Beyond R0 maximisation: on Pathogen Evolution and Environmental Dimensions

Sébastien Lion, Johan A.J. Metz

Trends in Ecology & Evolution

CellPress
REVIEWS

Review

Beyond R_0 Maximisation: On Pathogen Evolution and Environmental Dimensions

Sébastien Lion[1,*] and Johan A.J. Metz[2,3,4]

A widespread tenet is that evolution of pathogens maximises their basic reproduction ratio, R_0. The breakdown of this principle is typically discussed as exception. Here, we argue that a radically different stance is needed, based on evolutionarily stable strategy (ESS) arguments that take account of the 'dimension of the environmental feedback loop'. The R_0 maximisation paradigm requires this feedback loop to be one-dimensional, which notably excludes pathogen diversification. By contrast, almost all realistic ecological ingredients of host–pathogen interactions (density-dependent mortality, multiple infections, limited cross-immunity, multiple transmission routes, host heterogeneity, and spatial structure) will lead to multidimensional feedbacks.

Highlights

Contrary to established wisdom, selection in the long run rarely favours parasites that maximise their epidemiological basic reproduction ratio, R_0.

R_0 maximisation only occurs in models with simple forms of environmental feedback.

In realistic host–parasite interactions, ecological processes will commonly preclude R_0 maximisation.

The dimension of the environmental feedback loop here emerges as a unifying concept.

R_0 Maximisation and the Adaptive Theory of Virulence

The idea of R_0 maximisation is intimately linked with the development of the adaptive theory of virulence [1]. Virulence has long been thought of as a transient state in pathogen evolution, with avirulence being the expected long-term evolutionary endpoint [2–4], based on the rationale that harming the host would deplete the resources required by the pathogen. This 'classical wisdom' was challenged by modern adaptive explanations [1,5], according to which natural selection also can lead to an increase in virulence when this confers an indirect benefit to the pathogen. This happens, for example, when increasing virulence goes together with increasing transmission (the transmission–virulence trade-off hypothesis, reviewed in [6]). More generally, virulence may be connected to other disease parameters, such as recovery or within-host competitive ability. Virulence is predicted to evolve towards intermediate values whenever such connections are sufficiently strong.

The textbook explanation for evolution towards intermediate virulence assumes that long-term evolution results in maximising the following quantity (Equation 1),

$$R_0 = \frac{\beta S}{\mu + \alpha + \gamma} \qquad [1]$$

known as the basic reproduction ratio. In Equation 1, βS is the rate at which an infected host produces new infections in a susceptible population of density S; α is the virulence, equated to pathogen-induced mortality; μ is the mortality rate of uninfected hosts; and γ is the recovery rate. Equation 1 has great didactic power, because it immediately shows that, even though an increase in virulence has a direct negative effect on R_0, it can also have indirect positive effects if transmission increases with virulence, or recovery decreases with virulence. Thus, the virulence that maximises R_0 depends on the trade-off between virulence and other disease parameters. This idea has been extremely influential and has been shaping the theory of virulence evolution ever since [6–8]. However, the apparent simplicity of the argument obfuscates two caveats, as we discuss below. First, R_0 can only be written as Equation 1 under strong assumptions on the underlying epidemiological dynamics [9]. By contrast, the transmission–virulence trade-off

[1]Centre d'Écologie Fonctionnelle et Évolutive, CNRS, Université de Montpellier, Université Paul-Valéry Montpellier 3, EPHE, IRD, 1919, route de Mende, 34293 Montpellier Cedex 5, France
[2]Evolution and Ecology Program, International Institute of Applied Systems Analysis, A-2361 Laxenburg, Austria
[3]Mathematical Institute & Institute of Biology, Leiden University, P.O. Box 9512, 2300RA Leiden, The Netherlands
[4]Netherlands Centre for Biodiversity, Naturalis, P.O. Box 9517, 2300RA Leiden, The Netherlands

*Correspondence: sebastien.lion@cefe.cnrs.fr (S. Lion).

 https://doi.org/10.1016/j.tree.2018.02.004

hypothesis fits a larger class of epidemiological scenarios. Second, there is no guarantee that evolution selects for trait combinations maximising the R_0 of such a scenario: almost all realistic ecological ingredients of natural host–pathogen interactions flout the R_0 maximisation paradigm.

Although the theoretical literature has repeatedly emphasised these caveats [10–16], this has had less impact than deserved. The idea that pathogens evolve to maximise their R_0 is still a cornerstone of textbook discussions of virulence evolution. Thus, this idea remains widespread in the community, despite regular corroboration in discussions of the experimental evidence that this is far from general (e.g., [7]).

One possible explanation for this state of affairs is that empirical and theoretical examples where R_0 maximisation fails are typically discussed as exceptions, instead of from a general conceptual perspective. Here, we provide such a perspective through the notion of environmental feedback, that is, the effect of a mutant substitution on the ecology and, thus, the fitness of subsequent mutants. For example, the increase in frequency of a more virulent strain may cause the population density to decrease; this in turn leads to lower density-dependent mortality, which feeds back positively on the mutant fitness. We argue that a precise distinction between pathogen fitness and the epidemiological R_0 is a prerequisite for any discussion of the adaptive evolution of pathogens. We then discuss the main theoretical result that R_0 maximisation will only occur when the feedback through the environment is of a very simple kind, and illustrate this point by reviewing the evolutionary consequences of several realistic features of host–pathogen interactions. Throughout, we emphasise that, although R_0 maximisation may once have been a useful paradigm and may still be a good didactical tool, a more general conceptual framework based on **ESS** (see Glossary) arguments is needed for a proper understanding of the evolution of infectious diseases.

R_0 in Epidemiological Models

The general definition of R_0 in life-history theory is 'the average lifetime offspring number in a given environment'. In epidemiology, R_0 is typically defined as the average number of secondary infections produced by a single infected host in an otherwise uninfected host population [1,9,17–19] (see [20] for its historic roots). The emphasis on 'uninfected hosts' is crucial because R_0 is a function of not only the pathogen traits, X, but also of the environment, E, experienced by the pathogens. Thus, we may write $R_0(X|E)$, where in general X and E comprise more than one variable. For instance, the environment could collect the densities of susceptible and partially resistant hosts. The dependence on the environment reflects our intuition that pathogen spread will be hindered if the environment is less favourable, for instance if the frequency of resistant hosts is high.

For a pathogen to spread in an initially uninfected population, an infected individual must produce more than one secondary infection. Hence, the following condition must hold (Equation 2):

$$R_0(X|E_0) > 1 \qquad [2]$$

where E_0 is the environment produced by the dynamics of the host population in the absence of the pathogen. In the epidemiological literature, $R_0(X|E_0)$ is generally shortened as R_0. We follow this convention and write $R_0(X)$ for $R_0(X|E_0)$. To distinguish this from the more general case, we use $\mathscr{R}(X|E)$ to represent the basic reproduction ratio calculated in another environment E (Box 1).

Glossary

Evolutionarily stable strategy (ESS): a strategy that, if sufficiently common, creates an environment in which no alternative strategy can invade.

Invasion fitness: per-capita growth rate of a rare mutant strain in the environment created by the resident population. This can be written as a function of the traits and of the environment, $\rho(Y|\hat{E})$, or as a function of the mutant and resident traits, $s(Y|X)$.

Fitness proxy: any function of the traits and the environment that has the same sign as invasion fitness and, therefore, provides the same information about long-term evolution.

Fitness component: a property of the traits (and possibly the environment) that enters into the calculation of, but is not on its own sufficient to compute, a fitness proxy.

Optimisation principle: a function $\psi(X)$ of the traits such that, for any constraint on the traits, the ESSs can be calculated by maximising this function (for instance, R_0 in the classical SIR model).

Pessimisation principle: a function $\phi(E)$ of the environment that is minimised at an ESS, for any constraint on the traits (for instance, the density of susceptibles in the classical SIR model).

Effective dimension of the environmental feedback loop: the term dimension of the feedback loop refers to the number of environmental variables (such as the density of susceptible hosts) that are controlled by the population dynamics of the pathogen and influence $\mathscr{R}$ in different manners. However, for ESS calculations, only the sign of $\mathscr{R}$–1 matters. The term 'effective dimension' refers to the number of variables that independently influence this sign. In simple models, the effective dimension and the dimension are often equal, but in structured models, exceptions where the effective dimension is lower are commonplace.

Box 1. The Many Guises of R_0

The general argument we give in this article also extends to more general ecological scenarios. Indeed, although R_0 has become a cornerstone of epidemiological thinking, the historical roots of the concept are in demography and life-history theory. Here, we give a brief historical perspective to shed light on these connections.

R_0 in Demography

In epidemiology, the '0' in R_0 is often interpreted as referring to the uninfected population, but the notation in fact comes from human demography, where R_0 was first defined in [21] as the zeroth in a series of moments of the so-called 'reproduction kernel'; that is, the mean rate of producing children as a function of age.

R_0 in Life-History Theory

The R_0 concept was put to good use in life-history theory, where it is generally taken to be the life-time offspring production of ordinary individuals with a sequestered germ line. For general ecological scenarios, R_0 can be calculated as the dominant eigenvalue of the so-called 'next-generation operator', which, in the given environment, projects the state of the population from one generation to the next [9].

R_0 in Epidemiology

The calculation of R_0 in epidemiology proceeds in the same manner as in life-history theory. However, although it is a pathogen property, it is defined at a higher level, that of infected hosts. From a fundamental perspective, a population of infected hosts is a metapopulation of pathogens and, thus, the epidemiological R_0 corresponds with the R_0-like concepts for metapopulations, such as R_m [22–24] in evolutionary ecology.

On Notation

In the main text, we use different notations for the basic reproduction ratios computed in the pathogen-free population, $R_0(X)$, and in another environment where the host population is already infected by resident pathogen strains, $\mathcal{R}(X|E)$. This is done for clarity, but the common conceptual underpinning should be kept in mind.

In practice, the calculation of $R_0(X)$ as a function of pathogen parameters will lead to different expressions depending on the life cycle of the host–pathogen interaction that one considers. For instance, $R_0(X)$ does not take the same form for directly transmitted or vector-borne pathogens [9,19]. However, most discussions on pathogen evolution start with Equation 1, which is obtained using the classical susceptible–infected–recovered (SIR) epidemiological model (Box 2).

Let us assume that the traits of the pathogen may affect transmission (β), virulence (α), and recovery (γ), reflecting potential trade-offs between life-history traits [1,6]. Then, in the SIR model, $R_0(X)$ can be written as Equation 3:

$$R_0(X) = \frac{\beta(X)}{\mu + \alpha(X) + \gamma(X)} S_0 \qquad [3]$$

where S_0 is the equilibrium density of susceptible hosts in the absence of the pathogen (Box 2).

Equation 3 shows that, for the SIR model, R_0 equals the lifetime 'infection pressure' by an infected individual, $\beta(X)/[\mu + \alpha(X) + \gamma(X)]$, which is an individual-level property, multiplied by a single environmental variable, S_0 (the density of susceptible hosts in a pathogen-free population, which is a population-level property). This distinction between individual and population-level properties will prove essential in the following sections.

Epidemiological R_0 Is Not Pathogen Fitness

Evolution results from the competition between different strains, generally one or more resident strains and the mutants that they produce. This process is endlessly repeated as new mutants keep evolving and are either expelled or become new residents. Fitness is a measure of competitive prowess. In the pathogen-free environment, there is no competition among pathogens and, therefore, $R_0(X)$ cannot be expected to stand as a proxy for pathogen fitness

Box 2. The Standard SIR Model

The standard SIR model divides the host population into three compartments: susceptible (S), infected (I), and recovered (R) hosts. The model assumes that the disease is only transmitted horizontally through direct contacts with an infected host. Transitions between compartments are due to transmission and recovery events. Hosts can be removed from the population through mortality, while new susceptible hosts are created through reproduction. This is depicted in Figure I.

The dynamics of each class of hosts can then be captured by the following system of differential equations (Equations I–III):

$$\frac{dS}{dt} = b(S,I,R) - \mu S - \beta SI + \nu R \quad [I]$$

$$\frac{dI}{dt} = \beta SI - (\mu + \alpha + \gamma)I \quad [II]$$

$$\frac{dR}{dt} = \gamma I - (\mu + \nu)R \quad [III]$$

where $b(S,I,R)$ is the birth rate into the population, μ is the natural mortality, α represents pathogen-induced mortality (often equated to virulence in the theoretical literature), γ is the per-capita recovery rate, ν is the per-capita rate of immunity loss, and β is the transmissibility of the pathogen.

In a pathogen-free population, the demography of hosts will bring the host population to an equilibrium S_0. From Equation II, an initially rare infection will grow if (Equation IV):

$$\beta S_0 - (\mu + \alpha + \gamma) > 0, \quad [IV]$$

which can be rewritten as $R_0 > 1$ with $R_0 = \beta S_0/(\mu + \alpha + \gamma)$.

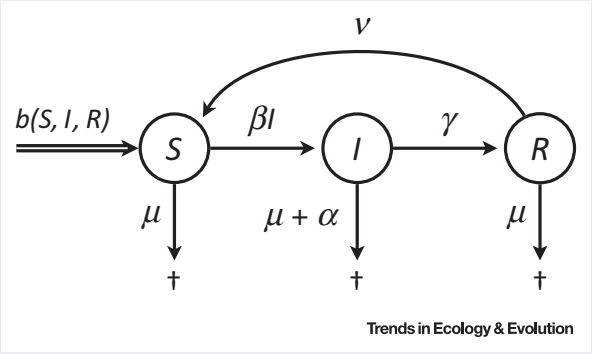

Figure I. Flow Diagram for the Standard SIR Model.

without a multitude of other assumptions. To study long-term evolution, we should rather use **invasion fitness**, defined as the per capita growth rate of the mutant population in a resident population that has reached its epidemiological attractor (Box 3). Alternatively, we can use a **fitness proxy,** such as $\mathcal{R}(Y|\hat{E}) - 1$, which has the same sign as invasion fitness. This fitness proxy also relies on a basic reproduction ratio, $\mathcal{R}(Y|\hat{E})$, but one that is measured in the environment determined by the resident pathogen strains, $\hat{E}$, instead of the pathogen-free environment, E_0.

In the simple SIR model discussed above, a mutant pathogen strain with traits Y will invade if (Equation 4; Box 3):

$$\mathcal{R}(Y|\hat{E}) = \frac{\beta(Y)}{\mu + \alpha(Y) + \gamma(Y)}\hat{S} > 1. \quad [4]$$

Box 3. How Should We Define Pathogen Fitness?

To make predictions about long-term evolution, the adaptive dynamics [25,26] framework provides us with a standardised procedure to calculate the fitness of pathogens. If the mutation rate is low, we may assume a separation of timescales between epidemiological and evolutionary dynamics. In other words, we may assume that the environment reaches an epidemiological attractor $\hat{E}(X)$ before a new mutation with trait value, say, Y occurs. With this assumption, the relevant measure of pathogen fitness is the invasion fitness, $\rho(Y|\hat{E})$, which measures the growth of the mutant population in a resident population that has reached its epidemiological attractor. Alternatively, we can use any fitness proxy that has the same sign as $\rho(Y|\hat{E})$. For instance, we can measure population increase in generation time and use $\ln \mathcal{R}(Y|\hat{E})$ or $\mathcal{R}(Y|\hat{E}) - 1$ as a fitness proxy.

Pathogen Fitness in the SIR Model

To fix ideas, let us return to the simple SIR model discussed in Box 1 in the main text. The epidemiological attractor is an endemic equilibrium $(\hat{S}, \hat{I}, \hat{R})$. From the dynamics of the density of hosts infected by the mutant parasite, we have, if we make the usual assumption that recovery from any strain confers immunity to all (Equation I):

$$\rho(Y|\hat{E}) = \beta(Y)\hat{S} - [\mu + \alpha(Y) + \gamma(Y)]. \quad [I]$$

The mutant strain invades if $\rho(Y|\hat{E}) > 0$. Alternatively, this condition can be rewritten as $\mathcal{R}(Y|\hat{E}) > 1$, where (Equation II):

$$\mathcal{R}(Y|\hat{E}) = \frac{\beta(Y)}{\mu + \alpha(Y) + \gamma(Y)}\hat{S}. \quad [II]$$

Although in the SIR model, there is no real practical benefit in using $\mathcal{R}$ instead of ρ, fitness proxies can often considerably simplify the calculations in more complicated ecological scenarios. (A further fitness proxy that in complicated situations is algebraically far simpler, but less interpretable, than R_0 can be found in [27].)

One Thousand and One Expressions for Pathogen Fitness

Equation II is only one of the many expressions for pathogen fitness derived in the theoretical literature when the simplistic assumptions underpinning the SIR model are relaxed. For instance, minor extensions of the SIR model often lead to expressions of the form (Equation III):

$$\mathcal{R}(Y|\hat{E}) = \frac{\beta(Y)\hat{S} + \tau(\hat{E})}{\mu + \alpha(Y) + \gamma(Y) + \delta(\hat{E})}, \quad [III]$$

where the environmental feedback affects both pathogen transmission [through the term $\tau(\hat{E})$] and the average lifetime of hosts infected by the mutant pathogen [through the term $\delta(\hat{E})$]. Examples include models with density-dependent mortality (Equation 7 in the main text), superinfection (Equation 8 in the main text), limited cross-immunity (Equation 9 in the main text) or vertical transmission (Equation 10 in the main text).

Equations 3 and 4 are misleadingly similar. The critical difference is that $\mathcal{R}$ is calculated in an environment characterised by the resident community of pathogen strains, $\hat{E}$, instead of the pathogen-free environment E_0. Moreover, a crucial property of this specific model is that the effect of the environment is captured by a single variable occurring multiplicatively, the equilibrium density of susceptible hosts, $\hat{S}$. Unfortunately this property, on which the R_0 maximisation paradigm hinges, is far from general.

Evolution Will Maximise R_0 Only in Very Simple Environments

The natural stops of evolution through repeated mutant substitutions are ESSs; that is, trait combinations making it impossible for alternative feasible combinations to invade. By definition, an ESS corresponds to a maximum of pathogen fitness in the corresponding environment. This implication extends to any fitness proxy, such as $\mathcal{R}(Y|\hat{E})$, when $\hat{E}$ is chosen to be the environment generated by the ESS. However, the statement 'evolution maximises R_0' is generally taken to mean that one can calculate the evolutionary endpoint by maximising $R_0(X)$, which is simply a function of X, the environment being fixed at its disease-free value E_0. Thus, it is taken for granted that the environment experienced by the mutant pathogen does

not matter, and that there exists a single type of pathogen that has maximal lifetime production of new infections per infected host in all possible environments. The examples detailed below show that we cannot in general expect the same pathogen type to perform best in both disease-free and already-infected populations.

A Necessary and Sufficient Condition for Evolution to Maximise R_0

To elucidate under which conditions the outcome of pathogen evolution can be determined by maximising the epidemiological R_0, it is helpful to turn to more general results on the conditions for the existence of an **optimisation principle**. The latter simply means a function of the traits, $\psi(X)$, such that that we can find potential ESSs by maximising this function. The question 'when does evolution maximise R_0?' then becomes 'when is $R_0(X)$ an optimisation principle?'. It turns out that this occurs if, and only if, the pathogen fitness can be written as Equation 5:

$$\mathscr{R}(Y|\hat{E}) = [R_0(Y)\phi(\hat{E})]^{q(Y,X)} \quad [5]$$

where q is a positive function of the traits [28]; that is, the effect of the environment can be summarised by a function $\phi(\hat{E})$, which multiplicatively affects the epidemiological basic reproduction ratio $R_0(Y)$. For instance, in the SIR model, we can simply obtain the fitness proxy $\mathscr{R}$ (Equation 4) by multiplying R_0 (Equation 3) with a function of the environment $\phi(\hat{E}) = \hat{S}/S_0$, so that Equation 5 is satisfied with $q(Y,X) = 1$.

If Equation 5 holds, the full ESS calculation is mathematically equivalent to maximising $R_0(X)$ [29,30]. To see this, note that, at the resident equilibrium, we have $\mathscr{R}(X|\hat{E}) = 1$, which implies $\phi(\hat{E}) = 1/R_0(X)$. Thus, $\mathscr{R}(Y|\hat{E}) > 1$ if, and only if, $R_0(Y) > R_0(X)$, which leads to the maximisation of R_0. Below, we shall see that Equation 5 can be used to quickly judge whether a given epidemiological model supports an optimisation principle.

Consequence 1: The Evolutionary Maximisation of R_0 Is Equivalent to the Minimisation of the Susceptible Density

The grand idea of R_0 maximisation has a more downbeat counterpart. Instead of looking at whether evolution maximises a function of the trait, $\psi(X)$, one may look at the impact of trait evolution on the environment, $\phi(\hat{E})$. For our baseline SIR model, we have $\phi(\hat{E}) = \hat{S}/S_0 = 1/R_0(X)$. Hence, maximising $R_0(X)$ is equivalent to minimising the equilibrium density of susceptible hosts (Equation 6):

$$\hat{S} = \frac{\mu + \alpha(X) + \gamma(X)}{\beta(X)}. \quad [6]$$

Any mutant that is favoured by evolution has a higher $R_0(X)$, but makes for a lower density of its resource. The process ends when the density of susceptible hosts is so low that no other mutant pathogen can invade. From the view of the pathogen, evolution thus leads to the worst attainable world, a result dubbed a '**pessimisation principle**' [29,30]. Pessimisation principles occur in all models with an optimisation principle. In a purely ecological context, they appear as the principle that, among species competing for a single resource, only the type survives that tolerates the lowest resource density. Similarly, SIR-type epidemiological models tell that a community of parasites will ultimately be dominated by the strain with the highest R_0 [1], which also results in the lowest susceptible density that allows the disease to persist.

The Dimension of the Environmental Feedback Loop

A crucial feature of Equation 5 is that the effect of the environment can be summed up by a single number, $\phi(\hat{E})$, such that increasing ϕ can only change the sign of $\mathscr{R} - 1$ from negative to positive [30]. An environmental feedback of this form is said to be effectively one-dimensional,

because only one variable is needed to describe the effect of the environment on the fitness sign. For instance, in the SIR model, increasing the density of susceptible hosts can only cause $\mathcal{R}$ to go from below 1 to above 1. In such simple environments, selection maximises a model-dependent function of the traits, $\psi(X)$, which only in the simplest scenarios will be $R_0(X)$ [30].

Conversely, any model for which the environmental feedback cannot be effectively summed up by only one variable does not allow for the ESS to be calculated through maximising R_0 [30]. Which is the case can be decided by checking whether the pathogen fitness $\mathcal{W}(Y|\hat{E})$ satisfies Equation 5. In the next section we review a diversity of biological mechanisms that generically give rise to multidimensional environmental feedback loops and thereby cause R_0 maximisation to break down. The long-term evolutionary outcome then can only be found from a full ESS calculation.

Consequence 2: R_0 Maximisation Excludes Diversification

If an optimisation principle exists (in particular if evolution maximises R_0), the evolutionary process is of the simplest kind: any mutant that increases the optimisation criterion goes to fixation until an ESS is reached, so that any ESS is an evolutionary attractor and vice versa [30]. This has one important corollary: polymorphisms are impossible. Thus, a prerequisite for the evolutionary diversification of pathogen populations is that evolution does not maximise anything and does not maximise R_0 in particular. Thus, the R_0 maximisation paradigm faces an immediate empirical challenge, because it is incompatible with any longer term coexistence of different pathogen strategies in nature.

Consequence 3: ESS Trait Values often Differ from Those Obtained from R_0 Maximisation

When the environmental feedback loop is not conducive to diversification, using R_0 maximisation to predict the endpoint of evolution usually leads to quantitative errors. In principle, the magnitude of such errors can be inferred from the structure of the model. Figure 1 shows a graphical tool for deducing what kind of influence the environmental feedback loop may exert. We start by noting that, under a trade-off between transmission and virulence, R_0 maximisation can be cast in a form corresponding to the so-called 'Marginal Value Theorem' [31], which allows the ESS to be found graphically, as depicted in Figure 1A. Suppose, for instance, that the effect of the environmental feedback loop affects the average time a mutant pathogen hangs on to an infected host [an effect captured in Equation III in Box 3 by the term $\delta(\hat{E})$]. This would happen, for instance, when a more virulent resident strain causes a decline in population density, which in turn decreases the density-dependent mortality rate experienced by a mutant parasite. In this case, the graphics tells that this feedback increases or decreases the ESS relative to the outcome of R_0 maximisation depending on whether, in the example under consideration, the added term is positive or negative (Figure 1B,C). The size of the error made by using R_0 maximisation instead of the full ESS calculation depends on the curvature of the trade-off (see Appendix S8 in the supplemental information online). If the value of virulence α_O^* that maximises R_0 lies on a fairly straight section of the trade-off, as in Figure 1B, any small shift from O to A will cause a large deviation of the ESS compared with α_O^*. By contrast, in Figure 1D, where the trade-off has a high curvature around α_O^*, the same shift from O to A will have negligible effect.

Most Biological Scenarios Jar with the R_0 Maximisation Paradigm

The preceding discussion gives a general argument for why the principle of R_0 maximisation can be expected to be misleading, either qualitatively or quantitatively, for most epidemiological scenarios. We now illustrate this general argument for a selection of more realistic biological scenarios. Using the SIR model as baseline, we highlight salient biological factors causing ESS

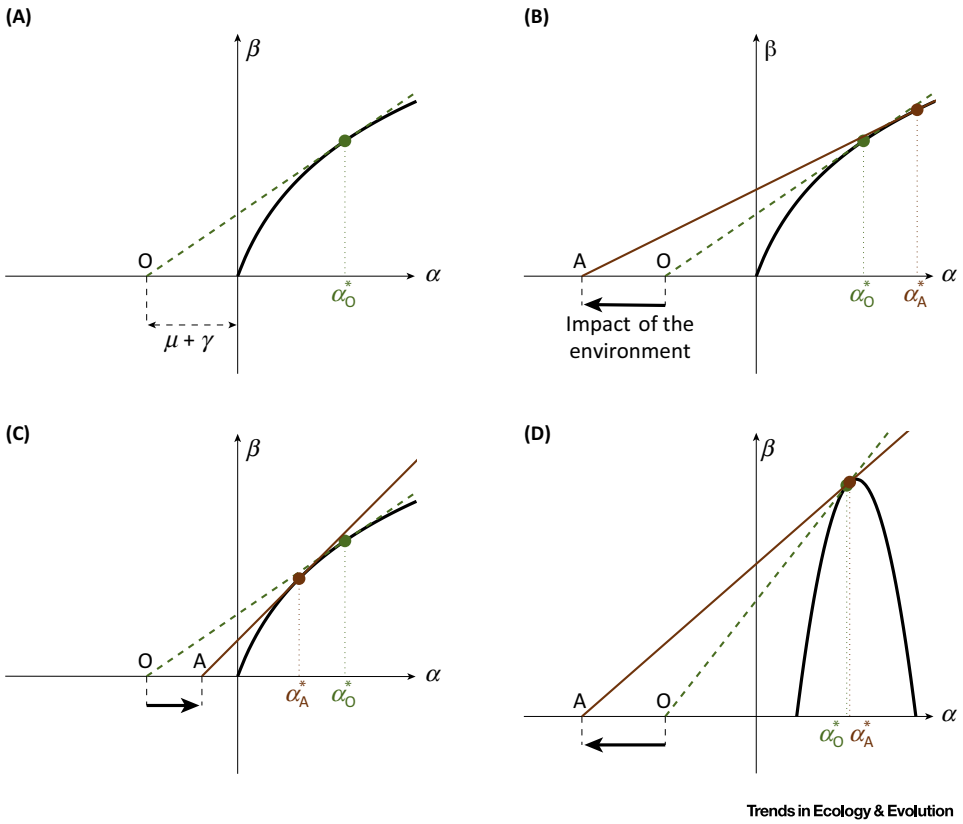

Figure 1. A Graphical Derivation of Quantitative Consequences of R_0 (Non-)Maximisation. (A) Assuming a simple trade-off between transmission (β) and virulence (α), R_0 maximisation in the susceptible–infected–recovered (SIR) model implies that the evolutionarily stable strategy (ESS) (α^*_O) can be found graphically by drawing the tangent at the trade-off curve that goes through the point O = $(-\mu-\gamma, 0)$. (B) With slightly different expressions for pathogen fitness, for instance as given by Equation III in Box 3, the ESS α^*_A will deviate from the prediction of R_0 maximisation due to the additional effect of the environmental feedback loop captured by the term $\delta(\hat{E})$. The tangent at the ESS then goes through the point $A = [-\mu - \gamma - \delta(\hat{E}), 0]$. If $\delta(\hat{E})$ is positive, the point A is to the left of point O and selection favours higher virulence than predicted by R_0 maximisation. (C) By contrast, a negative value of $\delta(\hat{E})$ leads to lower virulence at ESS. (D) The size of the discrepancy $\alpha^*_A - \alpha^*_O$ is inversely proportional to the curvature of the trade-off around the value of virulence α^*_O that maximises R_0 [compare with (B)].

predictions to deviate from the purported predictions coming from an R_0 maximisation (see e.g., [7,8] for reviews in the nontheoretical literature). The aim of our nonexhaustive review is to emphasise the unifying principle connecting these different scenarios, which is to be found in the dimension of the environmental feedback loop. To keep things simple, we use the classical assumption of a trade-off between transmission and virulence (reviewed in [6]) and focus on populations at endemic equilibrium (but see Appendix S7 in the supplemental information online for a discussion of nonequilibrium epidemiological attractors).

Density-Dependent Mortality

The classical SIR model assumes that density dependence only affects fecundity. However, density-dependent mortality has, for example, been identified as a key factor of the evolutionary dynamics of Marek's disease in poultry farms [32]. To take this into account, suppose now that μ is a function of the host densities, say $\mu = \mu_0 + \kappa N$, where $N = S + I + R$ is the total host density. Indicating the mutant properties with a prime, so that, for example,

$\mathscr{R}' = \mathscr{R}(\alpha', \beta' | S, I, R)$, we obtain the following fitness proxy (Equation 7):

$$\mathscr{R}' = \frac{\beta' \hat{S}}{\mu_0 + \kappa \hat{N} + \alpha' + \gamma}. \quad [7]$$

With this simple increment in ecological realism, the environmental feedback affects pathogen fitness in two contrasting ways: as before, pathogen transmission is proportional to the density of susceptible hosts, $\hat{S}$, but, in addition, the duration of infection also decreases with the total population density of the residents, $\hat{N}$, allowing the trait of the residents to exert an additional influence on the fitnesses of mutants. Thus, unless stringent assumptions are made, the effective dimension of the feedback loop is two, that is, there is no way we can sum up the effect of the environment by a single number as in Equation 5. As a result, evolution does not maximise any purported environment-independent fitness proxy. This may notably lead to evolutionary branching [12,14,33], but even when long-term evolution converges to a monomorphic ESS [11,34], the ESS will deviate from the value predicted by R_0 maximisation. Figure 1C graphically depicts this deviation. In this model, the effect of the environment is $\kappa\left(\hat{N} - S_0\right)$ (see Appendix S1 in the supplemental information online). If, as expected, the presence of the pathogens leads to a decrease in the total population size, $\hat{N}$, compared with the density of hosts in an uninfected population, S_0, the point A will be to the right of O and the evolutionarily stable (ES) virulence will be lower than the value that maximises R_0.

Multiple Infections

In nature, hosts are typically infected by several pathogen strains or species [35,36]. When different pathogen strains compete for within-host resources, higher levels of virulence can be selected for [37–39], a prediction backed up by some experimental results in malaria [40]. As an illustration, assume that hosts infected by strain i, if additionally infected by strain j, are then taken over with probability σ_{ji} following rapid within-host competition (so-called 'superinfection' [41]). For a monomorphic resident population, we only need to consider the resident (r) and mutant (m) strains. We then have the following fitness proxy (Equation 8; see also Appendix S2 in the supplemental information online; [39])

$$\mathscr{R}' = \frac{\beta'(\hat{S} + \sigma_{mr}\hat{I})}{\mu + \alpha' + \gamma + \sigma_{rm}\beta\hat{I}}. \quad [8]$$

The feedback of the environment acts through the densities of both susceptible and infected hosts. The total density of hosts that can be infected by a mutant pathogen, $\hat{S} + \sigma_{mr}\hat{I}$, acts as a first feedback variable, with a positive effect on the transmission of all mutant pathogens, the more so for mutants that are better at taking over a resident-infected host (high σ_{mr}). However, a high density of resident-infected hosts, $\hat{I}$, will also increase the risk of a resident take-over (through the term $\sigma_{rm}\beta\hat{I}$) for mutant-infected hosts, resulting in a reduced infection duration, the more so the better the resident is at such a take-over (high σ_{rm}). The presence of two independent feedback variables implies that the long-term evolutionary outcome cannot be predicted by a simple R_0 maximisation. Many theoretical studies have investigated the evolutionary consequences, with three main conclusions: first, superinfection models readily produce evolutionary branching leading to the coexistence of strains with different host exploitation strategies [41–44]. Second, even when diversification is impossible, the ES virulence will be typically higher than the value that maximises R_0, as captured by Figure 1B (point A is to the left of O). Third, the precise evolutionary outcome will generally be due to both the direct effect of within-host competitiveness and the indirect effect of the environmental feedback loop that comes from the take-over pressure by resident pathogens on mutant-infected hosts (see Appendix S2 in the supplemental information online for details).

Limited Cross-Immunity

The classical SIR model assumes full cross-immunity, such that recovered hosts are equally immune to all pathogen strains. However, if mutant pathogens can also infect hosts that have recovered from the resident infection, we obtain the following fitness proxy (Equation 9):

$$\mathscr{R}' = \frac{\beta'}{\mu + \alpha' + \gamma}(\hat{S} + [1 - c(\alpha', \alpha)]\hat{R}) \quad [9]$$

where $c(\alpha',\alpha)$ measures cross-immunity. Full cross-immunity implies $c = 1$, in which case Equation 9 satisfies Equation 5. A reasonable assumption is that cross-immunity is less for more dissimilar trait values. A detailed analysis (see Appendix S3 in the supplemental information online) then shows that the evolutionary dynamics will converge towards the value of virulence that maximises R_0, as in the SIR model with full cross-immunity. However, because c acts similar to a trait-dependent competition coefficient, this value can be a branching point at which the evolutionary path starts to diversify, leading to the coexistence of virulent and prudent pathogens. Several models incorporating limited cross-immunity have demonstrated such diversification (e.g., [45–47]). Hence, although the initial evolutionary dynamics may give the impression that R_0 is maximised, this is not predictive of long-term evolution.

Multiple Transmission Routes

So far, we have only considered pathogens with direct horizontal transmission. Multiple transmission routes are another ubiquitous factor causing an increase in the dimension of the environmental feedback loop. In pathogens with both horizontal and vertical transmission, selection has been found to favour pathogens with suboptimal values of R_0 [16,48–51]. To understand why, we extend the SIR model by allowing the pathogen to be transmitted vertically with probability ε. If $b_I'(N)$ denotes the density-dependent fecundity of hosts infected by the mutant strain, where N is the total population size, this leads to Equation 10 (see Appendix S4 in the supplemental information online for details):

$$\mathscr{R}' = \frac{\beta'\hat{S} + \varepsilon b_I'(\hat{N})}{\mu + \alpha' + \gamma}. \quad [10].$$

Thus, vertical transmission introduces a dependence of fitness on the total population density, in addition to the density of susceptible hosts, and we now have two independent feedback variables. Therefore, according to our general criterion, looking for an optimisation criterion is bound to fail. The key point is not the distinction between horizontal and vertical transmission but the different forms of density dependence introduced by each transmission route. In general, multiple transmission pathways (e.g., sexual versus nonsexual transmission [52] or direct versus environmental transmission [53,54]) introduce separate environmental feedback variables. This may lead to diversification of the pathogen population [52,54–56]. When there is no diversification, arguments similar to those of Figure 1 show that the ESS value of α is smaller than that coming from R_0 maximisation, with the size of the error again inversely proportional to the trade-off curvature (see Appendix S4 in the supplemental information online).

Host Heterogeneity

Most host populations exhibit among-host variation in quality or immune status. This heterogeneity can reflect genetic variation in host resistance or tolerance [57–59], sex-based dimorphism [60], nutritional status, infection history, senescence, environmental factors [61,62], different coinfections [37,63,64], or just different host species. Given that the reproductive potential of the pathogen is likely to differ between host classes, host heterogeneity will generally affect pathogen evolution [63], as shown in host populations with sexual dimorphism [65,66] or intermediate vaccination coverage [67,68]. In addition, because each class of host

potentially produces a separate environmental feedback variable, evolution will optimise some function of the traits only under very specific assumptions on the patterns of infection across classes (Box 4). In principle, host heterogeneity can favour evolutionary branching, because each host class may act as a potential niche for the pathogen. This effect is particularly strong when hosts and pathogens coevolve, in which case diversification in one species can readily lead to the co-diversification of the other species [71–74].

Box 4. Some Evolutionary Consequences of Host Heterogeneity

Pathogen evolution in heterogeneous host populations strongly depends on the pattern of infection across host classes. For a pathogen that can infect two classes of hosts (A and B), different cases can be distinguished.

Unbiased Transmission

Denoting τ_{ij} the transmission rate from class i to class j, this occurs if $\tau_{AA}\tau_{BB} = \tau_{AB}\tau_{BA}$. This property is satisfied in many models that assume that transmission is the product of infectivity and susceptibility; that is, $\tau_{ij} = \beta_i\sigma_j$, where σ_j is the susceptibility of host class j. Biologically, this means that pathogen propagules all pass through a common pool (*cf.* [69]). Then, pathogen fitness can be written as the sum of the basic reproduction ratios in each class of hosts [63,67] (Equation I)

$$\mathcal{R}' = \frac{\beta'_A}{\mu + \alpha'_A + \gamma'_A}\sigma_A\hat{S}_A + \frac{\beta'_B}{\mu + \alpha'_B + \gamma'_B}\sigma_B\hat{S}_B. \quad [I]$$

The fitness proxy depends on two environmental variables, which are the equilibrium densities of susceptible hosts in each class, $\hat{S}_A$ and $\hat{S}_B$. These are given by Equations II and III:

$$\hat{S}_A = \frac{\mu + \alpha_A + \gamma_A}{\sigma_A h/\hat{I}_A} \quad [II]$$

$$\hat{S}_B = \frac{\mu + \alpha_B + \gamma_B}{\sigma_B h/\hat{I}_B}. \quad [III]$$

where $h = \beta_A\hat{I}_A + \beta_B\hat{I}_B$ is the force of infection. We may then distinguish two cases: (i) if the two host classes only differ by their susceptibility to the disease, then pathogen fitness simplifies to the lifetime infectivity multiplied by the total density of susceptibles, $\sigma_A\hat{S}_A + \sigma_B\hat{S}_B$ [39]. If the susceptibilities are independent of the evolving traits, Equation 5 (in the main text) holds true. Thus, the ESS is unaffected by host heterogeneity and is predicted from simple R_0 maximisation using the unstructured SIR model; and (ii) if virulence is different in the two classes, the ESS is intermediate between the optimal virulences predicted from R_0 maximisation in each class in isolation [67,68]. However, there may still exist an optimisation principle if both $\hat{S}_A$ and $\hat{S}_B$ are decreasing functions of a single environmental variable, such as the force of infection h [14].

Biased Transmission

The above analysis breaks down if $\tau_{AA}\tau_{BB} \neq \tau_{AB}\tau_{BA}$. Then, pathogen fitness cannot be written as the sum of the contributions of each class [63]. This generically results in two-dimensional feedback loops, in which case there is no hope of finding a fitness proxy that is maximised by evolution.

Vector-Borne Diseases

A special case where R_0 maximisation can nevertheless do the job is when the two host classes are two host species that need to be exploited in strict alternation, so that $\tau_{AA} = \tau_{BB} = 0$. Then, we have Equation IV:

$$\mathcal{R}' = R'_0(Y)\sqrt{\frac{\hat{S}_A\hat{S}_B}{S_{0,A}S_{0,B}}} \quad [IV]$$

where $R'_0(Y)$ is the R_0 for a vector-borne mutant pathogen in the two-host population in the absence of the disease (see Appendix S5 in the supplemental information for details). Hence, Equation 5 in the main text holds true, and R_0 maximisation works, although the expression for R_0 is not the same as in the SIR model with direct transmission [16,19]. However, the existence of an additional transmission route will cause deviations from the predictions of R_0 maximisation. For instance, several vector-transmitted pathogens have also been shown to be transmitted vertically [70], either in the vertebrate host (e.g., *Plasmodium falciparum*) or in the vector (e.g., several arboviruses [71]).

Spatial Structure

In nature, patterns of local host and pathogen dispersal lead to the build-up of genetic and epidemiological structure, with deep implications for the evolutionary ecology of host–pathogen interactions [75–79]. Consider, for instance, that infectivity decreases with distance. Then, the effective density of susceptible hosts that can be infected by a focal host infected by a mutant pathogen, $[S|I']$, will be lower than the overall density of susceptible hosts in the population, S. This yields the following fitness proxy (Equation 11):

$$\mathscr{R}' = \frac{\beta'[S|I']}{\mu + \alpha' + \gamma}. \quad [11]$$

Although superficially similar to the nonspatial expression, Equation 11 hides a further complication. Given that transmission is mostly local, a mutant pathogen with a higher lifetime infection pressure will, on average, experience a lower density of susceptibles around it. Thus, $[S|I']$ depends on how the traits of the mutant influence the local epidemiological structure experienced by the carriers of the mutant pathogen. As a consequence, the environmental feedback loop generally can only be fully characterised by a large number of variables. However, not all is lost. If we further assume that the resident population is at equilibrium, the invasion condition can be written as in Equation 12:

$$\left(\frac{1}{R_0^{NS}} - \frac{1}{R'^{NS}_0}\right) + ([S|I'] - [S|I]) > 0 \quad [12]$$

where R_0^{NS} (resp. R'^{NS}_0) denotes the lifetime infection pressure exerted by resident (resp. mutant) pathogens in the corresponding nonspatial model [78]. The first term between brackets on its own would lead to the maximisation of the lifetime infection pressure predicted by the nonspatial model. The second term occurs since different pathogen strains experience different densities of susceptible hosts. Therefore, spatial structure is expected to affect the evolutionary outcome [78,80,81]. Further developments of in Equation 12 indicate that the deviation from R_0 maximisation is determined by the balance between genetic structure (local relatedness between pathogens infecting different hosts) and a measure of epidemiological structure for evolutionarily neutral mutants ([78,81]; see Appendix S6 in the supplemental information online for details).

Lessons for the Future

The R_0 maximisation principle is one of many examples in science where a specific result derived for a simple model, or under a particular simplifying assumption, has been promoted to canon status. In epidemiology, other examples include the transmission–virulence trade-off and the representation of virulence as disease-induced mortality, assumptions that underpin many theoretical models. One of our messages is that irreverence for tradition is a key element of scientific progress: we should not let habits or history stifle the development of new ideas. Further progress in the study of pathogen evolution requires explicitly accounting for environmental feedbacks. In this section, we discuss the implications for empirical studies and potential applications.

Should We Attempt to Measure Pathogen Fitness?

The conclusion that selection will only rarely maximise a 'measure of absolute fitness', such as $R_0(X)$, is of interest not only to theoretical biologists. Many empirical studies rely on the presumed measurement of some fitness proxy expected to be maximised by selection. This activity is seldomly informative. First, as we have seen, evolution only rarely satisfies an optimisation principle. Second, empirical measurements of fitness proxys are generally hard

to come by. This is even the case for R_0 and $\mathcal{R}$ since we have to take account of the demography of the full life cycle, which often includes parts that are hard to observe. Third, even if we know how to measure a valid fitness proxy, it is rarely possible to do more than measuring it in the current environment. Then, if the population mean does not sit close to the maximum of the proxy, either something went wrong or we stumbled on a case of fast ongoing evolution, and the result will probably not get reported. If the population mean does sit close to the maximum of the proxy, this tells us only that the population has roughly equilibrated to an ESS, but gives little information on the processes that have brought the population to this point, or where evolution will take the population after an imposed environmental change.

One could object that there is some experimental evidence of R_0 maximisation. However, only a relatively small number of experimental studies appear to support this paradigm. The myxomatosis epidemic in Australian rabbits has been used as such an example [1,82,83]. It is true that, initially, the population quickly settled to a virulence level that was relatively close to the value maximising the classical expression of R_0 [83,84]. However, the subsequent rise of resistance in Australian rabbits then selected for increased virulence [82]. These two phases of the epidemic are characterised by two different environmental feedbacks: during the early years, selection was mostly driven by a strongly curved transmission–virulence trade-off [84], while during the later years, host heterogeneity led to a two-dimensional feedback loop, which precludes R_0 maximisation (see Appendix S9 in the supplemental information online; see also [11]). Thus, the apparent maximisation of R_0 is only a transient state in the coevolution of the myxoma virus and its host. In a similar vein, it was shown the average set-point viral load of HIV in two human cohorts is close to the value that maximises R_0, calculated through an extension of Equation 1 to age-structured populations [85]. However, because the data from which the authors estimated the basic reproductive ratio incorporated the effect of environmental feedbacks, the authors probably estimated the fitness proxy $\mathcal{R}$ rather than the epidemiological R_0. Thus, an alternative interpretation of this result is that its fast evolution causes the HIV population to track a moving optimum of $\mathcal{R}(Y|E(t))$, with $E(t)$ the current environment (see Appendix S10 in the supplemental information online). To predict the outcome of interventions, what really matters is how a large treatment roll-out would impact the environmental feedback on HIV dynamics. This can only be achieved by combining careful empirical studies, as in [85], with the insights of more general ecological theory.

Rather than empirical support for the R_0 maximisation principle, we see these studies as an opportunity to infer conclusions about the form of the environmental feedback in these systems. In some cases, such studies may also help to identify approximate optimisation principles, which can be empirically useful when they exist. An interesting challenge for future theoretical research would be providing empiricists with a theoretical overview of the systems for which simple optimisation principles can be used, together with keys for their empirical identification (see Outstanding Questions). In general, however, trying to measure fitness will not necessarily be the best way to study the adaptive evolution of pathogens. Not only is it a difficult task, but the eventual benefit to our understanding may also often be disappointing. An alternative approach is to use simple models and ESS considerations to generate, and subsequently test, predictions phrased in terms of readily observable quantities, such as the average value of a trait or the frequency of an allele. In this perspective, fitness is best viewed as a theoretical device that can be used to make predictions on more directly measurable properties of biological systems.

Applications and Generalisations

Although for the sake of simplicity we have focussed on the evolution of virulence, there is more to host–pathogen evolution than just virulence. Our main message applies generally to life-history traits affecting the dynamics of host–pathogen interactions and, thus, pertains equally to other problems, such as the evolution of drug resistance or vaccine escape. All this has obvious practical implications for the short- and long-term management of infectious diseases, where one is interested in the evolutionary consequences of some external interference, such as treatments or control measures. For long-term predictions, we have to think beyond adaptations to observed circumstances and consider evolutionary changes of trait values in concert with the environmental changes induced by them. As we have shown, the principle of R_0 maximisation is then of limited use, and we need a more predictive theory, for which we gave some conceptual foundations. At the other extreme, it has long been known that, for short-term predictions, R_0 maximisation is misleading, because strains with higher per-capita growth rates but lower R_0 can be favoured transiently [38,86–88]. Hence, if we want to make predictions about the immediate consequences of a therapeutic intervention, we need to think carefully about how environmental feedbacks play out during transient epidemiological dynamics (see Outstanding Questions).

The message of this article is also relevant for more general problems in evolutionary biology. In fact, the line of argument that we followed here was developed for the evolution of life-history traits in general ecological systems [30]. The many pressing challenges facing evolutionary biologists are all characterised by multidimensional feedbacks between ecological and evolutionary dynamics. To understand the consequences of climate change, habitat fragmentation, or the harvesting of natural resources, an approach based on optimisation does not suffice.

Concluding Remarks

What should a first-principles-based view on the rationale of evolutionary epidemiology look like? For long-term predictions, we see ESS theory and its dynamic counterpart adaptive dynamics, both anchored in the concepts of invasion fitness and dynamical fitness landscapes, as its main pillars. R_0 optimisation did a great job in the early days, but should no longer keep its primacy in teaching and presumed applications since it only finds ESSes under very restrictive conditions. Therefore, emphasising it puts new generations of researchers in the wrong starting block. The challenges raised by emergent infectious diseases, to name but one of the many modern predicaments, require that we give our students the best possible conceptual starting point for tackling the world, and R_0 optimisation fails to fit that bill. The time is ripe for more accurate (and exciting!) approaches to pathogen evolution.

Acknowledgments

J.A.J.M. benefitted from support from the Chaire Modélisation Mathématique et Biodiversité of Veolia Environnement-École Polytechnique-Museum National d'Histoire Naturelle-Fondation X and from a visiting professorship funded by the University of Montpellier and CNRS. The following people very helpfully discussed the initial idea for this manuscript with us: N. Mideo, M. Sofonea, H. Heesterbeek, J. Wallinga, F. van den Bosch, and O. Diekmann (together with the Utrecht theoretical epidemiology discussion group). Two exceedingly good referees then changed its form beyond recognition.

Supplemental Information

Supplemental information associated with this article can be found, in the online version, at https://doi.org/10.1016/j.tree.2018.02.004.

References

1. Anderson, R.M. and May, R.M. (1982) Coevolution of hosts and parasites. *Parasitology* 85, 411–426
2. Smith, T. (1904) Some problems in the life history of pathogenic microorganisms. *Science* 20, 817–832

Outstanding Questions

How can we identify biological systems supporting approximate optimisation principles? For some systems, approximate optimisation principles may be sufficient to predict long-term evolution. Finding guidelines for identifying such systems could prove useful for empirical and experimental studies.

Can we construct useful fitness proxies from simple considerations of the life-cycle of host–parasite interactions? Finding good measures of fitness is a challenge for many empiricists. While the epidemiological R_0 cannot in general be expected to be a valid fitness proxy, a key motivation for further theoretical research is to provide disease ecologists with recipes to build fitness proxies from simple biological observations.

How important are host and parasite population structures in shaping selection on parasite traits? Given that population structure (such as age or spatial structure) can be expected to lead to higher-dimensional environmental feedbacks, we need to better understand to what extent and in which manner such structures influence the outcomes of evolution.

How do environmental feedbacks shape pathogen evolution during transient dynamics? We have assumed here that evolution is slow compared with ecology but, for many host–pathogen systems, evolution may be faster, or unfold on similar timescales. Disease management calls for theory of pathogen evolution during transient epidemiological dynamics.

3. Ball, G.H. (1943) Parasitism and evolution. *Am. Nat.* 77, 345–367
4. Méthot, P.-O. (2012) Why do parasites harm their host? On the origin and legacy of Theobald Smith's 'law of declining virulence' – 1900–1980. *History Philos. Life Sci.* 34, 561–601
5. Ewald, P.W. (1983) Host-parasite relations, vectors, and the evolution of disease severity. *Annu. Rev. Ecol. Syst.* 14, 465–485
6. Alizon, S. *et al.* (2009) Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* 22, 245–259
7. Ebert, D. and Herre, E.A. (1996) The evolution of parasitic diseases. *Parasitol. Today* 12, 96–101
8. Schmid-Hempel, P. (2011) *Evolutionary Parasitology: The Integrated Study of Infections, Immunology, Ecology and Genetics,* Oxford University Press
9. Diekmann, O. *et al.* (1990) On the definition and computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 28, 365–382
10. Bremermann, H.J. and Thieme, H.R. (1989) A competitive exclusion principle for pathogen virulence. *J. Math. Biol.* 27, 179–190
11. Dieckmann, U. (2002) Adaptive dynamics of pathogen-host interactions. In *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management* (Dieckmann, U. *et al.*, eds), pp. 39–59, Cambridge University Press
12. Dieckmann, U. and Metz, J.A.J. (2006) Surprising evolutionary predictions from enhanced ecological realism. *Theor. Popul. Biol.* 69, 263–281
13. Thieme, H.R. (2007) Pathogen competition and coexistence and the evolution of virulence. In *Mathematics for Life Science and Medicine* (Takeuchi, Y., ed.), pp. 123–153, Springer
14. Svennungsen, T.O. and Kisdi, É. (2009) Evolutionary branching of virulence in a single-infection model. *J. Theor. Biol.* 257, 408–418
15. Ferdy, J.-B. and Gandon, S. (2012) Evolution of virulence: intuitions and models. In *New Frontiers of Molecular Epidemiology of Infectious Diseases* (Morand, S. *et al.*, eds), pp. 103–121, Springer
16. Cortez, M.H. (2013) When does pathogen evolution maximize the basic reproductive number in well-mixed host-pathogen systems? *J. Math. Biol.* 67, 1533–1585
17. Macdonald, G. (1952) The analysis of equilibrium in malaria. *Trop. Dis. Bull.* 49, 813–829
18. Dietz, K. (1975) Transmission and control of arboviruses. In *Epidemiology* (Ludwig, D. and Cooke, K.L., eds), pp. 104–121, SIAM
19. Van den Driessche, P. and Watmough, J. (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* 180, 29–48
20. Heesterbeek, J.A.P. and Dietz, K. (1996) The concept of R_0 in epidemic theory. *Statistica Neerlandica* 50, 89–110
21. Dublin, L.I. and Lotka, A.J. (1925) On the true rate of natural increase of a population. *J. Am. Stat. Assoc.* 20, 305–339
22. Metz, J.A.J. and Gyllenberg, M. (2001) How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies. *Proc. R. Soc. B* 268, 499–508
23. Ajar, É. (2003) Analysis of disruptive selection in subdivided populations. *BMC Evol. Biol.* 3, 22
24. Massol, F. *et al.* (2009) The metapopulation fitness criterion: proof and perspectives. *Theor. Popul. Biol.* 75, 183–200
25. Geritz, S.A.H. *et al.* (1998) Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* 12, 35–57
26. Metz, J.A.J. (2012) Adaptive dynamics. In *Encyclopedia of Theoretical Ecology* (Hastings, A. and Gross, L.J., eds), pp. 7–17, California University Press
27. Metz, J.A.J. and Leimar, O. (2011) A simple fitness proxy for structured populations with continuous traits, with case studies on the evolution of haplo-diploids and genetic dimorphisms. *J. Biol. Dynamics* 5, 163–190
28. Metz, J.A. and Geritz, S.A. (2016) Frequency dependence 3.0: an attempt at codifying the evolutionary ecology perspective. *J. Math. Biol.* 72, 1011–1037
29. Mylius, S.D. and Diekmann, O. (1995) On evolutionarily stable life histories, optimization and the need to be specific about density dependence. *Oikos* 74, 218–224
30. Metz, J.A.J. *et al.* (2008) When does evolution optimize? *Evol. Ecol. Res.* 10, 629–654
31. Charnov, E.L. (1976) Optimal foraging: the marginal value theorem. *Theor. Popul. Biol.* 9, 129–136
32. Rozins, C. and Day, T. (2017) The industrialization of farming may be driving virulence evolution. *Evol. Appl.* 10, 189–198
33. Andreasen, V. and Pugliese, A. (1995) Pathogen coexistence induced by density-dependent host mortality. *J. Theor. Biol.* 177, 159–165
34. Pugliese, A. (2002) On the evolutionary coexistence of parasite strains. *Math. Biosci.* 177 (Supplement C), 355–375
35. Petney, T.N. and Andrews, R.H. (1998) Multiparasite communities in animals and humans: frequency, structure and pathogenic significance. *Int. J. Parasitol.* 28 (3), 377–393
36. Balmer, O. and Tanner, M. (2011) Prevalence and implications of multiple-strain infections. *Lancet Infect. Dis.* 11, 868–878
37. van Baalen, M. and Sabelis, M.W. (1995) The dynamics of multiple infection and the evolution of virulence. *Am. Nat.* 146 (6), 881–910
38. Frank, S.A. (1996) Models of parasite virulence. *Q. Rev. Biol.* 71 (1), 37–78
39. Gandon, S. *et al.* (2001) Host life history and the evolution of parasite virulence. *Evolution* 55 (5), 1056–1062
40. de Roode, J.C. *et al.* (2005) Virulence and competitive ability in genetically diverse malaria infections. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7624–7628
41. May, R.M. and Nowak, M.A. (1994) Superinfection, metapopulation dynamics, and the evolution of diversity. *J. Theor. Biol.* 170 (1), 95–114
42. Boldin, B. and Diekmann, O. (2008) Superinfections can induce evolutionarily stable coexistence of pathogens. *J. Math. Biol.* 56, 635–672
43. Boldin, B. *et al.* (2009) Superinfections and adaptive dynamics of pathogen virulence revisited: a critical function analysis. *Evol. Ecol. Res.* 11, 153–175
44. Adler, F.R. and Mosquera Losada, J. (2002) Super- and coinfection: filling the range. In *Adaptive Dynamics of Infectious Diseases: In Pursuit of Virulence Management, Cambridge Studies in Adaptive Dynamics* (Dieckmann, U. *et al.*, eds), pp. 138–149, Cambridge University Press
45. Adams, B. and Sasaki, A. (2007) Cross-immunity, invasion and coexistence of pathogen strains in epidemiological models with one-dimensional antigenic space. *Math. Biosci.* 210, 680–699
46. Alizon, S. and van Baalen, M. (2008) Multiple infections, immune dynamics and virulence evolution. *Am. Nat.* 172, E150–E158
47. Best, A. and Hoyle, A. (2013) A limited host immune range facilitates the creation and maintenance of diversity in parasite virulence. *Interface Focus* 3, 20130024
48. Nowak, M.A. (1991) The evolution of viruses: Competition between horizontal and vertical transmission of mobile genes. *J. Theor. Biol.* 150, 339–347
49. Lipsitch, M. *et al.* (1996) The evolution of virulence in pathogens with vertical and horizontal transmission. *Evolution* 50, 1729–1741
50. Messenger, S.L. *et al.* (1999) Virulence evolution in a virus obeys a trade-off. *Proc. R. Soc. B* 266, 397–404
51. Ferdy, J.B. and Godelle, B. (2005) Diversification of transmission modes and the evolution of mutualism. *Am. Nat.* 166, 613–627
52. Thrall, P.H.T. and Antonovics, J. (1997) Polymorphism in sexual versus non-sexual disease transmission. *Proc. R. Soc. B* 264, 581–587

53. Day, T. (2002) Virulence evolution via host exploitation and toxin production in spore-producing pathogens. *Ecol. Lett.* 5, 471–476
54. Boldin, B. and Kisdi, É. (2012) On the evolutionary dynamics of pathogens with direct and environmental transmission. *Evolution* 66, 2514–2527
55. Bernhauerová, V. and Berec, L. (2015) Role of trade-off between sexual and vertical routes for evolution of pathogen transmission. *Theor. Ecol.* 8, 23–36
56. Hamelin, F.M. *et al.* (2016) The evolution of plant virus transmission pathways. *J. Theor. Biol.* 396, 75–89
57. Dwyer, G. *et al.* (1997) Host heterogeneity in susceptibility and disease dynamics: tests of a mathematical model. *Am. Nat.* 150, 685–707
58. Råberg, L. *et al.* (2007) Disentangling genetic variation for resistance and tolerance to infectious diseases in animals. *Science* 318, 812–814
59. Keith, R. and Mitchell-Olds, T. (2013) Genetic variation for resistance to herbivores and plant pathogens: hypotheses, mechanisms and evolutionary implications. *Plant Pathol.* 62, 122–132
60. Nunn, C.L. *et al.* (2009) On sexual dimorphism in immune function. *Philos. Trans. R. Soc. B* 364, 61–69
61. Sorci, G. *et al.* (2013) Immunity and the emergence of virulent pathogens. *Infect. Genet. Evol.* 16, 441–446
62. Sorci, G. *et al.* (2013) Immune evasion: immunopathology and the regulation of the immune system. *Pathogens* 2, 71–91
63. Gandon, S. (2004) Evolution of multihost parasites. *Evolution* 58, 455–469
64. Lion, S. (2013) Multiple infections: kin selection and the evolutionary epidemiology of parasite traits. *J. Evol. Biol.* 26, 2107–2122
65. Cousineau, S.V. and Alizon, S. (2014) Parasite evolution in response to sex-based host heterogeneity in resistance and tolerance. *J. Evol. Biol.* 27, 2753–2766
66. Úbeda, F. and Jansen, V.A. (2016) The evolution of sex-specific virulence in infectious diseases. *Nat. Commun.* 7, 13849
67. Gandon, S. *et al.* (2001) Imperfect vaccines and the evolution of pathogen virulence. *Nature* 414, 751–756
68. Gandon, S. *et al.* (2003) Imperfect vaccination: some epidemiological and evolutionary consequences. *Proc. R. Soc. B* 270, 1129–1136
69. Rueffler, C. and Metz, J.A.J. (2013) Necessary and sufficient conditions for R_0 to be a sum of contributions of fertility loops. *J. Math. Biol.* 66, 1099–1122
70. Ebert, D. (2013) The epidemiology and evolution of symbionts with mixed-mode transmission. *Annu. Rev. Ecol. Evol. Syst.* 44, 623–643
71. Lequime, S. *et al.* (2016) Determinants of arbovirus vertical transmission in mosquitoes. *PLoS Pathog.* 12, e1005548
72. Pugliese, A. (2011) The role of host population heterogeneity in the evolution of virulence. *J. Biol. Dyn.* 5, 104–119
73. Best, A. *et al.* (2009) The implications of coevolutionary dynamics to host-parasite interactions. *Am. Nat.* 173, 779–791
74. Best, A. *et al.* (2010) The evolution of host-parasite range. *Am. Nat.* 176, 63–71
75. Greischar, M.A. and Koskella, B. (2007) A synthesis of experimental work on parasite local adaptation. *Ecol. Lett.* 10, 418–434
76. Jousimo, J. *et al.* (2014) Ecological and evolutionary effects of fragmentation on infectious disease dynamics. *Science* 344, 1289–1293
77. Tack, A.J. and Laine, A.-L. (2014) Spatial eco-evolutionary feedback in plant-pathogen interactions. *Eur. J. Plant Pathol.* 138, 667–677
78. Lion, S. and Gandon, S. (2015) Evolution of spatially structured host populations. *J. Evol. Biol.* 28, 10–28
79. Parratt, S.R. *et al.* (2016) Infectious disease dynamics in heterogeneous landscapes. *Annu. Rev. Ecol. Evol. Syst.* 47, 283–306
80. Boots, M. and Sasaki, A. (1999) 'Small worlds' and the evolution of virulence: infection occurs locally and at a distance. *Proc. R. Soc. B* 266, 1933–1938
81. Lion, S. and Boots, M. (2010) Are parasites 'prudent' in space? *Ecol. Lett.* 13, 1245–1255
82. Fenner, F. and Fantini, B. (1999) *Biological Control of Vertebrate Pests,* CABI Publishing
83. Mackinnon, M.J. *et al.* (2008) Virulence evolution in response to vaccination: the case of malaria. *Vaccine* 26 (Suppl 3), C42–C52
84. Massad, E. (1987) Transmission rates and the evolution of pathogenicity. *Evolution* 41, 1127–1130
85. Fraser, C. *et al.* (2007) Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17441–17446
86. Lenski, R.E. and May, R.M. (1994) The evolution of virulence in parasites and pathogens: reconciliation between two competing hypotheses. *J. Theor. Biol.* 169, 253–265
87. Day, T. and Gandon, S. (2007) Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* 10, 876–888
88. Bull, J.J. and Ebert, D. (2008) Invasion thresholds and the evolution of nonequilibrium virulence. *Evol. Appl.* 1, 172–182