Copyright © 1984 by the Genetics Society of America

THE COHESIVE POPULATION GENETICS OF MOLECULAR DRIVE

TOMOKO OHTA* AND GABRIEL A. DOVER[†]

**National Institute of Genetics, Mishima, 411 Japan, and* [†]*Department of Genetics, University of Cambridge, Cambridge CB2 3EH, England*

Manuscript received February 20, 1984
Revised copy accepted May 29, 1984

ABSTRACT

The long-term population genetics of multigene families is influenced by several biased and unbiased mechanisms of nonreciprocal exchanges (gene conversion, unequal exchanges, transposition) between member genes, often distributed on several chromosomes. These mechanisms cause fluctuations in the copy number of variant genes in an individual and lead to a gradual replacement of an original family of n genes (A) in N number of individuals by a variant gene (a). The process for spreading a variant gene through a family and through a population is called molecular drive. Consideration of the known slow rates of nonreciprocal exchanges predicts that the population variance in the copy number of gene a per individual is small at any given generation during molecular drive. Genotypes at a given generation are expected only to range over a small section of all possible genotypes from one extreme (n number of A) to the other (n number of a). A theory is developed for estimating the size of the population variance by using the concept of identity coefficients. In particular, the variance in the course of spreading of a single mutant gene of a multigene family was investigated in detail, and the theory of identity coefficients at the state of steady decay of genetic variability proved to be useful. Monte Carlo simulations and numerical analysis based on realistic rates of exchange in families of known size reveal the correctness of the theoretical prediction and also assess the effect of bias in turnover. The population dynamics of molecular drive in gradually increasing the mean copy number of a variant gene without the generation of a large variance (population cohesion) is of significance regarding potential interactions between natural selection and molecular drive.

MANY important aspects of phenotype of eukaryote species are known to be influenced by the products of genes that are in multiple copies in the genome (for reviews see DOVER 1982; OHTA 1983b; HOOD, CAMPBELL and ELGIN 1975; KEDES 1979; LONG and DAWID 1980; FEDOROFF 1979; JONES and KAFATOS 1982; HUNKAPILLER *et al.* 1982; JEFFREYS 1982; ARNHEIM 1983). The genetics of such multigene families is different from the Mendelian genetics of most single-copy genes and polygenic systems in that the members of a family are not wholly independent units of mutation, segregation and evolution (OHTA 1980; DOVER 1982). The genetics are circumscribed by the activities of three mechanisms of irregular DNA exchanges (gene conversion, unequal exchange and transposition) that cause nonreciprocal transfers of ge-

netic information between member genes, often irrespective of their chromosome distribution. The three mechanisms cause continual fluctuations in the frequencies of variant members of a family during the lifetime of an individual and, hence, can promote the gradual spread of one or another variant gene throughout a family (homogenization) and eventually throughout a population (fixation). The evolutionary process by which the genetic composition of a population is transformed as a consequence of irregular exchanges in its multigene and nongenic families has been called molecular drive (DOVER 1982; DOVER *et al.* 1982). Continuous cycles of family replacement by new variant genes are supported by the widespread observation that in all true families, irrespective of their size, chromosomal distribution and function, there is a much smaller within-species variation than between-species variation [see DOVER (1982), OHTA (1983b) and ARNHEIM (1983) for references]. This general observation is called concerted evolution. Molecular drive is a population genetics process, based on nonreciprocal exchanges between chromosomes, which attempts to explain how concerted evolution is achieved across all relevant chromosome lineages for any one family. The degree of within-species homogeneity and between-species heterogeneity varies widely between families, presumably as a consequence of variation in the rates of mutation, nonreciprocal exchanges, sexual recombination and selection.

To understand the important interactions between the internal forces of molecular drive and the external forces of Darwinian selection, it is necessary to quantify the differences in phenotype among individuals during the spread of a variant member gene in any particular family. The known slow rates of mutation (10^{-6} per gene per generation) and mechanisms of nonreciprocal exchange (10^{-5} – 10^{-2} per generation) lead to a specific prediction that at any given generation during molecular drive there will be a large genetic similarity between individuals with respect to the proportion of a family that has been replaced by a new variant gene (DOVER 1982). For example, it is to be expected that only a few extra copies of a variant gene would arise in the lifetime of an individual. The sexual process would then ensure that the chromosomes on which such copies were produced are distributed at random among individuals of the next generation. Hence, it would take many generations, depending on family and population sizes (OHTA and DOVER 1983), to replace effectively all preexisting copies of a family in all individuals. During this time no large differences in the copy number of the variant gene would exist between individuals at any generation. This pattern of change in the composition of a population with respect to a multigene family, by the internal forces of molecular drive, is analogous to the effect of external selection on a polygenic trait: a gradual change in the population mean while maintaining a small variance over many generations. The phenotypic cohesion of a population is expected to be maintained throughout a period of gradual biological transformation.

In the following sections, a theory is developed to estimate the population variance of the copy number of a mutant gene per individual during the process of spreading into the population. The theory of identity coefficients is

applied with special reference to the state of steady decay. The treatment follows the theory and notation of previous analysis of the population genetics of multigene families located on either one or two pairs of chromosomes (OHTA 1980, 1983b; OHTA and DOVER 1983).

VARIANCE UNDER GENE CONVERSION WITHOUT BIAS

In the following analysis we monitor the effects on variance of only one of the three mechanisms of turnover, namely, gene conversion. Gene conversion is a phenomenon by which two originally dissimilar domains of DNA end up with the sequence of one of them (FOGEL, MORTIMER and LUSNAK 1981). This is considered to take effect through mismatch correction of a heteroduplex formed by the invasion and displacement of a single strand in one helix by a single strand from the other (HOLLIDAY 1964; MESELSON and RADDING 1975; although see SZOSTAK *et al.* 1983 for alternative model). If only one heteroduplex is formed between two helices, then the conversion is asymmetric and a repair of the mismatch in either direction leads to conversion without bias [*i.e.*, prob (Aa gives A gamete) = prob (Aa gives a gamete)]. In this section we use the simplest model of asymmetric gene conversion without bias. In the next section we consider conversion with bias, which is a persistent discrimination during repair in favor of either A or a.

Let us assume initially that there are n genes that are either dispersed or tandemly arranged on each of a pair of homologous chromosomes. We assume that there are $n\lambda$ conversions per chromosome per generation, where λ is the rate of intrachromosome conversion. The model may also be expressed as follows. Assume one interaction event happens with probability μ per individual. Then, the chance that a particular gene is converted is $\lambda = \mu/4n$ (see NAGYLAKI 1984a for details). Under these assumptions, each gene is converted at the rate λ by any one of the remaining $(n - 1)$ genes on the same chromosome (OHTA 1982, 1983a, 1984a).

In addition to intrachromosome conversion we assume regular meiotic recombination between the pair of homologous chromosomes involving members of the family. Let β be the rate between adjacent members per generation so that recombination totals $(n - 1)\ \beta$ per family. In our study, conversion between homologous chromosomes (interchromosomal conversion) is not considered. Its effects are expected to increase gene identity and to exchange genes between chromosomes just as those of unequal interchromosomal crossing over (OHTA 1979). For a theoretical study of the cases in which gene homogenization is exclusively due to interchromosomal gene conversion, see NAGYLAKI (1984b).

In the following section, a method is developed to predict the population variation of the number of a new mutant member of a multigene family in a finite population with effective size N so that there are $2Nn$ genes in total. The method monitors accidental fluctuations both in the frequencies of the new mutant member (due to fluctuations in the directions of unbiased conversion) and in the frequencies of chromosomes in the population (genetic drift).

The method utilizes the theory of identity coefficients (OHTA 1982, 1983a,b): a coefficient defined as the probability of two randomly chosen genes of a family being identical. Strictly speaking, this refers to the probabilities of identity of any two conversion domains. The domain of conversion can be either shorter or longer than the gene itself. When it is shorter it is possible to evolve highly variable genes that are mosaics of different domains. This is observed in immunoglobulin (MIYATA *et al.* 1980; BALTIMORE 1981; OLLO and ROUGEON 1983) and other gene families (DOVER 1982; MUNZ *et al.* 1982). However, any one conversion domain will tend toward homogeneity between gene members. In our analysis we consider the domain to be the length of the gene (or the gene and its spacer in some tandemly arrayed families (COEN, STRACHAN and DOVER 1982; COEN and DOVER 1982) and that conversion takes place between any pair of genes chosen at random from the same chromosome.

The assumption that gene conversion occurs between random pairs of genes may be unrealistic in some families, for it may occur more frequently between closely linked genes than between distant ones. This may also be true when unequal crossing over is the mechanism of homogenization. Under such circumstances, identity coefficients become functions of chromosomal distance between genes (KIMURA and OHTA 1979; OHTA 1981). In the present analyses, the simple conversion model is considered in order to approximate the highly complicated process of spreading a mutant member gene throughout all of the family and throughout the population. The precise validity of the assumption can be left to future elaborations. Previous studies indicate that our assumption is valid for multigene families, or their individual subfamilies, having relatively uniform members such as those of rRNA, tRNA and histones, because the discrepancies between predictions of the conversion and unequal crossing over models disappear (OHTA 1983b), and gene rearrangements on the chromosome are more or less random in some families (FEDOROFF 1979; COEN and DOVER 1983). Thus, our study is mainly directed toward understanding the mutant dynamics of multigene families with uniform members or uniformly homogenized subfamilies within a family.

Assume that a mutation appeared in one copy of the population of $2Nn$ genes that are initially identical. In a previous report we have considered the time it would take for the mutant copy to spread by chance on the chromosome and into the population (OHTA 1983a; OHTA and DOVER 1983). We will consider the variance in the number of mutant copies per chromosome in the course of spreading. Here, the theory of identity coefficients is used with special reference to the state of steady decay. When the conversion rate is high and gene members are relatively uniform, it is expected that, starting from any condition, the state of steady decay is attained fairly rapidly, and once the state is reached, various quantities including the relative variance in copy number remain unaltered in their expectations (for the theory of steady decay in classical population genetics problems, see CROW and KIMURA 1970). In a later section, the adequacy of the theory will be examined through Monte Carlo simulations, by introducing a single mutant into the population and by calculating the variance in copy number in the course of spreading. This theory is

applicable to the cases even when the mutant will eventually disappear from the population, in so far as the mutant has stayed long enough in the population and the state of steady decay has been reached.

Let f be the average probability of identity between genes at the same locus (alleles), C_1 the average identity probability of genes at different loci on the same chromosome and C_2 that of two genes taken from different loci of two homologous chromosomes. To formulate the variance in terms of these identity coefficients, let x_i be the frequency of the mutant in the ith chromosome of the population of $2N$ homologous chromosomes. In other words, if m_i is the number of mutants on the ith chromosome, $x_i = m_i/n$. The average x_i of the population is

$$\bar{x} = \frac{\sum_{i=1}^{2N} x_i}{2N}, \tag{1}$$

and its expectation may be expressed

$$\mu_x = E(x_i), \tag{2}$$

where E denotes expectation. Next, we shall formulate the variance of x_i,

$$s_x^2 = \frac{1}{2N - 1} \left\{ \sum_i x_i^2 - \frac{\left(\sum_i x_i\right)^2}{2N} \right\}. \tag{3}$$

Its expectation is

$$\sigma_x^2 = E(s_x^2). \tag{4}$$

This variance can be defined in terms of identity coefficients from the following relationships between x_i and identity coefficients.

$$\begin{aligned} \frac{1}{n}\{(n-1)C_1 + 1\} &= E\{x_i^2 + (1 - x_i)^2\}, \\ &= E(2 \sum_i x_i^2 - 2 \sum_i x_i + 2N)/(2N) \end{aligned} \tag{5}$$

and

$$\begin{aligned} \frac{1}{n}\{(n-1)C_2 + f\} &= \underset{i \neq j}{E}\{x_i x_j + (1 - x_i)(1 - x_j)\} \\ &= E[2 \sum\sum_{i \neq j} x_i x_j - 2(2N - 1) \sum_i x_i \\ &\quad + 2N(2N - 1)]/[2N(2N - 1)] \end{aligned} \tag{6}$$

By subtracting (6) from (5), σ_x^2 becomes

$$\sigma_x^2 = \frac{1}{2n} \{(n-1)(C_1 - C_2) + (1 - f)\}. \tag{7}$$

A more convenient measure would be the variance of the number of a mutant gene per chromosome (m_i for the ith chromosome).

$$\sigma_m^2 = E(m_i^2) - \mu_m^2 = n^2\sigma_x^2. \tag{8}$$

The size of the variance depends on the mean number of m which is changing during fixation. Hence, we need to concentrate on the ratio of the variance to the mean, σ_m^2/μ_m. Let us call this ratio the relative variance, RV.

By using the theory of identity coefficients we can calculate RV approximately. When a single mutant gene appears, it will be lost in the majority of cases and will spread and eventually replace the previous gene only in a very small minority of cases. The theory of identity coefficients gives the average values of gene identity for all cases. Now, let **H** be the vector of nonidentity coefficients (OHTA and DOVER 1983),

$$\mathbf{H} = (1 - f,\ 1 - C_1,\ 1 - C_2). \tag{9}$$

Without further mutation, **H** decreases each generation, and the state of steady decay is eventually reached. At this state, **H** decreases with a constant rate which is obtained by the maximum eigenvalue of the transition matrix of **H** (see OHTA 1983a). The eigenvector corresponding to the largest eigenvalue gives the relative values of "nonidentity" coefficients at the state of steady decay. These values were obtained numerically and used to calculate the relative variance, $RV = \sigma_m^2/\mu_m$. The variance, σ_m^2, can be estimated from equations (7) and (8). It is difficult, however, to estimate the expected value of the relative variance, in the course of spreading, and so we use the following approximation formula which has been obtained empirically. This has been achieved by taking the ratio of σ_m^2 to a linear combination of the nonidentity coefficients, which may be called either "heterozygosity" (NEI 1975) or "virtual heterozygosity" (KIMURA 1983). If we calculate σ_m^2/μ_m by the following formula, the prediction fits remarkably well to the observed value obtained by Monte Carlo simulation experiments.

$$RV \approx \frac{(n-1)(C_1 - C_2) + (1 - f)}{2\{1 - f/n - (n-1)C_2/n\}}. \tag{10}$$

The numerator comes from equations (7) and (8), and the denominator is twice the heterozygosity. Note that we concentrate on the average value of relative variance in the course of spreading. It varies considerably from time to time and also decreases slightly as the frequency of the mutant increases.

Monte Carlo simulations show that this formula provides good predictions of the relative variance. The method of simulation is as by OHTA (1983a), *i.e.*, a single mutant is introduced into the population of $2N$ chromosomes and its fate is traced until it is either lost from the population or it is fixed in all loci of the family on all chromosomes, *i.e.*, the total number is $2Nn$. The relative variance, σ_m^2/μ_m is calculated at each generation, and the average value in the course of spreading is recorded if the mutant gene becomes fixed. The data are discarded if the mutant gene is lost.

Table 1 shows the comparison of the observed (from Monte Carlo) and the

TABLE 1

Comparison of observed (by Monte Carlo experiments) and theoretical (by equaton 10) relative variances

$(n-1)\beta$	Relative variance (σ_m^2/μ_m)	
	Monte Carlo	Theoretical
0	3.203 ± 1.163	4.210
0.1	1.871 ± 0.266	1.798
0.2	1.352 ± 0.369	1.309
0.3	0.899 ± 0.181	1.083
0.4	0.826 ± 0.157	0.952
0.5	0.886 ± 0.124	0.866

Parameters are $\lambda = 0.1$, $n = 20$, $N = 20$ and average ± standard deviation of five sample paths.

theoretical (from equation 10) relative variances. The observed value is the average with standard deviation of five sample paths, *i.e.*, five cases in which fixation occurs. The agreement between the observed and the expected values is good.

To assess the effect of regular recombination between homologous chromosomes (β), a numerical analysis was performed. One of the best understood examples is the (28S + 18S) ribosomal RNA gene family of Drosophila species (see COEN, STRACHAN and DOVER 1982 and COEN and DOVER 1982 for references), and the parameters of the analysis are chosen to be relevant to this and similar families (see ARNHEIM 1983). Figure 1 represents the results. The parameters are $n = 200$ and $N = 500$ with λ and β as shown in the figure. It is important to note that the products $N\lambda$ and $N\beta$ are realistic and that the results are applicable to a wide range of cases from $N = 5000$ and $\lambda = 0.01/20$ to $N = 50{,}000$ and $\lambda = 0.01/200$. As can be seen from the figure, regular recombination is quite effective in reducing the relative variance. The value of λ may look too high. However, note that we are approximating the joint effects of conversion and unequal crossing over by the simple conversion model. Therefore, these values are appropriate especially for considering the turnover process of multigene families with uniform gene members.

It has to be pointed out, however, that this treatment of β may result in overestimation of the effect of interchromosomal recombination. This is because, in the previous formulation of the transition equations of identity coefficients (OHTA 1982, 1983a,b), the average recombination value between randomly chosen pairs of loci is equated to $(n + 1)\ \beta/3$ (the arithmetic mean), under the assumption of no genetic correlation between gene identity and chromosomal distance. This assumption may overestimate the recombination value. For more exact analyses, the method of KIMURA and OHTA (1979), in which genetic correlation is formulated as a function of chromosomal distance, is required (see also OHTA 1981). It is possible that the harmonic mean of recombination frequency is more appropriate than the arithmetic mean as in the ordinary population genetic models of multilocus systems (LANDE 1976).

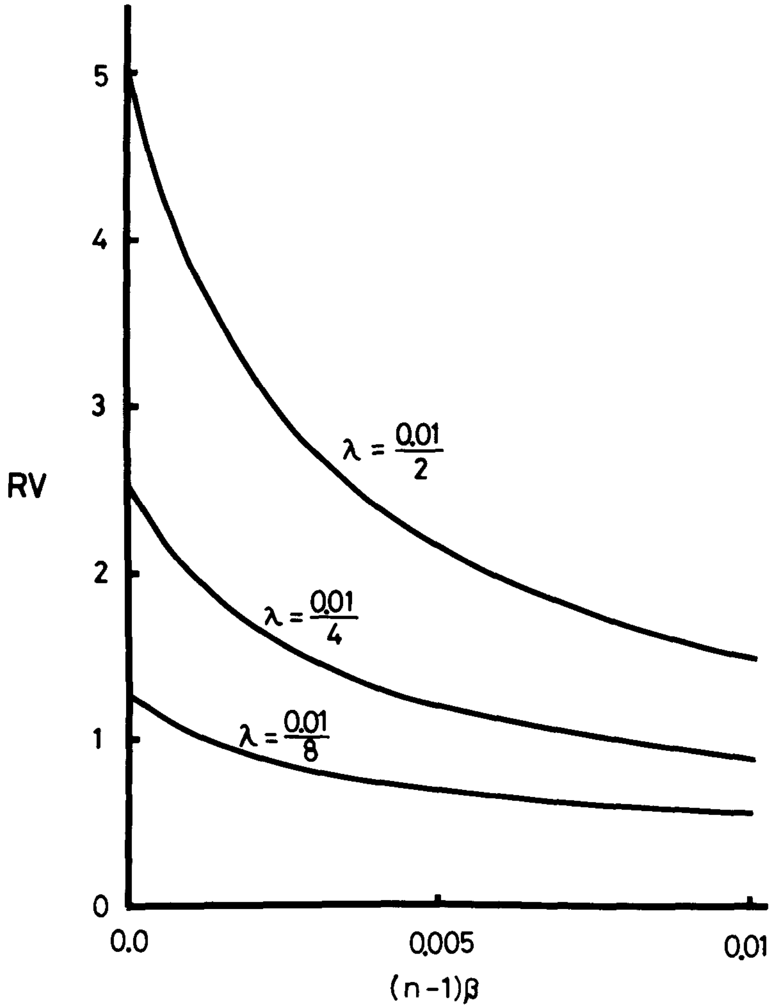

FIGURE 1.—Results of numerical calculations for the relative variance, $RV = \sigma_m^2/\mu_m$, obtained by equation (10), showing the effect of interchromosomal recombination rate $(n - 1)\beta$.

The problem is complex and, in the present paper, the previous simple equations of identity coefficients are used. The more detailed and exact analyses are left for future investigation. In a previous paper (OHTA 1983a), the numerical analysis contained an error, and the value of β needs to be multiplied by a factor, $(n - 1)/(n + 1)$, *i.e.*, the values of β for expected time until fixation in Tables 1 and 2, and in Figures 1 through 3 have to be multiplied by $(n - 1)/(n + 1)$. The amount of error is negligible when n is large but may be substantial when n is small (~5).

The effect of conversion rate on the relative variance is also examined, again with parameter values applicable to the ribosomal DNA family. Figure 2 shows the results. Numbers beside each curve are $(n - 1)\ \beta$. It is interesting to find that the relative variance increases almost linearly with the increase of conversion rate (λ).

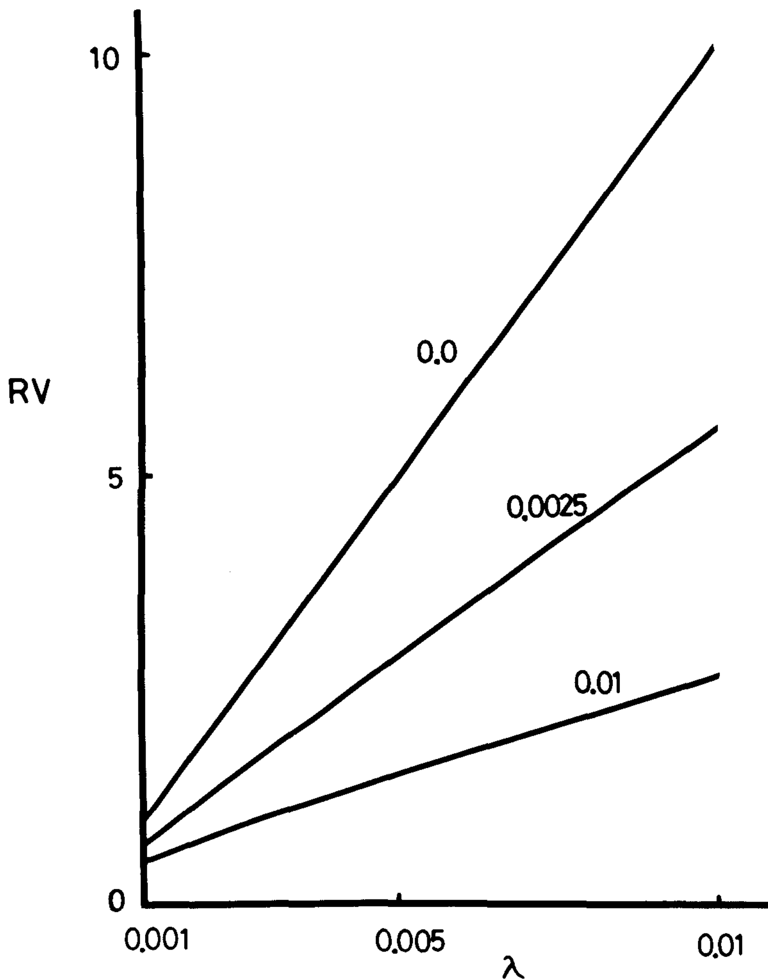

FIGURE 2.—Results of numerical calculations for the RV, showing the effect of conversion rate λ. Numbers beside each curve represent $(n - 1)\beta$. Other parameters are $n = 200$ and $N = 500$ for Figures 1 and 2.

Also examined is the effect of population size on RV. It is expected that RV gets larger as population size increases, because more kinds of chromosomes may survive in larger populations. Figure 3 shows such relationships. When $\beta = 0$, RV increases almost linearly, but not so for $\beta > 0$.

VARIANCE UNDER CONVERSION WITH BIAS

Biased conversion is invoked in molecular drive to explain the relatively fast homogenization of new variants throughout large DNA families in all individuals, during the short time of separation between many species (DOVER 1982). Some theoretical consequences of biased conversion in the single-locus case have been studied (LAMB and HELMI 1982; NAGYLAKI 1983; WALSH 1983). We performed Monte Carlo simulation experiments in order to describe the

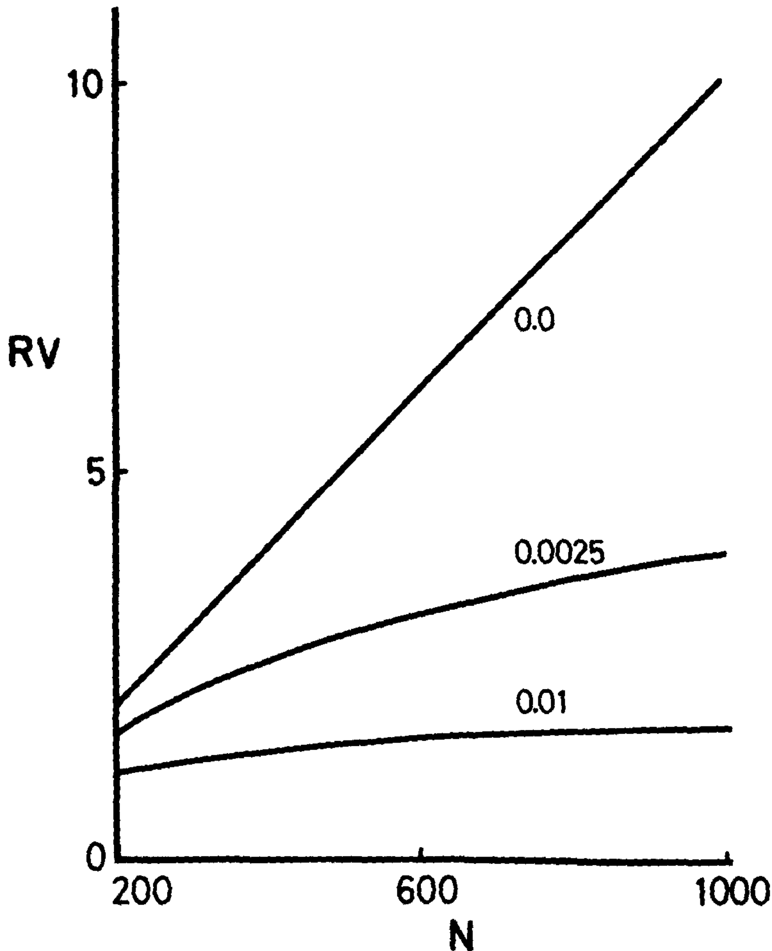

FIGURE 3.—Results of numerical calculations for the *RV*, showing the effect of population size *N*. Numbers beside each curve are $(n - 1)\beta$. Other parameters are $n = 200$ and $\lambda = 0.005$.

effect of biased conversion on the dynamics of mutant genes in multigene families. The method of the experiments was the same as before except that disparity in the direction of conversion is introduced. Gene conversion is carried out for each chromosome with probability $n\lambda$, by randomly choosing two genes. If the two chosen genes are mutant and nonmutant, then the probability that the mutant gene converts the other is set to $0.5 + c$. In other words, c is the proportion of gametes carrying a mutant gene that is in excess of one-half among the products of conversion between the mutant and nonmutant genes. The asymmetry coefficient as defined by NAGYLAKI and PETES (1982) becomes $(0.5 - c)/(0.5 + c)$. The coefficient of conversional advantage of the mutant in one generation may be expressed as follows, by letting s be the conversional advantage.

$$s = 2\lambda\{(0.5 + c) - (0.5 - c)\} = 4\lambda c. \tag{11}$$

TABLE 2

Relative variance and time until fixation for cases of biased gene conversion obtained by Monte Carlo experiments

c	λ: 0.1 n: 100 $(n-1)\beta$: 0	0.1 100 0.5	0.02 100 0	0.02 100 0.5	0.1 20 0	0.1 20 0.5
			Relative variance			
0.0	4.033 ± 1.695	0.772 ± 0.093	0.802 ± 0.176	0.377 ± 0.059	3.203 ± 1.163	0.886 ± 0.124
(Expected)	(4.036)	(0.889)	(0.805)	(0.370)	(4.196)	(0.854)
0.05	2.340 ± 1.173	0.799 ± 0.161	0.645 ± 0.166	0.348 ± 0.039	2.878 ± 1.630	0.793 ± 0.188
0.1	4.580 ± 3.599	0.904 ± 0.191	0.664 ± 0.250	0.376 ± 0.036	2.492 ± 1.298	0.762 ± 0.159
0.15	8.069 ± 0.513	0.957 ± 0.211	0.729 ± 0.273	0.337 ± 0.086	2.804 ± 1.582	0.808 ± 0.189
0.2	5.749 ± 4.536	0.954 ± 0.218	1.126 ± 0.517	0.390 ± 0.048	3.967 ± 2.691	0.827 ± 0.204
0.25	9.334 ± 10.285	1.113 ± 0.150	0.761 ± 0.281	0.392 ± 0.080	3.329 ± 2.730	0.991 ± 0.122
			Time until fixation			
0.0	952 ± 533	33,061 ± 28,574	3,992 ± 1,270	13,696 ± 5,443	235 ± 134	1,130 ± 489
(Expected)	(990)	(4,143)	(4,950)	(10,459)	(190)	(857)
0.05	456 ± 142	525 ± 110	1,831 ± 423	2,208 ± 447	170 ± 66	410 ± 78
0.1	245 ± 60	285 ± 31	970 ± 144	1,158 ± 197	169 ± 64	242 ± 49
0.15	171 ± 23	222 ± 17	862 ± 164	925 ± 161	128 ± 32	146 ± 25
0.2	140 ± 22	165 ± 16	620 ± 62	659 ± 59	93 ± 16	129 ± 13
0.25	117 ± 11	128 ± 17	506 ± 90	563 ± 75	87 ± 15	118 ± 11

Other parameters; $N = 20$ and average ± standard deviation of five sample paths for $c = 0$ and ten paths for $c > 0$. When $c = 0$, conversion is unbiased, and the theoretical values (by equation 10, and by equation 6 of OHTA 1983a) are given in parentheses.

With and without bias in conversion, the relative variance ($RV = \sigma_m^2/\mu_m$) and the time until fixation of a mutant were examined. Table 2 gives the results. Three cases examined are $\lambda = 0.1$ and $n = 100$, $\lambda = 0.02$ and $n = 100$ and, $\lambda = 0.1$ and $n = 20$ with and without interchromosomal recombination ($\beta = 0$ or $(n - 1)\beta = 0.5$). The conversional bias is changed from $c = 0$ to $c = 0.25$. When $c = 0$, the theoretical relative variance is obtained by equation (10) and is given in parentheses. From the table, it can be seen that the conversional advantage has a relatively small effect on the relative variance, although in some cases it appears that the variance is slightly increased by biased conversion.

On the other hand, the bias has a large effect on the time until fixation and on the probability of spreading of a mutant as suggested by NAGYLAKI and PETES (1982), NAGYLAKI (1983) and WALSH (1983) and as the present experiment clearly shows. Table 2 also gives the results for the time until fixation of the same Monte Carlo experiments. The theoretical values for the case of $c = 0$ is obtained by using the theory of OHTA (1983a). As can be seen from the table, biased conversion is quite effective for reducing the time. Let us examine the process of fixing a mutant gene in some detail. In terms of conversional advantage, s, the time until fixation within a *single* chromosome may be obtained. For our purpose, the fixation time is obtained deterministically, by regarding the gene family on a chromosome as a population, although, for an exact value, the method of NAGYLAKI and PETES (1982) is needed. Time until fixation (t_l) is, by noting that the rate of change of gene frequency is $dx/dt = sx(1 - x)$ (LAMB and HELMI 1982),

$$t_l \approx \int_{x=\frac{1}{n}}^{x=1-\frac{1}{n}} dt = \int_{\frac{1}{n}}^{1-\frac{1}{n}} \frac{1}{sx(1-x)}\, dx = \frac{2}{s}\, ln(n). \qquad (12)$$

Note that t_l is the time concerning a single chromosome. Hence, in consideration of the population dynamics, it is appropriate in those cases when intragenomic fixation time [equation (12)] is larger than that at the population level ($4N$ generations for a neutral mutant, see KIMURA and OHTA 1969), and when the rate of interchromosomal recombination is negligibly low (see OHTA 1983a). For example, when $\beta = 0$, $c = 0.25$, $\lambda = 0.02$ and $n = 100$, equation (12) gives $t_l = 460$, and the observed value in Table 2 is 506. Despite a limited applicability of this formula, it may be useful for considering large gene families that are clustered on a single pair of chromosomes. From Table 2, it can be seen also that the effect of interchromosomal recombination for increasing the time until fixation becomes less as the bias gets larger.

VARIANCE FOR GENE FAMILIES SUBDIVIDED INTO TWO PAIRS OF CHROMOSOMES

In this section, "cohesiveness" is examined for the cases in which a gene family is subdivided into two pairs of chromosomes *I* and *II* (Figure 4). Unbiased conversion is analyzed. Let us assume that n genes are tandemly arranged on a single chromosome so that a haploid set contains $2n$ genes be-

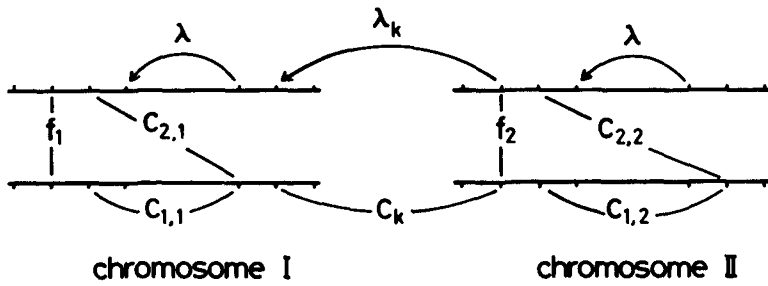

FIGURE 4.—Diagram showing the conversion model and the meaning of seven identity coefficients for a family dispersed into two chromosomes *I* and *II*.

longing to the family and $4Nn$ genes in the total population. Let λ_k be the rate per generation that a gene is converted by a gene on nonhomologous chromosomes (Figure 4). The rate of regular recombination is β as before.

We suppose that a mutation appears initially in a gene on one chromosome *I* in the population, and we consider the population dynamics of spread of the mutant gene. We define the following set of identity coefficients: f_1, $C_{1,1}$ and $C_{2,1}$ are the coefficients for chromosome *I*; f_2, $C_{1,2}$ and $C_{2,2}$ are those for chromosome *II* and C_k is that between genes on chromosomes *I* and *II* (Figure 4). The theory of transforming these identity coefficients from one generation to the next has been developed and some transient properties have been examined (OHTA and DOVER 1983). For convenience, let us define the non-identity coefficients as follows and call them heterozygosity as before.

$$\left.\begin{aligned} g_1 &= 1 - f_1 \\ H_{1,1} &= 1 - C_{1,1} \\ H_{2,1} &= 1 - C_{2,1} \\ H_k &= 1 - C_k \\ g_2 &= 1 - f_2 \\ H_{1,2} &= 1 - C_{1,2} \\ \text{and} \quad H_{2,2} &= 1 - C_{2,2}. \end{aligned}\right\} \quad (13)$$

Previously, we have studied the rapidity with which the ratio $H_{2,2}/H_{2,1}$, approaches unity under various values of λ_k relative to λ. At the state of steady decay, the heterozygosity coefficients for chromosomes *I* and *II* become the same, *i.e.*, $g_1 = g_2$, $H_{1,1} = H_{1,2}$ and $H_{2,1} = H_{2,2}$, and the transient property of this ratio reveals how rapidly the heterozygosity approaches the state of steady decay.

Here, we investigate the properties of heterozygosity coefficients at the state of steady decay and relate them to the mean square of the difference of the number of mutant genes between chromosomes *I* and *II*. Let x_i be the fre-

quency of the mutant of the ith chromosome I and y_j be that of the jth chromosome II in the population. We have, by noting that C_k is the identity coefficient between genes on chromosomes I and II,

$$\begin{aligned} C_k &= E\{x_i y_j + (1 - x_i)(1 - y_j)\} \\ &= E(2 \sum_{ij}\sum x_i y_j - 2N \sum_i x_i - 2N \sum_j y_j + 4N^2)/(4N^2). \end{aligned} \tag{14}$$

On the other hand, the mean square of the difference of average frequency of mutant between chromosomes I and II may be expressed as follows. We denote the expectation of the mean square, σ_{12}^2, and from equations (6) and (14),

$$\begin{aligned} \sigma_{12}^2 &= (\mu_x - \mu_y)^2 = E\{(x_i - y_j)^2\} \\ &= \frac{1}{2n}\{(n - 1)(C_{2,1} + C_{2,2}) + f_1 + f_2\} - C_k. \end{aligned} \tag{15}$$

At the state of steady decay, by letting $C_2 = C_{2,1} = C_{2,2}$, and $f = f_1 = f_2$,

$$\sigma_{12}^2 = \frac{1}{n}\{(n - 1)C_2 + f\} - C_k. \tag{16}$$

In terms of heterozygosity coefficients, by letting $g = g_1 = g_2$,

$$\sigma_{12}^2 = H_k - \frac{1}{n}\{(n - 1)H_2 + g\}. \tag{17}$$

σ_{12}^2 is divided by twice the heterozygosity to estimate relative value. It is further multiplied by n to have the value for the number of mutants per chromosome, *i.e.*, the relative mean square difference of average number of mutants between chromosomes I and II.

$$RV_{12} = \frac{nH_k - (n - 1)H_2 - g}{2\{1 - f/n - (n - 1)C_2/n\}}. \tag{18}$$

This formula may not be entirely appropriate as the expectation of the ratio when λ_k is very small. This is because the mutant may be completely lost from the chromosome I (or II) population while persisting in the chromosome II (or I) population. Nevertheless, it is a good measure for the purpose of comparing it with the RV among either of the pairs of homologous chromosomes (I or II).

Another measure of interest is the ratio, H_k/H_2, *i.e.*, heterozygosity measured between chromosomes I and II relative to that measured for genes on homologous chromosomes. The ratio represents relative amounts of heterozygosities at the state of steady decay. Relative values of heterozygosity coefficients at the state of steady decay were obtained by the eigenvector of the transition matrix corresponding to the maximum eigenvalue as in the previous section. Table 3 gives some examples of these quantities numerically obtained for various values of λ/λ_k. The relative variance of the number of mutants among the homologous chromosomes [equation (10)] is also shown for comparison.

TABLE 3

Results of numerical calculation for the RV, *the ratio of heterozygosity measured between chromosomes* I *and* II *relative to that measured between homologous chromosomes* (H_k/H_2) *and the relative mean square of the difference of the number of mutants per chromosome between chromosomes* I *and* II (RV_{12})

	$n = 50$			$n = 100$			$n = 200$		
λ/λ_k	RV	$\frac{H_k}{H_2}$	RV_{12}	RV	$\frac{H_k}{H_2}$	RV_{12}	RV	$\frac{H_k}{H_2}$	RV_{12}
4	1.417	1.013	0.378	1.282	1.006	0.344	1.092	1.003	0.296
8	1.747	1.034	0.912	1.511	1.015	0.786	1.216	1.006	0.634
16	2.022	1.083	2.128	1.685	1.034	1.739	1.299	1.013	1.334
32	2.210	1.191	4.832	1.796	1.074	3.747	1.349	1.027	2.760
64	2.320	1.436	10.970	1.860	1.159	8.017	1.376	1.056	5.682
128	2.375	2.013	25.418	1.894	1.347	17.420	1.390	1.117	11.766
256	2.397	3.359	59.148	1.910	1.782	39.187	1.397	1.248	24.855

Other parameters: $\lambda = 0.02$, $\beta = 0.0002$ and $N = 100$.

From the table, it can be seen that, as the conversion rate between genes on nonhomologous chromosomes, λ_k, becomes small, H_k/H_2 and RV_{12} become large. RV_{12} increases remarkably, indicating large differences in the numbers of the mutant gene between chromosomes *I* and *II*. On the contrary, *RV* is relatively insensitive to λ_k. In other words, the population variation is not much influenced by λ_k. This prediction of population cohesion holds also for the sum of the numbers of the mutant gene on chromosomes *I* and *II* ($n(x_i + y_i)$ for the *i*th gamete). This is because the two pairs of chromosomes segregate independently at meiosis, and the *RV* in the number of mutant genes is expected to be the same whether it is derived from the homologous chromosomes (either nx_i or ny_i) or from among the sum of two chromosomes ($n(x_i + y_i)$) in the population.

INTERACTION WITH NATURAL SELECTION

If natural selection favors a mutant gene of a multigene family, the variance of the number of mutants per genome is a critical factor for determining its effectiveness just as in ordinary quantitative genetics (MATHER and JINKS 1971). Let us consider a simple model of natural selection. Assume that selection acts in such a way that the fitness of the chromosomes with m_i mutants is

$$W_i = 1 + tm_i \tag{19}$$

where t is a selection coefficient that may be positive or negative. Let p_{m_i} be the frequency of the chromosome in the population with m_i mutants. Then, the change of p_{m_i} by selection in one generation is

$$\begin{aligned} \Delta p_{m_i} = p'_{m_i} - p_{m_i} &= \frac{1 + tm_i}{1 + t\bar{m}} p_{m_i} - p_{m_i} \\ &= \frac{t(m_i - \bar{m})p_{m_i}}{\bar{w}} \end{aligned} \tag{20}$$

where p'_{m_i} is the frequency after selection, and $\bar{w}$ is the mean fitness of the population. Thus, the change of μ_m by selection becomes,

$$\Delta\mu_m = E\{m_i(\Delta p_{m_i})\} = E\left\{\frac{tm_i(m_i - \bar{m})p_{m_i}}{\bar{w}}\right\} \\ \simeq \frac{t\sigma_m^2}{\bar{w}}. \tag{21}$$

This formula applies to the diploid selection model, if selection is additive, such that the fitness of a zygote with m_i copies of mutant on one haploid set and m_j copies on the other set is given by $W_{ij} = 1 + t(m_i + m_j)$. Equation (21) implies that the increase of the average number of mutants is $t\sigma_m^2/\bar{w}$ in one generation. The result is analogous to the so-called fundamental theorem of natural selection by FISHER (1930) and the secondary fundamental theorem by A. ROBERTSON (see FALCONER 1981). In this regard, it is interesting to note that the theory of quantitative genetics based on genotypic and phenotypic variances (*e.g.*, see BULMER 1980) remains unaltered, even if the underlying genetic mechanisms for spreading a mutant allele differ between the present multigene family model and the classical one (see next section).

Combining the prediction on the variance with our previous results, it can be said that the selection response becomes larger with either higher rates of conversion or with lower rates of regular recombination, if other parameters remain the same. This is because both rates affect the population variance for the number of mutant genes per individual. This model could apply to the elimination of deleterious mutations from the multigene family ($t < 0$). Some multigene families contain pseudogenes that are the results of accumulation of defective mutations. If the proportion of pseudogenes becomes too large, they could have deleterious effects on organisms, and the present model is a reasonable approximation of such effects for multigene families with uniform members. When selection is directional, the variance would not be much influenced by selection, and the present model would apply. For a complete quantitative assessment of the effectiveness of natural selection, however, further investigation is needed.

Our model treats the simple case in which fitness is a linear function of the number of the mutant gene in a family [equation (19)]. It is comparable to the model of natural selection which considers fitness as a function of identity coefficients and in which the identity probability may increase or decrease by selection (OHTA 1980). For gene families with diverse functions, such as those for the immunoglobulins and histocompatibility antigens, natural selection could be operating to lower identity coefficients, whereas for families coding for uniform products such as histones or rRNAs, identity coefficients may be strengthened by natural selection.

It is interesting to compare our results with those for transposable elements. CHARLESWORTH and CHARLESWORTH (1983) have shown that, with respect to the copy number per individual, the Poisson distribution is a sufficiently good approximation for transposons, with variance equal to the mean ($RV = 1$ in

our notation). This is because linkage is loose for any transposon family in diploid species, and linkage disequilibrium becomes negligible. The situation is comparable to the present case with sufficiently large recombination rate. It should be noted, however, that, from the numerical examples of Tables 1 and 2, and Figures 1 through 3, the variance may become less than the mean ($RV < 1$) when λ is relatively small and β is sufficiently large.

EVOLUTIONARY PERSPECTIVES

Our analysis reveals that the genotypic variance in the number of mutant genes that have spread in a multigene family consisting of relatively uniform members in each individual is expected to be small under realistic conditions of the rates of nonreciprocal exchanges between member genes. This means that the population mean of the ratio of wild-type to mutant genes per individual can be moved from an initial condition of n number of A genes to a final condition of n number of a genes per individual without the generation of a large variance at any given generation. The situation is analogous to a long-term shift in the population mean of a polygenic trait under natural selection. The analysis shows that the RV is the same, or less than, that expected of a Poisson distribution, under realistic rates of intrachromosome conversion and regular recombination. Under conditions in which there is no homologous recombination (or interchromosome conversion) coupled to high rates of intrachromosome conversion, the RV is higher than Poisson because the conditions are optimal for the maximum genetic differentiation between homologous chromosomes. For the case of gene families that are spread on two or more pairs of chromosomes, see the work by OHTA and DOVER (1983). Our models deal only with gene conversion; however, broadly similar dynamics of population change are expected from other mechanisms of nonreciprocal exchanges in multigene families such as unequal exchange and transposition. All three mechanisms operate within and between chromosomes (see DOVER 1982 for references) at rates that are considerably slower than the fast rate at which chromosomes are randomly distributed between individuals at each generation by the sexual process. Hence, there will be small differences between individuals with respect to the total number of a new mutant gene whether we consider a process of genetic replacement in a preexisting family or a gradual amplification of a family, *de novo*, such as the P element family involved with hybrid dysgenesis in *D. melanogaster* (BREGLIANO and KIDWELL 1983; ENGELS 1983).

The low population variance in the copy number of a mutant gene at any given generation during molecular drive is of consequence to natural selection. The situation is analogous to the ways in which natural selection interacts with continuous traits under polygenic control (MATHER and JINKS 1971). There is, however, an important distinction between the two systems. In the absence of selection, each individual locus of a polygenic system evolves independently and either reaches genetic equilibrium for a number of alternative alleles or becomes fixed for one or another allele, by genetic drift. The loci of a true

multigene family are not evolving independently, in that a unique variant allele can become fixed eventually in all loci distributed around the karyotype. Hence, in contrast to polygenic systems, a multigene family can evolve as a whole and change in composition from one extreme of n number of A genes per individual to the other extreme of n number of a genes per individual, at all loci without the aid of selection. The significance of a small phenotypic variance during this gradual process of genetic transformation of a population is that selection might tolerate the small concurrent changes that have taken place in the biology of the individuals at any given moment (DOVER 1982). Tolerance would be reinforced especially were the population to seek out more appropriate ecological niches during the period of change. Hence, small differences between individuals of any one population at any given generation, with respect to the degree of change in a family, may be of little consequence to fitness differentials, even though at the end of the process, when a mutant gene has completely spread through a population, selection might recognize a large fitness difference between such transformed individuals and those of another population. Hence, molecular drive, as an independent process for changing the genetic composition of a population, could contribute to the biological differentiation of populations.

One interesting example of a possible interaction between natural selection and molecular drive has recently come to light. This concerns the need for the products of single-copy genes, which bind to specific regions in a multigene family, to keep up with the continually changing composition of the sequences in such regions. For example, the known species specificities of RNA polymerase I complexes could have arisen as a consequence of selection for different and more efficient polymerases, or cofactors, needed to recognize the changing nature of the rDNA family promoters and other relevant sequences, as two populations differentiate in this respect. A gradual replacement of multiple rDNA promoters with a new sequence, in the cohesive manner effected by molecular drive, should allow sufficient time for selection to increase the frequency of new specific allele of *PolI* or its cofactors. Experimental evidence in favor of this direction of coevolution between promoters and polymerases, rather than the converse, is becoming available in the rDNA of diverse genera (COEN and DOVER 1982, 1983; ARNHEIM 1983; R. MIESFELD and N. ARNHEIM, personal communication; ONISHI, BERGLAND and REEDER 1984). This model of an interaction between selection and molecular drive might be generally applicable to other types of specific molecular interactions between single-copy gene products (RNA or protein) and multigene families, for example in the immunoglobulin family. This would imply a more widespread effect of DNA flux by nonreciprocal exchange mechanisms than simply the gene families in which it occurs.

The size of the variance in copy number of a variant gene as it spreads, and the extent to which selection tolerates such differences in a population, is expected to vary widely between families and species (DOVER 1982; Dover *et al.* 1982; FLAVELL 1982). Our models describe a small relative variance in families from 50 to 200 members under realistic rates of gene conversion and recombination. Sequence analysis of variation within two large noncoding fam-

ilies (>10,000 members) in Drosophila species reveals all of the expected stages of transition during replacement at individual nucleotide positions and confirms the continual homogenizing effects of the turnover mechanisms in these families in each of the species (T. STRACHAN, D. A. WEBB and G. A. DOVER, unpublished data). A quantification of these data and the extent to which they are predicted by our models is under investigation. The genetics of very small families (<10 members), distributed on one chromosome, might approach, however, the Mendelian genetics associated with single-copy genes, especially in cases such as the globin gene "family" in which the majority of genes (excluding the pairs of α's and γ's) no longer experience significant genetic exchanges among themselves. Such families are more akin to clusters (BODMER 1981) in which each gene seemingly evolves as an independent unit. In a similar vein, the genetics of very large families such as silk moth chorion genes (JONES and KAFATOS 1982), sea urchin histone genes (KEDES 1979; HENTSCHEL and BIRNSTIEL 1981), amphibian 5S rRNA genes (FORD and BROWN 1976) and mammalian immunoglobulin genes (HOOD, CAMPBELL and ELGIN 1975; HUNKAPILLER *et al.* 1983), which are divided into distinct subfamilies that affect the phenotype in different ways, is expected to be more elaborate (see OHTA 1984b for an improved model).

The evolutionary genetics of molecular drive, like that of Mendelian populations under selection and drift, can be modeled only in general terms. The specifics of the system and their ensuing interactions with selection and drift need to be examined case by case.

We are grateful for the advice and help of our colleagues MOTOO KIMURA, RICHARD B. FLAVELL and ENRICO COEN. Thanks are also due BRUCE S. WEIR, NICK BARTON and anonymous referees for their constructive criticisms. This work is supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan to T. O., and from the Science and Engineering Council (Great Britain) to G. A. D. Contribution no. 1569 from the National Institute of Genetics, Mishima, Japan.

LITERATURE CITED

ARNHEIM, N., 1983 Concerted evolution of multigene families. pp. 38–61. In: *Evolution of Genes and Proteins*, Edited by M. NEI and R. K. KOEHN. Sinauer Associates, Inc., Sunderland, Massachusetts.

BALTIMORE, D., 1981 Gene conversion: some implications for immunoglobulin genes. Cell **24:** 592–594.

BODMER, W. F., 1981 The William Allan Memorial Award Address: gene clusters, genome organization, and complex phenotypes: when the sequence is known, what will it mean? Am. J. Hum. Genet. **33:** 664–682.

BREGLIANO, J. -C. AND M. G. KIDWELL, 1983 Hybrid dysgenesis determinants. pp. 364–404. In: *Mobile Genetic Elements*, Edited by J. A. SHAPIRO. Academic Press, New York.

BULMER, M. G., 1980 *The Mathematical Theory of Quantitative Genetics.* Oxford University Press, Oxford.

CHARLESWORTH, B. and D. CHARLESWORTH, 1983 The population dynamics of transposable elements. Genet. Res. **42:** 1–27.

COEN, E. S. and G. A. DOVER, 1982 Multiple Pol I initiation sites in the rDNA spacers of *Drosophila melanogaster.* Nucl. Acid. Res. **10:** 7017–7026.

COEN, E. S. and G. A. DOVER, 1983 Unequal exchanges and the coevolution of X and Y rDNA arrays in *Drosophila melanogaster.* Cell **33:** 849–855.

COEN, E. S., S. STRACHAN and G. A. DOVER, 1982 Dynamics of concerted evolution of ribosomal DNA and histone gene families in the melanogaster species subgroup of *Drosophila*. J. Mol. Biol. **158:** 17–35.

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

DOVER, G., S. BROWN, E. COEN, J. DALLAS, T. STRACHAN and M. TRICK, 1982 The dynamics of genome evolution and species differentiation. pp. 343–372. In: *Genome Evolution*, Edited by G. A. DOVER and R. B. FLAVELL. Academic Press, New York.

DOVER, G. A., 1982 Molecular drive: a cohesive mode of species evolution. Nature **299:** 111–117.

ENGELS, W. R., 1983 The P family of transposable elements in *Drosophila*. Annu. Rev. Genet. **17:** 315–344.

FALCONER, D. S., 1981 *Introduction to Quantitative Genetics*, Ed. 2. Longman, New York.

FEDOROFF, N. V., 1979 On spacers. Cell **16:** 697–710.

FISHER, R. A., 1930 *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

FLAVELL, R. B., 1982 Sequence amplification, deletion and rearrangement: major sources of variation during genome evolution. pp. 301–324. In: *Genome Evolution*, Edited by G. A. DOVER and R. B. FLAVELL. Academic Press, New York.

FOGEL, S., R. K. MORTIMER and K. LUSNAK, 1981 Mechanisms of meiotic gene conversion or "wanderings on a foreign strand." pp. 289–339. In: *Molecular Biology of the Yeast Saccharomyces*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

FORD, P. J. and R. D. BROWN, 1976 Sequences of 5S ribosomal RNA from *Xenopus mulleri* and the evolution of 5S gene-coding sequences. Cell **8:** 485–493.

HENTSCHEL, C. C. and M. L. BIRNSTIEL, 1981 The organization and expression of histone gene families. Cell **25:** 301–313.

HOLLIDAY, R., 1964 A mechanism for gene conversion in fungi. Genet. Res. **5:** 282–304.

HOOD, L., J. H. CAMPBELL and S. C. R. ELGIN, 1975 The organization, expression, and evolution of antibody genes and other multigene families. Annu. Rev. Genet. **9:** 305–353.

HUNKAPILLER, T., H. HUANG, L. HOOD and J. H. CAMPBELL, 1982 The impact of modern genetics on evolutionary theory. pp. 164–189. In: *Perspectives on Evolution*, Edited by R. MILKMAN. Sinauer Associates, Inc., Sunderland, Massachusetts.

JEFFREYS, A., 1982 Evolution of globin genes. pp. 157–176. In: *Genome Evolution*, Edited by G. A. DOVER and R. B. FLAVELL. Academic Press, New York.

JONES, W. C. and F. C. KAFATOS, 1982 Accepted mutations in a gene family: evolutionary diversification of duplicated DNA. J. Mol. Evol. **19:** 87–103.

KEDES, L. H., 1979 Histone genes and histone messengers. Annu. Rev. Biochem. **48:** 837–870.

KIMURA, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, New York.

KIMURA, M. and T. OHTA, 1969 The average number of generations until fixation of a mutant gene in a finite population. Genetics **61:** 763–771.

KIMURA, M. and T. OHTA, 1979 Population genetics of multigene family with special reference to decrease of genetic correlation with distance between gene members on a chromosome. Proc. Natl. Acad. Sci. USA **76:** 4001–4005.

LAMB, B. C. and S. HELMI, 1982 The extent to which gene conversion can change allele frequencies in populations. Genet. Res. **39:** 199–217.

LANDE, R., 1976 The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. **26:** 221–235.

LONG, E. H. and I. B. DAWID, 1980 Repeated genes in eukaryotes. Annu. Rev. Biochem. **49:** 727–764.

MATHER, K. and J. L. JINKS, 1971 *Biometrical Genetics.* Chapman and Hall, London.

MESELSON, M. S. and C. M. RADDING, 1975 A general model for genetic recombination. Proc. Natl. Acad. Sci. USA **72:** 358–361.

MIYATA, T., T. YASUNAGA, Y. YAMAWAKI-KATAOKA, M. OBATA and T. HONJO, 1980 Nucleotide sequence divergence of mouse immunoglobulin γ_1 and γ_{2b} chain genes and the hypothesis of intervening sequence-mediated domain transfer. Proc. Natl. Acad. Sci. USA **77:** 2143–2147.

MUNZ, P., H. AMSTUTZ, J. KOHLI and U. LEUPOLD, 1982 Recombination between dispersed serine tRNA genes in *Schizosaccharomyces pombe.* Nature **300:** 225–231.

NAGYLAKI, T., 1983 Evolution of a large population under gene conversion. Proc. Natl. Acad. Sci. USA **80:** 5941–5945.

NAGYLAKI, T., 1984a The evolution of multigene families under intrachromosomal gene conversion. Genetics **106:** 529–548.

NAGYLAKI, T., 1984b Evolution of multigene families under interchromosomal gene conversion. Proc. Natl. Acad. Sci. USA **81:** 3796–3800.

NAGYLAKI. T. and T. D. PETES, 1982 Intrachromosomal gene conversion and the maintenance of sequence homogeneity among repeated genes. Genetics **100:** 315–337.

NEI, M., 1975 *Molecular Population Genetics and Evolution.* North-Holland and Elsevier, Amsterdam.

OHTA, T., 1979 An extension of a model for the evolution of multigene families by unequal crossing over. Genetics **91:** 591–607.

OHTA, T., 1980 *Evolution and Variation of Multigene Families: Lecture Notes in Biomathematics,* Vol. 37. Springer, New York.

OHTA, T., 1981 Further study on the genetic correlation between gene members of a multigene family. Genetics **99:** 555–571.

OHTA, T., 1982 Allelic and non-allelic homology of a supergene family. Proc. Natl. Acad. Sci. USA **79:** 3251–3254.

OHTA, T., 1983a Time until fixation of a mutant belonging to a multigene family. Genet. Res. **41:** 47–55.

OHTA, T., 1983b On the evolution of multigene families. Theor. Pop. Biol. **23:** 216–240.

OHTA, T., 1984a Some models of gene conversion for treating the evolution of multigene families. Genetics **106:** 517–528.

OHTA, T., 1984b Population genetics theory of concerted evolution and its application to the immunoglobulin V gene tree. J. Mol. Evol. In press.

OHTA, T. and G. A. DOVER, 1983 Population genetics of multigene families that are dispersed into two or more chromosomes. Proc. Natl. Acad. Sci. USA **80:** 4079–4083.

OLLO, R. and F. ROUGEON, 1983 Gene conversion and polymorphism: generation of mouse immunoblogulin γ2a chain alleles by differential gene conversion by γ2b chain gene. Cell **32:** 515–523.

ONISHI, T., C. BERGLAND and R. H. REEDER, 1984 On the mechanism of nucleolar dominance in mouse-human somatic cell hybrids. Proc. Natl. Acad. Sci. USA **81:** 484–487.

SZOSTAK, J. W., T. L. ORR-WEAVER, R. J. ROTHSTEIN and F. W. STAHL, 1983 The double-strand-break repair model for recombination. Cell **33:** 25–35.

WALSH, B., 1983 Role of biased gene conversion in one-locus neutral theory and genome evolution. Genetics **105:** 461–468.

Corresponding editor: B. S. WEIR