ORIGINAL ARTICLE

BRIDGING SCALES IN THE EVOLUTION OF INFECTIOUS DISEASE LIFE HISTORIES: THEORY

Troy Day,[1,2,3] **Samuel Alizon,**[4] **and Nicole Mideo**[5]

[1]*Department of Mathematics and Statistics, Jeffery Hall, Queen's University, Kingston, ON, K7L 3N6, Canada*

[2]*E-mail: tday@mast.queensu.ca*

[3]*Department of Biology, Queen's University Kingston, ON, K7L 3N6, Canada*

[4]*Laboratoire MIVEGEC (UMR CNRS 5290, IRD 224, UM1, UM2) 911 avenue Agropolis B.P. 64501 34394 Montpellier Cedex 5, France*

[5]*Centre for Immunity, Infection, and Evolution School of Biological Sciences, University of Edinburgh Edinburgh EH9 3JT, United Kingdom*

A significant goal of recent theoretical research on pathogen evolution has been to develop theory that bridges within- and between-host dynamics. The main approach used to date is one that nests within-host models of pathogen replication in models for the between-host spread of infectious diseases. Although this provides an elegant approach, it nevertheless suffers from some practical difficulties. In particular, the information required to satisfactorily model the mechanistic details of the within-host dynamics is not often available. Here, we present a theoretical approach that circumvents these difficulties by quantifying the relevant within-host factors in an empirically tractable way. The approach is closely related to quantitative genetic models for function-valued traits, and it also allows for the prediction of general characteristics of disease life history, including the timing of virulence, transmission, and host recovery. In a companion paper, we illustrate the approach by applying it to data from a model system of malaria.

KEY WORDS: Nested models, pathogen, pathogenesis, transmission-virulence tradeoff, virulence.

Studies of the dynamics of infectious diseases are complicated by the fact that there are at least two different scales on which important processes occur (Levin and Pimentel 1981; Levin and Bull 1994; Nowak and May 1994; van Baalen and Sabelis 1995; May and Nowak 1995; Frank 1996). First, pathogen replication and interactions with host resources and host defense mechanisms occur within individual hosts. These dynamics govern the extent to which the pathogen is transmissible, as well as the period of time over which the infection persists. Second, pathogens spread between hosts. It is the dynamical processes occurring at the population level that govern whether the pathogen can persist in the host population as a whole. These two levels of biological organization are clearly linked, because the within-host dynamics strongly influence the extent of pathogen transmission between hosts. In some cases the between-host dynamics can feed back to affect the within-host dynamics as well (Mideo et al. 2008).

A significant goal of recent theoretical research on pathogen evolution has been to develop theory that bridges these two scales of biological dynamics. From a mathematical standpoint, there is a relatively straightforward way in which this can be done (reviewed in Mideo et al. 2008). First an epidemiological model that tracks age of infection is developed. Models that keep track of infection age allow transmission rate, pathogen-induced mortality rate, and host recovery rate to vary over the course of an infection (Hethcote 2000). Then, a model for the within-host dynamics of the pathogen is developed, and this model is nested within the between-host epidemiological model by specifying how the

within-host dynamics of the pathogen, host defense mechanisms, and/or host resources affect the transmission, virulence, and recovery rates as a function of infection age (Sasaki and Iwasa 1991; Day 2001; Gilchrist and Sasaki 2002; Ganusov et al. 2002; André et al. 2003; Alizon and van Baalen 2005; André and Gandon 2006; Gilchrist and Coombs 2006; Coombs et al. 2007; Alizon and van Baalen 2008; Mideo et al. 2008).

The above approach yields a model that links the dynamics occurring at each of the two relevant scales, and such nested models can then be used to model pathogen evolution. The most commonly used technique for this purpose is an evolutionary invasion analysis. Once a resident pathogen has reached some form of endemic equilibrium, a mutant pathogen strain is introduced that differs from the resident in some the parameters affecting its within-host dynamics. This difference in parameter values will result in a pattern of transmission, virulence, and recovery during the infection that differs from that of the resident. As a result, the mutant pathogen strain will thereby display different epidemiological dynamics at the level of the host population, and this will determine its ability to invade. The primary goal of this analysis is to determine the evolutionarily stable pathogen strain.

The above approach is a very elegant theoretical solution for bridging scales of dynamics in models of pathogen evolution, but it suffers from some practical difficulties. Chief among these is that the information required to employ this approach is typically not available for the vast majority of infectious diseases of interest. For most pathogens, the mechanistic details of the within-host dynamics, and how genetic variation among pathogen strains affects these dynamics, are simply not yet known. Consequently, it is usually not yet possible to build the appropriate within-host model that is required.

In this article, we consider how to circumvent the above problem, and develop theory that bridges these two scales of biological organization without requiring a mechanistic understanding of the within-host dynamics. One main interest in models such as those described above is to make predictions about the evolution of disease life histories (i.e., the pattern of transmission, mortality, and recovery as a function of infection age; Day 2003). Mathematically, this amounts to making predictions about the evolution of function-valued traits (Kirkpatrick and Heckman 1989; Dieckmann et al. 2006; Parvinen et al. 2006). Roughly speaking, in the absence of coinfection, the epidemiological dynamics at the level of the host population determines the strength and direction of selection on these function-valued traits. The details of the within-host dynamics, and how genetic variation among pathogen strains affects these dynamics, then determines the constraints on the kinds of functions that can evolve. Our approach is motivated by that of Kirkpatrick and Heckman (1989) in quantitative genetics. Specifically, rather than characterizing these constraints

in terms of the mechanistic details of the within-host dynamics, we show how they can be adequately characterized via genetic covariance functions measured at the level of the host. In a companion paper, we illustrate the application of this approach to data from a model system of malaria (Mideo et al. 2011).

Development of the Modeling Approach

Our model combines ideas from mathematical ecology (Metz and Diekmann 1986; Gyllenberg and Hanski 1992) with results that come from applying concepts from theoretical population genetics to epidemiological models (Day and Proulx 2004; Day and Gandon 2006, 2007). Throughout this article, we suppose that each host can be infected by, at most, a single pathogen strain. Consequently, the results cannot be used to model the evolution of diseases such as HIV, for which infection with multiple genotypes and subsequent within-host evolution is common. In a future article, we will extend the results to account for multiple infections.

The approach presented here can be adapted to a variety of epidemiological settings but we focus on a very simple model to illustrate the technique. Suppose we have a single class of infected individuals that is structured by age of infection. Specifically, we use $I_i(a, t)$ to denote the number of hosts who are infected with strain i, and whose infection began at time $t - a$ (i.e., a denotes the "age of infection"). The quantities $I_i(a, t)$ will change through time, and quite generally we model these dynamics using the Mckendrick-von Foerster equation (McKendrick 1926; Hoppensteadt 1974) as

$$\frac{\partial I_i(a, t)}{\partial t} = -\frac{\partial I_i(a, t)}{\partial a} - D_i(a, t)I_i(a, t) \qquad (1)$$

with boundary condition $I_i(0, \ t) \ = \ \int_0^\infty B_i(a,t)I_i(a,t)\mathrm{d}a$ (Appendix A). Here $B_i(a, t)$ is the per capita rate at which new infections with strain i are created, and $D_i(a, t)$ is the rate at which such infections end, either through host death or recovery. Both depend on the age of infection, and they might also depend explicitly on time. For example, the function $B_i(a, t)$ will typically depend explicitly on time in epidemiological models because the rate of generation of new infections will depend on the density of susceptible hosts, and this will typically vary through time according to the epidemiological dynamics.

Now suppose that both B_i and D_i are determined by a function-valued trait that is expressed in infected individuals, and denote this trait by $\tau_i(a)$ (Appendix A). For concreteness, suppose that $\tau_i(a)$ represents the density of pathogen within the host, as a function of infection age. Within-host pathogen density will often affect pathogen transmission rate between hosts as well as host mortality and/or recovery rate, and therefore it is reasonable to expect that both B and D might be functions of τ; that is, $B(a, t) = b(\tau(a), t)$ and $D(a, t) = d(\tau(a), t)$ for some functions $b(\tau, t)$ and

$d(\tau, t)$. Strains will differ in their function τ, and thus they will differ in transmission rate, b, and mortality/recovery rate, d. We note that this formulation also allows for some forms of host heterogeneity and other factors that can influence B and D (e.g., stochasticity in inoculum size), in addition to pathogen genotype. For simplicity, we ignore these effects in the main text, but Appendix E illustrates how they enter into the calculations. Roughly speaking, in such cases we can simply view B_i and D_i as the average birth and death rates of infections with genotype i, where the average is calculated over the distribution of heterogeneity in question.

Defining $q_i(a, t)$ as the fraction of hosts infected with strain i that are of infection age a at time t, the evolution of the mean value of τ, as a function of infection age, is described by (Appendix B)

$$\dot{\bar{\tau}}(a, t) = \text{cov}[\tau_i(a), \tilde{r}_i(t)]. \tag{2}$$

Here

$$\tilde{r}_i(t) = \int_0^\infty b(\tau_i(s), t)q_i(s, t)\mathrm{d}s - \int_0^\infty d(\tau_i(s), t)q_i(s, t)\mathrm{d}s, \tag{3}$$

which is the per capita rate of change of hosts infected with strain i at time t, averaged over all infection ages (i.e., the fitness of strain i at time t). Throughout we use a "tilde" to denote an average over infection ages, and an "overbar" to denote an average over strains.

Equation (2) reveals the familiar result (Price 1970) that the rate of evolutionary change of the mean value of any trait (in this case the mean within-host pathogen density at infection age a; i.e., $\bar{\tau}(a, t)$) is determined by the covariance between the trait and fitness. Using equation (3), we can write equation (2) more explicitly as

$$\dot{\bar{\tau}}(a, t) = \int_0^\infty G_{\tau, bq}(a, s; t)\mathrm{d}s - \int_0^\infty G_{\tau, dq}(a, s; t)\mathrm{d}s, \tag{4}$$

where $G_{x,y}(a, s; t) = \text{cov}[x_i(a), y_i(s)]$ is the covariance between function x at infection age a and function y at infection age s over the distribution of all pathogen strains in the population at time t. Thus, the evolutionary dynamics of the average within-host pathogen density at infection age a (i.e., $\bar{\tau}(a, t)$) is composed of two parts: (1) the covariance between the within-host pathogen density at infection age a and the realized production of new infections at infection age s, $b(\tau(s), t)q(s, t)$, summed over all infection ages, s; (2) the covariance between the within-host pathogen density at infection age a and the realized loss rate of infected hosts at infection age s, $d(\tau(s), t)q(s, t)$, summed over all infection ages s.

Equation (4) applies quite generally, but considerably more progress can be made if we restrict attention to the case where the distribution of infection ages stabilizes quickly relative to the time scale on which the vital rates, $B(a, t)$ and $D(a, t)$, change.

This would require, for example, that the duration of an individual infection is short relative to the timescale on which these rates change. In this case, the time dynamics of $q_i(s, t)$ can be expected to reach a quasi-equilibrium value before much change in either $B(a, t)$ or $D(a, t)$ occurs, and we can approximate the per capita rate of change given in (3) by substituting in this quasi-equilibrium infection age distribution (Lande 1982; Kirkpatrick and Lofsvold 1989, 1992). Denoting this approximate value for strain i by $\rho[\tau_i](t)$, equation (3) then simplifies to the Euler–Lotka equation (Appendix C)

$$1 = \int_0^\infty e^{-a\rho[\tau_i](t)} e^{-\int_0^a d(\tau_i(s), t)\mathrm{d}s} b(\tau_i(a), t)\mathrm{d}a. \tag{5}$$

The square bracket notation $\rho[\tau_i](t)$ indicates that ρ is a scalar-valued operator on a function space (i.e., a functional), which is implicitly defined by equation (5). Also notice that, although the per capita rate of change is approximated using the quasi-equilibrium values of $q_i(s, t)$, it is still a function of time because the vital rates $B(a, t)$ and $D(a, t)$ in equation (5) are still functions of time. The quasi-equilibrium assumption simply means that the timescale on which these vital rates change is much longer than that on which $q_i(s, t)$ changes. Substituting this into expression (2) gives

$$\dot{\bar{\tau}}(a, t) = \text{cov}[\tau_i(a), \rho[\tau_i](t)]. \tag{6}$$

The operator $\rho[\cdot](t)$ is a nonlinear functional, meaning that fitness depends, in a nonlinear way, on the function-valued trait, $\tau(a)$. If the genetic variation in the pathogen population is not too large, however, then equation (6) can be approximated as

$$\dot{\bar{\tau}}(a, t) \approx \int_0^\infty \phi(s, t; \bar{\tau})G_{\tau, \tau}(a, s; t)\mathrm{d}s, \tag{7}$$

where $\phi(s, t; \bar{\tau})$ is the selection gradient on trait τ at infection age s and time t (Appendix D; Gomulkiewicz and Beder 1996; Beder and Gomulkiewicz 1998) and $G_{\tau, \tau}(a, s; t)$ is the autocovariance in trait τ between infection ages a and s. At a conceptual level, the selection gradient in equation (7) is simply the derivative of the fitness function with respect to the trait of interest, evaluated at the population mean trait value. Formally, it is the kernel of the linear integral operator representing the Fréchet derivative of the functional, $\rho[\cdot](t)$, evaluated at the population mean function, $\bar{\tau}$ (Appendix D).

Model (7) is analogous to that derived by Kirkpatrick and Heckman (1989) and Kirkpatrick and Lofsvold (1992) in quantitative genetics. The underlying assumptions in the derivation of (7) are, however, somewhat different. First, model (7) is not based on an assumption of Gaussian distributions of genotypes or phenotypes, but rather it allows for any discrete genotype distribution so long as the genetic variance is relatively small. Second, model (7) is derived in the context of an explicit ecological (in

this case epidemiological) model for the population dynamics. As a result, the fitness of different genotypes (as represented by the selection gradient, $\phi(s, t; \bar{\tau})$) is not an arbitrary, externally imposed component of the model. Rather, it is an emergent property of the underlying epidemiological dynamics. In other words, the selection gradient in this context is not an input that one explicitly specifies for the model, but instead is an output of the epidemiological dynamics. It is through this selection gradient that the evolutionary dynamics are linked to the epidemiological dynamics. That is the reason why the selection gradient depends on time, t.

Equation (7) has a useful interpretation. The selection gradient specifies how the epidemiological dynamics give rise to selection on the within-host pathogen density at each infection age, whereas the autocovariance function specifies how the within-host pathogen density at infection age a is genetically correlated with the density at infection age s. Therefore, according to (7), the within-host pathogen density at infection age a will evolve through direct selection acting at this age, as well as through indirect selection acting on the density at other infection ages, mediated by any genetic covariance between them.

The form of equation (7) for the evolutionary dynamics of the within-host pathogen density, τ, reveals how this modeling approach circumvents the needs for mechanistic knowledge of the within-host dynamics. All we need to know is the pattern of autocovariance in within-host pathogen density across infection ages. This can be estimated from measurements of the within-host density, without any reference to the mechanistic processes that gave rise to this covariance (see Mideo et al. 2011). At the same time, however, if we did know enough about the within-host dynamics, then we could still employ model (7) by simply using a model for the within-host dynamics to derive the resulting autocovariance structure. Below we consider examples that illustrate each of these possibilities.

Examples

We illustrate the utility of model (7) by focusing on a very simple Susceptible-Infected-Recovered (SIR) epidemiological model. In the first subsection, we consider the evolution of a single function-valued trait representing the within-host pathogen density, as a function of infection age. In the second subsection, we consider the evolution of disease life histories more generally, by allowing for the evolution of multiple function-valued traits.

(1) EVOLUTION OF WITHIN-HOST PATHOGEN DENSITY PROFILES

Consider a simple SIR model (Hethcote 2000) and assume that the epidemiological dynamics of infected hosts is given by equation (1). Suppose that there is a constant influx of suscepti-

ble individuals at rate θ, that all individuals suffer a constant per capita mortality rate of μ, and that pathogen transmission occurs according to the law of mass action with rate that depends on the within-host pathogen density at each infection age. Further, suppose that the virulence and recovery rate of infected individuals at any given infection age is independent of pathogen strain. Using $S(t)$ and $R(t)$ to denote the number of susceptible and recovered individuals at time t, we have

$$\frac{dS(t)}{dt} = \theta - \mu S(t) - \sum_i \int_0^\infty \beta(\tau_i(a)) S(t) I_i(a, t) da \quad (8a)$$

$$\frac{dR(t)}{dt} = \sum_i \int_0^\infty \gamma(a) I_i(a, t) da - \mu R(t), \quad (8b)$$

where i indexes pathogen strain. In terms of the general model in the preceding section, we have $b = \beta(\tau(a))S(t)$ for some function $\beta(\tau)$, and $d = \gamma(a) + v(a) + \mu$. The quantity $v(a)$ represents the pathogen-induced mortality rate at infection age a (i.e., virulence; Day 2002b), and the function $\gamma(a)$ is the recovery rate at infection age a. We stress, however, that both virulence and recovery rate are assumed to be independent of within-host pathogen density, and thus independent of strain type. Thus, evolutionary change in the within-host pathogen density profile results in a change in the transmission profile but not virulence or recovery rate.

Substituting the above expressions for b and d into equation (5), and calculating the selection gradient, $\phi(s, t; \bar{\tau})$ (i.e., the kernel of the Fréchet derivative of $\rho[\tau](t)$ with respect to τ, and evaluated at $\bar{\tau}$), gives

$$\phi(s, t; \bar{\tau}) = \frac{q(s, t)}{k(t)} \beta'(\bar{\tau}(s)) S(t), \quad (9)$$

where $k(t) = \int_0^\infty ab(\bar{\tau}(a), t) e^{-a\rho[\bar{\tau}](t)} l(a) da / \int_0^\infty e^{-a\rho[\bar{\tau}](t)} l(a) da$, where $l(a) = e^{-\int_0^a d(\bar{\tau}(h)) dh}$ (Kirkpatrick and Lofsvold 1989). Note that $k(t) = \tilde{ba}$ where recall, $\tilde{ba}$ is the expectation of ba over the distribution $q(a, t)$. Thus, $k(t)$ is a combined measure of transmission and generation time of infections. Roughly speaking, $k(t)$ is large when the product of the rate of generation of new infections and the generation time is large. We note that $k(t)$ can also be written as $k = q_0 T$, where $T = \int_0^\infty ab(\bar{\tau}(a), t) e^{-a\rho[\bar{\tau}](t)} l(a) da$ is the average age of the "parents" of all new infections at time t (i.e., the generation time), and $q_0 = \int_0^\infty e^{-a\rho[\bar{\tau}](t)} l(a) da$.

Substituting (9) into (7) then gives the evolutionary dynamics of τ as

$$\dot{\bar{\tau}}(a, t) \approx \frac{1}{k(t)} \int_0^\infty q(s, t) \beta'(\bar{\tau}(s)) S(t) G_{\tau, \tau}(a, s; t) ds. \quad (10)$$

The selection gradient in (10) shows that increased within-host pathogen density is always favored (we assume that $\beta' > 0$), and in a way that is proportional to the density of susceptible hosts. A large density of susceptible hosts selects more strongly

for higher within-host density because transmission is evolutionarily more valuable in this case. The strength of selection for increased transmission is weaker at late infection ages, however, because a smaller fraction of the population is actually exposed to this selection. This is reflected by the fact that the selection gradient is also proportional to the fraction of the infected population at each infection age, $q(s, t)$. Also, evolutionary change in within-host pathogen density (and thus transmission) at each infection age is not independent of its evolution at other infection ages because of genetic autocorrelation in pathogen density across infection ages, as described by the autocovariance function $G_{\tau,\tau}(a, s; t)$. Finally, the quantity $1/k(t)$ scales the rate of evolution.

In the absence of any negative genetic autocovariance (i.e., no trade-offs in transmission across infection ages), equation (10) predicts a perpetual evolutionary increase in within-host pathogen density and thus transmission at all infection ages. This is a familiar result from previous theory. If transmission rate is the only epidemiological quantity that is affected by pathogen genotype, then transmission will evolve to ever larger values.

The genetic autocovariance across infection ages will typically not be zero, however, because the within-host pathogen density is typically not independent from one time to the next, across pathogen strains. In this example, we will suppose that we know enough about the within-host dynamics to build an appropriate model for this, and we illustrate how the theory can be applied in this case. In fact, for illustrative purposes, we will consider three different within-host models to better convey how the mechanistic details of within-host dynamics are linked to patterns of autocovariance at the level of an infection.

The first model assumes that pathogen replication is controlled by lymphoctyes, and that lymphocytes grow exponentially and independently of pathogen density (e.g., André et al. 2003). Mathematically we have

$$\frac{d\tau(a)}{da} = r\tau(a) - \kappa x(a)\tau(a) \tag{11a}$$

$$\frac{dx(a)}{da} = \alpha x(a), \tag{11b}$$

where $x(a)$ is the lymphocyte density at infection age a, r is the pathogen replication rate, κ is a constant representing the susceptibility of pathogens to destruction by lymphocytes, and α is the growth rate of the lymphocyte population. We assume that pathogen strains differ in their replication rate, r, and thus model (11) yields a different function $\tau_i(a)$ for each r_i. An analogous analysis could be conducted if pathogens differed in other parameters instead.

The pattern of autocovariance in within-host pathogen density across infection ages arising from this model is given in Figure 1A. With this form of within-host dynamics, the autocovariance is always nonnegative across all infection ages—strains with high density (and thus high transmission) at one infection age will have high density at all ages. Thus, (10) predicts an evolutionary increase in within-host pathogen density at all infection ages if this is the appropriate description of the within-host dynamics (Fig. 1A).

The second model is a modification of (11) in which the rate of change of the lymphocyte population is also proportional to pathogen density (Gilchrist and Sasaki 2002):

$$\frac{d\tau(a)}{dt} = r\tau(a) - \kappa x(a)\tau(a) \tag{12a}$$

$$\frac{dx(a)}{dt} = \alpha x(a)\tau(a). \tag{12b}$$

Again, if we assume that pathogen strains differ in r, then we obtain strain-specific functions, $\tau(a)$. Despite the seemingly simple change in model structure, a qualitatively different pattern of autocovariance is now possible (Fig. 1B). It is no longer the case that strains can be ranked by their within-host densities at all infection ages. Rather, strains that have high density early in an infection have low density late in an infection and vice versa. As a result, the autocovariance changes sign across infection ages, meaning that there is a tradeoff between early and late transmission. Equation (10) therefore predicts that the pattern of evolution will differ depending on the relative strength of selection at different infection ages. When the number of infected individuals is increasing or constant, the infection age distribution, $q(s, t)$, will be skewed toward young infection ages. This places most of the "evolutionary weight" in (10) on early infection ages, and thus the trade-off results in evolution favoring strains with earlier transmission over late (Fig. 1B). Conversely, if the number of infected individuals is decreasing for a sustained period of time, then the infection age distribution will be skewed toward older infection ages. Selection then favors strains with late transmission over those with early transmission (results not shown).

Finally, the third model alters the immunological dynamics in yet a different way, and assumes that the lymphocyte population increases at a rate that is proportional to the pathogen density, but that is independent of the lymphocyte population. Lymphocytes are also assumed to decay at a constant per capita rate of v (e.g., Alizon and van Baalen 2005):

$$\frac{d\tau(a)}{dt} = r\tau(a) - \kappa x(a)\tau(a) \tag{13a}$$

$$\frac{dx(a)}{dt} = \alpha\tau(a) - vx(a). \tag{13b}$$

This model can display periodic waves of high pathogen density that ultimately stabilize to a nonzero equilibrium level (Alizon

Figure 1. Continued.
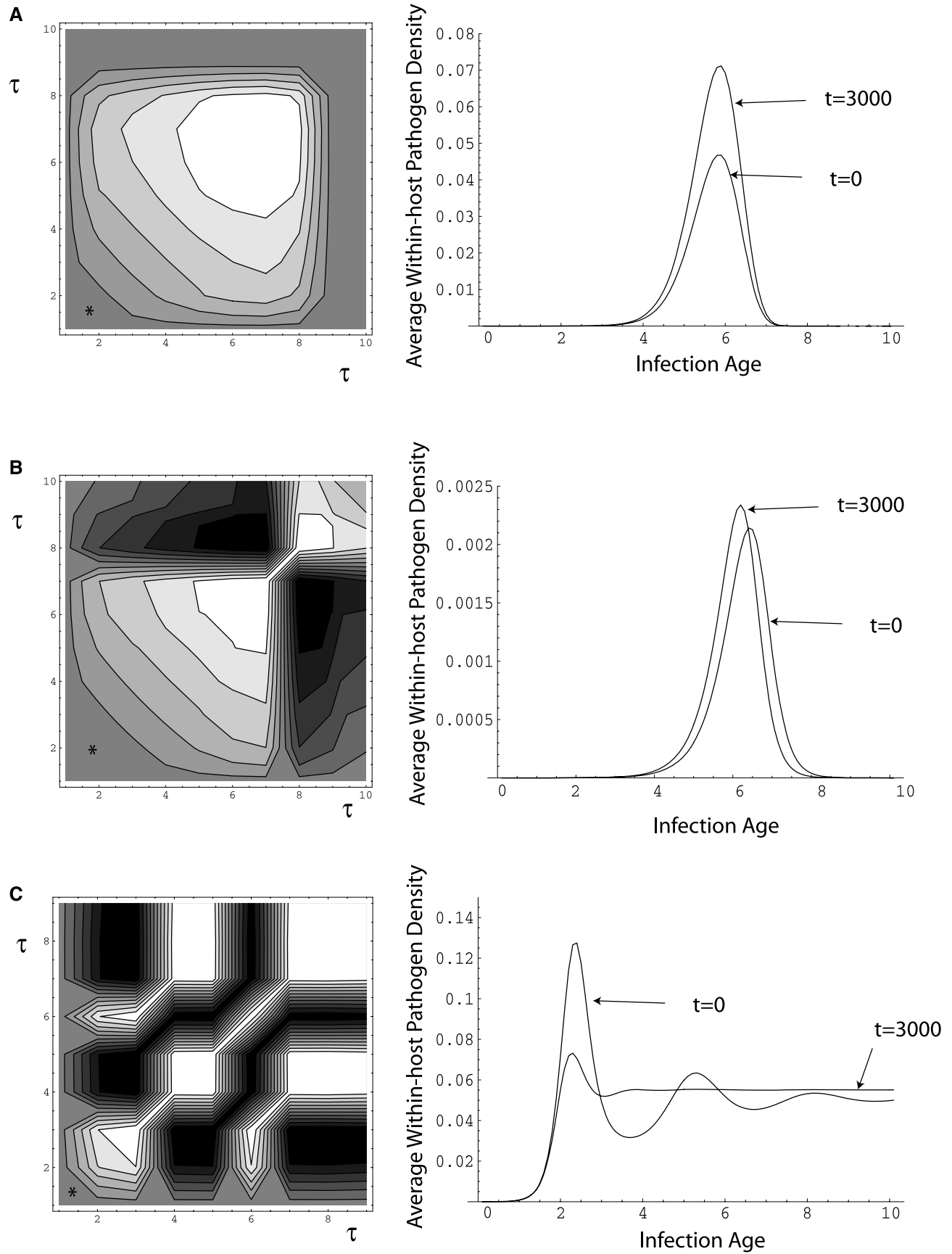
and van Baalen 2005). The phase of these oscillations changes with the parameters of the model, and as a result, the sign of the autocovariance across infection ages can also display a periodic pattern (Fig. 1C).

As with the second model, we obtain different evolutionary outcomes depending upon the relative strength of selection at each infection age (which is determined by the infection age distribution, $q(s, t)$). Furthermore, because of the increased complexity of this within-host model, a wider variety of disease life histories can now evolve, depending upon the epidemiological dynamics. As an example, Figure 1C depicts a case where the disease life history evolves to have sharp, periodic, waves of pathogen density when the epidemic is expanding ($t = 0$), but that evolves to have a constant, lower level, of pathogen density once an endemic equilibrium is reached ($t = 3000$). In an expanding epidemic, most infections are very young, and this thereby selects for very rapid within-host growth. Such high within-host growth rates also then result in marked oscillations in pathogen density during an infection. Once the prevalence of the disease has stabilized, however, the infection age distribution becomes less skewed toward young infection ages, and transmission from older infections thereby becomes important. This selects for slower within-host growth so that a more consistent within-host pathogen density is maintained at later infection ages (Fig. 1C).

(2) EVOLUTION OF DISEASE LIFE HISTORIES

The above results illustrate how within-host models can be used to determine the autocovariance structure required for making evolutionary predictions. In the absence of a within-host model, evolutionary predictions can nevertheless still be made, as long as we have an estimate of the required autocovariance structure. In this example, we illustrate this fact, and we also show how the above results can be readily extended to make predictions about the evolution of multiple function-valued traits.

Suppose that transmission rate and virulence (as a function of infection age) both differ among pathogen strains. Often we have no information about how the within-host dynamics of pathogen

replication, host resources, and host defense mechanisms give rise to strain-specific patterns of transmission and virulence during an infection. In such cases, we might then consider transmission rate and virulence to be, themselves, the function-valued traits of interest. We would then model the evolution of both the average transmission rate, $\bar{\beta}(a, t)$, and the average virulence, $\bar{v}(a, t)$, both being functions of infection age.

To illustrate, consider a simple SIR model

$$\frac{dS(t)}{dt} = \theta - \mu S(t) - \sum_i \int_0^\infty \beta_i(a) S(t) I_i(a, t) da \quad (14a)$$

$$\frac{dR(t)}{dt} = \sum_i \int_0^\infty \gamma(a) I_i(a, t) da - \mu R(t), \quad (14b)$$

where i indexes pathogen strain. In terms of the general model (1), we have $b = \beta(a)S(t)$ and $d = \gamma(a) + v(a) + \mu$ where the pair of functions, $\beta_i(a)$ and $v_i(a)$ describe the transmission rate and virulence of strain i as a function of infection age a. Substituting these expressions for b and d into equation (5), and calculating the selection gradient for multiple function-valued traits (Appendix D) gives

$$\phi_\beta(s, t; \bar{\beta}, \bar{v}) = \frac{q(s, t)}{k(t)} S(t) \quad (15a)$$

$$\phi_v(s, t; \bar{\beta}, \bar{v}) = -\frac{q(s, t)}{k(t)} \sigma(s, t), \quad (15b)$$

where $\sigma(s, t) = \int_s^\infty e^{-a\rho[\bar{\beta}, \bar{v}](t)} l(a) \bar{\beta}(a) S(t) da / e^{-s\rho[\bar{\beta}, \bar{v}](t)} l(s)$ is the reproductive value of an infection of age s at time t. Notice that the strength of selection on virulence at infection age s depends, not only on the weight of the infection age distribution at that age, but also on the reproductive value of infections of that age. This latter effect is absent from the selection gradient on transmission because evolutionary changes in virulence (i.e., mortality) affect reproductive output at all future infection ages as well, whereas evolutionary changes in transmission affect that infection age only.

Figure 1. Patterns of autocovariance in pathogen density, τ across infection ages, and the evolutionary dynamics of mean pathogen density as a function of infection age, for different models. An asterisk in the covariance plots indicates the shade corresponding to a covariance of zero (lighter=positive, darker=negative). (A) Model (11). Autocovariance plotted using two strains of equal frequency that differ in r. Parameter values for the within-host model are $r_1 = 3.2$, $r_2 = 3$, and $\kappa = 10^{-2}$, and $\alpha = 1$, with initial conditions $\tau(0) = 1$ and $x(0) = 1$. Parameter values for the between-host model are $\theta = 10$, $\mu = 0.01$, $v(a) \equiv 0.0005$, $\gamma = 0.0005$ and $\beta(\tau) = 6 \cdot 10^{-7}\tau$, with initial conditions: $I_i(0, 0) = 1$ and $I_i(a, 0) = 0$ for $a \neq 0$, along with $S_0 = 100$. (B) Model (12). Autocovariance plotted using two strains of equal frequency that differ in r. Parameter values for within-host model are $r_1 = 3.1$, $r_2 = 3$, and $\kappa = 1$, $\alpha = 10^{-5}$, with initial conditions $\tau(0) = 1$ and $x(0) = 1$. Parameter values for between-host model are $\theta = 10$, $\mu = 0.01$, $v(a) \equiv 0.0005$, $\gamma = 0.0005$, and $\beta(\tau) = 2 \cdot 10^{-6}\tau$, with initial conditions $I_1(0, 0) = 1$, $I_2(0, 0) = 10$, and $I_i(a, 0) = 0$ for $a \neq 0$, along with $S_0 = 100$. (C) Model (13). Autocovariance plotted using two strains of equal frequency that differ in v and κ. Parameter values for within-host model are $v_1 = 1$, $v_2 = 4$, $\kappa_1 = 10^{-4}$, $\kappa_2 = 10^{-3.45}$, and $r = 4.9$, $\alpha = 10^{-5}$, with initial conditions $\tau(0) = 10^6$ and $x(0) = 1$. Parameter values of between-host model are $\theta = 10000$, $\mu = 0.01$, $v(a) \equiv 0.0005$, $\gamma = 0.0005$, and $\beta(\tau) = 9.5 \cdot 10^{-17}\tau$, with initial conditions: $I_1(0, 0) = 1$, $I_2(0, 0) = 10$ and $I_i(a, 0) = 0$ for $a \neq 0$, along with $S_0 = 10^7$.

Substituting (15) into (D4) gives the evolutionary dynamics as

$$\dot{\bar{\beta}}(a, t) \approx \frac{1}{k(t)}\left(\int_0^\infty q(s, t)S(t)G_{\beta,\beta}(a, s; t)\mathrm{d}s \right.$$
$$\left. - \int_0^\infty q(s, t)\sigma(s, t)G_{\beta,v}(a, s; t)\mathrm{d}s \right) \qquad (16a)$$

$$\dot{\bar{v}}(a, t) \approx \frac{1}{k(t)}\left(\int_0^\infty q(s, t)S(t)G_{v,\beta}(a, s; t)\mathrm{d}s \right.$$
$$\left. - \int_0^\infty q(s, t)\sigma(s, t)G_{v,v}(a, s; t)\mathrm{d}s \right). \qquad (16b)$$

Model (16) reveals that the evolution of each function-valued trait now also depends on the cross-covariance, $G_{\beta,v}(a, s; t)$, between the two. For example, the first term in equation (16a) for the evolution of the average transmission rate is the same as that in equation (10) (except that $\beta' = 1$ now because the focal trait is β itself rather than τ). The second term in equation (16a) shows that there is now an additional component of indirect selection. Selection on virulence at infection age s can result in evolutionary changes in transmission at infection age a if there is any genetic covariance between the two. The cross-covariance function $G_{\beta,v}(a, s; t)$ describes such genetic correlations, and the second integral in (16a) essentially sums up the effects of this indirect selection on virulence across all infection ages.

The trade-off hypothesis is a cornerstone of theory on virulence evolution, and it postulates that, despite virulence always being selected against, intermediate levels nevertheless evolve because of a trade-off between transmission and virulence (Anderson and May 1982; Ewald 1983; Alizon et al. 2009). In genetic terms, the hypothesis postulates a positive genetic covariance between transmission and virulence—strains with high transmission also tend to induce high virulence (Day and Proulx 2004; Day and Gandon 2006, 2007). In the context of disease life-history evolution, however, the existence of such a trade-off (or the lack thereof) is a more subtle issue. Both transmission rate and virulence typically vary with infection age, and therefore describing any trade-off between the two is necessarily more complex, as it involves quantifying a cross-covariance function across all infection ages, rather than simply a single genetic covariance. More significantly, however, the existence of an evolutionary trade-off now depends, not only on the genetic cross-covariance structure of the pathogen, but also on how this interacts with the epidemiological dynamics.

To illustrate the subtlety of the nature of potential trade-offs, suppose we have two pathogen strains that display the transmission and virulence patterns depicted in Figure 2A. Virulence is constant across infection age for each strain, but transmission rate changes in a step-fashion for each strain at infection age between three and four. The autocovariance and cross-covariance functions are given in Figure 2B. The cross-covariance of transmission rate between ages zero and four with virulence at any age is zero because there is no genetic variation for transmission between ages zero and four. As a result, there is no trade-off across strains between transmission at these infection ages and virulence. On the other hand, the cross-covariance of transmission rate between ages four and 10 with virulence at any infection age is positive because the strain with the highest transmission rate between infection ages four and 10 also induces the highest virulence at all infection ages. Thus, there is a trade-off between transmission and virulence if we restrict attention to infection ages later than day four.

The above considerations suggest that if the majority of infections are quite young (as will be the case in an expanding epidemic), and evolution is determined largely by strain differences early in an infection, then virulence will evolve toward zero because there is effectively no trade-off (Fig. 3). On the other hand, in endemic situations (or cases where the number of infections is decreasing) the infection age distribution will be skewed toward later ages. Evolution will then be governed more by strain differences late in an infection, and the presence of a virulence-transmission trade-off at these later ages will thereby result in intermediate levels of virulence (Fig. 3). Thus, whether an evolutionary trade-off between virulence and transmission exists depends on an interesting interaction between the epidemiological dynamics and the genetic covariance structure of the pathogen population.

Discussion

Infectious disease dynamics occur on at least two important scales: the dynamics of within-host pathogen replication and its interaction with host resources and defense mechanisms, and the dynamics of between-host pathogen transmission. A goal of recent theoretical research in the evolutionary epidemiology of infectious disease has been to bridge these two dynamical scales within a common modeling approach. Although several interesting theoretical results have been developed along these lines (Sasaki and Iwasa 1991; Day 2001, 2002b, 2002a; Gilchrist and Sasaki 2002; Ganusov et al. 2002; André et al. 2003; Alizon and van Baalen 2005; Gilchrist and Coombs 2006; André and Gandon 2006; Coombs et al. 2007; Alizon and van Baalen 2008; Mideo et al. 2008), integration of the results with empirical research is still rather limited.

One reason for a divide between theoretical and empirical research stems from the difficulty researchers face when attempting to connect models with data. For many pathogens of interest the information required to apply the theory is simply not yet available. This is particularly true for models that aim to connect
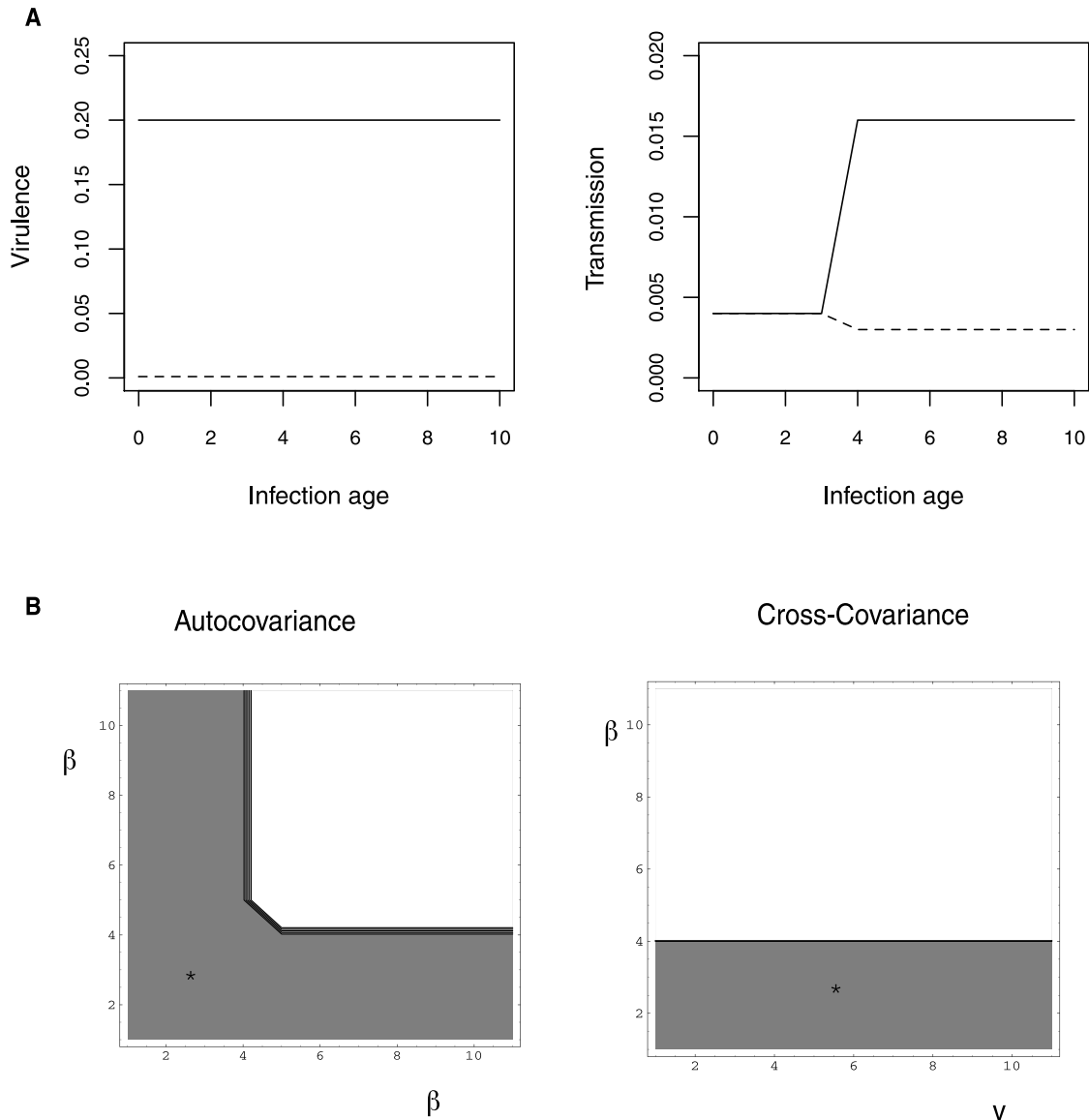
Figure 2. Patterns of virulence and transmission (A) for the two strains used in the model for the evolution of disease life histories, along with the corresponding patterns of autocovariance and cross-covariance (B). Solid and dashed lines in (A) corresponds to two different pathogen strains. Covariances in (B) are zero everywhere except where white.

different scales of biological dynamics. In the context of infectious diseases, it is often a lack of knowledge of the mechanistic details of within-host dynamics that prevents the application of theoretical results. These within-host dynamics are critical for understanding the evolutionary epidemiology of disease, however, because they are what determines the pattern of transmission, mortality, and recovery over the course of an infection, as well as how these disease life-history attributes vary across pathogen genotypes.

These difficulties have motivated the so-called function-valued trait approach presented here. These models aim to capture the relevant evolutionary information about the within-host aspects of disease in a phenomenological way, through the use of

statistical estimates of genetic covariance functions for the disease life-history traits of interest. Such an approach still allows one to predict the evolutionary dynamics of disease life histories, but it does so in a way that alleviates the need to understand the mechanistic processes through which these disease life-history traits arise.

The examples presented here illustrate the general techniques, and they also demonstrate how the approach can be used in cases where we do actually understand the within-host dynamics in a quantitative way. Furthermore, they also reveal some interesting general findings about the evolution of infectious diseases. For instance, these models show that even relatively simple within-host processes can give rise to rather complex patterns
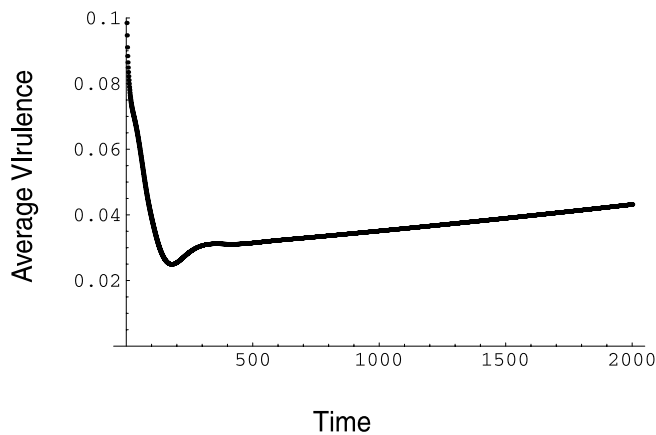
Figure 3. Pattern of evolution in the average level of virulence over time, using the genetic constraints described in Figure 2. From $t = 0$ until $t = 200$ the epidemic is expanding, after which it contracts somewhat and then settles to an endemic level from $t = 500$ onwards. Parameter values are as in Figure 2, along with $\theta = 10$, $\mu = 0.01$, $v_1(a) + \gamma \equiv 0.001$, $v_2(a) + \gamma \equiv 0.02$, with initial conditions: $I_i(0, 0) = 1$ and $I_i(a, 0) = 0$ for $a \neq 0$, along with $S_0 = 100$.

of genetic constraint on disease life-history evolution. They also demonstrate that quite small differences in within-host processes can lead to quite large differences in these patterns of genetic constraint. Interestingly, the models also reveal that the trade-off hypothesis from evolutionary epidemiology (Anderson and May 1982; Ewald 1983; Fraser et al. 2007; de Roode et al. 2008; Alizon et al. 2009) can be much more subtle than is usually appreciated. In contrast to most descriptions, the existence of an evolutionary trade-off between transmission and virulence is not solely a property of the host–pathogen interaction. Rather, the existence of a trade-off arises from an interplay between the pattern of genetic variation in the pathogen population, and the epidemiological dynamics of the disease at the level of the host population. The reason is that the epidemiological dynamics in the host population determine the age distribution of infections, and if patterns of genetic constraint in the pathogen population vary as a function of infection age, then changes in this age distribution will affect the consequences of genetic constraints as well.

Of course any real utility of the approach presented here must come from its application to real data. In principle this should be easier than with previous theoretical approaches, but there are still relatively few datasets available that are sufficient. The requisite data are beginning to appear, however, and we present an example of this as a proof of principle in a companion paper. Our hope is that these empirical examples will motivate the collection of further such data.

Finally, it is worth noting that the approach developed here is very closely related to models of function-valued traits in quantitative genetics (Kirkpatrick and Heckman 1989; Kirkpatrick and Lofsvold 1992; Gomulkiewicz and Beder 1996; Beder and Gomulkiewicz 1998; Kingsolver et al. 2001). In those models, the function-valued trait of interest is usually an organismal characteristic, such as an age-size relationship. The analogue of a within-host model in such examples would then be a model for the developmental processes that give rise to this size versus age function. As with the disease examples presented here, however, the information required to build such developmental models is typically not available, and therefore these approaches necessarily employ phenomenological descriptions of genetic covariance patterns as well.

This functioned-valued approach has become quite widely applied in the context of quantitative genetics, and the prospects for its application in the context of disease evolution should be at least as good. Indeed, estimates of the required genetic covariance functions might sometimes be even easier to obtain in the context of diseases because of shorter generation times and the lack of a need for complex breeding designs. Furthermore, for some diseases we do have the beginnings of an understanding of the details of within-host process. As a result, in these cases we might even be able to use this knowledge to derive the patterns of covariance from a mechanistic model of within-host dynamics. This would not only provide a second, independent, approach for quantifying these important functions, but it might also allow one to predict how these covariance structures evolve over time as well.

LITERATURE CITED

Alizon, S., and M. van Baalen. 2005. Emergence of a convex trade off between transmission and virulence. Am. Nat. 165:E155–E167.

———. 2008. Multiple infections, immune dynamics and virulence evolution. Am. Nat. 172:E150–E158.

Alizon, S., A. Hurford, N. Mideo, and M. van Baalen. 2009. Virulence evolution and the trade off hypothesis: history, current state of affairs and the future. J. Evol. Biol. 22:245–259.

Anderson, R. M., and R. M. May. 1982. Coevolution of hosts and parasites. Parasitology 85:411–426.

André, J.-B., J.-B. Ferdy, and B. Godelle. 2003. Within-host parasite dynamics, emerging trade off, and evolution of virulence with immune system. Evolution 57:1489–1497.

André, J.-B., and S. Gandon. 2006. Vaccination, within-host dynamics, and virulence evolution. Evolution 60:13–23.

Beder, J. H., and R. Gomulkiewicz. 1998. Computing the selection gradient and evolutionary response of an infinite-dimensional trait. J. Math. Biol. 36:299–319.

Coombs, D., M. A. Gilchrist, and C. L. Ball. 2007. Evaluating the importance of within-and between-host selection pressures on the evolution of chronic pathogens. Theor. Popul. Biol. 72:576–591.

Day, T. 2001. Parasite transmission modes and the evolution of virulence. Evolution 55:2389–2400.

———. 2002a. The evolution of virulence in vector-borne and directly transmitted parasites. Theor. Popul. Biol. 62:199–213.

———. 2002b. On the evolution of virulence and the relationship between various measures of mortality. Proc. R. Soc. Lond. B 269:1317–1323.

———. 2003. Virulence evolution and the timing of disease life-history events. Trends Ecol. Evol. 18:113–118.

Day, T., and S. Gandon. 2006. Insights from Price's equation into evolutionary epidemiology. Pp. 23–44 in Z. Feng, U. Dieckmann, and S. A. Levin, eds. Disease evolution: models, concepts, and data analyses. Vol. 71 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*. American Mathematical Society.

———. 2007. Applying population-genetic models in theoretical evolutionary epidemiology. Ecol. Lett. 10:876–888.

Day, T., and S. R. Proulx. 2004. A general theory for the evolutionary dynamics of virulence. Am. Nat. 163:E40–E63.

de Roode, J. C., A. J. Yates, and S. Altizer. 2008. Virulence-transmission trade offs and population divergence in virulence in a naturally occurring butterfly parasite. Proc. Natl. Acad. Sci. USA 105:7489–7494.

Dieckmann, U., M. Heino, and K. Parvinen. 2006. The adaptive dynamics of function-valued traits. J. Theor. Biol. 241:370–389.

Ewald, P. W. 1983. Host-parasite relations, vectors, and the evolution of disease severity. Annu. Rev. Ecol. Evol. Syst. 14:465–485.

Frank, S. A. 1996. Models of parasite virulence. Q. Rev. Biol. 71:37–78.

Fraser, C., T. D. Hollingsworth, R. Chapman, F. de Wolf, and W. P. Hanage. 2007. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. Proc. Natl. Acad. Sci. USA 104:17441–17446.

Ganusov, V. V., C. T. Bergstrom, and R. Antia. 2002. Within-host population dynamics and the evolution of microparasites in a heterogeneous host population. Evolution 52:213–223.

Gilchrist, M. A., and D. Coombs. 2006. Evolution of virulence: interdependence, constraints, and selection using nested models. Theor. Popul. Biol. 69:145–153.

Gilchrist, M. A., and A. Sasaki. 2002. Modeling host-parasite coevolution: a nested approach based on mechanistic models. J. Theor. Biol. 218:289–308.

Gomulkiewicz, R., and J. H. Beder. 1996. The selection gradient of an infinite-dimensional trait. SIAM J. Appl. Math. 56:509–523.

Gyllenberg, M., and I. Hanski. 1992. Single-species metapopulation dynamics—a structured model. Theor. Popul. Biol. 42:35–61.

Hethcote, H. W. 2000. The mathematics of infectious diseases. SIAM Rev. 42:599–653.

Hoppensteadt, F. 1974. An age dependent epidemic model. J. Franklin Institute 297:325–333.

Hutson, V., J. S. Pym, and M. J. Cloud. 2005. Applications of functional analysis and operator theory. Elsevier, Amsterdam.

Kingsolver, J. G., R. Gomulkiewicz, and P. A. Carter. 2001. Variation, selection and evolution of function-valued traits. Genetica 112:87–104.

Kirkpatrick, M., and N. Heckman. 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. J. Math. Biol. 27:429–450.

Kirkpatrick, M., and D. Lofsvold. 1989. The evolution of growth trajectories and other complex quantitative characters. Genome 31:778–783.

———. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954–971.

Lande, R. 1982. A quantitative genetic theory of life history evolution. Ecology 63:607–615.

Levin, B. R., and J. J. Bull. 1994. Short-sighted evolution and the virulence of pathogenic microorganisms. Trends Microbiol. 2:76–81.

Levin, S., and D. Pimentel. 1981. Selection of intermediate rates of increase in parasite-host systems. Am. Nat. 117:308–315.

May, R. M., and M. A. Nowak. 1995. Coinfection and the evolution of parasite virulence. Proc. R. Soc. Lond. B 261:209–215.

McKendrick, A. G. 1926. Applications of mathematics to medical problems. Proceedings of the Edinburgh Mathematical Society 44:98–130.

Metz, J., and O. Diekmann. 1986. The dynamics of physiologically structured populations. Springer, Berlin.

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

Mideo, N., W. A. Nelson, S. E. Reece, A. S. Bell, A. F. Read, and T. Day. 2011. Bridging scales in the evolution of infectious diseases: application. *In press*.

Nowak, M. A., and R. M. May. 1994. Superinfection and the evolution of parasite virulence. Proc. R. Soc. Lond. B 255:81–89.

Parvinen, K., U. Dieckmann, and M. Heino. 2006. Function-valued adaptive dynamics and the calculus of variations. J. Math. Biol. 52:1–26.

Price, G. R. 1970. Selection and covariance. Nature 227:520–521.

Sasaki, A., and Y. Iwasa. 1991. Optimal growth schedule of pathogens within a host: switching between lytic and latent cycles. Theor. Popul. Biol. 39:201–239.

van Baalen, M., and M. W. Sabelis. 1995. The dynamics of multiple infection and the evolution of virulence. Am. Nat. 146:881–910.

Appendix A - Some Technical Remarks

In general, to bridge the within- and between-host dynamics of disease transmission, one needs to account for the fact that the population of infections at the between-host level is structured by a number of different state variables at the within-host level (e.g., pathogen density, immune status, etc). Thus, one might use techniques for modeling physiologically structured populations (Metz and Diekmann 1986; Gyllenberg and Hanski 1992). The approach taken here is essentially a simplification of these techniques. In particular, we assume that "infection age" is the single variable required to characterize the structure of the infected population. This greatly simplifies both the analysis and its application to data, and it is appropriate provided that the dynamics of all the relevant within-host state variables are deterministic. In particular, when these dynamics are deterministic, the age of an infection (along with initial conditions) completely determines the within-host state. We note that most previous analyses that bridge these two scales of dynamics assume this type of deterministic dynamics as well, but it would be interesting to extend the results beyond this simplification.

Given that infection age is the sole means by which infections are structured, we focus on the evolution of traits that are functions

of infection age. We restrict attention to evolution within the Hilbert space of square-integrable functions on the interval (0, α); that is, $\mathcal{L}_2(0, \alpha)$. The value of α will differ depending on the function-valued trait of interest, but for notational simplicity we set $\alpha = \infty$. The norm of an element, $\tau(a) \in \mathcal{L}_2(0, \infty)$, is thus taken to be the $\mathcal{L}_2$ norm given by the Lebesgue integral $||\tau(a)|| = \int_0^\infty \tau(a)^2 \mathrm{d}a$ (Hutson et al. 2005).

Each pathogen genotype i encodes a specific function, $\tau_i(a) \in \mathcal{L}_2(0, \infty)$, and the fitness of this strain (defined as the per capita rate of change of hosts infected with this strain) is assumed to be determined by $\tau_i(a)$. The fitness of pathogen strain i can thus be viewed as the value obtained by applying an operator, $J[\cdot]$ (in this case a functional) to the element $\tau_i(a)$. Throughout we use square bracket notation to indicate functionals.

In many cases $J[\cdot]$ will be nonlinear, and we will want to approximate it with a linear operator near a function of interest. We say that the functional $J[\cdot]$ is Fréchet differentiable at τ^* if there exists a bounded linear operator, $L[\cdot]$ such that $J[\tau^* + h] = J[\tau^*] + L[\tau^*]h + o(h)$ as $h \to 0$ in the $\mathcal{L}_2$ norm. $L[\tau^*]$ is called the Fréchet derivative at τ^* and the notation $L[\tau^*]h$ denotes this linear operator applied to the element h.

Appendix B - Derivation of Evolutionary Dynamics

To begin, we need to track the dynamics of the frequency of strain i among infected individuals. Define $I_i^T(t) = \int_0^\infty I_i(a, t)\mathrm{d}a$ as the total number of hosts infected with strain i, $N(t) = \sum_i I_i^T(t)$ as the total number of infected hosts of all types, and

$$p_i(t) = I_i^T(t)/N(t) \tag{B1}$$

as the frequency of strain i among infected individuals. From equation (B1), the dynamics of p_i are

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = \frac{\mathrm{d}I_i^T/\mathrm{d}t}{N} - p_i \frac{\mathrm{d}N/\mathrm{d}t}{N}. \tag{B2}$$

To simplify (B2), we derive an expression for the time dynamics of I_i^T and N;

$$\frac{\mathrm{d}I_i^T}{\mathrm{d}t} = \int_0^\infty \frac{\partial I_i(a, t)}{\partial t} \mathrm{d}a \tag{B3a}$$

$$= -\int_0^\infty \frac{\partial I_i(a, t)}{\partial a} \mathrm{d}a - \int_0^\infty D_i(a, t) I_i(a, t) \mathrm{d}a \tag{B3b}$$

$$= I_i^T(t) \int_0^\infty B_i(a, t) q_i(a, t) \mathrm{d}a - I_i^T(t) \int_0^\infty D_i(a, t) q_i(a, t) \mathrm{d}a \tag{B3c}$$

$$= \tilde{r}_i(t) I_i^T(t) \tag{B3d}$$

and

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \sum_i \mathrm{d}I_i^T/\mathrm{d}t \tag{B4a}$$

$$= \bar{\tilde{r}}_i(t) N(t). \tag{B4b}$$

Substituting into (B2) then yields

$$\frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = p_i(t)(\tilde{r}_i(t) - \bar{\tilde{r}}_i(t)). \tag{B5}$$

The time dynamics of $\bar{\tau}(a, t)$ is therefore given by

$$\dot{\bar{\tau}}(a, t) = \sum_i \tau_i(a) p_i(t)(\tilde{r}_i(t) - \bar{\tilde{r}}_i(t)) \tag{B6a}$$

$$= \mathrm{cov}[\tau_i(a), \tilde{r}_i(t)]. \tag{B6b}$$

It is worth noting that the average value of $\tau(a, t)$ that satisfies the above equation is only one way in which the average of the function-valued trait might be calculated. In fact there are three common possibilities for how we might quantify the average value of the function, τ, at infection age a. First, we might take all infections of age a and calculate the average τ value of these infections. Second, we might take all newly produced infections and calculate the average τ value that they would express at age a. Third, we might take all current infections and calculate the average τ value that they would express at age a. Interestingly, to our knowledge no theory has been developed using approach 1 and, in fact, most life-history theory uses approach 2. The analysis above, however, uses approach 3. Our reason for this choice stems from our interest in developing a framework that can be readily applied to natural populations of infected hosts. In such cases it is usually impossible to ascertain the age of naturally occurring infections (unlike organismal age in life-history theory) and therefore there would be no way of restricting measurements to newly produced infections (or any other infection age for that matter). However, one might readily sample propagules from a random collection of all existing infections, and then conduct common garden experiments to measure the characteristics of the disease caused by these propagules at any infection age of interest. Nevertheless, an interesting avenue for future research is to explore the similarities and differences of theory developed for each of these different possibilities.

Appendix C - Derivation of Euler-Lotka Equation

Once a stable infection age distribution is reached, $q_i(s, t) = q_0 e^{-\int_0^s D_i(z,t)\mathrm{d}z} e^{-s\rho[\tau_i](t)}$ (e.g., see Day 2001 equation A4) and

$\tilde{r}_i(t) = \rho[\tau_i](t)$. Therefore, (3) becomes

$$\rho[\tau_i](t) = \int_0^\infty B_i(s,t)q_0 e^{-\int_0^s D_i(z,t)\mathrm{d}z} e^{-s\rho[\tau_i](t)}\mathrm{d}s$$
$$- \int_0^\infty D_i(s,t)q_0 e^{-\int_0^s D_i(z,t)\mathrm{d}z} e^{-s\rho[\tau_i](t)}\mathrm{d}s. \tag{C1}$$

Integrating the second term by parts gives

$$\rho[\tau_i](t) = q_0 \int_0^\infty B_i(s,t) e^{-\int_0^s D_i(z,t)\mathrm{d}z} e^{-s\rho[\tau_i](t)}\mathrm{d}s$$
$$- q_0 + \rho[\tau_i](t) \tag{C2}$$

or

$$1 = \int_0^\infty B_i(s,t) e^{-\int_0^s D_i(z,t)\mathrm{d}z} e^{-s\rho[\tau_i](t)}\mathrm{d}s \tag{C3}$$

which is equation (5).

Appendix D - A Linear Approximation

If the genetic variation in the pathogen population is not too large, in the sense that each genotype's trait, $\tau(s)$ is not too far from the population average, $\bar{\tau}(s)$, then we can approximate $\rho[\cdot](t)$ with a linear functional as $\rho[\tau(s)](t) \approx \rho[\bar{\tau}(s)](t) + L[\bar{\tau}(s)](t)\,(\tau(s) - \bar{\tau}(s))$ where $L[\bar{\tau}(s)](t)$ is the Fréchet derivative of ρ at time t (Hutson et al. 2005). Substituting this approximation into (2) yields

$$\dot{\bar{\tau}}(a,t) \approx \mathrm{cov}[\tau_i(a), \rho[\bar{\tau}(s)](t) + L[\bar{\tau}(s)](t)\,(\tau_i(s) - \bar{\tau}(s))] \tag{D1a}$$

$$= \mathrm{cov}[\tau_i(a), L[\bar{\tau}(s)](t)\,(\tau_i(s) - \bar{\tau}(s))] \tag{D1b}$$

$$= L[\bar{\tau}(s)](t)\,\mathrm{cov}[\tau_i(a), (\tau_i(s) - \bar{\tau}(s))] \tag{D1c}$$

or

$$\dot{\bar{\tau}}(a,t) \approx L[\bar{\tau}(s)](t)\,G_{\tau,\tau}(a,s;t). \tag{D2}$$

For most models of interest, the Fréchet derivatives will typically be linear integral operators of the form $L[\bar{\tau}(s)](t)h = \int_0^\infty \phi(s,t;\bar{\tau})h(s)\mathrm{d}s$ for some function $\phi(s,t;\bar{\tau})$, and therefore we have equation (7) of the text.

Finally, we note that equation (D2) can be generalized to situations in which the asymptotic growth rate, ρ, is dependent on multiple functions. For example, in the case where the asymptotic growth rate depends on two function (e.g., $\rho[\tau_i(a), \omega_i(a)]$), the evolutionary dynamics of $\bar{\tau}(a,t)$ and $\bar{\omega}(a,t)$ is given by

$$\dot{\bar{\tau}}(a,t) \approx L_\tau[\bar{\tau}(s), \bar{\omega}(s)](t)\,G_{\tau,\tau}(a,s;t)$$
$$+ L_\omega[\bar{\tau}(s), \bar{\omega}(s)](t)\,G_{\tau,\omega}(a,s;t) \tag{D3a}$$

$$\dot{\bar{\omega}}(a,t) \approx L_\tau[\bar{\tau}(s), \bar{\omega}(s)](t)\,G_{\omega,\tau}(a,s;t)$$
$$+ L_\omega[\bar{\tau}(s), \bar{\omega}(s)](t)\,G_{\omega,\omega}(a,s;t) \tag{D3b}$$

where $L_\tau[\bar{\tau}(s), \bar{\omega}(s)](t)$ and $L_\omega[\bar{\tau}(s), \bar{\omega}(s)](t)$ denote the Fréchet derivatives with respect to τ and ω, respectively. in the case that these derivatives are integral operators we have

$$\dot{\bar{\tau}}(a,t) \approx \int_0^\infty \phi_\tau(s,t;\bar{\tau},\bar{\omega})G_{\tau,\tau}(a,s;t)\mathrm{d}s$$
$$+ \int_0^\infty \phi_\omega(s,t;\bar{\tau},\bar{\omega})G_{\tau,\omega}(a,s;t)\mathrm{d}s \tag{D4a}$$

$$\dot{\bar{\omega}}(a,t) \approx \int_0^\infty \phi_\tau(s,t;\bar{\tau},\bar{\omega})G_{\omega,\tau}(a,s;t)\mathrm{d}s$$
$$+ \int_0^\infty \phi_\omega(s,t;\bar{\tau},\bar{\omega})G_{\omega,\omega}(a,s;t)\mathrm{d}s. \tag{D4b}$$

Appendix E - Host Heterogeneity and Environmental Noise

Throughout the main text, we assume that the "birth" and "death" rate functions, $B_i(a,t)$ and $D_i(a,t)$, are determined solely by pathogen genotype, i. Likewise, in the examples where we explicitly track the evolution of within-host density, $\tau_i(a)$, we assume that this is entirely determined by pathogen genotype, i, as well. In reality host heterogeneity and environmental noise will cause $\tau(a)$, or more generally $B(a,t)$ and $D(a,t)$, to vary even within a single genotype. Here we illustrate how such effects can be incorporated into the theory developed in the main text.

Suppose that both birth and death rates depend on some factor, e, in addition to genotype; that is, $B_{ie}(a,t)$ and $D_{ie}(a,t)$. The variable e might represent different types of hosts, or simply random environmental perturbations. For simplicity, we will refer to it as "noise" because we are interested in pathogen evolution that is averaged over this variation. Infections of type ie might change their noise status to infections of type $i\hat{e}$ over the course of an infection, and thus we begin by deriving an equation for the dynamics of the density of infections of type ie, denoted by $I_{ie}(a,t)$. Similar to equation (1) we have

$$\frac{\partial I_{ie}(a,t)}{\partial t} = -\frac{\partial I_{ie}(a,t)}{\partial a} - D_{ie}(a,t)I_{ie}(a,t)$$
$$- \xi(1 - \kappa_{ee})I_{ie}(a,t) + \xi \sum_{\hat{e}} \kappa_{\hat{e}e} I_{i\hat{e}}(a,t) \tag{E1}$$

with boundary condition $I_{ie}(0,t) = \int_0^\infty B_{ie}(a,t)I_{ie}(a,t)\mathrm{d}a$. Here ξ is the rate at which a potential change in the noise state occurs, and given this, $\kappa_{\hat{e}e}$ is the probability of changing from state $\hat{e}$ to e. We are primarily interested in the dynamics of all hosts infected with a specific genotype, and therefore defining $I_i(a,t) = \sum_e I_{ie}(a,t)$

as the total density of infections with genotype i, the dynamics of $I_i(a, t)$ can be derived as

$$\frac{\partial I_i(a, t)}{\partial t} = \sum_e \frac{\partial I_{ie}(a, t)}{\partial t} \quad \text{(E2a)}$$

$$= \sum_e \left(-\frac{\partial I_{ie}(a, t)}{\partial a} - D_{ie}(a, t)I_{ie}(a, t) \right. $$
$$\left. -\xi(1 - \kappa_{ee})I_{ie}(a, t) + \xi \sum_{\hat{e}} \kappa_{\hat{e}e}I_{i\hat{e}}(a, t) \right) \quad \text{(E2b)}$$

$$= -\frac{\partial I_i(a, t)}{\partial a} - D_i(a, t)I_i(a, t) \quad \text{(E2c)}$$

with boundary condition $I_i(0, t) = \int_0^\infty B_i(a, t)I_i(a, t)da$, and where $B_i(a, t)$ and $D_i(a, t)$ are interpreted as the birth and death rates of infections with genotype i averaged over the noise distribution; that is, $B_i(a, t) = \sum_e B_{ie}(a, t)\frac{I_{ie}(a,t)}{I_i(a,t)}$ and $D_i(a, t) = \sum_e D_{ie}(a, t)\frac{I_{ie}(a,t)}{I_i(a,t)}$. Equation (E2) is identical to equation (1) provided we interpret $B_i(a, t)$ and $D_i(a, t)$ in this way. Thus, the general results of the main text continue to hold with this new interpretation, provided that the frequency distribution of noise (i.e., $\frac{I_{ie}(a,t)}{I_i(a,t)}$, which is the frequency distribution of host types, random perturbations, etc.) does not change through time.

Likewise, suppose instead that the within-host pathogen density, $\tau_{ie}(a)$ was influenced by both genotype, i, and some noise factor, e. The above equations (E2) still continue to hold because B and D again vary with these two factors through their dependence on $\tau_{ie}(a)$. If we then use $\tau_i(a)$ to denote the τ function for genotype i, averaged over the noise distribution, the calculations of Appendix B also carry through unchanged.