Within-host viral evolution in a heterogeneous environment: insights into the HIV co-receptor switch

S. ALIZON* & B. BOLDIN†[1]

*Institut für Integrative Biologie, ETH, Universitätstrasse 16, Zürich, Switzerland
†Department of Mathematics and Statistics, University of Helsinki, Helsinki, Finland

Abstract

A virus infecting a host faces a heterogeneous and a spatially structured environment. Using a mathematical model that incorporates two types of target cells and spatial structuring, we investigate conditions for viral within-host diversification. We show that branching occurs for a wide range of parameters but that it always requires some spatial structure. Applying our model to the case of HIV, we show that it captures three main properties of the 'co-receptor switch' observed in many HIV infections: the initial dominance of virus strains that infect CCR5$^+$ cells, the late switch in some (but, importantly, not all) HIV infections and the associated drop in the number of uninfected T-cells. This suggests that the co-receptor switch could result from gradual adaptation of the virus population to target cell heterogeneity. More generally, we argue that evolutionary ecology can help us better understand the course of some infections.

Introduction

From the point of view of a pathogen, a host is a structured and a heterogeneous environment. In the case of HIV, for instance, the existence of spatial structure is supported by the fact that the virus is found in different tissues (e.g. lymph nodes, gut, brain, spleen), and environmental heterogeneity originates from the pathogen being able to exploit different types of immune cells (natural killer T-cells, macrophages, dendritic cells; Stebbing *et al.*, 2004). Different cells respond differently to infection, which is likely to affect viral evolution (Turner & Elena, 2000) as well as the course of an infection.

Habitat choice and specialization is a classical problem in evolutionary ecology (for a recent review, see Ravigné *et al.*, 2009). Here, we import this question to the within-host level by investigating the role of cell heterogeneity and spatial structure in the evolution of a viral pheno-

typic trait. In particular, we wish to determine the conditions that lead to within-host diversification in the virus population. The originality of our study lies in the type of ecological dynamics (viruses infecting cells) and in the simultaneous use of ecological and evolutionary theory to address questions of clinical importance.

Several theoretical studies have analysed the population dynamics of viruses in a structured and/or heterogeneous environment. Cuevas *et al.* (2003) develop a metapopulation model, which they test *in vitro* using vesicular stomatitis virus. They obtain environmental heterogeneity using three different types of host cells and show that high (resp. low) migration between patches leads to generalist (resp. specialist) viral strategies. They also find that the fitness of a virus is negatively correlated to the distance between the new tissue and the tissue the virus was adapted to. Kelly *et al.* (2003) develop a model that links population genetics and population dynamics to study viral rates of evolution in which they include multiple cell types. They conclude that target cell diversity favours the establishment of a persistent infection and can increase the amount of genetic diversity within the viral population. Funk *et al.* (2005) develop a target-cell limited model of viral infection in which they introduce spatial structure by assuming that each cell or

free virus has a location on a grid. They find that spatial structure can alter population dynamics by affecting the peak of the infection. They also find that viral clearance is more difficult to achieve when they assume environmental heterogeneity. Orive *et al.* (2005) analyse viral within-host dynamics in a source-sink perspective. They show that increased migration between within-host compartments makes the establishment of an infection more difficult. Holt & Barfield (2006) consider the importance of transient dynamics within a host and suggest that within-host heterogeneity can increase the importance of such dynamics. Some studies have focused on within-host heterogeneity in terms of drug concentration, especially in the case of HIV. The results show that compartmentalization increases the chances of drug resistance to evolve (Kepler & Perelson, 1998). This can also lead to different viral loads in the compartments, which has clinical implications (Ward *et al.*, 2009). Finally, as we discuss in more detail below some studies have focused on the role of target cell heterogeneity in the HIV co-receptor switch (Regoes & Bonhoeffer, 2002, 2005; Ribeiro *et al.*, 2006; Weinberger *et al.*, 2009).

We use an evolutionary dynamics framework, thus combining population dynamics and evolution of a viral phenotypic trait. Our mathematical model incorporates two types of target cells as well as spatial structure. As in previous models (Gilchrist *et al.*, 2004; Ball *et al.*, 2007; Boldin & Diekmann, 2008; Luciani & Alizon, 2009), we assume that each viral strain is defined by a continuous phenotypic trait, the 'cell exploitation strategy', which equals the rate of production of new viruses by an infected target cell of one of the cell types (see Fig. 1 and the Model). Our framework is general and can be applied to several host–pathogen systems. We will mostly refer to HIV because a lot of data is available and also because our

model captures some aspects of the co-receptor switch observed in many HIV infections. However, as we discuss later, cell heterogeneity and spatial structure are not the appanage of HIV infections but are observed also in other viral infections.

The co-receptor switch

HIV infects host immune cells and uses their machinery to replicate. To enter these cells, the virus first binds to CD4 receptors on the surface of the cell and then to co-receptors (Berger *et al.*, 1999). Early in an infection, HIV mainly infects host cells expressing the CCR5 chemokine co-receptor (i.e. CCR5[+] cells), whereas later in the infection, virus strains emerge that can infect cells expressing the CXCR4 co-receptor (i.e. CXCR4[+] cells). This change in target cell type preference is known as the co-receptor switch (Berger *et al.*, 1999; Moore *et al.*, 2004) and is linked with the coexistence of two 'variants' of viruses (R5 viruses and X4 viruses infecting CCR5[+] and CXCR4[+] cells, respectively). The switch occurs in more than 50% of the cases and has important clinical implications because it correlates with an acceleration in CD4[+] T-cell depletion and a progression to AIDS (Connor *et al.*, 1997). Note for the sake of completeness that some variants (R5X4) can use both the CCR5 and the CXCR4 co-receptors to enter a cell.

The co-receptor switch is puzzling in itself, because experimental studies show that only two or three changes in amino acids in one of the proteins of HIV are sufficient to convert an R5 virus into an X4 virus (Fouchier *et al.*, 1992). Given the high mutation rates of HIV (Coffin, 1995), why are not both variants detected after a couple of generations of cell infections? A plausible explanation for the late detection of X4 viruses is a change in selective pressures during the course of the infection so that strains that were counter-selected early in the infection are favoured later on.

This dynamical aspect of the co-receptor switch calls for mathematical modelling approaches (reviewed by Regoes & Bonhoeffer, 2005). Some models explain the late emergence of the X4 variant only through population dynamics of immune cells (i.e. without any evolution of the virus). Ribeiro *et al.* (2006) build a model where three HIV variants (R5, R5X4 and X4) with given trait values compete for target cells. They assume that CCR5[+] T-cells (which are 'memory' or 'activated' cells) result partly from the maturation of CXCR4[+] T-cells ('naive' or 'resting' cells) and show that the switch occurs at low CD4 T-cell counts because memory/activated cells (which they assume to divide 10 times more rapidly than naive/resting cells) become more rare than naive/resting cells. Although Ribeiro *et al.* do observe the switch for a particular choice of density-dependent functions and parameter values, it is not at all clear how sensitive the dynamics in their model is to changes in parameter values. Further analysis by Weinberger *et al.* (2009)
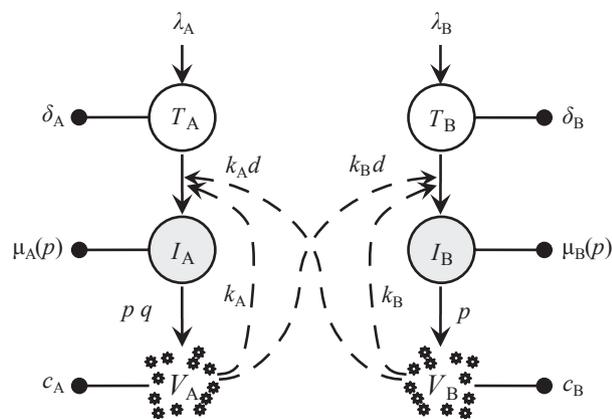
Fig. 1 The population dynamics of cells and free viruses. Target cells of one type (T_A or T_B, i.e. CCR5[+] or CXCR4[+] cells resp.) can become infected (I_A and I_B) by a free virus. We distinguish between free viruses produced by each type of cells (V_A and V_B). A virus strain is identified by its phenotypic trait (here the cell exploitation strategy, p). d reflects the spatial segregation between target cell types.

includes drug treatment and shows that the switch still occurs if some of the simplifying assumptions made by Ribeiro *et al.* are raised.

Other approaches assume that X4 viruses are recognized more easily by the immune system (Wodarz & Nowak, 1998). In this case, the switch comes from the erosion of the immune response over the course of the infection. Note that a common feature of all these models is that they assume beforehand that there exist two or three HIV variants with given traits [but see Regoes & Bonhoeffer (2002) for a study where the co-receptor preference of a virus is modelled as a continuous trait].

Finally, in their review, Regoes & Bonhoeffer (2005) propose a model with viral evolution. They assume that four mutations are required to go from an R5 to an X4 virus. They further assume that X4 viruses have a higher fitness than R5 viruses, which themselves have a higher fitness than all the intermediate mutants. Depending on the magnitude of the fitness costs, they observe early or late emergence of X4 viruses.

The main difference between our model and previous approaches is that it considers population dynamics and virus trait evolution simultaneously. Furthermore, it includes both cell type heterogeneity and spatial structure. We show how the co-receptor switch could originate from viral evolution in a heterogeneous environment within an infected host. Indeed, our model can account for three major aspects of the co-receptor switch: (i) the initial dominance of R5 viruses, (ii) the late occurrence of the switch in some (but not all) HIV-infected hosts and (iii) the associated drop and continued decline in the number of uninfected T-cells.

The model

Our model is inspired by that of Ball *et al.* (2007), who analyse the evolution of a viral trait (the production rate) in a context where viruses compete for one type of target cells in the host (for a general review, see Perelson, 2002). Here, as Regoes & Bonhoeffer (2002), we introduce a second type of target cells the viruses can infect (we label the two cell types as A and B). In addition to having multiple cell types, we also want to allow for a degree of spatial structuring. The problem is that this can rapidly lead to an intractable system (including the densities of susceptible and infected target cells of each type in two environments already requires eight variables). To limit the dimensionality of the system, we assume that each type of the cell is found in its own environment (Fig. 1). However, to counter this strong separation, we allow for viruses produced by cells in one of the environments to infect cells in the other environment (at a rate d times the rate of local infections). In the case where there is no spatial structure ($d = 1$), both types of cells are found in the same well-mixed

environment. With maximal spatial structure ($d = 0$), viruses can only infect cells of the same type that produced them, which implies a strict separation between the two environments. Finally, intermediate values ($0 < d < 1$) correspond to the case where free viruses are able to move to another environment and infect target cells there. To draw an analogy with ecology, d relates to the probability of surviving a dispersal event.

The population dynamics is governed by the following set of ODEs:

$$\frac{dT_A}{dt} = \lambda_A - \delta_A T_A - k_A V_A T_A - d k_A V_B T_A \quad (1a)$$

$$\frac{dT_B}{dt} = \lambda_B - \delta_B T_B - d k_B V_A T_B - k_B V_B T_B \quad (1b)$$

$$\frac{dI_A}{dt} = k_A V_A T_A + d k_A V_B T_A - \mu_A(p) I_A \quad (1c)$$

$$\frac{dI_B}{dt} = d k_B V_A T_B + k_B V_B T_B - \mu_B(p) I_B \quad (1d)$$

$$\frac{dV_A}{dt} = p q I_A - k_A V_A T_A - d k_B V_A T_B - c_A V_A \quad (1e)$$

$$\frac{dV_B}{dt} = p I_B - k_B V_B T_B - d k_A V_B T_A - c_B V_B. \quad (1f)$$

These equations describe the following processes. Uninfected target cells of type A and B (whose densities are denoted by T_A and T_B) die at constant per capita rates δ_A and δ_B and are produced at rates λ_A and λ_B, respectively. Free viruses are produced by infected target cells (I_A and I_B). The reproduction rate of the virus depends on the cell the virus is found in as well as the ability of a virus to exploit the cell. Here, p is the number of viral particles produced per unit of time by an infected cell of type B, and $p \cdot q$ is the same quantity in an infected cell of type A (the factor q denotes the cell-dependent component). These differences between cells are consistent with experimental data (e.g. for HIV, see Ariën *et al.*, 2006). In the following, p is the virus trait that varies among strains and is also referred to as the 'cell exploitation strategy'.

Cell infection leads to an increased cell's death rate (μ_A and μ_B), which we assume to depend on the virus strategy. The infection rate depends on the ability of the virus strain to infect each cell type (k_A and k_B) and on the within-host spatial structure (d). Infections across environments are less likely, therefore $0 \le d \le 1$.

Lastly, free viruses (with density V_A and V_B) are cleared from the host at a rate that depends on the environment (c_A and c_B). Note that viruses produced by each type of cells can have the same trait value p but they are produced in different environments (unless $d = 1$), which is why we model two compartments of free viruses.

Spatial segregation of viral population has been showed to occur for viral infections caused by hepatitis C virus (Roque-Afonso *et al.*, 2005), Epstein–Barr virus (Sitki-Green *et al.*, 2003) and HIV. Unsurprisingly, most of the data available concerns the latter virus. Spatial structure can lead to a drift between populations of the same lymphoid tissue (Frost *et al.*, 2001) or in the brain (González-Scarano & Martín-García, 2005). This compartmentation can even affect the expression of phenotypic traits such as replication rate, which seems to vary among different parts of the gut (van Marle *et al.*, 2007). In the specific case of the co-receptor switch, CCR5$^+$ and CXCR4$^+$ T-cells (which are the cells predominantly infected by HIV) dominate in different host tissues: the former are the majority in the gut and in the brain, whereas the latter dominate in the blood and in the spleen (see the Discussion). The possibility for spatial segregation between target cells is included in earlier models of viral evolution (e.g. Kepler & Perelson, 1998; Funk *et al.*, 2005; Orive *et al.*, 2005; Holt & Barfield, 2006) but not in models for co-receptor switch (Regoes & Bonhoeffer, 2002; Ribeiro *et al.*, 2006; Weinberger *et al.*, 2009).

As most studies of evolutionary dynamics, ours relies on a trade-off assumption. Following previous studies (Gilchrist *et al.*, 2004; Ball *et al.*, 2007; Boldin & Diekmann, 2008), we assume a trade-off between viral cell exploitation strategy p and infected cell's death rate μ such that viruses with intense exploitation strategies infect cells for a shorter time. This trade-off can be interpreted in terms of resource exploitation: by taking more of the resources from the cell, the virus accelerates cell's death. This trade-off can also be interpreted by the fact that cells infected with rapidly replicating viruses tend to be more easily killed by the immune system than cells infected by slow-replicating viruses (Luciani & Alizon, 2009). Several functional trade-off relationships $\mu(p)$ can capture the physiological cost we want to model. A mechanistic derivation of the trade-off function from underlying cellular processes would be possible, but is outside the scope of this study. We thus use exponential trade-off functions assumed by earlier studies (Gilchrist *et al.*, 2004; Ball *et al.*, 2007):

$$\mu_i(p) = \delta_i e^{\phi_i p}, \tag{2}$$

where i denotes the cell type (i.e. either A or B) and ϕ_i represents a scaling factor, reflecting the sensitivity of infected cells of type i to virus production. The cell exploitation strategy p is thus the trait that is subject to evolution. Table 1 summarizes all the parameters used in the model.

We emphasize that other trade-offs involving different life-history traits could also be considered. A recent study by Caraco & Wang (2008) on free-living pathogens (such as viruses) can help to understand how the choice of variables involved in a trade-off affects the optimal value of traits such as virulence, persistence or spore production rate. In our case, however, the optimal trait value is

Table 1 List of the variables and parameters of the model along with their default values. Cells of type A refer to CCR5$^+$ T-cells (i.e. activated/memory T-cells) and cells of type B refer to CXCR4$^+$ T-cells (i.e. resting/naive T-cells).

Notation	Description	Value
T_i	Number of target cells of type i (i.e. A or B)	Variable (μL^{-1})
l_i	Number of infected cells of type i	Variable (μL^{-1})
V_i	Number of viruses produced by cells of type i	Variable (μL^{-1})
p	Production rate of new viruses by an infected cell of type B (i.e. the cell exploitation strategy of a virus strain)	Variable (cell^{-1} day^{-1})
d	Spatial segregation between cell types	0.1
λ_A	Base-line production rate of target cells of type A	0.1 μL^{-1} day^{-1}
λ_B	Base-line production rate of target cells of type B	0.09 μL^{-1} day^{-1}
δ_A	Death rate of uninfected cells of type A	0.01 day^{-1}*
δ_B	Death rate of uninfected cells of type B	0.012 day^{-1}
μ_i	Death rate of infected cells of type i	See eqn 2
q	Scaling factor for the productivity of infected cells of type A	0.3
k	Virus infectivity on target cells of type A or B	0.00065 cell^{-1} day^{-1}*
c	Clearance rate of free viruses produced by cells of type A or B	3 day^{-1}*
ϕ_A	Sensitivity of infected cells of type A to virus production	0.003
ϕ_B	Sensitivity of infected cells of type B to virus production	0.01

*These default values are from Ball *et al.* (2007) and Stafford *et al.* (2000).

not the output of interest, and we focus more on population dynamics and evolutionary stability. In this case, the effect of the variables involved in the trade-off is less clear. What is known to have a large effect though is the shape of the trade-off function, and these effects can be studied using the critical function analysis (de Mazancourt & Dieckmann, 2004; Boldin *et al.*, 2009).

One of the simplifying assumptions we make in the main model is that only one virus trait evolves. In an extension of the model described in Supporting Information Appendix S1A, we include another virus trait (which we introduced as the viral variant; i.e. R5, X4 or R5X4) and let it co-evolve with viral cell exploitation strategy.

As Ball *et al.* (2007), we calibrate parameter values from HIV experimental data analysed by Stafford *et al.* (2000). However, these values are estimates only. The advantage of complementing the simulations with an analytical approach is that it allows us to perform a

bifurcation analysis, which consists of varying one (or several) parameters to identify critical parameter values that lead to a change in the evolutionary equilibrium of the system. As we show in bifurcation diagrams, the dynamics observed in simulations using a specific set of parameter values is robust to small parameter changes. In other words, our conclusions do not rely on the exactness of parameter values estimated by Stafford *et al.* (2000).

Adaptive dynamics approach

We use the adaptive dynamics (AD) approach (Geritz *et al.*, 1998) to investigate the evolutionary dynamics of viral cell exploitation strategy (p). Using the classical AD assumptions, we compute the basic reproduction number, $R_0(p_r, p_m)$, of a mutant virus strain (whose trait value is denoted p_m) that is introduced into the within-host milieu set by the resident strain (whose trait value is denoted p_r). Because mutant viruses can be introduced in any of the two cell environments (A or B) and because the 'offspring' of the mutant can be produced in either A or B (except when $d = 0$), we identify four reproduction numbers (R_{AA}, R_{AB}, R_{BA} and R_{BB}, fully described in the Supporting Information Appendix S1B). Classical invasion analysis (Diekmann & Heesterbeek, 2000) predicts that the fitness of the mutant strain, $R_0(p_r, p_m)$, is given by the largest eigenvalue of the matrix

$$\begin{pmatrix} R_{AA} & R_{AB} \\ R_{BA} & R_{BB} \end{pmatrix}.$$

The mutant can invade successfully if and only if $R_0(p_r, p_m) > 1$. If the invasion is successful, the mutant replaces the resident and becomes the new resident. Trait substitutions bring the value of p into the vicinity of an evolutionarily singular strategy (denoted by p^*), which is defined as the point where the selection gradient vanishes. Using the criteria based on second derivatives of $R_0(p_r, p_m)$ (Geritz *et al.*, 1998), we determine evolutionary stability of singular points, i.e. whether they are evolutionary stable strategies (ESS) or branching points.

Here, the expression for $R_0(p_r, p_m)$ is complicated, and we therefore identify the singular strategies and their evolutionary stability by means of a pairwise invasibility plot (PIP) and numerical computations – see Appendix S1B for further details. The explicit expression for $R_0(p_r, p_m)$ and numerical calculations of singular strategies furthermore allow us to plot bifurcation diagrams in which we investigate the occurrence of branching in various parameter regions.

In Appendix S1B, we also show that in the absence of spatial structure (if $d = 1$), branching can never occur. In fact, the case $d = 1$ yields an optimization model similar to that analysed by Ball *et al.* (2007), because in that case, the cell environment is essentially one-dimensional. The case where $\mu_A = \mu_B$ (i.e. when infection has the same detrimental effect on both types of cells) also yields an optimization model. In both of these cases, therefore, a single generalist strain is the evolutionary winner. Lastly, in the case when $d = 0$, the two cell environments are completely isolated. We then observe two separate evolutionary processes, and the evolutionarily stable strategy in each of the environments (A or B) is a trait that maximizes the fitness (or 'burst size') in its own environment.

Numerical simulations

Our second approach is based on hybrid deterministic-stochastic numerical simulations, which are derived from the framework developed by Luciani & Alizon (2009). In addition to evolutionary dynamics, simulations allow us to follow cell population dynamics.

The model we use for simulations is based on equation system 1, but we now allow for many viral strains (each with a different trait value of p) to coexist at a given time. The population dynamics (i.e. infection, death and births of cells and viruses) is calculated with a Runge–Kutta algorithm. Every time step, a virus strain present in the system can give birth to a mutant strain. This new strain is defined by its trait, which is drawn from a Gaussian distribution centred in the trait of the resident and of variance σ. The rate at which a mutation event occurs depends on the product of the mutation rate of the viral trait and of the number of new free viruses produced. Further details about the simulations are available in the Appendix S1B.

Performing simulations requires to set parameter values for each type of cell. We assume that the two types of cells, A and B, correspond to CCR5[+] (i.e. activated/memory) T-cells and CXCR4[+] (resting/naive) T-cells, respectively (Bleul *et al.*, 1997). These two types of cells have different population sizes (Moore *et al.*, 2004), which is consistent with a greater input ($\lambda_A > \lambda_B$) and a greater life-span of CCR5[+] T-cells ($\delta_B > \delta_A$). In the case of HIV, it is thought that T-cells infected by R5 viruses suffer less from an infection, i.e. that $\mu_A < \mu_B$, because the immune response is more efficient at detecting X4 viruses (see Moore *et al.*, 2004; Regoes & Bonhoeffer, 2005, for reviews). This can be included in the model by assuming that $\phi_A < \phi_B$. Finally, the number of viruses released by an infected cell before it dies, also known as the burst size, is thought to be higher in CCR5[+] than in CXCR4[+] T-cells in the case of HIV (Eckstein *et al.*, 2001). Mathematically, this can be written as

$$\frac{pq}{\mu_A(p)} > \frac{p}{\mu_B(p)}. \tag{3}$$

This means that viruses with the same strategy will always produce more viruses out of a cell of type A than out of a cell of type B. If $q \geq 1$, inequality 3 is always fulfilled because $\mu_A < \mu_B$. When $q < 1$, we restrict the values of p to the regions where the inequality holds (see Results).

Results

Using the adaptive dynamics approach, we derive a pairwise invasibility plot (or PIP, Fig. 2a) that reflects the possibility for a mutant viral strain (with trait p_m) to
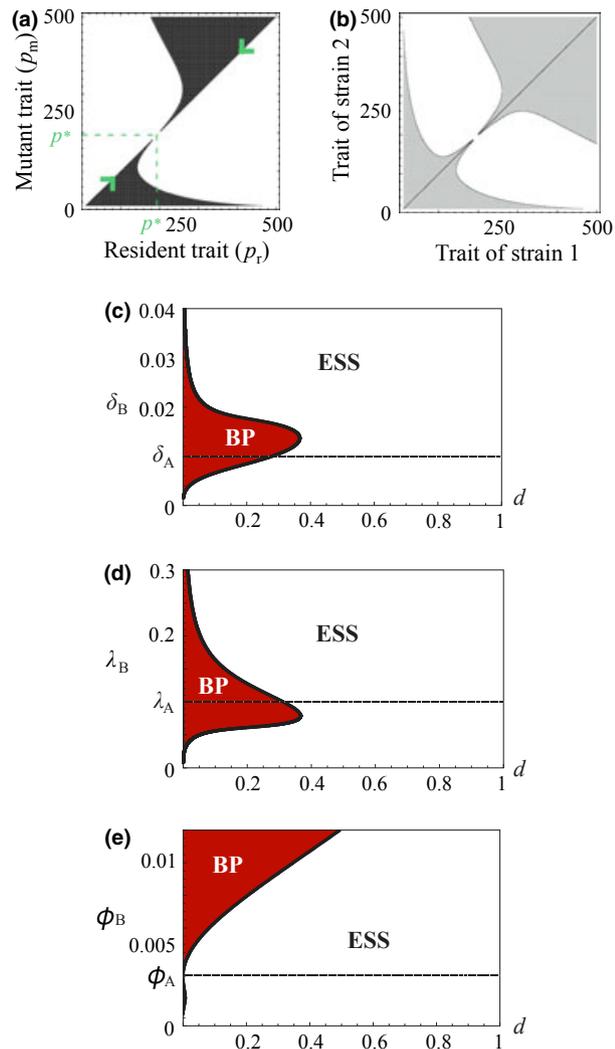
Fig. 2 Adaptive dynamics approach. (a) PIP showing that there is one evolutionarily singular strategy, which is convergent stable and evolutionarily unstable. In areas in black (resp. in white), the mutant has a negative (resp. positive) invasion fitness. Green arrows (light grey in print publication) indicate the direction of trait evolution, and the green/light grey dotted line indicates the evolutionary singular strategy (p^*). (b) Invasion plot showing that there is mutual invasibility (in white) near the singularity, which indicates a branching point; (c) Effect of the value of d and δ_B on the evolutionary stability of the singularity; (d) Effect of the value of d and λ_B on the evolutionary stability of the singularity and (e) Effect of the value of d and ϕ_B on the evolutionary stability of the singularity. Horizontal dashed lines show the default parameter value in the other cell type. BP stands for 'branching point', and ESS stands for 'evolutionarily stable strategy'. Parameter values are default unless stated otherwise.

invade a resident strain (with trait p_r). For chosen parameter values, we observe a unique singular strategy ($p^* = 192.2$), which is convergence stable. This strategy is evolutionarily unstable, because it can be invaded by both mutant with a higher and with a lower trait value. In Fig. 2b, white areas indicate regions of mutual invadability, and grey areas indicate competitive exclusion of one of the strains. This confirms that p^* is a branching point because not only can mutants with smaller or greater trait values invade p^* (Fig. 2a), but they can coexist after invasion (Fig. 2b).

To investigate the robustness of the evolutionary outcome to changes in parameter values, we perform a two-dimensional bifurcation analysis. We study the robustness of branching to parameter changes by varying d and another parameter (λ_B, δ_B or ϕ_B) whereas keeping the other parameters constant. Figure 2c–e show the parameter values that lead to a branching point (in red/dark grey in print publication) or to an ESS (in white). The bifurcation plots for the other parameters (c_B and q) are shown in the Appendix S1A.

We first observe that, for there to be an evolutionary branching, there needs to be spatial structuring of the two types of cells (i.e. $d < 1$). We show in Appendix S1B that if $d = 1$, the evolutionarily optimal strategy maximizes a weighted sum of the two burst sizes. Note that if the difference in sensitivity to the infection (which determines the death rate of an infected cell) between the cells is large (i.e. $\phi_B > \phi_A$), branching can be observed even for large values of d (Fig. 2e). In other words, increased differences between cell types can compensate for the proximity between cell types in the host and allow for evolutionary branching.

A second result is that the branching is more likely to occur if $\lambda_B < \lambda_A$ or if $\delta_B > \delta_A$ (but these are not necessary conditions). As discussed earlier, this is what one would expect if type A corresponds to CCR5[+] T-cells and type B corresponds to CXCR4[+] T-cells.

Course of the infection

In the numerical simulations, we first initiate the infection with a virus that has an intense cell exploitation strategy ($p = 300$). Figure 3a shows the trait value of all the strains present in the host at a given time. Early in the infection, there is a diversification of the initial strain but the population remains 'monomorphic' (with $250 < p < 350$). After approximately 2000 days, we observe the emergence of a second cluster of viruses that diverges from the initial morph, which leads to a 'dimorphic' population. The new morph is identified by low trait values ($100 < p < 200$). In the following, we show that this evolutionary branching (Geritz et al., 1998) is associated with the adaptation of virus morphs to the two cell types.

If we first consider the population dynamics (Fig. 3b), we see that CCR5[+] cells (in red/light grey) are predominantly infected early in the infection (the density of
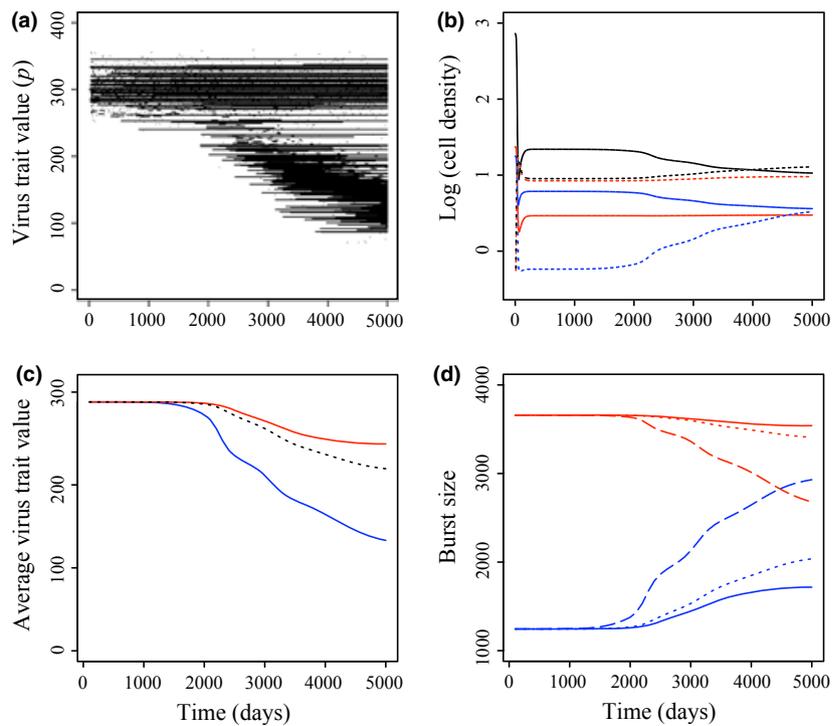
Fig. 3 Virus evolutionary dynamics for a rapidly replicating initial strain. (a) Production rate, p, of the strains present in the host at a given time; (b) Densities of uninfected (plain lines) and infected cells (dashed lines) for CCR5[+] cells (red/light grey in print publication), CXCR4[+] cells (blue/dark grey) and overall (black); (c) Average exploitation strategy in the host (in black), in CCR5[+] cells (red) and in CXCR4[+] cells (blue/dark grey); (d) Burst size in CCR5[+] cells (red/light grey) or in CXCR4[+] cells (blue/dark grey) of viruses with the host average value (dotted line), viruses found in CCR5[+] cells (plain line) and viruses found in CXCR4[+] cells (dashed line). $p = 300$ for the initial strain and parameter values are default.

target cells of type A is low, and the density of infected cells of type A is high). After the branching, the density of infected CXCR4[+] cells increases strongly (blue/dark grey doted line). Finally, the branching coincides with a strong acceleration in the decrease of the total density target cells (in black).

Figure 3c allows us to better understand the course of the infection because it shows that there is actually a segregation of the two virus morphs after the branching: viruses with a high value of p are found predominantly in CCR5[+] (in red/light grey) and viruses with a low value of p are found in CXCR4[+] cells (in blue/dark grey). The evolutionary dynamics shown in Fig. 3a can therefore be interpreted as an adaptation of the virus to the diversity in target cells. To confirm this interpretation, we show in Fig. 3d the burst size of a virus in a CXCR4[+] cell (in blue/dark grey) and in a CCR5[+] cell (in red/light grey). These values can be computed using the average trait of viruses found in CXCR4[+] cells (dashed line), CCR5[+] cells (plain line) and in the host on average (dotted lines). We find that, over time, a virus population adapts to the type of cell it infects predominantly. Note that the fitness of the average viral strain found in the host (dotted line) increases in CXCR4[+] cells (in blue/dark grey) but decreases in CCR5[+] cells (in red/light grey). This is because we initiate the infection with a strain that is better adapted to CCR5[+] cells. Finally, note that the burst size in cells of type A (in red/light grey) is greater than the burst size in cells of type B, which was one of the conditions to link the cell types A and B to CCR5[+] and CXCR4[+] cells, respectively (inequality 3).

In Fig. 4, we show that if the infection is initiated with a milder strain, which is better adapted to CXCR4[+] cells (i.e. has a low value of p, see the dashed blue/dark grey line on Fig. 3d), it takes so long for branching to occur that the host is likely to die before the branching occurs. Interestingly, we also observe that, even though the initial strain is better adapted to CXCR4[+] cells, it is CCR5[+] cells that are predominantly infected (Fig. 4b). This is attributed to the fact that, even for an ill-adapted strain, CCR5[+] cells are still a 'better' resource for the virus than CXCR4[+] cells (Fig. 4d). By a better resource, we mean that the burst size (the number of viruses produced by an infected cell before it dies) is higher.

A consequence of Figs 3 and 4 is that it is unlikely to observe a depletion of CXCR4[+] cells early in an infection. Regardless of whether the infecting strain is better adapted to CCR5[+] T-cells (high value of p) or to CXCR4[+] T-cells (low value of p), it is the CCR5[+] cells that are predominantly infected in the initial stages because they yield a larger burst size. To have CXCR4[+] cells infected in large numbers, it is required that the average viral trait value in the host comes in a vicinity of the branching point (p^*). Only when a strain with such a trait value initiates the infection, can we have early depletion of CXCR4[+] cells.

Discussion

Heterogeneity in the environment is known to be a driver of biological diversification (Levins, 1968; Schluter, 2000;
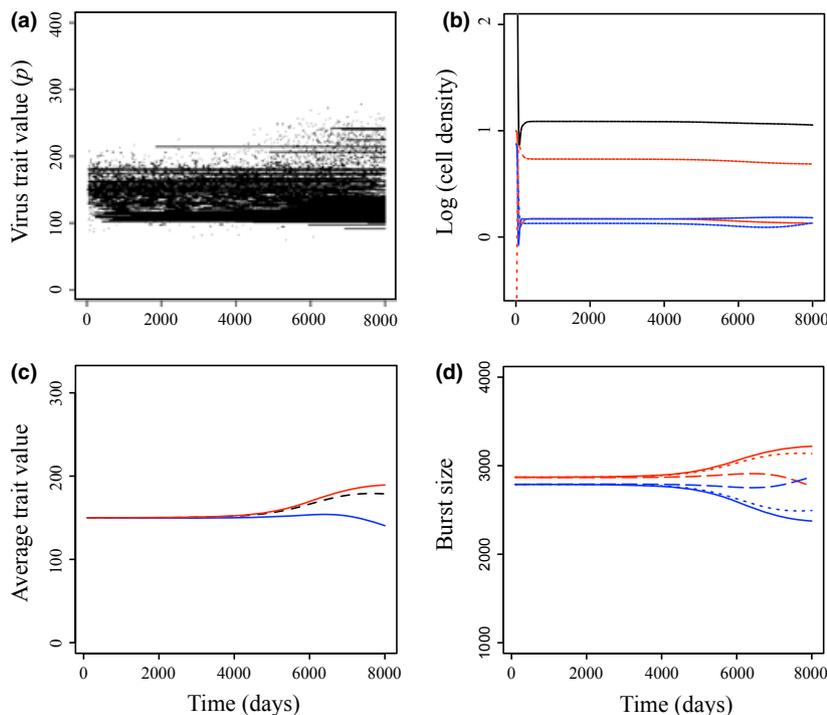
Fig. 4 Virus evolutionary dynamics for a mild initial strain. The panels code is the same as in Fig. 3. $p = 150$ for the initial strain and parameter values are default.

Ravigné *et al.*, 2009). In the case of viruses infecting cells of a host, we find that spatial structure and the presence of two cell types can lead to evolutionary branching with the separation between two viral morphs: viruses of one morph have an intense exploitation strategy and viruses of the other morph have a milder strategy. Bifurcation analysis shows that the occurrence of branching is robust to small parameter variations. Evolutionary branching only occurs if there is spatial structure (i.e. if $d < 1$). This result is in line with earlier results in evolutionary ecology (high migration rates between patches decrease the likeliness of evolutionary branching, as shown by Meszéna *et al.*, 1997) and results on virus evolution (Cuevas *et al.*, 2003). The value of d is likely to depend strongly on the clearance time of free viruses, which varies across viruses. The fact that a virus like HIV has a low survival time in the blood [approximately 3 days (Stafford *et al.*, 2000)] could lead to a higher degree of spatial structuring in the host (i.e. lower d) than in the case of a virus like hepatitis C that has a higher survival time [approximately 6 days (Neumann *et al.*, 1998)].

We complement our analytical approach with numerical simulations to study evolutionary and population dynamics during the course of an infection. We tune the parameters so that they fit the case where viruses infect CCR5$^+$ and CXCR4$^+$ T-cells. If a host is infected by a strain better adapted to CCR5$^+$ cells, specialization in the virus population occurs. Once the average trait value in the virus population comes near the branching point, the virus population splits into two morphs. As the infection

proceeds, each morph specializes in exploiting one type of cells. This evolutionary branching is accompanied by a pronounced drop and continued decline of uninfected target cell density because of an increased adaptation of the virus to the host. In the case where the host is infected by a strain better adapted to CXCR4$^+$ cells, we observe a similar scenario, but the whole process is much slower (and is thus not likely to occur before host's death). In other words, evolution in this case favours generalist strains. Interestingly, even though the virus has a low production rate in this case, it still predominantly infects CCR5$^+$ cells in the initial stages of an infection because they yield a higher burst size.

Insights into the co-receptor switch

We designed a general model, because it helps to obtain an understanding of underlying processes and offers insights into several host–pathogen systems. Because of its generality, our model does not encompass the whole complexity of an HIV infection. However, we think that the results we obtain in the case with CXCR4$^+$ and CCR5$^+$ T-cells can yield useful insights into the specific case of the co-receptor switch. This is supported by the fact that CCR5$^+$ and CXCR4$^+$ T-cells tend to dominate in different compartments: the former is more frequent in the gut (Meng *et al.*, 2002) and in the brain (Burkala *et al.*, 2005), whereas the latter is more frequent in the blood (Moore *et al.*, 2004) and in the spleen (Burkala *et al.*, 2005).

The evolutionary branching we find shares three important features with the co-receptor switch. First, the virus is initially found in CCR5$^+$ cells. We show that this is attributed to the fact that the burst size is always higher in these cells so that they are a better resource for the virus. Second, the switch (or branching) occurs late in some (but, importantly, not all) HIV infections. In our model, this is because of the difference between the trait value of the infecting strain and the value of the evolutionary singular strategy. In some cases, the time taken for the virus trait to evolve to the branching value is likely to exceed the duration of the infection. Third, the switch is associated with a drop in the total number of uninfected T-cells. Incidentally, the bifurcation analysis shows that branching occurs more easily in regions that correspond to biological differences between CCR5$^+$ and CXCR4$^+$ T-cells, i.e. between activated/memory cells and resting/naive cells.

In our main model, virus strains are distinguished solely by their cell exploitation strategy (i.e. their production rate in one type of cells), and we do not make an explicit assumption about the type of cell a virus can infect. In reality, one can tell from the RNA sequence of a virus the type of cells it can infect (R5 variants can only infect CCR5$^+$ T-cells, X4 variants can only infect CXCR4$^+$ T-cells, and R5X4 variants can infect both). Moreover, resting/naive cells express only the CXCR4 co-receptor, but activated/memory T-cells tend to express both CCR5 and CXCR4 co-receptors. The assumption we made here is that cell specificity of a virus does not have a major role, because it evolves rapidly (2 or 3 changes in amino acids are sufficient to change this cell specificity, Moore *et al.*, 2004). In other words, we assumed that the time scale on which the trait coding for the 'variant' of the virus (i.e. R5, X4, or R5X4) changes is fast compared to the time scale on which the trait p changes. This simplifying assumption already allows us to interpret the results of the basic model in the context of HIV co-receptor switch. However, to strengthen the link with the biology of HIV, we studied another model in which we modelled the variant of the virus (R5, X4 or R5X4), as well as the epxression of the co-receptors by the naive or the memory cells, explicitly (following Regoes & Bonhoeffer, 2002; Ribeiro *et al.*, 2006). In this second model, a virus strain is defined by two traits: its cell exploitation strategy (p) and the type of co-receptor it can use to enter a cell. In Fig. S1 of the Appendix S1A, we show that this more complicated model yields qualitatively similar results (the only difference has to do with the exact timing of the events). Thorougher analysis is always possible (for instance, one could investigate the effect of the linkage disequilibrium between the two virus traits) but, what we think is the important point here, is that key features of the co-receptor switch can be captured without specifying the variant a virus belongs to.

Before the discovery of the role of the CXCR4 and CCR5 as co-receptors for HIV, studies used to refer to a 'phenotypic switch' from nonsyncytium-inducing viruses (NSI) to syncytium-inducing viruses (SI). This switch comes from differences in replicative abilities of the virus (Moore *et al.*, 2004). It was shown later on that NSI viruses infect CCR5$^+$ T-cells and SI viruses infect CXCR4$^+$ T-cells. This means that the switch involves changes in target cell preference *and* in viral phenotypic traits. Current approaches use molecular criteria to define HIV variant (R5, X4 and R5X4) according to the type of cells they can infect. We show that these variants could be an emergent property of the evolutionary dynamics of a phenotypic trait (the virus production rate). The variant a virus belongs to, i.e. the type of immune cells it infects predominantly, depends not only on the genotype of the virus but also on its environment (the other viral strains and the cell densities).

Conclusion and perspectives

In the 1990s, models generated a rich discussion by suggesting that the combination of population dynamics and trait evolution could account for the course of the infection of HIV (Nowak *et al.*, 1990). Arguably, the role of virus evolution in the course of an HIV infection is still largely unclear (Rambaut *et al.*, 2004). This is illustrated by the co-receptor switch phenomenon because some models show that it can be explained with population dynamics (Ribeiro *et al.*, 2006) or by virus evolution only (Regoes & Bonhoeffer, 2005). We show that the combination of the two, i.e. population and evolutionary dynamics, yields interesting insights by stressing the importance of the within-host environment, i.e. the types and abundances of target cells and the phenotypic traits of the virus strains present in the host. In particular, our results suggest that the force driving the co-receptor switch could be the adaptation to different types of cells and that the dichotomy between R5 and X4 viruses could be the result of this process.

Many studies in adaptive dynamics tend to focus on long-term equilibria and their evolutionary stability, which has led to some criticism regarding the applicability of the results obtained using this framework (see e.g. Waxman & Gavrilets, 2005, and the resulting correspondence). Here, we show that the evolutionary branching can have a tangible biological interpretation. Moreover, we show that the population dynamics of cells driven by trait evolution echo clinical data. This suggests that within-host dynamics of rapidly evolving pathogens offers a possibility to test predictions made by adaptive dynamics models.

Our framework can be enriched in many ways. First, more than two types of target cells could be considered, which could potentially lead to more complex patterns in virus diversification. Second, the model could be

extended to incorporate explicit spatial structure (as in Funk *et al.*, 2005). This could lead to spatial self-structuring, which would allow for dynamics to emerge at the level of clusters of cells (Lion & van Baalen, 2008). Finally, a coherent sequel study would be to add another level of dynamics to the model by considering between-host dynamics. This would allow us to integrate selective pressures acting at different levels such that strains that transmit best from one host to the next need not be the strains that take over or dominate in the host. In the light of HIV dynamics, we find that, regardless of the trait that initiates the infection, there can be a co-receptor switch. However, we also find that the timing of the switch depends on the trait value. Because the switch goes along with a worsening of the patient's health, it will strongly affect the number of secondary infections generated (Fraser *et al.*, 2007), which calls for an explicit epidemiological model.

Our study shows that theory in evolutionary ecology can be applied to within-host dynamics and can enhance our understanding of the course of an infection. This has implications for animal intra-cellular pathogens in general. The existence of multiple types of infected cells that are spatially structured in the host has been shown for other human diseases than HIV, e.g. hepatitis C virus (Roque-Afonso *et al.*, 2005) or Epstein–Barr virus (Sitki-Green *et al.*, 2003). Our model could be extended to include more details for a specific pathogen and be used to predict the conditions under which one would expect viral diversification in the host.

Acknowledgments

References

Ariën, K.K., Gali, Y., El-Abdellati, A., Heyndrickx, L., Janssens, W. & Vanham, G. 2006. Replicative fitness of CCR5-using and CXCR4-using human immunodeficiency virus type 1 biological clones. *Virology* **347**: 65–74.

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.

Berger, E.A., Murphy, P.M. & Farber, J.M. 1999. Chemokine receptors as HIV-1 co-receptors: roles in viral entry, tropism, and disease. *Annu. Rev. Immunol.* **17**: 657–700.

Bleul, C.C., Wu, L., Hoxie, J.A., Springer, T.A. & Mackay, C.R. 1997. The HIV co-receptors CXCR4 and CCR5 are differentially expressed and regulated on human T lymphocytes. *Proc. Natl Acad. Sci. USA* **94**: 1925–1930.

Boldin, B. & Diekmann, O. 2008. Superinfections can induce evolutionarily stable coexistence of pathogens. *J. Math. Biol.* **56**: 635–672.

Boldin, B., Geritz, S. & Kisdi, E. 2009. Superinfections and adaptive dynamics of pathogen virulence revisited: a critical function analysis. *Evol. Ecol. Res.* **11**: 153–175.

Burkala, E.J., He, J., West, J.T., Wood, C. & Petito, C.K. 2005. Compartmentalization of HIV-1 in the central nervous system: role of the choroid plexus. *AIDS* **19**: 675–684.

Caraco, T. & Wang, I.N. 2008. Free-living pathogens: life-history constraints and strain competition. *J. Theor. Biol.* **250**: 569–579.

Coffin, J.M. 1995. HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. *Science* **267**: 483–489.

Connor, R.I., Sheridan, K.E., Ceradini, D., Choe, S. & Landau, N.R. 1997. Change in co-receptor use correlates with disease progression in HIV-1–infected individuals. *J. Exp. Med.* **185**: 621–628.

Cuevas, J.M., Moya, A. & Elena, S.F. 2003. Evolution of RNA virus in spatially structured heterogeneous environments. *J. Evol. Biol.* **16**: 456–466.

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

Diekmann, O. & Heesterbeek, J. 2000. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis, and Interpretation*. Wiley, New York.

Eckstein, D.A., Penn, M.L., Korin, Y.D., Scripture-Adams, D.D., Zack, J.A., Kreisberg, J.F., Roederer, M., Sherman, M.P., Chin, P.S. & Goldsmith, M.A. 2001. HIV-1 actively replicates in naive CD4$^+$ T-cells residing within human lymphoid tissues. *Immunity* **15**: 671–682.

Fouchier, R.A., Groenink, M., Kootstra, N.A., Tersmette, M., Huisman, H.G., Miedema, F. & Schuitemaker, H. 1992. Phenotype-associated sequence variation in the third variable domain of the Human Immunodeficiency Virus type 1 gp120 molecule. *J. Virol.* **66**: 3183–3187.

Fraser, C., Hollingsworth, T.D., Chapman, R., deWolf, 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.

Funk, G.A., Jansen, V.A.A., Bonhoeffer, S. & Killingback, T. 2005. Spatial models of virus-immune dynamics. *J. Theor. Biol.* **233**: 221–236.

Geritz, S.A.H., Kisdi, E., Meszéna, G. & Metz, J.A.J. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* **4**: 1–79.

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.

González-Scarano, F. & Martín-García, J. 2005. The neuropathogenesis of AIDS. *Nat. Rev. Immunol.* **5**: 69–81.

Holt, R.D. & Barfield, M. 2006. Within-host pathogen dynamics: Some ecological and evolutionary consequences of transients, dispersal mode, and within-host spatial heterogeneity. In: *Disease Evolution: Models, Concepts, and Data Analyses* (Z. Feng, U. Dieckmann & S.A. Levin, eds), vol. 71 of DIMACS Series

in Discrete Mathematics and Theoretical Computer Science, pp. 45–66. American Mathematical Society, Providence.

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.

Kepler, T.B. & Perelson, A.S. 1998. Drug concentration heterogeneity facilitates the evolution of drug resistance. *Proc. Natl Acad. Sci. USA* **95**: 11514–11519.

Levins, R. 1968. *Evolution in Changing Environments*. Princeton University Press, Princeton, NJ.

Lion, S. & van Baalen, M. 2008. Self-structuring in spatial evolutionary ecology. *Ecol. Lett.* **11**: 277–295.

Luciani, F. & Alizon, S. 2009. The evolutionary dynamics of a rapidly mutating virus within and between hosts: the case of hepatitis C virus. *PLoS Comput. Biol.* **5**: e1000565.

Meng, G., Wei, X., Wu, X., Sellers, M.T., Decker, J.M., Moldoveanu, Z., Orenstein, J.M., Graham, M.F., Kappes, J.C., Mestecky, J., Shaw, G.M. & Smith, P.D. 2002. Primary intestinal epithelial cells selectively transfer R5 HIV-1 to CCR5$^+$ cells. *Nat. Med.* **8**: 150–156.

Meszéna, G., Czibula, I. & Geritz, S.A.H. 1997. Adaptive dynamics in a 2-patch environment: a toy model for allopatric and parapatric speciation. *J. Biol. Syst.* **5**: 265–284.

Moore, J.P., Kitchen, S.G., Pugach, P. & Zack, J.A. 2004. The CCR5 and CXCR4 co-receptors–central to understanding the transmission and pathogenesis of Human Immunodeficiency Virus type 1 infection. *AIDS Res. Hum. Retroviruses* **20**: 111–126.

Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J. & Perelson, A.S. 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* **282**: 103–107.

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.

Orive, M.E., Stearns, M.N., Kelly, J.K., Barfield, M., Smith, M.S. & Holt, R.D. 2005. Viral infection in internally structured hosts. I. Conditions for persistent infection. *J. Theor. Biol.* **232**: 453–466.

Perelson, A.S. 2002. Modelling viral and immune system dynamics. *Nat. Rev. Immunol.* **2**: 28–36.

Rambaut, A., Posada, D., Crandall, K.A. & Holmes, E.C. 2004. The causes and consequences of HIV evolution. *Nat. Rev. Genet.* **5**: 52–61.

Ravigné, V., Dieckmann, U. & Olivieri, I. 2009. Live where you thrive: joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. *Am. Nat.* **174**: E141–E169.

Regoes, R. & Bonhoeffer, S. 2002. HIV co-receptor usage and drug treatment. *J. Theor. Biol.* **217**: 443–457.

Regoes, R.R. & Bonhoeffer, S. 2005. The HIV co-receptor switch: a population dynamical perspective. *Trends Microbiol.* **13**: 269–277.

Ribeiro, R.M., Hazenberg, M.D., Perelson, A.S. & Davenport, M.P. 2006. Naïve and memory cell turnover as drivers of CCR5-to-CXCR4 tropism switch in Human Immunodeficiency Virus type 1: implications for therapy. *J. Virol.* **80**: 802–809.

Roque-Afonso, A.M., Ducoulombier, D., Di Liberto, G., Kara, R., Gigou, M., Dussaix, E., Samuel, D. & Féray, C. 2005. Compartmentalization of hepatitis C virus genotypes between plasma and peripheral blood mononuclear cells. *J. Virol.* **79**: 6349–6357.

Schluter, D. 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, New York, USA.

Sitki-Green, D., Covington, M. & Raab-Traub, N. 2003. Compartmentalization and transmission of multiple Epstein–Barr virus strains in asymptomatic carriers. *J. Virol.* **77**: 1840–1847.

Stafford, M.A., Corey, L., Cao, Y., Daar, E.S., Ho, D.D. & Perelson, A.S. 2000. Modeling plasma virus concentration during primary HIV infection. *J. Theor. Biol.* **203**: 285–301.

Stebbing, J., Gazzard, B. & Douek, D.C. 2004. Where does HIV live? *N. Engl. J. Med.* **350**: 1872–1880.

Turner, P.E. & Elena, S.F. 2000. Cost of host radiation in an RNA virus. *Genetics* **156**: 1465–1470.

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.

Ward, Z.D., White, K.A.J. & van Voorn, G.A.K. 2009. Exploring the impact of target cell heterogeneity on HIV loads in a within-host model. *Epidemics* **1**: 168–174.

Waxman, D. & Gavrilets, S. 2005. 20 questions on adaptive dynamics. *J. Evol. Biol.* **18**: 1139–1154.

Weinberger, A.D., Perelson, A.S., Ribeiro, R.M. & Weinberger, L.S. 2009. Accelerated immunodeficiency by anti-CCR5 treatment in HIV infection. *PLoS Comput. Biol.* **5**: e1000467.

Wodarz, D. & Nowak, M.A. 1998. The effect of different immune responses on the evolution of virulent CXCR4-tropic HIV. *Proc. Biol. Sci.* **265**: 2149–2158.

Supporting information

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

Appendix S1 Supplementary results (A) and methods (B).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.