*J. theor. Biol.* (1992) **155**, 1–20



MINI REVIEW

# Variability of HIV Infections

MARTIN A. NOWAK

*Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, U.K.*

(*Received on* 30 *April* 1991, *Accepted on* 23 *September* 1991)

Genetic variation is the hallmark of infections with lentiviruses in general and the human immunodeficiency viruses (HIV-1, HIV-2) in particular. This article reviews both experimental evidence for the variability of the HIV genome during the course of an individual infection and mathematical models that outline the potential importance of antigenic variation as a major factor to drive disease progression. The essential idea is that the virus evades immune pressure by the continuous production of new mutants resistant to current immunological attack. This results in the accumulation of antigenic diversity during the asymptomatic period. The existence of an antigenic diversity threshold is derived from the asymmetric interaction between the virus quasispecies and the CD4 cell population: CD4 cells mount immune responses some of which are directed against specific HIV variants, but each virus strain can induce depletion of all CD4 cells and therefore impair immune responses regardless of their specificity. Therefore, increasing HIV diversity enables the virus population to escape from control by the immune system. In this context the observed genetic variability is responsible for the fact that the virus establishes a persistant infection without being cleared by the immune response and induces immunodeficiency disease after a long and variable incubation period. Mathematical biology has revealed a novel mechanism for viral pathogenesis.

## 1. Introduction: Variability of the HIV Genome

One of the striking features of the human immunodeficiency virus (HIV) is its high genetic variability. At any time point the virus population of an infected patient does not consist of a single uniform sequence, but rather of a broad distribution of different mutants (= a quasispecies). Quasispecies are very sensitive to selection pressures. One of the most obvious pressures *in vivo* is exerted by the immune response against the virus, which results in the favouring of new mutants that are not recognized by the current major immunological attack (= escape mutants). The analysis of the dynamical interaction between the HIV quasispecies and the immune system has provided a model for the development of acquired immunodeficiency disease (AIDS) after a long and variable incubation period with asymptomatic HIV infection (Nowak *et al.*, 1990; Nowak & May, 1991).

Viral replication is an error prone process. During reverse transcription of viral RNA into DNA errors occur at a rate of the order of $>10^{-4}$ per base per cycle (Preston *et al.*, 1988; Roberts *et al.*, 1988). For the HIV genome ($10^4$ bases) this implies a mutation rate of more than 1 base per genome; in other words the probability that there are no errors during reverse transcription of the whole genome is less



0022-5193/92/050001+20 $03.00/0 

than about 0·4. Recent studies also suggest the occurrence of recombination during reverse transcription. New viral RNA is produced from the DNA provirus by the RNA polymerase, a host cell enzyme that is also subject to high error rates because of a lack of any proofreading function.

These high mutation rates may represent an important factor for the virus to survive immunological attack and establish a persistent infection. In this context the existence of an optimal mutation rate that maximizes the rate of generating new resistant mutants has been proposed (Nowak, 1990); the theoretically calculated rate is in broad agreement with the observed mutation rate.

The first evidence for the genetic polymorphism of the HIV population in one infected individual was provided by restriction mapping of provirus DNA (Hahn *et al.*, 1986; Saag *et al.*, 1988; Fisher *et al.*, 1988). Subsequently a careful analysis using gene amplification by polymerase chain reaction (PCR) has confirmed the existence of genetic variation *in vivo* (Meyerhans *et al.*, 1989; Wain Hobson, 1989). This study compared sequence variation in the HIV tat gene obtained either directly from infected peripheral blood mononuclear cells (PBMC) or by cultivation of virus isolates. The virus mutant distribution was analysed at four sequential time points from a male homosexual patient over a 30-month period; approximately 20 clones were sequenced for each time point. There was extensive variability at each time point as well as a rapid turnover of the dominating variants at different time points. The quasispecies distribution found in infected PBMC was substantially different from the *in vitro* cultured virus populations. This shows that selection pressures exerted by *in vitro* cultivation can change the mutant distribution.

Sequence variation is not uniform throughout the genome. The gag and pol genes, which encode for the core proteins and the reverse transcriptase, are more conserved than the env gene that encodes for the envelope proteins gp41 and gp120. The envelope protein gp120 has a peculiar pattern of five hypervariable regions, V1 to V5. Of special interest is the hypervariable and immunodominant V3 loop (Rusche *et al.*, 1988; Goudsmit *et al.*, 1988). It has been identified as a target of neutralizing antibodies, cytotoxic and helper T lymphocyte response (Goudsmit, 1988; Palker, 1989). These immune responses are sensitive to sequence variation in this region. HIV may escape from antibody pressure by structural changes in the V3 loop due to mutations inside or outside the V3 region (Nara & Goudsmit, 1990*b*). A single amino acid substitution can restrict recognition by neutralizing antibodies (Looney *et al.*, 1989; McKeating *et al.*, 1989). The emergence of escape mutants under the selection pressure of neutralizing antibodies has also been demonstrated *in vitro* (Robert-Guroff *et al.*, 1986).

Sequence variation in the envelope protein was studied in a group of Scottish haemophiliacs who became infected during 1984 after exposure to the same batch of factor VIII (Simmonds *et al.*, 1990; Balfe *et al.*, 1990). Provirus sequences from infected PBMC were obtained during summer of 1989. There is extensive variation between sequences from the same individual and between different individuals. The number of different variants in each patient appears to be very large. Amino acid substitutions are found to accumulate at key positions that are important for immunological recognition (e.g. the antibody and CTL epitopes of V3). The ratio of

synonymous to replacement substitutions for the sequences of the V3 loop and flanking regions is 0·67. This means that a nucleotide substitution that changes the amino acid sequence is twice as likely to survive than is a silent mutation. This is the largest effect ever found in any genome and suggests strong positive selection for variation. In this context the observed sequence variation can be interpreted as an evolutionary response by HIV to evade destruction by the immune system. There is also extensive variation in the number and positions of potential N-linked glycosilation sites in the flanking regions of the immunodominant V3 loop. The V4 and V5 regions show an interesting sequence length polymorphism caused by insertions and deletions of multiples of three bases. The contribution of these mutations to the antigenicity of gp120 is uncertain. Within this region a neutralizing antibody and a cytotoxic T cell epitope have been reported (Sun *et al.*, 1989; Clerici *et al.*, 1991).

Sequence variation in the immunodominant V3 loop was also studied in six children who became infected by plasma transfusion from one donation of a single donor (Wolfs *et al.*, 1991). Provirus DNA was sequenced from both the donor and the children; two time points for each child with at least six clones. The mutant distribution is different for each child and different for the donor. In each child there seems to be a continuous change of the quasispecies distribution diverging from the initial inoculum into different directions. (Quasispecies move in a very high dimensional space.) The question remains whether this divergence represents a selective outgrowth of virus mutants present in the initial inoculum (in maybe very small amounts) or a true *de novo* mutation.

HIV is not only located in the circulating blood; it has also the ability to invade the nervous system and to establish a persistant infection in the brain. V3 domain sequences were obtained from post-mortem brain and spleen tissue of three perinatally infected children (Epstein *et al.*, 1991). Brain and spleen V3 sequences can diverge considerably within an individual patient. This tissue specific variability of the V3 loop may reflect the different selection pressures exerted by the need to maximize viral replication and to escape immune attack in different organs. Other studies have shown differences in the replication efficiencies, cell tropisms and antigenic properties of virus isolates from the brain, cerebrospinal fluid or PBMC (Koyanagi *et al.*, 1987; Cheng-Mayer *et al.*, 1988). Therefore, one may expect different mutant distributions in different cell types or organs, which are specifically adapted to local selection pressures. The brain is poorly accessible to immune surveillance and viral clearance; therefore antigenic variation is not driven by immune selection in the brain HIV population and one may expect rather homogeneous quasispecies in this organ. However, new antigenic variants may diffuse into the brain after they have been selected in other locations that are accessible to immunological pressures.

HIV-2 isolates exhibit biological and genome variability comparable to that observed for HIV-1 (Schulz *et al.*, 1990). Antigenic and genetic variation appears to be a common feature of lentivirus infections. It has been documented for visna virus (Clements *et al.*, 1980), equine infectious anaemia virus (Salinovich *et al.*, 1986) and caprine arthritis encephalitis virus (Ellis *et al.*, 1987). A recent study by Burns & Desrosiers (1991) has demonstrated extensive sequence variation *in vivo* after infection of two rhesus monkeys with molecularly cloned simian immunodeficiency virus

(SIV), the closest relative of HIV. This work also suggests selection pressure for changes in distinct variable regions of the envelope protein. Antigenic drift has been observed in chimpanzees experimentally infected with HIV-1 (Nara & Goudsmit, 1990*a*).

## 2. Variability of HIV Pathogenesis

The process of HIV infection and the development of immunodeficiency disease can be separated into three stages:

(1) Acute clinical illness during primary HIV infection occurs in 50–70% of infected patients, starts generally 2–4 weeks after infection and lasts from 1–2 weeks (Tindall & Cooper, 1991). The clinical manifestations are varied and include fever, neuropatic and dermatological symptoms. Virus can be isolated from PBMCs, cell free plasma, ceresbrospinal fluid and bone marrow cells. The high replication and widespread distribution of virus is followed by strong immunological responses, which results in a decrease of viral antigens to almost undetectable levels and a resolution of clinical symptoms.

(2) The second, chronic, phase (8–10 years on average) is characterized by low levels of HIV expression and only small pathological changes. Patients are generally asymptomatic. A study of 112 HIV infected haemophiliacs over 8 years shows a constant (linear) decline in the CD4 lymphocyte count during this phase and a strong correlation between the CD4 count and the cumulative probability to develop AIDS (Phillips *et al.*, 1989). Two other studies, however, have found rather constant CD4 levels and rapid decline only in the 1 or 2 years before progression to AIDS (Eyster *et al.*, 1987; Kaplan, 1988).

(3) The final phase is characterized by the development of ARC and AIDS. CD4 counts are low. Virus levels—both in terms of infected PBMC and free virus in the plasma—are about 100 times larger than in the asymptomatic stage (Ho *et al.*, 1989; Coombs *et al.*, 1989). The clinical symptoms are varied and characterized by opportunistic infections. The life expectation of AIDS patients is about 1 year.

What controls the three phases is a central but unanswered question. It is not understood why some people develop AIDS within 1 to 2 years after HIV infection, while others are still asymptomatic after 15 years. There is extensive variability in the rate of progression to disease.

Many different markers that correlate with disease progression have been discussed. Among these are high antigen titre, low CD4 cell count, low p24 antibody titre and high neopterin levels. In a recent study of 275 HIV-1 infected patients, a high rate of progression to AIDS was found to be associated with core antibody negativity and high antigenaemia (deWolf *et al.*, 1989). Another study of 238 HIV infected homosexual men revealed that patients with low p24 antibody titre and high levels of neopterin had the highest rate of disease progression (Sheppard *et al.*, 1991). The p24 antibody titres were extremely variable among cohort participants but relatively stable during the time of follow up (48 months). It is interesting to note that the level of p24 antibodies correlates with disease progression, while neutralizing antibody titres do not (Ascher & Sheppard, personal information). This paradox

could be explained if the anti p24 level reflects the patient's general ability to mount an immune response to HIV, but neutralizing antibody titres are too variable because of the variation in the V3 loop epitopes where some variants may induce more response than others. The neopterin level determines general immune activation (Fuchs *et al.*, 1988) as induced by HIV antigens or other pathogens.

In the following I would like to list some additional aspects that seem to be important for an understanding of the natural history of HIV infections.

There is continuous virus replication throughout the course of infection. This is supported (1) by the findings that all patients across the clinical spectrum have infectious free virus circulating in their plasma (if not in the early phase of zidovudine treatment) and (2) the observations that virus nucleotide sequences (amplified with PCR techniques) change continually over time of infection.

Most of the cell associated virus in the blood is found within CD4+ T cells (Schnittman *et al.*, 1989). But there is no proof that this is the major reservoir of infected cells in the body, nor is it clear which cells produce the free virus found in the plasma. It is, however, clearly established that HIV infection leads to a depletion of CD4 cells and impairment of CD4 cell function. This can happen by direct killing of infected cells or immune responses against gp120 bound to the CD4 receptor of uninfected cells or auto-immune responses induced by HIV. It has also been suggested that gp120 delivers a strong unspecific activation signal to CD4 positive cells that may be responsible for the impairment of immune function rather than direct killing of CD4 cells (Ascher & Sheppard, 1990). For the following mathematical model it is only necessary that HIV has evolved some means of destroying CD4 cell function.

Cytotoxic T lymphocytes (CTLs) appear at the end of the first acute phase, when virus titres fall, and remain strongly activated throughout the chronic phase of infection. CTL responses are also directed against rather conserved regions of the gag or pol gene products. These and other findings suggest their important role in limiting viral replication (Takahashi *et al.*, 1989; Phillips *et al.*, 1991).

Sera of infected patients contain antibodies capable of neutralizing HIV infectivity *in vitro* (Weiss *et al.*, 1985; Robert-Guroff *et al.*, 1985). In general these neutralizing antibodies appear to be rather strain specific and the high variability of the V3 loop makes this assumption plausible. However, the experiments of Weiss *et al.* (1986) have shown some cross reactivity of neutralizing antibodies between different strains of HIV. Within the V3 loop there are conserved sequence and structural elements and some antibodies directed at the V3 loop can cross neutralize some divergent strains of HIV (LaRosa *et al.*, 1990; Javaherian *et al.*, 1990). Besides the V3 loop there are other (more conserved) epitopes that can serve as targets for neutralizing antibodies (Michel *et al.*, 1988; Sun *et al.*, 1989; Dalgleish *et al.*, 1989). It has also been reported that some strains of HIV-2 can be cross neutralized by some HIV-1 sera (Schulz *et al.*, 1990). These findings suggest that neutralizing antibodies can be both strain specific and cross reactive, but the relative importance of these remains to be determined.

The development of HIV-related opportunistic infections is related to the level of CD4+ T cells. These diseases often represent the reactivation of quiescent infections

that were held in check by the patient's immune system before infection with HIV. Most of these infections are also found in people with impairment of immune function for other reasons. Therefore, opportunistic infections might be considered as a mere consequence of the depletion of CD4+ T cells, which is induced by HIV. But the interaction between HIV and other pathogens is more complex (because synergistic). HIV may be capable of (1) destroying immune memory and (2) increasing the level of chronic infections (by its immunosupressive function). Accumulation of foreign pathogens, on the other hand, may facilitate the activation of HIV infected T cells and thereby stimulate the virus to replicate. This could increase the overall rate of virus replication and CD4 cell destruction (McLean & Kirkwood, 1990; McLean & Nowak, 1991).

Different isolates of HIV may not only vary in their antigenic properties but also in their growth rates and ability to induce syncytia formation *in vitro*. More virulent, faster growing strains seem to evolve in the latter phase of the infection (Asjö *et al.*, 1986; Cheng-Mayer *et al.*, 1988; Tersmette *et al.*, 1989). This phenomenon may be a cause or a consequence of immunodeficiency.

## 3. Mathematical Models for the Genetic Variation of HIV During the Course of an Infection

The perspective of this section is to present mathematical models of increasing complexity for the dynamics associated with antigenic variation. First the simplest model is analysed and its biological implications are discussed. Then more complexity is added. This helps to understand the biological meaning of each assumption separately.

### 3.1. ANTIGENIC DRIFT

The replication dynamics of a single virus strain and its specific immune response can be described by the following equations

$$\frac{dv_i}{dt} = v_i(r - px_i) \tag{1}$$

$$\frac{dx_i}{dt} = kv_i. \tag{2}$$

Let $v_i$ and $x_i$ denote the density of the virus strain $i$ and of immune cells ("immune agents") specific to this strain. In the absence of immunity virus particles replicate with rate $r$. The term $pv_ix_i$ specifies killing of virus due to specific immune reactions directed at strain $i$. The specific immune cells, $x_i$, are produced at a constant rate proportional to the virus density, $v_i$.

The trajectories of this differential equation have the form

$$v_i(t) = v_i(0) + x_i(t)\left[r - \frac{p}{2}x_i(t)\right]\Big/k. \tag{3}$$

This holds for the initial conditions $v_i(0)$ and $x_i(0)=0$—i.e. we assume that the strain specific immune response to a newly arising strain is very small (zero) at the time when this strain appears. Each individual strain, $v_i$, increases as long as $x_i<r/p$. After the immune response has overcome this level the virus density decreases and goes to zero.

The total virus population can only survive if new resistant mutants are produced that escape from the specific immune response. The rate of production of escape mutants is assumed to be proportional to the replication rate of the virus population. The stochastic appearance of new mutants is represented in the following way: the probability that an escape mutant is produced in the time interval $[t, t+\mathrm{d}t]$ is given by $Pv(t)\,\mathrm{d}t$, where $P$ is a constant including the virus replication and mutation rate. The variable $v(t)=\sum v_i(t)$ denotes the total virus population at time $t$.

The diversification rate of the virus population can be characterized by the total number of escape mutants that are produced from any one strain. We have [using eqns (2) and (3)]

$$R=P\int_0^\infty v_i(t)\,\mathrm{d}t=Px_i(\infty)/k=P\frac{2r}{pk}. \tag{4}$$

Here again we have assumed that $x_i(0)$, $v_i(0)$ and $v(\infty)$ are very small (zero). The asymptotic behaviour of the total virus density, $v$, depends on the magnitude of the diversification potential $R$. If $R>1$ then $v$ tends to infinity; if $R<1$ then $v$ tends to zero. This means that each virus variant has to generate at least one escape variant before being suppressed by the immune system.

### 3.2. KILLING OF IMMUNE CELLS

In this section we include the possibility that the virus can impair immune function.

$$\frac{\mathrm{d}v_i}{\mathrm{d}t}=v_i(r-px_i) \qquad i=1,\ldots,n \tag{5}$$

$$\frac{\mathrm{d}x_i}{\mathrm{d}t}=kv_i-uvx_i \qquad i=1,\ldots,n. \tag{6}$$

The total number of strains, $n$, is increasing as new variants are produced. The term $uvx_i$ represents "killing" of immune cells induced by HIV. This term not only includes direct cell killing, but also any other adverse effects of HIV on the immune system. Note that specific immune cells can only be activated by specific viral strains ($v_i$) and are only reactive against their specific strain of HIV, but can be killed by all strains ($v$). This has the consequence that increasing antigenic diversity is advantageous for the virus population and that there exists a threshold in antigenic diversity that can be controlled by the immune system.

The immune system can control strain $i$ if $\dot{v}_i<0$, which is to say if $r-px_i<0$. The immune system can thus control all individual strains only if this holds for all $i$ ($i=1,\ldots,n$), which implies the restriction that

$$n<n_c=\frac{px}{r}. \tag{7}$$

The total density of specific immune cells, the sum $x=\sum x_i$, converges to $\hat{x}=k/u$ and therefore

$$n_c=pk/ru. \tag{8}$$

This establishes the potential existence of a viral antigenic diversity threshold. There exists an upper limit, $n_c$, of different strains that can be suppressed simultaneously by the immune system. Once the number of strains exceeds the threshold value the total virus density will eventually escape from control by the immune system. This concept has implications on disease progression in HIV infections. While the immune system may be able to control a virus population whose diversity is below the threshold, it fails to control a population above the diversity threshold. The long asymptomatic period may arise from the slow increase in antigenic diversity.

An appropriate measure of antigenic diversity is given by the Simpson index. The growth of the total virus population, $v=\sum v_i$, is given by

$$\dot{v}=v(r-pE) \tag{9}$$

where $E=\sum x_i v_i/v$ denotes the efficiency of the specific immune responses. Approximating $x_i$ by its "steady-state" value $\bar{x}_i=kv_i/uv$ leads to

$$\dot{v}=v\left(r-p\frac{k}{u}D\right) \tag{10}$$

where $D=\sum (v_i/v)^2$ is the ecologists' Simpson index, an inverse measure for diversity. Note that $\dot{v}>0$ is equivalent to $1/D>n_c$. This means that the inverse of the Simpson index exceeding the diversity threshold is a necessary and sufficient condition for virus growth. In other words, the antigenic diversity must be larger than $n_c$ strains in uniform distribution in order to obtain $\dot{v}>0$. In this context the Simpson index is a functional measure for antigenic diversity in this model.

Figure 1 illustrates the dynamical behaviour of this model. The infection is initially started with one strain, new strains are generated at the rate $Pv(t)\,\mathrm{d}t$. The total virus density is fluctuating around low levels for a long time before increasing dramatically. The diversity threshold in this model is $n_c=10$. The total virus density increases when the inverse of the Simpson index breaches this value. In the final phase, when the threshold is exceeded and the virus "escapes" control by the immune system, strains that earlier were suppressed can rise again. This happens as follows. After the immune system has downregulated a strain ($v_i$ very close to zero) there is no further induction of specific immune cells directed towards that particular strain. But these immune cells are steadily removed by the other viruses present in the system. After some time the frequency $x_i$ has dropped below the threshold ($x_i<r/p$) and $\dot{v}_i$ becomes positive again. This point can be important for the accumulation of disease in the final phase of infection.

The generation of new resistant mutants by replication errors is a stochastic event. Therefore, it can happen that the infection is cleared during the period of low virus concentration. This happens if the mutation rate to produce new resistant mutants is too low. On average we require that each virus strain has to produce at least one new escape mutant before being suppressed by the immune system.

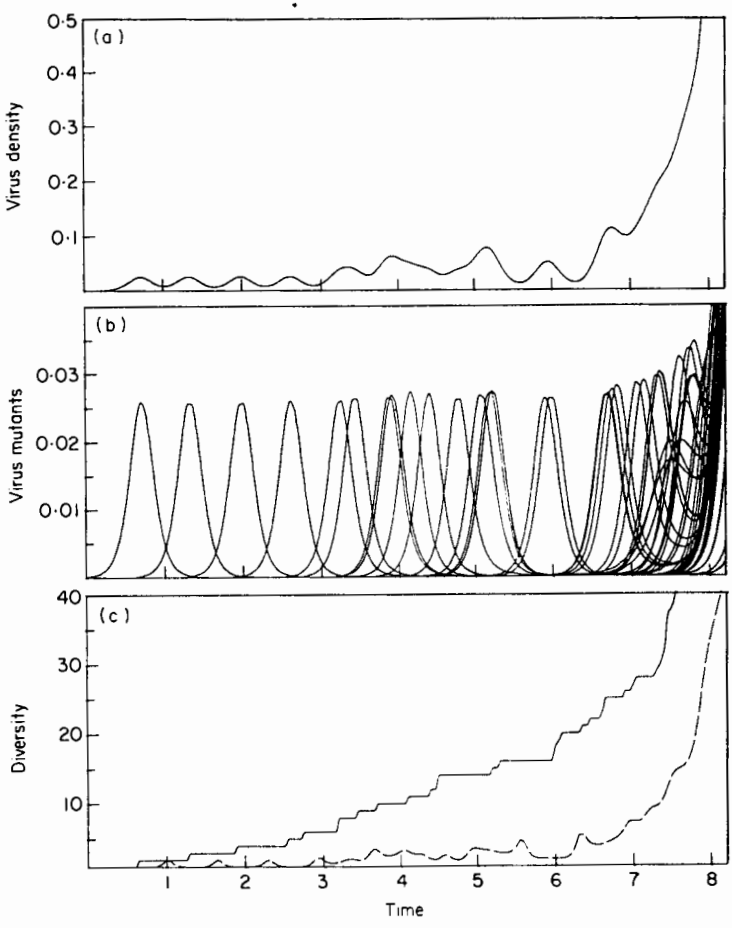



FIG. 1. A numerical simulation of the simplest model that includes antigenic variation of the virus population, the stimulation of specific immune responses and impairment of immune function as described by the eqns (5) and (6). (a) Total virus population size (virus density in arbitrary units); (b) abundance (density) of the individual HIV variants; there is no cut-off parameter, the densities of some strains are downregulated to very low levels—invisible in the plot—and may reappear later as immonological pressure decreases. Interestingly the reappearance of early strains later on in infection has been observed in several HIV infected patients. (c) The total number of strains (continuous line) and the inverse of the Simpson index (broken line) as a measure for population diversity. The total virus density is rather low for a long time before the virus population escapes from control by the immune system. There is no initial peak of viraemia, because this model does not include cross reactive immune responses. Therefore, the immune system does not learn to suppress subsequent strains more readily, all the different virus strains grow to the same level of abundance. The population diversity increases during the time of infection. (The effective diversity as measured by the Simpson index increases at a slower rate than the total number of strains.) This is a consequence of mutation and selection: replication errors during virus reproduction generate new escape mutants that may escape from the current immune response and are selected by the immune pressure. The breaching of the diversity threshold results in the final increase of viral replication. The parameter values are: $r=1$, $p=20$, $u=1$, $k=1$. The parameters have the dimension of reciprocal time (assuming dimensionless variables for virus and immune cell densities).

However, the complete extinction of particular strains by the immune system can become difficult if some "latently" infected reservoir cells exist, which have a long lifetime and are not easily accessible to immunological attack. This can be included in the model by

$$\frac{dv_i}{dt} = v_i(r - px_i - \alpha) + \beta w_i \tag{11}$$

$$\frac{dw_i}{dt} = \alpha v_i - \beta w_i - \gamma w_i. \tag{12}$$

Virus in the "reservoir cells" is denoted by $w_i$ and is produced at a rate proportional to the replicating virus, $\alpha v_i$. Reactivation occurs at a rate, $\beta w_i$ and the average lifetime of these cells is given by $1/\gamma$. The effect of these assumptions is an increased ability to maintain and accumulate diversity.

### 3.3. CROSS REACTIVE IMMUNE RESPONSES

Further realism can be added by including cross reactive immune responses that are directed against several strains simultaneously. In the simplest case we may think of subdividing the immune response to HIV into strain specific and cross reactive responses. For example neutralizing antibodies directed at the hypervariable V3 loop may be included in the class of strain specific responses, whereas CTLs directed at conserved regions in the gag proteins are rather cross reactive against different strains of HIV. This leads to the equations

$$\frac{dv_i}{dt} = v_i(r - sz - px_i) \qquad i = 1, \ldots, n \tag{13}$$

$$\frac{dx_i}{dt} = kv_i - uvx_i \qquad i = 1, \ldots, n \tag{14}$$

$$\frac{dz}{dt} = k'v - uvz. \tag{15}$$

To keep the mathematics simple we model the whole range of more or less cross reactive immune responses by taking into account only the two extreme cases. The terms $sz$ and $px_i$ represent unspecific and specific immune responses, respectively. These can be justified as follows. Let us assume that each virus strain $i$ induces the production of certain immune agents, a fraction of which is specifically directed only against that particular mutant strain (e.g. via the immunodominant loop), while the other fraction is directed against more conserved sites (e.g. the gag or pol gene products or conserved regions within the env protein) and hence is able to react with

several different mutant strains. $z$ is the number of immune cells activated against conserved regions, $x_i$ is the number of immune cells specifically against a particular strain and $x=\sum x_i$ denotes the total density of these "specific" immune cells. The relative magnitudes of the parameters $s$ and $p$ allow the regulation of the "cross reactivity" of the immune response to HIV.

The killing of immune cells by viral mechanisms is denoted by the terms $uvx_i$ and $uvz$; The densities of specific and unspecific immune cells converge towards the levels, $\hat{x}=k/u$ and $\hat{z}=k'/u$.

There are three different parameter regions according to the magnitude of specific and unspecific immune responses.

(1) The unspecific (= cross reactive) immune response is by itself able to suppress viral growth (i.e. $r<s\hat{z}$). There will be a rise in viral abundance following the initial infection, but once the unspecific immune response has been mounted the initial strain and all subsequently-evolved ones will be suppressed by this generalized response. In this case antigenic variation will never overwhelm the immune response. This may be the case in SIV infected African Green Monkeys and HIV infected chimpanzees. In both cases there is no development of immunodeficiency disease.

(2) The replicative capacity of a single strain can outrun both the specific and unspecific immune responses (i.e. $r>s\hat{z}+\hat{x}$). The immune system is not able to cope with any single strain. The initial viraemia is not suppressed by the immune response, there is no incubation period, no delay until the onset of disease. Antigenic variation is not necessary for the virus to escape from immune control. This seems to be the case for some acutely lethal variants of SIV (Fultz *et al.*, 1989; Dewhurst *et al.*, 1990).

(3) Between these two extremes lies the interesting region of dynamical behaviour, with its viral diversity threshold. This situation corresponds to individual viral strains having replication rates that can outrun the unspecific immune response, but not the combined effect of unspecific and specific immune responses (i.e. $p\hat{x}>r-s\hat{z}>0$). Only the continuous generation of new resistant strains enables the virus population to survive immunological attack. In this parameter region we observe the diversity threshold. The critical number of strains that can be suppressed by the immune system is obtained as

$$n_c=\frac{pk}{ru-sk'}. \tag{16}$$

The unspecific immune response in our model is responsible for the fact that the initial strains are growing to higher levels than the following escape mutants. Therefore, we obtain a peak of initial viraemia followed by a period with low virus abundance. Roughly speaking, the higher the effect of the unspecific response the higher the difference between the initial peak and the average virus density in the silent phase. A stronger unspecific (cross reactive) immune response is correlated with lower viral abundance in the incubation period and with an increased length of this period.

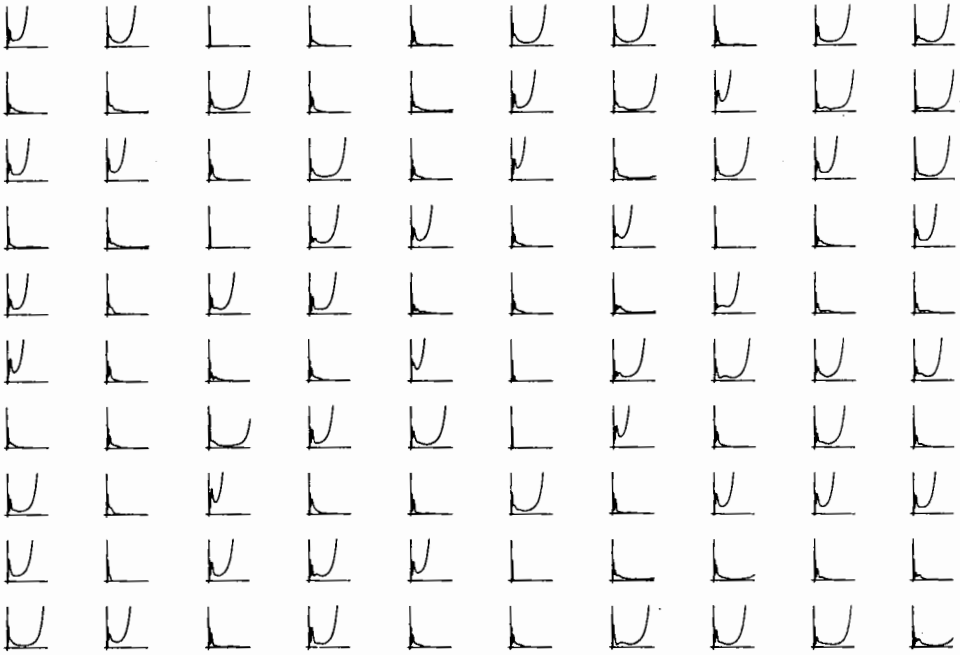


FIG. 2. Simulation of 100 individual HIV infections with the same parameter values. The only difference is the seed of the random number generator. It is of interest that the outcome of the infection and the rate of progression towards AIDS is extremely variable, although each patient has the same immunological parameters and the virus populations have the same replication and mutation rates. Some patients develop AIDS shortly after the initial viraemia, others are still asymptomatic after a long time and in some circumstances the immune response has wiped out the virus. The lengths of the incubation period depends on the stochastic emergence of new virus variants. The model defined by the differential eqns (13–15) can explain the long and variable asymptomatic period. Parameter values: $r=1$, $p=2$, $s=0{\cdot}93$, $u=3$, $k=k'=3$. Each plot shows virus population size (y-axis) against time since infection (x-axis).

Figure 2 illustrates the extensive variability in the rate of disease progression that can be generated by this model even with exactly the same set of parameters. The only difference in these pictures is the seed of the random number generator that governs the stochastic emergence of escape mutants. The simple process described by eqns (13–15) is able to explain the observed variability of rates and patterns of disease progression in HIV infections.

### 3.4. POPULATION DYNAMICS OF CD4 CELLS; VIRUS STRAINS WITH DIFFERENT REPLICATION RATES

The final model has four variables: $v_i$, $y$, $x_i$ and $z$ denoting, respectively, the densities of virus strain $i$, total CD4 cells, CD4 cells specific to strain $i$ and CD4 cells

that mount cross reactive responses to all strains. We obtain

$$\frac{dv_i}{dt} = v_i(r_i' + r_i y - s_i z - p_i x_i) \qquad i = 1, \ldots, n \tag{17}$$

$$\frac{dy}{dt} = K - dy - uvy \tag{18}$$

$$\frac{dx_i}{dt} = kv_i y - uvx_i \qquad i = 1, \ldots, n \tag{19}$$

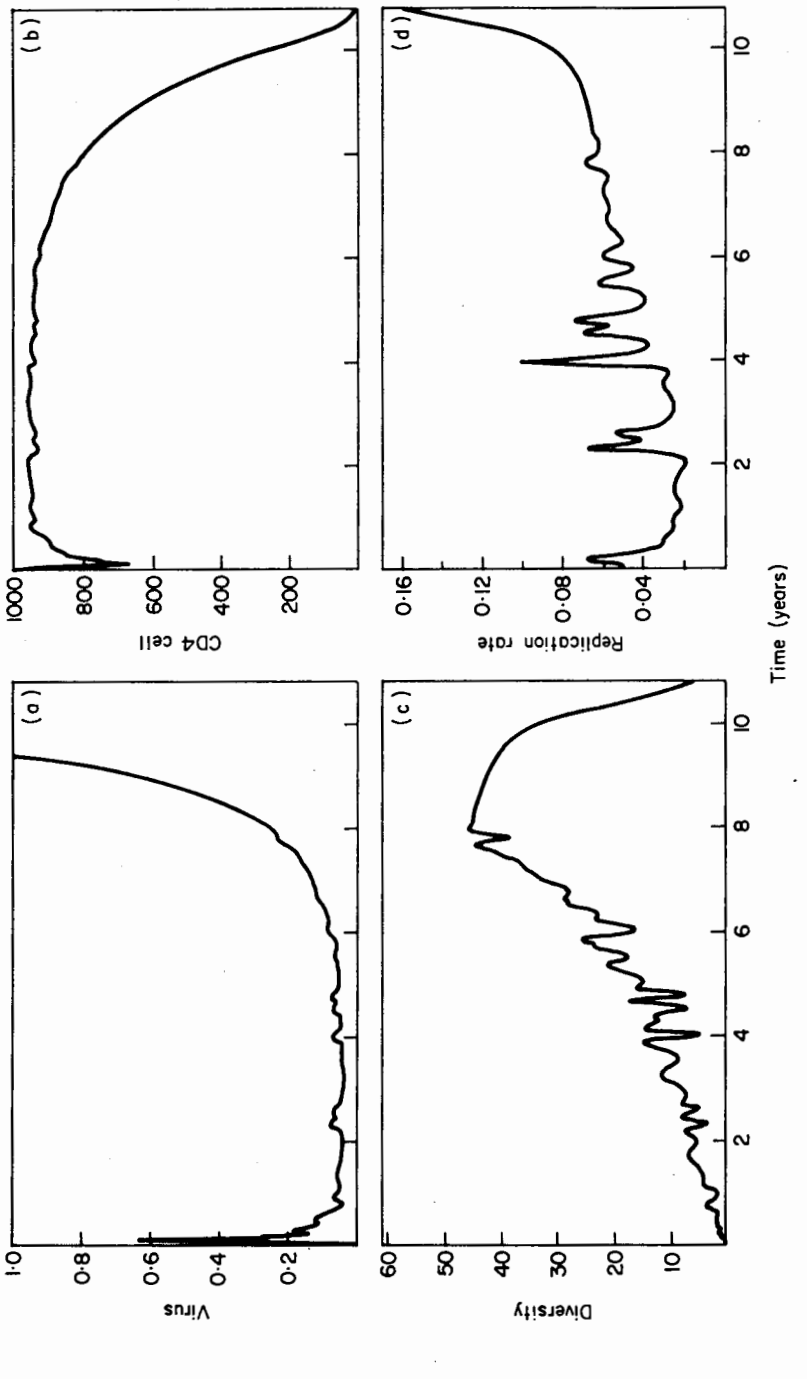
$$\frac{dz}{dt} = k'vy - uvz. \tag{20}$$

Virus replication consists of a term proportional to the density of CD4 cells, $r_i y$, and a constant background replication rate, $r_i'$, to denote replication of HIV in cells other than of the CD4 type (i.e. macrophages). Different virus variants have different replication rates and immunological parameters, $r_i$, $r_i'$, $s_i$, and $p_i$. CD4 cells are recruited (from the thymus) at a constant rate, $K$; removed at a constant rate, $dy$; and killed by the virus, $uvy$. The production of immune cells specific to HIV antigens is proportional to the total number of CD4 cells, according to the assumption that a certain (constant) fraction of all CD4 cells can serve as specific precursors for CD4 cells activated against HIV.

Figure 3 illustrates the dynamical behaviour of the eqns (17–20). The total virus density shows the typical picture with initial viraemia, a long period with low virus abundance and a final increase. The CD4 cell count is fluctuating and slightly decreasing during the asymptomatic phase. There is an oscillatory increase in population diversity, a peak when the virus population breaches the diversity threshold and a decline in the final phase when the impaired immune system cannot provide selection pressure strong enough to increase antigenic diversity further. The average replication rate of the virus population is slightly increasing during the intermediate phase; osillations correspond to the evolution of fast replicating strains that are subsequently suppressed by the immune system. In the final phase fast replicating strains dominate the population.

Figure 4 shows the individual virus mutants for the same simulation. Initially we observe high levels of virus, but the immune response is slowly increasing and finally manages to suppress the most abundant strain. But in the meanwhile new mutants have been created. The mini-outbreaks of higher virus levels correspond to the occurrence of newly-arisen neutralization-resistant mutants. After this long period where the virus is downregulated by the immune responses we observe a final increase in viral abundance. During the course of infection the number of mutant strains continually increases and finally exceeds the threshold value $n_c$ which finally results in the continuous rise of in viral abundance.

According to this model the evolution of faster replicating strains during infection is a consequence of the impairment of immune function. In the absence of HIV specific immunity there is selection for the fastest growing strains. There is no specific


(a)
Virus
1·0
0·8
0·6
0·4
0·2
(b)
CD4 cell
1000
800
600
400
200
(c)
Diversity
60
50
40
30
20
10
(d)
Replication rate
0·16
0·12
0·08
0·04
2
4
6
8
10
Time (years)

HIV mutant that causes AIDS. These seem to be supported by the isolation of fast replicating strains both during primary infection (before the onset of immunological pressure) and in the final phase (after sufficient impairment of immune function). The analysis of donor recipient pairs has not revealed any association between the presence of fast replicating, virulent strains in the donor and the rate of disease progression in the recipient (M. Busch, personal information).

### 3.5. FURTHER DEVELOPMENT OF THE MODEL

A more detailed description of the quasispecies structure of the virus population can be the focus of further mathematical research. The quasispecies concept was designed by Eigen (1971) and Eigen & Schuster (1979) to describe the erroneous replication of RNA molecules in the context of the origin of life or virus populations. A quasispecies is defined as the equilibrium distribution of mutants formed by a specific mutation selection process. HIV populations have been called quasispecies to indicate their genetic heterogeneity. However, the HIV quasispecies in an infected individual may never reach an equilibrium mutant distribution because of the rapidly changing selection pressures exerted by the specific immune responses. HIV quasispecies are adapting on an ever changing fitness landscape; the most successful and most abundant variant will become exposed to the strongest immune response. The discussed mathematical models are of phenotypic character. Different HIV variants are just labelled 1, 2, . . . and are defined to evoke different strain specific immune responses. There is no defined relation between different variants; no genetic distance. A subsequent model can be built upon a genotypic description of the virus population according to the equations

$$\frac{dv_i}{dt} = \sum_{j=1}^{n} r_j Q_{ij} v_j - s v_i - p v_i x_i \tag{21}$$

$$\frac{dx_i}{dt} = k v_i - u v x_i \tag{22}$$

FIG. 3. The dynamical properties of the model described by the eqns (17–20) that include CD4 cell dynamics, strain specific and cross reactive immune responses and virus strains with different replication capacities. (a) The total virus density shows an initial peak of viraemia followed by a long period with low virus abundance and a final increase. The y-axis shows virus population size (arbitrary units, e.g. number of virions per volume blood or so). (b) The CD4 cell population size is slightly decreasing during the asymptomatic phase and rapidly decreasing as the virus population replicates to high levels. The final CD4 cell depletion is a consequence of the high virus abundance. The y-axis represents CD4 cell counts (in arbitrary units, e.g. number of cells per volume of blood). (c) The population diversity, defined by the inverse of the Simpson index, displays a one humped pattern with a maximum just before the virus escapes control by the immune system. The antigenic diversity increases as long as the immune response selects for escape mutants. The diversity may decrease in AIDS patients. (d) The average replication rate of the virus population increases with progression towards disease. In the absence of a proper immune response to HIV in the final phase of infection the fastest replicating strains are selected. Parameter values: $K = 100$, $d = 1$, $k = k' = 0{\cdot}1$, $u = 1$, $r_i' = 3r_i$, $s_i = 9{\cdot}5r_i$, $p_i = 20r_i$; $r_i$ was taken from an exponential distribution with parameter 0·05. (All the parameter values have the dimension of reciprocal time- unit; 1/year.)

where $v_i$ and $x_i$ denote, respectively, the densities of virus mutant $i$ and CD4 cells specific to mutant $i$. The mutation matrix $Q_{ij}$ defines the probabilities to obtain variant $i$ from (erroneous) replication of variant $j$. This mutation matrix defines a genetic distance between different mutants. The eqns (21) and (22) describe the deterministic dynamics of an HIV quasispecies in the presence of specific immune responses.

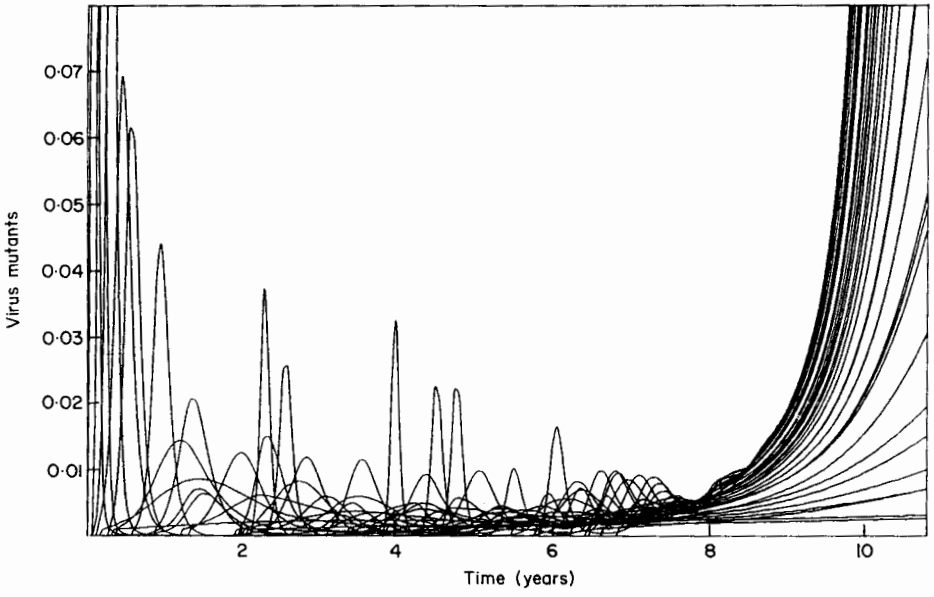



FIG. 4. Co-evolution and co-existence of many different virus strains during the time of infection. Initially the strains grow to high levels which may cause the clinical symptoms observed during primary HIV infection. The subsequently emerging escape mutants are suppressed at a faster rate, because of the action of cross reactive immune responses. Different virus strains grow to different levels according to their growth rates. The accumulation of viral diversity breaches the threshold after about 7 years in this simulation. In the final phase the fastest growing strains dominate the virus population. The y-axis indicates the relative concentration of different virus mutants.

## 4. Discussion

The theory presented here describes virus mutation and variability as essential for survival of the virus population in the presence of immune responses, and for the subsequent development of immunodeficiency disease. Each HIV infection represents a (unique) evolution of a series of different HIV quasispecies. This evolutionary process leads eventually to the development of AIDS after a long and variable incubation period, during which the balance between viral cytopathicity and the immune response is slowly shifted by increasing viral diversity. During the asymptomatic period the immune system itself drives diversification of the virus population

by continuous selection for new escape mutants. The accumulation of diversity is the cause of immunodeficiency disease. As the virus population breaches the diversity threshold the immune system becomes unable to control the virus. The consequence is extensive HIV replication, increasing virus load and rapidly decreasing CD4 cell numbers as ARC or AIDS are developed. Finally when severe immunodeficiency is established there is no longer a strong immune response to HIV which would drive diversification. The fastest replicating strains will outgrow other variants. Antigenic diversity may decrease in AIDS patients.

The concept of the diversity threshold does not arise because the immune system can only mount antibody or CTL responses against a limited variety of antigens. The diversity threshold is a consequence of the asymmetric interaction between HIV strains and the cells of the immune system. Each HIV strain can kill all CD4 cells and therefore impair immune responses regardless of their specificity, but each of these specific immune responses is only directed against a specific strain (that carries the specific antigen). As the number of strains increases this asymmetric killing is shifted in the direction of virus growth and CD4 cell killing.

Experimental support for the suggested pattern of population diversity comes from a study of two male homosexual patients who were followed since their infection in 1985 (Nowak *et al.*, 1991). Variation of the V3 domain was examined via cloned sequences derived from viral/genomic RNA (isolated from serum and amplified with PCR). The first sequence sample was taken during the first peak of viraemia. In both patients the genetic diversity is extremely low at this time. In one patient all V3 loops were identical (sample size: 11), in the other patient there were six identical V3 loops out of seven samples. Subsequently the diversity, as measured by the Simpson index of different V3 loops and the mean Hamming distance of the quasispecies distribution, increased during the asymptomatic phase in both patients. One patient developed AIDS after 55 months and received AZT treatment. This was followed by a decline in viral diversity.

The presented co-evolutionary process of the HIV quasispecies and the human immune system during the course of an individual infection is unique with respect to the time scale. The emergence of new escape mutants can occur within weeks and the immune system may require a similar time to respond. We are confronted with the complex interaction between two highly variable biological structures: (1) the HIV quasispecies under the pressure of the immune response and (2) the immune system exposed to mutating HIV antigens.

The intrinsic stochasticity of this process can make it difficult to find markers that clearly correlate with disease progression. For the same set of parameters the suggested mechanism can generate an extensive variability in the rates of disease progression. The length of the asymptomatic period depends on the magnitude of the diversity threshold and on the diversification rate. A strong immune response can tolerate a higher viral diversity.

This article has reviewed the experimental evidence available to date for the genetic variation of HIV, together with theoretical models that outline the importance of the genetic variation for the development of AIDS following a long and variable period of asymptomatic infection with HIV. The agreement between a number of

model predictions and experimental observations is encouraging at the moment. These are (1) an early peak in viraemia (primary HIV-1 infection) following infection; (2) a long and variable incubation period with low viral abundance for much of the period; (3) an increase of viral density in the final phase of infection as the failing immune system fails to control viral population growth (the appearance of the disease AIDS); (4) co-evolution and co-existence of many viral mutants in one infected person; (5) increasing population diversity during the asymptomatic phase and (6) a positive correlation between the presence of high replicative viral strains and the rate of progression to disease (AIDS).

The theoretical analysis has clearly established that antigenic variation of HIV does not only enable the virus population to remain persistant in the presence of a strong immune response but can also be responsible for disease progression. This is a new idea. While experimental evidence is accumulating that genetic (and antigenic) diversity increases between seroconversion and development of AIDS, the relative importance of this effect compared to other effects that may drive disease progression has to be established. One could think of experiments (e.g. in the SIV/macaque model) where repeated infection with diverse virus strains increases the rate of disease progression. Another possibility would be to find a correlation between antigenic diversity (or at least genetic diversity in immunodominant epitopes) and disease stage or rate of disease progression in a significant number of patients.

I would like to thank Angela McLean, Bob May, Roy Anderson and two referees for discussion and helpful comments.

## REFERENCES

Ascher, M. S. & Sheppard, H. W. (1990). AIDS as immune system activation. *J. AIDS* **3**, 177–191.

Asjö, B. & Fenyö, E. M. (1986). Replication properties of HIV from patients with varying severity of HIV infection. *Lancet* ii, 660.

Balfe, P., Simmonds, P., Ludlam, C. A., Bishop, J. O. & Leigh Brown, A. J. (1990). Concurrent evolution of HIV-1 in patients infected from the same source. *J. Virol.* **64**, 6221.

Burns, D. P. W. & Desrosiers, R. C. (1991). Selection of genetic variants of SIV in persistently infected rhesus monkeys. *J. Virol.*, in press.

Cheng-Mayer, C., Seto, D., Tateno, M. & Levy, J. A. (1988). Biological features of HIV-1 that correlate with variation in the host. *Science* **240**, 80–82.

Clements, J. E., Pedersen, F. S., Narayan, O. & Haseltine, W. A. (1980). Genomic changes associated with antigenic variation of visna virus during persistant infection. *PNAS* **77**, 4454–4458.

Clerici, M., Lucey, D. R., Zajac, R. A., Boswell, R. N., Gebel, H. M., Takahashi, H., Berzofsky, J. A. & Shearer, G. M. (1991). Detection of cytotoxic T lymphocytes specific for synthetic peptides of gp160 in HIV seropositive individuals. *J. Immunol.* **146**, 2214–2219.

Coombs, R. W., Collier, A. C. & Allain, J. P. (1989). Plasma viremia in HIV infection. *N. Engl. J. Med.* **321**, 1626–1631.

Dalgleish, A. G., Chanh, T. C., Kennedy, R. C., Kanada, P., Clapham, P. R. & Weiss, R. A. (1988). Neutralization of diverse HIV-1 strains by monoclonal antibodies raised against a gp41 synthetic peptide. *Virology* **165**, 209–215.

Dewhurst, S., Embretson, J. E., Anderson, D. C., Mullins, J. I. & Fultz, P. N. (1990). Sequence analysis and acute pathogenicity of molecularly cloned SIV. *Nature, Lond.* **345**, 636–639.

De Wolf, Lange, J., Houweling, J., Mulder, J., Beemster, J., Schellekens, P. T., Coutinho, R. A., Noordaa, J. & Goudsmit, J. (1989). Appearance of predictors of disease progression in relation to the development of AIDS. *AIDS* **3**, 563–569.

Eigen, M. (1971). Self organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **58**, 465–523.

EIGEN, M. & SCHUSTER, P. (1979). *The Hypercycle*. Berlin: Springer-Verlag.

ELLIS, T. M., WILCOX, G. E. & ROBINSON, W. F. (1987). Antigenic variation of CAEV during persistent infection of goats. *J. gen. Virol.* **68**, 3145–3152.

EPSTEIN, L. G., KUIKEN, C., BLUMBERG, B. M., HARTMAN, S., SHARER, L. R., CLEMENT, M. & GOUDSMIT, J. (1991). HIV-1 V3 domain variation in brain and spleen of children with AIDS: tissue-specific evolution within host-determined quasispecies. *Virology*, in press.

EYSTER, M. E., GAIL, M. H., BALLARD, J. O., AL-MONDHIRY, H. & GOEDERT, J. J. (1987). Natural history of HIV infections in haemophiliacs. *Ann. Intern. Med.* **107**, 1–6.

FISHER, A. G., ENSOLI, B., LOONEY, D., ROSE, A., GALLO, R. C., SAAG, M. S., SHAW, G. M., HAHN, B. H. & WONG-STAAL, F. (1988). Biologically diverse molecular variants within a single HIV-1 isolate. *Nature, Lond.* **334**, 444–447.

FUCHS, D. (1988). Neopterin as a marker for activated cell-mediated immunity: application in HIV infection. *Immunol. Today* **9**, 150–155.

FULTZ, P. N., MCCLURE, H. M., ANDERSON, D. C. & SWITZER, W. M. (1989). Identification and biological characterization of an acutely lethal variant of SIV from Sooty Mangabeys. *AIDS Res. Hum. Retroviruses* **5**, 397–409.

GOUDSMIT, J. (1988). Immunodominant B-cell epitopes of the HIV-1 envelope. *AIDS* **2** (Suppl. 1), S41–S45.

GOUDSMIT, J., DEBOUCK, C., MELOEN, R. H., SMIT, L., BAKKER, M., ASHER, D. M., WOLFF, A. V., GIBBS, C. J. & GAJDUSEK, D. C. (1988). HIV-1 neutralisation epitope with conserved architecture elicits early type-specific antibodies in experimentally infected chimpanzees. *PNAS* **85**, 4478–4482.

HAHN, B. H., SHAW, G. M., TAYLOR, M. E., REDFIELD, R. R., MARKHAM, P. D., SALAHUDDIN, S. Z., WONG-STAAL, F., GALLO, R. C., PARKS, E. S. & PARKS, W. P. (1986). Genetic variation in HTLV III/LAV over time in patients with AIDS or at risk for AIDS. *Science* **232**, 1548–1553.

HO, D. D., MOUGDIL, T. & ALAM, M. (1989). Quantitation of HIV-1 in the blood of infected persons. *N. Engl. J. Med.* **321**, 1621–1625.

JAVAHERIAN, K., LANGLOIS, A. J., LAROSA, G. J., PROFY, A. T., BOLOGNESI, D. P., HERLIHY, W. C., PUTNEY, S. D. & MATTHEWS, T. J. (1990). Broadly neutralizing antibodies elicited by the hypervariable neutralizing determinant of HIV-1. *Science* **250**, 1590–1592.

KAPLAN, J. E. (1988). A six-year follow up of HIV infected homosexual men with lymphadenopathy. *JAMA* **260**, 2649–2697.

KOYANAGI, Y., MILES, S., MITSUYASU, R. T., MERRIL, J. E., YINTERS, H. V. & CHEN, I. S. Y. (1987). Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms. *Science* **236**, 819–822.

LAROSA, G. J., DAVIDE, J. P., WEINHOLD, K., WATERBURY, J. A., PROFY, A. T., LEWIS, J. A., LANGLOISE, A. J., DRESSMAN, G. R., BOSWELL, R. N., SHADDUCK, P., HOLLEY, L. H., KARPLUS, M., BOLOGNESI, D. P., MATTHEWS, T. J., EMINI, E. A. & PUTNEY, S. D. (1990). Conserved sequence and structural elements in the HIV-1 principal neutralizing determinant. *Science* **249**, 932–935.

LOONEY, D. J., FISHER, A. G. & PUTNEY, S. D. (1988). Type restricted neutralization of molecular clones of HIV. *Science* **241**, 357–359.

MCKEATING, J. A., GOW, J., GOUDSMIT, J., PEARL, L. H., MULDER, C. & WEISS, R. A. (1989). Characterization of HIV-1 neutralization escape mutants. *AIDS* **3**, 777–784.

MCLEAN, A. R. & KIRKWOOD, T. B. L. (1990). A model of HIV infection in T helper cell clones. *J. theor. Biol.* **147**, 177–203.

MCLEAN, A. R. & NOWAK, M. A. (1991). The interaction between HIV and other pathogens. *J. theor. Biol.* **155**, 69–86.

MEYERHANS, A., CHEYNIER, R., ALBERT, J., SETH, M., KWOK, S., SNINSKY, J., MORFELDT-MANSON, L., ASJÖ, B. & WAIN-HOBSON, S. (1989). Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. *Cell* **58**, 901–910.

MICHEL, M., MANCINI, M., SOBCZAK, E., FARIER, Y., GUETARD, D., BAHRAUI, E. M. & TIOLLAIS, P. (1988), Induction of anti-HIV neutralizing antibodies in rabbits immunized with recombinant HIV-hepatitis B surface antigen particles. *PNAS* **85**, 7957–7161.

NARA, P. L. & GOUDSMIT, J. (1990*a*) In: *Retroviruses of Human AIDS and Related Animal Diseases* (Girard, M. & Valette, L., eds) pp. 203–215. Paris: Pasteur Vaccines.

NARA, P. L. & GOUDSMIT, J. (1990*b*). Neutralization resistant variants of HIV-1. In: *Vaccines* 90 (Brown, F., *et al.*, eds) New York: Cold Springs Harbour Laboratory.

NOWAK, M. A. (1990). HIV mutation rate. *Nature, Lond.* **347**, 522.

NOWAK, M. A., ANDERSON, R. M., MCLEAN, A. R., WOLFS, T., GOUDSMIT, J. & MAY, R. M. (1991). Antigenic diversity thresholds and the development of AIDS. *Science* **254**, 963–969.

NOWAK, M. A. & MAY, R. M. (1991). Mathematical biology of HIV infections: antigenic variation and diversity threshold. *Math. Biosci.* **106,** 1–21.

NOWAK, M. A., MAY, R. M. & ANDERSON, R. M. (1990). The evolutionary dynamics of HIV-1 quasi-species and the development of immunodeficiency disease. *AIDS* **4,** 1095.

PALKER, T. J., MATHEWS, T. J., LANGLOIS, A., TANNER, M. E., MARTIN, M. E., SCEARCE, R. M., KIM, J. E., BERZOFSKY, J. A., BOLOGNESI, D. & HAYNES, B. F. (1989). Polyvalent HIV synthetic immunogen comprised of envelope gp120 T helper cell sites and B cell neutralization epitopes. *J. Immunol.* **142,** 3612–3619.

PHILLIPS, A., LEE, C. A. & ELFORD, J. (1989). Prediction of progression to AIDS by analysis of CD4 lymphocyte counts in a haemophiliac cohort. *AIDS* **3,** 737–741.

PHILLIPS, R. E., ROWLAND-JONES, S., NIXON, D. F., GOTCH, F. M., EDWARDS, J. P., OGUNLESI, A. O., ELVIN, J. G., ROTHBARD, J. A., BANGHAM, C. R. M., RIZZA, C. R. & MCMICHAEL, A. J. (1991). HIV genetic variation that can escape cytotoxic T cell recognition. *Nature, Lond.* **354,** 453–459.

PRESTON, B. D., POIESZ, B. J. & LOEB, L. A. (1988). Fidelity of HIV-1 reverse transcriptase. *Science* **242,** 1168–1171.

ROBERT-GUROFF, M., BROWN, M. & GALLO, R. C. (1985). HTLV III neutralizing antibodies in patients with AIDS and AIDS related complex. *Nature, Lond.* **316,** 72–74.

ROBERT-GUROFF, M., REITZ, M. S., ROBEY, W. G. & GALLO, R. C. (1986). In vitro generation of an HTLV III variant by neutralizing antibody. *J. Immunol.* **137,** 3306–3309.

ROBERTS, J. D., BEBENEK, K. & KUNKEL, T. A. (1988). The accuracy of reverse transcriptase from HIV-1. *Science* **242,** 1171–1173.

RUSCHE, J. R., JAVAHERIAN, K., MCDANAL, C., PETRO, J., LYNN, D. L., GRIMAILA, R., LANGLOIS, A., GALLO, R. C., ARTHUR, L. O., FISCHINGER, P. J., BOLOGNESI, D. P., PUTNEY, S. D. & MATTHEWS, T. J. (1988). Antibodies that inhibit fusion of HIV infected cells bind a 24 amino acid sequence of the viral envelope gp120. *PNAS* **85,** 3198–3202.

SAAG, M. S., HAHN, B. H., GIBBONS, J., LI, Y., PARKS, E. S., PARKS, W. P. & SHAW, G. M. (1988). Extensive variation of HIV-1 in vivo. *Nature, Lond.* **334,** 440–444.

SALINOVICH, O., PAYNE, S. L., MONTELARO, R. C., HUSSAIN, K. A., ISSEL, C. J. & SCHNORR, K. L. (1986). Rapid emergence of novel antigenic and genetic variants of EIAV during persistent infection. *J. Virol.* **57,** 71–80.

SCHNITTMAN, S. M., PSALLIDOPOULOS, M. C., LANE, H. C., THOMPSON, L., BASELER, M., MUSSORI, F., FOX, C. H., SALZMAN, N. P. & FAUCI, A. S. (1989). The receiver for HIV-1 in human peripherial blood is a T cell that maintains expression of CD4. *Science* **245,** 305–308.

SCHULZ, T. F., WHITBY, D., HOAD, J. G., CORRAH, T., WHITTLE, H. & WEISS, R. A. (1990). Biological and molecular variability of HIV-2 isolates from The Gambia. *J. Virol.* **64,** 5177–5182.

SHEPPARD, H., ASCHER, M., MCRAE, B., ANDERSON, R. E., LANG, W. & ALLAIN, J. (1991). The initial immune response to HIV and immune system activation determine the outcome of HIV disease. *J. AIDS*, in press.

SIMMONDS, P., BALFE, P., PEUTHERER, J. F., LUDLAM, C. A., BISHOP, J. O. & LEIGH-BROWN, A. J. (1990). Analysis of sequence diversity in hypervariable regions of the external glycoprotein of HIV-1. *J. Virol.* **64,** 5840.

SUN, N., HO, D. D., SUN, C. R. Y., LIOU, R., WAYNE, G., FUNG, M. S. C., LI, X., TING, R. C., LEE, T., CHANG, N. & CHANG, T. (1989). Generation and characterization of monoclonal antibodies to the putative CD4 binding domain of HIV-1 gp120. *J. Virol.* **63,** 3579–3585.

TAKAHASHI, H., MERLI, S., PUTNEY, S. D., HOUGHTEN, R., MOSS, B., GERMAIN, R. W. & BERZOFSKY, J. A. (1989). A single ammino acid interchange yields reciprocal CTL specificities for HIV-1 gp160. *Science* **246,** 118–121.

TERSMETTE, M., GRUITERS, R. E. & DE WOLF, F. (1989). Evidence for a role of virulent HIV variants in the pathogenesis of AIDS: studies on sequential HIV isolates. *J. Virol.* **63,** 2218–2125.

TINDALL, B. & COOPER, D. A. (1991). Primary HIV infection. *AIDS* **5,** 1–14.

WAIN HOBSON, S. (1989) HIV genome variability in vivo. *AIDS* **3** (Suppl. 1), S13–S18.

WEISS, R. A., CLAPHAM, P. R., CHEINGSONG-POPOV, R., DALGLEISH, A. G., CARNE, C. A., WELLER, I. V. D., & TEDDER, R. S. (1985). Neutralization of HTLV III by sera of AIDS and AIDS risk patients. *Nature, Lond.* **316,** 69–72.

WEISS, R. A., CLAPHAM, P. R., WEBER, J. N., DALGLEISH, A. G., LASKY, L. A. & BERMAN, P. W. (1986). Variable and conserved antigens of HIV. *Nature, Lond.* **324,** 572–575.

WOLFS, T. W. F., DEJONG, J., BERG, H., TIJNAGEL, J., KRONE, W. & GOUDSMIT, J. (1991). Evolution of sequences encoding the principal neutralization epitope of HIV-1 is host dependent, rapid and continuous. *PNAS*, in press.