Stochastic Interplay between Mutation and Recombination during the Acquisition of Drug Resistance Mutations in Human Immunodeficiency Virus Type 1

Christian L. Althaus† and Sebastian Bonhoeffer*

Ecology & Evolution, ETH Zürich, ETH Zentrum CHN, CH-8092 Zürich, Switzerland

The emergence of drug resistance mutations in human immunodeficiency virus (HIV) has been a major setback in the treatment of infected patients. Besides the high mutation rate, recombination has been conjectured to have an important impact on the emergence of drug resistance. Population genetic theory suggests that in populations limited in size recombination may facilitate the acquisition of beneficial mutations. The viral population in an infected patient may indeed represent such a population limited in size, since current estimates of the effective population size range from 500 to 10^5. To address the effects of limited population size, we therefore expand a previously described deterministic population genetic model of HIV replication by incorporating the stochastic processes that occur in finite populations of infected cells. Using parameter estimates from the literature, we simulate the evolution of drug-resistant viral strains. The simulations show that recombination has only a minor effect on the rate of acquisition of drug resistance mutations in populations with effective population sizes as small as 1,000, since in these populations, viral strains typically fix beneficial mutations sequentially. However, for intermediate effective population sizes (10^4 to 10^5), recombination can accelerate the evolution of drug resistance by up to 25%. Furthermore, a reduction in population size caused by drug therapy can be overcome by a higher viral mutation rate, leading to a faster evolution of drug resistance.

Human immunodeficiency virus (HIV) infection is characterized by a high genetic diversity and a remarkable capacity of the virus to evolve rapidly in response to novel selection pressures such as drug therapy. These properties are mediated, at least in part, by the high mutation rate (3.4×10^{-5} per base pair) (27) and the fast turnover of actively infected cells (1 to 2 days) (18, 42). In addition, recombination has been conjectured as a further contributing factor (7). The capacity of recombination is common to all retroviruses and is due to template switching between the two genomic RNA strands during reverse transcription. Recombination can occur when the infecting virion is heterozygous (i.e., carries two distinct genomic strains). Such heterozygous virions are produced by cells that are coinfected by two distinct proviruses.

In vitro experiments show that recombination can lead to the rapid emergence of drug resistance (14, 21, 29). In these experiments, recombinant virus is typically produced by infecting cells with equal mixtures of two mutant strains at a high multiplicity of infection. These conditions favor the production of heterozygous virions and thus lead to the rapid emergence of multiple drug resistance through recombination. However, in vivo the conditions may be less favorable for recombination. First, recombination events will occur between the predominating wild type and the mutants, while recombination events between two mutants will be rare. Thus, it is not clear a priori whether recombination in vivo would more often combine than break up favorable combinations of mutations. Second, the multiplicity of infection may be considerably lower in vivo. Hence, fewer heterozygous virions may be produced.

To investigate the effect of recombination on the evolution of drug resistance, we recently developed a population genetic model of recombination in HIV (6). The model showed that whether recombination facilitates the evolution of drug resistance depends on how mutations interact to determine viral fitness. In particular the model suggested that if mutations conferring resistance interact synergistically (i.e., the combination of resistance mutations results in a higher fitness benefit than expected from the product of the benefits of single resistance mutations) recombination will break down these favorable combinations of mutations and therefore slow down the evolution of drug resistance. In population genetic theory, this interaction of single mutations is called positive epistasis. A recent large-scale data analysis of the fitness effects of drug resistance mutations gave evidence for a predominance of positive epistasis between resistance mutations in HIV-1 in the absence of drugs (3). These findings support the idea that recombination is unlikely to facilitate the preexistence of drug-resistant virus types prior to therapy. However, what type of epistatic interaction predominates between resistance mutations during therapy remains to be established.

A key assumption of our previous model was that the population size of infected cells is infinite and therefore that the transitions during a viral life cycle can be calculated deterministically (6). In agreement with population genetic theory (2, 34), the model showed that for infinite populations the effect of recombination depends critically on epistasis. However, in

* Corresponding author. Mailing address: Ecology & Evolution, ETH Zürich, ETH Zentrum CHN, CH-8092 Zürich, Switzerland. Phone: 41 1 632 7106. Fax: 41 1 632 12 71. E-mail: sebastian.bonhoeffer @env.ethz.ch.
† Present address: Theoretical Biology, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

TABLE 1. Estimates of the HIV-1 in vivo effective population size[a]

Study (references)	Patient group	Population	Gene(s)	N_e
Leigh Brown (22)	Untreated	Free virus	*env*	1,000–2,100
Nijhuis et al. (31)	Before and during therapy	Free virus	*env*	450–16,000
Rodrigo et al. (37)	Before and during therapy	Provirus	*env*	925–1,800
Rouzine and Coffin (38)	Untreated and on therapy	Free virus/provirus	*pro*	100,000
Seo et al. (39)	Untreated and on therapy	Free virus/provirus	*env*	1,500–5,500
Shriner et al. (40)	Untreated and on therapy	Free virus/provirus	*env*	1,000
Achaz et al. (1)	Untreated	Free virus	*gag-pol*	1,000–10,000

[a] Seven different studies estimate N_e between 500 and 10^5. Shriner et al. (40) show that the high number estimated by Rouzine and Coffin (38) decreases to aproximately 10^3 upon correcting for a biased sampling procedure. Therefore, the effective population size is probably between 10^3 and 10^4. Untreated patients show similar numbers to patients on therapy. Similar results were obtained by measuring plasma viral RNA or cell-associated viral DNA. Although different selection pressures act on *env* and *gag-pol*, it is interesting to note that similar estimates of N_e have been found based on data for both genes.

populations limited in size population genetic theory suggests that stochastic processes may play an important role in the outcome of recombination (34). The main difference in finite populations is that not all combinations of mutations are present at the same time in the viral population. If $N < 1/\mu$, with N being the population size and μ being the mutation rate per base pair, there will be less than one mutation on average per generation at a certain base pair and hence the presence or absence of any single-point mutant is governed by stochastic effects. However if $N > 1/\mu$, the dynamics of single-point mutants can be approximated deterministically (38).

Whether the dynamics of recombination can be approximated by deterministic models depends on a different population size, because it is important whether combinations of mutations are already present in the population. For the simplest case that a certain double mutant is produced each generation, the condition $N > 1/\mu^2$ must be met. Given a mutation rate of $\mu = 3.4 \times 10^{-5}$ per base pair (27), this corresponds to a population size of 8.7×10^8. The total size of HIV RNA-positive cells is large and estimates range from 10^7 to 10^8 (15). However, the number of infected cells that contribute their viral genetic information to successive generations is likely to be smaller. For example, not all RNA-producing cells may produce virions that can infect a new target cell. Therefore, the appropriate measure of population size is that of a genetically idealized population with which the actual population can be equated genetically, e.g., an idealized population that experiences the same amount of genetic drift. This population genetic concept is called the effective population size, N_e (17).

Current estimates of the effective population size of HIV in vivo range from 500 to 10^5 with most estimates between 10^3 and 10^4 (see Table 1). These estimates have been obtained both for treated and untreated patients. Moreover, the estimates are based on different genes. Since the difference between the total number of infected cells (10^7 to 10^8) and the estimated effective population size (500 to 10^5) is likely due to selection, differences between selection pressures in treated versus untreated patients or between different genes could in principle explain the discrepancies between these estimates. However, so far no clear trends have emerged from the currently available data. A further important point is that the effective population size is a concept that only has a well-defined meaning in the context of a chosen process of interest. In particular, the effective population size with regard to neutral or nearly neutral genetic variations will be different from

the effective population size in terms of rare beneficial or deleterious mutations. Hence, the current estimates of N_e, based on *env* and *gag-pol*, may underestimate N_e with regard to rare selected mutations. Another factor affecting effective population size is population structure. Frost et al. (10) used a metapopulation model to show that the population structure can lead to an effective population size that is considerably smaller than the total size of infected cells. However, this study was not able to account for effective population sizes as low as 10^3 and 10^4. Hence, further research will be needed to resolve discrepancies between current estimates of N_e.

As we are interested in the effect of stochasticity on the evolution of drug resistance, we focus here primarily on small population sizes (i.e., 1,000 to 10,000). A large body of population genetic literature has shown that stochastic effects can have an important influence on the effect of recombination in small or even large but finite populations (8, 9, 19, 28, 30, 33). These results prompt the consideration of finite population sizes in population genetic studies of HIV evolution. Leigh Brown and Richman (23) discussed the influence of a low effective population number on the evolution of drug resistance, and another study used a simple stochastic model to calculate the relative frequency of drug-resistant mutants prior to therapy (11). However, to our knowledge, there is currently no population genetic model of stochastic viral evolution that incorporates the combined effects of mutation and recombination. To this end, we expand a previously published deterministic model of HIV recombination (6) to study the stochastic effects resulting from finite populations of infected cells.

MATERIALS AND METHODS

Model. A virus contains two RNA strands that are transcribed into double-stranded DNA, called a provirus, which is then inserted into the cell nucleus. In our model, we consider proviruses consisting of two loci with two alleles (a/A and b/B). Thus, we have four types of proviruses, ab, Ab, aB, and AB, where lowercase letters denote drug-sensitive wild-type alleles and uppercase letters denote drug-resistant mutant alleles. There is a constant population of target cells, N, that are infected. These cells are either singly or doubly infected and therefore carry one or two proviruses. For simplicity we neglect cells that are infected with more than two proviruses. With f being the frequency of doubly infected cells, the total number of proviruses is $(1 + f)N$. We assume discrete generations and the provirus frequencies change after the completion of a full replication cycle. We divide the replication cycle into three steps, which are all formulated as stochastic processes and are discussed separately below. An illustration of one generation of the viral population is given in Fig. 1. The program was written in C++ and run under Linux and Mac OS X. The source code can be obtained freely on request from the authors.
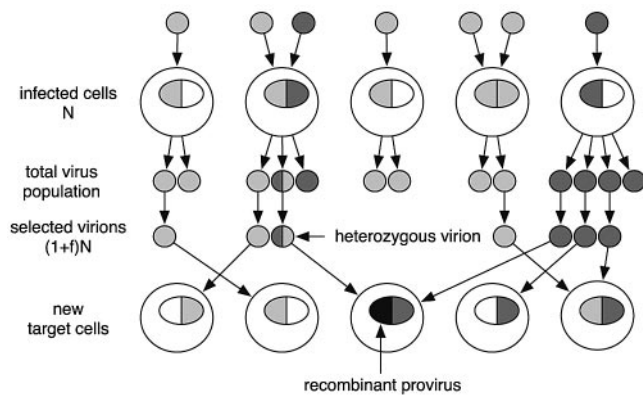
FIG. 1. Schematic illustration of a viral replication cycle in the computer model. A number of N cells are either singly or doubly infected. They contain one or two proviruses inserted in their nucleus. For simplicity, only two viral types are shown, either as light or dark gray. Infected cells produce virions according to the fitness of the inserted proviruses. From the total virus population, $(1 + f)N$ virions are selected that will infect N new target cells. The infecting virions are distributed randomly to the cells. A heterozygous virion can produce a recombinant provirus during reverse transcription, here shown in black.

Distribution of proviruses in the infected cell population. A total of $(1 - f)N$ cells are singly infected. Proviruses are chosen randomly according to their frequencies and distributed among these cells. The same is done for the $f N$ doubly infected cells, except that two proviruses are chosen for each cell. For two loci with two alleles, there are four different singly infected cell types and 10 different types of doubly infected cells. Note that not all of these types may be present at all times, since the population size of cells, N, is limited.

Selection of virions that infect target cells. Single infected cells produce homozygous virions that correspond to the provirus inserted in the cell. In contrast, doubly infected cells can generate heterozygous virions when the inserted proviruses are distinct from each other. We assume that the viral RNA strands are mixed randomly, and therefore half of the virions released from doubly infected cells are heterozygous. The number of virions released by each infected cell depends on the viral proteins produced by the cell and derived from the provirus or proviruses. Virions produced from doubly infected cells can therefore show phenotypic mixing (32) if the proviruses are distinct. We assume that the number of virions produced is proportional to the average fitness of the two proviruses. Since the viral burst size of infected cells is around 10^3 or higher (13, 41), we calculate the production of virions deterministically. The calculation of virions is best illustrated by an example. Cells that are double infected by two single-resistance viruses, Ab and aB, generate heterozygous virions, $AbaB$, with a frequency given by

$$V_{AbaB} = \frac{1}{V_{tot}} \frac{1}{2} C_{AbaB} \frac{w_{Ab} + w_{aB}}{2}$$

where C_{AbaB} is the frequency of cells double infected by the Ab and aB strains. The fitness of the virions is the average of the proviral fitnesses (defined as w_{Ab} and w_{aB}). Half of the produced virions will be heterozygous, assuming random segregation, and the result is normalized with V_{tot} to ensure that all virion frequencies sum up to one. From the resulting total virus population, $(1 + f)N$ virions are selected randomly to infect target cells (illustrated in Fig. 1), according to their frequency.

Transcription of inserted virions to proviruses. After a virion has bound to the surface of a target cell, both genomic RNA strands are released into the cell cytoplasm and the reverse transcriptase (RT) begins to synthesize proviral DNA. During reverse transcription, the RT molecule may jump back and forth between the two genomic RNA strands, thus producing a recombinant provirus. For a given virion, we calculate the probability of producing a certain provirus type as a function of the mutation rate, μ, and the recombination rate, r. To give an

example, the probability that the two RNA strands Ab and aB are transcribed into an AB provirus is given by

$$p_{AB} = \tfrac{1}{2}\{(1 - \mu)[(1 - r)\mu + r(1 - \mu)]\} + \tfrac{1}{2}\{\mu[(1 - r)(1 - \mu) + r\mu]\}. \quad (1)$$

The RT attaches to either strand with probability 0.5. At the first strand, it does not mutate the A allele with probability $1 - \mu$ and remains on the same strand with mutation at the second locus with probability $(1 - r)\mu$. Doing this for all four possible pathways of the RT, we get equation 1. Every single virion transcribes into one of the four different proviruses according to the calculated probabilities. This stochastic process generates a total of $(1 + f)N$ proviruses that then are distributed in the total cell population, N, as described in the first step.

RESULTS

Emergence of resistant mutants prior to therapy. Mutant virus types that cause resistance against drugs are constantly produced from the basal population through mutation. Most of these mutants carry a cost of resistance compared to the wild-type population (3). Therefore, resistant mutants are held in a mutation-selection balance prior to therapy. If they are already present at the start of therapy they may contribute to a rapid emergence of resistance (4, 5, 36). As an example, we calculated the expected preexistence frequencies of the M41L and T215Y resistance mutations against zidovudine. The levels of fitness of the mutants relative to the wild type are as follows: M41L, 80%; T215Y, 85%; and M41L/T215Y, 85% (16). Note that one nucleotide substitution is sufficient to generate the mutation at position 41, but two nucleotide substitutions are required to get the mutation at position 215. The preexistence frequencies that are equivalent to the probabilities of emergence in finite populations are shown in Table 2.

Since the interactions of the deleterious effects of the single mutants are nonmultiplicative, there is an epistatic interaction between these two positions. A mathematical definition of epistasis between two loci is $E = w_{ab}w_{AB} - w_{Ab}w_{aB}$, where w denotes the fitness of the corresponding mutants. Using the estimated fitness values, we get $E > 0$ and therefore positive epistasis. The T215Y mutant appears to have a compensating effect on the deleterious effect of the M41L mutant, and therefore the double mutant should be present at higher frequencies than expected from the deleterious effect of the single mutants. Recombination will break up this nonrandom association and thus reduce the frequency of the double mutant, but increase the frequencies of the single mutants. As seen in Table 2 the numerical effect is moderate. For $N_e = 10^4$, the total provirus

TABLE 2. Calculated frequencies of the M41L, T215Y, and M41L/T215Y resistance mutations to zidovudine in the absence of drugs[a]

Virus type	Frequency		Effect of recombination
	$r = 0$	$r = 0.5$	
Wild type	0.99976	0.99976	↓
M41L	0.00024	0.00024	↑
T215Y	9.598×10^{-9}	9.600×10^{-9}	↑
M41L/T215Y	5.087×10^{-12}	3.119×10^{-12}	↓

[a] The resistant mutants are held in a mutation-selection balance. The frequencies of these mutants are equivalent to the probabilities of emergence in finite populations. Recombination increases the frequencies of single mutants at the expense of the wild type and double mutant, but the numerical effect is moderate. Since Jung et al. (20) measured 75 to 80% of infected cells harboring more than one provirus, we set $f = 0.75$. The mutation rate was set to $\mu = 3 \times 10^{-5}$ per base pair.
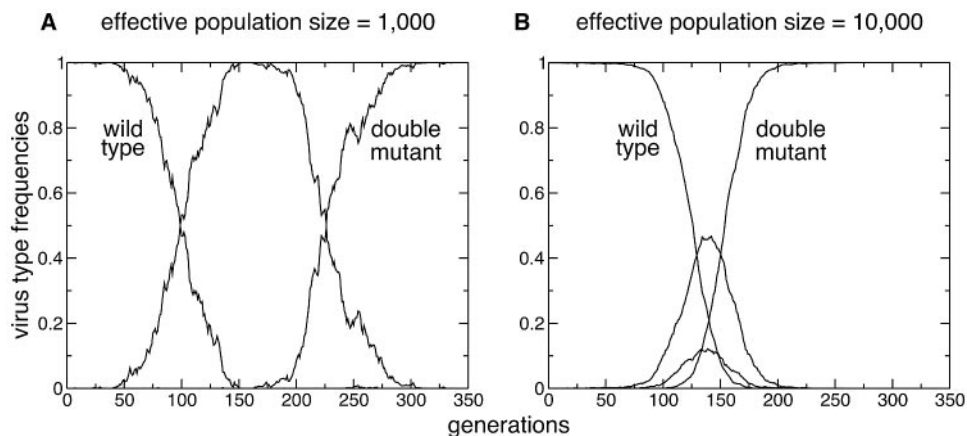
FIG. 2. Evolution of drug resistance. The homogenous wild-type population is replaced by the resistant double mutant. (A) In a small population of infected cells ($N_e = 10^3$), the first single mutation typically rises to fixation before the population acquires the second mutation. (B) In simulations with a larger population of infected cells ($N_e = 10^4$), the single mutations are typically both present at the same time. Therefore, recombination can combine the two single mutations and accelerate the evolution of the double mutant.

population is $(1 + f)N_e = 17, 500$. For this effective population size, only the M41L single mutant is likely to occur before therapy, representing on average about four proviruses. In contrast, the probability of T215Y or M41L/T215Y mutants being present prior to therapy would be very small.

Stochastic evolution of resistance. We measured the influence of recombination on the acquisition of drug resistance mutations during drug therapy. From our findings of preexistence frequencies, we can assume that at least one single mutant is missing at start of therapy. Recombination cannot generate the double mutant until both single mutants are present, and it is therefore appropriate to start the simulations from a homogenous wild-type population. We follow the dynamics of mutant virus types until the double mutant is fixed in the population. Double resistance can evolve in two different ways, as shown in Fig. 2. If the population size of infected cells is small, it is unlikely that both single mutations are present in the population at the same time. Instead, typically one of the

single mutants is produced by mutation and rises to fixation before the second mutation arises in the population. Thus, the double mutation occurs by sequential acquisition of both mutations. In this case, recombination has little effect on the evolution of drug resistance. In contrast, for large populations of infected cells, both single mutants are present at the same time. When one cell is infected by both of them, a double mutant can be produced through recombination. In this scenario, recombination can substantially accelerate the evolution of drug resistance. These findings are in line with classical population genetics theory, which showed that in populations of intermediate size recombination can accelerate the rate of fixation of beneficial alleles (9, 28, 30).

To measure the effect of recombination, we let the simulations run over different values of N_e and with different strengths of selection. The effect of recombination is expressed as the reduction in the time of fixation of the double-resistance mutant through recombination, given by $1 - T_{r = 0.5}/T_{r = 0}$,
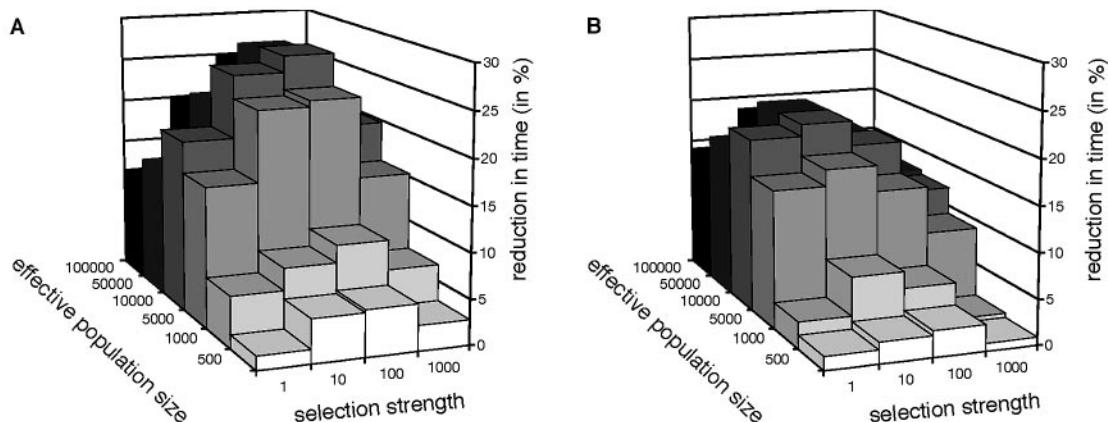

FIG. 3. Reduction in time to fixation of a double mutant through recombination as a function of different population sizes and selection strengths. (A) Recombination has a significant effect for a population size of 10^4 and intermediate selection strengths. However, at smaller population sizes, the effect is dramatically reduced. The following parameters were used: $f = 0.75$ and $r = 0.5$. (B) Using smaller parameters ($f = 0.25$ and $r = 0.1$) results in an overall decrease of the beneficial effect of recombination. In all simulations, the mutation rate was $\mu = 3 \times 10^{-5}$ per base pair.
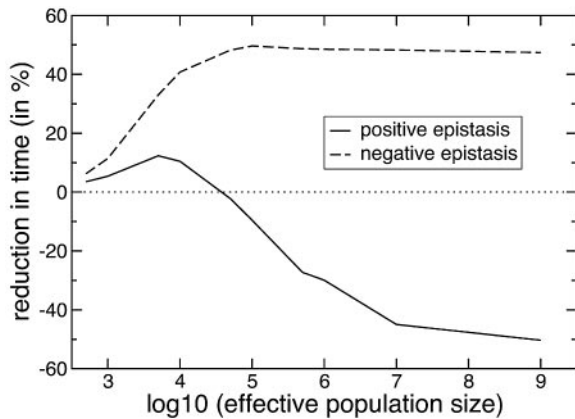
FIG. 4. Reduction in time to fixation of the double mutant through recombination as a function of the effective population size when there is positive or negative epistasis between resistance mutations. Drift effects dominate in populations smaller than 10^5, and in these populations, recombination reduces the time to fixation. For larger populations, genetic drift is smaller and epistatic interactions determine the population structure. In these populations, recombination only reduces the time to fixation for negative epistasis. For positive epistasis, recombination increases the time to fixation. The following parameters were used: $\mu = 3 \times 10^{-5}$, $r = 0.5$, $f = 0.75$. The fitness levels of the wild type and double mutant were set to 1 and 1.2, respectively. The fitness levels of both single mutants were 1.05 for positive epistasis and 1.15 for negative epistasis.

where T_r is the average time to fixation for a recombination rate, r. A reduction in fixation time of 0.2 thus means that recombination reduces the time to fixation by 20%. Fixation here means that the double-resistance mutants constitute at least 90% of the population. A full fixation of 100% is not appropriate since different virus types can be present due to recurrent mutation. The average time of fixation is measured in viral generations. Fitness of the wild type (w_{ab}) was set to 1. The single mutants are taken to have equal fitnesses, $w_{Ab} = w_{aB} = 1 + s$, and the double mutant is assumed to have fitness $w_{AB} = (1 + s)^2$. Thus, we assume no epistasis between mutations. Since the mutations confer resistance, they are beneficial during drug therapy compared to the wild type. Simulations were done with different strengths of selection, i.e., s was set to 1%, 10%, 100%, and 1,000%. We assume maximal recombination, $r = 0.5$, and we set the fraction of doubly infected cells to $f = 0.75$ (20). Figure 3A shows the reduction in time caused by recombination for different values of N_e and different strengths of selection. Recombination has a significant effect at population sizes of around 10^4 and at intermediate selection strengths, with the reduction in fixation time being approximately 25%. The effect of recombination is dramatically reduced with decreasing population sizes. At $N_e = 103$ the reduction in fixation time is only 5 to 10%. We also considered the effect of a smaller fraction of doubly infected cells and lower rates of recombination, since both may in reality be smaller than the values used so far. Since we use an estimate for the multiplicity of infection derived from solid tissue (20), where target cells are in close proximity to each other, it is possible that averaged across all infected cells the fraction of multiply infected cells may be smaller. Moreover, the recombination rate between two loci depends on the number of

nucleotides between them. Thus, mutations close to each other will have lower recombination rates than mutations further apart. Plot B in Fig. 3 shows the effect of recombination on fixation time with parameters $f = 0.25$ and $r = 0.1$. With these smaller parameter values, the effect on fixation time is less, with the maximum reduction being between 15 and 20%. For a small population size of 10^3, recombination will have a minor effect of 1 to 7%.

Epistasis in small population sizes. So far, we have excluded epistatic interactions between resistance mutations to examine the effects of recombination that arise in finite populations, independent of epistasis. However, epistasis between resistance mutations can generate a linkage disequilibrium that is reduced by recombination. Since there is evidence for a predominance of positive epistasis among resistance mutations in HIV-1 in the absence of therapy (3), we wanted to determine how the combined effects of epistasis and finite population sizes are affected by recombination. To this end, we measured again the reduction in fixation time compared to no recombination in two different fitness regimes. The fitness levels of the wild type and drug-resistant double mutant were set to 1 and 1.2, respectively. In simulations with positive epistasis, the fitness of each single mutant was set to 1.05. For simulations with negative epistasis, the fitness of each single mutant was set to 1.15. Figure 4 shows the reduction in time caused by recombination over different effective population sizes for these fitness values. Recombination reduces the time to fixation in small populations of around 10^3, since beneficial single mutants that arise after the start of therapy can be combined. The reduction is larger for negative epistasis because the single mutants are overrepresented and therefore are more likely to be combined. The effect of genetic drift decreases in larger population sizes. For $N_e > 10^5$, drift effects are small and epistatic selection determines the population structure. As described by Bretscher et al. (6), recombination only reduces the time to fixation for negative epistasis. For positive epistasis, there is a negative reduction, i.e., recombination increases the time to fixation of the double mutant.

Population size versus mutation rate. A correlation between the evolution of drug resistance and increased mutation rate has been observed in HIV (24). For example, zidovudine-resistant RT was found to increase the mutation rate 4.3-fold. Zidovudine itself increases the mutation rate of a wild-type RT by 7.6-fold (25). Together, a zidovudine-resistant RT in the presence of zidovudine led to a 24-fold increase in the virus mutant frequency (26). Concomitantly, a rapid decay in virus

TABLE 3. Effect of decreased population size and increased mutation rate on the evolution of drug resistance[a]

N_e (effective population size)	μ (mutation rate)	Time (fixation)
10,000	3×10^{-5}	212
1,000	3×10^{-4}	174
100,000	3×10^{-5}	179
10,000	3×10^{-4}	143

[a] Increased mutation rates caused by drug treatment result in a faster evolution of drug resistance even if the effective population size is reduced by the same factor or more. Two different scenarios are shown. The following parameters are used: $w_{ab} = 1$, $w_{Ab} = w_{aB} = 1.1$, $w_{AB} = 1.21$, $f = 0.75$, and $r = 0.5$.

load over several orders of magnitude can occur during therapy (18, 35, 42). It is unclear whether the decay in virus load affects the effective population size of virus or infected cells. The estimates of N_e shown in Table 1 were measured in untreated patients as well as during therapy. The ranges suggest that the effective population size can vary by roughly 1 order of magnitude. Therefore, we investigated the effect of increased mutation rates in populations with decreased effective population sizes during the acquisition of drug resistance mutations. Table 3 shows the average time to fixation in viral generations. The results show that increased mutation rates caused by drug treatment result in a faster evolution of drug resistance even if the effective population size is reduced by the same factor or even more. These results are in agreement with theoretical studies on the waiting time to produce a double mutant (8).

DISCUSSION

The results from our simulations show that whether recombination facilitates the acquisition of drug resistance mutations depends strongly on the effective population size of the virus. Recombination will only reduce the fixation time of a drug-resistant double mutant when both single mutants are present at a sufficiently high frequency at the same time. Consistent with these results, in in vitro experiments in which equal mixtures of two single-resistance viruses are used to infect cells, recombination leads to the rapid emergence of drug-resistant double mutants (14, 21, 29). However, provided the effective population size is small, the conditions in vivo may be very different, since the single parental mutants may be rare. Indeed, Nijhuis et al. (31) have shown in their study of stochastic processes during HIV-1 evolution that recombination is not observed in all patients.

Our simulations show that preexistence of resistant mutants is unlikely if the effective population size is indeed as small as 10^3 and, in particular, when more than one nucleotide substitution is required for resistance. In this scenario, the double-resistance mutations have to be produced during drug therapy. The acquisition of these mutations then occurs sequentially rather than simultaneously, and hence the effect of recombination in reducing the time to fixation of the double mutant is negligible. The effect of recombination is substantially higher for larger effective population sizes and maximal at effective population sizes around 10^4. Here, single mutants are likely to be present together but the double mutant is likely to be absent. Recombination then generates double mutants and therefore reduces the time to fixation by up to 25%. The effect of recombination decreases with decreasing mutation rate and decreasing fraction of multiply infected cells. The effective mutation rate may be lower than the current estimate of 3.4×10^{-5} per base pair (27), since deletions, insertions, and frameshift mutations generally do not generate resistance mutations (12). Furthermore, only specific nucleotide substitutions lead to a desired mutation and some resistance mutations need more than one nucleotide substitution as mentioned above. Given that the multiplicity of infection was measured in solid tissue (20), it is possible that the multiplicity of infection averaged across all infected cells is also somewhat lower.

Drug therapy generally decreases the viral population size and therefore likely decreases also the effective population size. On the other hand it is known that for some antiretroviral drugs both the drug and the resistance mutations induce an increase in the mutation rate. We have shown here that a reduction in effective viral population size as a result of drug treatment and the consequent slower evolution of drug resistance can be overcome by higher viral mutation rates.

Epistasis has a minor influence in small effective population sizes of infected cells. Negative epistasis causes a bigger reduction than positive epistasis in the time to fixation. Interestingly, for effective population sizes bigger than 10^5, the beneficial effect of recombination is cancelled out by positive epistasis. Here, the dynamics resemble the deterministic regime and recombination will break up double mutants rather than create them.

In summary, our simulations show that the effect of recombination on the emergence of drug resistance in HIV is likely to be small if the effective population size is as low as 1,000. However, recombination can increase the rate of evolution of multidrug resistance for larger effective population sizes with a maximum around 10^4. For very large effective population sizes ($>10^5$), the effect of recombination primarily depends on epistasis, which is in agreement with results for deterministic models of recombination. To develop a better quantitative understanding of the effect of recombination on the evolution of multidrug resistance in HIV, we will need more accurate estimates of the underlying biological parameters. In particular we need to resolve the discrepancies between current estimates of the effective population size in vivo, and we need more accurate estimates for the effective population size in terms of rare beneficial or deleterious mutations.

REFERENCES

1. **Achaz, G., S. Palmer, M. Kearney, F. Maldarelli, J. W. Mellors, J. M. Coffin, and J. Wakeley.** 2004. A robust measure of HIV-1 population turnover within chronically infected individuals. Mol. Biol. Evol. **21:**1902–1912.
2. **Barton, N. H., and B. Charlesworth.** 1998. Why sex and recombination? Science **281:**1986–1990.
3. **Bonhoeffer, S., C. Chappey, N. T. Parkin, J. M. Whitcomb, and C. J. Petropoulos.** 2004. Evidence for positive epistasis in HIV-1. Science **306:**1547–1550.
4. **Bonhoeffer, S., R. M. May, G. M. Shaw, and M. A. Nowak.** 1997. Virus dynamics and drug therapy. Proc. Natl. Acad. Sci. USA **94:**6971–6976.
5. **Bonhoeffer, S., and M. A. Nowak.** 1997. Pre-existence and emergence of drug resistance in HIV-1 infection. Proc. R. Soc. Lond. B Biol. Sci. **264:**631–637.
6. **Bretscher, M. T., C. L. Althaus, V. Muller, and S. Bonhoeffer.** 2004. Recombination in HIV and the evolution of drug resistance: for better or for worse? Bioessays **26:**180–188.
7. **Burke, D. S.** 1997. Recombination in HIV: an important viral evolutionary strategy. Emerg. Infect. Dis. **3:**253–259.
8. **Christiansen, F. B., S. P. Otto, A. Bergman, and M. W. Feldman.** 1998. Waiting with and without recombination: the time to production of a double mutant. Theor. Popul. Biol. **53:**199–215.
9. **Fisher, R. A.** 1930. The genetical theory of natural selection. Clarendon Press, Oxford, United Kingdom.
10. **Frost, S. D., M. J. Dumaurier, S. Wain-Hobson, and A. J. Brown.** 2001. Genetic drift and within-host metapopulation dynamics of HIV-1 infection. Proc. Natl. Acad. Sci. USA **98:**6975–6980.
11. **Frost, S. D. W., M. Nijhuis, R. Schuurman, C. A. B. Boucher, and A. J. L. Brown.** 2000. Evolution of lamivudine resistance in human immunodeficiency virus type 1-infected individuals: the relative roles of drift and selection. J. Virol. **74:**6262–6268.
12. **Fu, Y. X.** 2001. Estimating mutation rate and generation time from longitudinal samples of DNA sequences. Mol. Biol. Evol. **18:**620–626.

13. **Funk, G. A., M. Fischer, B. Joos, M. Opravil, H. F. Gunthard, B. Ledger-ber, and S. Bonhoeffer.** 2001. Quantification of in vivo replicative capacity of HIV-1 in different compartments of infected cells. J. Acquir. Immune Defic. Syndr. **26:**397–404.

14. **Gu, Z., Q. Gao, E. A. Faust, and M. A. Wainberg.** 1995. Possible involvement of cell fusion and viral recombination in generation of human immunodeficiency virus variants that display dual resistance to AZT and 3TC. J. Gen. Virol. **76:**2601–2605.

15. **Haase, A. T., K. Henry, M. Zupancic, G. Sedgewick, R. A. Faust, H. Melroe, W. Cavert, K. Gebhard, K. Staskus, Z. Q. Zhang, P. J. Dailey, H. H. J. Balfour, A. Erice, and A. S. Perelson.** 1996. Quantitative image analysis of HIV-1 infection in lymphoid tissue. Science **274:**985–989.

16. **Harrigan, P. R., S. Bloor, and B. A. Larder.** 1998. Relative replicative fitness of zidovudine-resistant human immunodeficiency virus type 1 isolates in vitro. J. Virol. **72:**3773–3778.

17. **Hartl, D. L., and A. G. Clark.** 1997. Principles of population genetics, 3rd ed. Sinauer Associates, Sunderland, Mass.

18. **Ho, D. D., A. U. Neumann, A. S. Perelson, W. Chen, J. M. Leonard, and M. Markowitz.** 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. Nature **373:**123–126.

19. **Iles, M. M., K. Walters, and C. Cannings.** 2003. Recombination can evolve in large finite populations given selection on sufficient loci. Genetics **165:**2249–2258.

20. **Jung, A., R. Maier, J.-P. Vartanian, G. Bocharov, V. Jung, U. Fischer, E. Meese, S. Wain-Hobson, and A. Meyerhans.** 2002. Multiply infected spleen cells in HIV patients. Nature **418:**144.

21. **Kellam, P., and B. A. Larder.** 1995. Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. J. Virol. **69:**669–674.

22. **Leigh Brown, A. J.** 1997. Analysis of HIV-1 env gene sequences reveals evidence for a low effective number in the viral population. Proc. Natl. Acad. Sci. USA **94:**1862–1865.

23. **Leigh Brown, A. J., and D. D. Richman.** 1997. HIV-1: gambling on the evolution of drug resistance? Nat. Med. **3:**268–271.

24. **Mansky, L. M.** 2002. HIV mutagenesis and the evolution of antiretroviral drug resistance. Drug. Resist. Updates **5:**219–223.

25. **Mansky, L. M., and L. C. Bernard.** 2000. 3′-Azido-3′-deoxythymidine (AZT) and AZT-resistant reverse transcriptase can increase the in vivo mutation rate of human immunodeficiency virus type 1. J. Virol. **74:**9532–9539.

26. **Mansky, L. M., D. K. Pearl, and L. C. Gajary.** 2002. Combination of drugs and drug-resistant reverse transcriptase results in a multiplicative increase of human immunodeficiency virus type 1 mutant frequencies. J. Virol. **76:**9253–9259.

27. **Mansky, L. M., and H. M. Temin.** 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.

28. **Morgan, T. H.** 1913. Heredity and sex. Columbia University Press, New York, N.Y.

29. **Moutouh, L., J. Corbeil, and D. D. Richman.** 1996. Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. Proc. Natl. Acad. Sci. USA **93:**6106–6111.

30. **Muller, H. J.** 1932. Some genetic aspects of sex. Am. Nat. **66:**118–138.

31. **Nijhuis, M., C. A. Boucher, P. Schipper, T. Leitner, R. Schuurman, and J. Albert.** 1998. Stochastic processes strongly influence HIV-1 evolution during suboptimal protease-inhibitor therapy. Proc. Natl. Acad. Sci. USA **95:**14441–14446.

32. **Novick, A., and L. Szilard.** 1951. Virus strains of identical phenotype but different genotype. Science **113:**34–35.

33. **Otto, S. P., and N. H. Barton.** 2001. Selection for recombination in small populations. Evol. Int. J. Org. Evol. **55:**1921–1931.

34. **Otto, S. P., and T. Lenormand.** 2002. Resolving the paradox of sex and recombination. Nat. Rev. Genet. **3:**252–261.

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

36. **Ribeiro, R. M., and S. Bonhoeffer.** 2000. Production of resistant HIV mutants during antiretroviral therapy. Proc. Natl. Acad. Sci. USA **97:**7681–7686.

37. **Rodrigo, A. G., E. G. Shpaer, E. L. Delwart, A. K. Iversen, M. V. Gallo, J. Brojatsch, M. S. Hirsch, B. D. Walker, and J. I. Mullins.** 1999. Coalescent estimates of HIV-1 generation time in vivo. Proc. Natl. Acad. Sci. USA **96:**2187–2191.

38. **Rouzine, I. M., and J. M. Coffin.** 1999. Linkage disequilibrium test implies a large effective population number for HIV in vivo. Proc. Natl. Acad. Sci. USA **96:**10758–10763.

39. **Seo, T.-K., J. L. Thorne, M. Hasegawa, and H. Kishino.** 2002. Estimation of effective population size of HIV-1 within a host: a pseudomaximum-likelihood approach. Genetics **160:**1283–1293.

40. **Shriner, D., R. Shankarappa, M. A. Jensen, D. C. Nickle, J. E. Mittler, J. B. Margolick, and J. I. Mullins.** 2004. Influence of random genetic drift on human immunodeficiency virus type 1 env evolution during chronic infection. Genetics **166:**1155–1164.

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

42. **Wei, X., S. K. Ghosh, M. E. Taylor, V. A. Johnson, E. A. Emini, P. Deutsch, J. D. Lifson, S. Bonhoeffer, M. A. Nowak, and B. H. Hahn.** 1995. Viral dynamics in human immunodeficiency virus type 1 infection. Nature **373:**117–122.