A Mathematical Model of HIV Infections and Vaccination against the Development of AIDS

Bo Li
1500012204

2016/7

Abstract

The human immunodeficiency virus (HIV) is a retrovirus that attacks human's immune system and leads to the consequent disease AIDS. Major features after the infection of HIV include salient antigenic variation of the virus and a long incubation period (on average 8-10 years) before AIDS is finally induced. These features are closely associated with some features of HIV, the most important one of which may be the high mutation capability of the virus.

Martin A. Nowak and Robert M. May studied the mathematical model of antigenic variation and diversiry threshold during HIV infections, which was published on a paper in 1991 [1]. Later, on the basis of this model, Nowak and Angela R. Mclean discussed and analyzed the vaccination against HIV mathematically [2]. In our work, we will follow and repeat the most contents of the two papers and have some discussions about the mathematical model.

Contents

1 Introduction

1.1 Fundamental Backrounds

1.1.1 Natural History of HIV-1 Infections

In the early period of HIV-1 infection, a viremia is often the typical symptom during which virus replication can be detected. Then a long and variable incubation period follows. Virus isolation becomes difficult and antigens are often undetectable during this so-called asymptomatic phase. Slow replication of viral abundance and constant or slow redution of CD4 cell numbers in patients are also characters of this period. Finally, the immune system may be damaged severely and can no longer inhibit the increase of HIVs and the occurrence of related symptoms.

1.1.2 Interactions between HIV and the Immune System and Genetic Varibility of HIV

HIV can infect CD4 cells, marcophages and dendritic cells, which are all vital for the function of immune system. HIV can lead to low level of CD4 cells by several mechanisms. At the same time, immune responses against HIV will be established after the infection, including humoral and cell-mediated responses. These responses seem to be responsible for the inhibition of HIV's replication during the early period and appearance of the asymptomatic period. Some responses are oriented against conserved epitopes, whereas some others seem to be more specific.

One remarkable property that HIV differs from other viruses in is its high genetic variability. This feature is resulted from the high replication rate of the virus and the high mutation rate during the replication cycle. When a single cell is infected by more than two strains of HIV, recombination between two genomes can occur (not considered in this model). In a word, the high variability of the virus genome is a striking feature and responsible for many phenomena of HIV infection, including the final defeat of the immune system.

1.2 Key Assumptions

These are the key assumptions of the mathematical model based on the biological knowledgement of HIV mentioned by the authors in the paper:

1. The continual evolution of new resistant viral mutants enables the total viral population (summation of all the strains) to evade the elimation by the immune system.

2. Immunological responses to HIV are composed of a specific response (to individual strains) and a unspecific response (acts agiainst all strains).

3. Each virus mutant can damage immune system regardless of their specificity to a particular mutant. Assumptions 2 and 3 characterize the asymmetric interaction between CD4 cells and virus.

2 The Basic Model

Here follows the basic equations to describe the dynamics of HIV abundance and immune response:

$$\dot{v}_i = v_i(r - sz - px_i), \quad i = 1, ..., n; \tag{1}$$

$$\dot{x}_i = kv_i - uvx_i, \quad i = 1, ..., n; \tag{2}$$

$$\dot{z} = k'v - uvz. \tag{3}$$

v_i is the population size of each virus strain i. Likewise, x_i for the specific immune response against strain i and, z for the non-specific general response. r, s, p, k, k' and u are parameters.

r is the duplication rate of the virus, and it can be considered as $r = bQ - d$, where b and d represent the birth and death rates, and Q donates the posibility that the replication is done without error. Equation (1) considers the natural increase, specific and non-specific response for each strain.

Immune cells x_i and z are assumed to be produced by the stimualtion of virus, which was indicated by kv_i and $k'v$ (here $v = \sum v_i$, and the same for $x = \sum x_i$) and elimated at rate uvx_i and uvz, respectively. Note that specific immune cells can only act against a particular strain but be killed by all strains of viruses.

Now consider the new mutants of virus. In the model, the total number of strains n is not constant. We assume in the time interval $(t, t + dt)$, the posibility that a new mutant is created is given by $bQ'v(t)dt$ (as $dt \to 0$), where Q' is the probability that a new escape mutant is created. Here we neglect the situation that the error leads to mutants that already exist in the system, because its possibility is too small.

By summing up equation (2) we get

$$\dot{x} = v(k - ux)$$

Assume $\dot{x}$ and $\dot{z}$ equals to 0, we get the equilibrium values $\hat{x} = k/u$ and $\hat{z} = k'/u$.

There are 3 situations for the parameter space:

1. $r - s\hat{z} < 0$. The unspecific immune itself is strong enough to suppress viral replication of all strains. There will be a reise in viral abundance following the initial infection, but the unspecific immune response will be developed and suppress all subsequently evolved strains after having been mounted the initial strain.

2. $r - s\hat{z} > p\hat{x}$, or to say $r - s\hat{z} - p\hat{x} > 0$. The replication rate r is very high and even the initial viral strain can defeat the specific and unspecific responses. v will tend to infinity in this model.

3. $0 < r - s\hat{z} < p\hat{x}$. In this situation, individual viral strains can outrun the unspecific immune response but not the combination of the two. Now we can establish the existence of a viral diversity threshold.

Consider the product of each strain $P = \prod v_i$. Take the logarithm

$$\ln P = \ln v_1 + \cdots + \ln v_n$$

Take the derivation

$$\frac{\dot{P}}{P} = \sum_{i=1}^{n} \frac{\dot{v_i}}{v_i} = \sum_{i=1}^{n} (r - sz - px_i) = n(r - sz) - px$$

If x and z are at their equilibrium and $n > n_c$ then $\dot{P} > 0$, where

$$n_c = p\frac{\hat{x}}{r - s\hat{z}} = \frac{pk}{ru - sk'}$$

According to mean value inequality, $v = \sum v_i \geqslant P$ grows monotonically. Thus the diversity threshold n_c is given.

Figure 1 and 2 are two typical simulations of the dynamics of the system using Heun's method. Note that in figure 1 the parameter bQ' is a little bigger, which implies the more possibility that HIV breaks the diversity threshold and gets rid of the control of the immune system.[1]

3 Selection Between Strains of Different Replicative Capacity

Now let's consider strains that have different replication rates and the selection between them. Assume there are two strains v_1 (with high replication capability) and v_2 (low replication capability). Then equations for the

[1] In the origin paper, time was measured by years. The authors didn't explain whether and how time axis was scaled. Here it is not scaled.

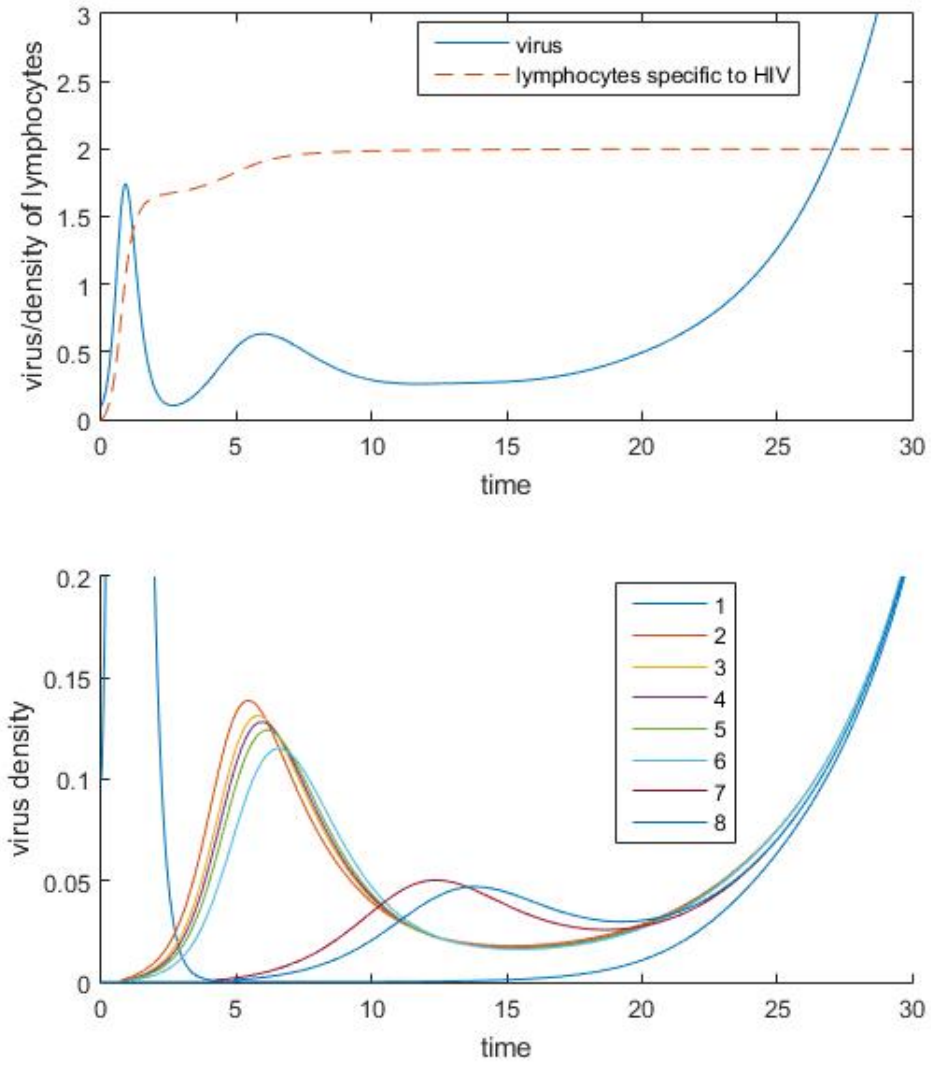

Figure 1: Simulation of the basic model. Values of the parameters are: $r = 5, s = 4.5, p = 5, k = k' = u = 1, bQ' = 2$, determine a diversity threshold $n_c = 10$. (Top) Lymphocytes specific to HIV $(x + z)$ and virus density (v) change with time. (Bottom) 8 strains' density. The virus outrun the diversity threshold and in the final phase a simultaneous increase of all strains can be observed. Note: the model is stochastic and it is not all the case that virus can break through the diversity threshold. The same for figure 2.

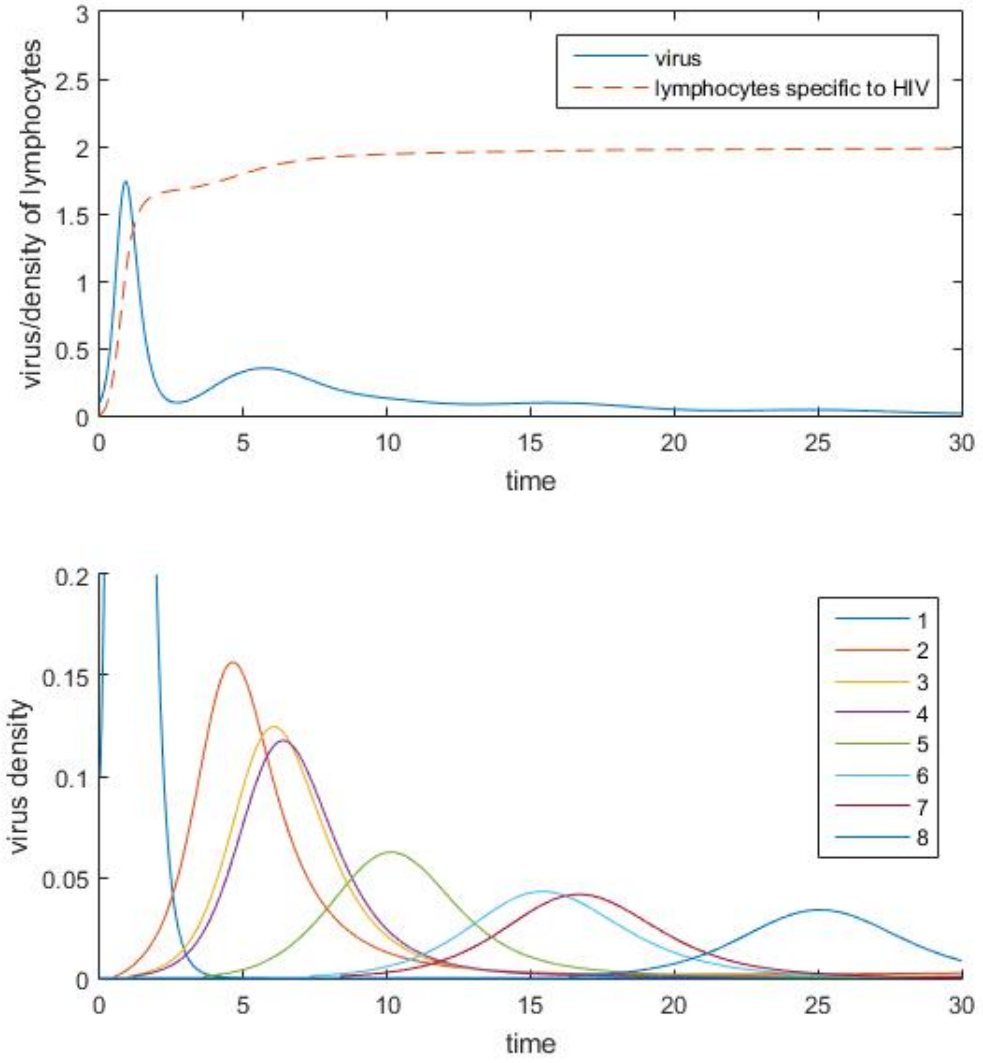

Figure 2: Simulatons of the basic model. All the parameters are the same as those in figure 1, except $bQ' = 1.75$. The virus is finally elimated in this figure.

selection model are

$$\dot{v}_{1,i} = v_{1,i}(r_1 - s_1 z - p_1 x_{1,i})$$
$$\dot{v}_{2,i} = v_{2,i}(r_2 - s_2 z - p_2 x_{2,i})$$
$$\dot{x}_{1,i} = k v_{1,i} - uv x_{1,i}$$
$$\dot{x}_{2,i} = k v_{2,i} - uv x_{2,i}$$
$$\dot{z} = v(k' - uz)$$

Here $v_1 = \sum v_{1,i}$, $v_2 = \sum v_{2,i}$, $x_1 = \sum x_{1,i}$, $x_2 = \sum x_{2,i}$, $v = v_1 + v_2$ and $x = x_1 + x_2$.

High replication rate should be associated with high immunosuppression because of more exposure to the immunological response. So r_1 is large then p_1, s_1 should be large. Figure 3 is one consequence of the selection model. We use the average replication rate $\bar{r} = \frac{v_1 r_1 + v_2 r_2}{v}$ to indicate the selction between the two strains. The initial densities of the two strains are equal.

During the first period of infection, the virus group is dominated by v_1 because of its high replication rate, and some peaks can be seen. With the establishment of immune response, it selects for the slow replicating strain v_2 because p_2, s_2 are smaller. This seems to be a relatively long and stable period in viral density and average replicative capacity. In the final phase, as viral diversity increases, strain 1 will be dominant again.

4 Estimation of Diversification Rate

Now we consider the average number of mutants produced by one single strain. We have assumed that the probability of the creation of an escape mutant by strain i in the time interval $[t, t+dt]$ is given by $bQ'v_i(t)dt$, so the average number produced from one strain is

$$R = bQ' \int_0^\infty v_i(t)dt$$

R cannot be calculated accurately, so we try to estimate it. Assume that the whole population of virus is dominated by some strain i, then we have

$$\dot{v}_i = v_i(r' - px_i), \dot{x}_i = v_i(k - ux_i),$$

where $r' = r - sz$. Get rid of v_i then

$$\dot{v}_i = \frac{r' - px_i}{k - ux_i}\dot{x}_i,$$

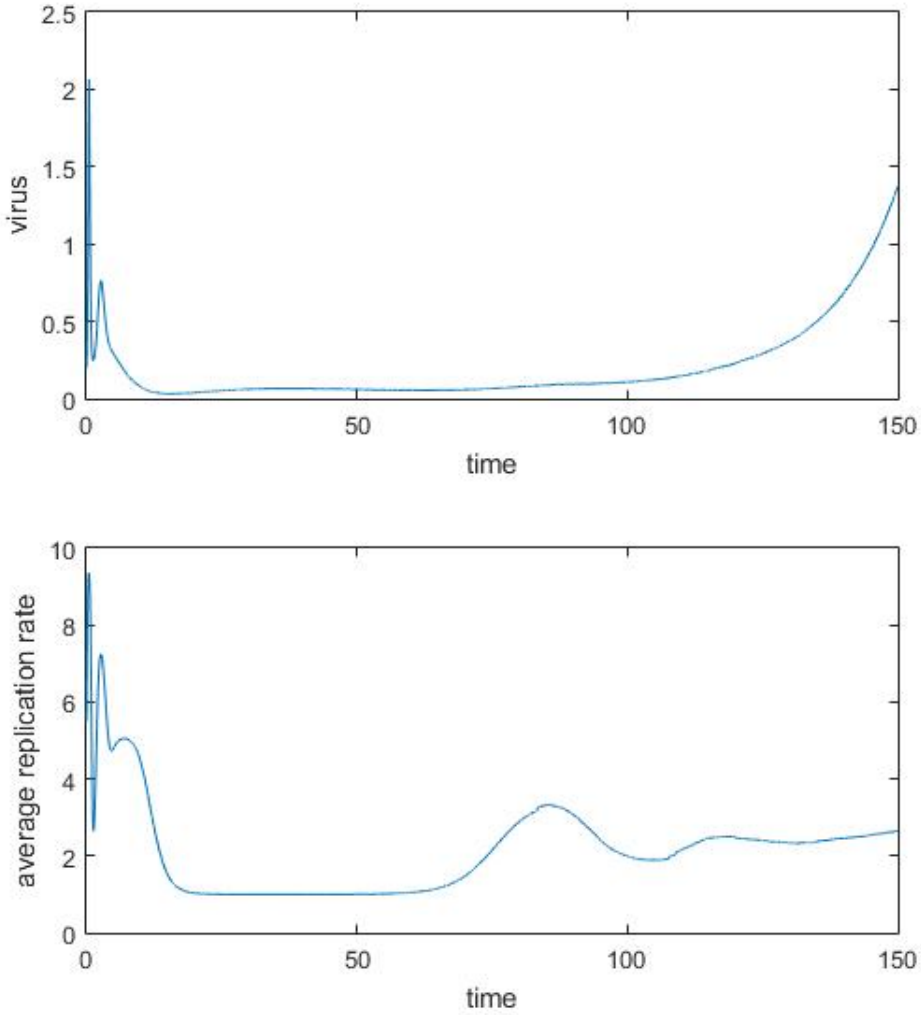

Figure 3: Numerical simulation of HIV infection by two strains of different replicative capabilities. (Top) Total viral density ($v_1 + v_2$) as a function of time. (Bottom) Average replication rate $\bar{r}$ as a functtion of time. Similar to figure 1 and 2, time axis here is not scaled. Parameter values: $r_1 = 10, r_2 = 1, s_1 = 9.7, s_2 = 0.9, p_1 = 20, p_2 = 1, k = k' = u = 1$. Here $bQ' = 3$ (not mentioned in the original paper). Note: chances are that: (1) The virus fails to reach the diversity threshold and is elimated. (2) Values of some variables exceed the range of float numbers and abnormal broken lines are produced. (3) The diversity exceeds the limit ($nmax$ in the code). When using the code, run it for a few times to verify all the results.

intergrate

$$v_i(t) = v_i(0) + \int_0^{x_i(t)} \frac{r' - px}{k - ux} dx = v_i(0) + \frac{p}{u} x_i - \frac{pk - r'u}{u^2} \ln \frac{k}{k - ux_i}.$$

We expect the final extinction of virus, that is to say, $v_i(t) \to 0$ as $t \to \infty$. Because $v_i(0)$ is also very small, so we get the transcendental equation

$$\frac{px_i(\infty)}{pk - r'u} = \frac{1}{u} \ln \frac{k}{k - ux_i(\infty)}.$$

So we get

$$\int_0^\infty v_i(t)dt = \int_0^{x_i(\infty)} \frac{dx}{k - ux} = \frac{1}{u} \ln \frac{k}{k - ux_i(\infty)}.$$

For reasonable (larger) values of n_c ($pk \gg r'u$), the solution for the transcendental equation converges quickly to $x_i(\infty) = 2r'/pk$, [2] so we obtain

$$\int_0^\infty v_i(t)dt = 2\frac{r'/k}{pk - r'u} = \frac{2}{ku}(\frac{r - s\hat{z}}{p\hat{x} + s\hat{z} - r}),$$

hence we find the estimation of R

$$R = \frac{2bQ'}{ku}(\frac{r - s\hat{z}}{p\hat{x} + s\hat{z} - r}).$$

5 Antigenic Drift as a Branching Process

In this section we will discuss the creation of new mutants and antigenic drift as a branching process. The concepts in this section are crucial for the discussion of vaccination against HIV in later sections.

We use the signal $P_i(t)$ to represent the probability that one certain strain produces i escape mutants in the time interval $[0, t]$. $t = 0$ here means the origin of the very strain that is considered. Consider the situation that no mutant is produced, obviously we have

$$P_0(t + dt) = P_0(t)P_0(dt) = P_0(t)[1 - bQ'v(t)dt].$$

When $dt \to 0$, we have

$$\dot{P}_0(t) = -bQ'v(t)P_0(t).$$

[2] The author didn't explain how this was achieved ... neither do I understand.

Using the initial condition $P_0(0) = 1$ and take the integration

$$P_0(t) = exp[-bQ' \int_0^t v(\tau)d\tau],$$

we appoint P_i to be $P_i(\infty)$, then

$$P_0 = exp[-bQ' \int_0^\infty v(\tau)d\tau] = e^{-R}.$$

Now consider $P_i(t)$ for $i \neq 0$. We use signal $R(t) = bQ' \int_0^t v(\tau)d\tau$ and $R = bQ' \int_0^\infty v(\tau)d\tau$, in fact $P_i(t)$ follows a Poisson distribution

$$P_i(t) = R^i(t)e^{-R(t)}/i!$$

Next we prove it by method of induction[3] $P_i(t + dt)$ can be decomposed as

$$P_i(t + dt) = P_{i-1}(t)bQ'v(t) + P_i(t)[1 - bQ'v(t)],$$

Take the limitation $dt \to 0$ and then

$$\dot{P}_i(t) + P_i(t)bQ'v(t) = P_{i-1}(t)bQ'v(t).$$

For $i = 0$, obviously the conclusion of Poisson distribution is correct. Now assume for $i = 0, ..., k - 1$, the conclusion is correct. For $n = k$, we have

$$\dot{P}_k(t) + P_k(t)bQ'v(t) = P_{k-1}(t)bQ'v(t).$$

multiplied by $e^{R(t)}$

$$e^{R(t)}[\dot{P}_k(t) + P_k(t)bQ'v(t)] = e^{R(t)}P_{k-1}(t)bQ'v(t) = R^{k-1}(t)bQ'v(t)/(k-1)!$$

notice that $\dot{R}(t) = bQ'v(t)$, so we get

$$e^{R(t)}dP_k(t) + de^{R(t)}P_k(t) = R^{k-1}(t)dR(t)/(k-1)!$$

and

$$d[e^{R(t)}P_k(t)] = dR^k(t)/k!$$

using the initial condition $R(0) = 0$ and $P_i(0) = 0 (i \neq 0)$, we have

$$P_i(t) = R^i(t)e^{-R(t)}/i!$$

and

$$P_i = R^i e^{-R}/i!.$$

[3]There's no proof in the original paper and this is my proof.

If there are n_t strains at generation t, then the number of strains in the next generation $t+1$ is given by

$$n_{t+1} = \sum_{j=1}^{n_t} k_j$$

where k_j is a random variable with distribution $Pr(k_j = i) = P_i$. A branching process is thus formed, and the probability-generating function of the process is

$$F(s) = \sum_{i=0}^{\infty} P_i s^i = e^{-R} \sum_{i=0}^{\infty} \frac{(Rs)^i}{i!} = e^{R(s-1)}.$$

If this process is started from one strain strain, the probability of extinction is given by the smallest positive root s_0 of the equation $s = F(s)$, thus $1 - s_0$ is the probability that the antigenic drift can continues forever. If starting with n strains, then the extinction probability reduces to s_0^n. Three curves for diffenent n values are shown in figure 4.

6 Immunization by Vaccination against HIV

In order to suppress the rapid mutation of HIV and progression towards AIDS, inducing broad immunological responses against different subtypes of viruses using a vaccination seems to be a viable approach. Considering the strong capability to create new mutants, the fraction, donated by f, that a vaccination can cover and act against becomes an essential question.

Now we specify the effect of immunization by increasing the specific immunity parameter k by a folds. We have estimated the value of R in section 4

$$R = \frac{2bQ'}{ku}\left(\frac{r - s\hat{z}}{p\hat{x} + s\hat{z} - r}\right)$$

by considering the form of the expression, we have

$$R' = R/a$$

where R' donates the number of escape mutants produced by the covered strains, thus on average

$$\bar{R} = (1-f)R + fR' = [1 - f(1 - \frac{1}{a})]R. \tag{4}$$

$\bar{R} < 1$ requires

$$f > f_c = \frac{1 - 1/R}{1 - 1/a} \tag{5}$$

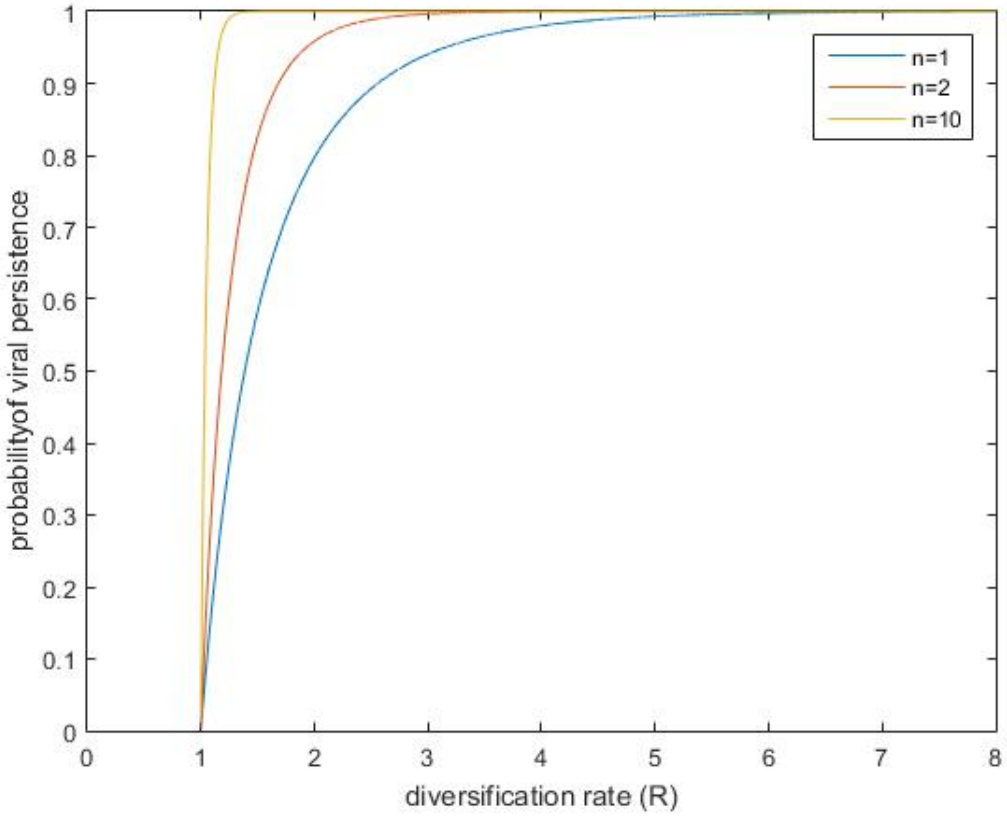

Figure 4: The probability of eventually developing AIDS as a function of diversification rate, R. We can directly learn from the figure that $R > 1$ is necessary for the development of AIDS, that is to say, one single strain must produce at least one escape mutant, which is a result that accords with common sense.

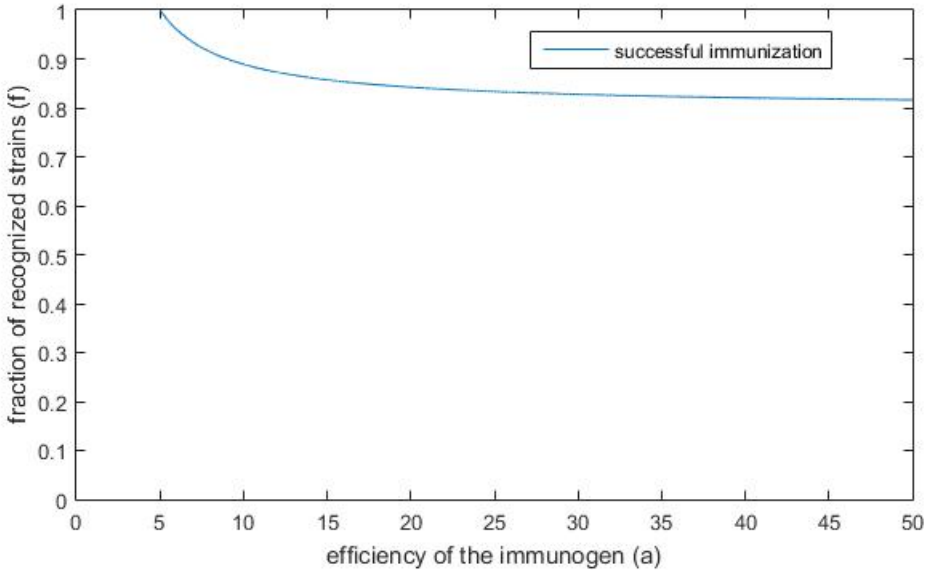

Figure 5: a is the factor indicates the immune response as a result of vaccination. For a certain value of R, $a > R$ is necessary for the successful defense of AIDS. Also, a fraction larger than f_c is needed. The region above the curve in the figure indicates successful immunization. In this example, $R = 5$.

to make sure that HIV is extinct in the final phase. The critical fraction f_c serves as a function of R and a. Note that a must be greater that R to obtain reasonable $f_c < 1$. Figure 5 shows f_c as a function of a.

7 The Timing of Post-Exposure Immunization

In this section, we discuss the importance of early immunotherapy against HIV. According to the branching process we discussed before, the average number of strains after t generations started with a single strain is given by R_t. So we can know the extinction probability directly

$$P(t) = s^{R^t}.$$

We would like to cauculate how long it will take for the extinction probability to be reduced to 0.1. Assume

$$s^{R^t} = 0.1$$

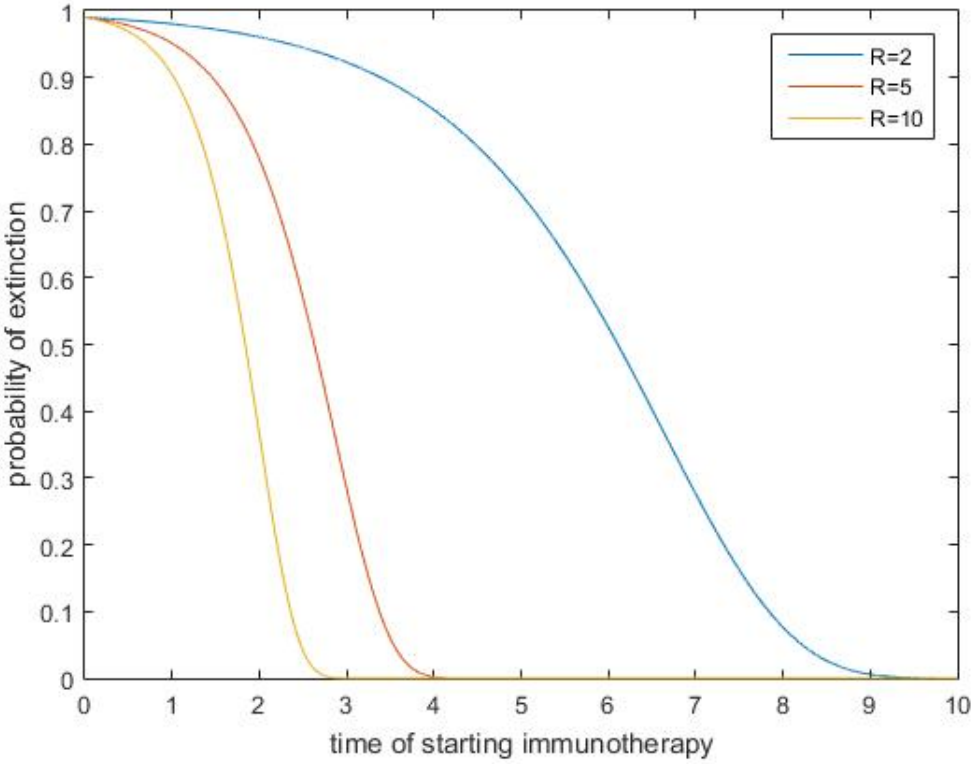

Figure 6: It is important to apply immunotherapy as soon as possible. Successful treatment can depend critically on the time when it is started. In this figure the extinction rate started from one strain $s_0 = 0.99$. Despite the high value of s_0, the total probability of extinction still decreases rapidly because of the form of the exponential function.

the solution is

$$t_c = -\frac{\lg \lg 1/s}{\lg R}.$$

The result shows that extinction probability quickly tends to 0 after a period of time defined by t_c (figure 6). f_c we discussed in last section is only a function of a and R, which is not directly associated with viral diversity. However, the more the viral diversity differs, the more strains of viruses are needed to be covered by a vaccine, which is a great challenge to the effect of therapy.

8 Discussion

Certainly, the models described here are just very simplificaion of real conditions in human bodies. Based on some simple backgrounds about HIV infection, three key assumptions are proposed to be the fundamental of the basic model, including asymmetric interactions between HIVs and immune system and high variation capability that leads to the escape of the immunological repression. The later one is an important concept that is throughout the whole model.

In the basic model, computer simulations explain the critical role that mutation plays in the infection. Also, diversity threshold n_c is characterized to be the criterion for whether HIV can finally get rid of the elimation of immune response or not. In the model of two strains of different replication capability, we can observe and explain the wave of two strains' fractions. All these may be responsible for the long incubation period without symptoms and breakdown of the immune system in the final phase.

By considering the antigenic drift as a branching process, we can discuss the extinction probability of the viral population. This process is also used to discuss immunization by vaccination against HIV infection. It is found that the effect of a vaccine depends essentially on its sphere and strength of immunological induction of specific immune responses. At the same time, the importance of early treatment can be easily characterized by some very simple calculations.

The model in the two papers partly indicates some mechanisms behind the features of HIV infection and its damage to human's immune system. It also has some significances of study of HIV vaccines, which is not completely successful by now and still consuming lots of researchers' efforts. Just as what the authors said at the end of their discussion: "It appears to be a fascinating task for the mathematical biology to present models that describe the dynamics and relative efficiency of these immune responses, and to provide a rationale for vaccine development."

References

[1] Nowak M A, May R M. Mathematical biology of HIV infections: antigenic variation and diversity threshold[J]. Mathematical Biosciences, 1991, 106(1):1-21.

[2] Nowak M A, Mclean A R. A mathematical model of vaccination against HIV to prevent the development of AIDS.[J]. Proceedings of the Royal Society B Biological Sciences, 1991, 246(1316):141-6.