PROBABILITY OF FIXATION OF NONFUNCTIONAL GENES AT DUPLICATE LOCI

MASATOSHI NEI* AND ARUN K. ROYCHOUDHURIY*

Division of Biological and Medical Sciences, Brown University, Providence, Rhode Island 02912

Haldane (1933) suggested that if a gene is duplicated by some mechanism such as chromosome doubling or unequal crossing over, one of the two resultant duplicate genes may become nonfunctional because of the fixation of a deleterious mutation (see also Ohno, Wolf, and Atkin 1968). Noting that gene duplication must have occurred frequently in the evolutionary process and that lethal mutations occur with an appreciable frequency in every generation, Nei (1969) postulated that a large number of nonfunctional genes have accumulated in the genomes of higher organisms. Recently, Ohta and Kimura (1971) estimated on theoretical grounds that more than 90% of the DNA in the mammalian genome is nonfunctional. Molecular data also suggest that the genomes of higher organisms include a large amount of noninformational DNA (Flamm, Walker, and McCallum 1969; Crick 1971; Yunis and Yasmineh 1971).

Nei's argument is based on a simple mathematical computation. Namely, a lethal or nonfunctional mutation occurring in one of the duplicate loci would be harmless and behave as a neutral or near-neutral gene in populations, as long as the other duplicate gene or genes function normally. Such a mutation will be fixed in the population with the probability of \( p = 1 / (2N) \) in diploids, where \( N \) is the size of the population. If such mutations occur at a locus with the frequency of \( u \) per generation, the total number of new mutations occurring in the population is \( 2Nu \) per generation, of which the proportion \( p \) will be fixed in the population. Thus, the rate of fixation of nonfunctional genes per generation is \( u \). Therefore, if there are many duplicate loci in the genome, an appreciable number of loci would become nonfunctional in a relatively short period of evolutionary time.

Strictly speaking, the above argument is not correct, particularly in large populations. In fact, Fisher (1935) showed, using a deterministic mathematical model, that in an infinitely large population, duplicate genes cannot become nonfunctional. This is because lethal mutations may occur at any locus and they are eliminated when all of the duplicate loci become homozygous for lethal genes. In relatively small populations, however, the fre-

* Present address: Center for Demographic and Population Genetics, University of Texas, Houston, Texas 77025.
frequency of lethal genes is subject to genetic random drift, and at one of the duplicate loci lethal genes may be absent or exist in low frequency for a certain period of time, during which lethal genes may be fixed at the other loci. Thus, the probability of fixation of nonfunctional genes at duplicate loci is expected to depend on population size.

Another factor which complicates our problem is the possible back-mutation of lethal genes to their normal alleles. If back-mutations occur, there is a small probability that a locus at which lethal genes have been fixed by random drift may become functional again. Experimental data suggest that the rate of back-mutations from nonfunctional genes is about one hundredth of the rate of forward mutations (Atwood, Schneider, and Ryan 1951). In this circumstance, the frequency of lethal genes at a locus is expected to attain some stable distribution at equilibrium in infinite populations. We are then interested in the probability that a lethal gene is temporarily fixed in the population at one of the duplicate gene loci at equilibrium. If this probability is high, the locus at which a lethal gene is fixed is essentially nonfunctional. We explore this probability below, considering two duplicate gene loci. A similar study was made by Nei (1970) in connection with the accumulation of nonfunctional genes in the Y chromosome.

**Probability of a Lethal Gene Being Fixed at Duplicate Gene Loci**

The probability that a lethal gene is fixed at one of the two duplicate loci can be obtained by studying the stationary distribution of gene frequency (Wright 1931). Let us consider a randomly mating population of size $N$ with diploid segregation. We assume that the effective population size is equal to $N$. Let $a$ and $b$ be the lethal alleles at duplicate loci $A$ and $B$, respectively, and $A$ and $B$ be their respective normal alleles. We designate by $x$ and $y$ the gene frequencies of $a$ and $b$, respectively. Let $s$ be the selection coefficient for genotype $aabb$ (table 1). It is 1 for completely lethal genes but $1 > s > 0.5$ for semilethal genes. Our main interest is in determining the probability of fixation of recessive lethals at a duplicate locus. There is, however, some evidence that a majority of recessive lethals are slightly deleterious in heterozygous condition (Crow and Temin 1964; Nei 1968), though some authors (e.g., Wallance 1966) claim, on the contrary, that they have a slight beneficial effect. Here we assume that the heterozygous effect of a lethal gene depends on the number of lethal genes carried by an individual, as given in table 1.

We further assume that there is linkage equilibrium between the two loci. This assumption appears to be satisfactory for an unlinked or loosely linked pair of loci in the present case (see Wright 1969). The effect of close linkage is discussed later. We designate by $u$ and $v$, respectively, the rates of lethal (forward) mutations and backward mutations per locus per generation. Unless $N$ is extremely small, the joint distribution of $x$ and $y$ at equilibrium is given by

\[ P(x, y) = \frac{1}{2N} \left( 1 - xy - 2sxy + s^2xy \right) \]

for $0 < x, y < 1$. The probability of fixation depends on the initial frequency of the lethal gene and the selection coefficient. For large $N$, the equilibrium frequencies of $x$ and $y$ can be obtained by solving the equations

\[ \frac{dx}{dt} = 2Nsx - 2Nxy \]

and

\[ \frac{dy}{dt} = 2Nsy - 2Nxy \]

with the initial conditions $x(0) = a$ and $y(0) = b$.
TABLE 1

FITNESSES AND FREQUENCIES OF NINE POSSIBLE GENOTYPES FOR TWO DUPLICATE GENES UNDER THE ASSUMPTION OF LINKAGE EQUILIBRIUM

<table>
<thead>
<tr>
<th>Genotype</th>
<th>AA</th>
<th>Aa</th>
<th>aa</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BB:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitness</td>
<td>1</td>
<td>$1 - h_1$</td>
<td>$1 - h_2$</td>
</tr>
<tr>
<td>Frequency</td>
<td>$(1 - x)^2(1 - y)^2$</td>
<td>$2x(1 - x)(1 - y)^2$</td>
<td>$x^2(1 - y)^2$</td>
</tr>
<tr>
<td><strong>Bb:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitness</td>
<td>$1 - h_1$</td>
<td>$1 - h_2$</td>
<td>$1 - h_3$</td>
</tr>
<tr>
<td>Frequency</td>
<td>$2(1 - x)^2y(1 - y)$</td>
<td>$4x(1 - x)y(1 - y)$</td>
<td>$2x^2y(1 - y)$</td>
</tr>
<tr>
<td><strong>bb:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fitness</td>
<td>$1 - h_2$</td>
<td>$1 - h_3$</td>
<td>$1 - s$</td>
</tr>
<tr>
<td>Frequency</td>
<td>$(1 - x)^2y^2$</td>
<td>$2x(1 - x)y^2$</td>
<td>$x^2y^2$</td>
</tr>
</tbody>
</table>

Note.—The A locus has a normal allele A and a nonfunctional gene a. Similarly, the B locus has a normal allele B and a nonfunctional allele b.

$$\phi(x,y) = C e^{-2N\phi(x,y)} (xy)^{U-1} \left[ (1 - x)(1 - y) \right]^{v-1},$$

where $U = 4Nu; V = 4Nv; \phi(x,y) = 2(1 - x)(1 - y)(x + y - 2xy)h_1 + [(x + y - 2xy)^2 + 2x(1 - x)y(1 - y)]h_2 + 2xy(x + y - 2xy)h_3 + x^2y^2s$; and $C$ is a constant such that $\int_0^1 \int_0^1 \phi(x,y) dxdy = 1$ (Wright 1937).

Following Wright (1931) and Kimura (1968), the probability that the wild-type alleles A and B are temporarily fixed at both loci is given by

$$f(0,0) = \int_0^{\frac{1}{2N}} \int_0^{\frac{1}{2N}} \phi(x,y) dxdy$$

$$= \frac{C}{U^2} \left( \frac{1}{2N} \right)^{2U},$$

approximately. Similarly, the probability that allele A is temporarily fixed at locus A and allele B at locus B $[f(0,1)]$ is given by

$$f(0,1) = \int_{\frac{1}{2N}}^1 \int_0^{\frac{1}{2N}} \phi(x,y) dxdy$$

$$= \frac{C}{UV} e^{-2N\phi_2} \left( \frac{1}{2N} \right)^{U+V}$$

approximately. The probability that allele a is temporarily fixed at locus A and allele B at locus B $[f(1,0)]$ is the same as $f(0,1)$ in the present case. Therefore, the probability that lethal genes are temporarily fixed at either of the two duplicate loci can be evaluated by using (3). If $s$ is not 1, there
is the possibility that at both loci semilethal genes \((s \geq 0.5)\) are fixed, but the probability of this event \([f(1,1)]\) is so small even in small populations that it can be neglected.

In practice, there are some difficulties in the evaluation of expression (3). As mentioned previously, \(C\) is given by

\[
C = \left[ \int_0^1 \int_0^1 P(x,y) \, dx \, dy \right]^{-1},
\]

where \(P(x,y) = e^{-2N\sigma(x,y)}(xy)^{U-1}[1 - x]^{(1-y)}\). The double integration in the above expression has to be evaluated numerically, but \(P(x,y)\) may become \(\infty\) at \(x = 0\) and 1 or \(y = 0\) and 1. This difficulty was removed by extending Kimura's (1963) method. Namely, \(C^{-1}\) may be written as

\[
C^{-1} = \int_0^1 \int_0^1 e^{-2N\sigma(x,y)}(xy)^{U-1}[1 - x]^{1-y} \, dx \, dy
\]

\[
= \int_0^1 y^{U-1}(1-y)^{V-1} \int_0^1 e^{-2N\sigma(x,y)}x^{U-1}(1-x)^{V-1} \, dx \, dy
\]

\[
= \int_0^1 [y^{U-1}(1-y)^{V-1}I(y) - y^{U-1}I(0)] \, dy + \frac{1}{U}I(0) + \frac{1}{V}I(1),
\]

where

\[
I(y) = \int_0^1 [e^{-2N\sigma(x,y)}x^{U-1}(1-x)^{V-1} - e^{-2N\sigma(0,y)}x^{U-1}
\]

\[
- e^{-2N\sigma(1,y)}(1-x)^{V-1}] \, dx + \frac{1}{U}e^{-2N\sigma(0,y)} + \frac{1}{V}e^{-2N\sigma(1,y)}.
\]

Therefore, if we define the following functions

\[
Q_1(x,y) = e^{-2N\sigma(x,y)}x^{U-1}(1-x)^{V-1}
\]

\[
- e^{-2N\sigma(0,y)}x^{U-1} - e^{-2N\sigma(1,y)}(1-x)^{V-1},
\]

and

\[
Q_2(y) = y^{U-1}(1-y)^{V-1}I(y) - y^{U-1}I(0) - (1-y)^{V-1}I(1),
\]

which are equal to the integrands in the right-hand sides of the expressions for \(I(y)\) and \(C^{-1}\) in the above, respectively, then \(Q_1(x,y)\) and \(Q_2(y)\) are finite over the entire range of integration. Here we assign \(-e^{-2N\sigma(1,y)}\) and \(-e^{-2N\sigma(0,y)}\), respectively, to \(Q_1(0,y)\) and \(Q_1(1,y)\), and \(-I(1)\) and \(-I(0)\) to \(Q_2(0)\) and \(Q_2(1)\). Thus, it is possible to evaluate \(I(y)\) and \(C^{-1}\). In the computation of \(C\) we used Gauss' methods of numerical integration.

We considered both completely recessive and partially recessive lethals.
In the former, \( h_1 = h_2 = h_3 = 0 \), while, in the latter, \( h_1 = 0.01 \), \( h_2 = 0.02 \), and \( h_3 = 0.03 \) were used. Experimental data in *Drosophila* suggest that the average reduction of fitness of lethal heterozygotes for a single nonduplicate locus is \( 0.01 \sim 0.02 \) approximately. In each of the above two cases complete lethals \((s = 1)\) and semilethals \((s = 0.5)\) were considered. The forward \((u)\) and backward \((v)\) mutation rates were assumed to be \( 10^{-5} \) and \( 10^{-7} \), respectively.

Results obtained are given in table 2. If lethal genes are completely recessive, the probability that a lethal gene is temporarily fixed either at the A or B locus is high when population size is smaller than 2,000, but low when population size is larger than this number. Since \( f(1,1) \) is practically \( 0, 1 - f(0,0) - f(0,1) - f(1,0) \) is the probability that lethal genes are segregating either at locus A or B or both. This probability is small when \( N \) is smaller than 100 but appreciably large when \( N \) is large, as expected. The probability of temporary fixation of lethal genes is barely affected by the value of \( s \) if this is larger than 0.5. On the other hand, if there is any appreciable selection against lethal heterozygotes, the probability is high only when population size is very small. If population size is equal to or larger than 1,000, the probability of temporary fixation of lethal genes at a duplicate locus is negligibly small. When \( N \) is around 1,000, the probability of wild-type alleles being fixed \([f(0,0)]\) is large, while if \( N \) is smaller than 100, \( f(0,1) + f(1,0) \) is larger than \( f(0,0) \). This is because in small populations heterozygous selection against lethal genes is ineffective relative to the effect of genetic random drift, while in large populations it becomes effective. Of course, as the population size increases, the probability that lethal genes are segregating either at locus A or B or both increases. The effect of the value of \( s \) is again very small.

In formula (2) or (3), fixation of an allele means that there is only one type of allele in the population. In large populations, however, it may be more appropriate to define fixation such that the frequency of an allele is

<table>
<thead>
<tr>
<th>( N )</th>
<th>( s = 1 )</th>
<th>( s = 0.5 )</th>
<th>( s = 1 )</th>
<th>( s = 0.5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>.0049</td>
<td>.9833</td>
<td>.0049</td>
<td>.9879</td>
</tr>
<tr>
<td>100</td>
<td>.0048</td>
<td>.9795</td>
<td>.0048</td>
<td>.9788</td>
</tr>
<tr>
<td>1,000</td>
<td>.0029</td>
<td>.7916</td>
<td>.0029</td>
<td>.7861</td>
</tr>
<tr>
<td>2,000</td>
<td>.0015</td>
<td>.5995</td>
<td>.0015</td>
<td>.5913</td>
</tr>
<tr>
<td>4,000</td>
<td>.0004</td>
<td>.3198</td>
<td>.0004</td>
<td>.3128</td>
</tr>
<tr>
<td>10,000</td>
<td>.0000</td>
<td>.0327</td>
<td>.0000</td>
<td>.0320</td>
</tr>
<tr>
<td>20,000</td>
<td>.0000</td>
<td>.0004</td>
<td>.0000</td>
<td>.0004</td>
</tr>
</tbody>
</table>

Note: \( f(0,0) \) stands for the probability that wild-type alleles are fixed at both of the A and B loci.
equal to or larger than $1 - q$, where $q$ is a small positive quantity. Then, $f(0,0)$, $f(0,1)$, and $f(1,0)$ are given by

$$f(0,0) = \int_0^q \int_0^q \phi(x,y) \, dx \, dy,$$

$$f(0,1) = \int_{1-q}^q \int_0^q \phi(x,y) \, dx \, dy,$$

$$f(1,0) = \int_0^q \int_{1-q}^q \phi(x,y) \, dx \, dy.$$

Putting $q = 0.001$, the above probabilities were computed for the case of $h_1 = h_2 = h_3 = 0$ and $s = 1$. Results are given in table 3. Values of $f(0,1) + f(1,0)$, and $f(0,0)$ are only slightly larger than those in table 2. Therefore, even if the above definition of fixation is used, our conclusions remain the same.

**EFFECT OF LINKAGE**

We assumed above that the gamete frequencies for two duplicate loci are in linkage equilibrium. This assumption, of course, is not fulfilled for tandem duplications due to unequal crossing over. In this case, however, the probability of a duplicate gene becoming nonfunctional is expected to be higher than in the case of linkage equilibrium. To see this, consider two linked duplicate gene loci $A$ and $B$, between which no recombination occurs. We again designate normal alleles by $A$ and $B$ and lethal alleles by $a$ and $b$. There are four possible types of chromosomes, that is, $AB, Ab, aB,$ and $ab$. First, assume that the lethal genes are completely recessive, so that selection occurs only against genotype $ab/ab$. In this case there is no selection against chromosome $Ab$ or $aB$. Therefore, if a mutation occurs from $A$ to $a$ or $B$ to $b$ in the chromosome $AB$, the mutant type $Ab$ or $aB$ would behave as a neutral mutation and be fixed in the population with the probability of $1/(2N)$, as long as the other locus remains unmutated. Since the forward mutation rate is much higher than the backward mutation rate, the chromosome type $AB$ will eventually be replaced by $Ab$ or $aB$ even in a large population, if it is finite. Therefore, close linkage is expected to increase the probability of a population being fixed with lethal genes.

**TABLE 3**

| Probability of Temporary Fixation of Lethal Genes at One of Two Duplicate Loci  
$[f(0,1) + f(1,0)]$  
and Probabilities of Temporary Fixation of Wild-Type Alleles at Both Loci  
$(h_1 = h_2 = h_3 = 0, \, s = 1)$ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Population Size</td>
<td>1,000</td>
<td>2,000</td>
<td>4,000</td>
<td>10,000</td>
<td>20,000</td>
</tr>
<tr>
<td>$f(0,0)$</td>
<td>.0031</td>
<td>.0019</td>
<td>.0007</td>
<td>.0000</td>
<td>.0000</td>
</tr>
<tr>
<td>$f(0,1) + f(1,0)$</td>
<td>.8135</td>
<td>.6684</td>
<td>.4429</td>
<td>.1071</td>
<td>.0073</td>
</tr>
</tbody>
</table>
If lethal genes show a deleterious effect in heterozygous condition, the above argument no longer holds. If the fitnesses of genotypes $AB/Ab$ and $Ab/Ab$ or $AB/aB$ and $aB/aB$ are $1 - h$ and $1 - 2h$, respectively, relative to that of $AB/AB$, the probability of fixation and an $Ab$ or $aB$ mutation is given by $p = 2h/(e^{2Nh} - 1)$, approximately, where $h$ is assumed to be small compared with 1 (Kimura 1957). Thus, if $4Nh$ is large, the probability is negligibly small compared with that of the case of $h = 0$, that is, $1/(2N)$. Therefore, if there is selection against lethal heterozygotes, linkage is not expected to increase the probability of temporary fixation of lethal genes at duplicate loci.

**DISCUSSION**

We have seen that the probability of a lethal gene being fixed at duplicate loci depends on the heterozygous effect of the lethal gene as well as on population size. The present study suggests that for one of the two duplicate genes to become nonfunctional the mutant gene must be completely recessive to the wild-type allele unless population size is very small ($\leq 100$). Furthermore, even for a completely recessive lethal, the population size should not be much larger than 2,000. The effective population size in natural populations appears to be often smaller than 2,000 (Nei 1970). Therefore, even if there are only two duplicate genes of the same function, one of them may become nonfunctional. In this context, note that the effective population size for deleterious genes is determined by the local size of the population rather than the total size, in contrast to the case of neutral genes, where the effective size is close to the total size when migration occurs (Kimura and Maruyama 1971).

As mentioned earlier, some authors claim that many lethal genes are slightly beneficial in heterozygous condition. If they show overdominance with respect to fitness, the probability of fixation of lethal genes at duplicate loci is greatly enhanced. In fact, in such a case (e.g., sickle-cell anemia gene) a population fixed with a normal allele at one locus but with an abnormal allele at the other would be more advantageous than a population fixed with normal alleles at both duplicate loci. Nevertheless, there is increasing evidence that overdominant lethal mutations are rather rare in nature (Nei 1968; Murata 1970).

If there are more than two duplicate loci of the same function, the probability of one of the loci becoming nonfunctional is, of course, higher than that in the case of two loci, since the probability of lethal mutations occurring at all duplicate loci rapidly decreases with increasing number of loci. Thus, in this case, even if the effective population size is larger than 2,000, some of the duplicate loci would become nonfunctional. If duplicate loci are completely linked with no recombination, and if lethal mutations are completely recessive, all the duplicate loci except one would eventually become nonfunctional, unless they have innovational mutations and acquire new gene functions which are beneficial to the organism. At any rate, if
there are many duplicate genes of the same function, it is expected that a majority of the genes become nonfunctional, even though they are not necessarily linked. This conclusion supports Nei’s (1969) postulate, which was based on a rough mathematical computation.

The importance of gene duplication in creating opportunities for new gene functions to evolve has long been recognized (Bridges 1936; Ohno 1970). However, as emphasized by Ohta and Kimura (1971), the chance of duplicate genes acquiring a new function must be much smaller than becoming nonfunctional, since mutation is a random event.

If recessive lethals at duplicate loci are completely sheltered from selection, owing to either a small population size or close linkage, the probability that a lethal gene is fixed at a locus by the $t$th generation is given by $P(1,t) = 1 - (4Nu + 1)e^{-ut}$, approximately, where $t$ is assumed to be large and the backward mutation is neglected (Crow and Kimura 1970, p. 395). Thus, if $N = 1,000$, $u = 10^{-5}$, and $t = 10,000$, $P(1,t)$ is 0.059; if $t = 100,000$, and $N$ and $u$ remain the same, $P(1,t)$ is 0.617. Therefore, the fixation of lethal genes at duplicate loci is a slow process, even if they are completely sheltered.

There is supporting evidence that some genes may become nonfunctional after duplication. *Triticum aestivum* (wheat) is a hexaploid and made up of three different genomes, A, B, and D. With an electrophoretic study of nullisomic plants of this species, Brewer, Sing, and Sears (1969) showed that the alkaline phosphatase gene (or genes) is located on the fourth homeologous (phylogenetically related) chromosomes in genomes B and D, but no enzyme band attributable to genome A is present. Since these three genomes are descendants from a common ancestor, this suggests that the gene in genome A has become nonfunctional or deleted after polyploidization. Brewer et al. (1969) studied 11 other enzymes, of which the genes seem to be present in all three genomes. There is also evidence that alcohol dehydrogenase and $\alpha$-amylase genes are possessed by all three genomes (Hart 1970; Nishikawa and Nobuhara 1971). Therefore, in the hexaploid wheat, only a minority of duplicate genes appear to have become nonfunctional. At the present time it is not clear whether this is due to a relatively short evolutionary time of this plant (compared with $1/u$ generations) or strong selection against those individuals which are homozygous for nonfunctional genes at one or two of the duplicate loci. (Wheat is a self-fertilizing plant.)

Other evidence comes from an experiment in *Drosophila melanogaster*. Kidwell (1972) studied the fixation of lethal genes in the Glued-Stubble region (16.8 centimorgans) of the third chromosome which had been kept heterozygous (*Gl-Sb*/+++) with population sizes of $8 \sim 48$. These populations were originally started to study the effectiveness of natural selection for reduced recombination. Tests of lethal genes revealed that at least one lethal gene was fixed on the non-*Gl-Sb* chromosome in five of the 10 populations studied within 60 generations. Lethal genes fixed on the *Gl-Sb* chromosome could not be detected because *Gl* and *Sb* are homozygous lethal. This result indicates that if there are two doses of genes of the same function,
one of them may become nonfunctional as long as the other gene functions normally. In this respect the inactivation of duplicate genes is the same as the inactivation of genes on sheltered chromosomes (see Nei 1970).

Ritossa and Spiegelman (1965) have shown that there are about 100 duplicate genes coding for ribosomal RNA in each nucleolar organizer region of the X and Y chromosomes in *D. melanogaster*. Furthermore, if there are 60 different kinds of genes for transfer RNA, the genome of this organism appears to contain about 13 duplicate loci of each gene (Ritossa, Atwood, and Spiegelman 1966). However, these duplicate genes are not expected to become nonfunctional, since a large amount of ribosomal and transfer RNA are required for protein synthesis. If lethal mutations occur at some of these loci, they are expected to reduce the fitness of heterozygotes, so that they will quickly be eliminated from the population. As pointed out by Crow and Kimura (1970), there may be centripetal selection operating with respect to the number of these genes.

Recently, DNA hybridization experiments have revealed a large amount of repeated DNA in the genome of higher organisms (Britten and Kohne 1968; Walker 1968). They do not appear to code for any protein (Flamm et al. 1969). Their function may be structural in the sense that they support informational genes, as emphasized by Yunis and Yasmineh (1971). It is possible that these DNAs have originated from some nonfunctionalized genes, though the evolution of this class of DNA is largely mysterious at present.

We have used the words nonfunctional genes and lethal genes interchangeably. In practice, nonfunctional genes would not necessarily be lethal genes. If a nonfunctional gene does not have any lethal effect, the probability of fixation of the gene is, of course, higher than that of lethal genes, as discussed by Nei (1970).

**SUMMARY**

The probability that a lethal or nonfunctional gene is fixed at autosomal duplicate gene loci is studied. It is shown that this probability is highly dependent on population size, heterozygous effect of lethal genes, and recombination value between duplicate genes. If effective population size is smaller than 2,000 and lethal genes are completely recessive, the probability that a lethal gene is fixed at one of two duplicate gene loci is high even if they are not linked. If the number of duplicate loci is large and they are linked, the chance that some of them become nonfunctional is even higher, provided lethal genes are completely recessive. This study supports the postulate that there are a large number of nonfunctional genes accumulated in the genomes of higher organisms.

**ACKNOWLEDGMENT**

This research was supported in part by Public Health Service grant GM-17719 and by National Science Foundation grant GB-21224. We thank Margaret Kidwell for valuable comments on the manuscript.
LITERATURE CITED


