The mean and variance of infection load reveal unseen host evolution

Abstract

Some hosts are more heavily infected than others, but this fact has been relatively neglected in the microparasite literature compared to the macroparasite literature. We develop a model that allows a host population's mean and variance of microparasite load to fluctuate dynamically over time, dependent on underlying processes. Our model uses empirically reasonable, non-linear functional forms, e.g., for the probability of infected hosts surviving as a function of load. Despite non-linearities, we derive an approximation that assumes a lognormal distribution of load, improving computational and analytical tractability. We show negative feedbacks between load mean and variance at equilibrium. These feedbacks determine how a change in the host-microparasite system shifts mean and variance in load, directly as well as indirectly. We parameterize this model with empirically supported parameter values for a focal frog-fungal pathogen system. We illustrate system changes through evolution of host defenses against microparasites that may drive evolutionary recovery from disease-induced declines. We find that different forms of host defenses have different implications for load, e.g., constitutive resistance has little impact on the standard deviation while inducible resistance decreases it, with different outcomes for host fitness, pathogen fitness, and our ability to infer underlying dynamics from population-level data.

Introduction

The outcomes of epidemics frequently hinge on how many hosts are infected and how heavily those hosts are infected. Classical microparasite (e.g., bacteria, viruses, fungi, etc.) epidemiology focused on the proportion of hosts that become infected ("prevalence"; Anderson and May 1981); beyond that, a significant body of empirical (Grodio et al. 2012; Sánchez et al. 2018; Schrock 2024; Walsman et al. 2022; Wilber et al.) and theoretical (Anderson and May 1978; Fowler and Déirdre Hollingsworth 2016; Grear and Hudson 2011; Kretzschmar and Adler 1993; May and Anderson 1978; Wilber et al. 2016a; Wilber et al. 2021) work has considered how heavily hosts are infected, especially in the macroparasite (e.g., helminths, arthropods, etc.) literature. A critical measure of the intensity of an infection, across a vast diversity of host-parasite systems, is the number of infectious agents (parasites, pathogens, etc.) infecting a host (“infection load”). Higher infection load can harm vulnerable hosts in myriad ways including by driving lower host fecundity (Lanciani 1975; Moller 1990), elevating host mortality (Acevedo et al. 2019; Wilber et al. 2016a; Wilber et al. 2016b), accelerating disease spread from the host (Acevedo et al. 2019; Walsman et al. 2022; Wilber et al. 2016a), and depressing rates of recovery from infection (Wilber et al. 2016a). As such, infection load can be a critical determinant of infection outcomes for an individual host as well as infection spread throughout a population and the total death toll from disease (Anderson and May 1978; May and Anderson 1978).

The importance of infection load motivates biologists to carefully monitor it across many host-pathogen systems. These data inspire, test, and help refine foundational disease theory on two key questions. First, does mean infection load reflect underlying biological mechanisms within hosts in characteristic ways, allowing us to infer how those mechanisms interact with load? Second, how does mean load affect population-level outcomes such as change in host

density or the (co-)evolution of hosts and/or pathogens (Gilchrist and Sasaki 2002; King and Bonsall 2017; Mideo et al. 2008)? In contrast to a host population's mean load, the dynamics of a host population's variance in load have received far less attention in the empirical and theoretical literature; the research that does exist has been confined largely to macroparasites (Adler and Kretzschmar 1992; Kretzschmar and Adler 1993). For microparasites, the dynamics of load in general, and variance in load especially, have received much less attention than for macroparasites. This is at odds with new insight that patterns in the variance in microparasite load can contain signatures of biological processes and significantly impact epidemiological dynamics (Vredenburg et al. 2010; Wilber et al. 2020), e.g., via superspreading of viruses (Beldomenico 2020). For example, variance in load will control how many hosts carry high enough loads to cause rapid mortality, critically modifying fitness, and perhaps evolution of hosts and pathogens, as well as population-level outcomes (Grear and Hudson 2011); further, variance in load may provide greater power to infer underlying, within-host dynamics than mean load alone. To address these gaps, we need additional theory for load-dependent dynamics in microparasite systems to explore how variance in load reflects within-host mechanisms and influences population-level outcomes.

The macroparasite literature demonstrated that variance in load, or the related aggregation of load on some hosts more than others, has strong implications for population dynamics and can respond to underlying biological mechanisms. Increased mortality with an increased infection load of macroparasites leads to a pattern of reduced aggregation, i.e., more similar infection loads across hosts (Anderson and Gordon 1982). Conversely, reproduction of macroparasites on a host can lead to increased aggregation (Grear and Hudson 2011). The degree of aggregation, in turn, can be critical for how macroparasites regulate host populations (Anderson and May 1978;

Kretzschmar and Adler 1993; May and Anderson 1978). Due to the importance of these and other biological mechanisms, aggregation patterns can reflect within-host mechanisms in wild populations (Wilber et al. 2020; Wilson et al. 2002). Macroparasite models of infection intensity have generally considered discrete numbers of macroparasites, simple functional forms for macroparasite-induced mortality, and no reproduction of macroparasites on hosts (though see May and Anderson 1978's model E for exponential macroparasite growth on hosts).

Building on the macroparasite literature, molecular advances motivate new theory for microparasites because many microparasites deviate from the assumptions of both macroparasite and classic microparasite models. Pathogen infection load, as quantified by molecular techniques (e.g., qPCR), can follow continuous distributions of infection load (Bonhoeffer et al. 2015; Euser et al. 2021; Yang et al. 2011; Yang et al. 2021), rather than the discrete distributions of macroparasite counts. Further, the availability of laboratory data on microparasites allows us to parameterize more biologically realistic, non-linear functions for the relationships between infection load and key processes such as pathogen growth rate or the mortality of infected hosts (e.g., Wilber et al. 2016a) contrasted with the linear functions often used in classic macroparasite theory. Further, the available data encourage us to also build upon and beyond classic, compartmental microparasite models that only track infected or not (Anderson and May 1981) because many microparasites grow within hosts slowly enough to drive heterogeneity in load, with large implications for population dynamics (Briggs et al. 2010).

Integral Projection Models (IPMs) track dynamic heterogeneity while also being easily parameterized with laboratory data (Coulson 2012; Merow et al. 2014; Metcalf et al. 2016; Wilber et al. 2016a). IPMs can track heterogeneity in multiple, continuously distributed variables in a population and have more recently been adapted for tracking infection load (Metcalf et al.

2016; Wilber et al. 2016a). Thus, IPMs offer a useful tool for modeling load-dependent dynamics in microparasites, helping us develop general theoretical insight into how variance in load reflects within-host mechanisms and influences population-level outcomes, like evolution. However, IPMs face challenges due to their high dimensionality (i.e., one state variable equation for every possible continuous value of infection load, which is technically infinity), which can make it difficult to draw conceptual, analytical insight from such a model. Simulation-based analyses of IPMs can be useful but the large number of equations makes running IPMs computationally slower than simpler host-microparasite models (particularly problematic for fitting models to data; Wilber et al. 2021). To develop more general theory regarding the effects of load heterogeneity on epidemiological dynamics in microparasite systems, we must address these limitations.

Here, we develop a reduced dimension IPM (rIPM) that explicitly tracks the dynamics of mean and variance in microparasite load. The rIPM assumes lognormally distributed parasite load (i.e., a moment-closure approach), simplifying an infinite dimensional IPM to a finite set of dynamical equations for the first and second moments of the load distribution. Moment-closure approaches can improve analytical tractability for conceptual insight as well as computational efficiency (Bolker and Pacala 1997). A previous model assumed a variable mean (the first moment of load) while neglecting higher moments (e.g., fixed variance, which is related to the second moment) of lognormally distributed microparasites (Wilber et al. 2021). However, in host-microparasite systems that are fluctuating (such as those experiencing epizootics or recovering from disease-induced declines following host evolution), fluctuating variance may be critical for understanding the dynamics. As our goal is to develop general theory on load-dependent microparasite dynamics beyond only enzootic systems (e.g., to explore how variance

in load and evolution interact), extending rIPMs past the first moment is essential. We perform this moment closure to include a dynamic second moment (related to variance) but no further because the non-linear functional forms for within-host processes, motivated by empirical data, make deriving rIPMs even for the second moment mathematically non-trivial. Our moment closure only requires microparasite load to be lognormally distributed within a group of comparable hosts, which may be true for diverse microparasite systems (Bonhoeffer et al. 2015; Euser et al. 2021; Yang et al. 2011; Yang et al. 2021). This suggests that the insights from our model, which accounts for dynamic mean and variance of lognormal load, could apply to many systems.

One relevant microparasite is *Batrachochytrium dendrobatidis* (*Bd*), the fungus that causes chytridiomycosis and poses a global threat to amphibian populations (Rollins-Smith 2020). *Bd* load distributions are consistently lognormal across diverse host species but with widely different levels of variance (Schrock 2024); variance in load can be closely linked to critical processes such as host mortality (Wu 2023) and epidemic stage (Vredenburg et al. 2010). These empirical results motivate new theory of fluctuating mean and variance for inferring underlying mechanisms and influencing population-level outcomes like evolution of hosts or microparasites. Thus, an rIPM that considers fluctuating variance represents not only a methodological advance but also a significant conceptual leap. It allows us to address new questions that cannot be answered efficiently with a full IPM or at all with a fixed-variance rIPM. We derive an rIPM with fluctuating variance, which provides clear, novel, analytical insights regarding how host traits shift the feedbacks between the mean and variance of log infection load.

Lastly, we examine how load dynamics may illuminate the evolution of host defense by endangered Mountain yellow-legged frogs (*Rana muscosa*) against *Bd*. These endangered frogs,

native to the California Sierra Nevada, experienced large population declines due to chytridiomycosis caused by *Bd*. However, some populations are now recovering, in part due to host evolution that reduces infection loads (Briggs et al. 2010; Knapp et al. 2016; Knapp et al. 2023). The infection dynamics of this system are well-characterized, empirically, allowing us to leverage biologically-relevant parameter values for simulations of disease-induced decline and evolutionary recovery via different forms of host defense evolution.

We illustrate our analytical insights by contrasting how different forms of host evolution (tolerance, resistance that avoids infection, constitutive resistance to pathogen growth, and inducible resistance to pathogen growth) have different signatures and implications during evolutionary recovery. In particular, we are able to draw a very fine distinction between population load distributions associated with constitutive versus inducible resistance to pathogen growth within the host. Constitutive defenses are always active while inducible defenses are triggered by a stimulus, often a natural enemy attack (Adler and Karban 1994), e.g., infection (Shudo and Iwasa 2001); the advantages and consequences of constitutive versus inducible defense may depend on epidemic context and warrant further exploration (Westra et al. 2015). Amphibians possess some constitutive and some inducible defenses against *Bd* with further work required to determine their relative importance (Grogan et al. 2018), particularly in the evolutionary recovery of Mountain yellow-legged frogs (Knapp et al. 2023). We find that constitutive resistance lowers the mean of log infection load but, through feedbacks, weakly raises the variance. In contrast, inducible resistance lowers both, with implications for host and pathogen fitness. Thus, our model shows that population-level data on the variance of load may be highly informative about underlying mechanisms of host defense evolution, e.g., for recovering Mountain yellow-legged frog populations.

Full IPM structure and parametrization

Our full IPM builds on a previous model (Wilber et al. 2021) that examined short-term dynamics of pathogen invasion and an epizootic but we introduce two crucial new features: host reproduction and evolution. The original IPM tracks the distribution of log-transformed pathogen load, ln(pathogen load) = x, at discrete time point, t. From t to t+1, susceptible hosts survive with some probability s_0 and become infected with a probability that depends on the transmission rate parameter, β, and the number of pathogens in the environment, Z. The probability that a newly infected individual has load x follows normal probability density G_0(x) with some constant mean and constant standard deviation. Infected hosts have some fixed probability of recovery, l, and a survival probability, s(x), that declines with increasing x. Empirically, s(x) is close to 1 at low x then suddenly declines toward zero as x approaches and surpasses the LD50 of load. At this LD50 value (μ_{LD50}), an infected individual has a 50% chance of dying in a single time step (one day in our parameterization) and the gradualness of this decline rises with higher, constant values of a shape parameter (σ_{LD50}); precisely, $s(\mathrm{x})=1-\Phi((x-\mu_{\mathrm{LD50}})/\sigma_{\mathrm{LD50}})$ where Φ is the cumulative distribution function of the standard normal. Infected hosts that survive and do not recover can experience a decrease or increase in load so that the probability of a given infected host having load x' in the next time step follows a normal density as a function of the current load x [$G(x'|x)$ is a normal density with mean $a+bx$ and fixed standard deviation σ_{G}]. This pathogen growth function means that the mean log load increases linearly (note a linear increase in log load is an exponential increase in load on the natural scale) near $x = 0$ and more slowly as x increases to a within-host carrying capacity, $a/(1-b)$, neglecting death and recovery (assuming $a > 0$, $1 > b > 0$). Every time step, infected hosts shed λ environmental pathogens for each pathogen they carry

(number shed is λe^x to convert back to the natural scale) and environmental pathogens have a probability v of persisting in the environment. We modify this model by incorporating host reproduction and multiple genotypes, enabling us to explore evolutionary questions relevant to our focal system.

Mountain yellow-legged frogs (*Rana mucosa*) are endangered due to multiple stressors including disease-induced population declines from *Bd*. These declines depend critically on pathogen load and load-dependent mortality in infected frogs (Vredenburg et al. 2010). Some populations, however, are recovering in part due to host evolution (Briggs et al. 2010; Knapp et al. 2016; Knapp et al. 2023). Therefore, we examine how different host genotypes may vary in their load-dependent dynamics, leading to host evolution. There are various possible modes of inheritance and dominance for forms of host defense that could be relevant in our focal frog system; because we lack strong *a priori* expectations for inheritance and dominance, we model host evolution as clonal competition (n = 10 genotypes in our simulations). Each clonally reproducing genotype, j, may differ from other genotypes in fecundity or defense against pathogens (parameters with a j subscript are considered as differing by host genotype) (see Rees and Ellner 2016 for continuous trait evolution in an IPM instead). The total number of hosts of a genotype (N_j, summing across infection status and loads) reproduce with some maximum, per-capita fecundity r_j, incorporating a probability (m) of mutation that produces offspring of a different genotype with equal probability [$m/(n-1)$ for each]. Fecundity is reduced in a density-dependent manner parameterized by q in a function similar to the density dependence of fitness in the Ricker population model [note that birth rate of genotype j is a function of the density of all genotypes due to mutation and crowding from total host density, $N(t)$]. Thus, an isolated population of host genotype j without disease has a carrying capacity [$\ln(r_j/(1-s_0))/q$]. Host

genotypes may differ in the transmission parameter, β_j, survival probability at pathogen load x [with probability $s_j(x)$ that can vary by genotype by varying $\mu_{LD50,\,j}$] or pathogen growth, $G_j(x'|x)$ (Eqs. 1b, 1c; G_j can vary by genotype by varying a_j or b_j). Infected hosts of all genotypes produce environmental pathogens, Z (Eq. 1d), depending on their loads. We do not consider genotypic variation in other processes such as mutation rates, survival probability of uninfected hosts (s_0), recovery from infection (l), or pathogen shedding (λ). We parameterize this model with biologically reasonable, focal parameter values for our focal system (see Table A1 for full list of values), taking advantage of the ease of parameterizing IPMs. These biologically reasonable parameter values ground the numerical portion of our analyses of model behaviors as well as providing a central starting point for our parameter search of model behaviors.

<u>Eq. 1a: Birthrate of genotype *j*</u>

$$f_j\left(N_{1:n}(t)\right) = [\underbrace{r_j N_j(t)(1-m)}_{\text{Birth less mutation}} + \underbrace{\frac{m}{n-1}\sum_{k=1,\ k\neq j}^{n} r_k N_k(t)}_{\text{Gains from mutation}}]\,\underbrace{e^{-qN(t)}}_{\text{Crowding}}$$

<u>Eq. 1b: Susceptible individuals of genotype *j*</u>

$$S_j(t+1) = \underbrace{f_j\left(N_{1:n}(t)\right)}_{\text{Birth}} + \underbrace{S_j(t)s_0 e^{-\beta_j Z(t)}}_{S_j\text{ survive, remain uninfected}} + \underbrace{\int_{-\infty}^{\infty} I_j(x,t)s_j(x)l\,dx}_{I_j\text{ survive, recover}}$$

<u>Eq. 1c: Infected individuals of genotype *j* and load *x'*</u>

$$I_j(x',t+1) = \underbrace{S_j(t)s_0[1-e^{-\beta_j Z(t)}]G_0(x')}_{S_j\text{ survive, get infected}} + \underbrace{\int_{-\infty}^{\infty} I_j(x,t)s_j(x)(1-l)G_j(x'|x)\,dx}_{I_j\text{ survive, do not recover}}$$

<u>Eq. 1d: Environmental pathogens</u>

$$Z(t+1) = \underbrace{\lambda\int_{-\infty}^{\infty} e^{x}\sum_{j=1}^{n} I_j(x,t)\,dx}_{\text{Pathogen shedding}} + \underbrace{vZ(t)}_{\text{Environmental persistence}}$$

Derivation of the reduced IPM (rIPM: lognormal distribution of pathogen load)

To improve the analytical tractability and computational speed of this model, we use a second moment closure approximation. This approach assumes that the distribution of log load x on each genotype j is normally distributed with mean $\mu_j(t)$ and standard deviation $\sigma_j(t)$ that can vary with time. In our focal frog system, the empirical distribution of log load follows a normal distribution fairly closely and more than other typical distributions (Schrock 2024), empirically supporting this approximation. A similar approach has been applied before (Wilber et al. 2021) with a first moment closure that tracked changing mean over time but fixed variance and no host reproduction or evolution while we consider dynamic variance, host reproduction, and host evolution. Here instead of tracking the density of infected individuals with the full distribution of log load x for each genotype, $I_j(x, t)$ (as the full IPM does), we track the first and second moments of x, $P_j(t) = \int_{-\infty}^{\infty} x I_j(x,t)dx$ and $Q_j(t) = \int_{-\infty}^{\infty} x^2 I_j(x,t)dx$ and the total density of infected individuals of genotype j, $I_j(t)$. Thus, we reduce the number of dimensions in the model by assuming x is normally distributed and replacing a very large number of state variables (one for the density of hosts with each continuous value of x, i.e., technically infinite) with three state variables [$I_j(t)$, $P_j(t)$, and $Q_j(t)$]. The result is a host population described by a much smaller number of state variables equal to 4n+1 with n genotypes [$S_j(t)$, $I_j(t)$, $P_j(t)$, $Q_j(t)$, $Z(t)$]. These P and Q quantities are closely related to the mean and standard deviation of the log pathogens per host at time t: $\mu_j(t) = P_j(t)/I_j(t)$ and $\sigma_j^2(t) = Q_j(t)/I_j(t) - \mu_j^2(t)$ [assuming $I_j(t) > 0$, even if very barely positive as in some simulations]. This dimension-reducing approximation gives the reduced IPM (rIPM; see Appendix for derivations and Eq. 2 for results). This derivation is complicated by the non-linear survival function of load, $s(x)$ that drops precipitously near the

LD50; but we show that the derivation can still be performed even if there are two non-linear functions, say $s(x)$ and recovery function $l(x)$ that decline at different thresholds (see Appendix).

To express the rIPM more compactly, we introduce some useful notation. We use $\varphi(\omega)$ for the probability density of the standard normal distribution at ω [$\varphi(\omega)=(1/\sqrt{(2\pi)})\, e^{-(\omega\wedge 2)/2}$]. We also introduce a notation closely related to survival, $\hat{s}$ (Eq. 2a) and another that captures critical information about the distribution of load relative to the LD50 load for a given genotype [$\alpha_j(t)$ in Eq. 2b]. Together, $\hat{s}(\alpha_j(t))$ gives the average survival probability of an infected host of genotype j from time t to time t+1 after accounting for the distribution of pathogen load for that genotype; e.g., low mean load but very high standard deviation would still mean that nearly 50% of infected hosts have a high probability of death. Note that while the $\hat{s}$ function does involve an integral, it involves the cumulative distribution function of the standard normal, which is highly supported in terms of speed and accuracy in many programming languages, especially compared to the specific integrals in Eq. 1.

<u>Eq. 2a. Useful notation: a survival-related function that is 1 - the cumulative standard normal</u>

$$\hat{s}(z) = 1 - \int_{-\infty}^{z} \varphi(x)dx$$

<u>Eq. 2b. Useful notation: an expression of a genotype's pathogen load distribution relative to LD50 ($\mu_{\mathrm{LD50,\,j}}$)</u>

$$\alpha_j(t) = \frac{\mu_j(t) - \mu_{\mathrm{LD50},j}}{\sqrt{\sigma_j^2(t) + \sigma_{\mathrm{LD50}}^2}}$$

<u>Eq. 2c for susceptible individuals of genotype j</u>

$$S_j(t+1) = \underbrace{f_j\big(N_{1:n}(t)\big)}_{\text{Birth}} + \underbrace{S_j(t)s_0e^{-\beta_j Z(t)}}_{S_j\text{ survive, remain uninfected}} + \underbrace{I_j(t)\hat{s}(\alpha_j(t))l}_{I_j\text{ survive, recover}}$$

<u>Eq. 2d for infected individuals of genotype j</u>

$$I_j(t+1) = \underbrace{S_j(t)s_0[1-e^{-\beta_j Z(t)}]}_{S_j \text{ survive, get infected}} + \underbrace{I_j(t)\hat{s}(\alpha_j(t))(1-l)}_{I_j \text{ surive, do not recover}}$$

<u>Eq. 2e for first moment of log load on infected individuals of genotype j ("0th" etc. refer to terms of pathogen growth of different order with respect to pathogen density dependence, b_j)</u>

$$P_j(t+1) = \underbrace{S_j(t)s_0[1-e^{-\beta_j Z(t)}]\mu_0}_{S_j \text{ survive, get infected}} +$$

$$\underbrace{I_j(t)\hat{s}\left(\alpha_j(t)\right)(1-l)}_{I_j \text{ survive, do not recover}} \left\{ \underbrace{a_j}_{0th} + \underbrace{b_j\left[\mu_j(t) - \frac{\sigma_j^2(t)\varphi\left(\alpha_j(t)\right)}{\hat{s}\left(\alpha_j(t)\right)\sqrt{\sigma_j^2(t)+\sigma_{\mathrm{LD50}}^2}}\right]}_{1st} \right\}$$

<u>Eq. 2f for second moment of log load on infected individuals of genotype j</u>

$$Q_j(t+1) = \underbrace{S_j(t)s_0\left[1-e^{-\beta_j Z(t)}\right](\sigma_0^2+\mu_0^2)}_{S_j \text{ survive, get infected}} + \underbrace{I_j(t)\hat{s}\left(\alpha_j(t)\right)(1-l)}_{I_j \text{ survive, do not recover}} \{\underbrace{\sigma_G^2 + a_j^2}_{0th} +$$

$$\underbrace{2a_j b_j\left[\mu_j(t) - \frac{\sigma_j^2(t)\varphi\left(\alpha_j(t)\right)}{\hat{s}\left(\alpha_j(t)\right)\sqrt{\sigma_j^2(t)+\sigma_{\mathrm{LD50}}^2}}\right]}_{1st} +$$

$$\underbrace{b_j^2[(\sigma_j^2(t)+\mu_j^2(t)) - \frac{\sigma_j^2(t)\varphi\left(\alpha_j(t)\right)\left(2\mu_j(t)\sigma_{\mathrm{LD50}}^2 + \left(\mu_{\mathrm{LD50},j}+\mu_j(t)\right)\sigma_j^2(t)\right)}{\hat{s}\left(\alpha_j(t)\right)\left(\sigma_j^2(t)+\sigma_{\mathrm{LD50}}^2\right)^{\frac{3}{2}}}]}_{2nd}\}$$

<u>Eq. 2g: Environmental pathogens</u>

$$Z(t+1) = \underbrace{\lambda\sum_{j=1}^{n} I_j(t)e^{\mu_j(t)+\frac{\sigma_j^2(t)}{2}}}_{\text{Pathogen shedding}} + \underbrace{vZ(t)}_{\text{Environmental persistence}}$$

<u>Analytical insight into the behavior of mean and standard deviation of pathogen load</u>

With the second moment closure, we were able to derive expressions that give key insight into the dynamics of the distribution of pathogen load. We found implicit expressions for the equilibrium mean (μ^*_j) and standard deviation (σ^*_j) of log load for a genotype j (Eq. 3):

<u>Eq. 3a: Equilibrium mean (i.e., first moment) of log load</u>

$$\mu_j^* = \underbrace{[1 - \hat{s}(\alpha_j^*)(1-l)]}_{\text{Prob. inf is lost}} \underbrace{\mu_0}_{\text{Initial}} +$$

$$\underbrace{\hat{s}(\alpha_j^*)(1-l)}_{\text{Prob. inf remains}} \underbrace{\{a_j + b_j\left[\mu_j^* - \frac{\sigma_j^{*2}\varphi(\alpha_j^*)}{\hat{s}(\alpha_j^*)\sqrt{\sigma_j^{*2} + \sigma_{\text{LD}50}^2}}\right]\}}_{\text{Contribution of within-host growth}}$$

<u>Eq. 3b: Equilibrium second moment of log load</u>

$$\sigma_j^{*2} + \mu_j^{*2} = \underbrace{[1 - \hat{s}(\alpha_j^*)(1-l)]}_{\text{Prob.inf is lost}} \underbrace{(\sigma_0^2 + \mu_0^2)}_{\text{Initial}} + \underbrace{\hat{s}(\alpha_j^*)(1-l)}_{\text{Prob. inf remains}} \{$$

$$\underbrace{\sigma_G^2 + a_j^2 + 2a_j b_j\left[\mu_j^* - \frac{\sigma_j^{*2}\varphi(\alpha_j^*)}{\hat{s}(\alpha_j^*)\sqrt{\sigma_j^{*2} + \sigma_{\text{LD}50}^2}}\right]}_{\text{Contribution of within-host growth}}$$

$$\underbrace{+ b_j^2[(\sigma_j^{*2} + \mu_j^{*2}) - \frac{\sigma_j^{*2}\varphi(\alpha_j^*)\left(2\mu_j^*\sigma_{\text{LD}50}^2 + (\mu_{\text{LD}50,j} + \mu_j^*)\sigma_j^{*2}\right)}{\hat{s}(\alpha_j^*)\left(\sigma_j^{*2} + \sigma_{\text{LD}50}^2\right)^{\frac{3}{2}}}]\}}_{\text{Contribution of within-host growth continued}}$$

The expressions for the first (i.e., mean) and second moments of log load reflect the balance between the influences of new infection and infection growth on the host. The first moment is the sum of initial mean load weighted by the probability of infection loss plus the contribution of within-host growth (see terms in Eq. 3a that mirror terms in Eq. 2e) weighted by the probability that infection remains (i.e., an infected individual survives and does not recover). The expression for the equilibrium value of the second moment follows the same structure (Eq. 3b). If we

assume no loss of infection, we can find explicitly that $\mu_j^* = a/(1-b)$ and $\sigma_j^* = \sigma_G/\sqrt{(1-b)}$. Because there is loss of infection, we retain the somewhat more complex expression in Eqs. 3a and 3b.

We analyze these expressions with implicit differentiation to find how any parameter (y_k) affects μ^*_j and σ^*_j, considered as functions of that fixed parameter. We write the right hand side of Eq. 3a as a function, $\mu^*_j = F(\mu^*_j, \sigma^*_j, y_k)$, and rearrange Eq. 3b and take the positive square root of the new right hand side to write $\sigma^*_j = K(\mu^*_j, \sigma^*_j, y_k)$.

<u>Eq. 4a: Pathways affecting equilibrium mean log load</u>

$$\frac{d\mu_j^*}{dy_k} = \underbrace{1/[\left(1-\frac{\partial F}{\partial \mu_j^*}\right)\left(1-\frac{\partial K}{\partial \sigma_j^*}\right)-\frac{\partial F}{\partial \sigma_j^*}\frac{\partial K}{\partial \mu_j^*}]}_{\text{Feedback scaling}}\,[\underbrace{\left(1-\frac{\partial K}{\partial \sigma_j^*}\right)}_{\text{Feedback scaling}}\underbrace{\frac{\partial F}{\partial y_k}}_{\text{Direct}}+\underbrace{\frac{\partial F}{\partial \sigma_j^*}\frac{\partial K}{\partial y_k}}_{\text{Indirect}}]$$

<u>Eq. 4b: Pathways affecting equilibrium SD of log load</u>

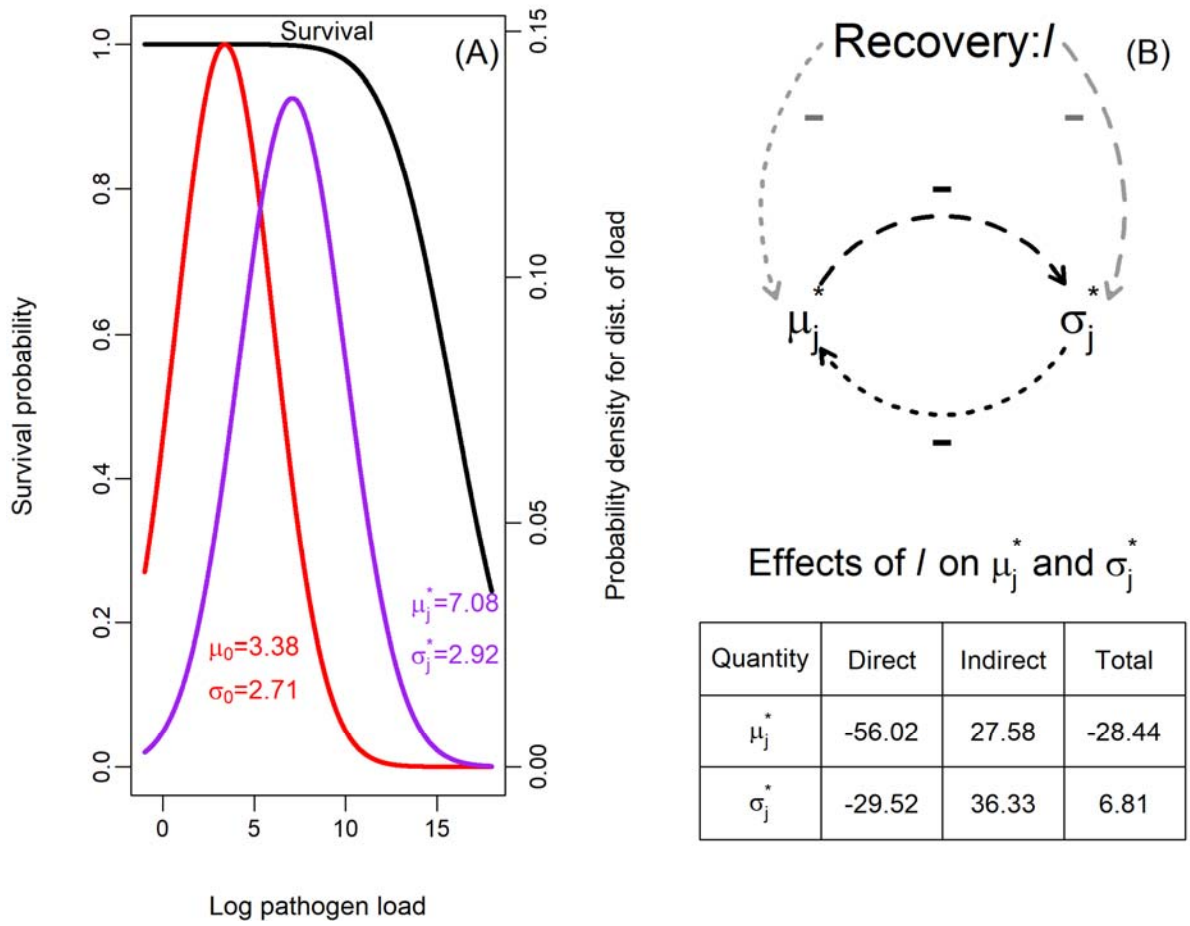
$$\frac{d\sigma_j^*}{dy_k} = \underbrace{1/[\left(1-\frac{\partial F}{\partial \mu_j^*}\right)\left(1-\frac{\partial K}{\partial \sigma_j^*}\right)-\frac{\partial F}{\partial \sigma_j^*}\frac{\partial K}{\partial \mu_j^*}]}_{\text{Feedback scaling}}\,[\underbrace{\left(1-\frac{\partial F}{\partial \mu_j^*}\right)}_{\text{Feedback scaling}}\underbrace{\frac{\partial K}{\partial y_k}}_{\text{Direct}}+\underbrace{\frac{\partial K}{\partial \mu_j^*}\frac{\partial F}{\partial y_k}}_{\text{Indirect}}]$$

As long as the terms we label "feedback scaling" remain positive (which they do across a broad parameter range, see the "Quantifying numerical robustness" section), we can interpret these expressions as how y_k affects μ^*_j directly ("direct" term in Eq. 4a) and indirectly by affecting σ^*_j ($\partial F/\partial \sigma^*_j\ \partial K/\partial y_k$ "indirect" term in Eq. 4a). The direct term, $\partial F/\partial y_k$, captures how a change in the parameter y_k directly changes the F function (where $F = \mu^*_j$). The indirect term comes from how y_k directly alters σ^*_j (i.e., $\partial K/\partial y_k$) multiplied by how a change in σ^*_j feeds back to change μ^*_j ($\partial F/\partial \sigma^*_j$). The same logic applies for direct and indirect effects on σ^*_j in Eq. 4b.

We find that pathogen growth and host death at high loads create negative feedbacks between μ^*_j and σ^*_j ($\partial F/\partial \sigma^*_j < 0$ and $\partial K/\partial \mu^*_j < 0$). These inequalities that represent negative feedbacks hold as long as the assumptions $\mu^*_j < \mu_{LD50,j}$, $\mu^*_j > \mu_0$, $\sigma^{*2}_j + \mu^{*2}_j > \sigma_0^2+\mu_0^2$, and $\mu^*_j - b_j\hat{s}(\alpha^*_j(t))(1-l)(b_j\mu^*_j+a_j)$ hold (see Appendix for proofs). For the most part, these assumptions are

biologically reasonable and likely to hold for many systems as they simply require that equilibrium mean log load is less than the LD50 (i.e., fewer than 50% of hosts die each time step) and that the first and second moments of log load are greater than those that correspond to initial infection (essentially guaranteed by pathogen growth within hosts). The last condition (which is sufficient but not necessary) is somewhat harder to interpret but essentially requires that the mean at equilibrium is larger than the growth of the mean ($b_j\mu^*_j+a_j$) reduced by recovery, death, and the density-dependence parameter (b); this condition is met at the wildtype's equilibrium. These negative feedbacks held across a broad parameter range (see the "Quantifying numerical robustness" section). Overall, we can interpret these negative feedbacks as arising largely because of growth of infection and the fact that the minority of hosts at equilibrium have lethal loads. Consider a perturbation that only *directly* increases μ^*_j; this increases the small fraction of high load hosts who are dying, cutting off more of those high loads and *indirectly* decreasing σ^*_j. Similarly, a perturbation that only *directly* increases σ^*_j also increases that small fraction of high load hosts who die, *indirectly* decreasing μ^*_j (Fig. 1A). Thus, a parameter that shifts both μ^*_j and σ^*_j also shifts their feedbacks; e.g., increased recovery rate (l) directly decreases μ^*_j and σ^*_j because both are constrained to be closer to initial infection levels; but lowering μ^*_j strongly reduces the survival-based constraints on σ^*_j, creating a net increase in σ^*_j (Fig. 1B). This example helps us understand how various parameters affect the negative feedbacks between μ^*_j and σ^*_j (Table 1). The signs of direct effects of these various parameters on μ^*_j and σ^*_j can be analytically guaranteed given biologically meaningful, sufficient but not always necessary conditions (see Appendix). Many of these conditions depend on the equilibrium, average probability of infected hosts surviving for one day, $\hat{s}(\alpha_j^*)$, not being too low; the most restrictive condition is $\hat{s}(\alpha_j^*) > 0.958$, which may be biologically reasonable given

that we find $\hat{s}(\alpha_j^*) = 0.984$ for our wild-type parameter values (see Table A1 for how biologically reasonable parameter values were found) and even higher for genotypes that evolve defense. Thus, we believe the signs of the direct effects that we found will hold in many biologically meaningful scenarios, giving us conceptual insight into load dynamics.



Quantity	Direct	Indirect	Total
μ_j^*	-56.02	27.58	-28.44
σ_j^*	-29.52	36.33	6.81

Figure 1. Load-dependent survival creates negative feedbacks between mean and standard deviation of log load. (A) The probability a host survives one day declines sharply with rising log pathogen load (black curve). The mean (μ_0) and standard deviation (σ_0) of pathogen load for the wild-type genotype are lower upon initial infection (red curve) than the equilibrium they arrive at (purple curve, μ^*_j and σ^*_j, is to the right and somewhat broader than red curve; $\mu_0 < \mu^*_j$, $\sigma_0 < \sigma^*_j$). (B) The survival curve imposes mutual negative feedbacks between μ^*_j and σ^*_j. A

perturbation that raises μ^*_j increases the minority of high load hosts who die quickly, shrinking the spread of the distribution (dashed black arrow). A perturbation that raises σ^*_j also leads to a larger minority of high load hosts dying quickly, shrinking the mean of the distribution (dotted black arrow). Thus, a change in the system (such as increasing recovery probability, l) changes μ^*_j and σ^*_j directly (gray arrows and quantified in column 2 of embedded table for parameter values corresponding to the wild-type genotype; e.g., for μ^*_j the direct effect is calculated as $\partial F/\partial l$ multiplied by the relevant feedback scaling terms in Eq. 4a) as well as indirectly through their feedbacks (for μ^*_j, this is $\partial F/\partial \sigma^*_j\ \partial K/\partial l$ multiplied by the relevant feedback scaling terms in Eq. 4a). The sum of a stronger negative direct effect and weaker positive indirect effect causes increasing l to decrease μ^*_j (column 3 of embedded table shows $d\mu^*_j/dl$ and $d\sigma^*_j/dl$); the sum of a weak negative direct effect and weak positive indirect effect causes increasing l to slightly increase σ^*_j (column 3).

Table 1. Analytical results for direct and numerical results for total effects of parameters on μ^*_j and σ^*_j. Superscripts in columns 2 and 3 denote conditions analytically found to be sufficient (though not always necessary) for that sign to hold (see Appendix for conditions and derivations); we assume the feedback scaling terms are always positive, as found in the numerical search. Columns 2 and 3 analytically determine the signs of the direct effects of parameter y_k on μ^*_j and σ^*_j respectively, given conditions. Columns 4 ($d\mu^*_j/dy_k$) and 5 ($d\sigma^*_j/dy_k$) quantify the total effects (direct effects+indirect effects, each multiplied by relevant feedback scalings from Eq. 4) found for the wild-type genotype in our focal parameter set (see Table A1 for derivation of parameter values). In our focal parameter set (and in all parameter sets we searched), the indirect effect on μ^*_j always has the opposite sign as the direct effect on σ^*_j, and

vice versa, due to mutual negative feedbacks (see Fig. 1). Parameters not listed do not affect μ^*_j and σ^*_j (e.g., note that β_j does not appear in Eqs. 3a, b).

Changing parameter y_k	**Direct effect on mean:** $\partial F/\partial y_k$	**Direct effect on SD:** $\partial K/\partial y_k$	**Total effect on mean** $d\mu^*_j/dy_k$	**Total effect on SD** $d\sigma^*_j/dy_k$
SD of load upon initial infection σ_0	0	≥0	-0.0729	0.0781
Mean load upon initial infection μ_0	≥0	≥0	0.205	-0.094
a_j contributes to pathogen growth rate at all loads	≥0	≥0[1, 2]	5.08	0.797
b_j contributes to pathogen	≥0[1, 2]	≥0[1, 2, 3, 4]	25.8	16.1

growth rate at high loads				
σ_{LD50} makes the survival function more gradual	≤ 0[1, 3, 5]	≤ 0[1, 3, 5, 6]	-0.642	0.0203
$\mu_{LD50,j}$ is the LD50 of the survival function	≥ 0[1, 5]	≥ 0[1, 3, 7]	0.424	0.0418
σ_G controls stochasticity in pathogen growth	0	≥ 0	-1.514	1.621
Recovery rate from infection *l*	≤ 0[5]	≤ 0[7]	-28.4	6.81

[1] $\mu^*_j < \mu_{LD50,j}$, i.e., infected host survival probability for one day is > 50%.

[2] $\mu^*_j > 0.8\sigma^*_j$, i.e., SD is not too much larger than mean.

[3] $\alpha^*_j < -1$, i.e., infected host survival probability for one day is > 84.1%.

[4] $\mu^*_j > 0.177\mu_{LD50,j}$, i.e., mean is not too much lower than LD50.

[5] $\mu^*_j > \mu_0$, i.e., infection grows so that mean log load is greater than the mean of new infections

[6] $\alpha^*_j < 3^{1/2}$, i.e., satisfied if infected host survival probability for one day is > 95.8%

[7] $\mu^{*2}_j + \sigma^{*2}_j > \mu^2_0 + \sigma^2_0$, i.e., infection grows so that the second moment of log load is greater than the second moment of new infections

Simulations of evolutionary recovery from disease-induced population decline

We modeled various forms of evolutionary recovery from disease-induced population decline to assess the insights our rIPM with dynamic variance can provide into how these processes affect the distribution of load. We initiated simulations of our rIPM with a wild-type host population at its disease-free, equilibrium density, no environmental pathogens, and low density of infected, wild-type hosts (1×10^{-6} m^{-3}) with a distribution of infection load corresponding to initial infection. Other genotypes were initiated with the same prevalence and initial distribution load distribution but at a total density lower by a factor of 1×10^{-6}. These other genotypes ($n = 10$) differed from the wild-type by a fitness advantage in terms of some pathogen defense trait (y_j ranging from the least defended, wild-type genotype, y_{wt}, to the most defended, $y_{defended}$); genotypes paid a fecundity cost of this defense $\{r_j = r_{wt}-(r_{wt} - r_{defended})[(y_{wt}-y_j)/(y_{wt}-y_{defended})]^w\}$. Thus, the wild-type genotype has the highest fecundity (r_{wt}); the fecundity cost per unit of defensive trait is small, at first, then accelerates rapidly driving a sudden decline in fecundity to the lowest fecundity of the most defended genotype ($r_{defended}$). These accelerating costs (accelerating because we chose $w > 1$, specifically, $w = 10$), generally select for an intermediate degree of host defense (Boots and Haraguchi 1999).

We considered cases of evolution of tolerance, avoidance resistance, constitutive resistance to pathogen growth, and inducible resistance to pathogen growth. For simplicity, genotypes were evenly spaced in the value of their defense traits. The magnitudes of variation in each defense trait was chosen to display evolutionary recovery of host abundance within 30-year simulations. Tolerance was modeled as higher LD50, allowing higher survival probability at the same pathogen loads [y_j ranging from $\mu_{\text{LD50, wt}}$ to $\mu_{\text{LD50, defended}} = \mu_{\text{LD50, wt}} + 4.05$]. Avoidance resistance was modeled as lower transmission rate to defended genotypes [y_j ranging from β_{wt} to $\beta_{\text{defended}} = 0.1\beta_{\text{wt}}$] that helps them avoid becoming infected. Resistance to pathogen growth reduces the within-host carrying capacity of pathogens [$a_j/(1-b_j)$]. Constitutive resistance lowered the growth rate of log load some degree regardless of current load [y_j ranging from a_{wt} to $a_{\text{defended}} = 0.1a_{\text{wt}}$], i.e., it was always on. Inducible resistance disproportionately lowered pathogen growth rate at high pathogen loads, also reducing the pathogen's with-host carrying capacity [y_j ranging from $1-b_{\text{wt}}$ to $1-b_{\text{defended}} = 2(1-b_{\text{wt}})$], i.e., inducible resistance takes some time to turn on after infection and thus is most likely to act on high load infections. We ran our simulations in R (R Core Team 2019) using the pomp package to simulate at higher speed by accessing the C language through Csnippets (King et al. 2024), see attached code for full details.

Analytical insights into the pathogen load distribution helps us understand how different forms of host evolution impact population-level patterns. Generally, tolerance-based defenses (Fig. 2) can be distinguished from resistance-based defenses (Figs. 3 and 4) by their effects on infection prevalence and/or the quantity of pathogens in the environment. Tolerance does not directly harm pathogen fitness, by definition, so it may increase prevalence and the quantity of pathogens in the environment, e.g., when highly infected hosts survive longer due to tolerance (Figs. 2B, F). This results in higher population mean and SD of log load (Figs. 2D, E) due to the

direct effects of tolerance on the mean and SD of individual genotypes (see Table 1 showing positive direct effects of higher $\mu_{LD50,\,j}$ on the μ^*_j and σ^*_j). In contrast, resistance evolution directly lowers pathogen fitness and will often decrease prevalence of infection and environmental pathogens. Further, different resistance mechanisms can have different effects on the distribution of pathogen load, highlighting the importance of considering various evolutionary strategies in the dynamics of pathogen load.

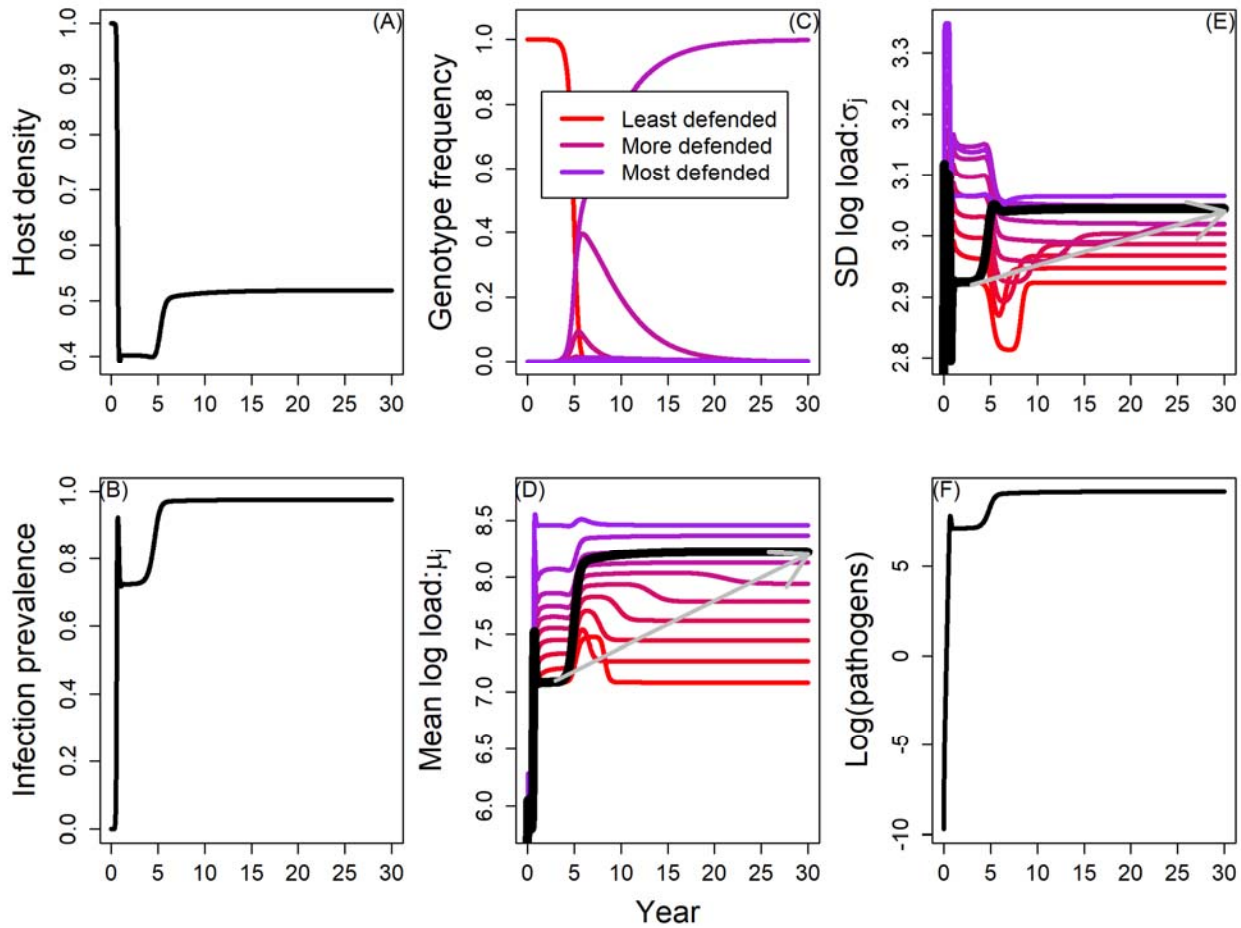


Figure 2. Tolerance (higher μ_{LD50}) raises mean and standard deviation (SD) of log load. (A) When pathogens invade, host density declines then recovers somewhat when hosts evolve tolerance. (B) Prevalence rises quickly with pathogen invasion then rises higher when hosts evolve higher tolerance. (C) Pathogen invasion leads to changes in host genotype frequencies as

selection favors genotypes defended with an intermediate degree of tolerance. (D) Higher tolerance corresponds to higher log loads (purple curves higher than red). As host populations evolve toward higher tolerance, mean log load rises (black curve), moving from the equilibrium level of wild type hosts to that of more tolerant hosts (difference emphasized by gray arrow). (E) More tolerant genotypes (purple curves) have higher SD of log load so that overall SD rises over time (black curve). (F) Pathogens in the environment rise with invasion then rise higher when hosts evolve tolerance.

While a reduction in prevalence and environmental pathogens can indicate host evolution of resistance, the mean and variance of log load can hold even more information. Avoidance resistance does not alter μ^*_j or σ^*_j and only slightly affects the transient dynamics of load (see that β_j does not appear in Eq. 3 and see Fig. A1). Stronger constitutive resistance (lower a_j) leads to lower μ^*_j (Fig. 3D) due to stronger, negative, direct effects compared to weaker, positive, indirect effects (see Table 1). Lower a_j slightly reduces σ^*_j (see ordering of colored curves in Fig. 3E) due to slightly stronger negative direct effects than positive indirect effects (see Table 1); however, as a_j decreases further, the positive indirect effects on σ^*_j become slightly stronger than the negative indirect effects, leading to a weak net increase in σ^*_j (see Fig. 3E and A2). Thus, constitutive resistance lowers mean log load but has a weak, mixed effect on the standard deviation, raising it. Inducible resistance (lower b_j), however, strongly decreases σ^*_j through strong direct effects (Fig. 4E), with all other responses being similar to constitutive (compare Figs. 3 and 4).

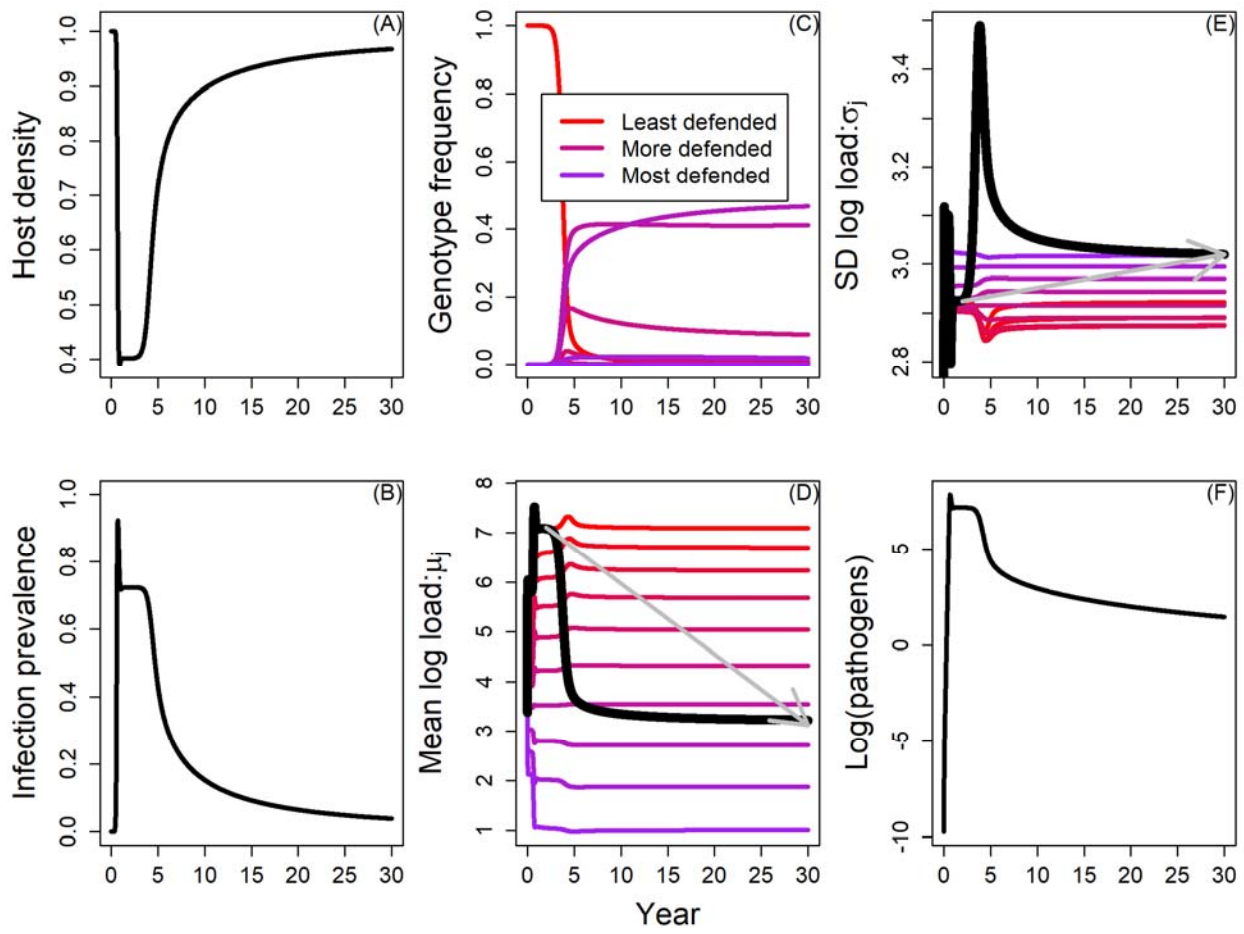


Figure 3. Constitutive resistance (lower a_j) lowers mean and raises standard deviation (SD) of log load. (A) When pathogens invade, host density declines then recovers when hosts evolve resistance. (B) Prevalence rises quickly with pathogen invasion then declines when hosts evolve resistance. (C) Pathogen invasion leads to changes in host genotype frequencies as selection favors genotypes defended with an intermediate degree of resistance. (D) Stronger constitutive resistance corresponds to lower log loads (purple curves lower than red). As host populations evolve toward stronger resistance, mean log load drops (black curve), moving from the equilibrium level of wild type hosts to that of more resistant hosts (difference emphasized by gray arrow). (E) More resistant genotypes (purple curves) have higher SD of log load so that overall SD (black curve) rises from near the equilibrium of the wild type to near the equilibrium

of defended genotypes (equilibrium difference emphasized by gray arrow). Further, there is a large, temporary spike in SD of log load during host evolution due to rising then falling genetic variance. (F) Pathogens in the environment rise with invasion then fall when hosts evolve resistance.

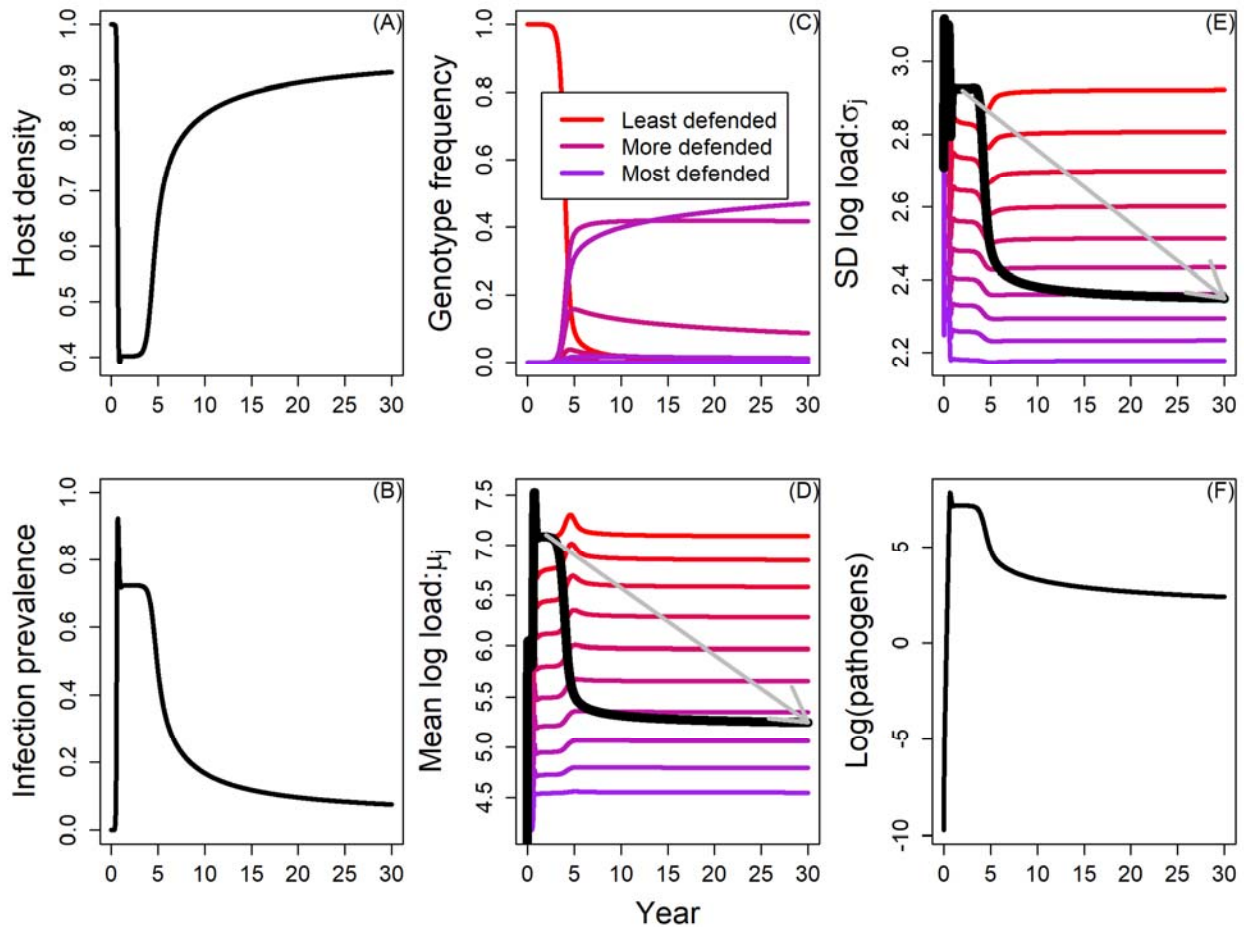


Figure 4. Inducible resistance (lower b_j, i.e., lower growth rate of high load infections) lowers both mean and standard deviation (SD) of log load. (A) When pathogens invade, host density declines then recovers when hosts evolve resistance. (B) Prevalence rises quickly with pathogen invasion then declines when hosts evolve resistance. (C) Pathogen invasion leads to changes in host genotype frequencies as selection favors genotypes defended with an

intermediate degree of resistance. (D) Stronger resistance corresponds to lower log loads (purple curves lower than red). As host populations evolve toward stronger inducible resistance, mean log load drops (black curve), moving from the equilibrium level of wild type hosts to that of more resistant hosts (equilibrium difference emphasized by gray arrow). (E) More resistant genotypes (purple curves) have lower SD of log load so that overall SD (black curve) drops from near the equilibrium of the wild type to near the equilibrium of defended genotypes (equilibrium difference emphasized by gray arrow). (F) Pathogens in the environment rise with invasion then fall when hosts evolve resistance.

The differences in the load dynamics of constitutive and inducible resistance hold potent implications for host fitness, pathogen fitness, and population-level disease spread. Consider, in general, the equilibrium mean and standard deviation of log loads for two host genotypes (*A* and *B*). Assume genotype *A* has a higher standard deviation ($\sigma^*_A > \sigma^*_B$); because mortality falls disproportionately on high load infections (assuming $\mu^*_A < \mu_{\mathrm{LD50}}$ and $\mu^*_B < \mu_{\mathrm{LD50}}$, i.e., fewer than 50% of infected hosts die each day), the higher standard deviation of *A* contributes to higher death rates because more hosts have high loads for genotype *A*. So, to achieve the same fitness while infected (assuming equal fecundities), genotype *A* must reduce mean log load further than genotype *B* [$\sigma^*_A > \sigma^*_B$, $\mu^*_B < \mu_{\mathrm{LD50}}$, $\mu^*_A < \mu_{\mathrm{LD50}}$, and $\hat{s}(\alpha^*_A) = \hat{s}(\alpha^*_B)$ requires $\mu^*_A < \mu^*_B$]. Assuming equal host fitness while infected, these two genotypes will differ in their per-capita rates of pathogen shedding, an important component of pathogen fitness. As long as these assumptions that are reasonable for our focal system, and perhaps other systems, are not violated, then increasing σ^*_A, even with commensurately decreasing μ^*_A, leads to higher shedding rates per infected host [$\mu_A^*+\sigma_A^{*2}/2 > \mu_B^*+\sigma_B^{*2}/2$; violation would occur if pathogen-induced mortality is

induced only by very high loads (high μ_{LD50} and low σ_{LD50}) while actual loads are distributed tightly at low loads (low μ^*_B and low σ^*_B); see Appendix for proof]. Higher pathogen shedding rates contribute directly to a higher basic reproductive number of pathogens spreading through a susceptible host population, an important metric of pathogen fitness. The basic reproductive number is $R_0 = S^* s_0 \beta' / (1-v) * \lambda \exp(\mu^* + \sigma^{*2}/2) / (1- \hat{s}(\alpha^*)(1-l))$ (shown by Wilber et al. 2021 , adapted by us, and numerically re-verified by our simulations). These results are not an artifact of our log transformation of load; for example, even when considering load on the natural scale and pathogen shedding proportional to load, the survival of infected hosts should generally be a non-linear function of load and thus depend both on the mean and the standard deviation of load and differ from the effect of standard deviation on pathogen shedding rates. Thus, the interactions of host survival and different forms of host defense evolution, constitutive with higher standard deviation vs. inducible with lower standard deviation, have important implications for host and pathogen fitness.

Quantifying numerical robustness

We tested the robustness of our core results (Figs. 2-4) for evolution of tolerance, constitutive resistance, and inducible resistance using simulations of the rIPM and the IPM across varying parameter values. We constructed 1,000 parameter sets drawn from a Latin Hypercube (Carnell 2022) with each parameter in Table A1 varying $\pm$50% from its focal value. Simulations followed the same procedure outlined above (see “Simulations of evolutionary recovery from disease-induced population decline” section) except we allowed the wildtype to reach its endemic equilibrium with pathogens before allowing host evolution so that we could

clearly quantify the effects of host evolution on mean and SD of log load. For simulations of the full IPM, we binned load into a wide yet finely divided range (see attached code).

Our core patterns for the effects of host evolution on the mean and SD of log load were robust to parameter variation and the approximate nature of our rIPM model. Our $\pm$50% parameter variation led to a very broad range of parameter sets, as shown by the fact that R_0 was 10 in our focal parameter set but <1 or >100 in 60.6% of the parameter sets in our search. Among the parameter sets that gave reasonable R_0 (greater than 1 but less than 100), the strong majority maintained our core patterns for the effects of different types of host evolution (see column 2 of Table 2). Typically, constitutive resistance evolution lowered equilibrium mean log load and raised the standard deviation (row 1, column 2 of Table 2); in the parameter sets that were an exception to this pattern, constitutive resistance still always lowered mean log load but also lowered the standard deviation which is not surprising since the effect on the standard deviation is weak and mixed. Evolution of inducible resistance nearly always lowered mean log load and the standard deviation (row 2, column 2 of Table 2); in the exceptional parameter sets, inducible resistance either raised mean log load or raised the standard deviation of log load but never both. Tolerance generally raised mean and standard deviation of log load (row 3, column 2 of Table 2); in the exceptional parameter sets, tolerance usually lowered the standard deviation instead of raising it (in a small minority of exceptional parameter sets and only in the full IPM, tolerance lowered the mean and raised the standard deviation). For all forms of host evolution, exceptional parameter sets deviated due to different strengths of direct and indirect effects but the signs of feedback scaling quantities, direct effects, indirect effects, and the feedbacks between μ^* and σ^* always remained the same as in our focal parameter set, reinforcing our conclusions about the broad potential relevance of our results. These results depend on allowing

the SD of log load to fluctuate, creating negative feedbacks between the mean and the standard deviation; thus, these core, conceptual results could not be obtained by the previous rIPM which only considered the first moment of load, keeping the standard deviation fixed. Further, we were able to separate these results into direct and indirect effects on the mean of standard deviation thanks to the improved analytical tractability of the rIPM over the full IPM. Thus, the rIPM we present here provides advantages over previous rIPMs and the full IPM.

Numerically, similar results came from the full IPM compared to the rIPM approximation (compare columns 2 and 3 of Table 2) with a good correlation between the models for when core patterns held or did not (column 4). The biggest deviation between the two models arose from different rates of when the weak, mixed effect of constitutive resistance on the standard deviation would be net positive or net negative. Additionally, the equilibrium results were generally similar between the rIPM and IPM across parameter sets (comparing wildtype only endemic equilibria) with the highest median absolute relative distance (absolute relative distance: $|X_{\mathrm{IPM}} - Xr_{\mathrm{IPM}}|/|X_{\mathrm{IPM}}|$) found for prevalence and the density of environmental pathogens (median absolute relative distances: host density 6.45%, prevalence 17.38%, mean log load 0.40%, SD log load 1.82%, log environmental pathogens 10.53%). Prevalence and log environmental pathogens, which showed the biggest deviation, were nearly always somewhat lower in the IPM than in the rIPM. Altogether, these results show that the rIPM approximates the IPM reasonably well and reveals robust patterns for the effects of host evolution on the distribution of pathogen load.

Table 2. Robustness of evolutionary patterns across parameter sets. Results from evolutionary simulations in the 1000 parameter sets found in the Latin hypercube ranging $\pm$50% around the focal parameters. A parameter set conformed to the pattern if lower *a* led to lower

mean and higher SD of pathogen load (row 1), lower *b* led to lower mean and lower SD (row 2), and higher μ_{LD50} led to higher mean and SD (row 3; we show % of parameter sets that met each pattern). The correlation column denotes the correlation between a parameter set conforming to the pattern for the rIPM (0 or 1) and the IPM (0 or 1).

Type of evolution	**% of parameter sets in rIPM**	**% of parameter sets in IPM**	**Correlation of rIPM and IPM**
Resistance that lowers growth rates (lower *a*). Pattern: μ^* down, σ^* up	71.2%	53.7%	0.835
Resistance that lowers growth rate of high loads (lower *b*). Pattern: μ^* down, σ^* down	99.1%	97.8%	0.970
Tolerance of higher loads (higher μ_{LD50}). Pattern: μ^* up, σ^* up	85.7%	74.3%	0.910

Discussion

rIPMs offer a powerful approach to integrating biological complexity of pathogen load dynamics, analytical tractability, and empirical parameterization, providing valuable conceptual insights. We were able to incorporate empirically supported and parameterized functional forms, e.g., a non-linear dependence of infected host death rate on log pathogen load, and derive implicit, analytical expressions for the mean and standard deviation of log load at equilibrium. Importantly, we found that the steep decline in host survival at high loads leads to negative

feedbacks between the mean and the standard deviation. Thus, any change (e.g., a change in a host trait) that directly affects mean and standard deviation also indirectly affects them by shifting their feedbacks (Table 1). Disentangling these direct and indirect pathways clarifies how the distribution of log load changes, e.g., evolution of inducible resistance (lower a) directly lowers the standard deviation but can actually cause it to increase due to the sum of direct and indirect effects (Fig. 3E).

Improved understanding of load dynamics also helps us recognize how the stochastic processes governing the standard deviation of load can become a battleground for evolution by hosts and/or microparasites. Our model confirms the predictions from classic compartmental models that tolerance (higher μ_{LD50} in our model) increases prevalence and environmental pathogens (Råberg et al. 2008), confirms the clear intuition that tolerance should increase mean log loads, and illustrates with greater novelty that tolerance increases the standard deviation of log load (Fig. 2). Resistance mechanisms, on the other hand, should decrease pathogen fitness (Råberg et al. 2008), as we found. We were able to differentiate resistance mechanisms even further. We found that constitutive resistance had a weak, mixed effect on the standard deviation, whereas inducible resistance clearly decreased it (Figs. 3 and 4). The higher standard deviation associated with constitutive resistance increases the likelihood of exceptionally high load infections, leading to host death. Even assuming compensatory changes in mean log load so that constitutive and inducible resistance led to the same death rate of infected hosts, we found that constitutive resistance should lead to higher pathogen shedding per infected host, increasing pathogen fitness. Previous compartmental models of constitutive vs. inducible resistance have considered how the two forms of resistance may differ in whether their costs and benefits are expressed while susceptible or infected (Boots and Best 2018; Westra et al. 2015); but we

assume both forms of resistance have constant costs and benefits while infected yet we still find they have different implications for the standard deviation, host fitness, and pathogen fitness. The implications of a dynamic distribution of pathogen load for pathogen fitness should feedback on disease spread and even, possibly, coevolution by pathogens.

Our work builds toward a greater understanding of how load dynamics shape host and microparasite fitness (and thus perhaps coevolution) at the within- and between-host scales. Previously, nested models have demonstrated that within-host dynamics, like pathogen growth, influence tradeoffs between transmission rate and virulence, impacting pathogen evolution at different epidemic stages (Day et al. 2011). And a previous rIPM showed that faster within-host growth can increase pathogen fitness through higher shedding or decrease it if host mortality becomes too high (Wilber et al. 2021). Our work emphasizes the importance of a dynamic standard deviation of load for both host and pathogen fitness.

These results suggest new avenues for studying host-pathogen coevolution. For example, evolution by host or microparasite will affect how strongly stochastic mechanisms can generate heterogeneity in load, altering the standard deviation that controls the probability of very high load infections that shed at high rates (good for pathogens) and die quickly (bad for hosts). Further, an rIPM could provide an ideal model for investigating more fully the likely critical feedbacks between the within-host and between-host levels, a criterion that makes such models particularly vital (Mideo et al. 2008). For example, high force of infection from the environment may increase the mean load of a new infection (would create a $\beta_j Z$ dependence of our G_0 function and allow avoidance resistance to alter equilibrium load dynamics). Further, pathogen growth rates may depend critically on dose (as dose has been found to influence many within-host dynamics: Kamiya et al. 2020) such that low dose infections are eliminated by hosts but

high dose infections grow out of control (we could readily model this by modifying our $a+bx$ growth parameters to have $a < 0$, $b > 1$). Further empirical work could determine the relevance of such functional forms while further theoretical work could determine their implications, e.g., possible bistability between endemic and disease-free equilibria.

In addition to a conceptual advance, our rIPM also provides a methodological advance for certain applications to merge theory and data. We demonstrate that investigators may be able to use the load dynamics of the mean *and* standard deviation to infer underlying processes, such as which form of defense hosts are evolving. Additionally, the reduced dimension of our rIPM, compared to the full IPM, will improve the efficiency of fitting a model to population-level data (Bolker and Pacala 1997) including host abundance, infection prevalence, and load dynamics. Because rIPMs link within-host and population-level processes, such a fitting process may be very useful for combining multiple data types. For example, laboratory data may provide informative prior probability distributions for some parameters, improving the ability of a model fitting process to infer other within-host processes from population-level data. In our focal system, Mountain yellow-legged frogs are likely evolving some form of resistance to pathogen growth (Briggs et al. 2010; Knapp et al. 2016; Knapp et al. 2023); ample data are collected on abundance, infection prevalence, and the distribution of *Bd* load in recovering, wild populations so that fitting an rIPM to these data may identify whether this resistance is more likely constitutive or inducible, helping target for future immunological investigations. Our model also serves as a resource for investigators interested in lognormally distributed pathogens as well as a blueprint for how investigators may derive helpful models for microparasites that follow other, continuous distributions (as measured, e.g., by qPCR). Further, our moment closure approach could be applied to track dynamic mean and variance of other lognormally distributed, ecological

variables outside of disease systems (e.g., body size), and is particularly useful when there are non-linear transition probabilities (e.g., survival) (see Stubberud et al. 2019 for an example of lognormal body size and non-linear survival).

Overall, our model results emphasize the value of considering load dynamics for critical, ongoing questions of microparasite diseases. Previous theory and theory-data integration have effectively utilized the mean of infection load, providing a strong theoretical foundation and many datasets to leverage when considering a dynamic standard deviation of infection load. rIPMs provide computationally efficient, empirically relevant, and analytically tractable tools for advancing our understanding of load dynamics for microparasites. Further, we can address how classic aggregation metrics change with these models to see how microparasite dynamics follow, or deviate from, patterns found for macroparasites. Improving our understanding of load dynamics will allow us to infer within-host processes more efficiently and build a more complete understanding of (co-)evolution of host-pathogen systems.

References

Acevedo, M. A., F. P. Dillemuth, A. J. Flick, M. J. Faldyn, and B. D. Elderd. 2019. Virulence-driven trade-offs in disease transmission: A meta-analysis. Evolution 73:636-647.

Adler, F., and M. Kretzschmar. 1992. Aggregation and stability in parasite—host models. Parasitology 104:199-205.

Adler, F. R., and R. Karban. 1994. Defended fortresses or moving targets? Another model of inducible defenses inspired by military metaphors. The American Naturalist 144:813-832.

Anderson, R., and D. Gordon. 1982. Processes influencing the distribution of parasite numbers within host populations with special emphasis on parasite-induced host mortalities. Parasitology 85:373-398.

Anderson, R. M., and R. M. May. 1978. Regulation and stability of host-parasite population interactions: I. Regulatory processes. The journal of animal ecology:219-247.

Anderson, R. M., and R. M. May. 1981. The population dynamics of microparasites and their invertebrate hosts. Philosophical Transactions of the Royal Society of London. B, Biological Sciences 291:451-524.

Beldomenico, P. M. 2020. Do superspreaders generate new superspreaders? A hypothesis to explain the propagation pattern of COVID-19. International Journal of Infectious Diseases 96:461-463.

Bolker, B., and S. W. Pacala. 1997. Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. Theoretical population biology 52:179-197.
Bonhoeffer, S., C. Fraser, and G. E. Leventhal. 2015. High heritability is compatible with the broad distribution of set point viral load in HIV carriers. PLoS pathogens 11:e1004634.
Boots, M., and A. Best. 2018. The evolution of constitutive and induced defences to infectious disease. Proceedings of the Royal Society B-Biological Sciences 285.
Boots, M., and Y. Haraguchi. 1999. The evolution of costly resistance in host-parasite systems. American Naturalist 153:359-370.
Briggs, C. J., R. A. Knapp, and V. T. Vredenburg. 2010. Enzootic and epizootic dynamics of the chytrid fungal pathogen of amphibians. Proceedings of the National Academy of Sciences 107:9695-9700.
Carnell, R. 2022. lhs: Latin Hypercube Samples. R package version 1.1.6.
Coulson, T. 2012. Integral projections models, their construction and use in posing hypotheses in ecology. Oikos 121:1337-1350.
Day, T., S. Alizon, and N. Mideo. 2011. Bridging scales in the evolution of infectious disease life histories: theory. Evolution: International Journal of Organic Evolution 65:3448-3461.
Euser, S., S. Aronson, I. Manders, S. van Lelyveld, B. Herpers, J. Sinnige, J. Kalpoe et al. 2021. SARS-CoV-2 viral-load distribution reveals that viral loads increase with age: a retrospective cross-sectional cohort study. International journal of epidemiology 50:1795-1803.
Fowler, A., and T. Déirdre Hollingsworth. 2016. The dynamics of Ascaris lumbricoides infections. Bulletin of mathematical biology 78:815-833.
Gilchrist, M. A., and A. Sasaki. 2002. Modeling host–parasite coevolution: a nested approach based on mechanistic models. Journal of Theoretical Biology 218:289-308.
Grear, D. A., and P. Hudson. 2011. The dynamics of macroparasite host-self-infection: a study of the patterns and processes of pinworm (Oxyuridae) aggregation. Parasitology 138:619-627.
Grodio, J. L., D. M. Hawley, E. E. Osnas, D. H. Ley, K. V. Dhondt, A. A. Dhondt, and K. A. Schat. 2012. Pathogenicity and immunogenicity of three Mycoplasma gallisepticum isolates in house finches (Carpodacus mexicanus). Veterinary Microbiology 155:53-61.
Grogan, L. F., J. Robert, L. Berger, L. F. Skerratt, B. C. Scheele, J. G. Castley, D. A. Newell et al. 2018. Review of the amphibian immune response to chytridiomycosis, and future directions. Frontiers in immunology 9:2536.
Kamiya, T., M. A. Greischar, D. S. Schneider, and N. Mideo. 2020. Uncovering drivers of dose-dependence and individual variation in malaria infection outcomes. PLOS Computational Biology 16:e1008211.
King, A. A., E. L. Ionides, C. M. Bretó, S. P. Ellner, M. J. Ferrari, S. Funk, S. G. Johnson et al. 2024. pomp: Statistical Inference for Partially Observed Markove Proccesses. (R package, version 5.6).
King, K. C., and M. B. Bonsall. 2017. The evolutionary and coevolutionary consequences of defensive microbes for host-parasite interactions. BMC evolutionary biology 17:1-12.
Knapp, R. A., G. M. Fellers, P. M. Kleeman, D. A. Miller, V. T. Vredenburg, E. B. Rosenblum, and C. J. Briggs. 2016. Large-scale recovery of an endangered amphibian despite ongoing exposure to multiple stressors. Proceedings of the National Academy of Sciences 113:11889-11894.
Knapp, R. A., M. Q. Wilber, A. Q. Byrne, M. B. Joseph, T. C. Smith, A. P. Rothstein, R. L. Grasso et al. 2023. Reintroduction of resistant frogs facilitates landscape-scale recovery in the presence of a lethal fungal disease. bioRxiv:2023.2005. 2022.541534.
Kretzschmar, M., and F. Adler. 1993. Aggregated distributions in models for patchy populations. Theoretical population biology 43:1-30.
Lanciani, C. 1975. Parasite-induced alterations in host reproduction and survival. Ecology 56:689-695.

May, R. M., and R. M. Anderson. 1978. Regulation and stability of host-parasite population interactions: II. Destabilizing processes. The Journal of Animal Ecology:249-267.

Merow, C., J. P. Dahlgren, C. J. E. Metcalf, D. Z. Childs, M. E. Evans, E. Jongejans, S. Record et al. 2014. Advancing population ecology with integral projection models: a practical guide. Methods in Ecology and Evolution 5:99-110.

Metcalf, C. J. E., A. L. Graham, M. Martinez-Bakker, and D. Z. Childs. 2016. Opportunities and challenges of Integral Projection Models for modelling host–parasite dynamics. Journal of Animal Ecology 85:343-355.

Mideo, N., S. Alizon, and T. Day. 2008. Linking within-and between-host dynamics in the evolutionary epidemiology of infectious diseases. Trends in Ecology & Evolution 23:511-517.

Moller, A. P. 1990. Effects of parasitism by a haematophagous mite on reproduction in the barn swallow. Ecology 71:2345-2357.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Råberg, L., A. L. Graham, and A. F. Read. 2008. Decomposing health: tolerance and resistance to parasites in animals. Philosophical Transactions of the Royal Society B: Biological Sciences 364:37-49.

Rees, M., and S. P. Ellner. 2016. Evolving integral projection models: evolutionary demography meets eco-evolutionary dynamics. Methods in Ecology and Evolution 7:157-170.

Rollins-Smith, L. A. 2020. Global amphibian declines, disease, and the ongoing battle between Batrachochytrium fungi and the immune system. Herpetologica 76:178-188.

Sánchez, C. A., D. J. Becker, C. S. Teitelbaum, P. Barriga, L. M. Brown, A. A. Majewska, R. J. Hall et al. 2018. On the relationship between body condition and parasite infection in wildlife: a review and meta-analysis. Ecology letters 21:1869-1884.

Schrock, S. 2024. Do fungi look like macroparasites? Quantifying the patterns and mechanisms of aggregation for a host-fungal parasite relationship. BioRxiv.

Shudo, E., and Y. Iwasa. 2001. Inducible defense against pathogens and parasites: optimal choice among multiple options. Journal of Theoretical Biology 209:233-247.

Stubberud, M. W., Y. Vindenes, L. A. Vøllestad, I. J. Winfield, N. C. Stenseth, and Ø. Langangen. 2019. Effects of size-and sex-selective harvesting: An integral projection model approach. Ecology and Evolution 9:12556-12570.

Vredenburg, V. T., R. A. Knapp, T. S. Tunstall, and C. J. Briggs. 2010. Dynamics of an emerging disease drive large-scale amphibian population extinctions. Proceedings of the National Academy of Sciences 107:9689-9694.

Walsman, J. C., M. J. Janecka, D. R. Clark, R. D. Kramp, F. Rovenolt, R. Patrick, R. S. Mohammed et al. 2022. Shoaling guppies evade predation but have deadlier parasites. Nature Ecology & Evolution:1-10.

Westra, E. R., S. van Houte, S. Oyesiku-Blakemore, B. Makin, J. M. Broniewski, A. Best, J. Bondy-Denomy et al. 2015. Parasite exposure drives selective evolution of constitutive versus inducible defense. Current biology 25:1043-1049.

Wilber, M. Q., C. J. Briggs, and P. T. Johnson. 2020. Disease's hidden death toll: Using parasite aggregation patterns to quantify landscape-level host mortality in a wildlife system. Journal of Animal Ecology.

Wilber, M. Q., K. E. Langwig, A. M. Kilpatrick, H. I. McCallum, and C. J. Briggs. 2016a. Integral projection models for host–parasite systems with an application to amphibian chytrid fungus. Methods in Ecology and Evolution 7:1182-1194.

Wilber, M. Q., F. Pfab, M. E. Ohmer, and C. J. Briggs. 2021. Integrating infection intensity into within-and between-host pathogen dynamics: Implications for invasion and virulence evolution. The American Naturalist 198:661-677.

Wilber, M. Q., S. B. Weinstein, and C. J. Briggs. 2016b. Detecting and quantifying parasite-induced host mortality from intensity data: method comparisons and limitations. International Journal for Parasitology 46:59-66.

Wilson, K., O. Bjørnstad, A. Dobson, S. Merler, G. Poglayen, S. Randolph, A. Read et al. 2002. Heterogeneities in macroparasite infections: patterns and processes, Pages 6-44 *in* P. J. R. Hudson, A.

Grenfell, B.T.

Heesterbeek, H.

Dobson, A., ed. The ecology of wildlife diseases, Oxford University Press.

Wu, N. C. 2023. Pathogen load predicts host functional disruption: A meta-analysis of an amphibian fungal panzootic. Functional Ecology 37:900-914.

Yang, J.-R., J. Lo, Y.-L. Ho, H.-S. Wu, and M.-T. Liu. 2011. Pandemic H1N1 and seasonal H3N2 influenza infection in the human population show different distributions of viral loads, which substantially affect the performance of rapid influenza tests. Virus research 155:163-167.

Yang, Q., T. K. Saldi, P. K. Gonzales, E. Lasda, C. J. Decker, K. L. Tat, M. R. Fink et al. 2021. Just 2% of SARS-CoV-2– positive individuals carry 90% of the virus circulating in communities. Proceedings of the National Academy of Sciences 118:e2104547118.