doi:10.1111/j.1420-9101.2009.01726.x

The Price equation framework to study disease within-host evolution

S. ALIZON

Institut für Integrative Biologie, ETH, Zürich, Switzerland

Keywords:

co-evolution;
immune system;
infectious diseases;
trade-offs;
within-host dynamics.

Abstract

The evolution of infectious diseases is known to affect epidemiological dynamics, but, for some viruses and bacteria, this evolution also takes place inside a host during the course of an infection. I develop an original approach to study intrahost evolutionary dynamics of quantitative disease traits. This approach can be expressed mathematically using the 'Price equation' framework recently developed in evolutionary epidemiology. This framework combines population genetics and within-host population dynamics models to identify trade-offs that affect disease intrahost evolution and to predict short-term evolutionary dynamics of life-history traits. I show that this can be applied to study the evolution of viruses competing for host cells or to study the coevolution between parasites and the immune system of the host. This framework can also easily incorporate experimental data. Studying intrahost evolutionary dynamics provides insight at the within-host level, because it allows us to better understand the course of chronic infections, and at the epidemiological level, because it helps to study multi-scale evolutionary processes. This framework can be used to address important biological issues, from immune escape to disease evolutionary response to treatments.

Introduction

Infectious diseases are arguably one of the best subjects on which to apply Darwin's natural selection theory; first, because parasites tend to evolve rapidly (due to short generation times and large population sizes, Price, 1977) and, second, because disease evolution raises important public health and socio-economical issues (Palumbi, 2001). The 1980s saw the emergence of a new field today known as 'evolutionary epidemiology' that takes into account disease evolution during an epidemic (Anderson & May, 1982). More recently, arguably because of the human immunodeficiency virus (HIV) pandemic, it has been realized that diseases also evolve continuously inside hosts (Holmes *et al.*, 1992) and that this evolution can shape the course of an infection (Nowak *et al.*, 1990). Here, I develop a framework that can combine theory and data to study intrahost evolutionary dynamics. This framework can be applied to micro-parasites that cause chronic infections [e.g. viruses such as HIV and hepatitis C (Alfonso *et al.*, 2005; Lemey *et al.*, 2006) or bacteria such as *Helicobacter pylori* (Suerbaum & Josenhans, 2007)], but also to other diseases such as cancer (Merlo *et al.*, 2006).

Currently, one of the most developed frameworks used to study within-host evolution is based on the quasispecies theory, that is, the evolution of a population of RNA (or DNA) sequences that replicate with a high mutation rate and where each sequence has a given fitness (Eigen *et al.*, 1988). Studies based on this theory use molecular data and/or mathematical tools to follow the evolution of the viral genome during an infection. There has been a lively debate in the field concerning the applicability of the quasispecies theory and highlighting its links with population genetics seems to be helpful for addressing the strengths and limitations of this theory (Wilke, 2005). A first limitation lies with the definition of the fitness landscape, which is usually assumed to be constant [but see Kamp (2003) for a discussion on how landscapes can

Correspondence: Samuel Alizon, Institut für Integrative Biologie, ETH, Universitätstrasse 16, 8092 Zürich, Switzerland.
Tel.: +41 44 632 8976; fax: +41 44 632 1271;
e-mail: samuel.alizon@env.ethz.ch

be shaped by the immune response]. A second limitation is that quantitative traits are difficult to study with the quasispecies framework. This second limitation is shared by many models, as most studies on within-host evolution mainly focus on qualitative traits, such as immune evasion or drug resistance (Goulder & Watkins, 2004). For instance, this is the case for the model Kelly *et al.* (2003) develop that links population genetics and population dynamics to study viral rates of evolution (see their study for further references of models of within-host evolution). This is also the case of many theoretical (Nowak *et al.*, 1990) and empirical studies (Shriner *et al.*, 2006) on HIV that investigate the effect of parasite diversity on the course of an infection.

Parasite replication rate provides a noticeable exception to the lack of experimental (and theoretical) studies on quantitative traits. This rate has been shown to vary among strains of the same virus species (Bocharov *et al.*, 2004; Dykes & Demeter, 2007). From a theoretical point of view, the dominant view is summarized by Levin & Bull (1994)'s 'short-sighted' evolution scenario. They argue that during an infection, strains with greater replication rates are always favoured, which can eventually lead to high rates of host mortality that are not adaptive at an epidemiological level. Several models have investigated this conflict between levels of selection. Bonhoeffer & Nowak (1994) assume that there is no within-host coexistence (a mutant disappears or takes over the host instantaneously) and their results confirm Levin and Bull's verbal reasoning. This model was extended to study the effect of the duration of the infection and the epidemiological feedbacks in further details (André & Godelle, 2006). It was also extended to include immune dynamics and, for eight different models of within-host dynamics, viral evolution always decreased host fitness (Iwasa *et al.*, 2005). Another body of theoretical work based on kin selection theory predicts that the evolutionary outcome is less straightforward if the co-infecting strains are related (Frank, 1996; Chao *et al.*, 2000; Brown *et al.*, 2002). Finally, recent models show that if one assumes a trade-off between viral production rate and the life span of an infected cell (Gilchrist *et al.*, 2004; Ball *et al.*, 2007), intrahost evolution favours strains with intermediate replication rates. The most noticeable feature of these models is that each of the many strains can have its own replication rate. However, these models do not study traits other than replication rate and they usually fail to predict short-term evolutionary dynamics of the trait. Also, modelling the stochastic emergence of mutants is often challenging.

The framework I develop here combines within-host population dynamics models and population genetics approaches to study disease intrahost evolution. Contrary to previous works, this new approach can be used to study the evolution of any parasite trait (both qualitative and quantitative). This is made possible by recent advances in evolutionary epidemiology that are based on the 'Price equation' (Day & Proulx, 2004; Day & Gandon, 2006, 2007; Gandon & Day, 2009). This Price equation framework has four advantages: (i) it helps to identify how (and which) trade-offs can affect within-host evolution; (ii) it allows for predicting the short-term evolutionary dynamics of a trait from the genetic composition of the parasite population in the host; (iii) it helps link theory and data; and (iv) it can be applied to most existing models of within-host population dynamics. To introduce the framework, I first present a simple case that corresponds to bacteria facing anti-microbial peptides. Then, I apply this framework to viruses competing for host cells. Finally, I show that the framework can be used with a general model to study parasite–lymphocyte coevolution.

Case 1: Gut bacteria facing anti-microbial peptides

Let us first consider a host in which a diverse parasite population is attacked by a uniform population of immune effectors (Fig. 1a). This is the case for gut or skin bacteria of mammals that face an innate immune response consisting of anti-microbial peptides that are synthesised constitutively (this would also be true in insects or even plants, Hancock & Diamond, 2000). Let us assume that the bacterial population is genetically diverse and that it is composed of n different strains. Each strain i is defined by the values of its traits (here, the replication rate and the sensitivity to the immune effectors). In this

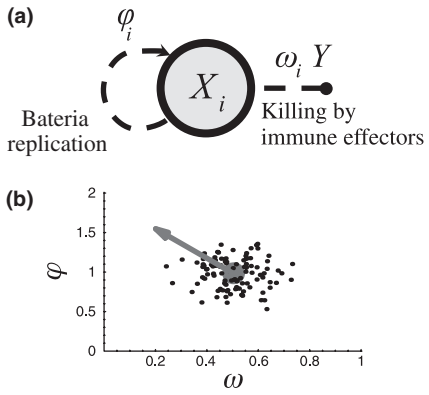

Fig. 1 (a) Model of bacteria facing constant immune pressure and (b) direction of short-term evolution depending on the genetic composition of the population. Each dot in (b) represents a bacterial strain in the host with its sensitivity to the immune response (ω) on the x-axis and its replication rate (φ) on the y-axis. The large grey dot represents the average value of the bacterial population and the arrow indicates the selection pressure on the traits given the current genetic composition of the population. Further details are available in Supplementary protocol S2.

simple case, the density dynamics of bacteria of strain i within the host are governed by the following equation (which is the equation of the prey in the classical Lotka–Volterra predator–prey model):

$$\frac{dX_i}{dt} = (\varphi_i - \omega_i Y)X_i \quad (1)$$

where Y is the density of immune effectors, X_i is the density of bacteria of strain i, φ_i is their replication rate and ω_i is their sensitivity to the immune effectors (notations are in Table 1).

There are no mutation terms from one strain to another in this equation. This simplification can be made because, if the mutation process is not biased (i.e. if the average trait of the mutant population is equal to the average trait of the resident population), the direction of the change of the mean trait values are not affected (see Case 2 below and Supplementary text S1). Also, note that the following analysis still holds if the immune response is a function of time ($Y(t)$), as long as it is independent from the bacteria density.

In Box 1, I use previous results (Day & Gandon, 2006) to derive the equations governing the dynamics of average values of the traits in the bacterial population. The main idea is to track changes in the proportion of each of the strains composing the bacterial population. For instance, an increase in the proportion of the strain with the highest growth rate will result in an increase in the average growth rate. More precisely, I find that:

Table 1 List of the variables and parameter used.

Case	Notation	Description
1	X_i	Density of bacteria of strain i
	n	Number of bacterial strains in the host
	Y	Density of immune effectors
	φ_i	Reproduction rate of bacteria of strain i
	ω_i	Sensitivity of bacteria of strain i to the immune effectors
	r_i	Fitness of parasites of bacterial strain i
2	S	Density of host uninfected target cells
	I_T	Total density of infected cells
	V_T	Total density of free viruses
	λ	host cell production rate
	d	death rate of target cells
	k_i	Infection rate of viruses of strain i
	p_i	Viral production rate of cells infected by strain i
	δ_i	Death rate of cells infected by strain i
	c_i	Clearance rate of free viruses of strain i
3	X_T	Total parasites density
	Y_T	Total lymphocyte density
	φ_i	Reproduction rate of parasites of strain i
	ω_{ij}	Sensitivity of parasites of strain i to lymphocytes of clone j
	a_{ij}	Activation rate of lymphocytes of clone j by parasites of strain i
	μ	Lymphocyte death rate

An i subscript indicates a value relative to one strain and a T subscript indicates a total population size (i.e. a sum over all the strains).

Box 1

This box describes the method used to derive the equation for the trait dynamics in the first case. The goal is to find the dynamics of average life-history trait values by using eqn 1 (see Supplementary text S1 for further details). I here focus on the replication rate but a similar approach can be derived for the sensitivity to the immune system. The first step is to write the equation governing the dynamics of the total density of the bacteria population (X_T):

$$\frac{dX_T}{dt} = (\bar{\varphi} - \bar{\omega}Y)X_T \quad (9)$$

where $\bar{\varphi}$ and $\bar{\omega}$ are the average values of replication rate and sensitivity to the immune effector.
By definition, we have

$$\bar{\varphi} = \sum_{i=1}^{n} p_i \varphi_i \quad (10)$$

where n is the number of parasite strains in the host, φ_i is the replication rate of strain i and p_i is the proportion of strain i defined by

$$p_i = \frac{X_i}{X_T} \quad (11)$$

From eqn 10, we have

$$\frac{d\bar{\varphi}}{dt} = \sum_{i=1}^{n} \frac{dp_i}{dt}\varphi_i + \sum_{i=1}^{n} p_i \frac{d\varphi_i}{dt} \quad (12)$$

If we assume that the traits of a given strain do not vary over time, then the derivative of φ_i with respect to time is zero and the second sum in eqn 12 vanishes.
From eqn 11, we get

$$\frac{dp_i}{dt} = \frac{1}{X_T}\frac{dX_i}{dt} - p_i \frac{1}{X_T}\frac{dX_T}{dt} \quad (13)$$

Using eqns 1 and 9, we can rewrite this equation as

$$\frac{dp_i}{dt} = (\varphi_i - \bar{\varphi})p_i - Y(\omega_i - \bar{\omega})p_i \quad (14)$$

Equation 12 then becomes

$$\frac{d\bar{\varphi}}{dt} = \sum_{i=1}^{n} p_i \varphi_i^2 - \bar{\varphi}^2 - Y\left(\sum_{i=1}^{n} p_i \omega_i \varphi_i - \bar{\omega}\bar{\varphi}\right) \quad (15)$$

Given that $\sum_{i=1}^{n} p_i \varphi_i^2 - \bar{\varphi}^2$ is the variance in φ and that $\sum_{i=1}^{n} p_i \omega_i \varphi_i - \bar{\omega}\bar{\varphi}$ is the covariance between φ and ω, this equation can be written more simply as in eqn 2a.

More generally, previous approaches (Day & Proulx, 2004; Day & Gandon, 2006) show that, for any trait θ, the dynamics of the average value of the trait are given by

$$\frac{d\bar{\theta}}{dt} = \mathrm{Cov}(\theta_i, r_i) - \eta(\bar{\theta} - \bar{\theta}_m) \quad (16)$$

where the Cov term is the covariance between the value of the trait in a strain and the fitness of this particular strain (r_i), η is the mutation rate, $\bar{\theta}$ is the average value of the trait and $\bar{\theta}_m$ is the average value of the trait that arise through mutations (see Supplementary text S1 for further details). If the mutation process is unbiased $\bar{\theta} = \bar{\theta}_m$ and the mutation term goes to zero. This formulation recalls the equation derived by Price (1970).

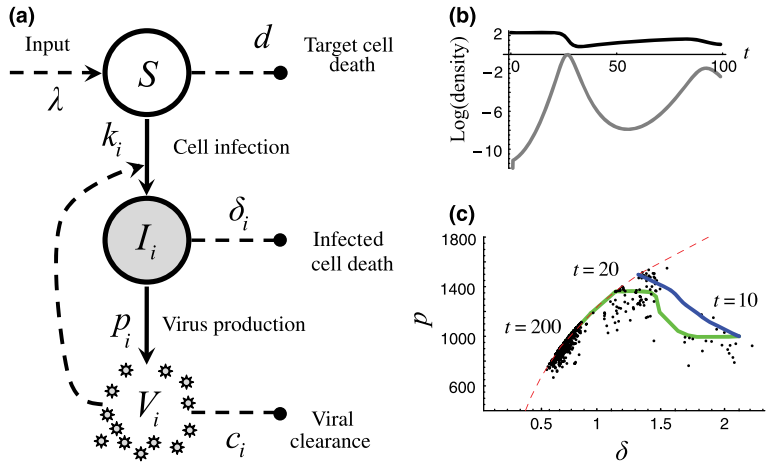

Fig. 2 Evolutionary dynamics in the case of resource competition: (a) model of viruses competing for host cells and (b) dynamics of total infected cells (I_T, grey line) and host target cells (S, black line) densities. (c) Trait evolutionary trajectories in the (δ,p) plane. The blue line (dark grey in the print edition) shows the evolutionary trajectory obtained using the Price equation model and the green line (light grey in the print edition) shows a trajectory obtained using the semi-stochastic model. The three clouds of dots represent the composition of the viral population at three different times (10, 20 and 200 days) in the stochastic model. Parameter values are similar to that used in Ball *et al.* (2007). Further details are available in Supplementary protocol S2.

$$\frac{\mathrm{d}\bar{\varphi}}{\mathrm{d}t} = \mathrm{Var}(\varphi) - Y\ \mathrm{Cov}(\varphi, \omega) \tag{2a}$$

$$\frac{\mathrm{d}\bar{\omega}}{\mathrm{d}t} = \mathrm{Cov}(\varphi, \omega) - Y\ \mathrm{Var}(\omega) \tag{2b}$$

where bars indicate average values (see Box 1), Var indicates the genetic variance of a trait and Cov a genetic covariance between two traits in the bacterial population.

What does this simple example tell us? As long as there is some variance in the bacterial population within the host, the average replication rate tends to increase (because the variance term in eqn 2a is always positive) and the resistance to the immune effectors tends to decrease (because of the variance in eqn 2b). Only a positive covariance between these two traits and the presence of immune effectors ($Y > 0$) can alter the evolutionary outcome. Thus, in this example, a trade-off between bacterial replication and destruction by the immune effector prevents the evolution towards very high replication rates.

This illustrates the first use of the Price equation framework: with the equations governing the dynamics of the trait, we have an overview of which trade-offs can affect the evolution of a trait and how. Here, the trade-off we identify, i.e. that replication increases the sensitivity to the immune response, has been used to study viral evolution (Bocharov *et al.*, 2004; Alizon, 2008). Other trade-offs could of course be considered. For instance, in a similar model that would include a maximum density of the total bacterial population that varies among strains, there could be a trade-off between resource utilization and replication rate, as shown for *Escherichia coli* (King *et al.*, 2004).

The second use of this framework is that, if we know the strain distribution in a host at a given time (and if we are able to measure the trait values of each of these strains), we have access to the variance and covariance values, that is, to the genetic (co)variance matrix (see Day & Gandon, 2007). Using eqns 2a and 2b, we can predict in which direction the system will evolve in the short-term (Fig. 1b). This is an advantage compared to current frameworks, in which getting this information would require to build a simulation with explicit dynamics of all the strains. This model even allows us to follow short-term evolutionary dynamics of the trait. For this, we need the equations describing the dynamics of the trait along with the equations describing changes in total bacterial population size and immune effectors. This is done in the next section with a case that offers richer dynamics.

Case 2: Viruses competing for host cells

Viruses need host cells to reproduce because they rely on the cell machinery to carry out their metabolism. Dynamics of viruses infecting host cells in order to produce new virus have been shown to be accurately described by ecological models of resource exploitation (Perelson, 2002). As shown in Fig. 2a (and in Supplementary text S1), susceptible cells (S), can be infected by free viruses of strain i (V_i) and become infected cells producing these viruses (I_i). The dynamics of the total densities of infected cells (I_T) and free virus (V_T) are given by

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \lambda - \mathrm{d}S - \bar{k}^V V_T S \tag{3a}$$

$$\frac{dI_T}{dt} = \bar{k}^V V_T S - \bar{\delta}^I I_T \tag{3b}$$

$$\frac{dV_T}{dt} = \bar{p}^I I_T - \bar{c}^V V_T \tag{3c}$$

where λ and d are the (constant) production and death rates of target cells, $\bar{k}^V$ is the average rate with which target cells become infected by free-living viruses, $\bar{p}^I$ is the average viral production rate of infected cells, $\bar{\delta}^I$ is the average death rate of infected cells and $\bar{c}^V$ is the average clearance rate of free-living viruses. The V and I superscript indicate if the average value is that of the free-living or within-cell viral population. Because the virus can be found inside the cells or free-living, it is necessary to follow the evolution of the trait in the two types of environments. The underlying idea is that if some strains are better adapted to one of the environments, it biases the strain distribution and thus affects the average trait value. To take this potential bias into account, I here use the epidemiological framework developed by Day & Gandon (2006) to study spore-producing diseases.

For simplicity, and following previous studies (Gilchrist *et al.*, 2004; Ball *et al.*, 2007), we will only focus here on the evolution of two traits: the average viral production rate ($\bar{p}$) and the average death rate of infected cells ($\bar{\delta}$). Also, we will assume that a cell can only be infected by a single viral strain. If we assume that the other parameters are constant and that the mutation process is not biased (see Box 1), the dynamics of the average values of the traits in each of the two environments are given by the following equations (see Supplementary text S1):

$$\frac{d\bar{p}^I}{dt} = -\mathrm{Cov}^I(p,\delta) + \frac{V_T S}{I_T}\left(\bar{p}^V - \bar{p}^I\right)k \tag{4a}$$

$$\frac{d\bar{p}^V}{dt} = \frac{I_T}{V_T}\left[\mathrm{Var}^I(p) + (\bar{p}^I - \bar{p}^V)\bar{p}^I\right] \tag{4b}$$

$$\frac{d\bar{\delta}^I}{dt} = -\mathrm{Var}^I(\delta) + \frac{V_T S}{I_T}\left(\bar{\delta}^V - \bar{\delta}^I\right)k \tag{4c}$$

$$\frac{d\bar{\delta}^V}{dt} = \frac{I_T}{V_T}\left(\mathrm{Cov}^I(\delta,p) + (\bar{\delta}^I - \bar{\delta}^V)\bar{p}^I\right) \tag{4d}$$

For a given trait, the equations in the two environments differ because the selective pressures differ. Each equation has two terms on its right hand side. The first term reflects the selective pressure due to the environment of the virus and the second term reflects virus migration from one environment to the other (when a cell becomes infected or when an infected cell releases viruses). The variance term in eqn 4b tells us that higher average production rates ($\bar{p}$) are always selected for in the free-living stages. Lower production rates can be selected for in the cells if this trait is positively correlated with infected cell death rate (eqn 4a). For the average infected cell death rate ($\bar{\delta}$), lower rates are always selected for within cells and higher rates can be selected for outside the cells if this also increases viral production rate.

The ratio of infected cells to free viruses has a strong influence on the evolution of the viral production rate. In a case where most viruses are within cells (either because it is an early stage of the infection or because they are latent strains), the I_T/V_T ratio is high and selection within a cell (eqn 4a) mostly depends on the covariance between the two traits of interest. On the contrary, outside the cells there is an intense selective pressure that favours strains with greater values of p (eqn 4b). Note that the migration term will homogenize values within and outside cells. In a case where most viruses are free, we have the exact opposite trend: selection for low p outside the cells and selection for high p within the cells. Finally, when the availability of host target cells (S) decreases, strains that exploit their cells for a longer time (i.e. with a low value of p) are favoured. Decreasing S has no direct effect on the selective pressure on free viruses.

Figure 2b shows the evolutionary dynamics of the traits using two different methods. The first method is based on the dynamical system derived using the Price equation framework (see below), which includes the seven equations in systems 3 and 4. The second method is a stochastic simulation with random mutation events, where the densities of each of the strains in the population are followed explicitly (see Supplementary protocol S2 for further details). In both of these cases, it is necessary to assume a trade-off relationship involving the traits of interest; otherwise the system evolves towards infinite production rates and zero mortality of infected cells. Here, following a model developed to study HIV infections (Ball *et al.*, 2007), I assume that increasing production rate (p) comes with a cost in terms of infected cell death rate (δ) (see Supplementary protocol S2). In other words, a strain cannot be more adapted than the trade-off allows (i.e. above the dashed curve in Fig. 2c) but it can be less adapted.

Game-theoretical frameworks can be used to infer the long-term evolutionary outcome of a system (Hammerstein, 1996; Geritz *et al.*, 1997), for example, finding evolutionary stable strategies (or ESS; Maynard-Smith & Price, 1973) and determining if they are convergent stable, that is, if the ESS can be reached (Christiansen, 1991). This long-term outcome depends on the shape of the trade-off curve (Levins, 1962; de Mazancourt & Dieckmann, 2004). In this specific example, the optimal production rate that maximizes the total number of viruses produced by a cell before it dies can be derived analytically (Gilchrist *et al.*, 2004; Ball *et al.*, 2007).

The Price equation framework allows us to analyse short-term dynamics. However, we need to make an assumption in order to give the system dynamic sufficiency, i.e. to be applied recursively over many generations of cells (for a detailed discussion, Gardner *et al.*, 2007). Here, we assume that changes in genetic variances and covariances can be neglected on the short-term, which is appropriate if the mutation process is not biased (see Supplementary text S1). This is known as the assumption of constant variance in quantitative genetics models. Other approaches could be used, for example, estimating higher moments of the trait distribution (Gardner *et al.*, 2007). The model predicts that the production rate will increase and that the death rate of an infected cell will decrease until the trade-off curve is reached (the blue curve, dark grey in the print edition, in Fig. 2c). The simulation was stopped when the population reached the trade-off curve because the boundary then biases the mutation process, which greatly affects the genetic variances and covariances. However, the trade-off curve represents the best production rate possible for a given infected cell death rate, therefore we expect the population to evolve on the trade-off curve to reach the potential ESS. In this particular case, this is confirmed by the stochastic model (the green curve, light grey in the print edition, in Fig. 2c), which shows an increase and then a decrease in production rate. The evolutionary trajectory in the stochastic model is more variable because the simulations are initiated with a single infecting strain and viral diversity has to build up from random mutations.

These results corroborate previous results showing that the early optimum strategy is to have a high value production rate because there are many susceptible cells to infect (Ball *et al.*, 2007). As the density of susceptible cells decreases (Fig. 2b), so does the optimal production rate. Similar results have also been derived in epidemiology to study virulence evolution (Lenski & May, 1994, Day & Proulx, 2004). Finally, in this case, the trajectory of the average trait values within the cell or outside the cells are almost identical (figure not shown) because of rapid diffusion from one environment to the other (due to high production rates). Viruses with lower production rates would exhibit more heterogeneity between the two environments.

The Price equation model can also be used to compare short-term viral evolutionary responses to different types of treatments. Figure 3a shows trait dynamics in the case of treatments targeting infected cells or free-living viruses. Treatments are modelled by adding an extra death term ($-\tau_I I_T$ or $-\tau_V V_T$) into eqns 5 or 6 (see Supplementary protocol S2). Targeting infected cells selects for higher production rates. This comes from the fact that the fitness of a strain strongly depends on the within-cell life stage. Similarly to vaccines that can select for higher virulence at an epidemiological level (Gandon *et al.*, 2001; Mackinnon & Read, 2004), this framework suggests that treatments can also select for higher viral production rates. Moreover, treatments targeting infected cells will be more efficient than treatments targeting free viruses (in terms of maintaining the density of target cells, Fig. 3b, and decreasing viral load, Fig. 3c) but they also select for higher production rates if the virus survives to the treatment. A possible extension of this model could be to treat the sensitivity to the treatment as a trait of a strain. This would be useful to study the evolution of the cost of resistance to treatments, which involves trade-offs between sensitivity to drugs and parasite replication rate.

Case 3: Pathogen–lymphocyte coevolution

The immune response acts as a major selective pressure on parasite evolution. Nonspecific immunity is discussed above, but, in many organisms, parasite cells or host-infected cells can be identified by specific immune cells, which triggers the proliferation of lymphocytes that are extremely efficient at fighting the infection. Predator–prey population dynamics models inspired from ecology can be used to describe parasites facing an immune response (Nowak & May, 2000; Alizon & van Baalen, 2008). Here, I model explicitly the whole

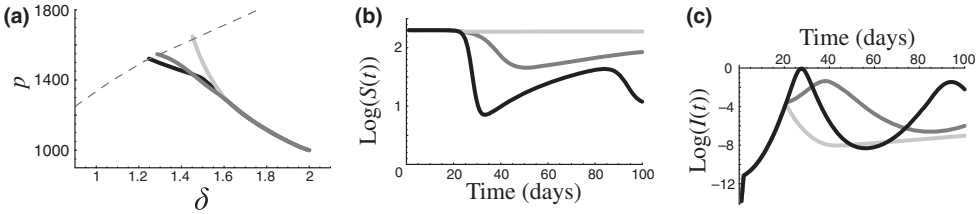

Fig. 3 Effect of anti-viral treatments on evolutionary dynamics in a case without treatment (in black, see also Fig. 2), in a case with a treatment targeting the free viral particles (in dark grey) and in a case targeting infected cells (in light grey). (a) Evolutionary dynamics of the traits, (b) density dynamics of host target cells and (c) density dynamics of free viral particles. Treatment begins after 20 days. Targeting infected cell is the most effective strategy to decrease viraemia but it also selects for higher viral production rates. Parameter values and further details on how the model with treatment is built are available in Supplementary protocol S2.

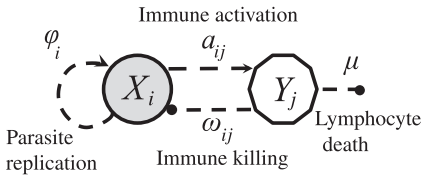

Fig. 4 Model of lymphocyte–parasite coevolution. The density of parasites of strain i is denoted X_i and the density of lymphocytes of clone j is denoted Y_j. Other notations are in Table 1.

diversity of the immune response, which introduces a coevolutionary process: parasites are selected depending on their growth rates and their antigens, whereas immune cells are selected depending on their receptor (see Fig. 4). More precisely, there are n parasite strains and m lymphocyte clones. Each lymphocyte clone can destroy a parasite strain with a given efficiency (σ_{ij}) and each parasite strain elicits the proliferation of each of the lymphocyte clones at a given rate (a_{ij}). Dynamics of parasite strains and lymphocyte densities are here given by the following equation system:

$$\frac{dX_i}{dt} = \left(\varphi_i - \sum_{j=1}^{m} \omega_{ij} Y_j\right) X_i \tag{5a}$$

$$\frac{dY_j}{dt} = \left(\sum_{i=1}^{n} a_{ij} X_i - \mu\right) Y_j \tag{5b}$$

where φ_i is the replication rate of parasites of strain i, a_{ij} is the activation rate of the lymphocytes of clone j by parasite strain i, ω_{ij} is the sensitivity of parasite strain i to lymphocytes of clone j, and μ the death rate of the lymphocytes. Lymphocytes are assumed to differ only in the way they are activated by, and in the way they destroy parasite strains. In a system with total specificity, we would have one lymphocyte clone for each parasite strain ($n = m$) and only the diagonal terms of the a and σ matrices would be nonzero. On the contrary, if cross-immunity is allowed, some parasite strains will activate more than one lymphocyte clone and some lymphocyte clones will destroy more than one parasite strain.

The equations governing variations in total parasite and lymphocyte densities (X_T and Y_T respectively) are

$$\frac{dX_T}{dt} = (\bar{\varphi} - \omega_{\bullet\bullet} Y_T) X_T \tag{6a}$$

$$\frac{dY_T}{dt} = (a_{\bullet\bullet} X_T - \mu) Y_T \tag{6b}$$

where $\bar{\varphi}$ is the average parasite replication rate, $\omega_{\bullet\bullet}$ is the average parasite sensitivity to the average lymphocyte population, $a_{\bullet\bullet}$ is the average lymphocyte proliferation rate for an average parasite population and μ is the lymphocyte death rate (here assumed to be constant among all the lymphocyte clones). In this case, both the parasite and the lymphocyte populations are diverse, which means that some traits can be expressed at four different levels: a parasite of strain i interacting with a lymphocyte of clone j (ω_{ij} and a_{ij}), a parasite of strain i interacting with an average lymphocyte of the lymphocyte population ($\omega_{i\bullet}$ and $a_{i\bullet}$), a lymphocyte of clone j interacting with an average parasite in the parasite population ($\omega_{\bullet j}$ and $a_{\bullet j}$) and, finally, the interaction between an average parasite and an average lymphocyte ($\omega_{\bullet\bullet}$ and $a_{\bullet\bullet}$).

As shown in supplementary material S1, if we assume that the mutation process is not biased, the dynamics of the parasite replication rate are given by the following equation:

$$\frac{d\bar{\varphi}}{dt} = \mathrm{Var}(\varphi_i) - Y_T \mathrm{Cov}_i(\varphi_i, \omega_{i\bullet}) \tag{7}$$

where φ_i is the replication rate of strain i and $\omega_{i\bullet}$ is the sensitivity of strain i to the average lymphocyte population. Cov_i indicates a covariance in the parasite population. As in the first case, variance in the replication rate in the parasite population selects for higher growth rates and covariance between the replication rate of a strain and its sensitivity to the average lymphocyte population ($\omega_{i\bullet}$) has the opposite effect.

For the evolutionary dynamics of other traits of interest (a and ω), deriving the equations is less straightforward. After some calculations (see Supplementary text S1), we find that

$$\frac{d\omega_{\bullet\bullet}}{dt} = X_T E_i\left[\mathrm{Cov}_j(\omega_{ij}, a_{\bullet j})\right] + \mathrm{Cov}(\omega_{i\bullet}, \varphi_i) - Y_T \mathrm{Var}(\omega_{i\bullet}) \tag{8a}$$

$$\frac{da_{\bullet\bullet}}{dt} = X_T \mathrm{Var}(a_{\bullet j}) + E_j\left[\mathrm{Cov}_i(\varphi_i, a_{ij})\right] - Y_T E_j\left[\mathrm{Cov}_i(a_{ij}, \omega_{i\bullet})\right] \tag{8b}$$

where E_i is an average over i.

These results are complex but this is not surprising since these equations encompass all the immunological changes that can occur during an infection, from parasite immune evasion to cross-reactivity of the immune response. Of course, finding all the variances and covariances involved in eqns 8a and 8b is likely to be impossible, but simplifying assumptions can be made. Also, some conclusions can already be drawn from these general equations.

To analyse eqns 8a and 8b, we need to bear in mind that some selective forces act through the parasite population and others act through the lymphocyte population. Equation 8a tells us that, as expected, variance in the parasite sensitivity to the average immune response ($\omega_{i\bullet}$) selects for lower average values of ω over all parasite strains. This evolution towards total escape from the immune response will not occur if decreasing ω decreases the replication rate of the strain

(φ_i) or if it increases immune proliferation (a). Finally the sum of covariances in eqn 8a (the first term from the right hand side) comes from the lymphocyte population: there is selection for higher values of ω if there are parasites (if $X_T > 0$) and if there is a positive correlation between the sensitivity of a lymphocyte clone j to the average parasite population ($a_{\bullet j}$) and the efficiency with which this clone kills parasites of strain i (ω_{ij}). This term actually reflects the cross-reactivity of the immune response. It contains an average over all parasite strains because a parasite strain can activate more than one lymphocyte clone and because a lymphocyte clone can target more than one parasite strain.

In eqn 8b, the first term comes from the lymphocytes. It implies that as long as the host is infected and there is variance in the lymphocyte population, there will be selection for higher activation rates. The second term comes from the evolution of the parasite. It tells us that the average activation rate increases if there is a positive correlation between lymphocyte activation rate and parasite growth rate. This can be seen as a manifestation of a coevolutionary arms race between hosts and parasites. Finally, the last term comes from the evolution of the parasite and it tells us that $a_{\bullet\bullet}$ decreases if there is a positive correlation between immune activation rate and killing rate.

With this formulation, the parasite density is followed directly, that is, the resources on which the virus feeds are not modelled explicitly. Possible extensions could be achieved by adding a term for host target cells or a more complicated growth function (e.g. logistic growth function). This would complicate the approach but not invalidate it. Nevertheless, this case with parasite–lymphocyte coevolution illustrates the breadth of the Price equation approach by showing that it can be applied to different models of intrahost dynamics.

Discussion

Many diseases evolve over the course of an infection; however, intrahost evolution is often assumed to be a discrete event. In the case of antibiotic resistance for instance, such an assumption implies that all the bacteria infecting a host suddenly become resistant. I develop an approach that combines population genetics and population dynamics to follow the dynamics of the evolutionary process. This is important for two reasons: first because intrahost evolutionary dynamics can affect the course of an infection and second because there can be complex feedbacks between selective pressures at the within- and at the between-host levels.

Evolutionary dynamics are intensively studied in epidemiology (Dieckmann *et al.*, 2002) but the specificity of intrahost dynamics, illustrated by the complexity of the immune system, calls for a specific approach. Current studies of within-host evolution are based either on the quasispecies framework or on game-theoretical frameworks. In the former, the fitness landscape is generally assumed to be constant and it is also difficult to track evolutionary trajectories of quantitative traits. In the latter, fitness values are relative and quantitative traits can be studied but it is more difficult to link results to data and short-term dynamics are often difficult to analyse. Kelly *et al.* (2003) also developed a model that links population genetics and population dynamics but they assume that within-host dynamics are at equilibrium and they only follow the accumulation of genetic diversity within the viral population.

The framework developed here is inspired from applications of the Price equation to epidemiology (Day & Proulx, 2004; Day & Gandon, 2006, 2007; Gandon & Day, 2009) and it allows one to study the intrahost evolution of quantitative traits, even if the cell population dynamics are not at equilibrium. First, this framework can be used to derive an equation that describes how (and which) trade-offs affect the evolution of the average value of a trait. Second, one can combine the equations for the traits dynamics and the equations for the cell population dynamics to predict the short-term evolutionary trajectories of traits of interest. These intrahost evolutionary dynamics cannot be predicted on the long term because, in order to make the system dynamically sufficient, we assume that genetic (co)variance in the population are constant (see Case 2 and Gardner *et al.*, 2007). As the parasite population evolves, its genetic composition will change. This will affect the genetic variances and covariances if the mutation process is biased (i.e. if the average trait of mutants differs from the average trait in the population), which is the case when the population reaches a trade-off boundary. In other words, this method cannot tell us if there is an evolutionary stable equilibrium (ESS) in the system and whether or not it can be reached. To make such long-term predictions with the Price equation framework, it would be necessary to find a way to update these covariances as the system evolves. This way, once the population has reached an ESS, the variances and covariances cancel each other so that equations driving the dynamics of the trait are all zero. The problem is that this updating is likely to be as complicated as a stochastic model that would track the population dynamics of all the parasite strains. However, this limitation of the model does not seem to be a major issue for two reasons. First, even though rates of within-host evolution can be high, it is possible that the parasite population may not have the time to reach its final state because the duration of an infection is limited (recovery or host death can occur). In this case, the short-term evolutionary dynamics are the only relevant dynamics. Second, if we are interested in the long-term outcome, we can use other frameworks, such as adaptive dynamics (Geritz *et al.*, 1997), which rely heavily on assumptions about shapes of trade-off curves (but see de Mazancourt & Dieckmann, 2004) and are more difficult to link to data than the Price equation framework.

There are several ways to extend this approach to address specific biological questions. For instance, replication rates might vary among populations of infected cells (this is the case for HIV in the gut, van Marle *et al.*, 2007). Such heterogeneity among host cells could be added to the second case developed here. In the same example, we also assumed that cells could not be infected by more than one viral strain. However, cell co-infections are known to occur and they can affect viral evolution, especially through recombination (Levy *et al.*, 2004). Finding a way to introduce co-infections in the Price equation framework is an open question, even in epidemiology. Another extension concerns the evolutionary response to treatments. The framework could be used to study the evolution of parasite resistance to treatments by introducing a quantitative trait. As shown in Day & Gandon (2007), it would also be possible to predict the rate of parasite evolution. Finally, the mutation process is here assumed to be unbiased. Experimental data on such biases could be introduced in this framework to better understand the effect of mutation on the course of an infection.

Existing data on intrahost evolution mostly concerns qualitative traits, but there are some data on the replication rate of viruses such as HIV (Dykes & Demeter, 2007; Tebit *et al.*, 2007). Current experimental studies tend not to follow changes in average trait values during the course of an infection. There is a practical reason for this: many rapidly mutating diseases (such as HIV or hepatitis C) are detected after the initial acute stage, which means we often lack part of the dynamics. Moreover, strain diversity can vary a lot depending on which organ the sample is taken from Frost *et al.*, (2001), which is likely to also affect the average trait value. However, in the case of HIV for instance, some approaches have been developed that allow one to estimate the production rate of a mixture of strains isolated from a patient *in vitro* (Dykes & Demeter, 2007). Studying samples collected over the whole duration of an infection would give us information on the evolution of the average value of a trait during the course of an infection. A way to test the predictions made by the model is to test qualitative predictions. For instance, the result that different types of treatments trigger different evolutionary responses does not require knowledge of the whole parasite diversity in the host but only the average trait values before and after the treatment. Overall, this approach calls for the collection of multiple samples from hosts at different stages of their infection, which could be done for diseases as closely monitored as HIV or hepatitis C.

Finally, this approach has implications for 'nested' models, which link within-host dynamics to epidemiological dynamics (Mideo *et al.*, 2008). These models are especially useful for tracking parasite evolution at the epidemiological level, but, currently, they ignore intrahost evolution. In the case of rapidly evolving diseases such as HIV, combining intrahost evolution and trade-offs at the epidemiological level (Fraser *et al.*, 2007) could prove to be essential for understanding the biology of the disease.

Acknowledgments

I thank S. Bonhoeffer, T. Day, A. Hurford, S. Lion and N. Mideo for suggestions and comments. I am funded by an ETH fellowship.

References

Alfonso, V., Mbayed, V.A., Sookoian, S. & Campos, R.H. 2005. Intra-host evolutionary dynamics of hepatitis C virus E2 in treated patients. *J. Gen. Virol.* **86**: 2781–2786.

Alizon, S. 2008. Transmission-recovery trade-offs to study parasite evolution. *Am. Nat.* **172**: E113–E121.

Alizon, S. & van Baalen, M. 2008. Acute or chronic? Within-host models with immune dynamics, infection outcome and parasite evolution. *Am. Nat.* **172**: E244–E256.

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

André, J.B. & Godelle, B. 2006. Within-host evolution and virulence in microparasites. *J. Theor. Biol.* **241**: 402–409.

Ball, C.L., Gilchrist, M.A. & Coombs, D. 2007. Modeling within-host evolution of HIV: mutation, competition and strain replacement. *Bull. Math. Biol.* **69**: 2361–2385.

Bocharov, G., Ludewig, B., Bertoletti, A., Klenerman, P., Junt, T., Krebs, P., Luzyanina, T., Fraser, C. & Anderson, R.M. 2004. Underwhelming the immune response: effect of slow virus growth on $CD8^+$-T-lymphocyte responses. *J. Virol.* **78**: 2247–2254.

Bonhoeffer, S. & Nowak, M.A. 1994. Intra-host versus inter-host selection: viral strategies of immune function impairment. *Proc. Natl. Acad. Sci. USA* **91**: 8062–8066.

Brown, S.P., Hochberg, M.E. & Grenfell, B.T. 2002. Does multiple infection select for raised virulence. *Trends Microbiol.* **10**: 401–405.

Chao, L., Hanley, K.A., Burch, C.L., Dahlberg, C. & Turner, P.E. 2000. Kin selection and parasite evolution: higher and lower virulence with hard and soft selection. *Q. Rev. Biol.* **75**: 261–275.

Christiansen, F.B. 1991. On conditions for evolutionary stability for a continuously varying character. *Am. Nat.* **138**: 37–50.

Day, T. & Gandon, S. 2006. Insights from Price's equation into evolutionary epidemiology. In: *Disease Evolution: Models, Concepts, and Data Analyses* (Z. Feng, U. Dieckmann & S.A. Levin, eds), pp. 23–44. Vol. 71 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society.

Day, T. & Gandon, S. 2007. Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* **10**: 876–888.

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

de Mazancourt, C. & Dieckmann, U. 2004. Trade-off geometries and frequency-dependent selection. *Am. Nat.* **164**: 765–778.

Dieckmann, U., Metz, J.A.J., Sabelis, M.W. & Sigmund, K. (eds) 2002. *Adaptive Dynamics of Infectious Diseases. In Pursuit of Virulence Management*. Cambridge Studies in Adaptive Dynamics. Cambridge University Press, Cambridge, UK.

Dykes, C. & Demeter, L.M. 2007. Clinical significance of human immunodeficiency virus type 1 replication fitness. *Clin. Microbiol. Rev.* **20**: 550–578.
Eigen, M., McCaskill, J. & Schuster, P. 1988. Molecular quasispecies. *J. Phys. Chem.* **92**: 6881–6891.
Frank, S.A. 1996. Models of parasite virulence. *Q. Rev. Biol.* **71**: 37–78.
Fraser, C., Hollingsworth, T.D., Chapman, R., de Wolf, F. & Hanage, W.P. 2007. Variation in HIV-1 set-point viral load: epidemiological analysis and an evolutionary hypothesis. *Proc. Natl. Acad. Sci. USA* **104**: 17441–17446.
Frost, S.D.W., Dumaurier, M.J., Wain-Hobson, S. & Brown, A.J.L. 2001. Genetic drift and within-host metapopulation dynamics of HIV-1 infection. *Proc. Natl. Acad. Sci. USA* **98**: 6975–6980.
Gandon, S. & Day, T. 2009. Evolutionary epidemiology and the dynamics of adaptation. *Evolution*, in press. doi: 10.1111/j.1558-5646.2009.00609.x.
Gandon, S., Mackinnon, M.J., Nee, S. & Read, A.F. 2001. Imperfect vaccines and the evolution of pathogen virulence. *Nature* **414**: 751–756.
Gardner, A., West, S.A. & Barton, N.H. 2007. The relation between multilocus population genetics and social evolution theory. *Am. Nat.* **169**: 207–226.
Geritz, S.A.H., Metz, J.A.J., Kisdi, E. & Meszéna, G. 1997. Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett.* **78**: 2024–2027.
Gilchrist, M.A., Coombs, D. & Perelson, A.S. 2004. Optimizing within-host viral fitness: infected cell lifespan and virion production rate. *J. Theor. Biol.* **229**: 281–288.
Goulder, P.J.R. & Watkins, D.I. 2004. HIV and SIV CTL escape: implications for vaccine design. *Nat. Rev. Immunol.* **4**: 630–640.
Hammerstein, P. 1996. Darwinian adaptation, population genetics and the streetcar theory of evolution. *J. Math. Biol.* **34**: 511–532.
Hancock, R.E. & Diamond, G. 2000. The role of cationic antimicrobial peptides in innate host defences. *Trends Microbiol.* **8**: 402–410.
Holmes, E.C., Zhang, L.Q., Simmonds, P., Ludlam, C.A. & Brown, A.J. 1992. Convergent and divergent sequence evolution in the surface envelope glycoprotein of human immunodeficiency virus type 1 within a single infected patient. *Proc. Natl. Acad. Sci. USA* **89**: 4835–4839.
Iwasa, Y., Michor, F. & Nowak, M.A. 2005. Virus evolution within patients increases pathogenicity. *J. Theor. Biol.* **232**: 17–26.
Kamp, C. 2003. A quasispecies approach to viral evolution in the context of an adaptive immune system. *Microbes Infect.* **5**: 1397–1405.
Kelly, J.K., Williamson, S., Orive, M.E., Smith, M.S. & Holt, R.D. 2003. Linking dynamical and population genetic models of persistent viral infection. *Am. Nat.* **162**: 14–28.
King, T., Ishihama, A., Kori, A. & Ferenci, T. 2004. A regulatory trade-off as a source of strain variation in the species *Escherichia coli*. *J. Bacteriol.* **186**: 5614–5620.
Lemey, P., Rambaut, A. & Pybus, O.G. 2006. HIV evolutionary dynamics within and among hosts. *AIDS Rev.* **8**: 125–140.
Lenski, R.E. & May, R.M. 1994. The evolution of virulence in parasites and pathogens - reconciliation between two competing hypotheses. *J. Theor. Biol.* **169**: 253–265.
Levin, B.R. & Bull, J.J. 1994. Short-sighted evolution and the virulence of pathogenic microorganisms. *Trends Microbiol.* **2**: 76–81.
Levins, R. 1962. Theory of fitness in a heterogeneous environment. I. The fitness set and adaptive function. *Am. Nat.* **96**: 361–373.
Levy, D.N., Aldrovandi, G.M., Kutsch, O. & Shaw, G.M. 2004. Dynamics of HIV-1 recombination in its natural target cells. *Proc. Natl. Acad. Sci. USA* **101**: 4204–4209.
Mackinnon, M.J. & Read, A.F. 2004. Immunity promotes virulence evolution in a malaria model. *PLoS. Biol.* **2**: 1286–1292.
Maynard-Smith, J. & Price, G.R. 1973. The logic of animal conflict. *Nature* **246**: 15–18.
Merlo, L.M.F., Pepper, J.W., Reid, B.J. & Maley, C.C. 2006. Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* **6**: 924–935.
Mideo, N., Alizon, S. & Day, T. 2008. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol. Evol.* **23**: 511–517.
Nowak, M.A. & May, R.M. 2000. *Virus Dynamics: Mathematical Principles of Immunology and Virology*. Oxford University Press, Oxford, UK.
Nowak, M.A., May, R.M. & Anderson, R.M. 1990. The evolutionary dynamics of HIV-1 quasispecies and the development of immunodeficiency disease. *AIDS* **4**: 1095–1103.
Palumbi, S.R. 2001. Humans as the world's greatest evolutionary force. *Science* **293**: 1786–1790.
Perelson, A.S. 2002. Modelling viral and immune system dynamics. *Nat. Rev. Immunol.* **2**: 28–36.
Price, G.R. 1970. Selection and covariance. *Nature* **227**: 520–521.
Price, P.W. 1977. General concepts on the evolutionary biology of parasites. *Evolution* **31**: 405–420.
Shriner, D., Liu, Y., Nickle, D.C. & Mullins, J.I. 2006. Evolution of intrahost HIV-1 genetic diversity during chronic infection. *Evolution* **60**: 1165–1176.
Suerbaum, S. & Josenhans, C. 2007. Helicobacter pylori evolution and phenotypic diversification in a changing host. *Nat. Rev. Microbiol.* **5**: 441–452.
Tebit, D.M, Nankya, I., Arts, E.J. & Gao, Y. 2007. HIV diversity, recombination and disease progression: how does fitness ''fit'' into the puzzle? *AIDS Rev.* **9**: 75–87.
van Marle, G., Gill, M.J., Kolodka, D., McManus, L., Grant, T. & Church, D.L. 2007. Compartmentalization of the gut viral reservoir in HIV-1 infected patients. *Retrovirology* **4**: 87.
Wilke, C.O. 2005. Quasispecies theory in the context of population genetics. BMC Evol. *Biol.* **5**: 44.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Supplementary text S1 Calculations involving the Price equation.

Supplementary protocol S2 Simulating intrahost evolutionary dynamics.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Received 9 January 2009; revised 9 February 2009; accepted 10 February 2009