PERSISTENCE OF COMMON ALLELES IN TWO RELATED POPULATIONS OR SPECIES

WEN-HSIUNG LI AND MASATOSHI NEI

Center for Demographic and Population Genetics, University of Texas Health Science Center at Houston, Houston, Texas 77025

Manuscript received January 8, 1977
Revised copy received April 26, 1977

ABSTRACT

Mathematical studies are conducted on three problems that arise in molecular population genetics. (1) The time required for a particular allele to become extinct in a population under the effects of mutation, selection, and random genetic drift is studied. In the absence of selection, the mean extinction time of an allele with an initial frequency close to 1 is of the order of the reciprocal of the mutation rate when $4Nv << 1$, where N is the effective population size and v is the mutation rate per generation. Advantageous mutations reduce the extinction time considerably, whereas deleterious mutations increase it tremendously even if the effect on fitness is very slight. (2) Mathematical formulae are derived for the distribution and the moments of extinction time of a particular allele from one or both of two related populations or species under the assumption of no selection. When $4Nv << 1$, the mean extinction time is about half that for a single population, if the two populations are descended from a common original stock. (3) The expected number as well as the proportion of common neutral alleles shared by two related species at the tth generation after their separation are studied. It is shown that if $4Nv$ is small, the two species are expected to share a high proportion of common alleles even $4N$ generations after separation. In addition to the above mathematical studies, the implications of our results for the common alleles at protein loci in related Drosophila species and for the degeneration of unused characters in cave animals are discussed.

WHEN two populations are reproductively isolated, their genes are gradually differentiated from each other, owing to mutation, selection, and random genetic drift, and the number of common alleles shared by the two populations declines. We may then ask: How long does a particular allele that was present in the ancestral population persist in the two descendant populations? What is the expected number of common alleles shared by the two populations at the tth generation after population splitting? What is the probability that the two populations are monomorphic for the same allele at the tth generation? The last problem has already been studied by NEI and LI (1975) under the assumption of neutral mutation, but the first two questions have yet to be answered. These problems are of considerable importance in evolutionary studies. Since the allelic differences at a structural locus can now be studied by examining the

amino acid sequence or the electrophoretic mobility of the protein produced, the mathematical theory developed may be directly applicable to real data.

In the present paper we shall attempt to answer the first two questions. Before answering these questions, however, we need to study a simpler problem: How long does it take for a particular allele to become extinct in a *single* population? For those alleles which are segregating in the initial population, this problem has been studied by KIMURA and OHTA (1969a,b), disregarding the effect of mutation. In practice, however, the effect of mutation cannot be neglected in long-term evolution. Thus, an allele may be temporarily fixed in a population, but in the presence of mutation pressure it will eventually be eliminated from the population, unless it is maintained by some strong selective force. Note that at the molecular level new mutations are almost always different from the extant alleles. The extinction time of an allele in a population under the effect of irreversible mutation can be studied by using the method of EWENS (1964) and NAGYLAKI (1974). Furthermore, using a different method, NEI (1976) obtained the distribution as well as the mean and variance of the fixation time of a neutral allele when mutation is recurrent. This distribution is directly applicable to the present problem if the direction of mutation is reversed. In the present paper we shall consider both neutral and selective genes.

Although the primary purpose of this paper is to study the persistence of common alleles in two related populations, the theory developed will also be important for understanding some other problems such as degeneration of unused characters and the role of slightly deleterious mutations in evolution (OHTA's (1974) hypothesis).

EXTINCTION TIME OF AN ALLELE IN A POPULATION

Let us consider a random mating population of effective size N. We denote the allele under consideration by A. We assume that this allele mutates to other alleles, but mutation from other alleles to this allele is practically negligible. We consider all alleles other than A as a single class and denote them by a. Let v be the mutation rate per gene per generation. Therefore, A mutates to a irreversibly at the rate of v per generation. For simplicity, we consider only genic selection and assume that the relative fitnesses of genotypes AA, Aa, and aa are 1, $1+s$, and $+2s$, respectively. Thus, $s>0$ ($s<0$) means that all new mutations have a selective advantage (disadvantage) of s over A, whereas $s=0$ means that new mutations are all neutral. Now let the initial frequency of A be p and the frequency at generation t be x. Our situation is then equivalent to EWENS' (1964) diallelic model: A mutates to a irreversibly and absorption occurs only at $x=0$. The mean sojourn time, $\tau(p,x)$, of A at a particular gene frequency class x, starting from the initial frequency of p to its extinction, *i.e.*, $x=0$, has been studied by EWENS (1964) and NAGYLAKI (1974). The mean extinction time is given by $T(p)=\int_0^1 \tau(p,x)\,dx$, or more explicitly

$$T(p)=4N\left[\int_0^p \frac{g(0,x)e^{-Sx}}{x(1-x)^{1-M}}\,dx+g(0,p)\int_p^1 \frac{e^{-Sx}}{x(1-x)^{1-M}}\,dx\right], \qquad (1)$$

where $M = 4Nv$, $S = 4Ns$, and

$$g(a,b) = \int_a^b e^{Sx}(1-x)^{-M}dx \ .$$

When $p = 1$ and mutations are neutral, formula (1) is simplified and given by

$$T(1) = \sum_{i=1}^{\infty} \frac{4N}{i(M+i-1)} \ . \tag{2}$$

As mentioned earlier, allele A may become temporarily fixed in the population, owing to random genetic drift. It is therefore interesting to know the expected number of generations in which the frequency of A is higher than $1 - \alpha$, where α is a small quantity. We denote this expected number of generations by $t(p, 1-\alpha)$ and call it the mean sojourn time at the monomorphic class. This mean sojourn time is given by

$$t(p, 1-\alpha) = \int_{1-\alpha}^{1} \tau(p,x)dx \ , \tag{3a}$$

which becomes

$$t(1, 1-\alpha) = 4N \int_{1-\alpha}^{1} \frac{g(0,x)e^{-Sx}}{x(1-x)^{1-M}} dx \ , \tag{3b}$$

when $p = 1$. If $M < 1$, formula (3b) can be approximated by

$$t(1, 1-\alpha) \approx \frac{4N\alpha^M}{M} \int_0^1 (1-x)^{-M} e^{-S(1-x)}dx \ , \tag{4}$$

In the case of advantageous mutations ($S > 0$), the above formula may be written as

$$t(1, 1-\alpha) \approx 4NS(\alpha S)^M M^{-1}\Gamma(1-M)P(1-M,S) \ ,$$

where $P(\cdot,\cdot)$ denotes the incomplete gamma function (ABRAMOWITZ and STEGUN 1964). For neutral mutations ($S = 0$), formula (3a) reduces to

$$t(p, 1-\alpha) \approx \frac{4N}{M(1-M)} \alpha^M[1-(1-p)^{1-M}] \ , \qquad M \neq 1, \tag{5a}$$

$$= 4N \log_e(1-p) \log_e(1-\alpha) \ , \qquad M = 1, \tag{5b}$$

when $p \leq 1 - \alpha$, and

$$t(1, 1-\alpha) \approx \frac{4N}{1-M} \left[\frac{\alpha^M}{M} - \alpha\right] \ , \qquad M \neq 1, \tag{5c}$$

$$\approx 4N\alpha[1-\log_e\alpha] \ , \qquad M = 1, \tag{5d}$$

when $p = 1$.

The higher moments of extinction time may be obtained by following the method of NAGYLAKI (1974) (see also MARUYAMA and KIMURA 1971), but they are too complicated to be of practical use. In the case of neutral mutations, how-

ever, the probability distribution as well as the higher moments of extinction time have been obtained by NEI (1976), though the direction of mutation has to be reversed. Namely, the probability that A becomes lost by the tth generation is given by

$$f(p,0;t) = 1 - p\sum_{i=1}^{\infty}(-1)^{i-1}\frac{\Gamma(i-1+M)\,(M+2i-1)}{\Gamma(M)i!} \times F(1-i, i+M, M, 1-p)\exp\{-\lambda_i t\} \quad . \tag{6}$$

In the above formula, $F(\cdot,\cdot,\cdot,\cdot)$ denotes the hypergeometric function and $\lambda_i = i(M+i-1)/4N$. By using this distribution, it can be shown that the nth moment of extinction time is

$$\mu'_n(p) = n!p\sum_{i=1}^{\infty}(-1)^{i-1}\frac{\Gamma(i-1+M)\,(M+2i-1)}{\Gamma(M)i!(\lambda_i)^n} \times F(1-i, i+M, M, 1-p) \quad , \tag{7}$$

(NEI 1976). The above series converges for all M if $p < 1$, though the rate of convergence is very slow if M is large and p is close to 1. However, if $p = 1$, it converges only if M is smaller than 3.

We note that for the case of neutral mutations, the mean extinction time of allele A can be computed either by formula (1) or by formula (7). Theoretically, these two formulae should give the same value, because they are based on the same diffusion model, though derived in different ways. It can be shown that they are identical for the following two values of M:

$$T(1) = \mu'_1(1) = 4N\pi^2/6, \quad \text{if } M = 1 \quad ,$$
$$T(1) = \mu'_1(1) = 4N; \quad \text{if } M = 2 \quad .$$

In addition, for $M << 1$, both formulae give the same approximate value $4N + 1/v$ if $p = 1$. It seems, however, difficult to show that they are identical in general.

Figure 1 shows the probability distribution of extinction time for $p = 0.5$, assuming $4Nv = 0.1$. The distribution resembles the gamma distribution and is very similar to that of the conditional fixation time of a single neutral mutation (KIMURA 1970), though the mean is about 6 times larger compared with the latter case (see Table 2). There is virtually no probability of A becoming extinct before $0.1N$ generations. The probability density, however, increases sharply as t becomes larger than $0.1N$ generations and reaches its maximum at $t = 1.2N$ generations. It then declines, first rapidly and then gradually. Thus, the probability distribution has a long tail. The shape of the probability distribution of extinction time for $p = 1$ is similar to that for $p = 0.5$ (see Figure 2, NEI 1976), though the mean is much larger and the maximum density is located at $t = 8N$ generations. Numerical computations have shown that the shapes of the distributions for $4Nv = 0.01$ are very similar to those for $4Nv = 0.1$, but they are more flatly distributed and have a larger mean.

Tables 1 and 2 show the mean extinction times and their decompositions

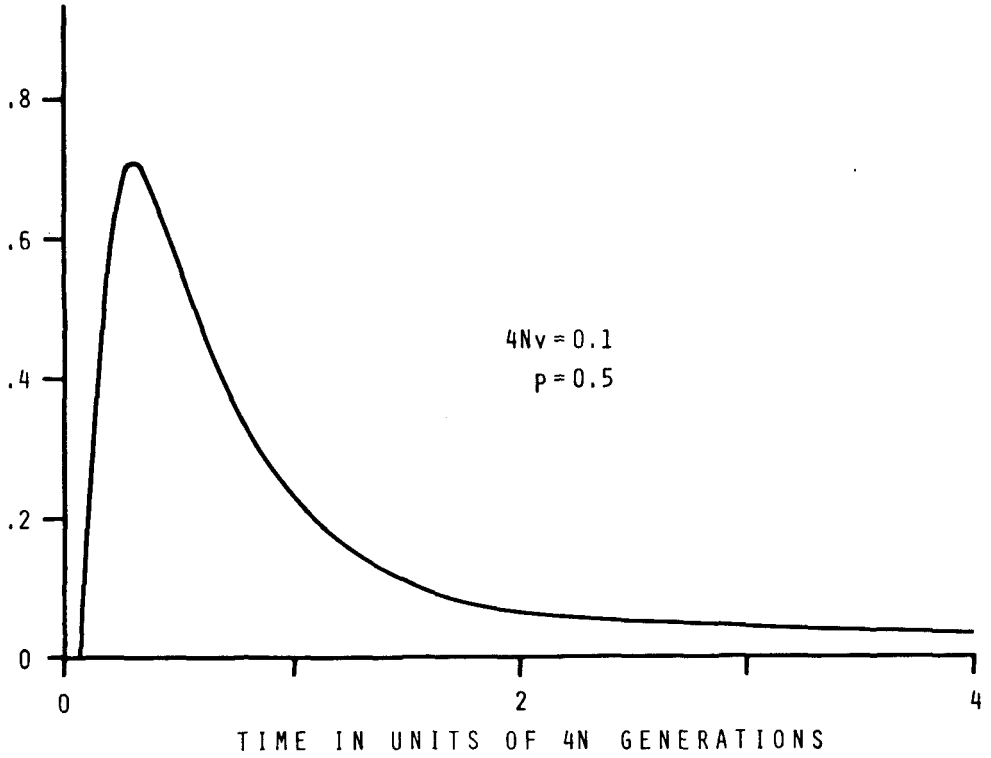

FIGURE 1.—Probability distribution of extinction time of a neutral allele under neutral mutation pressure for the case of $p = 0.5$ and $4Nv = 0.1$. Time is measured in units of $4N$ generations. This distribution has mean $\mu_1 = 5.80$, standard deviation $\sigma = 8.76$, skewness $\gamma_1 = 2.63$, and kurtosis $\gamma_2 = 9.6$.

into mean sojourn times for several gene frequency intervals for neutral and advantageous mutations. It is clear that in the case of $p = 1$ both mutation and selection have a strong effect on the mean extinction time (Table 1). Let us first consider the effects of the variations in $M = 4Nv$ and $S = 4Ns$ when N is fixed. In the case of neutral mutations the mean extinction time for a given value of N is roughly inversely proportional to mutation rate if $M \leq 1$. Advantageous mutation decreases the extinction time, as expected. If $S = 100$ and $M = 0.001$, the mean extinction time is about 100 times shorter than that for the case of $S = 0$ and $M = 0.001$. It is interesting to note that for the three cases of (1) $S = 0$; $M = 0.1$, (2) $S = 10$; $M = 0.012$, and (3) $S = 100$; $M = 0.001$, the mean extinction time is nearly the same. Yet if we look at the sojourn times for the gene frequency intervals (0.099), (0.99, 0.999), and (0.999, 1), there is a tendency for the proportion of the time spent for (0.999, 1), i.e. $t(1, 0.999)$ in (3a), to increase as S increases. This tendency occurs because selection is not very effective in this interval and gene frequency change is largely determined by mutation and random genetic drift.

When mutation rate is fixed, but population size varies, some caution is re-

TABLE 1

Mean extinction times and their decompositions into mean sojourn times in three gene frequency intervals with the initial frequency of p = 1

Selection intensity	Mutation pressure	Mean extinction time*	Mean sojourn times* (0,0.99)	(0.99,0.999)	(0.999,1)
S†= 0	M‡= 1	1.65	1.59	0.048	0.008
	M = 0.1	10.94	3.94	1.44	5.56
	M = 0.01	101.0	4.53	2.20	94.3
	M = 0.001	1001.0	4.60	2.29	994.1
S = 10	M = 1	0.544	0.514	0.025	0.005
	M = 0.1	1.53	0.700	0.172	0.662
	M = 0.01	10.42	0.732	0.224	9.46
	M = 0.001	99.07	0.736	0.231	98.1
S = 100	M = 1	0.104	0.0911	0.0107	0.003
	M = 0.1	0.197	0.0969	0.0205	0.080
	M = 0.01	1.05	0.0976	0.0228	0.934
	M = 0.001	9.61	0.0977	0.0231	9.49

* Time measured in units of $4N$ generations.
† $S = 4Ns$.
‡ $M = 4Nv$.

quired in the interpretation of Table 1 (and 2), since the unit of the extinction time in this table is $4N$ generations. Thus, in a small population the absolute mean extinction time may be small even if the value in this table is relatively large. If we note this property, it is clear that the mean extinction time for the case of $S = 0$ and $4N << 1/v$ is almost independent of population size and approximately given by $1/v$. When there is selection, the mean extinction time

TABLE 2

Mean extinction times and their decompositions into mean sojourn times in four gene frequency intervals with the initial frequency of p = 0.5

Selection intensity	Mutation pressure	Mean extinction time*	(0,0.5)	Mean sojourn times* (0.5,0.99)	0.99,0.999)	(0.999,1)
S†= 0	M‡= 1	1.06	0.582	0.474	0.006	0.0007
	M = 0.1	5.80	0.68	1.86	0.67	2.59
	M = 0.01	50.8	0.69	2.25	1.10	46.8
	M = 0.001	500.8	0.69	2.29	1.15	496.7
S = 10	M = 1	0.29	0.26	0.029	10^{-5}	10^{-6}
	M = 0.1	0.32	0.27	0.042	0.001	0.004
	M = 0.01	0.39	0.28	0.045	0.002	0.06
	M = 0.001	0.99	0.28	0.045	0.002	0.67
S = 100	M = 1	0.052	0.051	0.0004	6×10^{-26}	4×10^{-27}
	M = 0.1	0.052	0.052	0.0004	4×10^{-24}	10^{-23}
	M = 0.01	0.052	0.052	0.0004	6×10^{-24}	2×10^{-22}
	M = 0.001	0.052	0.052	0.0004	7×10^{-24}	2×10^{-20}

* Time measured in units of $4N$ generations.
† $S = 4Ns$.
‡ $M = 4Nv$.

generally decreases as N increases. Thus, in the case of $S = 10$ and $M = 0.1$, $T(1)$ is $10 \times 4N$ generations, whereas in the case of $S = 100$ and $M = 0.1$ where N is 10 times larger, $T(1)$ is $0.8N$ generations. Decomposition of $T(1)$ into its sojourn times shows that the reduction in $T(1)$ in large populations is due to the reduction of $t(1, 0.999)$. Namely, in large populations the mean sojourn time at the monomorphic class with $\alpha = 0.001$ is much shorter than in small populations. However, the absolute sojourn time for the interval (0, 0.99) is larger for the case of $S = 100$ and $M = 0.1$ than for the case of $S = 10$ and $M = 0.01$. Of course, for a given value of M selection always reduces the sojourn time for any gene frequency interval and consequently the mean extinction time.

When the initial gene frequency is intermediate, selection seems to have a stronger effect on extinction time than mutation (Table 2). In the case of $p = 0.5$ and $S = 100$, the mean extinction time for a given value of N is virtually the same for all the mutation rates considered. This is because the advantageous mutation existing in the original population is almost always fixed rather quickly by selection. When S is as small as 10, however, the effect of mutation is appreciably large. In the absence of selection the extinction time for $p = 0.5$ is about half that for $p = 1$ when $M << 1$. This is because for $M << 1$ and $p > 0.4$, $\mu_1'(p)$ in (7) is approximately given by $p(M + 1)/v$. An intuitive explanation of this formula is as follows: When the initial gene frequency of allele A is p and M is small, allele A is quickly eliminated from the population with a probability of about $1 - p$, whereas it is temporarily fixed with a probability of about p. In the latter event it takes about $(M + 1)/v$ generations for the allele to be eventually lost (NEI 1976). Thus, we have $\mu_1'(p) = p(M + 1)/v$ approximately.

Recently, OHTA (1974) emphasized the importance of slightly deleterious mutations in molecular evolution. She postulates that the level of protein polymorphism in large populations is determined mainly by mutation-selection balance, whereas allele substitution in evolution occurs by random genetic drift. In this hypothesis the rate of allele substitution is expected to be slowed down in large populations. How long then does it take for an allele to be replaced by slightly deleterious genes under irreversible mutation? This question can be answered by studying the mean extinction time for negative values of S. This mean extinction time is presented in Table 3 for various values of S and M. It is clear that if S is smaller than -50, the extinction time is very large even if M is

TABLE 3

Mean times for a type allele to be replaced by slightly deleterious alleles under recurrent mutation. The initial gene frequency is* 1

M‡\S†	−2	−5	−10	−20	−50	−100
10	0.35	0.41	0.41	4.3	6.1×10^{8}	1.2×10^{27}
1	2.79	9.93	281.38	1.3×10^{6}	2.1×10^{18}	2.5×10^{39}
0.1	30.91	251.23	1.7×10^{3}	1.7×10^{8}	6.7×10^{20}	1.6×10^{42}

* Time measured in units of $4N$ generations.
† $S = 4Ns$.
‡ $M = 4Nv$.

as large as 10. This suggests that the replacement of a type allele by slightly deleterious alleles is an extremely slow process in large populations.

PERSISTENCE OF AN ALLELE IN TWO POPULATIONS

Let us now consider how long a particular allele persists in two populations derived from a foundation stock. In this and following sections, we shall consider only neutral alleles, since selection introduces a number of difficulties. We shall also assume that the two descendant populations are reproductively isolated immediately after separation. We denote by N_1 and N_2 the effective sizes of populations 1 and 2, respectively. We also denote by x and y the frequency of a particular allele, A, in the populations 1 and 2, respectively, with $x = p$ and $y = q$ at $t = 0$. Let $f(p,0;t)$ be the probability that A becomes lost in population 1 by the tth generation and $g(q,0;t)$ be the corresponding probability in population 2. Then, the probability that A still exists in both populations at generation t is given by

$$Q(p,q;t) = [1 - f(p,0;t)]\,[1 - g(q,0;t)]\ . \tag{8}$$

The probabilities f and g can be obtained from (6) by a proper substitution of parameters. The probability given by (8) may be called the probability of common existence of A.

This probability includes the cases where A exists in a low frequency in one or both of the two populations. In practice, however, low frequency alleles may not be detected by usual surveys of allele frequencies. In this case it would be more appropriate to define the probability of common existence by considering only those cases where the frequency of A is equal to or higher than α in both populations, in which α is a small quantity such as 0.01. Namely, at generation t

$$Q(p,q;\alpha;t) = P(x \geq \alpha,p;t)P(y \geq \alpha,q;t)\ , \tag{9}$$

where $P(x \geq \alpha,p;t)$ is the probability of x being equal to or larger than α in population 1. It is given by

$$\begin{aligned} P(x \geq \alpha,p;t) = p(1-\alpha)^{M_1} \sum_{i=0}^{\infty} & \frac{(M_1+1+2i)\Gamma(M_1+1+i)\Gamma(M_1+i)}{i!(i+1)!\Gamma^2(M_1)M_1} \\ & \times F(-i,i+M_1+1,M_1,1-p) \\ & \times F(-i,i+M_1+1,M_1+1,1-\alpha)e^{-[(i+1)(M_1+i)/4N_1]t}\ , \end{aligned} \tag{10}$$

where $M_1 = 4N_1v$ (see formula (3) of NEI and LI (1975)). $P(y \geq \alpha, q;t)$ is the corresponding probability for population 2.

For a large t with $(2v + 1/2N_i)t >> 1$ $(i = 1,2)$, (9) is given by

$$Q(p,q;\alpha;t) = (1-\alpha)^{M_1+M_2}(M_1+1)\,(M_2+1)pq\,e^{-2vt} \tag{11}$$

approximately. The corresponding formula for $Q(p,q;t)$ is $(M_1 + 1)(M_2 + 1)pge^{-2vt}$. Thus, if $M_1 + M_2$ and α are small, the difference between $Q(p,q;t)$ and $Q(p,q;\alpha;t)$ is very small. Indeed, numerical computations have shown that, for any value of t, $Q(p,q;\alpha;t)$ is only slightly smaller than $Q(p,q;t)$, if $M_1 + M_2 \leq 1$ and $\alpha \leq 0.01$.

TABLE 4

Probabilities of common existence of a particular allele in two populations
p is the initial allele frequency in the ancestral population

Time in generations		0.4N	4N	40N	400N
$p=1$	$M=1$	1.000	0.409	6.7×10^{-9}	$<10^{-86}$
	$M=0.1$	1.000	0.944	0.160	2.4×10^{-9}
	$M=0.01$	1.000	0.995	0.834	0.137
$p=0.5$	$M=1$	0.929	0.113	1.7×10^{-9}	$<10^{-86}$
	$M=0.1$	0.947	0.309	0.040	6.1×10^{-10}
	$M=0.01$	0.949	0.335	0.208	0.034

Table 4 gives the values of $Q(p,q;\alpha;t)$ for the case of $N_1=N_2=N$, $p=q$, and $\alpha=0.01$. It is seen that an allele with an initial gene frequency of 0.5 or higher may exist in both populations for a very long time. For example, if $p=1$ and $M=0.1$, the probability of common existence at the $40N$th generation is 0.16. Thus, in organisms with $N=10^6$, two related species may possess the common allele even 40 million generations after separation. As expected, the probability of common existence decreases faster with increasing time when M is large and p is small than when M is small and p is large.

If we use expression (8), the mean time for A to become lost from one or both of the two populations can easily be obtained. Let us assume $p=q$ and $N_1=N_2=N$ for simplicity. Then, the probability density, $G(p,0;t)$, that A is lost from one or both populations at generation t is given by

$$G(p,0;t)=\frac{\partial}{\partial t}[1-\{1-f(p,0;t)\}^2]$$
$$=2[1-f(p,0;t)]\frac{\partial f(p,0;t)}{\partial t}\ . \tag{12}$$

The nth moment of this probability distribution is

$$E(t^n)=2\int_0^\infty t^n[1-f(p,0;t)]\frac{\partial f(p,0;t)}{\partial t}\,dt$$
$$=n!2p^2\sum_{j=1}^{\infty}\sum_{i=1}^{\infty}(-1)^{i+j}\frac{\Gamma(j-1+M)\Gamma(i-1+M)(M+2j-1)(M+2i-1)}{\Gamma^2(M)i!j!(\lambda_i+\lambda_j)^{n+1}}\lambda_j$$
$$\times F(1-j,j+M,M,1-p)F(1-i,i+M,M,1-p)\ . \tag{13}$$

This formula is subject to the same condition of covergence as that of formula (7). If $p=1$ and M is about 0.1 or less, the first two moments are given by

$$E(t)\approx(M+1)^2/(2v)\ , \tag{14a}$$
$$E(t^2)\approx(M+1)^2/(2v^2)\ . \tag{14b}$$

Therefore, in this case the mean time for A to be lost from one or both populations is about half the expected extinction time in a single population.

NUMBER OF COMMON ALLELES SHARED BY TWO POPULATIONS

Formula (8) gives the probability that a particular allele exists in two related populations. Therefore, if there are k alleles in the initial populations, the expected number of common alleles that exist at the tth generation is given by $\sum_{i=1}^{k}[1-f(p_i,0;t)][1-g(q_i,0;t)]$, where p_i and q_i are the initial frequencies of the ith allele (A_i) in populations 1 and 2, respectively. In practice, however, it is more meaningful to consider the number of common alleles in samples of given sizes. Suppose that m_1 individuals ($2m_1$ genes) are sampled from population 1 and m_2 individuals are sampled from population 2. The probability density that the frequency of A_i in population 1 is x at generation t is given by

$$\begin{aligned}\phi(p_i,x;t) = p_i \sum_{j=0}^{\infty} & \frac{(M_1+2j+1)\Gamma(M_1+j+1)\Gamma(M_1+j)}{j!(j+1)!\Gamma^2(M_1)} \\ & \times F(-j,M_1+j+1,M_1,1-p_i) \\ & \times (1-x)^{M_1-1}F(-j,M_1+j+1,M_1,1-x)\exp\left\{-\frac{(j+1)(M_1+j)t}{4N_1}\right\}\end{aligned}$$

(*cf.* CROW and KIMURA 1970; NEI and LI 1975). Thus the probability that A_i appears in a sample of $2m_1$ genes, is given by

$$\begin{aligned}P_1(p_i,t;2m_1) &= \int_0^1 [1-(1-x)^{2m_1}]\phi(p_i,x;t)\,dx \\ &= p_i \sum_{j=0}^{\infty} \frac{(M_1+2j+1)\Gamma(M_1+j)}{(j+1)!\Gamma(M_1)} F(-j,M_1+j+1,M_1,1-p_i) \\ &\quad \times \left\{(-1)^j + \frac{\Gamma(M_1+j+1)}{j!\Gamma(M_1)} \sum_{l=0}^{j} \frac{(-j)_l(M_1+j+1)_l}{(M_1)_l l!(M_1+2m_1+l)}\right\} \\ &\quad \times \exp\left\{-\frac{(j+1)(M_1+j)t}{4N_1}\right\}, \qquad (15)\end{aligned}$$

where $(a)_l = a(a+1)\ldots(a+l-1)$. The probability, $P_2(q_i,t;2m_2)$, that A_i appears in a sample of $2m_2$ genes from population 2 can be computed in the same way. Therefore, the expected number of common alleles in the samples is given by

$$n(t) = \sum_{i=1}^{k} P_1(p_i,t;2m_1)P_2(q_i,t;2m_2)\,. \qquad (16)$$

If $N_1 = N_2 = N$ and the ancestral population was in equilibrium with respect to mutation and genetic drift with $M = 4Nv$, and this equilibrium is maintained

in each descendant population, we can obtain an explicit formula for (16). In this case the initial distribution of the number of alleles at different frequencies is given by $\Phi(p) = Mp^{-1}(1-p)^{M-1}$ (KIMURA and CROW 1964; see also WRIGHT 1949). Therefore,

$$n(t) = \int_0^1 Mp(1-p)^{M-1}\Big[\sum_{j=0}^{\infty} \frac{(M+2j+1)\Gamma(M+j)}{(j+1)!\Gamma(M)} F(-j,M+j+1,M,1-p) \times \Big\{(-1)^j + \frac{\Gamma(M+j+1)}{j!\Gamma(M)} \sum_{l=0}^{j} \frac{(-j)_l (M+j+1)_l}{(M)_l l!(M+2m+l)} \times \exp\Big\{-\frac{(j+1)(M+j)}{4N} t\Big\}\Big]^2 dp \ , \tag{17}$$

where $m_1 = m_2 = m$ is assumed.

Some numerical values of $n(t)$ obtained by (17) are given in Table 5. When M is large, $n(t)$ in the early generations is considerably large, as expected. In this case, however, $n(t)$ decreases rather rapidly as t increases. On the other hand, if M is small, $n(t)$ in the early generations is small but declines less rapidly with increasing t. In Table 5 the ratio of the expected number of common alleles to the total number of alleles in the two populations is also presented. It is computed by $n(t)/[2n_a - n(t)]$, where n_a is the expected number of alleles in a sample of $2m$ genes in one population. As expected, this ratio declines with increasing t but very slowly if M is small. For example, if $M = 0.01$, the ratio is about 90 percent even at the $4N$th generation. This indicates that the proportion of common alleles shared by the two populations remains high for a long time after their separation.

TABLE 5

Expected numbers of common alleles in samples of m *individuals from each of populations 1 and 2*

Time in generations		0.04N	0.4N	4N	40N	400N
$M = 1$	$m = 20$	3.25(0.61)	2.01(0.31)	0.26(0.03)	4×10^{-9}	$<10^{-86}$
	$m = 50$	3.79(0.58)	2.13(0.26)	0.26(0.03)	4×10^{-9}	$<10^{-86}$
	$m = 100$	4.05(0.53)	2.17(0.23)	0.27(0.02)	4×10^{-9}	$<10^{-86}$
	$m = 500$	4.33(0.41)	2.21(0.17)	0.27(0.02)	4×10^{-9}	$<10^{-86}$
$M = 0.1$	$m = 20$	1.34(0.91)	1.21(0.75)	0.93(0.49)	0.15(0.06)	$<10^{-9}$
	$m = 50$	1.39(0.86)	1.23(0.69)	0.93(0.45)	0.15(0.05)	$<10^{-9}$
	$m = 100$	1.42(0.83)	1.23(0.64)	0.93(0.42)	0.15(0.05)	$<10^{-9}$
	$m = 500$	1.45(0.72)	1.23(0.55)	0.93(0.37)	0.15(0.05)	$<10^{-9}$
$M = 0.01$	$m = 20$	1.03(0.98)	1.02(0.96)	0.99(0.91)	0.82(0.65)	0.13(0.07)
	$m = 50$	1.04(0.98)	1.02(0.94)	0.99(0.89)	0.82(0.64)	0.13(0.07)
	$m = 100$	1.04(0.96)	1.02(0.92)	0.99(0.88)	0.82(0.63)	0.13(0.07)
	$m = 500$	1.04(0.95)	1.02(0.91)	0.99(0.86)	0.82(0.62)	0.13(0.06)

The values in parenthess denote the value of $r(t)$.

DISCUSSION

Recent electrophoretic surveys of protein polymorphism have shown that two related species often share many common alleles (*e.g.*, SELANDER, HUNT and YANG 1969; AYALA *et al.* 1974). Some investigators interpreted these observations as an indication of selective maintenance of the alleles, since they must have been maintained in the population for a long time. The present study, however, shows that two related species may share common alleles for a long time after their separation even if there is no selection. In *Drosophila willistoni* group species the average heterozygosity (H) for protein loci is about 0.16, so that M is estimated to be about 0.2 from the formula $H = M/(M + 1)$. Under the assumption of neutral mutations, KIMURA and OHTA (1971) estimated that the mutation rate for electrophoretically detectable alleles is roughly 10^{-7} per locus per *year*. Many Drosophila species seem to have about 10 generations per year. Therefore, the mutation rate per *generation* is estimated to be 10^{-8}. Thus, if the netural mutation hypothesis is correct, the effective population size for the long-term evolution in *Drosophila willistoni* group is estimated to be about 5×10^6 (see NEI, MARUYAMA and CHAKRABORTY 1975 for the bottleneck effect). It can be shown that $r(t)$ is 0.15 for $M = 0.2$ and $t = 4N$. Therefore, two species in this Drosophila group may share an appreciable proportion of common alleles even 2×10^7 generations (2 million years) after separation.

In this connection it should be noted that our theoretical value of $r(t)$ is an underestimate when applied to electrophoretic data. This is because at the level of electrophoretically detectable proteins some back mutation may occur, contrary to our assumption of no back mutation. Recently there have been some efforts to study protein polymorphism in detail by using heat denaturation and other techniques (BERNSTEIN, THROCKMORTON and HUBBY 1973; SINGH, HUBBY and THROCKMORTON 1975). Our computation would apply more appropriately to data obtained by these techniques or by amino acid sequencing.

In a recent paper COYNE (1976) showed that at the xanthine dehydrogenase locus the proportion of alleles in common between *Drosophila pseudoobscura* and *D. persimilis* was $3/11 = 0.273$ when the ordinary electrophoresis was used, whereas it was $3/47 = 0.064$ when a more detailed study was made by using different gel concentrations and pH values (four levels of screening). It is interesting to note that the latter value is not far from what we expect under the neutral mutation theory if the time after divergence between these two species is of the order of $1.5N$ generations (see NEI and LI (1975) for the discussion of the divergence time between these two species). Namely, the M value for this locus when the four levels of screening were used seems to be about 2. since the average heterozygosity for the two species has been estimated to be 0.66. (The Drosophila xanthine dehydrogenase is an unusually large protein.) In the case of $M = 2$ and $m = 30$ (60 homozygous lines) we obtain $r(t) = 0.045$ for $t = 1.5N$, using formula (17). This is close to the observed value. Of course, data from a single locus are not very reliable for this type of argument, since the proportion of common alleles is expected to have a large sampling variance.

Our results on the extinction probability and mean extinction time of an allele

in a single population have a direct bearing on the problem of degeneration of unused characters in evolution. It is well known that many cave-dwelling animals lack pigmentation and eyesight. It has been controversial, however, whether these characters degenerated simply because of mutation pressure or because the individuals lacking these characters had a selective advantage in cave conditions (WRIGHT 1964; BARR 1968). If we know the time when a cave population was formed, this problem can be studied by using our mathematical formulae. For example, using AVISE and SELANDER's (1972) data on protein polymorphism, CHAKRABORTY and NEI (1974) estimated that the cave population of the characid fish *Astyanax mexicanus* in Pachon, Mexico, was formed about 710,000 ± 460,000 years ago. If we assume that the generation time of this fish is 5 years (P. SADOGLU, personal communication), this evolutionary time corresponds to about 140,000 generations. On the other hand, if the rate of mutations to *lethal or other nonfunctional alleles* is 10^{-5} per locus per generation, the mean time for the original allele to become lost from the population is about 100,000 generations, since the effective size of this population is very small (less than 200). Therefore, it is possible to explain the degeneration of eyes and pigmentation in this cave fish in terms of mutation pressure without assuming any selective advantage.

In the past the process of degeneration of unused characters has been studied mainly by using deterministic models. However, since population size is generally small when this type of evolutionary change occurs, stochastic models are much more appropriate. In the absence of selection the probability that the original allele becomes lost from the population by generation t is given by (6). When $p \geq 0.4$ and $M << 1$, this formula reduces to

$$f(p,0;t) = 1 - pe^{-vt}$$

approximately (see also CROW and KIMURA 1970). In the above example of *Astyanax mexicanus* this probability becomes 0.75 if we assume $v = 10^{-5}$, $p = 1$, and $t = 140{,}000$. Thus, the probability of loss of gene function is very high. On the other hand, if t is 14,000, the probability becomes 0.13. This indicates that even in such a short evolutionary time as 14,000 generations the loss of eyes and pigmentation may occur with an appreciable probability.

We thank the referees for their comments. This work was supported by grants from the Public Health Service and the National Science Foundation.

LITERATURE CITED

ABRAMOWITZ, M. and I. A. STEGUN, 1964 *Handbook of Mathematical Functions with formulas, graphs, and mathematical tables.* U.S. Dept. Commerce, Washington, D.C.

AVISE, J. C. and R. K. SELANDER, 1972 Evolutionary genetics of cave-dwelling fishes of the genus *Astyanax*. Evolution **26**: 1–19.

AYALA, F. J., M. L. TRACEY, L. G. BARR, J. F. MCDONALD and S. PÉREZ-SALAS, 1974 Genetic variation in natural populations of five Drosophila species and the hypothesis of the selective neutrality of protein polymorphisms. Genetics **77**: 343–384.

BARR, T. C., JR., 1968 Cave ecology and the evolution of troglobites. Evol. Biol. **2**: 35–102.

BERNSTEIN, S. C., L. H. THROCKMORTON and J. L. HUBBY, 1973 Still more genetic variability in natural populations. Proc. Natl. Acad. Sci. U.S. **70**: 3928–3931.

CHAKRABORTY, R. and M. NEI, 1974 Dynamics of gene differentiation between incompletely isolated populations of unequal sizes. Theor. Pop. Biol. **5**: 460–469.

COYNE, J. A., 1976 Lack of genic similarity between two sibling species of Drosophila as revealed by varied techniques. Genetics **84**: 593–607.

CROW, J. F. and M. KIMURA, 1970 *An Introduction to Population Genetics Theory*. Harper, New York.

EWENS, W. J., 1964 The pseudo-transient distribution and its use in genetics. J. Appl. Prob. **1**: 141–156.

KIMURA, M., 1970 The length of time required for a selectively neutral mutant to reach fixation through random frequency drift in a finite population. Genet. Res. **15**: 131–133.

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

KIMURA, M. and T. OHTA, 1969a The average number of generations until fixation of a mutant gene in a finite population. Genetics **61**: 763–771. ——, 1969b The average number of generations until extinction of an individual mutant gene in a finite population. Genetics **63**: 701–709. ——, 1971 Protein polymorphism as a phase of molecular evolution. Nature **229**: 467–469.

MARUYAMA, T. and M. KIMURA, 1971 Some methods for treating continuous stochastic processes in population genetics. Japan. J. Genet. **46**: 407–410.

NAGYLAKI, T., 1974 The moments of stochastic integrals and the distribution of sojourn times. Proc. Natl. Acad. Sci. U.S. **71**: 746–749.

NEI, M., 1976 Mathematical models of speciation and genetic distance. In: *Population Genetics and Ecology*. Edited by S. KARLIN and E. NEVO. Academic Press, New York.

NEI, M. and W.-H. LI, 1975 Probability of identical monomorphism in related species. Genet. Res. **25**: 31–43.

NEI, M., T. MARUYAMA and R. CHAKRABORTY, 1975 The bottleneck effect and genetic variability in populations. Evolution **29**: 1–10.

OHTA, T., 1974 Mutational pressure as main cause of molecular evolution. Nature **252**: 351–354.

SELANDER, R. K., W. G. HUNT and S. Y. YANG, 1969 Protein polymorphism and genic heterozygosity in two European subspecies of the house mouse. Evolution **23**: 379–390.

SINGH, R. S., J. L. HUBBY and L. H. THROCKMORTON, 1975 The study of genic variation by electrophoretic and heat denaturation techniques at the octanol dehydrogenase locus in members of the *Drosophila virilis* group. Genetics **80**: 637–650.

WRIGHT, S., 1949 Genetics of populations. Encyclopedia Britannica **10**: 111, 111A-D, 112.

WRIGHT, S., 1964 Pleiotropy in the evolution of structural reduction and of dominance. Am. Naturalist **98**: 65–69.

Corresponding editor: W. J. EWENS