Copyright © 1999 by the Genetics Society of America

Overdominant Alleles in a Population of Variable Size

Montgomery Slatkin and Christina A. Muirhead

Department of Integrative Biology, University of California, Berkeley, California 94720-3140

Manuscript received November 16, 1998
Accepted for publication February 26, 1999

ABSTRACT

An approximate method is developed to predict the number of strongly overdominant alleles in a population of which the size varies with time. The approximation relies on the strong-selection weak-mutation (SSWM) method introduced by J. H. Gillespie and leads to a Markov chain model that describes the number of common alleles in the population. The parameters of the transition matrix of the Markov chain depend in a simple way on the population size. For a population of constant size, the Markov chain leads to results that are nearly the same as those of N. Takahata. The Markov chain allows the prediction of the numbers of common alleles during and after a population bottleneck and the numbers of alleles surviving from before a bottleneck. This method is also adapted to modeling the case in which there are two classes of alleles, with one class causing a reduction in fitness relative to the other class. Very slight selection against one class can strongly affect the relative frequencies of the two classes and the relative ages of alleles in each class.

STRONG balancing selection can result from overdominance in fitness or from disassortative mating of the kind created by self-incompatibility systems in plants. Models of balancing selection are difficult to analyze completely because of the large number of alleles present, so numerous approximations have been used to provide some quantitative understanding of how a balance is achieved between mutation, selection, and genetic drift. Approximations that have been used—including those of Wright (1939), Kimura and Crow (1964), Sasaki (1989, 1992), Takahata (1990), and Vekemans and Slatkin (1994)—have all assumed that a stochastic equilibrium was reached in a population of constant size. Yet many applications, such as the estimation of the ancestral population size of modern humans (Klein *et al.* 1993; Ayala 1995), are to populations that have been variable in size. Of particular concern in this and other applications is the effects of a bottleneck in size. The problem is to obtain quantitative results without resorting to extensive computer simulations. Simulations are important and useful, but they cannot usually show how a model's behavior depends on simple combinations of parameters.

In this article, we introduce an approximate analytic theory that allows us to obtain relatively simple results for populations of variable size and to generalize models of balancing selection to allow for slight differences in relative fitness of different classes of alleles. Our approximate method is based on the strong-selection weak-mutation (abbreviated as SSWM and pronounced "swim") approximation of Gillespie (1984, 1991). This approximation leads naturally to a Markov chain for the number of common alleles in a population and provides a way to compute the rates of gain and loss of common alleles when population size varies. Our approximation is the same as that used previously by Sasaki (1989, 1992), and our results for a population of constant size are the same as his.

The Markov chain can be further approximated by an ordinary differential equation for the expected numbers of alleles and thus can provide a rough prediction of the rate at which alleles are gained and lost because of a bottleneck. We verify that our method provides an accurate approximation by comparing our predictions with computer simulations. Our results are very similar to and consistent with those of Takahata (1990) for populations of constant size and with those of Takahata (1993) for a bottleneck in population size.

ONE SELECTIVE CLASS OF ALLELES

Analytic theory: Throughout, we consider a single autosomal locus in a randomly mating diploid population in which the population size, $N(t)$, may vary with time. We assume that mutations occur at rate u and that every mutation is to an allele not previously found in the population—the infinite alleles model of Kimura and Crow (1964). We let i be the number of common alleles in the population at any time, where a common allele is one with a frequency near its deterministic equilibrium frequency, as described later. We assume overdominance in fitness, in which every heterozygote has a fitness of $1 + s$ relative to every homozygote. This parameterization is different from that of Takahata

Corresponding author: Montgomery Slatkin, Department of Integrative Biology, University of California, 3060 Valley Life Sciences Bldg., Berkeley, CA 94720-3140. E-mail: slatkin@socrates.berkeley.edu

(1990), who assumed every homozygote has a fitness $1 - s$ relative to every heterozygote. In the diffusion limit (large population size and small s), on which both Takahata's and our theories depend, there is no difference in the parameterizations. A value of s in Takahata's parameterization corresponds to $s/(1 - s)$ in ours.

New mutations arise and if they become common, they persist in the population as common alleles for a long time. This situation is suitable for Gillespie's (1984, 1991) SSWM approximation. We treat i, the number of alleles, as a random variable for which transitions from one time to the next can be modeled by a Markov chain. In fact, the Markov chain we use when there is one class of alleles is of a particularly simple kind because it allows for an increase or decrease in i by only one. Consider a population with i common alleles and assume that they are at the deterministic equilibrium with each having a frequency $1/i$. Now suppose that a new allele appears by mutation. Let the frequency of the new mutant be z. The change in z is approximately

$$\Delta z = sz\left(\frac{1}{i} - z\right) \cong \frac{sz}{i} \tag{1}$$

when z is small. That is, the advantage that a rare mutant gains because of overdominance decreases with i because the proportion of the population that is heterozygous increases with i. We assume that the probability that a new mutant increases from one copy to become common is the same as the probability of fixation of a new advantageous mutant with a fixed selection coefficient. With a selection coefficient of s', the fixation probability is $\sim 2s'$ if N is large. Hence, the probability that a new mutant becomes a common allele is approximately $2s/i$. This approximation ignores the decrease in the frequencies of the other common alleles that must occur as the new mutant becomes common.

With a mutation rate u per generation, there are on average $2N(t)u$ mutations per generation, so the probability per generation that a new mutant appears and becomes common is

$$\lambda_i(t) = \frac{4N(t)us}{i}, \tag{2}$$

which is the probability per generation that i increases to $i + 1$. In using the population size at t to determine $\lambda_i(t)$, we are assuming that if a mutant is to become common, it does so instantaneously on the time scale of interest. If an allele does not become common, we assume that it is lost from the population so quickly that it can be ignored.

We also need the probability of loss of a common allele. Takahata (1990, Equation 6) found that in a population of constant size, the average time to loss of an allele in frequency x is approximately

$$\bar{t}(x) = 2Ne^{Sm^2}\sqrt{\frac{2\pi}{S^3m^4}} \tag{3}$$

in a population of constant size N, where $S = 2Ns$ and $m = F$, the homozygosity. Equation 3 assumes that Ns is large and i is relatively small. When those conditions are not satisfied, as can be the case immediately after an extreme reduction in population size, (3) no longer provides a good approximation and $\bar{t}(x)$ must be computed numerically from the equation given on p. 2420 of Takahata (1990). We found that, in agreement with Sasaki (1992, Equation 8), for very large values of S a better analytic approximation is given by (3) without the 2 under the square root. But for values of S of practical interest in models of human populations (50 to 200) a number between 1 and 2 provides a better approximation (Takahata 1993; N. Takahata, personal communication).

When i alleles are equally frequent, $F = 1/i$. Equation 3 differs from Takahata's equation because time is measured in generations and not in units of $2N$ generations. When S is large, common alleles are held relatively close to their expected frequency, $x = 1/i$, with only a weak force of genetic drift pulling them toward the boundary ($x = 0$). This stochastic process is similar to that described by Newman *et al.* (1985), in which the transition from being common to being lost occurs in a very short time. In that case, it is reasonable to approximate the process of loss of a common allele as a Poisson process occurring with an instantaneous rate $1/\bar{t}(x)$.

When population size varies with time, it is still reasonable to use the assumption that loss of an allele occurs instantaneously with the probability of loss per generation determined by $N(t)$, provided that $2N(t)s$ is large enough that the selection dominates genetic drift. In that case, the instantaneous rate of loss is still $1/\bar{t}(x)$ with S replaced by $2N(t)s$. Because there are i alleles that are liable to be lost by drift, the probability of loss per generation is

$$\mu_i(t) = \frac{1}{2N(t)i}e^{-2N(t)s/i^2}\sqrt{\frac{[2N(t)s]^3}{2\pi}} \tag{4}$$

once we replace x by $1/i$. The probability per generation that i does not change is $1 - \mu_i(t) - \lambda_i(t)$.

We can compare our approximation with Takahata's (1990) by assuming a constant population size N and finding the value of i for which $\mu_i = \lambda_i$. The result is

$$\hat{i} = \sqrt{2S/\ln\left(\frac{S}{32\pi M^2}\right)}, \tag{5}$$

where $M = Nu$. This expression differs from Takahata's (1990, Equation 5) only by having 32 instead of 16 in the denominator of the logarithm, and it differs only slightly from the comparable result of Sasaki (1989, 1992). Numerically, the difference between Takahata's and our approximate results is very small, thus supporting the idea that our method based on a Markov chain is similar in accuracy to Takahata's. Both approxi-

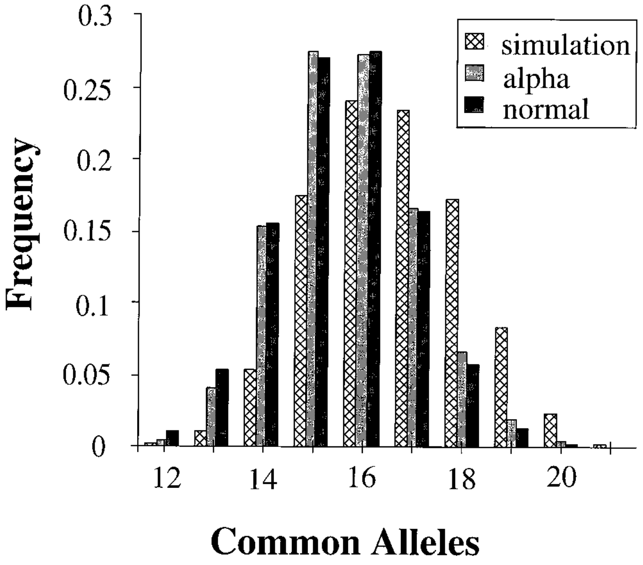

Figure 1.—Equilibrium distribution of the number of common alleles in a population of constant size. Cross-hatched bars represent results from 1000 simulated populations ($N = 5000$, $u = 0.00001$, $s = 0.1$). The solid bars are the expected equilibrium distribution of the Markov chain (light gray) and an approximation of this distribution (dark gray) fitted to a normal distribution using mean $\hat{i}$ (Equation 5) and variance $\hat{i}^3/(2S)$ (see text).

mations provide a good fit to simulation results (shown later in Figure 2).

We can use the Markov chain to predict the stationary probability distribution of i in the population, α_i. Because of the simple form of the chain, we can use standard theory to find α_i

$$\alpha_i = \alpha_1 \frac{\lambda_1 \ldots \lambda_{i-1}}{\mu_2 \ldots \mu_i}, \tag{6}$$

where α_1 is chosen so that the α_i sum to 1 (Ewens 1979, p. 74). Equation 6 provides a good fit to the simulations (presented later in Figure 1). Figure 1 also shows that α_i is very close to the values of a normal distribution with mean equal to $\hat{i}$ and variance equal to $\hat{i}^3/(2S)$. Thus, it is not necessary to use (6) for most practical purposes.

Simulation results: To test the accuracy of our analytic approximation, we developed a computer simulation of our model. We assumed a population of N diploid individuals. A generation consisted of mutation, viability selection, and random mating. Beginning with N zygotes, the program generated a Poisson distributed random number of mutations with mean $2Nu$. For each mutation, one of the $2N$ copies was chosen at random to mutate to an allele not found in the population. It was possible for the same copy to mutate more than once, but because $2Nu$ was one or less in our simulations, double mutations were extremely rare. Viability selection was modeled by computing the deterministic change in allele frequencies given the allele frequencies and selection parameter. Then random mating was modeled by randomly drawing $2N$ new alleles from a multinomial distribution with allele frequencies given by the values after selection. We began each simulation with the population initially fixed for one allele and waited until a stochastic equilibrium was reached before recording summary statistics. The waiting time depended primarily on the mutation rate and was always $<1/u$ generations.

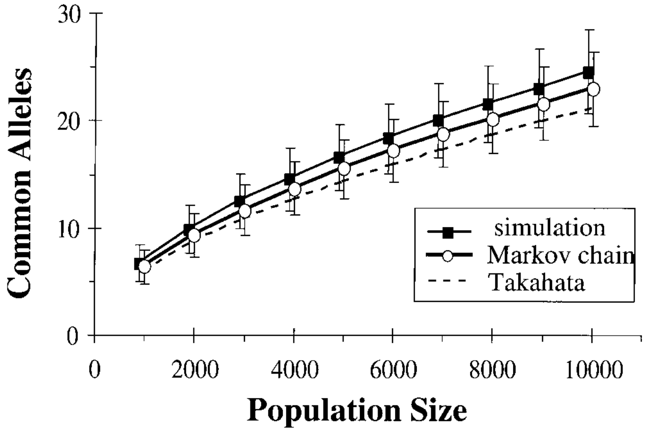

Figure 2.—Accuracy of predictions for a range of (constant) population sizes (for all cases, $s = 0.1$, $u = 0.00001$). Solid squares are from simulations (1000 replicates), open circles from Markov chain prediction. Error bars represent two standard deviations. The dashed line shows a different approximation to the mean number of common alleles (from Takahata 1990).

In the simulations, we recorded the numbers of common alleles at each time. We used Takahata's (1990) threshold for a common allele, his value $\delta = 1/(4Nsm)$, where $m = F - u/s$ and F is the homozygosity. We found that the exact value of this threshold made little difference. Figure 1 shows some typical results for the distribution of i, the number of common alleles across replicates, compared to normal distribution with the mean and variance computed from the Markov chain approximation. The fit is close although not perfect. There are slightly more common alleles found than predicted. Figure 2 shows the average number of common alleles as a function of N. These results show that there is little difference between the predictions of Equation 5 and Takahata's (1990) result (his Equation 5) and that both provide good approximations to the simulated results. Although in Figure 2 the results from the Markov chain are slightly closer to the simulation results, that is not always the case for other combinations of parameters.

POPULATION BOTTLENECK

We can use our method to predict the response to changes in population size. These predictions can be

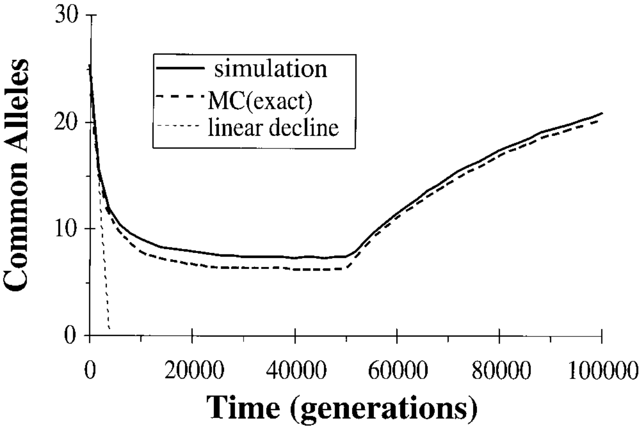

Figure 3.—Accuracy of Markov chain approach under dynamic conditions. At time 0, the initial (equilibrium) population of size $N = 100{,}000$ ($u = 0.000001$, $s = 0.01$) experiences a bottleneck of size $N_{\mathrm{bot}} = 10{,}000$ lasting for 50,000 generations, at which point the population size rebounds instantaneously to the initial size. The solid curve shows the average number of common alleles in (100) simulated populations undergoing this simple bottleneck, the heavy dashed curve shows the Markov chain prediction for the expected number of alleles over the same time course. The light dashed line represents a linear approximation to the rate of allele loss, which is based on the probability of loss in the first generation of the bottleneck.

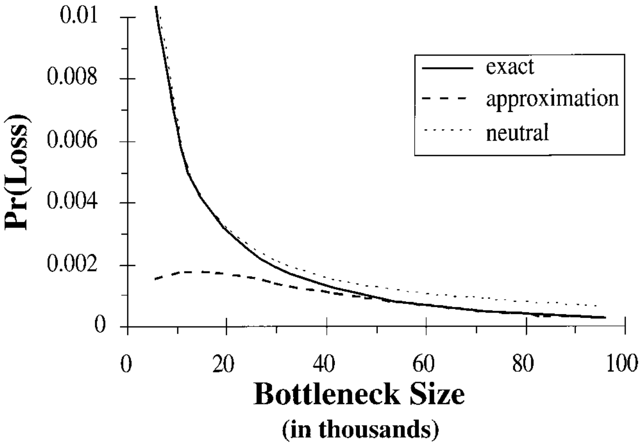

Figure 4.—Boundary conditions for behavior of alleles under a population bottleneck. The initial probability of losing an allele from a recently bottlenecked population is shown (original size $N_{\mathrm{init}} = 100{,}000$, $u = 0.000001$, $s = 0.01$), which varies according to the size of the bottleneck. The solid line represents an exact integral representation of this probability, the dashed line a closed-form approximation to this integral (Takahata 1990). For small bottlenecks, overdominant alleles behave approximately as if neutral (dotted line), and there is a transition zone in which neither the closed-form approximation nor the neutral expression can adequately predict the probability of allele loss.

obtained in two ways. The first is to use the solution to the equation that treats i as a deterministic variable with λ_i and μ_i as the deterministic rates of increase and decrease. That is, we use the equation

$$\frac{di(t)}{dt} = \lambda_i(t) - \mu_i(t)$$

$$= \frac{4N(t)us}{i} - \frac{1}{2N(t)i}e^{-2N(t)s/i^2}\sqrt{\frac{4N^3(t)s^3}{\pi}}, \quad (7)$$

which does not appear to have a simple analytic solution for arbitrary $N(t)$. We can get an idea of the time scale of response to changes in population size by assuming that the population is initially at size N_0 and that i is at its equilibrium given by (5) with $N = N_0$. We call that value of i, i_0. Then the population size changes abruptly to N_1, which may be larger or smaller. The initial rate of increase or decrease in i is given by the right-hand side of (7) with $N(t) = N_1$ and $i = i_0$. That is,

$$\frac{di(t)}{dt} \cong \lambda_i(i_0, N_1) - \mu_i(i_0, N_1), \quad (8)$$

where $\lambda_i(i_0,N_1)$ and $\mu_l(i_0,N_1)$ are the values at $i = i_0$ and $N = N_1$. For small t this approximate equation has the solution

$$i(t) = i_0 + [\lambda_i(i_0, N_1) - \mu_i(i_0, N_1)]t = i_0 \pm \frac{t}{T}, \quad (9)$$

where T is introduced to define a time scale associated with the initial increase or decrease in i. If there were a bottleneck in population size, meaning that $N_1 \ll N_0$, then the population size would have to be reduced for at least T generations for a substantial number of common alleles to be lost. We found that if N_1 is small enough, Equation 3 cannot be used to compute μ and numerical analysis is necessary.

The second and more accurate way to predict the effect of a change in population size is to iterate the Markov chain using the time-dependent values of λ_i and μ_i. Figure 3 compares the simulation results with the results from iterating the Markov chain and with the linear approximation, Equation 9.

In many discussions of the effect of a population bottleneck on variability at an overdominant locus, an important question is how rapidly are alleles lost as a result of the bottleneck. Our results support the conclusions of Takahata (1993). Figure 4 shows the probability of loss of an allele per generation, comparing the exact result based on the evaluation of $\bar{t}(x)$ by numerical integration with values based on assuming no selection and values based on Equation 3. We can see that the probability of loss under selection is always smaller than under neutrality, as expected. Overdominant selection retards the loss of common alleles. The analytic approximation, Equation 3, is quite accurate for a reduction in size by a factor of two but not for a larger reduction. The neutral approximation is accurate if the reduction is by a factor of five or more.

Takahata (1993) discusses the quantitative effects of a bottleneck in population size on the number of common alleles and discusses the survival of allelic lineages from before the bottleneck. Our results are consistent with Takahata's. The probability of loss of a lineage from before the bottleneck to any later time is, as Takahata (1993) argues, the integral of $\bar{t}(x)$ over that time period. Because all common alleles are equivalent in survival probability, the probability distribution of the number of lineages surviving from before the bottleneck is a binomial distribution with the probability being the probability of survival of a single lineage and the sample size being the number of lineages immediately before the bottleneck. The distribution of the number of new lineages is then obtained by subtraction.

Both Takahata's (1993) and our models of a bottleneck in population size assume a diffusion limit in which at most one common allele can be lost per generation. That assumption is appropriate if a bottleneck is not too sudden or small, and it appears to be appropriate for modeling the history of human populations. If the reduction in population size is extreme and rapid as might happen in the colonization of an island or other isolated region, then many alleles could be lost in a single generation. In that case, a single generation of random sampling would have to be included in the analysis to allow for the loss of several alleles in one generation.

These results support the discussions of Takahata (1990) and Ayala (1995) that for a bottleneck to lead to a substantial reduction in the number of alleles the ratio of the length of time during which the bottleneck occurs and the size during the bottleneck must exceed one, provided that the reduction in size is sufficiently great that the neutral approximation can be used. When the neutral approximation is not valid, the rate of loss is somewhat smaller and more alleles survive the bottleneck than expected on the basis of neutral theory. Although the differences between the rates of loss for neutral and overdominant alleles appear small in Figure 4, the rate for neutral alleles is two to three times that for overdominant alleles.

We have not been able to obtain a simple criterion for the ranges of parameter values for which the asymptotic approximation for strong selection, Equation 4, applies and when the asymptotic approximation for neutral alleles applies. The exact result depends in a complex way on i. The integral is easy to evaluate using a mathematical algebra program, so that is the best way to determine whether either of the asymptotic approximations can be used in any case.

TWO SELECTIVE CLASSES OF ALLELES

Our analysis of overdominant selection follows a long tradition of assuming equal fitnesses of all homozygotes and all heterozygotes. Even that case has proved difficult enough to analyze. The method we have developed allows us to test the robustness of results based on assuming homogeneity of selection on heterozygotes and homozygotes. We explore the possibility that there is heterogeneity in selection and show that relatively small differences in fitness can lead to large differences from the results previously obtained. We use a fitness scheme that leads to relatively simple algebra, but other assumptions about heterogeneity in fitness can be analyzed in the same way. Satta (1997) carried out computer simulations of a different model of heterogeneity in fitness (her asymmetric balancing selection model), in which the relative fitness of a heterozygote increased with the number of codons that differed between two alleles.

We assume that there are two classes of alleles, with i different alleles in class A and j different alleles in class a. The relative fitnesses are

$$\begin{aligned} A_kA_k&: \quad 1; \\ A_kA_l&: \quad 1+s,\ k\neq l; \\ a_ka_k&: \quad 1-2r; \\ a_ka_l&: \quad 1+s-2r,\ k\neq l; \\ A_ka_l&: \quad 1+s-r; \end{aligned} \qquad (10)$$

where A_k represents an allele in class A and a_k represents an allele in class a. That is, alleles in class a carry an intrinsic disadvantage that reduces fitness additively by r. This reduction in fitness could be caused by differences in pleiotropic effects of alleles in the two classes.

We first investigate the deterministic theory of this selection model. We let x be the equilibrium frequency of each allele in class A and y be the equilibrium frequency of each allele in class a ($ix + jy = 1$). It is relatively easy to show that at equilibrium

$$x = \frac{s+jr}{s(i+j)}, \quad y = \frac{s-ir}{s(i+j)}, \qquad (11)$$

which implies that $r < s/i$ for class a alleles to be present at all. In fact, r must be much smaller than this critical value for each a allele to be in high enough frequency for the SSWM approximation to be valid.

Using the same approach as in the previous section, we assume that the population contains i class A alleles, each in frequency x, and j class a alleles, each in frequency y, and then find the change in frequency of a new mutant of each type. It is straightforward to show that if z_A is the frequency of a new class A mutant and z_a is the frequency of a new class a mutant, then

$$\begin{aligned} \Delta z_A &\cong z_A\,(Fs + jyr) = sz_Am_A \\ \Delta z_a &\cong z_a\,(Fs - ixr) = sz_am_a, \end{aligned} \qquad (12a)$$

where

$$m_A = F + jy\frac{r}{s}, \quad m_a = F - ix\frac{r}{s}, \qquad (12b)$$

$F = ix^2 + jy^2$ is the homozygosity and s and r are assumed to be small. Therefore, the probability that a new class

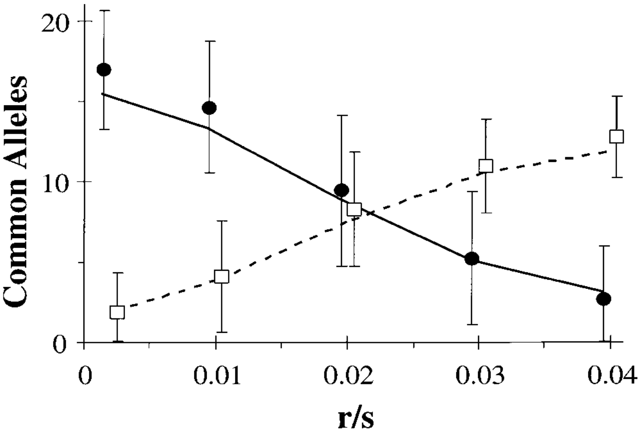

Figure 5.—Prediction of allele number in two selective classes. The number of common alleles (in 100 simulation runs) of type *A* (open squares) and *a* (closed circles) ±2 SD, plotted for a range of selective differences (r/s) between the two classes and compared to the theoretical prediction of Equation 15 (dashed line, type *A*; solid line, type *a*; $N = 10{,}000$, $u = 0.000001$, $v = 0.00001$, $s = 0.05$).

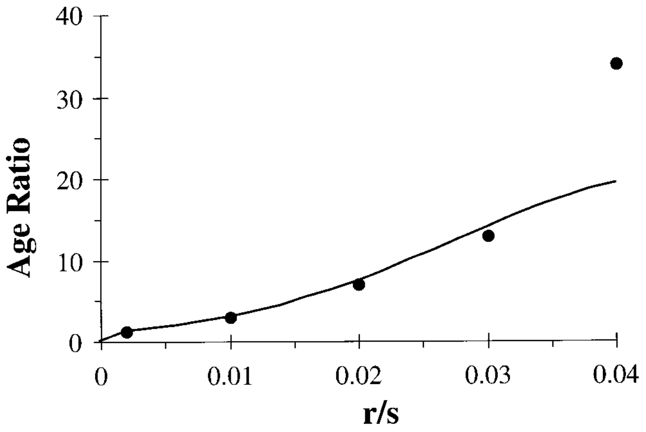

Figure 6.—Ratio of average ages of two selective classes of overdominant alleles. The ratio of the average age of type *A* alleles to the average age of type *a* alleles is plotted as a function of the selective difference (r/s) between the two classes (for all cases, $N = 10{,}000$, $u = v = 0.000001$, $s = 0.05$). The curve is the ratio predicted by the model, and the points represent sample simulation results, each point representing 100 simulation runs. For each type, the average allele age was computed as the mean average age across all (polymorphic) simulated populations, and average age within populations was weighted by copy number.

A or class *a* allele becomes common is approximately $2sm_A$ and $2sm_a$. If the mutation rate to new class *A* alleles is u and the rate to class *a* alleles is v, then the probabilities of i and j increasing by one are approximately

$$\lambda_i^{(A)} = 4N(t)usm_A, \quad \lambda_j^{(a)} = 4N(t)vsm_a. \tag{13}$$

The expected time to loss of a class *A* allele initially at frequency x, $\bar{t}_A(x)$, is given by Equation 3, where now x takes the value in Equation 11 and m is replaced by m_A, and the expected time to loss of a class *a* allele initially at frequency y defined in (11), $\bar{t}_a(x)$, is given by (3) with y replacing x and m_a replacing m. The probabilities per generation that i and j are reduced by 1 are then

$$\mu_i{}^{(A)} = \frac{i}{2N} e^{-2Nsm_A^2} \sqrt{\frac{(2Ns)^3 m_A^4}{2\pi}} \tag{14a}$$

$$\mu_j{}^{(a)} = \frac{j}{2N} e^{-2Nsm_a^2} \sqrt{\frac{(2Ns)^2 m_a^4}{2\pi}}, \tag{14b}$$

where the possible time dependence of N is not shown.

If the population size is constant, we can estimate the equilibrium numbers of alleles in both classes by solving the pair of equations,

$$\lambda_i{}^{(A)} = \mu_i{}^{(A)}, \quad \lambda_j{}^{(a)} = \mu_j{}^{(a)} \tag{15}$$

for i and j. There seems to be no analytic solution, but they are easily solved using a mathematical algebra program. If $r = 0$, then $m_A = m_a = F$, so (15) is solved by (4) with u replaced by $u + v$ and i replaced by $i + j$ and $i/(i + j) = u/(u + v)$. Figure 5 shows the dependence of the solutions to (15) on r/s for one pair of mutation rates.

We tested the accuracy of the Markov chain approximation for the case of two classes of alleles. The simulation program was the same as described above except for the modifications necessary to account for two classes of alleles. As shown in Figure 5, the average numbers of alleles are as predicted by the analytic theory.

If the population is constant in size, the average time to loss of alleles in the two classes is approximately the average allele age. The average ages of class *A* and *a* alleles can be quite different, even for relatively small values of the ratio r/s. By substituting into (3), we find

$$\frac{\bar{t}_A(x)}{\bar{t}_a(y)} = \frac{y^2 e^{S/x^2}}{x^2 e^{S/y^2}}, \tag{16}$$

where x and y are given by (11). The exponential dependence on $1/x^2$ and $1/y^2$ makes this ratio sensitive to small differences in x and y. Figure 6 shows the ratio of average ages as a function of r/s.

Figure 6 also shows the ratio of allele ages found in the simulations. For relatively small values of r/s the Markov chain accurately predicts the ratio of average ages (and both average ages as well), but for larger r/s the ratio of ages is even larger than predicted because alleles in the *a* class are so rare that their loss is no longer slowed by overdominant selection. Note that even very small values of r/s can result in ratios of average ages of five or larger.

DISCUSSION

We have shown that the SWMM approximation of Gillespie (1984, 1991) provides a simple, accurate way to predict the numbers of common alleles when there is strong overdominant selection. This approach allows us to make simple predictions about the loss of alleles during a bottleneck in population size and to explore the consequences of slight fitness differences among overdominant alleles. The cases we examined were chosen to illustrate the effects of varying population size and heterogeneous selection in simple contexts but the method can be easily generalized to provide approximate results for more complex and realistic cases. The success of the method depends only on the decoupling of the dynamics of new mutations from the dynamics of common alleles.

Our results are of relevance to the problem of understanding the effects of a bottleneck in population size on the number of alleles at a locus such as those in the major histocompatibility complex (MHC) in humans and other mammals. Many loci in the MHC are highly polymorphic and the high level of polymorphism is thought to be maintained by overdominance in fitness (Hughes and Nei 1988). Reduced levels of variation in the MHC can be caused by a population bottleneck (Ayala 1995; Ellegren *et al.* 1996; Miller *et al.* 1997; Wenink *et al.* 1998). In general, more overdominant alleles will survive a bottleneck than would be expected on the basis of neutral theory, but the neutral approximation will accurately predict the initial rate of loss of overdominant alleles after a large reduction in population size (Takahata 1993).

Our results are also of relevance to discussions of the relative ages of overdominant alleles. Under Takahata's (1990) theory, common alleles follow a coalescent model with a rescaled population size. According to that theory, the relative ages of alleles follow a geometric distribution. Our results with two selective classes of alleles show that even very slight heterogeneity in selective effects of different alleles can lead to large deviations from a geometric distribution. Ratios of ages of 5 or 10 are easily obtained when there is much less than a 1% difference in relative fitnesses. Slight differences in relative fitness are amplified by the long persistence time of both classes of alleles. Our model of two selective classes is not the only one possible; a similar approach could be used for three or more selective classes. The result for two selective classes tells us that, unfortunately, pleiotropic effects far too small to measure could cause substantial deviations from the coalescent approximation of Takahata (1990).

This research was supported in part by U.S. Public Health Service grant R01-GM40282 to M.S. C.A.M. also received support from U.S. Public Health Service grant T32-GM07127-24. We thank J. Gillespie and M. Turelli for helpful discussions of this topic, M. Uyenoyama for telling us about Sasaki's previous work on the SSWM approximation, and A. Sasaki for providing additional information about his Ph.D. dissertation.

LITERATURE CITED

Ayala, F. J., 1995 The myth of Eve: molecular biology and human origins. Science **270:** 1930–1936.

Ellegren, H., S. Mikko, K. Wallin and L. Anderson, 1996 Limited polymorphism at major histocompatibility complex (MHC) loci in the Swedish moose *A. alces*. Mol. Ecol. **5:** 3–9.

Ewens, W. J., 1979 *Mathematical Population Genetics*. Springer-Verlag, New York.

Gillespie, J. H., 1984 Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. **121:** 691–708.

Gillespie, J. H., 1991 *The Causes of Molecular Evolution*. Oxford University Press, Oxford.

Hughes, A. L., and M. Nei, 1988 Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335:** 167–170.

Kimura, M., and J. F. Crow, 1964 The number of alleles that can be maintained in a finite population. Genetics **49:** 725–738.

Klein, J., N. Takahata and F. J. Ayala, 1993 MHC polymorphism and human origins. Sci. Am. **269:** 78–83.

Miller, K. M., R. E. Withler and T. D. Beacham, 1997 Molecular evolution at Mhc genes in two populations of chinook salmon Oncorhynchus tshawytscha. Mol. Ecol. **6:** 937–954.

Newman, C. M., J. E. Cohen and C. Kipnis, 1985 Neo-darwinian evolution implies punctuated equilibria. Nature **315:** 400–401.

Sasaki, A., 1989 Evolution of Pathogen Strategies. Ph.D. Thesis, Kyushu University, Fukuoka, Japan.

Sasaki, A., 1992 The evolution of host and pathogen genes under epidemiological interaction, pp. 247–263 in *Population Paleo-Genetics*, edited by N. Takahata. Japan Scientific Society Press, Tokyo.

Satta, Y., 1997 Effects of intra-locus recombination of HLA polymorphism. Hereditas **127:** 105–112.

Takahata, N., 1990 A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc. Natl. Acad. Sci. USA **87:** 2419–2423.

Takahata, N., 1993 Evolutionary genetics of human paleo-populations, pp. 1–21 in *Mechanisms of Molecular Evolution*, edited by N. Takahata. Sinauer Associates, Sunderland, MA.

Vekemans, X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics **137:** 1157–1165.

Wenink, P. W., A. F. Groen, M. E. Roelke-Parker and H. H. T. Prins, 1998 African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. Mol. Ecol. **7:** 1315–1322.

Wright, S., 1939 The distribution of self-sterility alleles in populations. Genetics **24:** 538–552.

Communicating editor: N. Takahata