RECOMBINATION MODIFICATION IN A
FLUCTUATING ENVIRONMENT

BRIAN CHARLESWORTH

*School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG,
Sussex, England*

ABSTRACT

This paper examines the theory of the evolution of increased recombination
between two loci subjected to interactive selection in a temporally fluctuating
environment. Both cyclical and stochastic environments are considered. It is
shown that temporal variation in the linkage disequilibrium coefficient for the
pair of selected loci, due to fluctuations in the selective values of the genotypes
at these loci, can give rise to selection in favor of modifier genes increasing
recombination. The equilibrium level of recombination established in a given
population depends on several factors; it is highest for intermediate values of
the environmental periodicity or autocorrelation, for cases when the modifier
genes are themselves closely linked to the selected loci, and for high levels of
environmental variation. In general, it seems that the rate of modification of
recombination values by this process will be low except when the modifiers are
tightly linked to the selected loci. The possible evolutionary significance of this
process is discussed in relation to observations on genetic systems of plants and
animals.

THERE have been several recent theoretical studies of the fate of selectively
neutral modifiers of recombination introduced into multi-locus (usually two-
locus) systems in randomly mating populations at equilibrium under selection
(NEI 1967, 1969; FELDMAN 1972; DEAKIN 1972; CHARLESWORTH and CHARLES-
WORTH 1973; KARLIN and McGREGOR 1974). These studies show that modifiers
which reduce recombination values tend to spread when there is linkage disequi-
librium among the selected loci, and thus support the conjecture of FISHER (1930)
that there is a selective advantage to reducing recombination rates in populations
segregating for interacting genes. Moreover, LEWONTIN (1971) showed that a
multi-locus population achieves its maximum mean fitness when there is no
recombination among the selected loci, provided that there are constant genotypic
fitnesses and random mating. This suggests that multi-locus populations in con-
stant environments tend to evolve towards a state where there is zero recombi-
nation. This suggestion is strengthened by the considerations given in APPENDIX 1
of this paper, which demonstrate that a multi-locus system at a stable equilibrium
with no recombination is either stable to the introduction of new modifiers induc-
ing recombination, or incorporates them at a rate proportional at most to the
square of their frequency (i.e. they are effectively neutral).

All this raises the question of why, as TURNER (1967) put it, the genome does not congeal. In this paper, I wish to explore a possible answer which has been discussed recently by several authors (e.g. WILLIAMS 1966; MAYNARD SMITH 1971; GHISELIN 1974): the environment does not remain constant from generation to generation, but fluctuates in such a way that the combinations of alleles which are favored by selection change with time. In these circumstances, it seems intuitively likely that some degree of recombination will be selectively favored. (An alternative mechanism, which does not involve the necesssity of environmental change has been proposed by STROBECK et al. (1976).

I shall consider a model with two selected loci with alleles A, a and B, b, respectively. There are thus four possible gamete types AB, Ab, aB and ab, with frequencies $x_1^{(k)}$, $x_2^{(k)}$, $x_3^{(k)}$ and $x_4^{(k)}$ respectively in generation k. Marginal fitnesses can be assigned to each gamete (see Appendix 1); let $w_i^{(k)}$ be the marginal fitness of gamete i in generation k. In the absence of modifiers of recombination, the recombination fraction for loci A and B can be represented by a constant, R. In generation k, the degree of nonrandom combination between alleles at the two loci is measured by the linkage disequilibrium parameter, $D_k = x_1^{(k)} x_4^{(k)} - x_2^{(k)} x_3^{(k)}$. The recurrence relations for changes in gamete frequencies are standard and will not be given here. In the remainder of this paper it will be assumed that a randomly mating, infinite two-locus population which is exposed to a fluctuating environment comes to a steady-state where both loci are maintained polymorphic and where gametic frequencies vary in response to the changing genotypic fitnesses. The fates of inversions and of genic modifiers of the amount of recombination between the selected loci, introduced into such a population, will be examined. I shall start by considering inversions, as the results derived for this case provide a useful basis for understanding the fate of modifiers of recombination of less extreme effect.

Selection on a new inversion in a fluctuating environment

An inversion introduced into gamete type i will be assumed to remain permanently associated with that gamete as a result of the suppression of recombination in inversion heterozygotes. The inversion is also assumed not to affect viability or reproductive performance. Let the frequency of the inversion in generation k be y_k; the inversion will initially be rare so that y_k^2 can be neglected in comparison with y_k. Given these assumptions, the formulae of DEAKIN (1972) and CHARLESWORTH and CHARLESWORTH (1973) can be generalized to give

$$\frac{y_n}{y_o} = \prod_{k=x_0}^{n-1} \frac{w_i^{(k)}}{\bar{w}_k} \tag{1}$$

where $\bar{w}_k$ is the population mean fitness in generation k. For a cyclically varying environment, equation (1) implies (HALDANE and JAYAKAR 1963) that the inversion will eventually increase in frequency if the geometric mean (G.M.) of $w_i^k/\bar{w}_k$ over the equilibrium set of values generated by the cycle exceeds 1. If the G.M. is less than 1, the inversion will be eliminated. For the case when the

environment follows a stationary stochastic process we have the corresponding conditions that the inversion will achieve a frequency outside the neighborhood of $\gamma=0$ with a probability close to 1 when $E\{ln\ w_i^{(k)}/\bar{w}_k\}>0$ (the expectation is taken over the steady-state distribution of values of $w_i^{(k)}/\bar{w}_k$ generated by the environmental distribution); when $E\{ln\ w_i^{(k)}/\bar{w}_k\}<0$, γ_k occupies a value in the neighborhood of 0 with probability close to 1 (GILLESPIE 1973; KARLIN and LIEBERMAN 1974).

It is obvious that if the conditions for elimination of an inversion are satisfied for inversions introduced into *each* of the four gamete types, the population resists the introduction of this particular type of genetic element reducing recombination. If the condition for incorporation of an inversion is satisfied for even one of the gamete types, then the population cannot be said to be resistant to all genetic elements reducing recombination. Resistance to the introduction of an inversion in all four gametes is thus necessary, but not sufficient, for a population with a given value of R to be stable to the introduction of modifiers reducing recombination, and provides a useful criterion for finding selective systems which might favor the evolution of greater R. In the remainder of this section, we shall be concerned with evaluating the conditions under which this criterion is satisfied.

Consider first the simple case when the matrix of fitnesses of the genotypes at loci A and B alternates between two values, in a cycle with a fixed period. When the environmental period is long, so that the population spends a long time in a given environment before the switch-over occurs, the population (in the absence of an inversion) will spend most of its time at or near an equilibrium corresponding to the currrent fitness matrix. In simple cases, therefore, the population will tend to alternate between 2 equilibria corresponding to the 2 fitness matrices. Let $\hat{x}_i^j$ and $\hat{w}_i^j$ $(j=1,2)$ represent the equilibrium frequency and marginal fitness of gamete i in a population kept in a constant environment of type j. We conclude that the G.M. of $w_i^{(k)}/\bar{w}_k$ for a population in a fluctuating environment of long period is approximated by $\sqrt{\hat{w}_i^1\hat{w}_i^2/\hat{\bar{w}}_1\hat{\bar{w}}_2}$, where $\hat{\bar{w}}_1$ and $\hat{\bar{w}}_2$ are the equilibrium mean fitnesses for populations kept in constant environments of types 1 and 2 respectively. For the population to resist the introduction of an inversion, this quantity must be less than 1 for each i. If $\hat{D}_1$ and $\hat{D}_2$ are both the same sign (where $\hat{D}_j = \hat{x}_1^j\hat{x}_4^j - \hat{x}_2^j\hat{x}_3^j$), it is impossible to satisfy this condition, since we have $\hat{w}_1^j$, $\hat{w}_4^j \geqslant \hat{\bar{w}}_j \geqslant \hat{w}_2^j$, $\hat{w}_3^j$ when $\hat{D}_j \geqslant 0$, and the reverse inequalities when $\hat{D}_j \leqslant 0$. For an inversion always to be eliminated under these assumptions, it is necessary that the sign of $\hat{D}$ for an equilibrium population in a constant environment is determined by the fitness values for that environment, and that the alternative fitness values encountered with varying environments would give $\hat{D}$'s with different signs for populations kept in uniform environments of one or other type. This automatically rules out certain classes of fitness matrix such as multiplicative fitnesses, the symmetric model of LEWONTIN and KOJIMA (1960), and optimizing selection models. With the symmetric model of BODMER and FELSENSTEIN (1967) (see Table 1), the sign of $\hat{D}$ in an equilibrium population with

B. CHARLESWORTH

TABLE 1

The symmetric fitness model

	AA	Aa	aa
BB	$1-\delta$	$1-\beta$	$1-\alpha$
Bb	$1-\gamma$	1	$1-\gamma$
bb	$1-\alpha$	$1-\beta$	$1-\delta$

small R can be shown analytically to be the same as the sign of $\alpha-\delta$ when the fitness interactions are such that $\beta+\gamma<\alpha,\delta$ (BODMER and FELSENSTEIN 1967). Numerical solutions show that to be true also for arbitrary R in many cases. Hence, variation in the sign of $\alpha-\delta$ is, in this case, both necessary and sufficient for the population to resist the introduction of an inversion.

All this is for environments of large period. As the period of the environment is reduced, the population will spend more of its time in the transient states connecting the two alternative equilibria. It seems reasonable to assume that the ratio $w_i^{(k)}/\bar{w}_k$ for the transient states will lie between the two extreme values for the alternative equilibria, so that the G.M. will be closer to 1 for environments of small period.

These conclusions were checked numerically as follows. From equation (1), it can be seen that the geometric mean of $w_i^{(k)}/\bar{w}_k$ approximates the geometric mean of γ_n/γ_{n-1}, the ratio of the frequencies of a rare inversion introduced into gamete type i in successive generations. The above theory can then be tested by computing the G.M.'s of both these quantities for specific cases. The results of one such calculation are shown in Table 2. The fitness matrices for the two alternative environments were both of the symmetric type (Table 1 above); β and γ were held constant at values of 0.25 and 0.10 respectively while

TABLE 2

Selection on a new inversion in a periodic environment

Period	$R=0.05$		$R=0.10$	
	G.M.$\{w_1^{(k)}/\bar{w}_k\}$	G.M.$\{\gamma_n/\gamma_{n-1}\}$	G.M.$\{w_1^{(k)}/\bar{w}_k\}$	G.M.$\{\gamma_n/\gamma_{n-1}\}$
1	1.000	1.000	1.001	1.000
5	0.998	0.997	0.995	0.995
10	0.992	0.993	0.987	0.987
25	0.979	0.981	0.977	0.977
50	0.974	0.975	0.972	0.973
∞	0.967	——	0.969	——
Period	$R=0.25$		$R=0.50$	
1	1.002	1.001	1.004	1.002
5	0.991	0.991	0.988	0.985
10	0.983	0.982	0.984	0.985
25	0.978	0.979	0.982	0.980
50	0.976	0.976	0.981	0.980
∞	0.974	——	0.980	——

α and δ were switched from 1, 0.30 to 0.30, 1 respectively. Because of the symmetry of the selective model, only the results for gamete type 1 (AB) are given. The G.M.'s for $w_1^{(k)}/\bar{w}_k$ were calculated from the values over 1,000 generations for two-locus populations without any inversions, started with frequencies $x_1 = x_4 = 0.27$; $x_2 = x_3 = 0.23$. (The populations rapidly approached the steady-state cycle of gamete frequencies, so that the error in using these G.M.'s as estimates of the steady-state values is negligible.) The values for a period of infinity are calculated using the formula for environments of long period given above. As can be seen from the table, the G.M.'s of $w_1^{(k)}/\bar{w}_k$ increase towards the values for an infinite period as the period increases. For an environment of period 1, there is actually weak selection in favor of an inversion. These results agree well with the results based on direct calculation of the G.M.'s of γ_n/γ_{n-1} for an inversion introduced with an initial frequency of 1%. These G.M.'s are based on the inversion frequencies from generations 100 to 200. Except for these cases in which the G.M. of $w_1^{(k)}/\bar{w}_k$ is so close to 1 that terms of order γ_k^2 cannot be neglected safely, it will be seen that equation (1) appears to predict accurately the progress of the inversion.

These sorts of considerations can easily be extended to cyclical changes in selective values which are more complex than the simple switching of two fitness matrices. For example, with the alternation of three different fitness matrices at long intervals, the G.M.'s of $w_i^{(k)}/\bar{w}_k$ are approximated by the products of the equilibrium values for the three different environments, and these cannot be less than 1 for each gamete if D has the same value in each environment. The basic conclusion remains true: a necessary condition for a two-locus population to resist the introduction of an inversion is that the variation between environments is such that there should be variation in the sign of D between the equilibria in the alternative environments.

These conclusions depend on the assumption that the alternative environments each generate a stable two-locus polymorphism. It is conceivable that situations might exist in which the alternative environments would not support such stable equilibria: for example, one equilibrium might be AB/AB and the other ab/ab. APPENDIX 2 gives an approximate analysis of the selection coefficient of a new inversion, which does not depend on this assumption. Because of the approximations involved in deriving the conclusions in APPENDIX 2, notably that changes in gamete frequencies are slow, this analysis is more appropriate for the case of a stochastically varying environment than for one in which there are sudden switches in selective values, and so it will be considered here under that heading. APPENDIX 2 should be consulted for the definitions of symbols used here.

As discussed in APPENDIX 2, for there to be a negative selection coefficient on an inversion in AB when $<D> > 0$, the variance of D must normally be so large that D will take negative values some of the time. If $<D> = 0$, variance in D automatically implies the occurrence of negative D values. If $<D> < 0$, then s_1 is usually < 0 regardless of the value of V_D: the corresponding formulae for s_2 and s_3 yield the condition that for s_2 and s_3 also to be < 0, D must take positive values some of the time. These considerations thus agree with the conclusions

developed above: for the population to resist the introduction of a new inversion, the nature of selection must usually be such that the sign of the linkage disequilibrium between the selected loci varies in time. From formulae such as (A.2.3) this is clearly not a sufficient condition for there to be immunity to inversions: if the mean value of D is sufficiently great in absolute value, there may still be selection in favor of a new inversion despite the varying environment. For a given $<D>$, the greater V_D the greater is the likelihood of selection against new inversions.

The approximations of APPENDIX 2 were tested numerically by running two-locus populations without inversions in stochastically varying environments to obtain estimates of $<D>$, V_D etc., and then calculating the progress of new inversions in order to compare the value of the selection coefficients obtained directly with that given by the theoretical formulae involving $<D>$, etc. The stochastic environment most studied was one using the symmetric fitness model in which α and δ in any generation were represented by $\bar{a} + \varepsilon$ and $\bar{\delta} - \varepsilon$, where ε was a variable following a stationary Gaussian process with zero mean, variance σ^2 and autocorrelation ρ. (If one of $\bar{a} + \varepsilon$ or $\bar{\delta} - \varepsilon$ were < 0, ε for both fitnesses were replaced by ε', such that the negative fitness value was made zero.) Some results of this sort with $\bar{a} = \bar{\delta} = 0.65$, $\beta = 0.25$ and $\gamma = 0.10$ are shown in Table 3. In this case $<D>$ is expected to be zero, and little change in gene frequencies is expected, from the symmetry of the selective model. Also, in a selective model of this sort little correlation between D and $\bar{w}$ is to be expected. Equation (A.2.4) with the term involving Cov $(D, \bar{w})$ omitted was therefore used in computing the theoretical selection coefficients. Each population was run for 2,000 generations, with and without an inversion in AB, and 200 replicates of each run were made. V_D was estimated as the average within-run, between-generation variance in D. The standard errors between runs of $<D>$ and of the theoretical and computed selection coefficients are given in the Table. Within the limits of these standard errors there is reasonably good agreement between most of the theoretical and directly computed selection coefficients, although a general tendency for the theoretical selection coefficient to be smaller than the observed is noticeable, particularly for low autocorrelations.

It will be seen from the table that, as expected from the form of equation (A.2.4), the magnitude of the selection against a new inversion increases with the recombination fraction and with V_D. The latter is, as expected, an increasing function of σ^2, the environmental variance, and also of ρ. (Clearly, for high values of ρ, this relationship must be reversed, since $\rho = + 1$ would generate no environmental change.) This latter relation corresponds with the dependence of the selection pressure against an inversion on the environmental period in the cyclical environment case.

Selection on genic modifiers of recombination

In a constant environment, a genic modifier which reduces the recombination fraction of a pair of selected loci A and B, and which is itself capable of recombination with A and B, is selected for by virtue of its ability to establish an associ-

TABLE 3

Selection on a new inversion in a stochastic environment

R	$<D>$	V_D	s_1	
			Theoretical	Computed directly
$\rho = +0.8$				
0.05	—2.8 ± 8.1	7.3	—14.8 ± 3.3	— 9.9 ± 3.0
0.10	—5.4 ± 6.4	4.7	—23.2 ± 5.1	—18.6 ± 4.3
0.25	—0.9 ± 3.8	2.2	—21.6 ± 6.2	—16.4 ± 5.1
0.50	—2.5 ± 2.3	0.9	—29.1 ± 8.6	—22.9 ± 5.8
$\rho = -0.2$				
0.05	0.7 ± 2.6	1.4	— 2.1 ± 1.3	1.3 ± 0.9
0.10	—0.1 ± 1.7	1.0	— 3.7 ± 1.3	— 1.3 ± 1.2
0.25	—0.9 ± 1.0	0.8	— 7.2 ± 1.7	— 4.5 ± 1.4
0.50	0.0 ± 0.6	0.7	—12.6 ± 2.2	— 4.6 ± 1.5

R	$<D>$	V_D	s_1	
			Theoretical	Computed directly
$\rho = +0.2$				
0.05	1.5 ± 3.7	2.4	— 3.7 ± 1.4	— 1.8 ± 1.4
0.10	0.8 ± 2.7	1.8	— 5.6 ± 2.1	— 3.3 ± 1.8
0.25	—1.8 ± 1.7	1.2	—12.1 ± 2.7	— 9.0 ± 2.2
0.50	0.5 ± 0.9	0.8	—12.1 ± 3.4	— 6.4 ± 2.3
$\rho = -0.8$				
0.05	—0.3 ± 1.1	0.4	— 0.8 ± 0.4	2.2 ± 0.4
0.10	0.0 ± 0.7	0.4	— 1.2 ± 0.5	1.2 ± 0.5
0.25	—0.4 ± 0.4	0.4	— 3.4 ± 0.7	— 0.8 ± 0.6
0.50	—0.2 — 0.2	0.5	—10.3 ± 0.9	— 0.6 ± 0.6

$\sigma = 0.10$. All the values of $<D>$, V_D and s_1 are $\times 10^4$.

ation with whichever pair of gametes have marginal fitnesses higher than the mean (NEI 1967, 1969). In an environment in which sometimes the gametes AB and ab and sometimes Ab and aB are favored, a genic modifier will be able to switch across to the selectively advantageous gamete types whenever the environment changes; the looser the linkage of the modifier to the selected loci, the faster this switchover will occur. An inversion is, of course, irreversibly tied to one gamete type. One would therefore expect that populations in a varying environment which resist the introduction of an inversion may still incorporate genic modifiers which reduce recombination, and that the looser the linkage between the modifier locus and the selected loci, the more likely it is that an allele reducing recombination will spread. Furthermore, the longer the period of a cyclical environment, or the higher the autocorrelation of a stochastic environment, the greater the opportunity for a crossover reducer to become associated with the currently favored gametes before the environment changes. This will be counterbalanced by the tendency, noted in the previous section, for the strength of selec-

tion against an element associated with one particular gamete type to increase with the environmental period or autocorrelation. The strength of this selection is also a function of the variance of D.

These considerations suggest that:

(a) If variation in recombination is due to segregation of alleles at a modifier locus with a given degree of linkage to the selected loci, the equilibrium value of R, R_E, which is established in a given selective system, will vary with the environmental period or autocorrelation, such that the highest value of R_E is attained in environments with intermediate values of the period or autocorrelation.

(b) The looser the linkage of the modifier locus to the selected loci in a given selective system and a given pattern of environmental variation, the lower the value of R_E.

(c) In a stochastically varying environment, the greater the environmental variance, and hence V_D, the higher R_E, for a given degree of linkage of the modifier locus and a given autocorrelation. Similarly, in a cyclical environment R_E will be higher the greater the differences in values of D between the equilibria in the alternative environments.

We may also note that the selection pressure for recombination modification in a fluctuating environment is likely to be low for all but very tightly linked modifiers, since it is known to be low in system in constant environments and, furthermore, in this case the direction of the selection pressure is constantly fluctuating.

These conclusions were borne out by numerical calculations of systems with two selected loci A and B, as before, with a modifier of recombination with alleles M and m. The gene order in all the calculations was assumed to be ABM. Calculations were made for systems with two types of modifier locus.

(i) *Unlinked modifier*. Here the recombination fraction for B and M was set equal to 50%, and there was no interference.

(ii) *Linked modifier*. The distance B-M was 0.05, and the coefficient of coincidence was 0.2.

In order to locate the approximate value of R_E for a given selective system, pattern of environmental change, and degree of linkage of B and M, the following procedure was adopted. Populations were started with 99% of the allele m (always in the gamete AB) and run for 400 generations. The output was examined to see whether M was increasing or being eliminated, allowing the initial hitch-hiking effects due to the association with AB to die out. For each system the following types of populations were run (the R values in mm, Mm and MM are represented as elements in a 3-vector): (.5, .25, .12), (.25, .12, .06), (.06, .12, .25) and (.12, .25, .50). If, for example, R_E for a given system is located between .12 and .25, M should increase in each case. If R_E is located between .06 and .12, M should increase in the first three cases, but decrease in the last.

The results of calculations using cyclical environments and the fitness matrices employed in constructing Table 2 are shown in Table 4. It will be seen that the conclusions outlined above seem to hold good with this example. The rates of change in modifier frequency are, as expected, slow, even with linkage. For example, an unlinked modifier increasing R from .06 to .25 in an environment

TABLE 4

Equilibrium values of recombination established in a periodic environment

	Equilibrium recombination value (R_E)	
Period	Linked modifier	Unlinked modifier
1	< 0.06	< 0.06
2	0.50	0.50
3	0.50	0.50
4	0.50	0.50
5	0.50	0.50
10	$0.12 < R_E < 0.25$	$R_E \approx 0.12$
25	$0.06 < R_E < 0.12$	< 0.06
50	< 0.06	< 0.06

of period 5 changed from 1% to 2.6% in 400 generations, and a linked modifier with the same effect from 1% to 3.5%. Similar results were obtained with stochastically varying environments. Selective modification of the recombination fraction was exceedingly slow in these cases. For example, Table 5 shows the course of change in the mean recombination fraction for a linked modifier which increases R from .12 to .50 in a stochastic environment of the type of Table 3 with $\sigma = 0.2$ and $\rho = 0.2$.

DISCUSSION

The results presented above yield the conclusions that

(a) Increased recombination between two selected loci may be favored when the environment varies in such a way that the linkage disequilibrium varies in time. If there is no temporal variaton in the sign of D, it seems unlikely that there will often be selection for increased recombination, although the present analysis does not absolutely rule this out.[1] MAYNARD SMITH (1971) arrived by

TABLE 5

Progress of a modifier which increases recombination from 0.12 to 0.50 in a stochastic environment

Generation	Modifier frequency	Mean R
0	0.010	0.123
2,000	0.020	0.125
4,000	0.041	0.131
6,000	0.077	0.141
8,000	0.143	0.160
10,000	0.225	0.185
12,000	0.308	0.217
14,000	0.387	0.238

[1] Some recent simulations of the model of fluctuating environments proposed by STURTEVANT and MATHER (1938) show that selection for increased recombination is in fact possible without changes in the sign of D. This model involves environments in which fixation would occur if the population persisted in them, in contrast to the cases studied in detail in this paper.

a less formal argument at a somewhat similar condition for the evolution of sex, viewed as a means of promoting recombination.

(b) The selection pressure in favor of increased recombination is likely to be rather weak for genic modifiers of recombination values, although there can be reasonably strong selection against crossover suppressors, such as inversions, which are tightly linked to the selected loci (Table 2).

(c) The amount of recombination which is established by this process is heavily dependent on the degree of linkage between the selected loci and the modifiers of recombination: tight linkage between these two groups will favor a higher level of recombination than loose linkage. There is thus no such thing as a selectively optimal level of recombination for a given environment; the level of recombination which is produced in a given chromosome region will depend on the proximity of that region to available modifier genes.

(d) The equilibrium level of recombination in a system is affected, not only by the strength of the temporal variation in the environment but also by the period or autocorrelation of the environmental changes. In general, it seems that either a very short or very long period (or, equivalently, a low or high auto-corrrelation) will favor low recombination values.

We may now consider how far these results can help to interpret comparative data on genetic systems in different species of animals and plants. One observation which has been firmly established for certain species of Drosophila, but not for others (CARSON 1965), is that ecologically marginal populations tend to have a lower level of inversion heterozygosity than central populations of the same species, even to the extent of becoming homokaryotypic for one or more inversions for which the central populations are polymorphic. *Drosophila robusta* is one of the best-studied examples of this phenomenon (CARSON 1955, 1959). If one is prepared to assume that marginal populations are more subject to environmental fluctuations than central populations, then the mechanism described in this paper might be held responsible for these observations. This is in line with the ideas of CARSON, who believes that inversion homozygosity in marginal populations is due to selection for increased recombination levels. It is obvious that monomorphism for inversions is consistent with the model described here, in which new inversions are eliminated by populations subjected to fluctuating environments; in such cases, one would also expect to find old-established inversion polymorphisms to break down when the population is exposed to a fluctuating environment. Similarly, the frequency of an inversion could be reduced by a population shift from a constant environment with a fixed equilibrium frequency to an environment which, say, alternates between the fitness matrix corresponding to the original equilibrium, and a fitness matrix in which the inversion would be eliminated in a constant environment. Of course, one cannot rule out the alternative explanation of DOBZHANSKY (DA CUNHA and DOBZHANSKY 1954) that inversion polymorphism is more frequent in central areas because of the greater degree of habitat diversity there, which is assumed to promote greater

genetic variability. Recent data from electrophoretic surveys does not show any greater genic variability in central than in marginal populations of Drosophila (PRAKASH 1973), but these loci may not behave in the same way as those responsible for selection pressures on inversions.

Data on chiasma frequencies and chromosome numbers in plants are not so easy to interpret. One might expect at first sight that a high chiasma frequency and/or chromosome number would generally be favored in species associated with unstable environments, e.g. in annuals adapted to colonizing situations. In fact, it is often found that perennials and woody plants have higher chiasma frequencies or chromosome numbers than related cross-pollinated annuals (STEBBINS 1958; GRANT 1958). For instance, STEBBINS notes that in the cross-pollinated members of the group he studied (the tribe Cichorieae of the Compositae), low chromosome numbers are characteristic of species living in unstable habitats such as roadsides and fields. This corresponds roughly to the division between perennial and annual species. One should bear in mind, however, the importance of the periodicity of the environment as a determinant of the selectively favored recombination fraction. It might easily be the case that for annual species, this year's habitat is usually much like the last, and that the same evolutionarily important environmental factors such as climate are operating on both annuals and perennials. If this were so, then annual species would be exposed to environmental fluctuations of much longer period in terms of generations than perennials, which would lead to lower recombination values being favored in annuals. Little attention seems to have been given previously to the role of environmental periodicity or autocorrelation; it is clearly dangerous to disregard the possible influence of this factor when interpreting comparative data.

In trying to evaluate the possible validity of this theory as an explanation of the evolution and maintenance of non-zero recombination in higher organisms, one should bear in mind that it suffers from two important difficulties. In the first place, it is clear from the numerical results quoted earlier that, even with strong selection, the rate of increase in recombination values will be low. If polymorphic loci with no changes in the value of D are being maintained in the same region of chromosome as loci whose D values are fluctuating, then there will be a conflict between the opposing directions of selection on modifiers of recombination in the chromosome region concerned, and the net effect will depend on which force is stronger. Secondly, as pointed out by MAYNARD SMITH (1971) in connection with the evolution of sex, the type of selection discussed here, which produces changes in the sign of the linkage disequilibrium, is difficult to model in terms of phenotypic effects of the loci concerned. One has to assume that each locus is responding to a certain feature of the environment, such as temperature or humidity, and that the correlation between the possible environmental states changes from time to time. For example, one might have A/A adapted to heat, a/a to cold, B/B to dry and b/b to wet. If in some generations the environment tends to be divided into patches which are either hot and wet or cold and dry, and in others, hot and dry or cold and wet, then one has the basis for the type of

shift in the direction of gene interaction necessary for the model to work. As MAYNARD SMITH comments, this type of selective situation is probably not widespread and, although some groups of genes in some parts of the chromosomes might be exposed to such selection, it seems unlikely that recombination would be maintained throughout the whole genome.

APPENDIX 1

Evolutionary stability of zero-recombination multi-locus systems in constant environments.

Consider an infinite, random-mating population segregating at several loci. There will be some number n of gamete types which can be produced by all possible combinations of the alleles represented in the population when it is at a stable equilibrium with no recombination between the selected loci. The actual number of allele combinations represented in the population at the equilibrium in question is $m \leqslant n$. Let the frequency of the ith gamete type represented in the population be $x_i(>0)$, and let the fitness of the genotype formed by the ith and jth gamete types present in the equilibrium population be w_{ij}. The marginal fitness of gamete i is thus $w_i = \sum_j x_j w_{ij}$, and the mean fitness of the population is $\bar{w} = \sum_i x_i w_i$. Since the population is in equilibrium with no recombination, $\bar{w} = w_i$ for all i corresponding to gametes present in the equilibrium population. It is convenient to use the subscript l for gamete types not represented in the equilibrium population. For the equilibrium to be stable, we must have $w_l = \sum_j x_j w_{jl} < \bar{w}$ for each l.

We now introduce into this population a modifier gene which induces recombination between the selected loci. If any gametes are not represented in the original population, they will be produced as a result of this recombination. Let the frequency of the modifier gene among gametes of type r (r taken over all possible gamete types) be γ_r. The total frequency of the modifier is $\gamma = \sum_r \gamma_r$. It will be assumed that γ is sufficiently small that γ^2 is negligible. Providing that the modifier is not recessive, it is easily seen that, neglecting $0(\gamma^2)$,

$$\bar{w}\Delta\gamma = \sum_r \gamma_r w_r - \bar{w}\gamma$$
$$= \sum_l \gamma_l (w_l - \bar{w}) \leqslant 0 \tag{A.1.1}$$

The equality signs holds only if all possible gamete types are represented in the initial equilibrium population. If this is the case, we conclude that a rare modifier can increase at most at a rate proportional to the square of its frequency, which implies that it will have a nearly zero chance of establishing itself in a large population. If some gametes are unrepresented in the initial population, the modifier will be eliminated by selection; in this case, zero recombination corresponds to what MAYNARD SMITH (1972) has called an evolutionary stable strategy. It should be noted that no assumption of frequency-independent genotypic fitnesses is involved in this proof, in contrast to LEWONTIN's (1971) mean fitness argument.

APPENDIX 2

Selection on a new inversion in a fluctuating environment

In this appendix an approximate analysis is made of the selection coefficient on a new inversion in a fluctuating environment. In a two-locus population without any inversions, we have in any generation (dropping the generation subscript k):

$$\bar{w} \Delta x_1 = x_1 (w_1 - \bar{w}) - RD$$

so that
$$ln \frac{w_1}{\bar{w}} = ln \left[1 + \frac{RD}{x_1\bar{w}} + \frac{\Delta x_1}{x_1} \right] \tag{A.2.1}$$

The average selection coefficient for a new inversion introduced into a gamete of type AB is defined here (using angle brackets to denote expectations over the stationary process) as

$$s_1 = \left< ln\, \frac{w_1}{\bar{w}} \right> = \left< \frac{RD}{x_1\bar{w}} \right> + \left< \frac{\Delta x_1}{x_1} \right> - \frac{1}{2} \left< \frac{R^2D^2}{x_1^2\bar{w}^2} + \frac{(\Delta x_1)^2}{x_1^2} + \frac{2\,R\,D\Delta x_1}{x_1^2\bar{w}} \right>$$
$$+\; \ldots$$

(A.2.2)

The terms in this expansion can be further analyzed as follows. Noting that $x_1 = p\,q + D$, where p and q are the frequencies of alleles A and B respectively, we have

$$\left< \frac{RD}{x_1\bar{w}} \right> \approx \frac{R}{\langle x_1 \rangle \langle \bar{w} \rangle} \left\{ \langle D \rangle \left[1 + \frac{V_{x_1}}{\langle x_1 \rangle^2} + \frac{V_{\bar{w}}}{\langle \bar{w} \rangle^2} + \frac{Cov(x_1,\bar{w})}{\langle x_1 \rangle \langle \bar{w} \rangle} \right] \right.$$
$$\left. - \frac{[V_D + Cov\,(D,pq)]}{\langle x_1 \rangle} - \frac{Cov\,(D,\bar{w})}{\langle \bar{w} \rangle} \right\} ,$$

where V denotes variance and Cov covariance.

Since terms such as $\langle D \rangle$, $V_{x_1}/\langle x_1 \rangle^2$ etc. are likely to be small, their cross-products may be ignored to a good approximation. In addition, if we let b_1 and b_2 be the linear regression coefficients of pq and $\bar{w}$ on D, we can write the corresponding covariance terms as $V_D b_1$ and $V_D b_2$. We therefore obtain:

$$\left< \frac{RD}{x_1\bar{w}} \right> \approx \frac{R}{\langle x_1 \rangle^2 \langle \bar{w} \rangle^2} \left\{ \langle D \rangle \langle x_1 \rangle \langle \bar{w} \rangle - V_D (\langle \bar{w} \rangle [1 + b_1] + \langle x_1 \rangle b_2) \right\}.$$

We also have:

$$\left< \frac{\Delta x_1}{x_1} \right> \approx \langle \Delta ln\, x_1 \rangle = 0 ,$$

neglecting terms in $(\Delta x_1)^2$, in accordance with our assumption that changes in gamete frequencies are slow.

Similarly,

$$\left< \frac{D^2}{x_1^2\,\bar{w}^2} \right> \approx \frac{\langle D \rangle^2 + V_D}{\langle x_1 \rangle^2 \langle \bar{w} \rangle^2} ,$$

$$\left< \frac{D\,\Delta x_1}{x_1^2\,\bar{w}} \right> \approx \frac{V_{\Delta x_1}\,(b_3 - \langle D \rangle b_4)}{\langle x_1 \rangle^2 \langle \bar{w} \rangle} ,$$

where b_3 and b_4 are the linear regression coefficients of D and $\bar{w}$ respectively on Δx_1. Provided that these are not many times greater than unity in absolute value, which would imply a rather strong relationship between the variables, this term may therefore be neglected as being of order $(\Delta x_1)^2$.

We therefore obtain the final expression for s_1:

$$s_1 \approx \frac{R}{\langle x_1 \rangle^2 \langle \bar{w} \rangle^2} \left\{ \langle D \rangle \left[\langle x_1 \rangle \langle \bar{w} \rangle + \frac{R}{2} \langle D \rangle \right] \right.$$
$$\left. - V_D \left[\langle \bar{w} \rangle (1 + b_1) + \langle x_1 \rangle b_2 + \frac{R}{2} \right] \right\}$$

(A.2.3)

Similar formulae can be obtained for the selection coefficients s_2, s_3 and s_4 for inversions introduced into the other gamete types, replacing p with $(1-p)$, q with $(1-q)$ and D with $-D$, as appropriate.

For processes in which $\langle D \rangle = 0$ and the allele frequencies do not change, as in the computations reported in the text, this equation can be simplified further to

$$s_1 \approx - \frac{R\,V_D}{\langle x_1 \rangle^2 \langle \bar{w} \rangle^2} \left\{ \langle \bar{w} \rangle + \langle x_1 \rangle\, b_2 + \frac{R}{2} \right\}$$

(A.2.4)

The following heuristic argument can be used to show that, for processes in which D is approximately normally distributed, a necessary condition for $s_i < 0$ ($i = 1$–4) will usually be that D changes sign occasionally. Assume arbitrarily that $<D> > 0$. For $s_1 < 0$, we have from equation (A.2.3)

$$<D>^2 \left[1 + \frac{R}{2<\bar{w}>} \right] < <x_1><D> + \frac{R<D>^2}{2<\bar{w}>}$$

$$< V_D \left[1 + b_1 + \frac{<x_1>}{<\bar{w}>} b_2 + \frac{R}{2<\bar{w}>} \right].$$

Unless at least one of b_1 and b_2 is substantially greater than unity, implying somewhat strong correlations between D and the gene frequencies or $\bar{w}$, $s_1 < 0$ will require V_D to be so large in relation to $<D>^2$ that D is negative a significant proportion of the time, if its distribution is approximately normal. A similar argument can be applied to the other gamete types to cope with the case with $<D> < 0$ (see text)

One final point is worth mentioning. The accuracy of neglecting terms in $(\Delta x_1)^2$ may be questioned, particularly when D and V_D are small. For example, if D were always zero, variations in gene frequency alone would generate a negative value of s_1 in equation (A.2.2) through the terms in $(\Delta x_1)^2/x_1^2$. Although this would formally give a selection coefficient against a new inversion, it merely reflects the vicissitudes of an element permanently associated with one gametic type which is fluctuating in frequency, and has nothing to do with selection for increased recombination. This is evident from the absence of R from the relevant terms. In the case with D always zero, a genic modifier has no effect on the frequencies of gametes at the selected loci, and so will eventually cease to change in frequency once any initial hitch-hiking effects have died out. From the point of view of selection for increased recombination, the $(\Delta x_1)^2$ terms may thus be legitimately ignored.

LITERATURE CITED

BODMER, W. F. and J. FELSENSTEIN, 1967 Linkage and selection: theoretical analysis of the deterministic two-locus random mating model. Genetics **57**: 237–265.

CARSON, H. L., 1955 The genetic characteristics of marginal populations of *Drosophila*. Cold Spring Harbor Symp. Quant. Biol. **20**: 276–287. ——, 1959 Genetic conditions which promote or retard the formation of species. Cold Spring Harbor Symp. Quant. Biol. **24**: 87–105. ——, 1965 Chromosomal morphism in geographically widespread species of Drosophila. *The genetics of colonizing species*. Academic Press, Inc., New York.

CHARLESWORTH, B. and D. CHARLESWORTH, 1973 Selection of new inversions in multi-locus genetic systems. Genetical Research **21**: 167–183.

DA CUNHA, A. B. and T. DOBZHANSKY, 1954 A further study of chromosomal polymorphism in *Drosophila willistoni* in relation to the environment. Evolution **8**: 119–134.

DEAKIN, M. A. B., 1972 A model for inversion polymorphism. J. Theoret. Biol. **35**: 191–212.

FELDMAN, M. W., 1972 Selection for linkage modification. 1. Random mating populations. Theoret. Popul. Biol. **3**: 324–346.

FISHER, R. A., 1930 *The genetical theory of natural selection*. Clarendon Press, Oxford.

GHISELIN, M. T., 1974 *The economy of nature and the evolution of sex*. University of California Press, Berkeley.

GILLESPIE, J. H., 1973 Polymorphism in random environments. Theoret. Popul. Biol. **4**: 193–195.

GRANT, V., 1958 The regulation of recombination in plants. Cold Spring Harbor Symp. Quant. Biol. **23**: 337–363.

HALDANE, J. B. S. and S. D. JAYAKAR, 1963 Polymorphism due to selection of varying direction. J. Genet. **58**: 237–242.

KARLIN, S., and U. LIEBERMAN, 1974 Random temporal variation in selection intensities: case of large population size. Theoret. Popul. Biol. **6**: 355–382.

KARLIN, S. and J. MCGREGOR, 1974 Towards a theory of the evolution of modifier genes. Theoret. Popul. Biol. **5**: 59–103.

LEWONTIN, R. C., 1971 The effect of genetic linkage on the mean fitness of a population. Proc. Natl. Acad. Sci. U.S. **68**: 984–986.

LEWONTIN, R. C. and K. KOJIMA, 1960 The evolutionary dynamics of complex polymorphisms. Evolution **14**: 458–472.

MAYNARD SMITH, J., 1971 What use is sex? J. Theoret. Biol. **30**: 319–335. ——, 1972 *On evolution.* Edinburgh University Press, Edinburgh.

NEI, M., 1967 Modification of linkage intensity by natural selection. Genetics **57**: 625–641. ——, 1969 Linkage modification and sex difference in recombination. Genetics **63**: 681–689.

PRAKASH, S., 1973 Patterns of gene variation in central and marginal populations of *Drosophila robusta.* Genetics **75**: 347–369.

STEBBINS, G. L., 1958 Longevity, habitat and release of genetic variability in the higher plants. Cold Spring Harbor Symp. Quant. Biol. **23**: 365–377.

STROBECK, C., J. MAYNARD SMITH and B. CHARLESWORTH, 1976 The effects of hitch-hiking on a gene for recombination. Genetics **82**: ......–.......

STURTEVANT, A. H. and K. MATHER, 1938 The interrelations of inversions, heterosis and recombination. Amer. Nat. **72**: 447–452.

TURNER, J. R. G., 1967 Why does the genotype not congeal? Evolution **21**: 645–656.

WILLIAMS, G. C., 1966 *Adaptation and natural selection.* Princeton University Press. Princeton.

Corresponding editor: R. C. LEWONTIN