A Mathematical Model of Continuous HIV Mutations Eluding Immune Defence

D.P. WILSON* and D.L.S. McELWAIN

Centre in Statistical Science and Industrial Mathematics, Queensland University of Technology, Brisbane, Australia

We develop a description of HIV mutations based upon a continuum representation of the fitness of the virus, including the interaction of the virus with both specific Th1 lymphocytes as well as cross-reactive cells. This deterministic model allows a straightforward measure of the diversity of viral population and reproduces the observed increase in diversity as the disease progresses in an untreated patient. We use the diversity threshold theory, extending the modelling to track mutations on a continuum. When the diversity threshold is exceeded, the host immune system collapses.

INTRODUCTION

The development of AIDS is associated with the depletion of the most crucial cell type of the immune system, the $CD4^+$ helper T lymphocyte. This occurs after a period of clinical latency varying in length from $1-2$ years to more than 15 years. The human immunodeficiency virus (HIV) and the immune system interact dynamically over the entire course of infection. HIV replicates in large numbers each day and also contributes to the destruction of cells of the immune system. This growth is usually balanced by a strong defensive response which keeps the virus in check. However, eventually the balance of power is transferred so that HIV gains the upper hand and causes this severe damage to the immune system. This defines the period of disease known as AIDS.

As with all viruses, HIV cannot reproduce by itself. It relies on a host to assist reproduction. HIV targets $CD4^+$ T cells which are the most abundant white blood cells of the immune system. Since $CD4^+$ T cells play a key role in orchestrating both the humoral and the cellular immune response, the infection and destruction of this cell type by HIV has a devastating impact on the immune response. When HIV enters the body, it targets cells with $CD4^+$ receptors. The gp120 protein on the viral particle binds to the $CD4^+$ T cell receptors and the HIV particle injects its core. HIV, being a retrovirus carries a copy of its RNA, which is first transcribed into DNA. After the processes of reverse transcription, integration of viral DNA into cellular DNA, translation forming polypeptides and protease, and cleaving of polypeptides by protease into functional HIV proteins, the infected cell is stimulated to produce copies of the virus, releasing hundreds, or even thousands, of free virus particles that can infect other $CD4^+$ T-cells. We assume that peripheral blood CD4 counts (generally 1000 per μl in healthy persons) are a good indicator of CD4 densities in the body. Since the $CD4^+$ cells are depleted over time, strengthening cytotoxic responses cannot occur. The loss of immune competence enables normally controllable diseases to cause opportunistic infections. In addition to a reduced CD4 count, one of the reasons for the specific immune response dysfunction is that viral epitopes also change resulting in escape from immunological response.

Initially, the transformation of immune-sensitive to resistant genotypes occurs by the generation of mutations primarily due to reverse transcriptase errors. HIV's RNA genome is reverse transcribed into DNA when the virus replicates. Amplification of mutant genotypes by selection, and then fixation of the resistant population aids in the problem of resistance. These mutations are changes in DNA structure, and, therefore, changes in protein and phenotype. The mutations we consider are caused by spontaneous errors during DNA replication (base substitutions), or frame shift mutation (addition or deletion of a base). They are heritable and may be beneficial, neutral or detrimental to an organism. In contrast to most other infectious agents, HIV combines a very high mutation rate

*Corresponding author. E-mail: d.wilson@fsc.qut.edu.au

with a very large production of virions. The extreme heterogeneity and diversity of HIV makes the design of effective vaccines extremely difficult (Domingo *et al.*, 1997). A virus' biological fitness determines its survival. It is this phenomenon, of viral mutation, causing escape from immunological attack, that we address in this paper.

The understanding of the dynamics of antigenic escape from immunological response has been that a mutation may enable the virus to have a selective advantage. Mansky and Temin (1995) found that the level of genetic variation of HIV is high relative to similar retroviruses, although not as high as earlier predicted. They were able to estimate the mutation rate for HIV (see below).

There is measurable variation among the replication rates of HIV strains in $CD4^+$ cells (Nowak *et al.*, 1991). Mathematical models for the interaction of HIV and $CD4^+$ suggest that evolutionary forces may drive selection for more virulent strains (May *et al.*, 1990; DeBoer and Boerlijst, 1994; Maree *et al.*, 2000). The observations that the apparent diversity of viral strains increases with time (Balfe *et al.*, 1990; Holmes *et al.*, 1992) has eventually lead to the conclusion that evolutionary dynamics of HIV quasispecies is not a consequence but a cause of immunodeficiency (Nowak and McMichael, 1995). In this paper, we suggest a novel approach to considering viral diversity and the progression of mutation towards diversity threshold at which time AIDS results.

The number of immunologically different mutants appears to be very large. Genetic variation is not uniform throughout the HIV genome. There is more variability in the *env* gene than in the *gag* or *pol* genes. Within the *env* gene there are five hypervariable regions. Much attention has been paid to the third hypervariable region, the V3 loop because it has immunodominant properties. The combinatorial possibilities of the 19 variable amino acids in the V3 loop of the *env* protein is 19^{20}, and furthermore the shape of the loop can also be altered by mutations in other parts of the envelope protein (Nowak *et al.*, 1991). For information on genetic and antigenic variation of the V3 loop refer to (Simmonds *et al.*, 1990; Moore and Nara, 1991; Wolfs *et al.*, 1991).

The complexity of the retroviral lifecycle results in difficulty in estimating the rate of mutation. However, Mansky and Temin (1995) state that the error rate has been estimated as 3.4×10^{-5} per base per replication cycle. Since the average HIV genome has approximately 10^4 bases, the average changes per genome is approximately 0.34 per replication cycle (Colgrove and Japour, 1991). In addition to the occurrence of large antigenic variation that occurs due to this probability, a very large number of virions are produced each day during the asymptomatic period of HIV progression before AIDS (Perelson and Nelson, 1999).

During early stages of infection the viral strains that are most prevalent in initial infection dominate the dynamics. As the immune response to HIV is initiated virus diversification results. This antigenic diversity makes it difficult for the immune system to control the different mutants simultaneously and the virus *runs ahead* of the immune response. This is also because there is an asymmetric interaction between immunological specificity and viral diversity. Martin Nowak was the first to point out this asymmetry with the diversity threshold model (See Nowak, 1990; 1992; Nowak *et al.*, 1991; Nowak and May, 1992; 1993). That is, each viral strain can infect all immune cells with $CD4^+$ protein, but individual strain-specific immune cells can only attack specific virus strains. The outcome of mutation is that each viral genome must be viewed as being different from any other but at the same time viral strains may be thought of as populations of closely related genomes. The immune system is then persistently confronted with new antigen targets before it has built up a defence to antigen already present.

It has also been suggested that viral genetic diversity is caused by the host immune response. Findings by Lukashov *et al.* (1995) and Wolinsky *et al.* (1996) indicate that HIV undergoes adaptive evolution *in vivo*, in response to selective pressure exerted by the host immune system. Evidence for immune-response dependence of virus mutation can be observed by noting that viral genetic diversity is greatest in clinically healthy individuals and much less in persons with AIDS, where the immune response has become severely impaired.

Many have mathematically modelled virus mutation in a discrete form in an attempt to understand this phenomenon (see, for example, Colgrove and Japour, 1991; DeBoer and Boerlijst, 1994; Kirschner and Webb, 1997). But since there are a large number of mutations and each is considered to be very closely related to its *parent* strain, it seems that a reasonable proposal is to model mutation in a continuous form. This gives rise to this novel continuous model. The model of Nowak *et al.* (1991) involving the tracking of many viral variants is extended here whereby antigenic diversity and immune response are considered as a continuum. In our model, we assume that all possible mutants will exist at all times, even if at minute concentrations, and the viral distribution will change form or shift as mutation occurs.

MODELLING VIRAL MUTATION

In this section, we present the essential mutation concepts for our mathematical model that explores antigenic drift of HIV. Mutations are assumed to involve small changes in the gene sequence of a viral strain. These small sequence changes may alter the viral properties considerably. However, we consider viruses associated with a neighbourhood of sequences in sequence space around a given sequence to have very closely related properties. Here, we are assuming the existence of criteria for defining a neighbourhood of gene sequences in sequence space such that all viruses containing sequences in this space have quantitative properties in a measurable

neighbourhood. The properties of a virus are influenced by many aspects associated with the complicated sequence. We suggest that all viral strains, influenced multi-dimensionally, can be ordered and mapped onto a one-dimensional continuum. In reality amino acids are discrete and mutations are obviously not continuous. But for reasons of elegance, we develop a continuous mutation model. In our model, we let $V(t)$ be the total virus population, and $v(r,t)$ be the population of virus identified by parameter r at time t. The identifier for a viral strain, r, is an attribute measuring the strain's fitness. We regard the non-negative parameter r as a continuous variable. We refer to fitness as an ordered measure of competitive advantage. One way to define fitness is in terms of the basic reproduction ratio of the population (Anderson and May 1979a,b; 1991; Bonhoeffer *et al.* 1997; 2002). For mathematical convergence and to reduce our assumptions, we define the domain of r to be finite, such that $0 \le r \le M$. Thus defining fitness in terms of the standard reproduction ratio would not be appropriate in our model due to the consequent complexity.

The total virus load at time t is given by

$$V(t) = \int_0^M v(r, t) \, dr. \tag{1}$$

The time rate of change of viral strain r is given by

$$\frac{\partial v}{\partial t} = \lambda \int_0^M k(r, r')v(r', t) \, dr' - \lambda \int_0^M k(r', r)v(r, t) \, dr'$$

$$+ \mu(r, t). \tag{2}$$

Here, λ is the replication rate per unit time and $k(r, r')$ is the probability that viral strain r' mutates to viral strain r. We consider mutations to and away from viral strain r. The function, $\mu(r,t)$, describes the net growth rate of the viral strain with fitness r. This function is described later.

Since a virion's genome changes only slightly as it mutates, we assume that this affects the fitness parameter, r, incrementally. Here, we assume that small genome mutations influence small changes in the corresponding viral fitness properties although it is possible that this results in a large change in viral fitness. Then for $r' \ne r$ we assume the form, $\alpha e^{-\beta(r - r')^2}$, for the probability of mutation from strain r to strain r', where α and β are to be determined.

However, for $r = r'$, there is no mutation. Thus we include an *atom* $P_0(r)$, for $k(r,r)$, obtaining a quasi-continuous distribution. We have already resolved that the probability that a genome will change in a replication cycle is 0.34. Then $P_0 = 0.66$, independent of r. We obtain a probability distribution of the form

$$k(r, r') = \begin{cases} P_0, & r = r' \\ \alpha e^{-\beta(r-r')^2}, & r \ne r'. \end{cases}$$

Equation (2) becomes

$$\frac{\partial v(r, t)}{\partial t} = \lambda \int_0^M k(r, r')v(r', t) \, dr' - \lambda(1 - P_0)v(r, t)$$

$$+ \mu(r, t). \tag{3}$$

To determine α, we consider $\lim_{r \to r'} k(r, r') = \alpha$, and suggest a parameter θ, $0 \le \theta \le 1$ such that $\alpha = \theta P_0$. If there is no mutation, v is independent of r, $\theta = 0$ and we set $P_0 = 1$. Then $\partial v / \partial t = \mu(r, t)$. If $\theta = 1$, the probability density function is continuous suggesting that the mutations occur readily from strain r, however, the mutations will be closely related to the parent strain in fitness properties. The smaller the parameter θ takes, the more likely it is that a mutant will have properties away from the parent strain. Mutation occurs during the reverse transcription process in an infected immune cell and then mutant virion will be released from the infected cell. Thus, we indirectly model mutation amongst released virions by modelling species shifts in the viral population distribution.

Since k is a probability density function, we require

$$\int_0^M k(r, r') \, dr' + k(r, r) = 1.$$

Thus, $\beta = \beta(r)$ and

$$\frac{\theta P_0}{2} \sqrt{\frac{\pi}{\beta(r)}} \Big(\mathrm{erf}\big(r\sqrt{\beta(r)}\big) - \mathrm{erf}\big((r - M)\sqrt{\beta(r)}\big) \Big)$$

$$+ P_0 - 1 = 0.$$

This equation allows the determination of the value of β, given a value of r. The function, $\beta(r)$ will be symmetric about $r = M/2$. A plot of β versus r is shown in Fig. 1. It should be intuitive that β is a function of r as illustrated by plots of $k(r,r')$ for fixed r in Fig. 2. We choose values
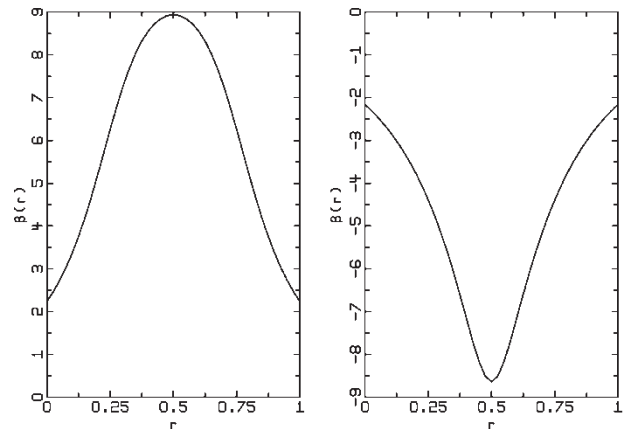

FIGURE 1 Plot of the function $\beta(r)$ determining spread of virus mutation. Here, $M = 1$. (a) $\theta = 0.9$, (b) $\theta = 0.2$.
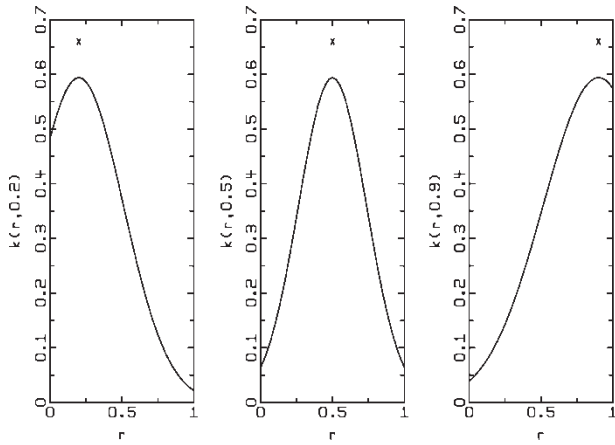
FIGURE 2 Plot of the probability density function $k(r,r')$ for fixed values of r. Here, $M = 1$ and $\theta = 0.9$. (a) $r' = 0.2$, (b) $r' = 0.5$ and (c) $r' = 0.9$.

for the parameter θ so that $\beta > 0$ for realism. If θ is too low, β will become negative indicating higher likelihood of mutation to a fitness farther away from the *parent* strain.

MODELLING IMMUNE INTERACTION WITH ANTIGEN

The immune system is capable of generating vast diversity in its recognition of foreign antigens and distinguishing subtle differences between them. Lymphocytes adapt to recognize a particular molecular configuration on cells to be attacked. Lymphocytes are generated with random receptors for antigen. A T-lymphocyte has of the order of 10^5 receptors each with an identical specificity for a particular antigen. Goldsby *et al.* (2000) state there are of the order of 10^9 antigenic specificities in the immune system although this is later diminished through a selection process in the thymus ensuring that there are only non-self receptors. Once an antigen with particular specificity is recognized by a T-lymphocyte, the lymphocytes become activated causing clonal expansion, so that a much higher proportion of T-lymphocytes will be able to recognize any further infection, providing the basis of immunological memory. Since there is great diversity in the immune specificity, we consider $s(r,t)$ to be the total number of T lymphocytes with specificity for viral antigen with fitness between r and $r + dr$ at time t.

Null cells, a type of lymphocyte that generally makes up about 5–10% of all lymphocytes in human blood (Goldsby *et al.*, 2000), do not express the membrane molecules and receptors that distinguish T and B lymphocytes. They lack immunological memory and specificity but have shown cytotoxic behaviour against a variety of antigens. In addition, when a mutation occurs, the viral epitope may be conserved so that a lymphocyte species with committed specificity may respond against multiple strains. Phagocytes, activated

macrophages and other $CD4^+$ cells that can mount immunological attack against different antigenic strains are known as *cross-reactive cells*. Here, we assume cross-reactive cells have negligible influence on our investigation of the evolution of mutation (see Nowak and May (2000) for modelling with cross-reactive cells). In our model, the population of immune cells (majority T lymphocytes), T, at time t is given by

$$T(t) = \int_0^M s(r, t)\,dr. \qquad (4)$$

Many models distinguish between effector cells and target cells, generally CD8 and CD4 cells, respectively. Since the population of effector cells are directly influenced by and signalled by target cells we combine the two immune cell types into the one population. Time rate parameters will incorporate the proportion of cells that are of each type. We assume that the strain specific population is given by the equation

$$\frac{\partial s(r,t)}{\partial t} = s_1 + P_s(r, t) - \delta_S s(r, t) - k_1 V(t)s(r, t), \qquad (5)$$

where s_1, δ_S and k_1 are constants. The parameter s_1 represent the source of immune cells originating in the thymus for the strain specific immune population. Although there are various algebraic expressions that may be used for the source of immune cells, for simplification we take the source, s_1, as constant. The parameter, δ_S, represents the rates of natural death for immune cells. The constant, k_1, represents the rate that the immune cells are infected with virus. The term $P_s(r,t)$ represents the immune response of clonal expansion from existing cells (mainly memory cells), once activated. We assume activation occurs by interaction with virion. When an antigen binds to an antigenically committed T (or B) lymphocyte, the cell is stimulated to divide repeatedly into a clone of cells with the same antigenic specificity as the original parent cell. Specificity and immunological memory are consequences of clonal expansion. We assume that HIV-induced clonal expansion is limited by an immune cell-dependent reaction and consequently is governed by a mass action term. We describe the proliferation as mass action saturated by the concentration of antigen presenting sites. Thus, we take $P_s(r,t) = p_1 v(r,t)s(r,t)/[c_1 + v(r,t)]$. A similar form has been used by Kirschner (1996). Here, p_1 is constant and c_1 is a saturation constant giving the concentration of antigenic sites at which the rate of proliferation is half-maximal.

We denote the total population of immune cells that have become infected at time t by $I(t)$ where

$$I(t) = \int_0^M i(r, t)\,dr, \qquad (6)$$

with $i(r, t)$ representing the population at time t of cells infected with virus of fitness r. We assume that these

species populations are governed by

$$\frac{\partial i(r,t)}{\partial t} = S_i(t)v(r,t) - \delta_I i(r,t) - p_2 i(r,t)$$

$$- I_m(r,t)i(r,t). \tag{7}$$

Here, $S_i(t)$ is a source rate denoting infection of a healthy immune cell and δ_I is a natural death rate. As infected cells are activated we assume they die, releasing infectious virion. This occurs at rate p_2, incorporating activation of the infected cell by a large range of cell types to release its virion. N virions are released and will become a source for the viral population of type "r". The function, $I_m(r,t)$, represents the rate of cytotoxic effect the immune system generates against virus-infected cells.

Strains with greater fitness parameter values have a competitive advantage. We could consider many combinations of factors to influence the selective advantage of a viral strain. However, here we regard the immune response against an antigenic variant to be the influencing factor determining strain fitness in an attempt to model immune evasion by means of evolutionary drift. We assume that the cytotoxic immune response function is given by $I_m(r,t) = \eta(r)s(r,t)$. Here, we consider the immune response to be of the form given by the decreasing function $\eta(r) = \eta_0 r_0^n/(r_0^n + r^n)$, where η_0, n and r_0 are constants. The parameter η_0 is the maximum immune response against the strains of lowest fitness and r_0 is the fitness value in which the immune response is half maximal. The parameter n is a shape factor.

The total source rate of infected cells is given by

$$S_i(t) = V(t)\int_0^M k_1 s(r,t)\,\mathrm{d}r = k_1 T(t)V(t)$$

$$= k_1 T(t)\int_0^M v(r,t)\,\mathrm{d}r,$$

where k_1 has been described above. Then, we take the source of cells that are infected with viral strain r to be $k_1 T(t)v(r,t)$.

MODEL TERMS FOR CHANGE IN VIRAL POPULATION

We have seen that our model presents the total change in viral load by the term $\int_0^M \mu(r,t)\,\mathrm{d}r$. We now specify the components of $\mu(r,t)$ to describe the aspects of the HIV lifecycle that effect its change in population.

We consider that the rate of change of virus concentration is influenced by a source of new virions released from activated infected cells and loss due to natural death. We assume the rate of change of viral strain

"r" to be given by

$$\frac{\partial v}{\partial t} = \lambda \int_0^M k(r,r')v(r',t)\,\mathrm{d}r' - \lambda(1 - P_0)v(r,t)$$

$$+ Np_3 i(r,t) - \delta_V v(r,t). \tag{8}$$

Here the parameter δ_V is the rate of natural death and N is the average number of virions that are released from an infected cell. Although variant strains are released at this stage of the viral lifecycle, since mutation occurs inside the host cell when reverse transcription errors arise, for simplicity we model mutation exterior to the host cell amongst the viral dynamics.

INITIAL CONDITIONS

We assume that the initial inoculum of virus consists of one strain of V_0 per ml and that the initial strain has the weakest fitness, which for convenience we have taken to be $r = 0$. Then,

$$v(r,0) = \begin{cases} V_0, & 0 \le r \le \delta r \\ 0, & r > \delta r. \end{cases}$$

Since we model on a continuum, immediately after inoculation all strains will exist, sourced from the initial strain, even if at minute concentrations. We assume that initially there are no infected immune cells so $i(r,0) = 0$.

Typically, a healthy person will have an average of 1000 CD4$^+$ cells per μl of blood and so we take $T(0) = 1000$. In the maturation process for lymphocytes the random gene rearrangements generate an enormous diversity of antigenic specificities before its contact with antigen. Since generation of specificity is random, before inoculation we assume a uniform distribution of strain specific cells and so $s(r,0) = T(0)/M$.

MEASURING DIVERSITY OF MODEL HIV STRAINS

To measure how genetic variation within the HIV quasispecies changes with time we consider the ecologist's Simpson Index (Goldberg and Estabrook, 1998). A diversity index is a measure of species diversity in a *community*. Diversity indices provide more information about community composition than simply species richness (i.e. the number of species present). They also take the relative abundances of different species into account. Simpson's Diversity Index, D, is a simple measure that characterizes species diversity in a community. In our case of continuous HIV strains, we take

$$D_v = D_v(t) = \frac{[V(t)]^2}{\int_0^M [v(r,t)]^2\,\mathrm{d}r},$$

so that initially $D_v(0) = \delta r \to 0$ for infinitesimal δr. We also track a diversity index for immune strain specific cells to indicate how the immune response is adapting to the presented antigen. Then, we let

$$D_s = D_s(t) = \frac{[T(t)]^2}{\int_0^M [s(r,t)]^2 \, dr},$$

so that initially $D_s(0) = M$. Using our adapted Simpson's diversity index for the continuous case, we see that $0 \leq D \leq M$ where a small value of D corresponds to little diversity.

Antigenic change is important to the success of the virus. We would expect to see antigenic change in our model and, therefore, see the value of D_v increase to resemble the high mutation rate. Over time, antigenic drift will occur whereby, in a practical sense, gradual mutation of the gene sequences for the glycoproteins will occur at a constantly slow rate and there will be a shift in the distribution of viral dominance. The value of D_v will reflect this shift. Accordingly, the strain specific immune cells will adapt to confront the presenting challenge of HIV strains. In effect, since most CD4 immune cells do not contribute to the destruction of the virus but are merely targets for the virus, because of the asymmetric dynamics, they, in fact, make a negative impact on the system as a whole. The distribution of strain specific cells will attempt to reflect the distribution of viral strains and the adapted Simpson's diversity index for immune diversity, D_s, will track the immune system's evolution.
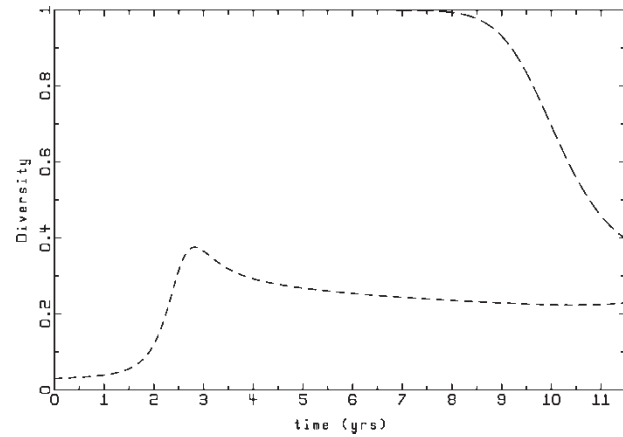

FIGURE 3   Simulation showing change in adapted Simpson's Diversity Index for viral population and specific immune cell population with time. D_s (long dashed line), D_v (shorter dashed line).

RESULTS

Prior to infection we expect initial conditions to be the constant population levels for all species and, therefore, from Eq. (5), $0 = s_1 - \delta_s T_0$. Thus, $\delta_s = s_1/T_0$. The parameter values are shown in Table I. Some parameter values may vary from individual to individual. The values of parameters are taken from other models (Kirschner, 1996; Perelson *et al.*, 1996; Kirschner and Webb, 1997). Numerical simulations can now be carried out and the output is presented in Figs. 3–7.

In our simulations in which immune cells are depleted we see an increasing antigenic diversity until a threshold level, a competitive reduction and then approximately constant diversity thereafter (Fig. 3). The rapid change in the dynamics of the virus diversity occurs prior to immune collapse. The immune distribution is not specifically targeted enough to handle the relatively peaked viral distribution. Towards the end of our simulation we observe a decrease in immune specificity to combat the virus but it

TABLE I   Parameter values for HIV mathematical model

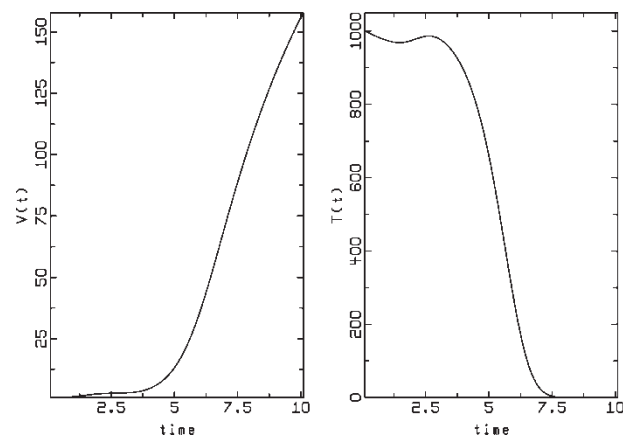| Parameters and Constants	Values
M = Upper range of domain of all measured fitness's	1
V_0 = Initial inoculation concentration of virus	10^{-3}
P_0 = Probability that HIV genome will not mutate in a replication cycle	0.66
θ = Measure of amount of mutation	$0 \leq \theta \leq 1$
α = Coefficient of mutation function	$\alpha = \theta P_0$
λ = Rate of mutation/replication cycle	0.38
N = Average number of virions released upon bursting of infected cell	1000
δ_V = Rate of natural death of HIV	3.33
δ_S = Rate of natural death of strain specific immune cells	0.01
δ_I = Rate of natural death of HIV-infected immune cells	0.5
η_0 = Maximal rate of infected cell death due to cytotoxic immune cells	0.001
r_0 = Fitness value for which the immune response is half maximal	$\frac{M}{2}$
n = Shape factor for the immune response function	4
s_1 = Total source of strain specific immune cells from the thymus	10
p_1 = Maximal proliferation rate constant (Strain-specific cells)	0.01
p_2 = Release of virion from infected cells rate constant	0.001
c_1 = Viral half-maximal saturation constant for proliferation	100
k_1 = Rate of HIV-infection of immune cells	0.001

FIGURE 4   Simulation of model equations accounting for immunological memory. Resembles the typical course of infection in an untreated patient. (a) Total viral load with time. (b) Total immune cell population with time.
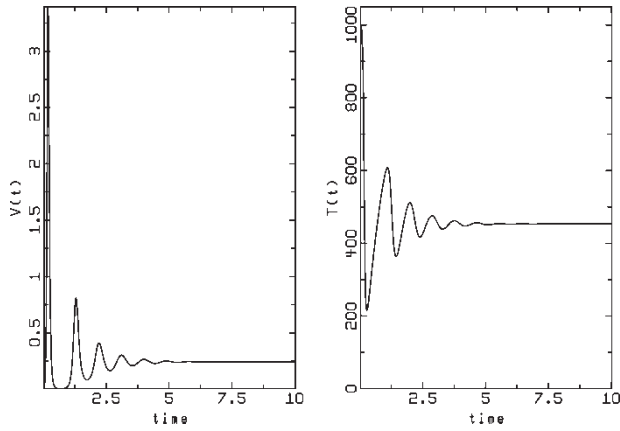
FIGURE 5   Simulation of the dynamics without mutation. (a) Total viral load with time. (b) Total immune cell population with time.

is too little-too late. We suggest that to control the viral load, the immune diversity would need to be as small as the viral diversity (obviously evolved to specificity around the peak viral distribution). Then, extraneous immune cells are not purely target cells for HIV but effector cells against the virus.

DISCUSSION

A key cause for HIV's devastating impact is the generation of mutants primarily due to reverse transcription errors enabling the virus to escape from immunological response. The level of genetic variation of HIV is high and the number of immunological mutants is very large. Immune resistance is thought to be induced by mutation and so has been the topic of much study. The mathematically modelling of a finite number of mutants in a discrete form has been carried out by other researchers (for example, May and Nowak). However, we consider mutations on a continuum to more accurately represent the antigenic diversity.
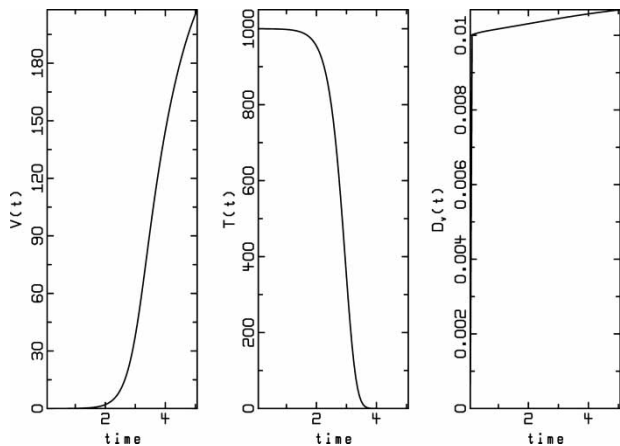

FIGURE 6   Simulation of the dynamics with low effector immune response. (a) Total viral load with time. (b) Total immune cell population with time. (c) Adapted Simpson's diversity index with time.
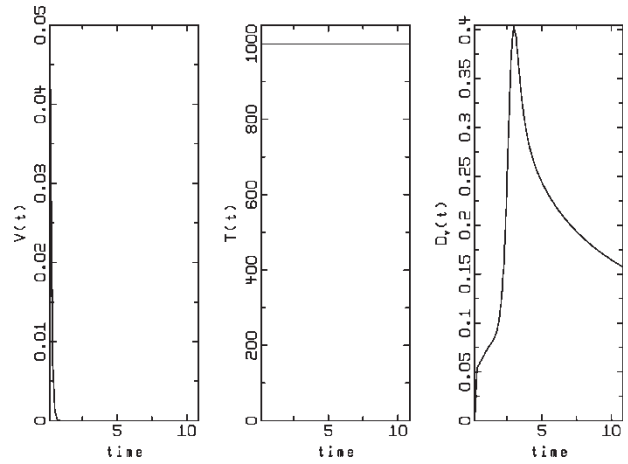

FIGURE 7   Simulation of the dynamics with low rate of infection of immune cells. (a) Total viral load with time. (b) Total immune cell population with time. (c) Adapted Simpson's Diversity Index with time (see text for description).

Our model adequately represents the natural course of HIV in a typical untreated patient (See Fig. 4). We have varied parameters to identify key aspects leading to AIDS. Our model suggests a number of conclusions about the dynamics of HIV as it interacts with the immune system and the evolutionary drift of the virus frequency to evade the immune response, eventually resulting in AIDS.

Nowak and colleagues have proposed the existence of an *antigenic diversity threshold*, in which the genetic diversity of HIV is ever-expanding and eventually exhausts the immune system's capacity to respond, resulting in immune collapse (Nowak and May 1990; Nowak *et al.* 1991, 1995). Although a well-presented theory, direct support for this theory has been elusive. Our model incorporates the assumptions of the diversity threshold theory and consequently our numerical solutions show this. We suggest that if mutation is rapid, which we assume it is *in vivo*, then the virus will run ahead of the immune system's capacity to withhold the virus' impact. We measure diversity with an adapted Simpson's index. We suggest that to control the viral load, the immune diversity would need to be as small as the viral diversity.

Without mutation the immune system can control the impact of HIV infection and the level of immune cells remains at a steady value, although at a compromised level in comparison with the pre-infection level, as shown in Fig. 5. Thus, our model leads us to conclude that mutation is important to the pathogenesis of disease. If mutation occurs readily and there is a high probability that mutant strains are closely related to the parent strain in regards to fitness, then evolutionary drift to escape immune response will occur but it will occur quite slowly. In fact, in this case large antigenic drift will not occur because there are high levels of strains in a local area, which actually forces immune evolutionary drift through proliferation. Proliferation, regardless of its extent, will not be adequate to prevent immune collapse. However, if mutation from a strain has a high likelihood of obtaining a significantly

different fitness, then evolutionary drift will occur rapidly. It will occur at a rate, which is too rapid for the immune system to adapt and the virus will eventually dominate the system, resulting in AIDS. Here, the rate of immune depletion is noticeable immediately and immune cell levels continue decreasing steadily, as opposed to most situations in which there is a slower decrease in immune cells during the majority of the dynamics before a rapid decrease.

Proliferation of naive cells to stimulate memory cells is crucial in the immune defence. Our model suggests that if the proliferation of immune cells is high, then, despite the increased proliferation enabling immune control over the antigenic drift in fitness of antigens for a prolonged time, the immune response will never be strong enough to contain the virus and AIDS will result. Here we observe immune specificity frequencies drift with the antigenic drift yet it is still not adequate to withhold the virus. If the proliferation of immune cells is low then viral dominance occurs early in infection and immune cell depletion will occur quickly.

If the effector immune response is low then there is not a strong drive in the evolutionary drift of HIV frequencies. However, the killing of infected cells is reduced enabling the virus to replicate quite freely. Because the cytotoxic effect of the immune system is low the virus maintains control and depletes immune levels. It does this despite only a small drift in diversity (see Fig. 6). In contrast, if the effector immune response is high then there is a greater evolutionary drive to evade the response but less virus replication due to killing of infection. For a very large effector response our model suggests it is theoretically possible to eliminate the virus. Here there is a strong drive for drift of HIV frequencies to higher fitness but the high fitness values are still able to be controlled by the high effector response. In practice, we suggest that immune cytotoxic levels can not be driven this high, but we suggest that it is possible to eliminate disease with a high rate of clearance of infected cells.

Our model also suggests that if infection of target cells occurs rapidly then immune cell levels are depleted early and the virus soon dominates the dynamics, and does so without full antigenic drift. However, if the rate of infection is small then large antigenic diversity will result, with viral frequencies drifting to high fitness values, but the virus levels may be contained and, in fact, cleared by the immune system (shown in Fig. 7). If infection of target cells can be reduced, the virus will not have its devastating impact. In practice, drugs such as protease inhibitors have shown these effects. However, over an extended period of time, protease inhibitors will not completely eliminate the virus.

We suggest that under normal circumstances all of the factors we have presented that hinder immune control of HIV will exist in a typically infected individual. Our model suggests that building immune defence to remove HIV by strengthening T cell proliferation is not a viable strategy for preventing the onset of AIDS. However,

diminishing the rate of infection of target cells by HIV, or highly increasing clearance of infected cells can lead to HIV eradication theoretically. Our model suggests that mutation of HIV is a crucial aspect of its dominance over the immune system. We support the antigenic diversity threshold theory. Our continuous model for antigenic drift gives rise to this threshold, leading to the onset of AIDS.

References

Anderson, R.M. and May, R.M. (1979a) "Population biology of infectious diseases. ii", *Nature* **280**, 455.
Anderson, R.M. and May, R.M. (1979b) "Population biology of infectious diseases: part 1", *Nature* **280**, 361–367.
Anderson, R.M. and May, R.M. (1991) *Infectious Diseases of Humans* (Oxford University Press).
Balfe, P., Simmonds, P., Ludlam, C.A., Bishop, J.O. and Brown, A.J.L. (1990) "Concurrent evolution of human-immunodeficiency-virus type-1 in patients infected from the same source-rate of sequence change and low-frequency of inactivating mutations", *J. Virol.* **64**, 6221–6233.
Bonhoeffer, S., May, R.M., Shaw, G.M. and Nowak, M.A. (1997) "Virus dynamics and drug therapy", *Proc. Natl Acad. Sci. USA* **94**, 6971–6976.
Bonhoeffer, S., Barbour, A.D. and Boer, R.J.D. (2002) "Procedures for reliable estimation of viral fitness from time-series data", *Proc. R. Soc. Lond. B. Biol. Sci.* **1503**, 1887–1893.
Colgrove, R. and Japour, A. (1991) "A combinatorial ledge: reverse transcriptase fidelity, total body viral burden, and the implications of multiple-drug HIV therapy for the evolution of antiviral resistance", *Antivir. Res.* **41**, 45–56.
DeBoer, R.J. and Boerlijst, M.C. (1994) "Diversity and virulence threshholds in AIDS", *Proc. Natl Acad. Sci. USA* **94**, 544–548.
Domingo, E., Menendez-Arias, L. and Holland, J.J. (1997) "Rna virus fitness", *Med. Virol.* **7**, 87–96.
Goldberg, D.E. and Estabrook, G.F. (1998) "Separating the effects of number of individuals sampled and competition on species", *J. Ecol.* **86**, 983–988.
Goldsby, R.A., Kindt, T.J. and Osbourne, B.A. (2000) *Kuby Immunology*, Fourth Edition (W.H. Freeman and Company).
Holmes, E., Zhang, L.Q., Simmonds, P., Ludlam, C.A. and Brown, A.J.L. (1992) "Convergent and divergent sequence evolution in the surface envelope glycoprotein of hiv-1 within a single infected patient", *Proc. Natl Acad. Sci. USA* **89**, 4835–4839.
Kirschner, D. (1996) "Using mathematics to understand HIV immune dynamics", *Notices AMS* **43**, 191–202.
Kirschner, D.E. and Webb, G.F. (1997) "Resistance, remission and qualitative differences in HIV chemotherapy", *Emerg. Infect. Dis.* **3**, 273–283.
Lukashov, V.V., Kuiken, C.L. and Goudsmit, J. (1995) "Intrahost human immunodeficiency virus type 1 evolution is related to length of the immunocompetent period", *J. Virol.* **69**, 6911–6916.
Mansky, L.M. and Temin, H.M. (1995) "Lower *in vivo* mutation rate of human immunodeficiency Virus type 1 than that predicted from the fidelity of purified reverse transcriptase", *J. Virol.* **69**, 5087–5094.
Maree, A.F.M., Keulen, W., Boucher, C.A.B. and DeBoer, R.J. (2000) "Estimating relative fitness in viral competition experiments", *J. Virol.* **74**, 11067–11072.
May, R.M., Nowak, M.A. and Anderson, R.M. (1990) "The evolutionary dynamics of hiv-1 quasispecies and the development of immuno-deficiency disease", *AIDS* **4**, 1095–1103.

Moore, J.P. and Nara, P.L. (1991) "The role of the v3 loop of gp120 in hiv infection", *AIDS* **5**(Suppl. 2), S21–S33.

Nowak, M. (1990) "HIV mutation-rate", *Nature* **347**, 522.

Nowak, M.A. (1992) "Variability of hiv infections", *J. Theor. Biol.* **155**, 1–20.

Nowak, M.A. and May, R.M. (1990) "Mathematical biology of hiv infections: antigenic variation and diversity threshold", *Math. Biosci.* **106**, 1–21.

Nowak, M.A. and May, R.M. (1992) "Coexistence and competition in hiv infections", *J. Theor. Biol.* **159**, 329–342.

Nowak, M. and May, R.M. (1993) "Aids pathogenesis—mathematical models of hiv and siv infection", *J. Theor. Biol.* **159**, 329–342.

Nowak, M.A. and McMichael, A.J. (1995) "How hiv defeats the immune system: the process is a competition between the virus and the body's defenses'", *Sci. Am.* August 1995, 42.

Nowak, M.A., Anderson, R.M., McLean, A.R., Wolfs, T.F.W., Goudsmit, J. and May, R.M. (1991) "Antigenic diversity thresholds and the development of AIDS", *Science* **254**, 963–969.

Nowak, M.A., May, R.M., Phillips, R.E., Rowland-Jones, S., Lalloo, D.G., McAdam, S., Klenerman, P., Koppe, B., Sigmund, K., Bangham, C.R. and McMichael, A.J. (1995) "Antigenic oscillations and shifting immunodominance in hiv-1 infections", *Nature* **375**, 606–611.

Perelson, A.S. and Nelson, P.W. (1999) "Mathematical analysis of HIV-1 dynamics *in vivo*", *SIAM Rev.* **41**, 3–44.

Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M. and Ho, D.D. (1996) "HIV-1 dynamics *in vivo*: virion clearance rate, infected cell lifespan and viral generation time", *Science* **271**, 1582–1586.

Simmonds, P., Balfe, P., Peutherer, J.F., Ludlam, C.A., Bishop, J.O. and Brown, A.J.L. (1990) "Analysis of sequence diversity in hypervariable regions of the external glycoprotein of human immunodeficiency virus type 1", *J. Virol.* **64**, 5840–5850.

Wolfs, T.F.W., Zwart, G., Bakker, M., Valk, M., Kuiken, C.L. and Goudsmit, J. (1991) "Naturally occurring mutations within the hiv-1 v3 genomic rna lead to antigenic variation dependent on a single amino-acid substitution", *Virology* **185**, 195–205.

Wolinsky, S.M., Korber, B.T., Neumann, A.U., Daniels, M., Kunstman, K.J., Whetsell, A.J., Furtado, M.R., Cao, Y., Ho, D.D., Safrit, J.T. and Koup, J. (1996) "Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection", *Science* **272**, 537–542.