, Fisher BE, Thomas DL, Cox AL, Ray SC (2012) Spontaneous clearance of primary acute hepatitis C virus infection correlated with high initial viral RNA level and rapid HVR1 evolution. Hepatology 55: 1684–1691. doi: 10.1002/hep.25575 PMID: 22234804

57. Dustin LB, Rice CM (2007) Flying Under the Radar: The Immunobiology of Hepatitis C. Annu Rev Immunol 25: 71–99. doi: 10.1146/annurev.immunol.25.022106.141602 PMID: 17067278

58. Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, et al. (2009) Genetic variation in *IL28B* predicts hepatitis C treatment-induced viral clearance. Nature 461: 399–401. doi: 10.1038/nature08309 PMID: 19684573

59. Thomas DL, Thio CL, Martin MP, Qi Y, Ge D, et al. (2009) Genetic variation in *IL28B* and spontaneous clearance of hepatitis C virus. Nature 461: 798–801. doi: 10.1038/nature08463 PMID: 19759533

60. Rauch A, Kutalik Z, Descombes P, Cai T, Di Iulio J, et al. (2010) Genetic Variation in *IL28B* Is Associated With Chronic Hepatitis C and Treatment Failure: A Genome-Wide Association Study. Gastroenterology 138: 1338–1345.e7. doi: 10.1053/j.gastro.2009.12.056 PMID: 20060832

61. Hartfield M, Bull R, White PA, Lloyd A, Luciani F, et al. (2014) Evidence that hepatitis C virus genome partly controls infection outcome. Evol Appl 7: 533–547. doi: 10.1111/eva.12151 PMID: 24944567

62. Jackson AL, Loeb LA (1998) The mutation rate and cancer. Genetics 148: 1483–1490. PMID: 9560368

63. Pardoll DM (2012) The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 12: 252–264. doi: 10.1038/nrc3239 PMID: 22437870