

Recombination favors the evolution of drug resistance in HIV-1 during antiretroviral therapy

Antonio Carvajal-Rodríguez^{a,b,*}, Keith A. Crandall^a, David Posada^b

^aDepartment of Microbiology and Molecular Biology, Brigham Young University, 84602 Provo, UT, USA

^bDepartamento de Bioquímica, Genética e Inmunología. Universidad de Vigo, 36310 Vigo, Spain

Received 16 October 2006; received in revised form 5 February 2007; accepted 7 February 2007

Available online 12 February 2007

Abstract

We studied the relationship between recombination and the fixation time of multiple drug resistance mutations after HIV-1 drug therapy, under a set of different realistic scenarios. We have generalized a previous model by Bretscher et al. [Bretscher, Althaus, Muller, Bonhoeffer, 2004. Recombination in HIV and the evolution of drug resistance: for better or for worse? *Bioessays* 26(2), 180–188] in order to explore different implementations of phenotypic mixing and more realistic demographic and selective regimes. Using computer simulations we show that the effect of recombination on the evolution of drug resistance depends strongly on the intensity of selection, as well as on the viral population size. Under the high selection pressure expected during antiretroviral therapy, the strength of the Hill–Robertson effect increases and recombination favors the evolution of resistance under a wide range of population sizes, independently of the sign of the epistatic interaction. Our results suggest that recombination plays an important role in the evolution of drug resistance in HIV-1 under various realistic scenarios. These findings could be taken into account in order to develop optimal HIV-1 drug treatments.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Recombination; Drug resistance; Hill–Robertson effect; Epistasis; HIV-1

1. Introduction

Several empirical studies have related recombination with the development of HIV-1 resistance to drug therapy (Gu et al., 1995; Kellam and Larder, 1995; McDowell et al., 1998; Moutouh et al., 1996). Indeed, population genetics theory predicts that recombination can favor adaptation in finite populations. The Hill–Robertson effect (HRE) (Hill and Robertson, 1966) implies that in finite populations negative disequilibrium accumulates on average (i.e., beneficial and deleterious alleles are found together on the same chromosome more often than expected given their frequencies) because of drift and mutation, reducing the probability of fixation of beneficial alleles, and therefore, the efficacy of selection (Felsenstein, 1974; Fisher, 1930; Muller, 1964). Indeed, the accumulation of negative disequilibrium not only occurs because of the stochastic nature of mutation and sampling in

finite populations, but can also take place when fully susceptible genotypes are less fit than expected from the product of the fitness at each locus (i.e., negative epistasis). Importantly, genetic recombination reduces the strength of the HRE, increasing the rate of fixation of beneficial alleles (McVean and Charlesworth, 2000; Roze and Barton, 2006). In a multilocus system under negative linkage disequilibrium, recombination breaks up the unpaired combinations of resistant and susceptible alleles, generating both fully resistant genotypes, that will spread rapidly within the population, and fully susceptible genotypes, that will be eliminated fairly quickly. Therefore, recombination should favor the emergence of multidrug resistance, that is, the accumulation of multiple mutations that lead to resistance to one or several drugs.

While specific mathematical models have suggested that recombination is an important mechanism for the generation of HIV-1 multidrug resistance (Rouzine and Coffin, 2005), recently, Bretscher et al. (2004) and Althaus and Bonhoeffer (2005) disputed the generality of this result. These authors used computer simulations to study the interaction of mutation, selection, effective population size (the number of individuals

* Corresponding author. Tel.: +34 986 813828; fax: +34 986 812556.

E-mail address: acraaj@uvigo.es (A. Carvajal-Rodríguez).

in a population who contribute offspring to the next generation) and epistasis on the evolution of multidrug resistance in HIV-1. They concluded that while recombination favors the acquisition of multidrug resistance for low and (especially for) intermediate effective population sizes regardless of epistasis, and for large effective population sizes ($>10^5$) when epistasis is negative, recombination severely impairs the development of multidrug resistance for large effective population sizes ($>10^5$) when epistasis is positive.

This last result might be very relevant because estimates of HIV-1 effective population size range from 500 to 10^5 (Althaus and Bonhoeffer, 2005; Rouzine and Coffin, 1999), and epistasis in HIV-1 might be positive, at least in the absence of drug therapy (Bonhoeffer et al., 2004; but see Wang et al., 2006). However, we believe that these results depend on the particular model implemented, which we can relax in order to provide a more realistic view of HIV-1 evolution. First of all, the selection coefficient used by Althaus and Bonhoeffer (2005) with positive epistasis was very low ($s = 0.1$), considering that under drug therapy we expect a high selection pressure (Finzi et al., 1999; Ribeiro et al., 1998; Rouzine and Coffin, 2005). Second, they used a simple function for phenotype mixing (when two proviruses are transcribed simultaneously within the same infected cell and the newly assembled virions contain mixed proteins from both parental proviruses), assuming that the fitness of a virion in a double-infected cell is always the arithmetic mean of the fitness of both parental proviruses. However, it seems more realistic to incorporate some uncertainty in this process. Third, they used a constant effective population size, while the effective population size of HIV-1 varies throughout infection, and especially under highly active antiretroviral therapy, where the population size can drop more than one order of magnitude (Rodrigo, 1999; Rouzine and Coffin, 1999; Rouzine et al., 2001).

In addition, Fraser (2005) has recently extended the deterministic model of Bretscher et al. (2004) to incorporate viral dynamics, concluding that the impact of recombination on the evolution depends on the frequency of co-infection: if co-infection falls rapidly with viral load, then recombination should have no effect, whereas if the frequency of co-infection is constant (as in Bretscher et al., 2004), then recombination slows down the emergence of drug resistance. Note that these results contradict those by Bretscher et al. (2004).

Here, and in order to better understand the role of recombination on the evolution of HIV-1 drug resistance, we have generalized the deterministic and stochastic models in (Althaus and Bonhoeffer, 2005; Bretscher et al., 2004) to include more realistic conditions. We explore different implementations of phenotypic mixing and higher selection coefficients, under infinite and finite population sizes, with constant or variable demographics. We demonstrate that the accumulation of negative linkage disequilibrium due to a strong HRE explains how recombination facilitates the evolution of drug resistance during drug therapy in HIV-1 regardless epistasis. Finally, we also study the role of pre-existent resistant mutants, which should be an important factor for the acquisition of drug resistance (Ribeiro and Bonhoeffer, 2000).

2. Materials and methods

The models implemented in this study are based on a two-locus two-allele deterministic model in Bretscher et al. (2004). We will first describe the basic features of this model, and then explain in detail our extensions.

2.1. Deterministic model

We have four types of proviruses, *ab*, *aB*, *Ab* and *AB*, where lowercase denotes drug-sensitive wild-type alleles and uppercase indicates drug-resistant mutant alleles. Under drug therapy, provirus type *ab* is fully sensitive, *aB* and *Ab* are partially drug resistant, and *AB* is fully resistant. The change in provirus frequencies during the replication cycle is best described by dividing this cycle in three steps: cell infection, virion release, and provirus transcription.

2.1.1. Infection of cells

It is assumed that all cells will be infected by either one or two proviruses. Let f be the frequency of cells infected with two proviruses (i.e., f is the frequency of co-infection). The frequency (C) of a cell infected with a single provirus type is the product of the frequency (P) of the corresponding provirus type and the probability of single infection ($1 - f$). For example, $C_{ab} = (1 - f) \times P_{ab}$. Similarly, the frequency of double-infected cells is given by the product of the frequencies of both proviruses and the probability of double infection (f). For example, $C_{abAB} = 2 \times f \times P_{ab} \times P_{AB}$. With four provirus types, there are four types of single-infected and 10 types of double-infected cells.

2.1.2. Release of virions

The frequency of homozygous and heterozygous virions released by the cell population is determined from the cell frequencies. It is assumed that both single- and double-infected cells release the same amount of virions, and that released virions infect target cells depending on their fitness. For example, the frequency of the infecting *ABAB* virion is given by $V_{ABAB} = [1/V] [W_{AB} (C_{AB} + C_{ABAB})] + 0.25 [g(W_{AB}, W_{Ab}) C_{ABAb} + g(W_{AB}, W_{aB}) C_{ABaB} + g(W_{Ab}, W_{ab}) C_{ABab}]$, where W_{AB} , W_{Ab} , W_{aB} , W_{ab} are the fitness values of the corresponding proviruses, C_{AB} , C_{ABAB} , C_{ABAb} , C_{ABaB} and C_{ABab} are the corresponding single or double-infected cell frequencies, g is a function that will depend on the phenotypic mixing implementation (see below), and V is a normalization factor to ensure that all infecting virion frequencies sum to 1. Provirus fitness during therapy is computed under a multiplicative model in which the wild-type has fitness $1 - s$. Genotypes carrying one mutant allele have fitness $(1 - s - E)^{1/2}$ and the double mutant has fitness 1. E is the epistatic value.

2.1.3. Transcription of inserted virions to proviruses

After all released virions have infected a target cell, the frequencies of proviruses in the next generation are computed from that of virions as a function of the mutation rate μ and the recombination rate r . Mutation is considered only during the reverse transcription process.

2.2. Extensions to the basic model

2.2.1. Phenotype mixing

The fitness of emerging virions depends on both the parental provirus fitness and the cellular type. This separation of phenotype and genotype is known as *phenotypic mixing* (Brenner, 1957; Novick and Szilard, 1951). In real life, phenotypic mixing occurs when two proviruses that are being transcribed simultaneously within the same infected cell mix their corresponding viral proteins in the assembly of new virions. It is possible then that the fitness of a released virion does not necessarily reflect the genomic RNA that it carries, but that it is rather somewhere in between that of both parental proviruses. The specific implementation of the phenotypic mixing in our simulations is described by the function g . In the *Arithmetic* model, which is exactly the same as the one described by Bretscher et al. (2004), g is the arithmetic mean. In the *Geometric* model virion fitness is the geometric mean of the fitness of both parental proviruses, implying that, in the absence of epistasis, the fitness of virions emerging from double-infected cells with provirus AB and ab will be the same as the fitness of virions emerging from double-infected cells with provirus Ab and aB . Therefore, this model is unique in that there are no fitness differences between the two fully resistant double mutants AB/ab and Ab/aB . In the *Minimum* and *Maximum* models, virion fitness is computed simply as the minimum or the maximum, respectively, of the two parental proviruses. In the *Random* model, the most realistic model considered here, we assume that phenotypic mixing does not always occur. Instead, we consider four types of possible events, that the virion inherits the viral proteins and the genomic RNA from either parent, or that the virion inherits the viral proteins from one parent and the genomic RNA from the other, with equal probability for each type of event. Finally, if there is no phenotypic mixing then the fitness of the virion is computed assuming a multiplicative model. In this case, the fitness of the double homozygote virion $ABAB$ is the product $W_{AB} \times W_{AB}$. The same scheme applies to all provirus genotypes.

2.3. Stochastic implementation

We introduce stochasticity by considering that different cell types can contribute a variable amount of virions in each generation, and that there is a finite amount of infecting virions (N , the effective population size). First, we select a cell type according to its frequency. Second, a virion is released by the selected cell type contingent on its probability of being produced. Third, this virion will survive and become infecting depending on its fitness. For example, the probability of selecting a cell co-infected by proviruses AB and ab is C_{ABab} . From this cell we release a virion $ABAB$ with probability $1/4$ that it will survive and become infecting with probability $g(W_{AB}, W_{ab})$. We repeat this process until N released virions have survived until infection.

In addition, because drug therapy affects viral load, we also consider a model with variable effective population size. In this case, we repeat the first two steps above until N_0 virions have

been released. From these, N_t ($\leq N_0$) virions will survive and become infecting depending on their fitness. The expected number of infecting virions at a given generation will be simply the sum of the fitnesses of the N_0 previously released virions. As a consequence, N_t will vary across generations, but always below N_0 . N_0 is a constant specified at the beginning of the simulation that represents the maximum number of virions released by the infected cells at any generation.

2.4. Computer simulations

To determine how recombination affects the rate at which multidrug resistance evolves during drug therapy, we started the simulations from a population consisting only of wild-type proviruses and measured the number of generations until 90% of the provirus population was fully resistant (*fixation time*). In the deterministic simulations, the double infection rate per cell per generation (f) was 0.3 or 0.75 (Jung et al., 2002), the mutation rate (μ) was 3×10^{-5} mutations per locus per generation (Mansky and Temin, 1995), the selection coefficient (s) was 0.1, 0.5 or 0.9 (Finzi et al., 1999; Ribeiro et al., 1998; Rouzine and Coffin, 2005), epistasis values (E) were -0.05 , 0.0, or 0.05 (Althaus and Bonhoeffer, 2005; Bretscher et al., 2004), and the recombination rate (r) was 0.0, 0.01, 0.1, 0.2, 0.3, 0.4, or 0.5 recombination events per locus per generation (Shriner et al., 2004). In addition, we ran some cases with $s = 0.01, 0.2, 0.3, 0.4, 0.6, 0.7$ and 0.8. We also studied the effect of recombination on the fixation time when single resistant mutations pre-exist at a frequency of 0.1%. In the stochastic simulations, we used the same parameter values as above, except that we only explored $r = 0.0$ and 0.1, plus a set of distinct effective population sizes (N or N_0): $10^2, 10^3, 10^4, 10^5$ and 10^6 . We carried out 1000 replicates for each set of conditions, but for $N = 10^6$ we only ran 100. The number of replicates was always enough to maintain standard errors below 1%. To test our stochastic implementation, we ran an asymptotic case with $N = 10^8$, which should mimic the behavior of the deterministic model. We measured the recombination effect on the evolution of drug resistance as the percent reduction on fixation time for a specific recombination rate compared to the case of no recombination, or $R_e = 100(1 - T_r/T_0)$, where T_r is the fixation time with recombination rate r and T_0 is the fixation time without recombination. Thus, R_e will be positive when recombination reduces fixation time and negative when recombination increases it. Simulations were implemented with a computer program written in C. A copy of the source code is freely available upon request from the authors.

3. Results

3.1. Deterministic evolution of multidrug resistance

Without epistasis, $s = 0.1$, and with 75% of cells doubly infected, the effect of recombination during therapy varied depending on the phenotypic mixing implementation (Fig. 1). Under the *Minimum* model recombination reduced the fixation

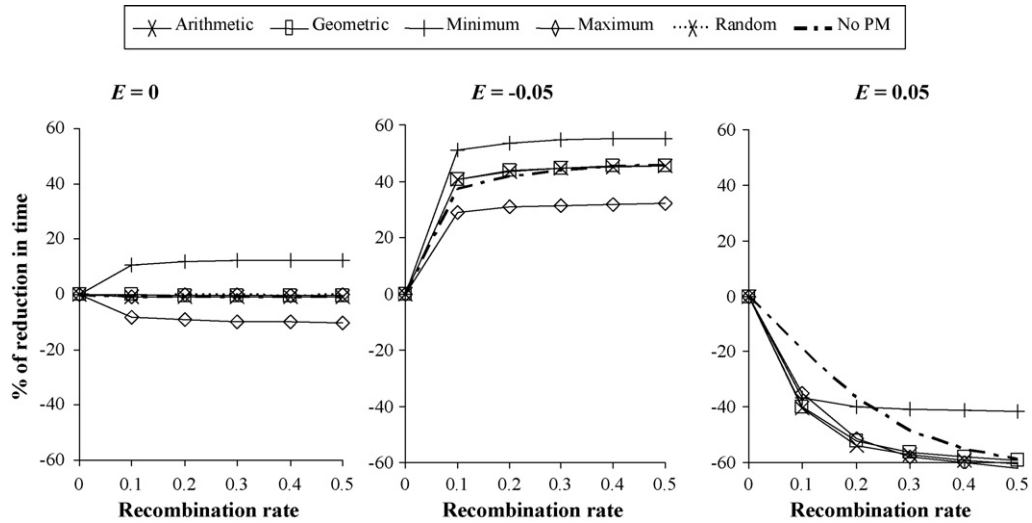


Fig. 1. Effect of recombination on the evolution of multidrug resistance in the deterministic case. The y-axis indicates the % of reduction of fixation time produced by a given recombination rate with respect to the case of no recombination. Here, $f = 0.75$ and $s = 0.1$. E : epistasis; PM: phenotypic mixing.

time, while for the *Maximum* model the effect was the opposite. When epistasis was negative, recombination significantly reduced the fixation time of the double-resistant proviruses regardless of the implementation of phenotypic mixing, especially for the *Minimum* model. When epistasis was positive, recombination significantly increased the fixation

time for all models, although this effect was smaller for the *Minimum* model or in the absence of phenotypic mixing. Other values of the selection coefficient resulted in the same trends (not shown), although smaller selection coefficient values implied longer times to fixation and vice versa). Importantly, when the frequency of double-infected cell was low ($f = 0.3$),

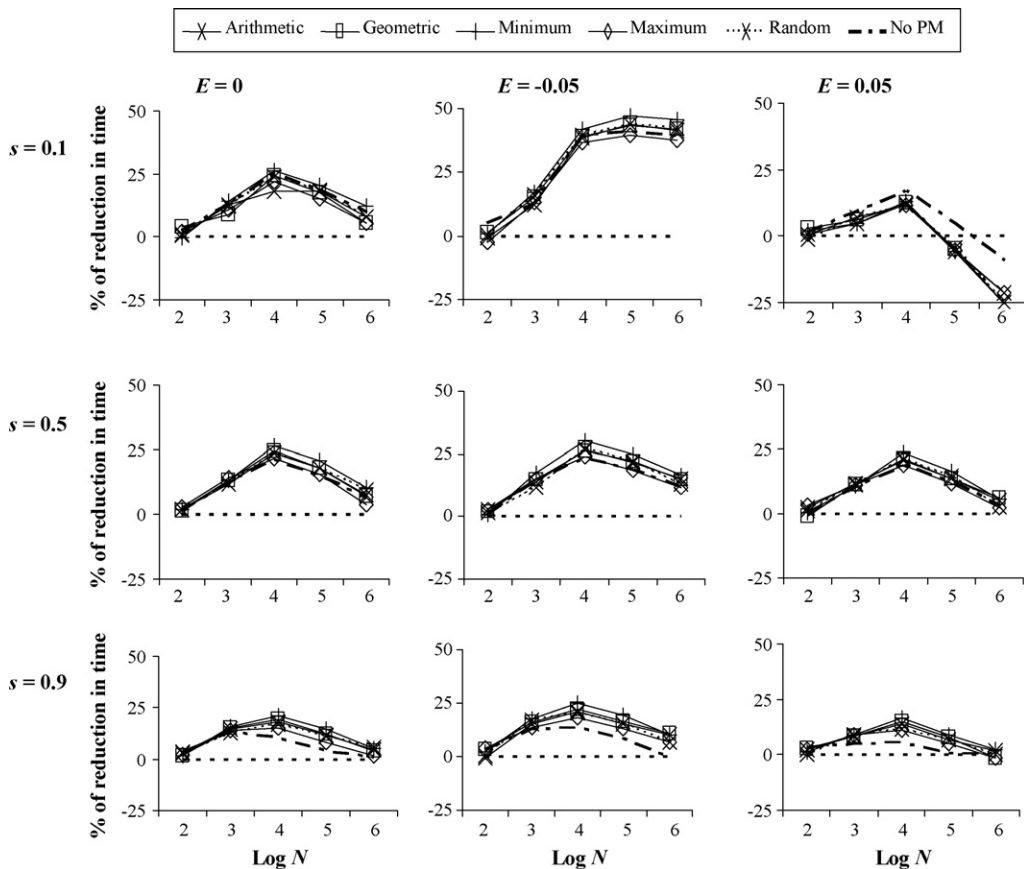


Fig. 2. Effect of recombination on the evolution multidrug resistance with constant population size. The y-axis indicates the % of reduction onto fixation time under a recombination rate of 0.1 with respect to the case of no recombination. Here, $f = 0.75$. E : epistasis; s : selection coefficient; $\text{log } N$: logarithm of the effective population size.

the results obtained were similar (not shown), although the differences between the distinct phenotypic mixing implementations were weaker.

3.2. Stochastic evolution of multidrug resistance

The introduction of a finite effective population size had a remarkable effect. When population size was constant and selection pressure was low ($s = 0.1$) (Fig. 2, first row) with zero or negative epistasis, recombination decreased the fixation time for a wide range of population sizes ($N = 10^2 - 10^6$) with a maximum at intermediate sizes. When epistasis was positive, recombination slightly decreased the fixation time for small and intermediate population sizes, but increased it when $N > 10^4$. Remarkably, with higher selection pressures ($s = 0.5$ or 0.9) (Fig. 2, second and last rows), recombination always decreased the fixation time regardless of the sign of the epistatic interaction, especially when $N = 10^4$. Phenotypic mixing had practically no effect. When we used a very large effective population size ($N = 10^8$), results matched exactly those obtained under the deterministic model (not shown), therefore validating the implementation of the stochastic model.

With variable population size (Fig. 3), results were very similar to the constant case, although it seemed that the largest decreases in fixation time were obtained under slightly larger

population sizes, especially for high selective pressures. Again, phenotypic mixing had practically no effect.

3.3. Pre-existence of resistant mutations

In the deterministic case, the presence of resistant mutations (0.1%) before drug therapy had no effect compared to the non-pre-existence case (results not shown). In contrast, with finite population sizes the effect of recombination was magnified with respect to non-pre-existence (Fig. 4), but with similar trends. Under strong selection, demographics seemed to play a significant role.

4. Discussion

4.1. Deterministic versus stochastic models

Only under a deterministic model is negative epistasis needed for recombination to accelerate fixation time, as previously shown by Bretscher et al. (2004). This is true unless phenotypic mixing always confers the virion the fitness of the weakest parent, which does not seem very plausible. Although in the study of HIV-1 dynamics a general controversy exists regarding whether deterministic or stochastic models should be considered (Garnett, 2002; Leigh Brown, 1997; Rouzine and Coffin, 1999;

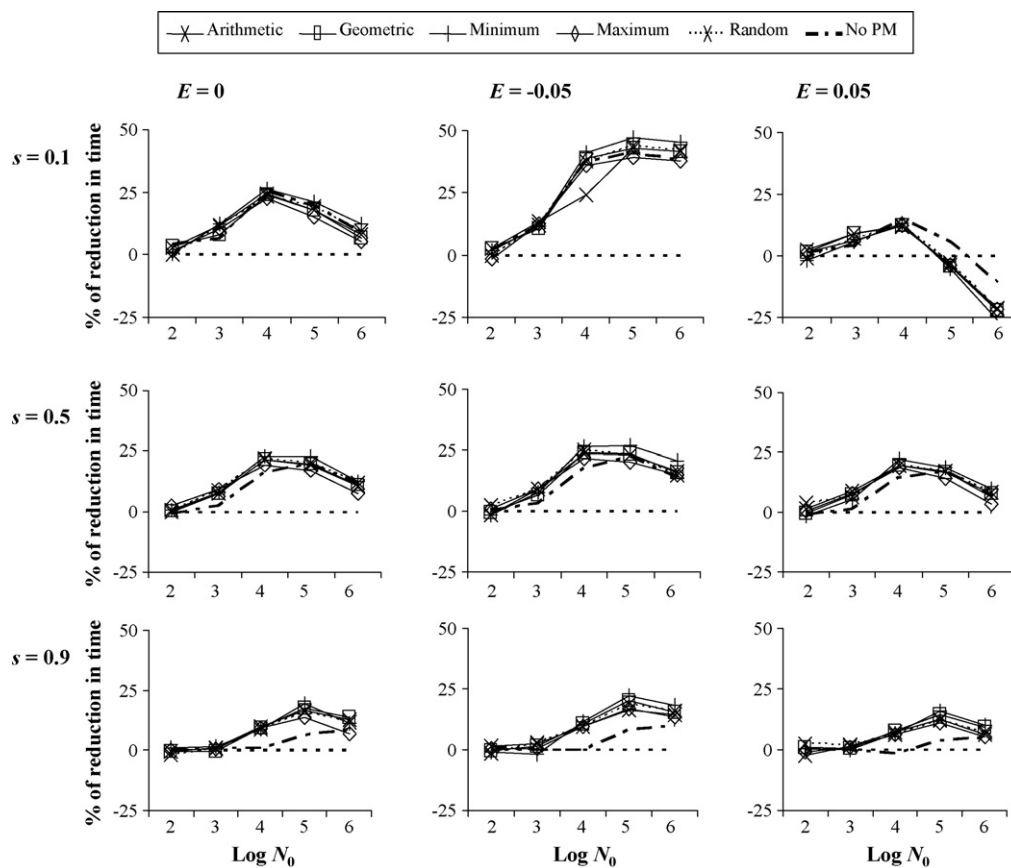


Fig. 3. Effect of recombination on the evolution multidrug resistance with variable population size. The y-axis indicates the % of reduction onto fixation time under a recombination rate of 0.1 with respect to the case of no recombination. Here, $f = 0.75$. E : epistasis; s : selection coefficient; $\log N_0$: logarithm of the initial effective population size.

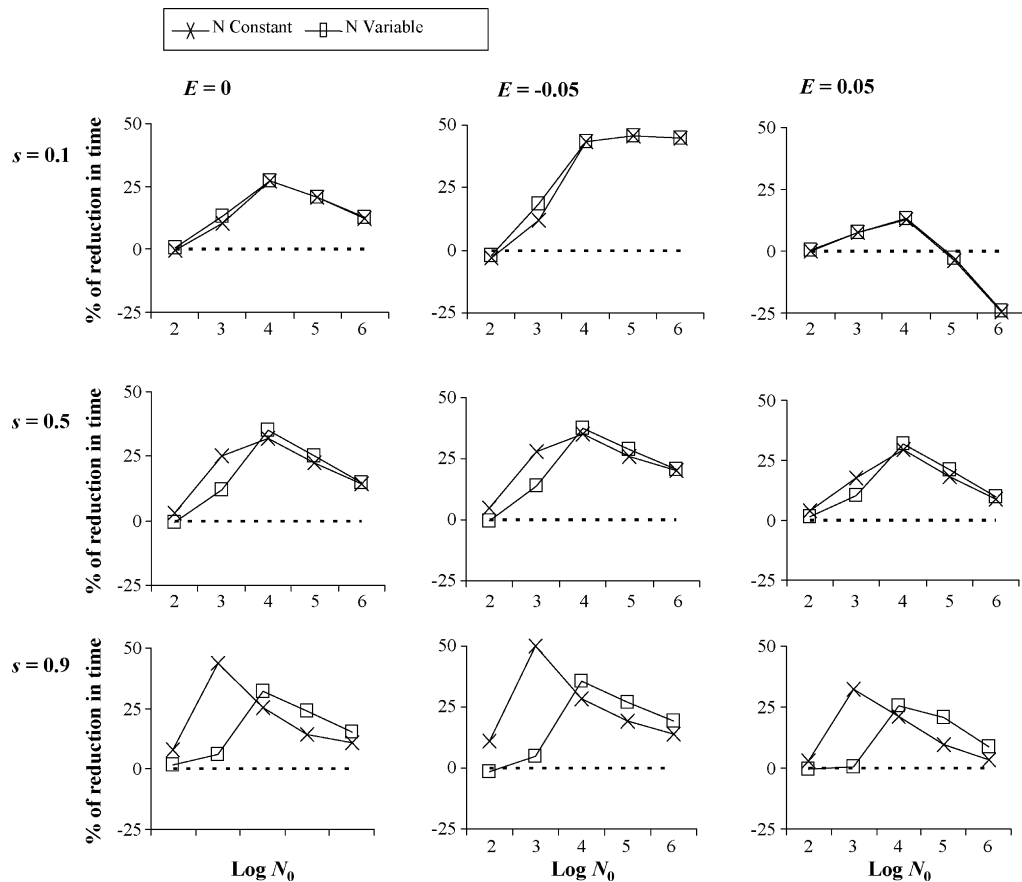


Fig. 4. Effect of recombination on the evolution multidrug resistance when resistant mutants exist before drug therapy (pre-existence). The y-axis indicates the % of reduction onto fixation time under a recombination rate of 0.1 with respect to the case of no recombination. Here, the initial frequency of mutants is 0.1%, phenotypic mixing is *Random*, and $f = 0.75$; E : epistasis; s : selection coefficient; $\log N_0$: logarithm of the initial effective population size.

Tan and Wu, 1998), we consider that the evolution of HIV-1 is better described by a stochastic process because low effective population sizes are expected due to compartmentalization (Nickle et al., 2003), to variability in progeny number between infected cells, to population expansion from a small initial inoculum, and from variable selective pressures (Leigh Brown, 1997). Furthermore, stochasticity is especially important under antiretroviral therapy, when the number of infected cells can drop by one or more orders of magnitude (Ho et al., 1995; Rodrigo, 1999; Wei et al., 1995).

From a theoretical point of view, the effect of recombination on the evolution of drug resistance can be formally explained by analyzing the models in terms of linkage disequilibrium (D) and the phenotypic mixing effect (see Appendix A, for details). A key difference between the deterministic and the stochastic situation is that in the former the frequency of a mutant will never go to zero. Therefore, an insignificant fraction of double resistant provirus could persist and subsequently be selected until fixation. However, in the stochastic models new mutations are often lost quickly by chance. This will generate random deviations, positive or negative, from linkage equilibrium. In the absence of selection, these deviations cancel out and D averaged across generations (or loci) equals 0. However, in the presence of selection, positive deviations from linkage equilibrium will be rapidly fixed or eliminated. On the other hand, negative D deviations are not so quickly solved by selection, because of the

intermediate fitness of single mutants (a fitter allele in a less fit background). This is known as the Hill–Robertson effect (HRE) (Barton and Otto, 2005; Hill and Robertson, 1966). The accumulation of negative disequilibrium can be significant even in fairly large populations as long as the allele frequencies are initially low and subject to drift (Barton and Otto, 2005). Thus, the HRE predicts that genetic drift will interact with directional selection to produce negative D . Such negative D is all that recombination needs to favor beneficial mutations, and, as predicted by Eq. (A1) (see Appendix A), it will allow recombination to speed up the emergence of multidrug resistance. Furthermore, the form of epistasis is not critical to the advantage of recombination in finite populations, as recently suggested (Keightley and Otto, 2006), if selection is strong.

4.2. Phenotypic mixing

Bretscher et al. (2004) used the average of the parental fitnesses to include phenotypic mixing in their model. It seemed reasonable to think the outcome of phenotypic mixing might not always be the same due to the stochasticity of this process. However, in our simulations the particular implementation of phenotypic mixing had only a clear effect in the deterministic case, and even in this case only the two extreme models (*Minimum* and *Maximum*) departed from the average behavior. In fact, results under the most realistic *Random* model were

very similar to those obtained under the *Arithmetic* model of Bretscher et al. (2004). That the particular implementation of phenotypic mixing did not influence the results of the stochastic simulations is important, and makes the biological interpretation of the simulations more straightforward.

4.3. The frequency of co-infection

Importantly, the specific frequency f of double-infected cells did not change the results.

Although a smaller f resulted in a reduction of the effective recombination rate, recombination still favored drug resistance. It is very important to note that a reduction of the recombination rate does not necessarily reduce its effect on the fixation time, unless this rate is 0 (i.e., the effect of $r = 0.5$ is very similar to that of $r = 0.1$). Our results contrast to those obtained by Fraser (2005), who claimed that with low co-infection rates recombination does not have an effect on drug resistance. However, such results were obtained using a deterministic model under low selective pressures, which is not a realistic scenario under drug therapy.

4.4. Pre-existence of resistant mutations

Pre-existence of mutations had an effect on the role of recombination independently of epistasis, especially with the stochastic models. In particular, pre-existent mutations magnified the effect of recombination reducing the fixation time of double mutants, and therefore favor the evolution of drug resistance. This is because the pre-existence of single resistant mutants generates an initial negative linkage disequilibrium that adds to that already generated by the HRE. Under strong selection, the different meaning of N_0 under constant and variable population sizes becomes evident. In the constant case $N_0 = 10^3$ represents the optimal population size for negative disequilibrium to accumulate in the population under this model. In the variable case $N_0 = 10^4$ represents the optimal maximum population size, which under high selective pressure will be seldom reached.

These results agree well with a previous study that showed that even a very modest amount of recombination is more efficient at generating highly fit viral genomes than mutation alone, provided the pre-existence of the more fit alleles (Rouzine and Coffin, 2005). This result is important because it suggests that if resistant mutations appear before drug therapy the fixation of resistance in the viral population could be very rapid. In addition, if transmission of drug resistance mutations increases in a population, the effect of pre-existent mutations will be even more relevant.

4.5. The effect of recombination on the evolution of drug resistance

We have shown that the effect of recombination on the evolution of multidrug resistance depends strongly not only on the viral population size but also on the intensity of selection. With higher selection pressure, the importance of the HRE prevails over epistasis, and recombination always diminishes the fixation time of resistance under a wide range of population

sizes. Indeed, we expect selection to be very strong during antiretroviral therapy (Finzi et al., 1999; Ribeiro et al., 1998; Rouzine and Coffin, 2005). In this way, we reconcile HIV-1 evolution with previous evidence (Gu et al., 1995; Kellam and Larder, 1995; McDowell et al., 1998; Moutouh et al., 1996; Rouzine and Coffin, 2005) and with the common belief regarding the adaptive role of recombination (Felsenstein, 1974; Fisher, 1930; Hill and Robertson, 1966; Keightley and Otto, 2006; McVean and Charlesworth, 2000; Muller, 1964; Roze and Barton, 2006). Our results suggest that, everything else being equal, a combination of drugs that target close regions of the HIV-1 genome will be less susceptible to failure due to the emergence of drug resistance because of a reduced recombination rate between target loci.

Acknowledgments

We thank Armando Caballero for very useful comments on a previous version of this manuscript. This work was supported by grant R01-GM66276 from the US National Institutes of Health (AC-R, KAC, DP), grant BFU2004-02700 of the Spanish Ministry of Education and Science (DP) and by the “Ramón y Cajal” programme of the Spanish government (DP). AC-R is currently funded by an Isidro Parga Pondal research fellowship from Xunta de Galicia (Spain).

Appendix A

Results here and from previous work (Althaus and Bonhoeffer, 2005; Bretscher et al., 2004) can be explained by analyzing the models in terms of linkage disequilibrium (D) and the phenotypic mixing effect (δ).

Linkage disequilibrium in generation t , $D(t)$, can be defined as the difference between the product of frequencies of coupled and uncoupled provirus genotypes: $D(t) = P_{AB}(t)P_{ab}(t) - P_{Ab}(t)P_{aB}(t)$. After equilibrium is reached, the final D value accumulated through generations is D_{eq} .

The fitness of each virion is computed assuming a phenotypic mixing function g such that $g(W_{AB}, W_{ab})$ is the fitness of virions coming from a cell co-infected with genotypes AB and ab , and $g(W_{Ab}, W_{aB})$ is the fitness of virions coming from a cell co-infected with genotypes Ab and aB . In the absence of phenotypic mixing, this function represents directly the fitness of a double heterozygous virion.

The effect of recombination at generation t , on the frequency of double mutant provirus can be expressed as:

$$P_{AB}(t) = X(t-1) - Yf[D(t-1)g(W_{Ab}, W_{aB}) + P_{AB}(t-1)P_{ab}(t-1)\delta] \quad (A1)$$

where $X(t-1)$ is a long term that includes cell frequencies, provirus fitnesses and mutation rate at generation $t-1$ (we will ignore it because it is independent of the recombination rate r), f is the frequency of double-infected cells, and $Y = (1 - 2\mu)^2 r / 2$, and $\delta = g(W_{AB}, W_{ab}) - g(W_{Ab}, W_{aB})$.

Results obtained in the simulations can be explained upon inspection of Eq. (A1). Whenever $D_{eq} < 0$ or $\delta < 0$,

recombination will increase the double mutant frequency, decreasing fixation time (the time until 90% of the provirus population is fully resistant). The opposite will be true when $D_{eq} > 0$ or $\delta > 0$. Moreover, negative epistasis will decrease δ and positive epistasis will increase it.

In the deterministic case, and in the absence of epistasis, $D_{eq} = 0$ (Felsenstein, 1965; Lewontin, 1974), and the only possible effect is associated with δ . If $\delta < 0$ (Minimum model), recombination will reduce fixation time, while the opposite will be true if $\delta > 0$ (Maximum model). If there is epistasis, D_{eq} will have the same sign as the epistatic interaction (Eshel and Feldman, 1970; Felsenstein, 1965). Because of this, negative epistasis will reduce fixation time, while positive epistasis will increase it.

When population size is finite, or when mutations pre-exist, D_{eq} is negative under strong selection because of the HRE, even with positive epistasis. Note that the effect of δ will be less important because the magnitude of the multiplying factor in Eq. (A1) is much smaller for δ than for D . Thus, under the simulated conditions of antiretroviral therapy, i.e., finite population size, strong selection pressure and null or low initial frequency of beneficial alleles, negative linkage disequilibrium will accumulate, providing the context for recombination to favor the emergence of drug resistance.

References

- Althaus, C.L., Bonhoeffer, S., 2005. Stochastic interplay between mutation and recombination during the acquisition of drug resistance mutations in human immunodeficiency virus type 1. *J. Virol.* 79 (21), 13572–13578.
- Barton, N.H., Otto, S.P., 2005. Evolution of recombination due to random drift. *Genetics* 169 (4), 2353–2370.
- Bonhoeffer, S., Chappey, C., Parkin, N.T., Whitcomb, J.M., Petropoulos, C.J., 2004. Evidence for positive epistasis in HIV-1. *Science* 306 (5701), 1547–1550.
- Brenner, S., 1957. Genetic control and phenotypic mixing of the adsorption cofactor requirement in bacteriophage-T2 and bacteriophage-T4. *Virology* 3 (3), 560–574.
- Bretscher, M.T., Althaus, C.L., Muller, V., Bonhoeffer, S., 2004. Recombination in HIV and the evolution of drug resistance: for better or for worse? *Bioessays* 26 (2), 180–188.
- Eshel, I., Feldman, M.W., 1970. On the evolutionary effect of recombination. *Theor. Popul. Biol.* 1 (1), 88–100.
- Felsenstein, J., 1965. Effect of linkage on directional selection. *Genetics* 52 (2P1), 349–363.
- Felsenstein, J., 1974. The evolutionary advantage of recombination. *Genetics* 78 (2), 737–756.
- Finzi, D., Blankson, J., Siliciano, J.D., Margolick, J.B., Chadwick, K., Pierson, T., Smith, K., Lisziewicz, J., Lori, F., Flexner, C., Quinn, T.C., Chaisson, R.E., Rosenberg, E., Walker, B., Gange, S., Gallant, J., Siliciano, R.F., 1999. Latent infection of CD4(+) T cells provides a mechanism for lifelong persistence of HIV-1, even in patients on effective combination therapy. *Nat. Med.* 5 (5), 512–517.
- Fisher, R.A., 1930. *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
- Fraser, C., 2005. HIV recombination: what is the impact on antiretroviral therapy? *J.R. Soc. Interface* 2 (5), 489–503.
- Garnett, G.P., 2002. An introduction to mathematical models in sexually transmitted disease epidemiology. *Sexually Transmitted Infect.* 78 (1), 7–12.
- Gu, Z., Gao, Q., Faust, E.A., Wainberg, M.A., 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 (10), 2601–2605.
- Hill, W.G., Robertson, A., 1966. The effect of linkage on limits to artificial selection. *Genet. Res.* 8 (3), 269–294.
- Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M., 1995. Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* 373, 123–126.
- Jung, A., Maier, R., Vartanian, J.P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S., Meyerhans, A., 2002. Multiply infected spleen cells in HIV patients. *Nature* 418 (6894), 144.
- Keightley, P.D., Otto, S.P., 2006. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* 443 (7107), 89–92.
- Kellam, P., Larder, B.A., 1995. Retroviral recombination can lead to linkage of reverse transcriptase mutations that confer increased zidovudine resistance. *J. Virol.* 69 (2), 669–674.
- 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. U.S.A.* 94, 1862–1865.
- Lewontin, R.C., 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.
- Mansky, L.M., Temin, H.M., 1995. Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol.* 69 (8), 5087–5094.
- McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G., Goff, S., Holub, E.B., Dangi, J.L., 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell* 10 (11), 1861–1874.
- McVean, G.A., Charlesworth, B., 2000. The effects of Hill–Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. *Genetics* 155 (2), 929–944.
- Moutouh, L., Corbeil, J., Richman, D.D., 1996. Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc. Natl. Acad. Sci. U.S.A.* 93 (12), 6106–6111.
- Muller, H.J., 1964. The relation of recombination to mutational advance. *Mutat. Res.* 106, 2–9.
- Nickle, D.C., Jensen, M.A., Shriner, D., Brodie, S.J., Frenkel, L.M., Mittler, J.E., Mullins, J.I., 2003. Evolutionary indicators of human immunodeficiency virus type 1 reservoirs and compartments. *J. Virol.* 77 (9), 5540–5546.
- Novick, A., Szilard, L., 1951. Virus strains of identical phenotype but different genotype. *Science* 113 (2924), 34–35.
- Ribeiro, R.M., Bonhoeffer, S., 2000. Production of resistant HIV mutants during antiretroviral therapy. *Proc. Natl. Acad. Sci. U.S.A.* 97 (14), 7681–7686.
- Ribeiro, R.M., Bonhoeffer, S., Nowak, M.A., 1998. The frequency of resistant mutant virus before antiviral therapy. *AIDS* 12 (5), 461–465.
- Rodrigo, A., 1999. HIV evolutionary genetics. *Proc. Natl. Acad. Sci.* 96, 10559–10561.
- Rouzine, I.M., Coffin, J.M., 1999. Linkage disequilibrium test implies a large effective population number for HIV *in vivo*. *Proc. Natl. Acad. Sci.* 96, 10758–10763.
- Rouzine, I.M., Coffin, J.M., 2005. Evolution of human immunodeficiency virus under selection and weak recombination. *Genetics* 170 (1), 7–18.
- Rouzine, I.M., Rodrigo, A., Coffin, J.M., 2001. Transition between stochastic evolution and deterministic evolution in the presence of selection: general theory and application to virology. *Microbiol. Mol. Biol. Rev.* 65 (1), 151–184.
- Roze, D., Barton, N.H., 2006. The Hill–Robertson effect and the evolution of recombination. *Genetics* 173 (3), 1793–1811.
- Shriner, D., Rodrigo, A.G., Nickle, D.C., Mullins, J.I., 2004. Pervasive genomic recombination of HIV-1 *in vivo*. *Genetics* 167 (4), 1573–1583.
- Tan, W.Y., Wu, H.L., 1998. Stochastic modeling of the dynamics of CD4(+) T-cell infection by HIV and some Monte Carlo studies. *Math. Biosci.* 147 (2), 173–205.
- Wang, K., Mittler, J.E., Samudrala, R., 2006. Comment on “Evidence for positive epistasis in HIV-1”. *Science* 312 (5775), 848 (author reply, 848).
- Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emami, E.A., Deutsch, P., Lifson, J., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., Saag, M.S., Shaw, G.M., 1995. Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117–122.