Genet. Res., Camb. (1972), **20**, *pp.* 1–7
With 2 text-figures
Printed in Great Britain

Fixation time of overdominant alleles influenced by random fluctuation of selection intensity*

BY TOMOKO OHTA AND MOTOO KIMURA
National Institute of Genetics, Mishima, Japan

(*Received* 8 *September* 1971)

SUMMARY

It was demonstrated that the number of generations until fixation or loss of an overdominant alleles is influenced by random fluctuation of selection coefficients. When $2\bar{s} < V_s$, where $\bar{s}$ is the mean selection coefficient against either homozygote and V_s is the between-generation variance of the selection coefficient, overdominance generally accelerates rather than retards fixation of segregating alleles. This finding should have important bearing on our consideration of the behaviour of polymorphic variants which are nearly neutral but have very slight overdominance. When the population size (N_e) is extremely large, not only $N_e\bar{s}$ but also $\bar{s}/V_s$ have to be considered in discussing the effectiveness of overdominance.

1. INTRODUCTION

The average number of generations until a mutant gene becomes fixed in the population is important for our understanding of the causes of genetic polymorphism. Since overdominance is the most popular mechanism proposed for the explanation of polymorphisms, one might naturally expect that overdominant alleles spend much more time until fixation as compared with the neutral case. However, as Robertson (1962) has shown, overdominance is ineffective as a factor retarding fixation when the selection coefficient is small and/or the equilibrium frequency is outside the range 0·2–0·8. A similar conclusion was obtained by Ewens & Thomson (1970) in terms of mean time retardation factor. If N_e is the effective population size and s_1 and s_2 are the selection coefficients against the two homozygotes, then $N_e(s_1+s_2)$ must be fairly large to keep the polymorphic state significantly longer, especially when the equilibrium frequency differs from 0·5.

In real biological populations it is likely that the selection coefficients are not constant but fluctuate from generation to generation due to changes in environment and genetic background. Recently, Crow (1971) pointed out, in relation to molecular polymorphism, that if a mutant is neutral on the average but has a fluctuating selection coefficient, then the time until fixation is shortened, especially when N_eV_s is large, where V_s is the between-generation variance of selection coefficient of the mutant (assuming genic selection). Also Ohta (1971) showed that the fixation probability of a mutant is influenced by random fluctuation of selection intensity

* Contribution no. 852 from the National Institute of Genetics, Mishima, Shizuoka-ken, 411, Japan. Aided in part by a Grant-in-Aid from the Ministry of Education, Japan.

when the absolute value of the average selection coefficient does not exceed the variance. Such a condition should apply to mutants having very small effects in a large population.

In our present paper we shall examine the effect of overdominance as a factor retarding fixation when the selection intensity fluctuates from generation to generation. It will be shown that the overdominance is again ineffective when the variances are larger than the means of selection coefficients.

2. BASIC THEORY AND RESULTS

Let us consider a diploid population having the 'variance' effective number N_e (for the meaning of N_e, see Crow & Kimura, 1970, p. 352). We assume that a pair of overdominant alleles, A_1 and A_2 are segregating in the population. Let the fitnesses of the three genotypes be as follows:

Genotype ...	A_1A_1	A_1A_2	A_2A_2.
Fitness ...	$1-s_1$	1	$1-s_2$.

The selection coefficients s_1 and s_2 are not constant but fluctuate from generation to generation with mean $\bar{s}\,(>0)$ and variance V_s. We assume that there is no correlation between s_1 and s_2.

We shall evaluate here the mean time until loss or fixation of A_1, starting from the initial frequency 1/2. Let us denote it by $T(\frac{1}{2})$. The general solution for the time until either fixation or loss starting from an initial frequency p has been given by Ewens (1963) and Kimura (1971) as follows;

$$T(p) = u(p)\int_p^1 \psi(\xi)\{1-u(\xi)\}\,d\xi + \{1-u(p)\}\int_0^p \psi(\xi)\,u(\xi)\,d\xi, \tag{1}$$

where

$$u(p) = \int_0^p G(x)\,dx \Big/ \int_0^1 G(x)\,dx$$

and

$$\psi(\xi) = 2\int_0^1 G(x)\,dx/\{V_{\delta\xi}G(\xi)\},$$

where

$$G(x) = \exp\left\{-\int_0^x \frac{2M_{\delta y}}{V_{\delta y}}\,dy\right\}.$$

$M_{\delta y}$ and $V_{\delta y}$ are the mean and the variance of the change per generation of gene frequency y.

In the present case, $M_{\delta y}$ and $V_{\delta y}$ are respectively

$$M_{\delta y} = \bar{s}(1-2y)\,y(1-y)$$

and

$$V_{\delta y} = V_s\{(1-y)^2+y^2\}y^2(1-y)^2 + \frac{y(1-y)}{2N_e}. \tag{2}$$

Then it can be shown that

$$G(x) = \exp\left\{-K\log\frac{1+\alpha x(1-x)}{1-\beta x(1-x)}\right\}, \tag{3}$$

https://doi.org/10.1017/S0016672300013549 Published online by Cambridge University Press

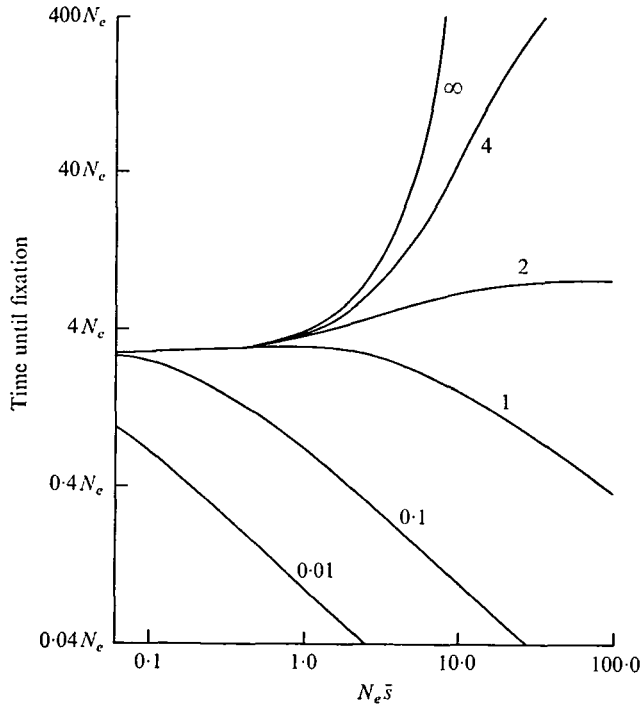

Fig. 1. Time until fixation of an overdominant allele starting from the initial frequency of 1/2 as a function of $N_e\bar{s}$. The figure beside each curve shows the value of K which is equal to $4N_e\bar{s}/\sqrt{[2N_eV_s(8+2N_eV_s)]}$. The relation $K = \infty$ implies $V_s = 0$.

where

$$K = \frac{A}{\sqrt{[B(8+B)]}},$$

$$\alpha = \frac{\sqrt{(2B)}}{\sqrt{(1+\frac{1}{8}B)-\sqrt{(\frac{1}{8}B)}}}$$

and

$$\beta = \frac{\sqrt{(2B)}}{\sqrt{(1+\frac{1}{8}B)+\sqrt{(\frac{1}{8}B)}}}$$

with $A = 4N_e\bar{s}$ and $B = 2N_eV_s$. $T(\frac{1}{2})$ can be expressed as

$$T(\tfrac{1}{2}) = 2\int_0^{\frac{1}{2}} \frac{\int_0^{\xi} G(x)\,dx}{V_{s\xi}G(\xi)}d\xi. \tag{4}$$

This equation (4) was integrated numerically using a computer. Figs. 1 and 2 give the results of integration. Fig. 1 shows the relationship between $T(\frac{1}{2})$ and $N_e\bar{s}$ for various values of $K = 4N_e\bar{s}/\sqrt{[2N_eV_s(8+2N_eV_s)]}$. Similarly Fig. 2 shows the relationship between $T(\frac{1}{2})$ and K for various values of $N_e\bar{s}$. Note here that for completely neutral alleles $T(\frac{1}{2}) = 4N_e\log_e 2 \approx 2{\cdot}8N_e$. From the figures it can be seen that when K is smaller than unity the overdominance is generally ineffective for maintaining the polymorphism. The condition $K < 1$ implies roughly $V_s > 2\bar{s}$ when $N_eV_s \gg 1$.

https://doi.org/10.1017/S0016672300013549 Published online by Cambridge University Press

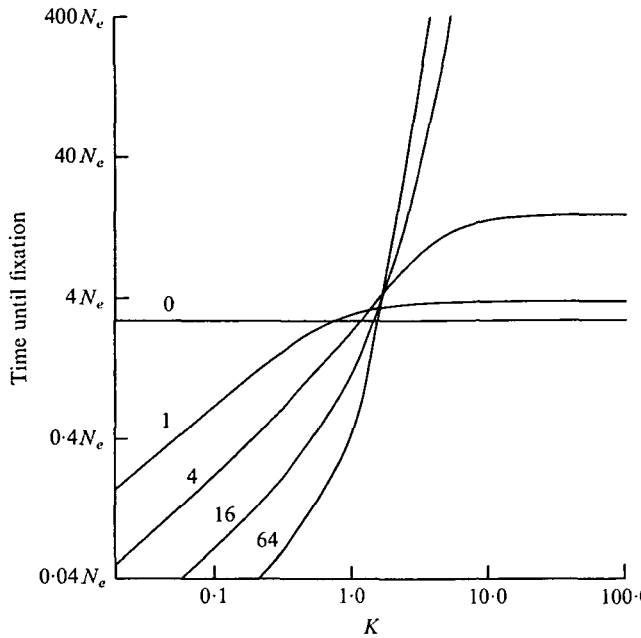

Fig. 2. Time until fixation of an overdominant allele starting from the initial frequency 1/2 as a function of K ($= 4N_e\bar{s}/\sqrt{[2N_e V_s(8+2N_e V_s)]}$. The figure beside each curve shows the value of $N_e\bar{s}$. The straight line with $N_e\bar{s} = 0$ shows the case with $\bar{s} = V_s = 0$.

In the present analysis we have assumed a most favourable situation for maintaining a polymorphism, namely $\bar{s}_1 = \bar{s}_2$ with equilibrium gene frequency at 0·5. Yet, we obtained the result that the fixation time becomes shorter with overdominance as compared with the neutral case if $K < 1$.

Dr A. Robertson (personal communication) has pointed out that a similar result can be obtained through much simpler approach as follows: The change of gene frequency per generation by selection can be expressed as

$$\begin{aligned}\Delta p &= -p(1-p)\,[s_1 p - s_2(1-p)] \\ &= -p(1-p)\,2\bar{s}(p-\tfrac{1}{2}) - p(1-p)\,[p\,\delta s_1 - (1-p)\,\delta s_2],\end{aligned}$$

where p is the frequency of A_1 and $\delta s_i = s_i - \bar{s}$. Since the variance of gene frequency is

$$V_p = (p-\tfrac{1}{2})^2,$$

the change per generation of V_p by selection is approximately

$$\begin{aligned}\Delta V_p &= 2(p-\tfrac{1}{2})\,\Delta p + (\Delta p)^2 \\ &= -\bar{s}V_p + V_s/32,\end{aligned}$$

where we assume that deviation of p from 1/2 is small. At equilibrium, this will be balanced by the effect of random sampling of gametes that increases the variance by the amount $p(1-p)/2N_e \approx 1/8N_e$, so that

$$-\bar{s}V_p + V_s/32 + 1/(8N_e) = 0.$$

https://doi.org/10.1017/S0016672300013549 Published online by Cambridge University Press

Table 1. *Comparison of theoretical prediction and the result of Monte Carlo experiment for the time until fixation of overdominant alleles*

(The effective size of the population is 40 and each experimental value is the average of 200 replications. For details of the experimental procedure, see text.)

$\bar{s}$	V_s	Theoretical prediction	Monte Carlo experiment
0·05	0·025	165·3	156·1
	0.1	89·6	79·8
	0·2	56·1	53·8
0·1	0·025	371·2	347·2
	0·1	144·6	139·5
	0·2	76·2	70·6
0·2	0·025	3498·8	2698·4
	0·1	537·8	446·8
	0·2	169·7	158·5

Thus we obtain

$$V_p = \frac{1}{8N_e\bar{s}} + \frac{V_s}{32\bar{s}}.$$

By comparing this with the variance of the uniform distribution – that is, $V_p = 1/12$ – and assuming that $N_e\bar{s} \gg 1$, we conclude that the overdominance is ineffective if $V_s/(32\bar{s}) > 1/12$ or

$$V_s > 8\bar{s}/3 = 2{\cdot}67\bar{s},$$

which is slightly stronger than $V_s > 2\bar{s}$, which we obtained above.

3. MONTE CARLO EXPERIMENTS

Monte Carlo experiments were carried out in order to check the validity of the above analysis. The procedure of the experiments follows essentially the one used by Hill & Robertson (1966). In each generation selection coefficients s_1 and s_2 are determined by generating two random numbers which follow the normal distribution having the mean $\bar{s}$ and the variance V_s. This was done using the subroutine RAND 10 in our computer TOSBAC 3400. Selection is done deterministically using the formula

$$\delta x = [s_2(1-x) - s_1 x]\, x(1-x)$$

for the change of the frequency of A_1. After selection, zygotic frequencies of three genotypes are determined by random union of gametes. Sampling is carried out by generating pseudo-random numbers with rectangular distribution (RAND 20 in TOSBAC 3400). Namely, if f_i's are zygotic frequencies and if a random number lies between $\sum_{i=1}^{j-1} f_i$ and $\sum_{i=1}^{j} f_i$, then one individual with jth genotype is sampled (if less than f_1, the first genotype). Sampling is repeated N_e times to obtain individuals in the next generation. Each experiment is continued until fixation of one or the other allele.

Table 1 shows the results of the experiments. Three levels of $N_e\bar{s}$ and three levels of N_eV_s have been used. The effective size of the population was 40, and 200 replications were made for each set of parameters. As seen from the table, the agreement

between the theoretical prediction and the result of Monte Carlo experiment is on the whole satisfactory. However, theoretical predictions appear to overestimate slightly the true value.

4. DISCUSSION

As shown by the present analysis, overdominance is ineffective as a factor retarding fixation when the variance of the selection coefficient is larger than twice the mean. This means that the effect of variation in fitness is important only when the average selection coefficient is already small, say less than 10^{-3}. However, this should have an important bearing on our understanding of the borderline case between practically neutral and definitely overdominant polymorphisms. As pointed out by Crow (1972) and Ohta (1972), gene mutations having very small effect must play an important role on the amino acid substitutions of proteins and it is possible that the intensity of selection for such mutants does not stay constant over many generations but fluctuates from generation to generation due to change in environment or the genetic background.

Overdominance at the molecular level may be considered to be interallelic complementation as discussed by Fincham (1966). From the standpoint of population genetics, overdominance may be understood in terms of differential effects of alleles on many components of fitness (Ohta & Kimura, 1971). For example, if A_1A_1 homozygote reduces the fertility by s_1, and A_2A_2 reduces the viability by s_2, then overdominance will result with respect to the overall fitness. Such homozygous effects may fluctuate independently of each other. Furthermore, it is possible that the smaller the effect, the more likely they are subject to random fluctuation.

In discussing the role of balancing selection for the maintenance of protein polymorphisms, overdominance with very small selection coefficients such as 10^{-5} or 10^{-6} has sometimes been assumed as operating on a very large number of loci. It is claimed that such a weak overdominance can still work effectively in a species like *Drosophila*, whose population size is so large that the value of $N_e\bar{s}$ can still be large. However, in such a case, not only $N_e\bar{s}$ but also the relative magnitude of the mean and the variance of the selection coefficients has to be considered. For example, in the symmetric model, if $\bar{s} = 10^{-5}/2$ and $V_s \geqslant 10^{-5}$ (i.e. $\sigma_s \geqslant 1/300$), the overdominance does not work effectively.

If the same set of alleles are segregating over millions of generations, or in a very large area such as a whole continent as reported by Prakash, Lewontin & Hubby (1969) and Ayala *et al.* (1970), it is hard to imagine that the selective force remains constant throughout space or time. It is likely that the selection intensity varies from time to time or from region to region. Thus, the variation in fitness has to be considered when we discuss the effectiveness of overdominance having very small selection coefficients.

Also, we should point out that the associative overdominance due to linkage disequilibrium (Ohta & Kimura, 1969, 1970, 1971) is in some cases ineffective as a retardation factor because linkage will increase variance in fitness values. More detailed study on this subject will be published elsewhere.

We would like to thank Dr Alan Robertson for constructive criticism.

REFERENCES

AYALA, F. J., MOURÃO, C. A., PÉREZ-SALAS, S., RICHMOND, R. & DOBZHANSKY, TH. (1970). Enzyme variability in the *Drosophila willistoni* group. I. Genetic differentiation among sibling species. *Proceedings of the National Academy of Sciences, U.S.A.* **67**, 225–232.

CROW, J. F. (1972). Darwinian and non-Darwinian evolution. *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability*. (In the Press.)

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

EWENS, W. J. (1963). The diffusion equation and a pseudo-distribution in genetics. *Journal of the Royal Statistical Society* B **25**, 405–412.

EWENS, W. J. & THOMSON, G. (1970). Heterozygote selective advantage. *Annals of Human Genetics* **33**, 365–376.

FINCHAM, J. R. S. (1966). *Genetic Complementation*. New York: Benjamin.

HILL, W. G. & ROBERTSON, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research* **8**, 269–294

KIMURA, M. (1971). Theoretical foundation of population genetics at the molecular level. *Theoretical Population Biology* **2**, 174–208.

OHTA, T. (1972). Fixation probability of a mutant influenced by random fluctuation of selection intensity. *Genetical Research* **19**, 33–38.

OHTA, T. & KIMURA, M. (1969). Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. *Genetics* **63**, 229–238.

OHTA, T. & KIMURA, M. (1970). Development of associative overdominance through linkage disequilibrium in finite populations. *Genetical Research* **16**, 165–177.

OHTA, T. & KIMURA, M. (1971). Behaviour of neutral mutants influenced by associated overdominant loci in finite populations. *Genetics* **69**, 247–260.

PRAKASH, S., LEWONTIN, R. C. & HUBBY, J. L. (1969). A molecular approach to the study of genetic heterozygosity in natural populations. IV. Pattern of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. *Genetics* **61**, 841–858.

ROBERTSON, A. (1962). Selection for heterozygotes in small populations. *Genetics* **47**, 1291–1300.