

The coreceptor mutation CCR5 Δ 32 influences the dynamics of HIV epidemics and is selected for by HIV

Amy D. Sullivan*, Janis Wigginton, and Denise Kirschner†

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109-0620

Communicated by Avner Friedman, University of Minnesota, Minneapolis, MN, June 27, 2001 (received for review May 30, 2000)

We explore the impact of a host genetic factor on heterosexual HIV epidemics by using a deterministic mathematical model. A protective allele unequally distributed across populations is exemplified in our models by the 32-bp deletion in the host-cell chemokine receptor CCR5, CCR5 Δ 32. Individuals homozygous for CCR5 Δ 32 are protected against HIV infection whereas those heterozygous for CCR5 Δ 32 have lower pre-AIDS viral loads and delayed progression to AIDS. CCR5 Δ 32 may limit HIV spread by decreasing the probability of both risk of infection and infectiousness. In this work, we characterize epidemic HIV within three dynamic subpopulations: CCR5/CCR5 (homozygous, wild type), CCR5/CCR5 Δ 32 (heterozygous), and CCR5 Δ 32/CCR5 Δ 32 (homozygous, mutant). Our results indicate that prevalence of HIV/AIDS is greater in populations lacking the CCR5 Δ 32 alleles (homozygous wild types only) as compared with populations that include people heterozygous or homozygous for CCR5 Δ 32. Also, we show that HIV can provide selective pressure for CCR5 Δ 32, increasing the frequency of this allele.

Nineteen million people have died of AIDS since the discovery of HIV in the 1980s. In 1999 alone, 5.4 million people were newly infected with HIV (ref. 1 and <http://www.unaids.org/epidemicupdate/report/Epireport.html>). (For brevity, HIV-1 is referred to as HIV in this paper.) Sub-Saharan Africa has been hardest hit, with more than 20% of the general population HIV-positive in some countries (2, 3). In comparison, heterosexual epidemics in developed, market-economy countries have not reached such severe levels. Factors contributing to the severity of the epidemic in economically developing countries abound, including economic, health, and social differences such as high levels of sexually transmitted diseases and a lack of prevention programs. However, the staggering rate at which the epidemic has spread in sub-Saharan Africa has not been adequately explained. The rate and severity of this epidemic also could indicate a greater underlying susceptibility to HIV attributable not only to sexually transmitted disease, economics, etc., but also to other more ubiquitous factors such as host genetics (4, 5).

To exemplify the contribution of such a host genetic factor to HIV prevalence trends, we consider a well-characterized 32-bp deletion in the host-cell chemokine receptor CCR5, CCR5 Δ 32. When HIV binds to host cells, it uses the CD4 receptor on the surface of host immune cells together with a coreceptor, mainly the CCR5 and CXCR4 chemokine receptors (6). Homozygous mutations for this 32-bp deletion offer almost complete protection from HIV infection, and heterozygous mutations are associated with lower pre-AIDS viral loads and delayed progression to AIDS (7–14). CCR5 Δ 32 generally is found in populations of European descent, with allelic frequencies ranging from 0 to 0.29 (13). African and Asian populations studied outside the United States or Europe appear to lack the CCR5 Δ 32 allele, with an allelic frequency of almost zero (5, 13). Thus, to understand the effects of a protective allele, we use a mathematical model to track prevalence of HIV in populations with or without CCR5 Δ 32 heterozygous and homozygous people and also to follow the CCR5 Δ 32 allelic frequency.

We hypothesize that CCR5 Δ 32 limits epidemic HIV by decreasing infection rates, and we evaluate the relative contributions to this by the probability of infection and duration of infectivity. To capture HIV infection as a chronic infectious disease together with

vertical transmission occurring in untreated mothers, we model a dynamic population (i.e., populations that vary in growth rates because of fluctuations in birth or death rates) based on realistic demographic characteristics (18). This scenario also allows tracking of the allelic frequencies over time. This work considers how a specific host genetic factor affecting HIV infectivity and viremia at the individual level might influence the epidemic in a dynamic population and how HIV exerts selective pressure, altering the frequency of this mutant allele.

CCR5 is a host-cell chemokine receptor, which is also used as a coreceptor by R5 strains of HIV that are generally acquired during sexual transmission (6, 19–25). As infection progresses to AIDS the virus expands its repertoire of potential coreceptors to include other CC-family and CXC-family receptors in roughly 50% of patients (19, 26, 27). CCR5 Δ 32 was identified in HIV-resistant people (28). Benefits to individuals from the mutation in this allele are as follows. Persons homozygous for the CCR5 Δ 32 mutation are almost nonexistent in HIV-infected populations (11, 12) (see ref. 13 for review). Persons heterozygous for the mutant allele (CCR5 W/ Δ 32) tend to have lower pre-AIDS viral loads. Aside from the beneficial effects that lower viral loads may have for individuals, there is also an altruistic effect, as transmission rates are reduced for individuals with low viral loads (as compared with, for example, AZT and other studies; ref. 29). Finally, individuals heterozygous for the mutant allele (CCR5 W/ Δ 32) also have a slower progression to AIDS than those homozygous for the wild-type allele (CCR5 W/W) (7–10), remaining in the population 2 years longer, on average. Interestingly, the dearth of information on HIV disease progression in people homozygous for the CCR5 Δ 32 allele (CCR5 Δ 32/ Δ 32) stems from the rarity of HIV infection in this group (4, 12, 28). However, in case reports of HIV-infected CCR5 Δ 32/ Δ 32 homozygotes, a rapid decline in CD4⁺ T cells and a high viremia are observed, likely because of initial infection with a more aggressive viral strain (such as X4 or R5X4) (30).

The Model

Because we are most concerned with understanding the severity of the epidemic in developing countries where the majority of infection is heterosexual, we consider a purely heterosexual model. To model the effects of the allele in the population, we examine the rate of HIV spread by using an enhanced susceptible-infected-AIDS model of epidemic HIV (for review see ref. 31). Our model compares two population scenarios: a CCR5 wild-type population and one with CCR5 Δ 32 heterozygotes and homozygotes in addition to the wild type. To model the scenario where there are only wild-type individuals present in the population (i.e., CCR5 W/W), we track the sexually active susceptibles at time t [$S_{i,j}(t)$], where $i = 1$ refers to genotype (CCR5 W/W only in this case) and j is either the male or female subpopulation. We also track those who are

*Present address: Centers for Disease Control and Prevention Epidemiology Program Office, State Branch Oregon Health Division, 800 NE Oregon Street, Suite 772, Portland, OR 97232.

†To whom reprint requests should be addressed. E-mail: kirschne@umich.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

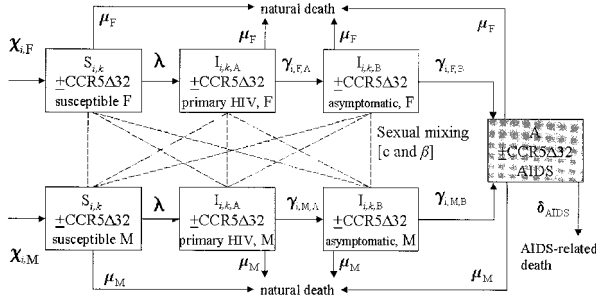


Fig. 1. A schematic representation of the basic compartmental HIV epidemic model. The criss-cross lines indicate the sexual mixing between different compartments. Each of these interactions has a positive probability of taking place; they also incorporate individual rates of transmission indicated as λ , but in full notation is $\lambda_{i,j,k \rightarrow i,j}$ where i,j,k is the phenotype of the infected partner and i,j is the phenotype of the susceptible partner. Also shown are the different rates of disease progression, $\gamma_{i,j,k}$, that vary according to genotype, gender, and stage. Thus, the interactions between different genotypes, genders, and stages are associated with a unique probability of HIV infection. M, male; F, female.

HIV-positive at time t not yet having AIDS in $I_{i,j,k}(t)$ where k refers to stage of HIV infection [primary (A) or asymptomatic (B)]. The total number of individuals with AIDS at time t are tracked in $A(t)$. The source population are children, $\chi_{i,j}(t)$, who mature into the sexually active population at time t (Fig. 1, Table 1). We compare the model of a population lacking the CCR5 Δ 32 allele to a demographically similar population with a high frequency of the allele. When genetic heterogeneity is included, male and female subpopulations are each further divided into three distinct genotypic groups, yielding six susceptible subpopulations, $[S_{i,j}(t)]$, where i ranges from 1 to 3, where 1 = CCR5W/W; 2 = CCR5 W/ Δ 32; 3 = CCR5 Δ 32/ Δ 32]. The infected classes, $I_{i,j,k}(t)$, also increase in number to account for these new genotype compartments. In both settings we assume there is no treatment available and no knowledge of HIV status by people in the early acute and middle asymptomatic stages (both conditions exist in much of sub-Saharan Africa). In addition, we assume that sexual mixing in the population occurs randomly with respect to genotype and HIV disease status, all HIV-infected people eventually progress to AIDS, and no barrier contraceptives are used. These last assumptions reflect both economic and social conditions.

Parameter Estimates for the Model. Estimates for rates that govern the interactions depicted in Fig. 1 were derived from the extensive literature on HIV. Our parameters and their estimates are summarized in Tables 2–4. The general form of the equations describing the rates of transition between population classes as depicted in Fig. 1 are summarized as follows:

$$\frac{dS_{i,j}(t)}{dt} = \chi_{i,j}(t) - \mu_j S_{i,j}(t) - \lambda_{i,j,\hat{k} \rightarrow i,j} S_{i,j}(t), \quad [1]$$

Table 1. Children's genotype

Parents	Mother			
Father	W/W	W/W	W/ Δ 32	Δ 32/ Δ 32
	W/ Δ 32	$\chi_{1,j}$ $\chi_{2,j}$	$\chi_{1,j}$ $\chi_{2,j}$ $\chi_{3,j}$	$\chi_{2,j}$ $\chi_{3,j}$
	Δ 32/ Δ 32	$\chi_{2,j}$	$\chi_{2,j}$ $\chi_{3,j}$	$\chi_{3,j}$

$\chi_{1,j}$ = wild-type children; (W/W); $\chi_{2,j}$ = heterozygous children (W/ Δ 32); $\chi_{3,j}$ = homozygous children (Δ 32/ Δ 32) of gender j . Children's genotypes are determined by using Mendelian inheritance patterns.

$$\frac{dI_{i,j,A}(t)}{dt} = \lambda_{i,j,\hat{k} \rightarrow i,j} S_{i,j}(t) - \mu_j I_{i,j,A}(t) - \gamma_{i,j,A} I_{i,j,A}(t), \quad [2]$$

$$\frac{dI_{i,j,B}(t)}{dt} = \gamma_{i,j,A} I_{i,j,A}(t) - \mu_j I_{i,j,B}(t) - \gamma_{i,j,B} I_{i,j,B}(t), \quad [3]$$

$$\frac{dA(t)}{dt} = \gamma_{i,j,B} \left(\sum_{i=1}^3 I_{i,F,B}(t) + I_{i,M,B}(t) \right) - \mu_A A(t) - \delta_{AIDS} A(t), \quad [4]$$

where, in addition to previously defined populations and rates (with i equals genotype, j equals gender, and k equals stage of infection, either A or B), μ_j , represents the non-AIDS (natural) death rate for males and females respectively, and μ_A is estimated by the average ($\mu_F + \mu_M/2$). This approximation allows us to simplify the model (only one AIDS compartment) without compromising the results, as most people with AIDS die of AIDS (δ_{AIDS}) and very few of other causes (μ_A). These estimates include values that affect infectivity ($\lambda_{i,j,\hat{k} \rightarrow i,j}$), transmission ($\beta_{i,j,\hat{k} \rightarrow i,j}$), and disease progression ($\gamma_{i,j,k}$) where the i,j,\hat{k} notation represents the genotype, gender, and stage of infection of the infected partner, and $j \neq \hat{j}$.

The effects of the CCR5 W/ Δ 32 and CCR5 Δ 32/ Δ 32 genotypes are included in our model through both the per-capita probabilities of infection, $\lambda_{i,j,\hat{k} \rightarrow i,j}$, and the progression rates, $\gamma_{i,j,k}$. The infectivity coefficients, $\lambda_{i,j,\hat{k} \rightarrow i,j}$, are calculated for each population subgroup based on the following: likelihood of HIV transmission in a sexual encounter between a susceptible and an infected ($\beta_{i,j,\hat{k} \rightarrow i,j}$) person; formation of new partnerships (c_j); number of contacts in a given partnership (ϕ_j); and probability of encountering an infected individual ($I_{i,j,\hat{k}}/N_j$). The formula representing this probability of infection is

$$\lambda_{i,j,\hat{k} \rightarrow i,j} = \frac{c_j \cdot \phi_j}{N_j} \left[\sum_{\hat{i}, \hat{k}} \beta_{\hat{i},j,\hat{k} \rightarrow i,j} I_{\hat{i},j,\hat{k}} \right],$$

where $j \neq \hat{j}$ is either male or female. N_j represents the total population of gender \hat{j} (this does not include those with AIDS in the simulations).

The average rate of partner acquisition, c_j , includes the mean plus the variance to mean ratio of the relevant distribution of partner-change rates to capture the small number of high-risk people: $c_j = m_j + (\sigma_j^2/m_j)$ where the mean (m_j) and variance (σ_j^2) are annual figures for new partnerships only (32). These means are estimated from Ugandan data for the number of heterosexual partners in the past year (33) and the number of nonregular heterosexual partners (i.e., spouses or long-term partners) in the past year (34). In these sexual activity surveys, men invariably have more new partnerships; thus, we assumed that they would have fewer average contacts per partnership than women (a higher rate of new partner acquisition means fewer sexual contacts with a given partner; ref. 35). To incorporate this assumption in our model, the male contacts/partnership, ϕ_M , was reduced by 20%. In a given population, the numbers of heterosexual interactions must equate between males and females. The balancing equation applied here is $S_{AF} m_F N_F = S_{AM} m_M N_M$, where S_{Aj} are the percent sexually active and N_j are the total in the populations for gender j . To specify changes in partner acquisition, we apply a male flexibility mechanism, holding the female rate of acquisition constant and allowing the male rates to vary (36, 37).

Transmission probabilities. The effect of a genetic factor in a model of HIV transmission can be included by reducing the transmission coefficient. The probabilities of transmission per contact with an infected partner, $\beta_{i,j,\hat{k} \rightarrow i,j}$, have been estimated in the literature (see ref. 38 for estimates in minimally treated groups). We want to capture a decreased risk in transmission based on genotype (ref. 39, Table 2). No studies have directly evaluated differences in infectivity between HIV-infected CCR5 W/ Δ 32

Table 2. Transmission probabilities

HIV-infected partner (<i>i, j, k</i>)	Susceptible partner (<i>i, j</i>) (<i>j</i> to <i>j</i>)	Susceptible partner (<i>i, j</i>)		
		W/W	W/Δ32	Δ32/Δ32
Acute/primary				
W/W or Δ32/Δ32	M to F	0.040	0.040	0.00040
	F to M	0.020	0.020	0.00020
W/Δ32	M to F	0.030	0.030	0.00030
	F to M	0.015	0.015	0.00015
Asymptomatic				
W/W or Δ32/Δ32	M to F	0.0010	0.0010	10×10^{-6}
	F to M	0.0005	0.0005	5×10^{-6}
W/Δ32	M to F	0.0005	0.0005	5×10^{-6}
	F to M	0.00025	0.00025	2.5×10^{-6}

Listed are the different transmission probabilities ($\beta_{i,j,k \rightarrow i,j}$) for random sexual mixing between persons where *i, j, k* is the phenotype of the infected partner and *i, j* is the phenotype of the susceptible partner. M, male; F, female.

heterozygotes and HIV-infected CCR5 wild types. Thus, we base estimates for reduced transmission on studies of groups with various HIV serum viral loads (40), HTLV-I/II viral loads (41), and a study of the effect of AZT treatment on transmission (29). We decrease transmission probabilities for infecting CCR5Δ32/Δ32 persons by 100-fold to reflect the rarity of infections in these persons. However, we assume that infected CCR5Δ32/Δ32 homozygotes can infect susceptibles at a rate similar to CCR5W/W homozygotes, as the former generally have high viremias (ref. 30, Table 2). We also assume that male-to-female transmission is twice as efficient as female-to-male transmission (up to a 9-fold difference has been reported; ref. 42) (ref. 43, Table 2).

Given the assumption of no treatment, the high burden of disease in people with AIDS is assumed to greatly limit their sexual activity. Our initial model excludes people with AIDS from the sexually active groups. Subsequently, we allow persons with AIDS to be sexually active, fixing their transmission rates (β_{AIDS}) to be the same across all CCR5 genotypes, and lower than transmission rates for primary-stage infection (as the viral burden on average is not as high as during the acute phase), and larger than transmission rates for asymptomatic-stage infection (as the viral burden characteristically increases during the end stage of disease).

Disease progression. We assume three stages of HIV infection: primary (acute, stage A), asymptomatic HIV (stage B), and AIDS. The rates of transition through the first two stages are denoted by $\gamma_{i,j,k}$, where *i* represents genotype, *j* is male/female, and *k* represents either stage A or stage B. Transition rates through each of these stages are assumed to be inversely proportional to the duration of that stage; however, other distributions are possible (31, 44, 45). Although viral loads generally peak in the first 2 months of infection, steady-state viral loads are established several months beyond this (46). For group A, the primary HIV-infecteds, duration is assumed to be 3.5 months. Based on results from European cohort studies (7–10), the beneficial effects of the CCR5 W/Δ32 genotype are observed mainly in the asymptomatic years of HIV infection; ≈ 7 years after seroconversion survival rates appear to be quite similar between heterozygous and homozygous individuals. We also assume that CCR5Δ32/Δ32-infected individuals and wild-type individuals progress similarly, and that men and women progress through each disease stage at the same rate. Given these observations, and that survival after infection may be shorter in untreated populations, we choose the duration time in stage B to be 6 years for wild-type individuals and 8 years for heterozygous individuals. Transition through AIDS, δ_{AIDS} , is inversely proportional to the duration of AIDS. We estimate this value to be 1 year for the time from onset of AIDS to death. The progression rates are summarized in Table 3.

Table 3. Progression rates

Genotype	Disease stage	Males/females
W/W	A	3.5
	B	0.16667
W/Δ32	A	3.5
	B	0.125
Δ32/Δ32	A	3.5
	B	0.16667

Shown are the rates of progression, $\gamma_{i,j,k}$ reflecting the different rates at which persons progress through different stages of disease by genotype, gender, and disease stage.

Demographic Setting. Demographic parameters are based on data from Malawi, Zimbabwe, and Botswana (3, 47). Estimated birth and child mortality rates are used to calculate the annual numbers of children ($\chi_{i,j}$) maturing into the potentially sexually active, susceptible group at the age of 15 years (3). For example, in the case where the mother is CCR5 wild type and the father is CCR5 wild type or heterozygous, the number of CCR5 W/W children is calculated as follows [suppressing (*t*) notation]: $\chi_{1,j} =$

$$B_r \sum_k \left[S_{1,F} \frac{(S_{1,M} + I_{1,M,k})}{N_M} + \left[(0.5) S_{1,F} \frac{(S_{2,M} + I_{2,M,k})}{N_M} \right] + P_v \left(\frac{(I_{1,F,k}(S_{1,M} + I_{1,M,k}))}{N_M} + \left[(0.5) I_{1,F,k} \frac{(S_{2,M} + I_{2,M,k})}{N_M} \right] \right) \right],$$

where the probability of HIV vertical transmission, $1 - p_v$, and the birthrate, B_r , are both included in the equations together with the Mendelian inheritance values as presented in Table 1. The generalized version of this equation (i.e., $\chi_{i,j}$) can account for six categories of children (including gender and genotype). We assume that all children of all genotypes are at risk, although we can relax this condition if data become available to support vertical protection (e.g., ref. 48). All infected children are assumed to die before age 15. Before entering the susceptible group at age 15, there is additional loss because of mortality from all non-AIDS causes occurring less than 15 years of age at a rate of $\mu_\chi \times \chi_{i,j}$ (where μ_χ is the mortality under 15 years of age). Children then enter the population as susceptibles at an annual rate, $\sigma_j \times \chi_{i,j}/15$, where σ_j distributes the children 51% females and 49% males. All parameters and their values are summarized in Table 4.

Prevalence of HIV

Demographics and Model Validation. The model was validated by using parameters estimated from available demographic data. Simulations were run in the absence of HIV infection to compare the model with known population growth rates. Infection was subsequently introduced with an initial low HIV prevalence of 0.5% to capture early epidemic behavior.

In deciding on our initial values for parameters during infection, we use Joint United Nations Programme on HIV/AIDS national prevalence data for Malawi, Zimbabwe, and Botswana. Nationwide seroprevalence of HIV in these countries varies from $\approx 11\%$ to over 20% (3), although there may be considerable variation within given subpopulations (2, 49).

In the absence of HIV infection, the annual percent population growth rate in the model is $\approx 2.5\%$, predicting the present-day values for an average of sub-Saharan African cities (data not shown). To validate the model with HIV infection, we compare our simulation of the HIV epidemic to existing prevalence data for Kenya and Mozambique (<http://www.who.int/emc-hiv/factsheets/pdfs/kenya.pdf> and ref. 51). Prevalence data collected from

Table 4. Parameter values

Parameter	Definition	Value
μ_F, μ_M	All-cause mortality for adult females (males)	0.015 (0.016) per year
μ_X	All-cause childhood mortality (<15 years of age)	0.01 per year
B_r	Birthrate	0.25 per woman per year
SA_F	Percent females acquiring new partners (sexual activity)	10%
SA_M	Percent males acquiring new partners (sexual activity)	25%
$m_F(\sigma_F^2)$	Mean (variance) no. of new partners for females	1.8 (1.2) per year
σ_M^2	Variance in no. of new partners for males	5.5 per year
$1 - p_v$	Probability of vertical transmission	0.30 per birth
$I_{i,j,k}(0)$	Initial total population HIV-positive	0.50%
$\chi_{i,j}(0)$	Initial total children in population (<15 years of age)	45%
$W/W(0)$	Initial total wild types (W/W) in population	80%
$W/\Delta32(0)$	Initial total heterozygotes ($W/\Delta32$) in population	19%
$\Delta32/\Delta32(0)$	Initial total homozygotes ($\Delta32/\Delta32$) in population	1%
$r_M(r_F)$	Initial percent males (females) in total population	49% (51%)
ϕ_F, ϕ_M	Number of sexual contacts a female (male) has	30 (24) per partner
$\varepsilon_{i,j,k}$	% effect of mutation on transmission rates (see Table 2)	$0 < \varepsilon_{i,j,k} < 1$
δ	Death rate for AIDS population	1.0 per year
q	Allelic frequency of $\Delta32$ allele	0.105573

Shown are the parameter values for parameters other than the transmission probabilities (Table 2) and the progression rates (Table 3). Each were estimated from data as described in text.

these countries follow similar trajectories to those predicted by our model (Fig. 2).

Effects of the Allele on Prevalence. After validating the model in the wild type-only population, both CCR5 $\Delta32$ heterozygous and homozygous people are included. Parameter values for HIV transmission, duration of illness, and numbers of contacts per partner are assumed to be the same within both settings. We then calculate HIV/AIDS prevalence among adults for total HIV/AIDS cases.

Although CCR5 $\Delta32/\Delta32$ homozygosity is rarely seen in HIV-positive populations (prevalence ranges between 0 and 0.004%), 1–20% of people in HIV-negative populations of European descent are homozygous. Thus, to evaluate the potential impact of CCR5 $\Delta32$, we estimate there are 19% CCR5 $W/\Delta32$ heterozygous and 1% CCR5 $\Delta32/\Delta32$ homozygous people in our population. These values are in Hardy-Weinberg equilibrium with an allelic frequency of the mutation as 0.105573.

Fig. 3 shows the prevalence of HIV in two populations: one lacking the mutant CCR5 allele and another carrying that allele. In the population lacking the protective mutation, prevalence increases logarithmically for the first 35 years of the epidemic, reaching 18% before leveling off.

In contrast, when a proportion of the population carries the CCR5 $\Delta32$ allele, the epidemic increases more slowly, but still logarithmically, for the first 50 years, and HIV/AIDS prevalence reaches $\approx 12\%$ (Fig. 3). Prevalence begins to decline slowly after 70 years.

In the above simulations we assume that people with AIDS are not sexually active. However, when these individuals are included in the sexually active population the severity of the epidemic increases considerably (data not shown). Consistent with our initial simulations, prevalences are still relatively lower in the presence of the CCR5 mutation.

Because some parameters (e.g., rate constants) are difficult to estimate based on available data, we implement an uncertainty analysis to assess the variability in the model outcomes caused by any inaccuracies in estimates of the parameter values with regard to the effect of the allelic mutation. For these analyses we use Latin hypercube sampling, as described in refs. 52–56. Our uncertainty and sensitivity analyses focus on infectivity vs. duration of infectiousness. To this end, we assess the effects on the dynamics of the epidemic for a range of values of the parameters governing trans-

mission and progression rates: $\beta_{i,j,k \rightarrow i,j}$ and $\gamma_{i,j,k}$. All other parameters are held constant. These results are presented as an interval band about the average simulation for the population carrying the CCR5 $\Delta32$ allele (Fig. 3). Although there is variability in the model outcomes, the analysis indicates that the overall model predictions are consistent for a wide range of transmission and progression rates. Further, most of the variation observed in the outcome is because of the transmission rates for both heterosexual males and females in the primary stage of infection ($\beta_{2,M,A \rightarrow i,F}$, $\beta_{2,F,A \rightarrow i,M}$). As mentioned above, we assume lower viral loads correlate with reduced infectivity; thus, the reduction in viral load in heterozygotes has a major influence on disease spread.

HIV Induces Selective Pressure on Genotype Frequency

To observe changes in the frequency of the CCR5 $\Delta32$ allele in a setting with HIV infection as compared with the Hardy-Weinberg

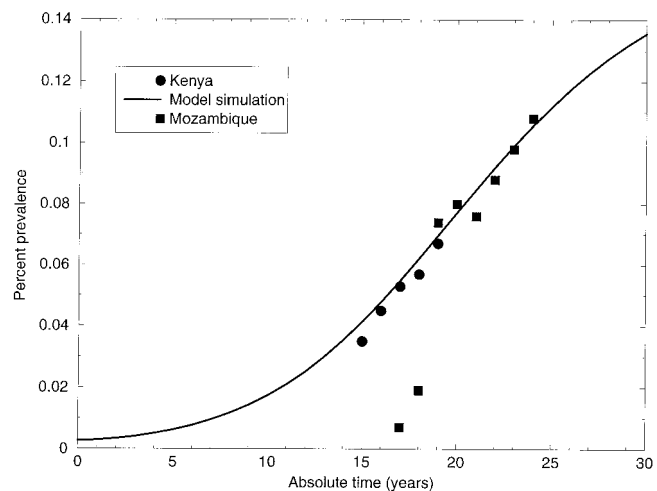


Fig. 2. Model simulation of HIV infection in a population lacking the protective CCR5 $\Delta32$ allele compared with national data from Kenya (healthy adults) and Mozambique (blood donors, ref. 17). The simulated population incorporates parameter estimates from sub-Saharan African demographics. Note the two outlier points from the Mozambique data were likely caused by underreporting in the early stages of the epidemic.

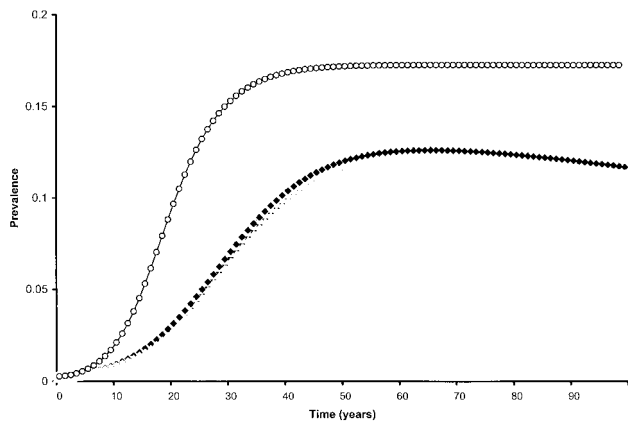


Fig. 3. Prevalence of HIV/AIDS in the adult population as predicted by the model. The top curve (○) indicates prevalence in a population lacking the protective allele. We compare that to a population with 19% heterozygous and 1% homozygous for the allele (implying an allelic frequency of 0.105573. Confidence interval bands (light gray) are shown around the median simulation (◆) providing a range of uncertainty in evaluating parameters for the effect of the mutation on the infectivity and the duration of asymptomatic HIV for heterozygotes.

equilibrium in the absence of HIV, we follow changes in the total number of CCR5 Δ 32 heterozygotes and homozygotes over 1,000 years (Fig. 4). We initially perform simulations in the absence of HIV infection as a negative control to show there is not significant selection of the allele in the absence of infection. To determine how long it would take for the allelic frequency to reach present-day levels (e.g., $q = 0.105573$), we initiate this simulation for 1,000 years with a very small allelic frequency ($q = 0.00105$). In the absence of HIV, the allelic frequency is maintained in equilibrium as shown by the constant proportions of CCR5 Δ 32 heterozygotes and homozygotes (Fig. 4, solid lines). The selection for CCR5 Δ 32 in the presence of HIV is seen in comparison (Fig. 4, dashed lines). We expand the time frame of this simulation to 2,000 years to view the point at which the frequency reaches present levels (where $q \sim 0.105573$ at year = 1200). Note that the allelic frequency increases for $\approx 1,600$ years before leveling off.

Discussion

This study illustrates how populations can differ in susceptibility to epidemic HIV/AIDS depending on a ubiquitous attribute such as a prevailing genotype. We have examined heterosexual HIV epidemics by using mathematical models to assess HIV transmission in dynamic populations either with or without CCR5 Δ 32 heterozygous and homozygous persons. The most susceptible population lacks the protective mutation in CCR5. In less susceptible populations, the majority of persons carrying the CCR5 Δ 32 allele are heterozygotes. We explore the hypothesis that lower viral loads (CCR5 Δ 32 heterozygotes) or resistance to infection (CCR5 Δ 32 homozygotes) observed in persons with this coreceptor mutation ultimately can influence HIV epidemic trends. Two contrasting influences of the protective CCR5 allele are conceivable: it may limit the epidemic by decreasing the probability of infection because of lower viral loads in infected heterozygotes, or it may exacerbate the epidemic by extending the time that infectious individuals remain in the sexually active population. Our results strongly suggest the former. Thus, the absence of this allele in Africa could explain the severity of HIV disease as compared with populations where the allele is present.

We also observed that HIV can provide selective pressure for the CCR5 Δ 32 allele within a population, increasing the allelic frequency. Other influences may have additionally selected for this allele. Infectious diseases such as plague and small pox have been postulated to select for CCR5 Δ 32 (57, 58). For plague, relatively

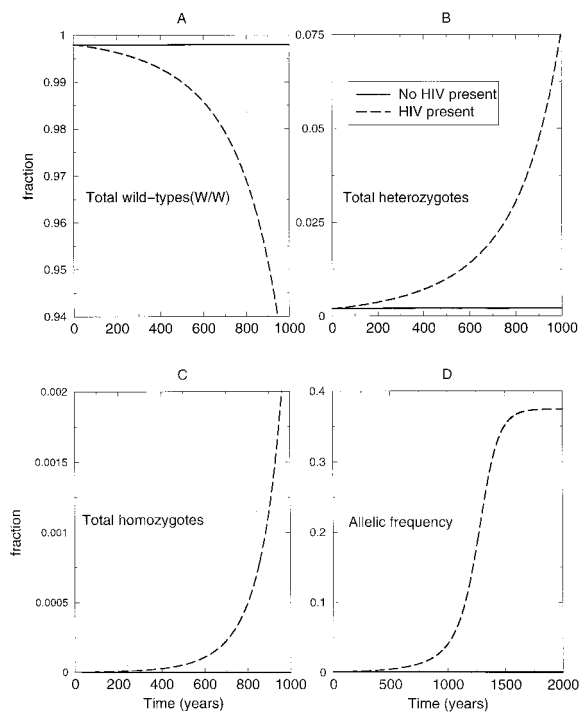


Fig. 4. Effects of HIV-1 on selection of the CCR5 Δ 32 allele. The Hardy-Weinberg equilibrium level is represented in the no-infection simulation (solid lines) for each population. Divergence from the original Hardy-Weinberg equilibrium is shown to occur in the simulations that include HIV infection (dashed lines). Fraction of the total subpopulations are presented: (A) wild types (W/W), (B) heterozygotes (W/ Δ 32), and (C) homozygotes (Δ 32/ Δ 32). Note that we initiate this simulation with a much lower allelic frequency (0.00105) than used in the rest of the study to better exemplify the actual selective effect over a 1,000-year time scale. (D) The allelic selection effect over a 2,000-year time scale.

high levels of CCR5 Δ 32 are believed to have arisen within $\approx 4,000$ years, accounting for the prevalence of the mutation only in populations of European descent. Smallpox virus uses the CC-coreceptor, indicating that direct selection for mutations in CCR5 may have offered resistance to smallpox. Given the differences in the epidemic rates of plague (59), smallpox, and HIV, it is difficult to directly compare our results to these findings. However, our model suggests that the CCR5 Δ 32 mutation could have reached its present allelic frequency in Northern Europe within this time frame if selected for by a disease with virulence patterns similar to HIV. Our results further support the idea that HIV has been only recently introduced as a pathogen into African populations, as the frequency of the protective allele is almost zero, and our model predicts that selection of the mutant allele in this population by HIV alone takes at least 1,000 years. This prediction is distinct from the frequency of the CCR5 Δ 32 allele in European populations, where pathogens that may have influenced its frequency (e.g., *Yersinia pestis*) have been present for much longer.

Two mathematical models have considered the role of parasite and host genetic heterogeneity with regard to susceptibility to another pathogen, namely malaria (60, 61). In each it was determined that heterogeneity of host resistance facilitates the maintenance of diversity in parasite virulence. Given our underlying interest in the coevolution of pathogen and host, we focus on changes in a host protective mutation, holding the virulence of the pathogen constant over time.

Even within our focus on host protective mutations, numerous genetic factors, beneficial or detrimental, could potentially influence epidemics. Other genetically determined host factors affecting HIV susceptibility and disease progression include a CCR5 A/A to

G/G promoter polymorphism (62), a CCR2 point mutation (11, 63), and a mutation in the CXCR4 ligand (64). The CCR2b mutation, CCR264I, is found in linkage with at least one CCR5 promoter polymorphism (65) and is prevalent in populations where CCR5 Δ 32 is nonexistent, such as sub-Saharan Africa (63). However, as none of these mutations have been consistently shown to be as protective as the CCR5 Δ 32 allele, we simplified our model to incorporate only the effect of CCR5 Δ 32. Subsequent models could be constructed from our model to account for the complexity of multiple protective alleles. It is interesting to note that our model predicts that even if CCR264I is present at high frequencies in Africa, its protective effects may not augment the lack of a protective allele such as CCR5 Δ 32.

Although our models demonstrate that genetic factors can contribute to the high prevalence of HIV in sub-Saharan Africa, demographic factors are also clearly important in this region. Our models explicitly incorporated such factors, for example, lack of treatment availability. Additional factors were implicitly controlled for by varying only the presence of the CCR5 Δ 32 allele. More complex models eventually could include interactions with infectious diseases that serve as cofactors in HIV transmission. The role of high sexually transmitted disease prevalences in HIV infection

has long been discussed, especially in relation to core populations (15, 50, 66). Malaria, too, might influence HIV transmission, as it is associated with transient increases in semen HIV viral loads and thus could increase the susceptibility of the population to epidemic HIV (16).

In assessing the HIV/AIDS epidemic, considerable attention has been paid to the influence of core groups in driving sexually transmitted disease epidemics. Our results also highlight how characteristics more uniformly distributed in a population can affect susceptibility. We observed that the genotypic profile of a population affects its susceptibility to epidemic HIV/AIDS. Additional studies are needed to better characterize the influence of these genetic determinants on HIV transmission, as they may be crucial in estimating the severity of the epidemic in some populations. This information can influence the design of treatment strategies as well as point to the urgency for education and prevention programs.

We thank Mark Krosky, Katia Koelle, and Kevin Chung for programming and technical assistance. We also thank Drs. V. J. DiRita, P. Kazanjian, and S. M. Blower for helpful comments and discussions. We thank the reviewers for extremely insightful comments.

- Weiss, H. A., & Hawkes, S. (2001) *Leprosy Rev.* **72**, 92–98.
- Taha, T. E., Dallabetta, G. A., Hoover, D. R., Chipangwi, J. D., & Mtimavalye, L. A. R. (1998) *AIDS* **12**, 197–203.
- World Health Organization/UNAIDS (1998) *AIDS Epidemic Update* (World Health Organization, Geneva), pp. 1–17.
- D'Souza, M. P., & Harden, V. A. (1996) *Nat. Med.* **2**, 1293–1300.
- Martinson, J. J., Chapman, N. H., Rees, D. C., Liu, Y. T., & Clegg, J. B. (1997) *Nat. Genet.* **16**, 100–103.
- Roos, M. T. L., Lange, J. M. A., deGoede, R. E. Y., Miedema, P. T., Tersmette, F., Coutinho, M., & Schellekens, R. A. (1992) *J. Infect. Dis.* **165**, 427–432.
- Garred, P., Eugen-Olsen, J., Iversen, A. K. N., Benfield, T. L., Sveigaard, A., & Hofmann, B. (1997) *Lancet* **349**, 1884.
- Katzenstein, T. L., Eugen-Olsen, J., Hofman, B., Benfield, T., Pedersen, C., Iversen, A. K., Sorensen, A. M., Garred, P., Koppellus, U., Sveigaard, A., & Gerstoft, J. (1997) *J. Acquired Immune Defic. Syndr. Hum. Retrovirol.* **16**, 10–14.
- deRoda, H., Meyer, K., Katzenstein, W., & Dean, M. (1996) *Science* **273**, 1856–1862.
- Meyer, L., Magierowska, M., Hubert, J. B., Rouzioux, C., Deveau, C., Sanson, F., Debre, P., Delfraissy, J. F., & Theodorou, I. (1997) *AIDS* **11**, F73–F78.
- Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, D. A., Lomb, G. A., Goedert, J. J., O'Brien, T. R., Jacobson, L. P., Kaslow, R., et al. (1997) *Science* **277**, 959–965.
- Samson, M., Libert, F., Doranz, B. J., Rucker, J., Liesnard, C., Farber, C. M., Saragosti, S., Lapoumeroulie, C., Cognaux, J., Forcille, C., et al. (1996) *Nature (London)* **382**, 722–725.
- McNicholl, J. M., Smith, D. K., Qari, S. H., & Hodge, T. (1997) *Emerging Infect. Dis.* **3**, 261–271.
- Michael, N. L., Chang, G., Louie, L. G., Mascola, J. R., Dondero, D., Bix, D. L., & Sheppard, H. W. (1997) *Nat. Med.* **3**, 338–340.
- Mayaud, P., Mosha, F., Todd, J., Balira, R., Mgara, J., West, B., Rusizoka, M., Mwijarubi, E., Gabone, R., Gavyole, A., et al. (1997) *AIDS* **11**, 1873–1880.
- Hoffman, I. F., Jere, C. S., Taylor, T. E., Munthali, P., & Dyer, J. R. (1998) *AIDS* **13**, 487–494.
- U.S. Bureau of the Census (1999) *HIV/AIDS Surveillance Database* (Population Division, International Programs Center, Washington, DC).
- Anderson, R. M., May, R. M., & McLean, A. R. (1988) *Nature (London)* **332**, 228–234.
- Berger, E. A., Doms, R. W., Fenyo, E. M., Korber, B. T., Littman, D. R., Moore, J. P., Sattentau, Q. J., Schuitemaker, H., Sodroski, J., & Weiss, R. A. (1998) *Nature (London)* **391**, 240.
- Alkhatib, G., Broder, C. C., & Berger, E. A. (1996) *J. Virol.* **70**, 5487–5494.
- Choe, H., Farzan, M., Sun, Y., Sullivan, N., Rollins, B., Ponath, P. D., Wu, L., Mackay, C. R., LaRosa, G., Newman, W., et al. (1996) *Cell* **85**, 1135–1148.
- Deng, H., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhardt, M., Di Marzio, P., Marmor, S., Sutton, R. E., Hill, C. M., et al. (1996) *Nature (London)* **381**, 661–666.
- Doranz, B. J., Rucker, J., Yi, Y., Smyth, R. J., Samsom, M., Peiper, M., Parmentier, S. C., Collman, R. G., & Doms, R. W. (1996) *Cell* **85**, 1149–1158.
- Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., Cayan, C., Maddon, P. J., Koup, R. A., Moore, J. P., & Paxton, W. A. (1996) *Nature (London)* **381**, 667–673.
- Zhu, T., Mo, H., Wang, N., Nam, D. S., Cao, Y., Koup, R. A., & Ho, D. D. (1993) *Science* **261**, 1179–1181.
- Bjorndal, A., Deng, H., Jansson, M., Fiore, J. R., Colognesi, C., Karlsson, A., Albert, J., Scarlatti, G., Littman, D. R., & Fenyo, E. M. (1997) *J. Virol.* **71**, 7478–7487.
- Conner, R. L., Sheridan, K. E., Ceradini, D., Choe, S., & Landau, N. R. (1997) *J. Exp. Med.* **185**, 621–628.
- Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhmann, H., Koup, R. A., & Landau, N. R. (1996) *Cell* **86**, 367–377.
- Mussico, M., Lazzarin, A., Nicolosi, A., Gasparini, M., Costigliola, P., Arici, C., & Saracco, A. (1994) *Arch. Intern. Med. (Moscow)* **154**, 1971–1976.
- Michael, N. L., Nelson, J. A., KewalRamani, V. N., Chang, G., O'Brien, S. J., Mascola, J. R., Volzky, B., Louder, M., White, G. C., & Littman, D. R., et al. (1998) *J. Virol.* **72**, 6040–6047.
- Hethcote, H. W., & Yorke, J. A. (1984) *Gonorrhea Transmission Dynamics and Control* (Springer, Berlin).
- Anderson, R. M., & May, R. M. (1988) *Nature (London)* **333**, 514–522.
- Asiimwe-Okiror, G., Opio, A. A., Musinguzi, J., Madraa, E., Tembo, G., & Carael, M. (1997) *AIDS* **11**, 1757–1763.
- Carael, M., Cleland, J., Deheneffe, J. C., Ferry, B., & Ingham, R. (1995) *AIDS* **9**, 1171–1175.
- Blower, S. M., & Boe, C. (1993) *J. AIDS* **6**, 1347–1352.
- Kirschner, D. (1996) *J. Appl. Math.* **56**, 143–166.
- Le Pont, F., & Blower, S. (1991) *J. AIDS* **4**, 987–999.
- Kim, M. Y., & Lagakos, S. W. (1990) *Ann. Epidemiol.* **1**, 117–128.
- Anderson, R. M., & May, R. M. (1992) *Infectious Disease of Humans: Dynamics and Control* (Oxford Univ. Press, Oxford).
- Ragni, M. V., Faruki, H., & Kingsley, L. A. (1998) *J. Acquired Immune Defic. Syndr.* **17**, 42–45.
- Kaplan, J. E., Khabbaz, R. F., Murphy, E. L., Hermansen, S., Roberts, C., Lal, R., Heneine, W., Wright, D., Matijas, L., Thomson, R., et al. (1996) *J. Acquired Immune Defic. Syndr. Hum. Retrovirol.* **12**, 193–201.
- Padian, N. S., Shiboski, S. C., Glass, S. O., & Vittinghoff, E. (1997) *Am. J. Edu.* **146**, 350–357.
- Leynaert, B., Downs, A. M., & de Vincenzi, I. (1998) *Am. J. Edu.* **148**, 88–96.
- Garnett, G. P., & Anderson, R. M. (1995) *J. Acquired Immune Defic. Syndr.* **9**, 500–513.
- Stigum, H., Magnus, P., Harris, J. R., Samualson, S. O., & Bakketeig, L. S. (1997) *Am. J. Edu.* **145**, 636–643.
- Ho, D. D., Neumann, A. U., Perelson, A. S., Chen, W., Leonard, J. M., & Markowitz, M. (1995) *Nature (London)* **373**, 123–126.
- World Resources Institute (1999) *World Resources (1998–1999)* (Oxford Univ. Press, Oxford).
- Kostrikis, L. G., Neumann, A. U., Thomson, B., Korber, B. T., McHardy, P., Karanicas, R., Deutsch, L., Huang, Y., Lew, J. F., McIntosh, K., et al. (1999) *J. Virol.* **73**, 10264–10271.
- Low-Beer, D., Stoneburner, R. L., & Mukulu, A. (1997) *Nat. Med.* **3**, 553–557.
- Grosskurth, H., Mosha, F., Todd, J., Senkoro, K., Newell, J., Klokke, A., Chantalucha, J., West, B., Mayaud, P., & Gavyole, A. (1995) *AIDS* **9**, 927–934.
- Melo, J., Beby-Defaux, A., Faria, C., Guiraud, G., Folgosa, E., Barreto, A., & Agius, G. (2000) *J. AIDS* **23**, 203–204.
- Iman, R. L., Helton, J. C., & Campbell, J. E. (1981) *J. Quality Technol.* **13**, 174–183.
- Iman, R. L., Helton, J. C., & Campbell, J. E. (1981) *J. Quality Technol.* **13**, 232–240.
- Blower, S. M., & Dowlatabadi, H. (1994) *Int. Stat. Rev.* **62**, 229–243.
- Porco, T. C., & Blower, S. M. (1998) *Theor. Popul. Biol.* **54**, 117–132.
- Blower, S. M., Porco, T. C., & Darby, G. (1998) *Nat. Med.* **4**, 673–678.
- Libert, F., Cochaux, P., Beckman, G., Samson, M., Aksenova, M., Cao, A., Czeizel, A., Claustres, M., de la Rua, C., Ferrari, M., et al. (1998) *Hum. Mol. Genet.* **7**, 399–406.
- Lalani, A. S., Masters, J., Zeng, W., Barrett, J., Pannu, R., Everett, H., Arendt, C. W., & McFadden, G. (1999) *Science* **286**, 1968–1971.
- Kermack, W. O., & McKendrick, A. G. (1927) *Proc. R. Soc. London* **261**, 700–721.
- Gupta, S., & Hill, A. V. S. (1995) *Proc. R. Soc. London Ser. B* **260**, 271–277.
- Ruwende, C., Khoo, S. C., Snow, R. W., Yates, S. N. R., Kwiatkowski, D., Gupta, S., Warn, P., Allsopp, C. E., Gilbert, S. C., & Peschu, N. (1995) *Nature (London)* **376**, 246–249.
- McDermott, D. H., Zimmerman, P. A., Guignard, F., Kleeberger, C. A., Leitman, S. F., & Murphy, P. M. (1998) *Lancet* **352**, 866–870.
- Kostrikis, L. G., Huang, Y., Moore, J. P., Wolinsky, S. M., Zhang, L., Guo, Y., Deutsch, L., Phair, J., Neumann, A. U., & Ho, D. D. (1998) *Nat. Med.* **4**, 350–353.
- Winkler, C., Modi, W., Smith, M. W., Nelson, G. W., Wu, X., Carrington, M., Dean, M., Honjo, T., Tashiro, K., Yabe, D., et al. (1998) *Science* **279**, 389–393.
- Martinson, J. J., Hong, L., Karanicas, R., Moore, J. P., & Kostrikis, L. G. (2000) *AIDS* **14**, 483–489.
- Vernazza, P. L., Eron, J. J., Fiscus, S. A., & Cohen, M. S. (1999) *AIDS* **13**, 155–166.