SELECTION ON RECOMBINATION IN CLINES

D. CHARLESWORTH* AND B. CHARLESWORTH

Laboratory of Environmental Mutagenesis, National Institute of Environmental Health Sciences, P. O. Box 12233, Research Triangle Park, North Carolina 27709

Manuscript received June 23, 1978
Revised copy received September 29, 1978

ABSTRACT

Computer runs have been done to examine SLATKIN's (1975) model for selection on recombination rates in linear sets of populations with environmental changes affecting two loci. In order to determine whether the suggested selection pressures on recombination do, in fact, exist, we follow the changes in frequency at a third locus that is polymorphic for alleles affecting the recombination rate between the two selected loci. With haploid or diploid selection models, there can be selection for increased recombination if the parameter values are chosen suitably, but changes in parameter values often lead to changes in the direction of selection, so that decreased recombination is favored. The selection for increased recombination is usually weak, while that for decreased recombination is frequently much stronger. Weaker selection on the selected loci often leads to increasing selection for decreased recombination.

IT has recently been suggested that increased recombination can be selected for in the following situation. A population is subdivided, but there is migration between the sub-populations. There are two loci under selection, and each is assumed to respond to a different environmental feature. The environmental difference that affects selection at one locus (say A) is assumed to occur at a different place from that affecting the other locus (B). For example, in a linear array of populations, the environment might be such that alleles A_1 and B_1 are favored by selection at the left-hand end. After some point, allele A_2 is favored compared with A_1, but selection at the B locus is unchanged. Finally, after some other point, B_2 and A_2 are both favored. SLATKIN (1975) showed that in this system the effect of selection on the two loci depends on the linkage between them. If the loci are tightly linked, the position of the cline that is generated at one locus is shifted from its position when only that locus is segregating. There is linkage disequilibrium in and around the region between the two environmental change-over points. In such a situation, in which (at least in the intermediate populations) each locus is perturbed from its equilibrium by its association with the other locus, it seems plausible that there might be selection for increased recombination such as that operating in the model of STROBECK, MAYNARD SMITH and CHARLESWORTH (1976).

This suggestion was discussed by MAYNARD SMITH (1977), who expressed the model in the following terms. To the left of the first environmental discontinuity,

* Present address: School of Biological Sciences, University of Sussex, Brighton, Sussex BN1 9QG, England.

the gamete A_1B_1 is favored. At the other extreme, A_2B_2 is favored, while in the intermediate sub-populations A_2B_1 is favored. Now, if we consider the intermediate populations, it is obvious that the genotype A_1B_1/A_2B_2 can generate the favored gamete type A_2B_1 by recombination, and thus there might be selection for increased recombination. To give an intuitive idea of how the selection works, MAYNARD SMITH (1977) described a simplified version of the above model, with just three sub-populations, and with selection in the haploid phase such that *only* gamete A_1B_1 survives in environment 1, *only* A_2B_1 in environment 2, and *only* A_2B_2 in environment 3. It is easy to see that this model gives selection for recombination, since no progeny of A_1B_1/A_2B_2 other than recombinants can survive in environment 2.

Both the haploid and diploid models described above stress the fact that the A_1B_1/A_2B_2 genotype can generate A_2B_1 gametes by recombination, and that this may give selection in favor of recombination, but do not mention that A_1B_2 gametes will also be produced; this gamete is disadvantageous in all environments, and there may therefore also be an opposing selection pressure, tending to decrease recombination, acting at the same time. It is not obvious what the net outcome of these pressures would be in any given situation. We have therefore done computer calculations to examine these models, using a three-locus system, two loci being the selected loci (A and B), and a third locus (C) modifying the recombination fraction between A and B. The genes were assumed to be in the order A-B-C, with C either linked or unlinked to B.

THE COMPUTER MODEL

These runs involved purely deterministic computer calculations, of up to 101 sub-populations, with migration between them. The sequence of events in one generation was as follows. Starting with a set of gamete frequencies in each sub-population, migration occurred, generating a new set of gamete frequencies for each location. The probability of migration over one interval between populations was called the migration parameter, m. In the three population runs, and in some of the runs with more than three populations, migration was over at most one interval (*i.e.*, a stepping-stone model). With more than three sub-populations, more extensive migration was sometimes allowed; the probability of migration from each population to a population n places away was assumed to be m^n, for $n \leqq k$, where k was the upper limit of the distance migrated. (The end populations were assumed to send out fewer migrants than the central ones, since migration could occur in only one direction, for some distances at least, and we did not increase the migration from these populations to compensate for this.) Note that migration of gametes is assumed, and that this is not exactly equivalent to adult migration.

For each sub-population taken singly, the gamete arrays after migration were operated on by a sub-routine that performed the transition to the next generation, taking into account the appropriate diploid fitness values and the genetic process for three loci. This sub-routine had previously been tested thoroughly. When

haploid selection models were studied, the fitnesses of all diploid genotypes were assumed to be equal, and an additional sub-routine for haploid selection was applied after the new gamete frequencies had been generated.

This sequence of events was run with only the selected loci segregating for enough time (about 50–200 generations) to reach the equilibrium state. Then a modifier of recombination was introduced and its progress followed. In all the runs, the modifier was introduced at a low frequency in just one sub-population.

HAPLOID SELECTION MODELS

We assumed multiplicative effects of the A and B loci, with substitution of one allele by another changing the fitness at that locus by an amount s (Table 1). Table 2 shows some results of runs of an unlinked modifier increasing recombination; C_1C_1 zygotes were assumed to have no recombination, C_2C_2 to have a recombination fraction of 50%, and C_1C_2 of 25%. A stepping-stone model of migration was used. A frequency of 0.05 of C_2 was introduced into the central sub-population at generation 100, and its asymptotic rate of change in frequency, $\Delta p/p$, was estimated. As the changes were very slow, the rates per 20 generations are shown in the table. After initial perturbations lasting in most cases for less than 100 generations, the asymptotic rates of change were effectively reached. These were, of course, the same in all sub-populations and independent of the

TABLE 1

Haploid fitnesses in the different environments in the haploid selection model

Environment	A_1B_1	A_2B_1	A_1B_2	A_2B_2
1	1	$1-s$	$1-s$	$(1-s)^2$
2	$1-s$	1	$(1-s)^2$	$1-s$
3	$(1-s)^2$	$1-s$	$1-s$	1

TABLE 2

Rates of change in frequency per 20 generations of an unlinked modifier increasing recombination between loci A *and* B *from 0 to 50% (with intermediate dominance)*

s	Number of populations	m 0.01	0.1	0.2	0.3	0.5
1.0	3	0.00004	0.00454	0.0204	0.0517	>0.113
0.5	3	0.00002	0.00232	0.00786	0.01374	>0.0103
	7	—	0.00099	0.00349	0.00672	0.0109
0.3	3	~0	0.00107	0.00076	—	—0.0309
	7	—	0.00050	0.00083	—0.00059	—0.01775
0.2	3	~0	0.00014	—0.00358	—	—
	7	—	0.00014	—0.00105	—	—

The haploid selection scheme was used, and one central population was exposed to environment 2.

— indicates that the run with this parameter set was not done.

gamete into which the modifier was introduced, so that only one value will be given in what follows.

The effects of changes in m, the migration parameter, and the selection coefficient for the selected loci, s, are complex. When $s = 1$, selection for increased recombination was stronger, the higher m, and there was always selection for increased recombination, as would be expected. Decreasing s decreased the strength of selection for increased recombination, in all cases tested, and the rate of change in frequency of the modifier was often highest when the migration rate was given some intermediate value; the rate of change was sometimes negative in sign in these cases, indicating selection for zero recombination between the selected loci. The effect of changing the number of populations from three to seven was usually to weaken the selection, though a few parameter sets gave the opposite result. We shall return to a detailed consideration of these results in the DISCUSSION.

The changes in the modifier frequency in all these runs were very slow, except with $s = 1$. It was therefore not possible to run the runs to equilibrium. When the modifier was eliminated, no further runs were needed, but when it increased in frequency, one may ask whether there would still be an increase if the initial recombination fraction were not zero, but some higher value, or whether there is an intermediate equilibrium recombination fraction between zero and 0.5 that would be reached under selection. In all those cases of Table 2 in which a positive rate of change in frequency of the modifier was found, we found that a modifier increasing the recombination fraction from any initial value chosen increased in frequency, while a modifier decreasing recombination was always selected against. There is thus no intermediate equilibrium recombination fraction in any of the cases tested.

DIPLOID SELECTION

For diploid selection, we used the multiplicative scheme shown in Table 3. This was chosen because, if s is equal to 0.5, it is the diploid analogue, assuming intermediate dominance, of the haploid model used in the previous section, with $s = 1$. Tables 4 and 5 show some representative results with a modifier having the same effects as in the haploid case, and with a stepping-stone migration model. In general, selection for recombination is even weaker than with haploid selection, but similar effects of changing the parameters are seen.

We have run a number of variants of the above scheme. For example, making the modifier recessive instead of semi-dominant gave much weaker selection than that shown in Tables 4 and 5. A linked, semi-dominant modifier, however, changed in frequency considerably faster than an unlinked one, though in all cases studied the changes were in the same direction. Tables 4 and 5 give the results for an unlinked modifier, because this gave a much quicker approach to the asymptotic rate of increase. The highest rate of increase that we obtained with diploid selection was found with a modifier having a recombination fraction of 0.05 with the B locus (the closest linkage tried), with seven populations

arranged as in Table 5, and with $s = 0.2$ and $m = 0.5$; these parameter values gave a rate of increase of the modifier of 0.0044 per 20 generations.

DISCUSSION

The results presented above can be understood in a general, qualitative way in terms of two opposing forces, one giving selection for increased, and one for

TABLE 3

Diploid fitness scheme with intermediate dominance at both loci and multiplicative interactions between the loci

	A_1A_1	A_1A_2	A_2A_2
		Environment 1	
B_1B_1	1	$1-s$	$1-s$
B_1B_2	$1-s$	$(1-s)^2$	$(1-s)(1-2s)$
B_2B_2	$1-2s$	$(1-s)(1-2s)$	$(1-2s)^2$
		Environment 2	
B_1B_1	$1-2s$	$1-s$	1
B_1B_2	$(1-s)(1-2s)$	$(1-s)^2$	$1-s$
B_2B_2	$(1-2s)^2$	$(1-s)(1-2s)$	$1-2s$
		Environment 3	
B_1B_1	$(1-2s)^2$	$(1-s)(1-2s)$	$1-2s$
B_1B_2	$(1-s)(1-2s)$	$(1-s)^2$	$1-s$
B_2B_2	$1-2s$	$1-s$	1

TABLE 4

Rates of change in frequency per 20 generations of a modifier as in Table 2, except that the diploid selection scheme was used

		m			
s	Number of populations	0.1	0.2	0.3	0.5
0.5	3	0.00047	0.00115	0.00060	—0.01143
	5	0.00028	0.00073	0.00077	—0.00342
	7	—	0.00052	0.00054	0.00604
0.45	3	0.00048	0.00086	—0.00071	—0.01708
	5	0.00029	0.00063	0.00026	—0.00606
	7	0.00020	0.00044	0.00019	—0.00390
0.4	3	0.00045	0.00035	—0.00253	—
	5	0.00028	0.00043	—0.00052	—0.00936
	7	0.00020	0.00030	—0.00033	—
0.2	3	—0.00027	—0.00425	—	—
	5	—0.00007	—0.00201	—0.00755	—0.02874
	7	—	—0.00136	—0.00502	—0.01992
0.1	5	—0.00068	—0.00472	—0.00970	—
	7	—0.00046	—0.00340	—0.00722	—
0.02	5	—	—0.00049	—0.00037	—
	7	—	—0.00074	—0.00084	—

TABLE 5

As Table 4, except that there were three central sub-populations subjected to environment 2

s	Number of populations	m 0.1	0.2	0.3	0.5
0.5	5	0.00004	0.00005	0.00016	0.00060
	7	—	0.00003	0.00012	—0.02212
0.45	5	0.00007	0.00008	0.00026	0.00082
	7	—	0.00006	0.00019	0.00072
0.4	5	0.00009	0.00013	0.00038	0.00102
	7	0.00001	0.00009	0.00030	0.00100
	11	—	—	—	0.00061
0.2	5	0.00010	0.00043	0.00064	0.00036
	7	0.00008	0.00043	0.00090	0.00151
	11	~0	0.00027	0.00059	0.00117
0.1	5	0.00016	0.00020	0.00057	~0
	7	0.00016	0.00047	0.00055	0.00006
	11	—	0.00034	0.00051	0.00042
	15	—	0.00022	0.00037	0.00034
	25	—	0.00015	0.00022	—

decreased, recombination between the selected loci, as suggested in the introduction to this paper. With strong selection on the A and B loci (high s values), the component that selects for increased recombination usually predominates, but when s is lower, the selection is often weaker, or may even work against recombination. The pattern of the environmental change-over points is very important in this respect, as the comparison between Tables 4 and 5 shows.

The reasons for the effects seen can be interpreted, at least in part, in the following terms. The selection pressures on a modifier of the recombination fraction between loci A and B derive almost exclusively from the progeny of the A_1B_1/A_2B_2 double heterozygotes; we can ignore the repulsion heterozygotes, because the A_1B_2 gamete is always rare, being at a disadvantage in all three environments. The effect of recombination in the coupling double heterozygote is to generate A_2B_1 and A_1B_2 gametes, and of these A_2B_1 is favored in environment 2. Hence, if there is no region subject to environment 2, recombination will be selected against, as suggested by SLATKIN (1975) in the first part of his paper in which he studies the case when the environmental changeover points for both loci coincide. If, however, a number of sub-populations are subjected to environment 2, the outcome with respect to recombination modification will depend on the size of this region, in relation to the region of overlap between the two clines. It is, for example, obvious that if the clines do not overlap (*i.e.*, the A locus is fixed in the region occupied by the B locus cline, or *vice versa*), then recombination makes no difference. In order to understand what is happening in models of this sort, we must therefore consider the degree of overlap of the clines, in relation to the region occupied by environment 2.

With these general ideas in mind, we can make sense of the changes that are

produced by changing the s and m parameter values. Consider first the effects of changing s. Stronger selection on the A and B loci will, other things being equal, tend to narrow the clines, and will thus decrease their degree of overlap. Thus, in Table 4, when s is small, so that the clines overlap considerably, there is selection for reduced recombination, which is stronger with $s = 0.1$ than with $s = 0.2$, though it weakens again if s is made sufficiently small ($s = 0.02$, for example). When $s = 0.4$ and m is small, selection favors increased recombination, and the strength of selection increases with increasing s. When s becomes very high, however, we would expect the area of overlap to become so small that selection on a modifier of recombination would become ineffective. Table 4 shows a slight weakening of the selection for increased recombination, when $m = 0.1$, and Table 5 shows similar, but stronger, effects. As expected, this effect is seen most clearly when migration is limited; when m is large, the clines will overlap even when s is also large. Table 5 illustrates the effect of increasing the size of the region subjected to environment 2. Now, more of the critical region of overlap between the clines is exposed to environment 2, which explains the observation that there is a greater tendency for selection to favor increased recombination in this situation, compared with the situation in Table 4.

These ideas can be summarized by the generalization that large s and small m are unfavorable to selection on modifiers that either increase or decrease recombination, while small s and large m tend to give selection for reduced recombination, and selection for increased recombination will occur in some intermediate region. It is not easy, however, to put these generalizations in quantitative terms, or to give them predictive value. SLATKIN (1975) has attempted to do this by the use of "characteristic lengths." The characteristic length, l_c, for a system that generates a cline, is defined as $l/\sqrt{s}$, where l is the standard deviation of the migration distribution (SLATKIN 1973). For a single-locus cline, l_c gives an approximate measure of the width of the cline, so that one might expect that it could be used to assess the probable extent of overlap between clines set up at two loci. Of course, this would be most likely to work well with two unlinked loci, and might be quite unsatisfactory for linked genes. Nevertheless, SLATKIN (1975) suggests that a cline at one locus will be affected by linkage to another locus when the distance between the environmental change-over points affecting the two loci is of the same order of magnitude as l_c, or less than this. MAYNARD SMITH (1977) has even suggested that when the boundaries are separated by a distance of approximately l_c, there will be selection for increased recombination, but if they are too close together, selection for reduced recombination will predominate.

The runs that we have done permit us to examine the question of whether l_c predicts the outcome of selection on recombination. All the runs of Tables 4 and 5 involve a stepping-stone model of migration, so that we have $l_c^2 = m/s$. For the ranges of m and s values used, l_c ranges between 2.24 (for $m = 0.5$, $s = 0.1$) and 0.447 (for $m = 0.1$, $s = 0.5$). Clearly, this is not a large range. Nevertheless, in both Tables 4 and 5, changing s or m can affect the outcome of the runs in

highly important ways. The system may change from one that promotes recombination to one that selects for zero recombination, when one parameter is changed, as for example in Table 4 with $m = 0.2$ when s is changed from 0.4 to 0.2; this corresponds to a change in l_c from 0.707 to 1. Table 6 shows some further examples, with a wider range of l_c values, and with 15 populations so that edge effects are minimized. The table shows that selection on the modifier does not become negligible until the number of sub-populations exposed to environment 2 is considerably in excess of l_c. In many cases, selection changes from negative to positive in sign when the number of sub-populations exposed to environment 2 is of the order of l_c, so that l_c could be used as a crude guideline in determining conditions favorable for selection for increased recombination. In general, we can conclude that, other things being equal, some function of m/s is a determining factor, and that, as we saw above, decreasing s can lead to increased selection for reduced recombination.

In addition to the effects of the parameters m and s, which we have tried to interpret above, Tables 4 and 5 show the effect of changing the number of sub-populations flanking a fixed set of sub-populations. When the array of populations is very large, so that there are no edge effects due to the end populations, we would predict that the effect of additional flanking populations would be to give slower changes, since it would increase the number of populations in which recombination is essentially neutral. With small arrays of populations, as in Tables 4 and 5, adding flanking populations does indeed tend to decrease the observed rates of change of the modifier, but changes in the opposite direction are occasionally seen. These are presumably due to edge effects; they are seen most often when m is rather high, in accordance with this interpretation. There is thus no suggestion that larger arrays of populations would generate stronger selection for recombination. We can therefore conclude that the selection for increased recombination generated in this model is always weak, though as mentioned above it is stronger for linked modifiers than for the unlinked cases shown in the tables.

TABLE 6

Rates of change in frequency of a modifier as in Table 4, but with a variable number of central sub-populations exposed to environment 2

Number of central sub-populations	Maximum distance of migration (k) 1		3	
	m=0.3	m=0.5	m=0.3	m=0.5
1	−0.00329	−0.00593	−0.00903	−0.01927
3	0.00037	0.00034	−0.00088	−0.01101
5	0.00018	0.00042	0.00054	−0.00211
7	0.00005	0.00019	0.00041	−0.00001
9	0.00001	0.00007	0.00018	0.00005
l_c	1.73	2.24	3.00	5.12

There were 15 sub-populations, and s was equal to 0.1. k is defined in the section on the computer model.

We thank J. FELSENSTEIN and M. SLAKIN for helpful comments on this work.

LITERATURE CITED

MAYNARD SMITH, J., 1977 Why the genome does not congeal. Nature **268**: 693–696.

SLATKIN, M., 1973 Gene flow and selection in a cline. Genetics **75**:: 733–756. ——, 1975 Gene flow and selection in a two-locus system. Genetics **81**: 787–802.

STROBECK, C., J. MAYNARD SMITH and B. CHARLESWORTH, 1976 The effects of hitch-hiking on a gene for recombination. Genetics **82**: 547–558.

Corresponding editor: M. NEI