Inversion polymorphism in a two-locus genetic system

By BRIAN CHARLESWORTH

Department of Genetics, University of Liverpool,

Liverpool, L69 3BX

SUMMARY

The types of equilibria possible with an inversion in a two-locus system are considered, and their stability properties investigated. With complete suppression of crossing-over in inversion heterozygotes, there are three possible types of stable equilibria; which of these is reached by a new inversion depends on the fitness effects of the two loci concerned. With one of these equilibria, basically involving cumulative overdominance of the selected loci, the inverted and standard sequences are genetically homogeneous and differ with respect to both loci. With the other types of equilibrium, the standard sequence remains heterogeneous for one or both loci. It is shown that this situation may lead to variations in karyotypic fitnesses when the inversion is changing in frequency. It is also found that, with certain fitness relationships, two alternative stable equilibria may coexist; the final frequency reached by an inversion may therefore depend on the population's history.

The effects of double crossing-over in inversion heterozygotes were also investigated, and it was shown that the equilibria with double crossing-over are closely related to the corresponding equilibria without it, except that both sequences are more heterogeneous genetically. Within each sequence there is almost complete linkage equilibrium between the selected loci, although both are in linkage disequilibrium with the inversion itself. It was also found that, with double crossing-over, the population tends to remain for many thousands of generations in a state of quasi-equilibrium. In this state, the inversion tends not to return to its original frequency after a perturbation; also, it may remain for a long time relatively homogeneous genetically, especially when rare.

These results were compared with those from experiments and observations on inversion polymorphisms.

I. INTRODUCTION

Fisher (1930) suggested that natural selection favours an increase in the degree of linkage between two polymorphic loci maintained heterozygous by interactive selection. He did not develop any detailed model, but later work by many geneticists has largely substantiated this idea and clarified the conditions under which reduction in crossing-over will be favoured in multi-locus systems. The work of Feldman (1972) is especially noteworthy in this respect. He considered the fate of a genic modifier of crossing-over, introduced at a low frequency into a population at a stable

equilibrium with two loci heterozygous. Only modifiers reducing crossing-over can spread; there will be selection for such modifiers if, and only if, the original two-locus polymorphism is such that the gamete frequencies deviate from those expected on the basis of random combination of alleles (i.e. if there is linkage disequilibrium). A similar conclusion was reached by Charlesworth & Charlesworth (1973) for the case of a new inversion in a multi-locus system near equilibrium: the inversion will spread if introduced into a gamete whose marginal fitness is higher than the population mean fitness. This is possible only if there is linkage disequilibrium.

The conditions for the spread of a gene reducing crossing-over, or an inversion, are thus fairly clear; the final state attained by the population is not so clear. Feldman's work suggests that a dominant or semi-dominant modifier gene will usually spread through the population; the fate of an inversion or a gene which completely suppresses crossing-over when heterozygous will depend on the gene-content of the gamete in which it occurs initially, on the nature of the fitness inter-actions between the loci concerned, and on the genetic make-up of the population into which it is introduced.

The problem of the nature of the equilibrium attained by an inversion has been examined previously by Haldane (1957), Turner (1970) and Deakin (1972). Haldane assumed that the inversion would completely replace the standard sequence with the same gene content; Turner pointed out that this is not necessarily the case. Deakin made a more thorough-going analysis of the equilibrium states possible for an inversion in a two-locus genetic system, and showed that several different types of equilibria are possible. He did not, however, discuss the stability of these equilibria, and the relationship of this to the nature of the fitness interactions between the two loci involved. This is the problem with which I shall be concerned in this paper.

I will discuss the conditions for the existence and local stability of the various possible types of inversion polymorphism, and try to relate these to the conditions for the maintenance of the various types of two-locus polymorphisms in the absence of inversions. In this way, it is possible to build up a fairly complete picture of the fate of an inversion introduced into a two-locus system, especially with low amounts of recombination. I will pay particular attention to the symmetric fitness model studied by Bodmer & Felsenstein (1967) and Karlin & Feldman (1970), because of the resulting algebraic simplification. I will also consider the effects of double crossing-over on the equilibrium attained by an inversion. An infinitely large random mating population with discrete generations and constant genotypic fitnesses will be assumed.

2. THE BASIC MODEL

I shall be concerned with a genetic system consisting of two loci, each with two alleles $(A, a$ and $B, b)$. There will be four gametic types in such a system: AB, Ab, aB and ab. Let the frequencies of these in a given generation be x_1, x_2, x_3 and x_4 respectively. I will also consider an inversion introduced into an AB gamete, with which it remains associated indefinitely (I assume that the inversion completely suppresses crossing-over). The inversion-carrying gametes can be represented as BA,

with frequency y. The total frequency of gametes containing genes A and B is $z = x_1 + y$. The inversion is also assumed to have no direct effect on the fitness of its carriers, so that the fitnesses of BA genotypes are the same as for the corresponding AB genotypes. Let the relative fitness of the genotype formed from the gametes with frequencies x_i and x_j be w_{ij}. Note that $w_{ij} = w_{ji}$, and $w_{14} = w_{23}$ (in the absence of position effects on fitness). The marginal fitness for the gamete with frequency x_i is defined as

$$w_{i\cdot} = \sum_{j=1}^{4} x_j w_{ij} + y w_{1i}. \tag{1}$$

The mean fitness of the population, $\overline{w}$, is given by

$$\overline{w} = \sum_{i=1}^{4} x_i w_{i\cdot} + y w_{1\cdot}. \tag{2}$$

Let the recombination fraction for the pair of loci A and B be R. If we define a linkage disequilibrium parameter as $D = x_1 x_4 - x_2 x_3$, we obtain the following expressions for the gamete frequencies in the next generation, y', x_1', etc.:

$$\left.\begin{aligned} \overline{w} y' &= y w_{1\cdot}, \\ \overline{w} x_i' &= x_i w_{i\cdot} \pm R D w_{14}, \end{aligned}\right\} \tag{3}$$

where the sign of RDw_{14} is positive for $i = 2$ and 3, and otherwise negative.

These expressions follow directly by straightforward modification of the standard equations of two-locus systems (Lewontin & Kojima, 1960).

The equilibrium frequencies of the five types of gametes ($\hat{y}$, $\hat{x}_1$, etc.) must therefore satisfy the following equations:

$$\left.\begin{aligned} \hat{\overline{w}} \hat{y} &= \hat{y} \hat{w}_{1\cdot}, \\ \hat{\overline{w}} \hat{x}_i &= \hat{x}_i \hat{w}_{i\cdot} \pm R\hat{D} w_{14}. \end{aligned}\right\} \tag{4}$$

As discussed by Deakin (1972), it is easy to determine the nature of the possible equilibria with BA present ($1 \geqslant \hat{y} > 0$). It follows from the first of equations (4) that $\hat{\overline{w}} = \hat{w}_{1\cdot}$. This in turn implies that $\hat{D} = 0$. This situation can arise in five different ways:

(i) $\hat{x}_2 = \hat{x}_3 = \hat{x}_4 = 0$. This is an equilibrium with only BA, or with BA and AB, which can be symbolized as BA/AB.

(ii) $\hat{x}_1 = \hat{x}_2 = \hat{x}_3 = 0$. This equilibrium has only BA and ab present (BA/ab).

(iii) $\hat{x}_1 = \hat{x}_3 = 0$. Only BA, Ab and ab are present ($BA/Ab/ab$). There is also an analogous equilibrium of type $BA/aB/ab$.

(iv) $\hat{x}_3 = \hat{x}_4 = 0$. Only BA, AB and Ab are present ($BA/AB/Ab$). There is an analogous equilibrium $BA/AB/aB$.

(v) $\hat{x}_1, \hat{x}_2, \hat{x}_3, \hat{x}_4 > 0$. All gametic types are present ($BA/AB/Ab/aB/ab$).

Equilibria (i) and (iv) are peculiar in that the relative frequencies of BA and AB are indeterminate; the inversion is not strictly maintained by selection in these cases, and they are thus unlikely to be of any great biological importance. The possibility that a new inversion may reach one of these equilibria cannot be excluded *a priori*, however.

I shall now discuss the conditions for the existence and stability of these

equilibria, with special reference to the symmetric fitness model of Table 1, introduced by Bodmer & Felsenstein (1967), and whose properties have been extensively analysed by them and by Karlin & Feldman (1970).

Table 1. *The symmetric fitness model*

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

3. ANALYSIS OF THE EQUILIBRIA

(i) *The general fitness model with small R*

In addition to knowing the conditions for existence and local stability of the inversion equilibria just described, it is of great interest to know which of them can be reached from the various possible types of two-locus polymorphisms lacking an inversion. For small R, it is possible to obtain a fairly complete analysis of this problem, which I shall give below before turning to a consideration of the stability conditions with arbitrary R. The argument can be divided into four stages.

(1) Let us first consider the conditions for the existence and stability of the two-locus equilibria in which a new BA inversion will be selected. R will be assumed to be small. It was shown by Charlesworth & Charlesworth (1973) that a new BA inversion will spread (when introduced into a population at or near a stable two-locus polymorphism) if and only if the population is in linkage disequilibrium with AB and ab present in excess of their frequencies under random combination. For small R, the conditions for the existence and stability of such equilibria can be determined by considering the effects of small amounts of recombination on the possible stable equilibria which can exist with $R = 0$. Using the type of argument formalized by Karlin & McGregor (1972), all the stable equilibria with small R have a one-to-one correspondence with neighbouring $R = 0$ (multiple allele) equilibria: the conditions for existence and stability in terms of the w_{ij} are the same in both cases.

The following types of equilibrium with AB and ab in excess of random combination are possible with low values of R:

(*a*) An equilibrium with AB and ab predominating. This equilibrium is generated from the AB/ab equilibrium with $R = 0$, by allowing a small amount of recombination.

(*b*) A pair of equilibria generated from the $AB/Ab/ab$ and $AB/aB/ab$ equilibria with $R = 0$. Here, aB or Ab has a low frequency.

(*c*) An equilibrium generated from the $R = 0$ equilibrium with all four gametes present. If $D > 0$ in the $R = 0$ equilibrium, it will be positive for small R. This follows by continuity.

From standard multiple allele theory, it follows that a stable equilibrium of type (*a*) is mutually exclusive with stable equilibria of types (*b*) and (*c*); two stable equilibria of type (*b*) may coexist, but are incompatible with a stable type (*c*) equilibrium.

(2) Having established the nature of the possible types of two-locus equilibria which favour the spread of a new BA inversion, we may now consider the conditions for existence of equilibria with BA present. In equations (4), $\hat{D} = 0$. The frequencies $\hat{z}$, $\hat{x}_2$, $\hat{x}_3$ and $\hat{x}_4$ are thus identical to the frequencies $\hat{x}_1$, $\hat{x}_2$, etc., in the analogous two-locus equilibrium lacking BA, with $R = 0$ and the same set of fitnesses. I shall refer to this equilibrium as the 'associated' $R = 0$ two-locus equilibrium. The BA/AB equilibrium is thus associated with a population fixed for AB, the BA/ab equilibrium with an AB/ab equilibrium, and so on. It should be noted that an equilibrium with all four standard sequence gametes plus the inversion can exist if and only if the associated $R = 0$ two-locus equilibrium has $\hat{D} > 0$. This is because the inversion equilibrium requires $\hat{x}_1\hat{x}_4 - \hat{x}_2\hat{x}_3 = 0$, and $\hat{x}_1$ here is necessarily lower than in the $R = 0$ two-locus equilibrium.

(3) The stability of the inversion equilibria can, with small R, be determined by consideration of the properties of the associated $R = 0$ equilibria. The argument is as follows. I will first consider the equilibrium where Ab, aB and ab are each present, and then indicate how the other cases can be treated. The stability properties can be studied in the standard way, by considering the eigenvalues of the matrix of partial differential coefficients of the functions y', z', x_2', x_3' and x_4', with respect to the variables y, z, x_2, etc., evaluated at the equilibrium point. Let this matrix be $\mathbf{B}$. Let $\mathbf{b}$ be a column vector with elements b, $-b$, $-b$ and b ($b = R\hat{x}_4 w_{14}/\hat{\bar{w}}$), and $\mathbf{c}'$ be a row vector with elements $w_{11}/\hat{\bar{w}}, w_{12}/\hat{\bar{w}} \ldots w_{14}/\hat{\bar{w}}$. Then we have

$$\mathbf{B} = \begin{pmatrix} 1 & \mathbf{c}' \\ \mathbf{b} & \mathbf{A}^* \end{pmatrix}, \tag{5}$$

where the elements of the 4×4 matrix $\mathbf{A}^*$ are equal to those of the matrix $\mathbf{A}$ which describes the stability behaviour of the associated $R = 0$ two-locus equilibrium, plus terms of order R. Note that $a_{ii} = 1 + w_{ii}/\hat{\bar{w}}$ and $a_{ij} = w_{ij}/\hat{\bar{w}}$ ($i \neq j$).

The characteristic equation of $\mathbf{B}$ is thus

$$\det(\mathbf{B} - \lambda\mathbf{I}) = (1-\lambda)\det(\mathbf{A}^* - \lambda\mathbf{I}) - b(\det_1 + \det_2 - \det_3 - \det_4) = 0, \tag{6}$$

where

$$\det_1 = \begin{vmatrix} a_{11}-1 & a_{12} & a_{13} & a_{14} \\ a_{21}^* & a_{22}^*-\lambda & a_{23}^* & a_{24}^* \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix},$$

$$\det_2 = \begin{vmatrix} a_{11}-1 & a_{12} & a_{13} & a_{14} \\ a_{11}^*-\lambda & a_{12}^* & a_{13}^* & a_{14}^* \\ a_{31}^* & a_{32}^* & a_{33}^*-\lambda & a_{34}^* \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix}, \quad \text{etc.}$$

When $R = 0$, $\mathbf{B}$ has one eigenvalue equal to 1, and the others equal to those of $\mathbf{A}$. Clearly, if $\mathbf{A}$ has eigenvalues greater than one in absolute value, then (for small R) the same will hold for $\mathbf{A}^*$ and $\mathbf{B}$, so that the inversion equilibrium will be unstable. If $\mathbf{A}$ has all its eigenvalues less than 1, then for small R the system with the inversion will be stable provided that the differential coefficient of λ with respect to R is

negative at the point $R = 0$, $\lambda = 1$. This is always the case, as is demonstrated below.

Using the rule for the differentiation of an implicit function, equation (6) gives:

$$\frac{\partial \lambda}{\partial R} = -\frac{\partial \det(\mathbf{B} - \lambda \mathbf{I})}{\partial R} \bigg/ \frac{\partial \det(\mathbf{B} - \lambda \mathbf{I})}{\partial \lambda}.$$

We obtain (after some reduction):

$$\left(\frac{\partial \lambda}{\partial R}\right)_{\substack{R=0 \\ \lambda=1}} = -1. \tag{7}$$

This proves what is required: the stability properties of the inversion equilibrium with small R are the same as those of the associated $R = 0$ two-locus equilibrium.

When one or more of gametes Ab, aB and ab are absent at equilibrium, it is obvious that (when R is small) the population is stable against the introduction of the absent gametes only if the associated $R = 0$ two-locus equilibrium is also stable. If this is the case, the gamete or gametes concerned will eventually be eliminated from both classes of population, and it is thus sufficient to consider the stability properties of the system in their absence. This can be done by removing from matrix $\mathbf{B}$ the rows and columns corresponding to the absent gametes. The above argument is otherwise unchanged, and leads to the same conclusion.

(4) We have seen that, for small R, the inversion equilibria have the same existence and stability conditions as the associated $R = 0$ two-locus equilibria. Also (see point (1)), with small R certain of these $R = 0$ equilibria generate stable two-locus polymorphisms which favour the spread of a new BA inversion. From what has been said under the earlier points, those equilibria where the inversion frequency is stable under selection (equilibria of types (ii), (iii) or (v)) are associated with stable two-locus polymorphisms (equilibria (a), (b), and (c) respectively, of point (1)), favouring the spread of BA, whereas equilibria of types (i) and (iv) are not. This suggests strongly that, with small R, the only type of equilibrium which can be reached by a new BA inversion occurring in a population with a stable two-locus equilibrium is that which is associated in the above sense with the two-locus polymorphism concerned – a BA/ab polymorphism will be reached from a type (a) two-locus polymorphism, a $BA/Ab/ab$ polymorphism from a type (b) population, and $BA/AB/Ab/aB/ab$ from type (c). This conclusion is borne out by all the numerical examples which I have studied. The other types of inversion equilibria seem unlikely to be reached, except when genetic drift is effective in changing gamete frequencies.

I will now consider in turn each of the five possible equilibria with respect to the conditions for their existence and local stability properties with arbitrary R.

(ii) *Local stability behaviour of the inversion polymorphisms*
with arbitrary R

BA/AB equilibrium

It is obvious that the frequencies of BA and AB are not affected by selection, in the absence of the other types of gamete. It is possible, however, that a population may reach this type of equilibrium from nearby points, so that it is worth determining its stability to the introduction of Ab, aB and ab at small frequencies. If the linearized recurrence relations for the system are written out, it can be shown that the system is stable if and only if:

$$1 > \frac{w_{14}(1 - \hat{x}_1 R)}{w_{11}}, \quad \frac{w_{12}}{w_{11}}, \quad \frac{w_{13}}{w_{11}}. \tag{8}$$

It is easy to see from this expression that equilibria with the inversion present are stable to the introduction of Ab, aB and ab only when the equilibrium with AB only present is stable $(\hat{x}_1 = 1)$, and may be unstable (when $\hat{x}_1$ is low enough) even if this equilibrium is stable. As discussed above, with small R it is unlikely that this type of equilibrium will be reached by an inversion introduced into a population with a stable two-locus polymorphism but it is just conceivable that this could happen with large R. This type of equilibrium is therefore of little importance, and certainly cannot correspond to any of the observed cases of inversion polymorphisms stabilized by selection.

BA/ab equilibrium

This case has previously been discussed by Haldane (1957). It is obvious that, for this polymorphism to exist and be stable, homozygotes for AB and ab must have lower fitnesses than AB/ab, i.e. $w_{11}, w_{44} < w_{14}$. In the case of the symmetric fitness model, this means $\delta > 0$. The stability of this equilibrium, against the introduction of gametes AB, Ab and aB at low frequencies can be tested in the standard way. We find that the system is unstable if either

$$\hat{y}w_{12} + \hat{x}_4 w_{24} > \hat{\bar{w}},$$

or

$$\hat{y}w_{13} + \hat{x}_4 w_{34} > \hat{\bar{w}}. \tag{9}$$

If the converse inequalities hold, the system is stable, provided that $w_{11}, w_{44} < w_{14}$. These conditions are identical with those derived by Haldane (1957), who only considered the effects of the introduction of Ab or aB singly, however.

These conditions have been fully discussed by Haldane. With the symmetric fitness model, they reduce to $\gamma + \beta > \delta$, i.e. the effect on fitness of making both A and B homozygous is less than the sum of the effects of making each homozygous on its own.

The conditions $w_{11}, w_{44} < w_{14}$ and the converse of inequalities (9) correspond to the conditions for the equilibrium AB/ab to be stable with $R = 0$. As discussed earlier, this implies that for small R there will be a single stable two-locus equilibrium favouring the spread of a BA inversion introduced at a low frequency, and that such an inversion will end up at the BA/ab equilibrium. Numerical examples support this, and suggest that it may often be true with high values of R as well.

BA/Ab/ab equilibrium

The analysis of the conditions for the existence and local stability of this equilibrium is set out in Appendix (*a*). For the symmetric fitness model, the conclusions can be expressed in a relatively simple form, as follows: there will be a locally stable equilibrium of this sort if and only if

$$0 < \delta > \beta + \gamma, \quad \beta, \gamma < \alpha, \quad \beta[\delta - (\beta - \gamma)] < 0, \left.\right\}$$
$$\gamma[\delta - (\gamma - \beta)] < 0, \quad \alpha\delta < (\beta - \gamma)^2. \quad\left.\right\} \tag{10}$$

These conditions can be only satisfied when either β and γ are opposite in sign, or when α, β and γ are all negative. This implies that this type of equilibrium can be stable only when the double heterozygote is not the fittest possible genotype, i.e. when at least one of the genes shows underdominance when the other is heterozygous.

Returning to the general fitnesses case, it has already been shown that a stable equilibrium of this type is associated (for small R) with a stable two-locus polymorphism favouring the spread of a new BA inversion, and located near the $AB/Ab/ab$ equilibrium. The only other such two-locus polymorphism which is possible under these circumstances is one near the $AB/aB/ab$ equilibrium, which is associated with a $BA/aB/ab$ inversion equilibrium. This suggests that, with small R, this type of inversion equilibrium will be reached when BA gametes are introduced into a population at or near the associated two-locus equilibrium.

It should be noted that, since the conditions for the existence of a stable inversion polymorphism of this type are independent of R, such a stable inversion equilibrium may exist even when the associated two-locus polymorphism is unstable or non-existent because R is too large. An inversion can only arrive at this equilibrium in such a case from a population which is initially not in equilibrium.

BA/aB/ab equilibrium

This can be analysed in the same way as the $BA/Ab/ab$ case, with suitable changes in the relevant fitness parameters. With the symmetric fitness model, it is easy to see that the existence and stability criteria are the same as for the $BA/Ab/ab$ equilibrium, except that β and γ are interchanged in the equations for the equilibrium gamete frequencies. In this case, therefore, the inversion equilibrium and its associated two-locus polymorphism will co-exist with the corresponding $BA/Ab/ab$ equilibrium and its associated two-locus polymorphism. The fate of a BA gamete introduced into a population will therefore depend on which two-locus equilibrium the population is situated at.

Numerical examples of these types of equilibria were investigated. Using a symmetric fitness model with $\alpha = 0.5$, $\beta = 0.25$, $\gamma = -0.15$ and $\delta = 0.30$, it was found that the two-locus equilibria lacking the inversion were stable provided that $R < 0.021$. For $R = 0.020$, there was a selection coefficient of 0.001 (calculated by the method of Charlesworth & Charlesworth (1973)) for an inversion introduced into an AB gamete in the $AB/Ab/ab$ equilibrium population. There was rather stronger

selection (0·009) for an inversion in the $AB/aB/ab$ population, although the final frequency of the inversion was only 0·09 as compared with 0·56 in the former.

BA/AB/Ab and BA/AB/aB equilibria

Reasons have been given earlier for regarding these equilibria as being of little biological interest. As with the BA/AB type of equilibrium, it is possible that these equilibria could be reached from nearby points, so that it is worth discussing their stability to the introduction of the other gametic types. We need only consider the $BA/AB/Ab$ equilibrium in detail. Let $\hat{w}_{3\cdot} = \hat{z}w_{13} + \hat{x}_2 w_{14}$ be the marginal fitness of newly-introduced aB gametes, and $\hat{w}_{4\cdot} = \hat{z}w_{14} + \hat{x}_2 w_{24}$ be the same for ab. Following the method of analysis of Bodmer & Felsenstein (1967) for the stability of similar equilibria in two-locus systems, it is easy to show that the population will be stable to the introduction of Ab and ab if $\hat{w}_{3\cdot}, \hat{w}_{4\cdot} < \hat{\hat{w}}$, and unstable if $\hat{w}_{3\cdot}, \hat{w}_{4\cdot} > \hat{\hat{w}}$. If $\hat{w}_{3\cdot} > \hat{\hat{w}} > \hat{w}_{4\cdot}$ or vice-versa, the equilibrium will be unstable only if

$$R[\hat{\hat{w}} - (\hat{x}_1 \hat{w}_{3\cdot} + \hat{x}_2 \hat{w}_{4\cdot})] < (\hat{w}_{3\cdot} - \hat{\hat{w}})(\hat{\hat{w}} - \hat{w}_{4\cdot}). \tag{11}$$

This may be compared with the corresponding expression for the equilibrium AB/Ab, where $\hat{x}_1$ is necessarily larger than when BA is present, but the other quantities have the same values. It is therefore possible for the equilibrium AB/Ab to be unstable while the $BA/AB/Ab$ equilibrium is stable, for $\hat{x}_1$ small enough and R large enough. If the two-locus equilibrium is stable, then the inversion equilibrium will be stable also.

In the case $\hat{w}_{3\cdot}, \hat{w}_{4\cdot} < \hat{\hat{w}}$, AB/Ab is stable; in the converse case, it is unstable.

These results imply that an inversion may be attracted to a $BA/AB/Ab$ type equilibrium if random events, or the population's history, have carried gamete frequencies to a suitable point in its neighbourhood.

BA/AB/Ab/aB/ab equilibrium

For small R, the arguments developed previously demonstrate that this equilibrium will exist and be stable under the same fitness conditions as must be satisfied by the 4-allele system made up of alleles corresponding to $BA + AB$, Ab, aB and ab. If these conditions are satisfied, there will be an associated two-locus stable equilibrium, with linkage disequilibrium in favour of AB.

For large R, the existence of this inversion equilibrium is unaffected, but the stability properties may be. Unfortunately, I have been unable to obtain a complete stability analysis. A useful necessary condition can, however, be derived for the symmetric fitness case. I shall now consider this case in detail.

The equilibrium frequencies of the gametes follow from the two-locus equilibrium frequencies with $R = 0$. With the symmetric fitness model, the gene frequencies of A and B must both be $\frac{1}{2}$, and the gamete frequencies such that $\hat{z} = \hat{x}_4$ and $\hat{x}_2 = \hat{x}_3$. Equation (7) of Bodmer & Felsenstein (1967) shows that the equilibrium value of $\xi = zx_4 - x_2 x_3$ is given by the expression

$$64l\xi^3 - 16m\xi^2 - 4l\xi + m + 0, \tag{12}$$

where $l = 2(\beta + \gamma) - (\alpha + \delta)$ and $m = \delta - \alpha$.

Equation (12) has solutions $\xi = \pm \frac{1}{4}$ and (provided $l \neq 0$) $\xi = m/4l$. We are only interested in the second type of equilibrium, since the $\xi = \frac{1}{4}$ case corresponds to the BA/ab equilibrium already considered, and $\xi = -\frac{1}{4}$ corresponds to $\hat{z} = \hat{x}_4 = 0$. Noting that $\hat{z} = \hat{x}_4 = \frac{1}{4} + \xi$, we can use the relation $\hat{x}_1 \hat{x}_4 = \hat{x}_2 \hat{x}_3$ to obtain the equilibrium frequencies:

$$\left.\begin{array}{r} \hat{y} = m/(l+m), \\ \hat{x}_1 = (l-m)^2/4l(l+m), \\ \hat{x}_2 = \hat{x}_3 = (l-m)/4l, \\ \hat{x}_4 = (l+m)/4l. \end{array}\right\} \tag{13}$$

It can be shown that the only possible sets of selection coefficients which are compatible with these relations, and with $\hat{y} > 0$, are those with $\alpha < \delta < \beta + \gamma$ or $\alpha > \delta > \beta + \gamma$.

The stability of this type of equilibrium can, in principle, be tested in the standard way by determining the eigenvalues of the linearized system. Unfortunately, this leads to a quartic equation. A necessary condition for stability for arbitrary R can be obtained by considering perturbations which conserve the symmetry of the frequencies z, x_1, etc., i.e. which leave gene frequencies unchanged. The analysis of this case is given in Appendix (b). The conclusion reached is that the equilibrium can be stable only if $\beta + \gamma < \delta$: the system is unstable to perturbations which conserve gene frequencies if $\beta + \gamma > \delta$.

This type of stable equilibrium cannot, therefore, co-exist with a stable BA/ab equilibrium, which requires $\beta + \gamma > \delta$, regardless of the value of R. For small R, it has been shown previously that stable $BA/Ab/ab$ and $BA/aB/ab$ equilibria are incompatible with a stable equilibrium of the present type; for large R, it is possible that they may be compatible. I have, however, been unable to find a numerical example of this. It is easy to find numerical examples in which the associated two-locus equilibrium is stable for all R, and where a new BA inversion increases towards its expected equilibrium frequency.

4. THE EFFECTS OF DOUBLE CROSSING-OVER

(i) *General considerations*

Up to now, I have assumed that the inversion completely suppresses crossing-over when heterozygous, so that the gametes carrying it are always genetically AB. In practice, all the products of two-strand double crossing-over and one of the cross-over strands produced by three-strand double crossing-over are recoverable from inversion heterozygotes, so that it is possible for genetic material to be exchanged between the inverted (In) and standard (ST) sequences. This means that the In gametes will gradually acquire representatives of all the gametic types in the system and, as will be seen, linkage equilibrium between loci A and B within both sequences tends to develop. This considerably complicates the analysis of the population genetics of an inversion. But, as I will show below, it is possible to achieve at least a qualitative understanding of what happens with double crossing-over in terms of the equilibria attained in its absence. I will only consider the equilibria where the inversion is actively maintained by selection (cases (ii), (iii) and (v) above).

In the first place, it is clear from genetic studies of crossing-over in inversion heterozygotes (Sturtevant, 1926; Sturtevant & Beadle, 1936; Philip *et al.* 1944; Levine, 1956) that double cross-overs within inversions of the size normally encountered in natural populations of *Drosophila* are extremely rare events. For example, Philip *et al.* report a frequency of 1 in 8000 for a long inversion of the *X* chromosome of *Drosophila subobscura*. Since there is no crossing-over in males of many *Drosophila* species, the effective rate of double crossing-over as far as population dynamics is concerned will be half this amount (Charlesworth & Charlesworth, 1973).

Since the production by mutation of a new inversion is a unique event, gamete types *Ab*, *aB* and *ab* can enter *In* gametes only as a result of double crossing-over in heterozygotes for *BA* and *ST* (ignoring mutation). Because of the extreme rarity of double cross-overs, as just discussed, the initial frequencies of *bA*, etc., will be an order of magnitude lower than that of *BA*. It will therefore take much longer for these to increase in frequency under selection than for *BA*, so that the population will move into the neighbourhood of the equilibrium which it would attain without double crossing-over, before they can reach appreciable frequencies.

This implies that the effect of double crossing-over on inversion polymorphisms can be understood by considering the perturbations produced by introducing *bA*, etc., into the equilibrium with only *BA* represented in *In* gametes, ignoring the very small effect of the process of double crossing-over itself. Write η_2, η_3 and η_4 for the frequencies of *bA*, *Ba* and *ba* respectively. Neglecting second-order terms, we get:

$$\left.\begin{aligned} \hat{\bar{w}}\eta_2' &= \eta_2\hat{w}_{2.} + \eta_4 R\hat{y}w_{14}, \\ \hat{\bar{w}}\eta_3' &= \eta_3\hat{w}_{3.} + \eta_4 R\hat{y}w_{14}, \\ \hat{\bar{w}}\eta_4' &= \eta_4(\hat{w}_{4.} - R\hat{y}w_{14}). \end{aligned}\right\} \tag{14}$$

These equations imply that the fate of the new gametic types depends on the type of equilibrium into which they are introduced.

(*a*) *BA/ab equilibrium.* For the original equilibrium to be stable, it was found that we need $\hat{w}_{2.}, \hat{w}_{3.} < \hat{\bar{w}} = \hat{w}_{4.}$. All three types of gamete will therefore tend to be eliminated under the influence of selection or of crossing-over in homokaryotypes. With double crossing-over producing them at a low rate, we can expect there to be an equilibrium with *In* overwhelmingly *BA* and *ST* mostly *ab*, with a sprinkling of the other types of gamete in each sequence.

(*b*) *BA/Ab/ab equilibrium.* Analysis of equations (14) shows that in this case η_3 and η_4 tend to decrease under the influence of selection, and crossing-over in homokaryotypes, whereas *bA* becomes neutral with respect to first-order terms. The results of Appendix (*a*) show that *AB* and *aB* are selected against in the *ST* gametes, if the original equilibrium is stable. With double crossing-over, *bA* can increase in frequency at a rate of the order of the amount of double crossing-over, while *Ba* and *ba* will tend to low equilibrium frequencies, as will *AB* and *aB*. Now Deakin (1972) showed that, in the absence of double crossing-over but with more than one gamete type in *In*, there is a neutral curve of equilibria. This curve is such that the frequencies of $AB+BA$, $Ab+bA$, etc., are equal to their equilibrium values with

only BA present in In, and that there is linkage equilibrium within both the In and ST sets of gametes (in this case because In and ST segregate only at a single locus). Furthermore, with small R it is easy to see from the recurrence relations derived by Deakin that the population will return to a point on the neutral curve after a perturbation from it, if the original equilibrium with only BA in In is stable. It follows that (certainly with small R, and probably with arbitrary R) the gamete frequencies with double crossing-over will move slowly along a path close to the appropriate curve of equilibria, until the various forces come into balance. In the present case, we have seen that AB is kept at a low frequency, so that BA will be present at equilibrium in a relatively high frequency, to give the required total frequency of $AB + BA$. Ab and bA will be present in relatively high frequencies; aB and Ba will have low frequencies; ab will have a high frequency, but ba a low one.

(c) *$BA/AB/Ab/aB/ab$ equilibrium*. In this case, equations (14) show that bA and Ba are neutral as far as first-order terms are concerned, while ba tends to be eliminated. With double crossing-over, therefore, we may expect a low equilibrium frequency of ba, and a relatively high frequency of ab. In order for linkage equilibrium to be maintained within the In and ST gametes, there must be a high frequency of BA, and a low frequency of AB at equilibrium. A fuller analysis of this case for the symmetric fitness model, using the equations developed below, is given in Appendix (c).

(ii) *Numerical results*

These conclusions can be compared with the results of direct calculations of population trajectories with double crossing-over. Such trajectories are calculated as follows. Double cross-overs will produce different results depending on where they fall inside the inversion with respect to the locations of A and B. The inversion is assumed to be paracentric, with loci A and B included in it, A being to the left of B. Call the stretch of chromosome between the left breakpoint and A, region 1; that between A and B, region 2, and that between B and the right breakpoint, region 3.

Cross-overs in regions 1 and 2 result in the exchange of alleles at locus A between ST and In; cross-overs in regions 2 and 3 result in the exchange of locus B alleles; cross-overs in 1 and 3 result in the exchange of alleles at both loci. Let the probabilities of these three classes of events be R_1^*, R_2^* and R_3^* respectively. Let y_i ($i = 1$ to 4) represent the frequencies of BA, bA, etc. Define linkage disequilibrium parameters

$$\left.\begin{aligned} D_x &= x_1 x_4 - x_2 x_3, \\ D_y &= y_1 y_4 - y_2 y_3, \end{aligned}\right\} \tag{15}$$

and marginal fitnesses

$$\left.\begin{aligned} w_{ix} &= \sum_j x_j w_{ij}, \\ w_{iy} &= \sum_j y_j w_{ij}, \\ w_{i.} &= w_{ix} + w_{iy}. \end{aligned}\right\} \tag{16}$$

The new frequency of the ith type of ST gamete is given by the expression

$$\overline{w}x'_i = x_i w_i \pm w_{14} R D_x + \sum_{j=1}^{3} D^*_{ij} R^*_j \qquad (17)$$

where the sign of $w_{14} R D_x$ is positive for $i = 2$ and 3 and otherwise negative; the D^*_{ij} are given as the matrix of coefficients in Table 2.

To calculate the new frequencies of the In gametes, x_i and y_i are simply interchanged in equation (17) and in the expressions in Table 2.

Table 2. *Coefficients D^*_{ij} for calculating the effect of double crossing-over on gamete frequencies*

3 (12)–1 (34)	2 (13)–1 (24)	1
4 (12)–2 (34)	1 (24)–2 (13)	2
1 (34)–3 (12)	4 (13)–3 (24)	3
2 (34)–4 (12)	3 (24)–4 (13)	4

The terms of form $i(jk)$ refer to $x_i(y_j w_{ij} + y_k w_{ik})$; the terms of form i refer to $y_i w_{ix} - x_i w_{iy}$.

Computer calculations based on these equations were used to determine the trajectory of a population with a BA inversion introduced with an initial frequency of 0·005 into a two-locus system at one of the stable equilibria described earlier. The results were found to be in broad agreement with the theoretical predictions made above. With the BA/ab type of equilibrium, equilibrium is reached fairly fast and, even with a high rate of double crossing-over, selection is effective in keeping the double cross-over gametes down to very low frequencies. With the $BA/aB/ab$ type of equilibrium, the fitness values used earlier (p. 266), with $R = 0·02$, were used in calculations. Populations with the R^*_i all equal to 10^{-4} or 10^{-5} were run. In both cases, it was found that after 25 000 generations bA and ba are still present at very low frequencies. The population is evidently in a state of quasi-equilibrium, and is moving slowly (changing only in the 5th decimal place) along a path near the curve with $D_x = D_y = 0$, as predicted. With double crossing-over at a total rate of 3×10^{-4} the frequency of A in In is reduced to about 0·5 by generation 25 000, whereas B is unaffected. After a further 25 000 generations, the equilibrium was reached, with only slightly different frequencies from those at generation 25 000. Among the In gametes, BA, Ba and ba predominate, with frequencies of 0·09, 0·13 and 0·01 respectively, whereas among ST gametes we have chiefly aB and ab, with frequencies of 0·23 and 0·54. As would be expected, the effect of double crossing-over is less marked with the lower rate and the system is nowhere near equilibrium after 25 000 generations, although changing very slowly. With this type of system, In is more heterozygous than ST for A, but less so for B.

Part (a) of Table 3 shows the results at generation 25 000 for a symmetric fitness model which generates a $BA/AB/Ab/aB/ab$ type of equilibrium in the absence of double crossing-over, with a high inversion frequency (0·43). With the higher rate of double crossing-over, the system reaches equilibrium after about 5000 generations; it is virtually at equilibrium at 25 000 generations with the lower rate. It will

18-2

be seen that the conclusions of Appendix (*c*) are confirmed by this example: BA and ab are the predominant gametes, and the frequencies $\hat{x}_2$, $\hat{x}_3$, $\hat{y}_2$ and $\hat{y}_3$ are all equal as a result of the fact that $R_1^* = R_2^*$. In runs with $R_1^* \neq R_2^*$, it was found that these frequencies become modified in the way predicted by Appendix (*c*): with $R_1^* > R_2^*$, we find $\hat{x}_3 = \hat{y}_2 > \hat{x}_2 = \hat{y}_3$, and with $R_1^* < R_3^*$, the opposite inequality. As can be seen from the table, at equilibrium in this case both In and ST are moderately heterozygous at both loci, and there is almost complete linkage equilibrium within gene arrangements ($D_x = D_y \approx 0$).

Table 3. *Equilibrium and quasi-equilibrium with double crossing-over*

(*a*) $\alpha = 0.70$, $\beta = 0.25$, $\gamma = 0.10$, $\delta = 0.40$ (high frequency inversion)

	Gamete frequencies				Gene frequencies		Heterozygosities*	
	AB	Ab	aB	ab	A	B	A	B
			$R = 0.10$, $R_i^* = 0.0001$ ($i = 1, 2, 3$)					
In	0.431	0.033	0.033	0.003	0.928	0.928	0.134	0.134
ST	0.003	0.033	0.033	0.431	0.072	0.072	0.134	0.134
			$D_x = D_y = 0.0001$					
			$R = 0.10$, $R_i^* = 0.00001$ ($i = 1, 2, 3$)					
In	0.435	0.031	0.030	0.002	0.936	0.934	0.120	0.123
ST	0.002	0.032	0.033	0.435	0.068	0.070	0.127	0.130
			$D_x = 0.0001$, $D_y = 0.0002$					

(*b*) $\alpha = 0.71$, $\beta = 0.40$, $\gamma = 0.21$, $\delta = 0.70$ (low frequency inversion)

	Gamete frequencies				Gene frequencies		Heterozygosities	
	AB	Ab	aB	ab	A	B	A	B
			$R = 0.10$, $R_i^* = 0.0001$ ($i = 1, 2, 3$)					
In	0.112	0.053	0.051	0.025	0.685	0.676	0.432	0.438
ST	0.149	0.186	0.188	0.236	0.441	0.444	0.493	0.494
			$D_x = 0.0001$, $D_y = 0.0002$					
			$R = 0.10$, $R_i^* = 0.0001$ ($i = 1, 2, 3$)					
In	0.066	0.011	0.010	0.002	0.865	0.854	0.234	0.250
ST	0.197	0.226	0.227	0.261	0.464	0.465	0.497	0.498
			$D_x = 0.0000$, $D_y = 0.0001$					

* The heterozygosities are calculated by treating the In and ST gametes as if drawn from two separate populations in Hardy–Weinberg equilibrium.

Part (*b*) of Table 3 shows the results at generation 25 000 for a fitness matrix which generates the same type of equilibrium, but with a low (0·05) inversion frequency, in the absence of double crossing-over. As might be expected, double crossing-over takes longer to produce an effect in this case, and the system is far from equilibrium after 25 000 generations. With the low rate of double crossing-over, it can be seen that at this point In is much less heterozygous for both loci than ST. This picture is not altered much after a further 25 000 generations, although the frequency of In increases from 0·09 to 0·12 as the double cross-over gametes accumulate.

The results of this section show that the nature of inversion equilibria with double crossing-over can be fairly well understood in terms of the corresponding equilibria without. A similar type of analysis can, of course, be applied to the case where either one or both loci are situated outside the inversion, but close enough for crossing-over between them and the inversion to be greatly reduced. If A is located outside the inversion, and B inside, then the only events to occur with non-negligible frequency are the exchange of alleles at locus A between In and ST and the exchange of B, i.e. $R_1^* > R_2^* \gg R_3^*$. If both A and B are located to the left of the inversion, then the commonest events will be the exchange of A and the exchange of both A and B, i.e. $R_1^*, R_3^* \gg R_2^*$.

DISCUSSION

The results derived in the previous parts of this paper have a direct bearing on some of the results obtained from experimental studies of inversion polymorphism. (It should, however, be stressed that two-locus theory is without doubt grossly inadequate as a model of the genetic systems underlying inversion polymorphism, and can be at most only a guide to the possible sorts of behaviour of real systems.)

In the first place, it has been shown that there are systems of fitness interactions which can generate more than one stable equilibrium with the inversion present (see the discussion of the $BA/Ab/ab$ and $BA/aB/ab$ equilibria); which final frequency the inversion arrives at is determined by the initial conditions of the population. This finding is of interest in the light of experimental results such as those of Dobzhansky & Pavlovsky (1957), which demonstrate the existence of alternative stable equilibrium inversion frequencies in the same population. The present results imply that interactions between genes contained in the inversion and the genetic background outside it need not be involved in this phenomenon.

The fact that the standard sequence may be genetically heterogeneous at equilibrium even without double crossing-over in inversion heterozygotes, and that both the inversion and the standard sequences may be heterogeneous when double crossing-over occurs, suggests that if the frequency of the inversion is perturbed experimentally from equilibrium, there will be consequent shifts in the frequencies of the various gamete types as the system returns to equilibrium. The fitnesses of the inversion and standard homokaryotypes (relative to that of the heterokaryotype), which are weighted averages of the appropriate genotypic fitnesses, may therefore change as the gamete frequencies move back to their equilibrium levels. A numerical example of this (for a $BA/AB/Ab/aB/ab$ equilibrium without double crossing-over) is shown in Table 4. It can be seen that there are, in fact, quite substantial changes in karyotypic fitnesses with changes in inversion frequency. The fitness of the inversion homokaryotype is highest when the inversion is below its equilibrium frequency, and lowest when it is much higher. The fitness of the standard homokaryotype, on the other hand, seems to have a maximum below the equilibrium point, and to fall off on either side. Changes in karyotypic fitnesses in cage experiments with *Drosophila* populations have been reported several times

(Dobzhansky & Levene, 1951; Kojima & Tobari, 1969; Watanabe *et al.* 1970), and it is possible that this type of mechanism may be involved in some of these cases.

When double crossing-over is non-negligible, perturbation experiments may give peculiar results which are superficially similar to the effects discussed in the first point. Because of the existence of neutral curves of equilibria, with fitness matrices other than those which generate BA/ab type equilibria double crossing-over results in quasi-equilibria, with gamete frequencies changing slowly along the equilibrium curve. If the system is perturbed from the curve, for example by experimentally changing the frequencies of In and ST without altering the relative frequencies of

Table 4. *Variation in fitnesses of the homokaryotypes as a result of genetic heterogeneity of the standard sequence*

$$\alpha = 0.85,\ \beta = 0.25,\ \gamma = 0.10,\ \delta = 0.65$$
$$\hat{y} = 0.200,\ \hat{x}_1 = 0.113,\ \hat{x}_2 = \hat{x}_3 = 0.187,\ \hat{x}_4 = 0.313$$

		Fitnesses	
Gen.	Inversion frequency	Inversion homokaryotype	Standard homokaryotype
(a) Initial inversion frequency 0.85			
5	0.476	0.383	0.694
25	0.294	0.398	0.763
45	0.259	0.407	0.800
65	0.239	0.412	0.816
85	0.227	0.415	0.829
105	0.219	0.418	0.838

The final homokaryotypic fitnesses are 0.423 and 0.856 for the inversion and standard sequences respectively.

		(b) Initial inversion frequency 0.05	
5	0.069	0.442	0.912
25	0.094	0.450	0.933
45	0.110	0.447	0.923
65	0.125	0.443	0.913
85	0.139	0.439	0.904
105	0.151	0.436	0.895

the various gamete types within each gene arrangement, it seems likely that the population will merely return to the nearest point on the equilibrium curve, and not to the original inversion frequency. (In any case, it can only return to the original state when the population was initially in true equilibrium rather than quasi-equilibrium.) Within the time ordinarily available for experimental studies, the same population perturbed to different inversion frequencies may apparently stabilize at different 'equilibrium' frequencies. This type of behaviour may theoretically be distinguished from that discussed under the first point by two criteria. First, in this case there is probably a continuum of 'equilibria' rather than a number of discrete stable points. Secondly, no new gamete types need be introduced (for example, by using a population of mixed geographic origin), in order to shift the population to a new equilibrium, but this will often be the case with the earlier type of effect.

Two examples of this type of behaviour are shown in Fig. 1. The upper graph shows the behaviour of the equilibrium population of Table 3 (*a*) (with the higher level of double crossing-over), when the inversion is increased to a frequency of 0·8 or decreased to 0·2 from its equilibrium level of 0·5. It will be seen that, within the sort of time-interval usual in experiments on artificial populations of *Drosophila*, the inversion appears to stabilize at two different frequencies according to its starting frequency. Only very gradually does it approach its true equilibrium level.
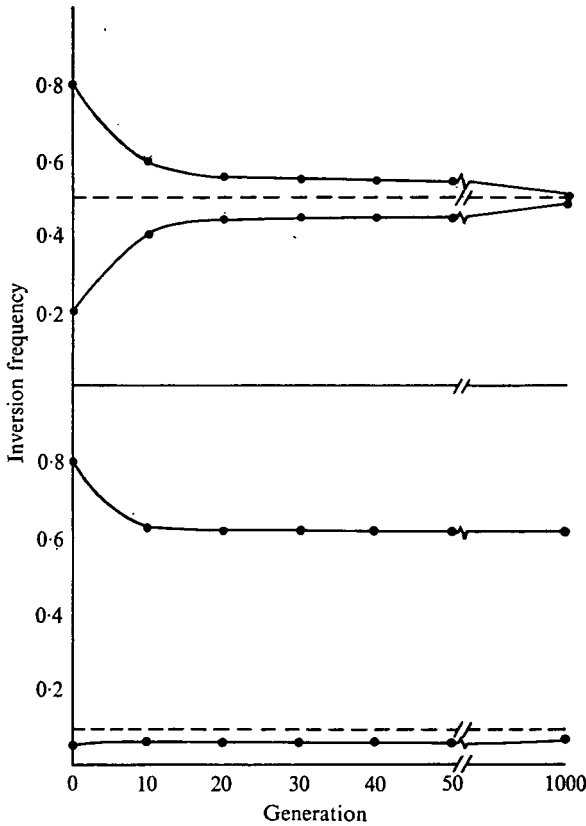

Fig. 1. The results of pertubation of an inversion, subject to double crossing-over when heterozygous, from its equilibrium or quasi-equilibrium point (indicated by the dashed line). Further details in text.

More striking behaviour is exhibited by systems which have not reached equilibrium under double crossing-over initially. This is illustrated by the lower graph of Fig. 1, where the quasi-equilibrium population of Table 3 (*b*) (with the lower level of double crossing-over) is perturbed from an initial inversion frequency of 0·09 to 0·8 and 0·05. In the first case, the inversion appears to stabilize at 0·62; in the second, it barely moves from the frequency to which it was perturbed.

If this type of fitness interaction is at all common, therefore, we would expect to find that experimental populations of *Drosophila* started with different initial

inversion frequencies might reach different, apparently stable frequencies. The most thorough investigation known to me which sheds light on this point is that of Watanabe *et al.* (1970) on *Drosophila pseudoobscura*. Populations were started with different initial frequencies of the gene arrangements *ST*, *AR*, *CH* and *PP*. The population trajectories indicated that at most only *ST* and *AR* would remain segregating. One population became fixed for *AR*. Among the others, three were started with a low (approx. 0·05) frequency and had final frequencies, predicted from the population's trajectory, of 0·40, 0·29 and 0·36. Two were started with a high frequency (approx. 0·50) of *AR*, and had predicted final frequencies of 0·50 and 0·60. Two which were started with an intermediate frequency (0·25) had predicted final frequencies of 0·42 and 0·35. While it is perfectly possible that these results are due to statistical fluctuations, the differences between the three classes of population are in the direction predicted on this theory. It would clearly be of great interest to examine this question on a larger scale, particularly using low-frequency inversions which are more likely to be in quasi-equilibrium only, since a certain amount can be inferred about the nature of the fitness interactions among the genes involved in the inversion polymorphism from this type of behaviour. If the inversion is maintained as a result of selective differences which generate a cumulative advantage of heterozygosity, then one would not expect such behaviour. If the advantage of heterozygosity falls off with increasing heterozygosity, then this type of behaviour is to be expected.

We can now turn to a rather different aspect of inversion polymorphism. Recent work on the relations between protein variation and inversion polymorphism (Prakash & Lewontin, 1968, 1971; Prakash & Merritt, 1972; Kojima, Gillespie & Tobari, 1970; Mukai, Mettler & Chigusa, 1971) has demonstrated that loci located in or near inversions tend to show non-random associations of alleles with polymorphic inversions. It is interesting to compare the results of such observations with the theoretical conclusions derived in this paper.

In the first place, it is easy to see how differentiation in gene content between different gene arrangements can arise as a result of non-allelic interactions in fitness. Any of the classes of inversion polymorphism discussed in this paper would show this effect for at least one of loci *A* and *B* if they could be followed experimentally. With double crossing-over, it is also easy to see that this differentiation need not be absolute, provided that the fitness interactions are not of the type giving *BA/ab* equilibria. As we have seen, the double cross-over gametes are selected against with this sort of interaction, and are rare at equilibrium. If it turns out usually to be the case that genes located in or near inversions do not show almost absolute associations of alleles with gene arrangements, it can perhaps be considered unlikely that this form of interactive selection (which involves a cumulative advantage of over-dominance on a linear scale) is commonly involved in maintaining gene frequencies.

Another feature of the equilibria or quasi-equilibria when double crossing-over is taken into account is that, although alleles show non-random associations with gene arrangements, within a given sequence there is almost complete linkage equilibrium between *A* and *B* (see Table 3). Failure to detect linkage disequilibrium for protein

variants within a given gene arrangement when there are significant associations between alleles and gene arrangements (e.g. Prakash & Merritt, 1972), is thus quite consistent with the maintenance of the latter by selection.

Prakash & Merritt (1972) comment on the fact that low frequency inversions tend to be less heterozygous than the standard sequence. This is explicable in terms of the kind of result displayed in the last entry of Table 3. With low double cross-over frequencies and the appropriate type of fitness matrix, the system may take an enormous number of generations to get any where near the final equilibrium, where *In* is as heterozygous as *ST*. Furthermore, in a real population, stochastic events will considerably retard the operation of a slow process such as double crossing-over; as we have seen, the double cross-over gametes are at best neutral (with respect to the first order of their frequencies), and so will require many repeated occurrences before having a reasonable chance of establishing themselves. The assumption that the population is not at equilibrium in these cases is therefore a reasonable one, particularly for a low frequency inversion where the effect of double crossing-over is minimal. Again, if this type of observation turns out to be the rule, it would suggest that the *BA/ab* type of fitness interaction is rarely involved in the maintenance of variation.

These considerations demonstrate that the observations which have been made on associations between inversions and protein variants are consistent with a selective basis for the maintenance of the latter. They do not, of course, necessarily exclude a 'neutral mutation' interpretation of the data.

APPENDIX

(a) *Analysis of the BA/Ab/ab equilibrium*

The system with only these 3 gametes present behaves like a 3-allele system, and any equilibria must therefore satisfy the same conditions of existence and stability, as given, for example, by Crow & Kimura (1970, p. 277). The equilibrium frequencies $(\hat{y}, \hat{x}_2$ and $\hat{x}_4)$ can thus be obtained straightforwardly. The 3-allele stability conditions provide only necessary conditions for stability of the inversion polymorphism. Given that they are met, we can investigate its stability to the introduction of AB and aB at low frequencies. This yields the condition that the population is stable to this perturbation if and only if $\hat{w}_{3.} < \hat{\bar{w}}$ where

$$\hat{w}_{3.} = \hat{y}w_{13} + \hat{x}_2 w_{14} + \hat{x}_4 w_{34}.$$

With the symmetric fitness model, the conditions become somewhat simplified. The 3-allele conditions reduce to:

$$\left.\begin{aligned}
\Delta_1 &= \delta(\alpha - \gamma) + \beta(\gamma - \beta) > 0, \\
\Delta_2 &= \delta[\delta - (\beta + \gamma)] > 0, \\
\Delta_3 &= \delta(\alpha - \beta) + \gamma(\beta - \gamma) > 0, \\
\delta &> 0,
\end{aligned}\right\} \tag{A 1}$$

with $\hat{y}, \hat{x}_2$ and $\hat{x}_4$ being in the ratio $\Delta_1 : \Delta_2 : \Delta_3$.

The second and fourth inequalities together imply $\beta + \gamma < \delta$. The condition $\hat{w}_{3.} < \hat{\bar{w}}$ yields, after some reduction, the condition:

$$\alpha\delta < (\gamma - \beta)^2. \tag{A.2}$$

Combination of this condition with those of (A.1) gives the remainder of the set of equations (10).

(b) Stability analysis of the BA/AB/Ab/aB/ab equilibrium (symmetric fitness model)

It will be assumed that gene frequencies at both loci remain constant at 1/2. (It is easily shown by a local analysis that gene frequencies remain constant if the system is subjected to a perturbation from equilibrium which conserves gene frequencies in the initial generation.) Given this assumption, the system reduces to one of two dimensions, and the following matrix of coefficients governing the local stability behaviour is obtained:

$$\begin{bmatrix} 1 + \dfrac{(\beta + \gamma - \delta)}{\hat{\bar{w}}} & \dfrac{(\beta + \gamma - \delta)}{\hat{\bar{w}}} \\ \dfrac{(\beta + \gamma - \delta) - R(\hat{x}_1 + 2\hat{x}_2)}{\hat{\bar{w}}} & 1 + \dfrac{(\beta + \gamma - \delta) - R(1 - \hat{y})}{\hat{\bar{w}}} \end{bmatrix}.$$

This matrix can be shown to have real eigenvalues, at least one of which is greater in modulus than 1 if $\beta + \gamma > \delta$, proving that $\beta + \gamma < \delta$ is a necessary condition for stability.

(c) Effect of double crossing-over on the BA/AB/Ab/ab equilibrium with the symmetric fitness model

A constraint on the nature of the equilibrium is provided by the conditions for the changes in the total frequencies of the various gene combinations (the frequencies $x_i + y_i = z_i$) to be zero. Let $D^* = w_{14}(x_2 y_3 + x_3 y_2 - x_1 y_4 - x_4 y_1)$. Then we have, from equation (17) and Table 2, the equilibrium conditions:

$$\left. \begin{array}{l} \hat{z}_4(\hat{w}_{4.} - \hat{\bar{w}}) = \hat{z}_1(\hat{w}_{1.} - \hat{\bar{w}}) = R(\hat{D}_x + \hat{D}_y) - (R_1^* + R_2^*)\hat{D}^*, \\ \hat{z}_3(\hat{w}_{3.} - \hat{\bar{w}}) = \hat{z}_2(\hat{w}_{2.} - \hat{\bar{w}}) = -R(\hat{D}_x + \hat{D}_y) + (R_1^* + R_2^*)\hat{D}^*. \end{array} \right\} \tag{A.3}$$

It follows from these equations that, if the R_i^* are small and we neglect squared terms in them, we obtain the approximate relations:

$$\left. \begin{array}{l} \hat{w}_{1.} = \hat{w}_{4.}, \\ \hat{w}_{2.} = \hat{w}_{3.}. \end{array} \right\} \tag{A.4}$$

It can be shown that these equations imply that $\hat{z}_1 = \hat{z}_4$ and $\hat{z}_2 = \hat{z}_3$, except in the unlikely event $\alpha\delta = (\gamma - \beta)^2$.

A further constraint can be derived from the condition for there to be no change in the total frequencies of the *In* and *ST* gametes at equilibrium. Equation (17) and Table 2 imply that at equilibrium:

$$\frac{\sum \hat{x}_i \hat{w}_{i.}}{\sum \hat{x}_i} = \frac{\sum \hat{y}_i \hat{w}_{i.}}{\sum \hat{y}_i}. \tag{A.5}$$

Using equations (A. 4), (A. 5) reduces to

$$\frac{\hat{x}_1 + \hat{x}_4}{\hat{y}_1 + \hat{y}_4} = \frac{\hat{x}_2 + \hat{x}_3}{\hat{y}_2 + \hat{y}_3}. \tag{A. 6}$$

We also have (from the nature of the neutral curve of equilibria) the approximate relationships

$$\left. \begin{array}{l} \hat{x}_1 \hat{x}_4 = \hat{x}_2 \hat{x}_3, \\ \hat{y}_1 \hat{y}_4 = \hat{y}_2 \hat{y}_3. \end{array} \right\} \tag{A. 7}$$

The first of these gives:

$$(\hat{z}_1 - \hat{y}_4)(\hat{z}_1 - \hat{y}_1) = (\hat{z}_2 - \hat{y}_2)(\hat{z}_2 - \hat{y}_3).$$

Using the second of equations (A. 7), this gives

$$\hat{y}_2 + \hat{y}_3 = [(\hat{z}_2^2 - \hat{z}_1^2) + \hat{z}_1(\hat{y}_1 + \hat{y}_4)]/\hat{z}_1. \tag{A. 8}$$

Substituting into equation (A. 6) and rearranging, we obtain

$$\hat{y}_1 + \hat{y}_4 = \hat{z}_1.$$

Since by definition $\hat{z}_1 = \hat{x}_1 + \hat{y}_1$, this implies that

$$\left. \begin{array}{l} \hat{x}_1 = \hat{y}_4, \\ \hat{x}_4 = \hat{y}_1. \end{array} \right\} \tag{A. 9}$$

Similarly, we also have

$$\left. \begin{array}{l} \hat{x}_2 = \hat{y}_3, \\ \hat{x}_3 = \hat{y}_2. \end{array} \right\} \tag{A. 10}$$

The equilibrium is thus highly symmetrical, with the frequency of the inversion being 0·5. (This conclusion is of course dependent on the symmetric fitness model assumed.) Furthermore, we may expect $\hat{x}_1$ and $\hat{y}_4$ to have small values since equations (14) imply that the initial progress of ba is retarded by the effects of single crossing-over. Since $\hat{z}_1 > \hat{z}_2$, the majority of effective double cross-overs will take place in BA/ab individuals. If R_1^* and R_2^* are approximately equal (i.e. if the two genes are roughly equal distances from the nearest breakpoints), it is easily seen that double crossing-over contributes about equally to x_2' and x_3'. At equilibrium, therefore, we may expect $\hat{x}_2 = \hat{y}_3 \approx \hat{x}_3 = \hat{y}_2$. If $R_1^* > R_2^*$, then x_3' receives larger contributions than x_2', so that we expect $\hat{x}_3 = \hat{y}_2 > \hat{x}_2 = \hat{y}_3$. The reverse holds if $R_1^* < R_2^*$.

REFERENCES

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

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

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

DEAKIN, M. A. B. (1972). A model for inversion polymorphism. *Journal of Theoretical Biology* **35**, 191–212.

DOBZHANSKY, T. & LEVENE, H. (1951). Development of heterosis through natural selection in experimental populations of *Drosophila pseudoobscura*. *American Naturalist* **58**, 591–604.

DOBZHANSKY, T. & PAVLOVSKY, O. (1957). An experimental study of interaction between genetic drift and natural selection. *Evolution* **11**, 311–319.

FELDMAN, M. W. (1972). Selection for linkage modification. 1. Random mating populations. *Theoretical Population Biology* **3**, 324–346.

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

HALDANE, J. B. S. (1957). The conditions for co-adaptation in polymorphism for inversions. *Journal of Genetics* **55**, 218–225.

KARLIN, S. & FELDMAN, M. W. (1970). Linkage and selection: two locus symmetric viability model. *Theoretical Population Biology* **1**, 39–71.

KARLIN, S. & McGREGOR, J. (1972). Polymorphisms for genetic and ecological systems with weak coupling. *Theoretical Population Biology* **3**, 210–238.

KOJIMA, K. & TOBARI, Y. N. (1969). Selective modes associated with karyotypes in *Drosophila ananassae*. II. Heterosis and frequency-dependent selection. *Genetics* **63**, 639–651.

KOJIMA, K., GILLESPIE, J. & TOBARI, Y. N. (1970). A profile of *Drosophila* species' enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities, and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. *Biochemical Genetics* **4**, 627–637.

LEVINE, R. P. (1956). Crossing over and inversions in coadapted systems. *American Naturalist* **90**, 41–46.

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

MUKAI, T., METTLER, L. E. & CHIGUSA, S. I. (1971). Linkage disequilibrium in a local population of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences, U.S.A.* **68**, 1056–1069.

PHILIP, U., RENDEL, J. M., SPURWAY, H. & HALDANE, J. B. S. (1944). Genetics and karyology of *Drosophila subobscura*. *Nature* **154**, 260–262.

PRAKASH, S. & LEWONTIN, R. C. (1968). A molecular study of genic heterozygosity in natural populations. III. Direct evidence of coadaptation in gene arrangements of *Drosophila*. *Proceedings of the National Academy of Sciences, U.S.A.* **59**, 398–405.

PRAKASH, S. & LEWONTIN, R. C. (1971). A molecular approach to the study of genic heterozygosity in natural populations. V. Further direct evidence of coadaptation in inversions of *Drosophila*. *Genetics* **69**, 405–408.

PRAKASH, S. & MERRITT, R. B. (1972). Direct evidence of genic differentiation between sex ratio and standard gene arrangements of X chromosome in *Drosophila pseudoobscura*. *Genetics* **72**, 169–175.

STURTEVANT, A. H. (1926). A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biologisches Zentralblatt* **46**, 697–702.

STURTEVANT, A. H. & BEADLE, G. W. (1936). The relations of inversions in the X chromosome of *Drosophila melanogaster* to crossing over and disjunction. *Genetics* **21**, 544–604.

TURNER, J. R. G. (1970). Some properties of two locus systems with epistatic selection. *Genetics* **64**, 147–155.

WATANABE, T., ANDERSON, W. W., DOBZHANSKY, T. & PAVLOVSKY, O. (1970). Selection in experimental populations of *Drosophila pseudoobscura* with different initial frequencies of chromosomal variants. *Genetical Research* **15**, 123–129.