Dynamics of gametic disequilibria between loci linked to chromosome inversions: the recombination-redistributing effect of inversions

ARCADIO NAVARRO1*, ESTHER BETRÁN1, CARLOS ZAPATA2 AND ALFREDO RUIZ1
1 Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
2 Departamento de Biología Fundamental, Universidade de Santiago de Compostela, 15706 Santiago de Compostela, Spain
(Received 6 March 1995 and in revised form 24 July 1995)

Summary
The total gametic disequilibrium between two loci linked to polymorphic inversions can be partitioned into two types of components: within and between chromosome arrangements. The within components depend on the gametic disequilibrium within each chromosome arrangement. The between components depend on the locus-inversion disequilibrium. This partitioning has practical applications and is indispensable for studying the dynamics of these systems because inversions greatly reduce recombination in the heterokaryotypes while allowing free, and sometimes different, recombination in each of the homokaryotypes. We provide equations for the per generation change of the various disequilibria for systems with two and three chromosome arrangements, and the general recursive equations predicting the disequilibria after any number of generations for the case of two arrangements. Simulation studies were carried out using different values of the recombination parameters and all possible initial conditions. The results show a complex convergence to linkage equilibrium in inversion systems. The various disequilibria can have local maxima and minima while approaching equilibrium and, moreover, their dynamics cannot be described, in general, using a single parameter, i.e. an effective recombination rate. We conclude that the effects of inversions on gametic disequilibria must be carefully considered when dealing with disequilibria in inversion systems. The formulae provided in this paper can be used for such purpose.

1. Introduction
Chromosome inversions not only diminish but also redistribute recombination among the various karyotypes. In the heterokaryotypes, single crossovers within the inversion loop give rise to non-functional or non-viable aneuploid meiotic products. Recombinants between loci included within the inverted segment (or between one locus and the inversion itself) result from double crossovers and gene conversion events only. Since these processes take place at very low rates, of the order of $10^{-5}$ or $10^{-4}$ per locus per generation (Chovnick, 1973; Ishii & Charlesworth, 1977), recombination is drastically reduced, yet not completely excluded, in the heterokaryotypes. In the homokaryotypes, on the other hand, recombination is freely allowed and when inversions shift the position of genes along the chromosome, recombination rates differ among homokaryotypes for different arrangements.

Obviously, the effects on recombination exerted by chromosome inversions should be taken into account when measuring and testing gametic disequilibria between loci. The usual way to do so is to calculate the association considering only those chromosomes with a particular arrangement (Charlesworth et al. 1979; Fontdevila et al. 1983; Zapata et al. 1986; Knibb et al. 1987; Knibb & Barker, 1988; Zapata & Alvarez, 1992), i.e. the disequilibrium within chromosome arrangements. When there are locus-inversion disequilibria, associations between loci result also from the pooling of different chromosome arrangements. This disequilibrium can be termed disequilibrium between chromosome arrangements. The total disequilibrium between two loci, A and B, linked to a polymorphic chromosome, C, may be partitioned into within and between components as follows (Ruiz et al. 1991):

$$D_{AB} = u_1 D_{AB(C_1)} + u_2 D_{AB(C_2)} + \frac{D_{AC}D_{BC}}{u_1 u_2}, \quad (1)$$

* Corresponding authors: Arcadio Navarro, Department de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. Phone: (3)581-2729. Fax: (3)581-2387. E-Mail: Arcadi@CC.UAB.ES

https://doi.org/10.1017/S0016672300033486 Published online by Cambridge University Press
where $u_1$ and $u_2$ are the frequencies of the two chromosome arrangements, $C_1$ and $C_2$; $D$ refers to Lewontin–Kojima (1960) measure of gametic disequilibrium; $D_{BICD_0}$ is the disequilibrium between $A$ and $B$ within arrangement $C_1$ (producing the within components of total disequilibrium); and $D_{AC}$ and $D_{BC}$ are the locus-inversion disequilibria (producing the between component, $D_{BICD} = (D_{ACD_{BC}})/(u_1u_2)$).

Locus-inversion disequilibria may be accounted for by several factors. The coadaptation hypothesis postulates that inversions protect allele combinations that are favourable or adaptive from recombination (Dobzhansky, 1970; Lewontin, 1974; Krimbas & Loukas, 1980). This proposal is supported by theoretical work on the origin and increase in frequency of inversions. Multilocus models show that locus-inversion disequilibria are generated when an inversion trapping an allele combination with high fitness increases in frequency (Charlesworth & Charlesworth, 1973). In two-locus, two-allele models with selection, when the system finally reaches equilibrium, locus-inversion disequilibria are maximal if recombination in heterokaryotypes is zero (Deakin, 1972; Charlesworth, 1974; Deakin & Teague, 1974). There are, however, other non-selective explanations. Ishii & Charlesworth (1977) derived the rate of decay of locus-inversion disequilibria and pointed out that these disequilibria may be historical if the associations generated by chance when the inversion arose are kept long enough over time by the recombination-reducing effect of inversions. Additionally, random drift may also produce locus-inversion disequilibria if the effective population size is small enough (Nei & Li, 1975, 1980; Strobeck, 1983).

Gametic disequilibria within chromosome arrangements may also arise by selective and non-selective causes, although relatively little work has been done on this particular subject. Theoretical work on the origin of inversions (Deakin, 1972; Charlesworth, 1974; Deakin & Teague, 1974) predicts that in two-locus, two-allele systems gametic disequilibrium within chromosome arrangements will be non-existent when the inversion reaches equilibrium if recombination in the heterokaryotypes is zero. However, in more complex situations, epistatic selection may conceivably generate this kind of disequilibrium. The existence of within disequilibria is implicit in the model of recombination-induced heterosis proposed by Wasserman (1968) in which several coadapted haplotypes are assumed to segregate inside each of the arrangements. A formal and explicit treatment of the rate of decay of disequilibria within chromosome arrangements has not been carried out. The rate of decay of the total disequilibrium between two loci linked to polymorphic inversions has not been worked out either.

In this paper, we first extend the partition of the total disequilibrium between two loci to the general case of multiple alleles and chromosome arrangements. Then, we derive the rate of decay of the different components in an ideal (infinite and panmictic) population for systems of two multiallelic loci linked to chromosomes polymorphic for two or three arrangements, a common type of inversion polymorphism in Drosophila (Sperlich & Pfriem, 1986; Krimbas & Powell, 1992). Finally, we explore the dynamics of gametic disequilibria in these systems by means of deterministic simulations.

2. The model

We analysed the dynamics of gametic disequilibria in a system of two multiallelic loci, $A$ and $B$, linked to a chromosome, $C$, segregating for two or three arrangements. Recombination was the only evolutionary factor considered in our model. That is, we assumed the population to be infinite and panmictic, generations to be discrete, and selection, mutation and migration to be absent. The only parameters that may change in this model from one generation to the next $(t+1)$ are gametic frequencies and, therefore, gametic disequilibria. Allelic frequencies are constant in the population taken as a whole but may change within a particular chromosome arrangement. We used the measures of disequilibrium $D$ (Lewontin & Kojima, 1960) and $D' = D/D_{max}$ (Lewontin, 1964).

Two loci may be linked to an inversion in several ways. Here we consider three possibilities: both loci, only one, or none of them located within the inverted chromosome segment (Fig. 1). Four recombination parameters are required to provide a full description of the system dynamics. Let $r_A$ ($r_B$) denote the recombination frequency between locus $A$ ($B$) and the inversion, and $r_{12}$ and $r_3$ that between $A$ and $B$ in the two possible homokaryotypes, $C_1C_1$ and $C_2C_2$, respectively. Assuming no interference, the recombination frequency between $A$ and $B$ in the heterokaryotypes, $r_{12}$, is equal to $r_A + r_B - 2r_A r_B$. In the case of three chromosome arrangements which differ by two overlapping inversions, the two loci may be located in several different positions also. Here we considered six cases (cases in which the two loci are located beyond the same breakpoint of an inversion are assimilable to the cases with two arrangements). Those cases may be described using nine recombination parameters. Let $r_A$ ($r_B$) be the recombination frequency between locus $A$ ($B$) and inversion $C_1C_2$; $r'_A$ ($r'_B$) that with inversion $C_2C_3$; and $r''_A$ ($r''_B$) that with inversion $C_1C_3$. The frequencies of recombination between $A$ and $B$ in the three homokaryotypes, $C_1C_1$, $C_2C_2$, and $C_3C_3$, will be denoted as $r_1$, $r_2$, and $r_3$, respectively. The recombination frequencies between $A$ and $B$ in the three heterokaryotypes can be defined, assuming no interference, as above: $r_{12} = r_A + r_B - 2r_A r_B$; $r_{56} = r'_A + r'_B - 2r'_A r'_B$; and $r_{13} = r''_A + r''_B - 2r''_A r''_B$. All these recombination parameters include any processes, such as crossing-over and gene conversion events.
Gametic disequilibria and chromosome inversions

3. Results

(i) Partition of the gametic disequilibrium between two loci into components: generalization to multiple alleles and chromosome arrangements

Consider two loci, A and B, with m and n alleles, respectively \((A_1, A_2, \ldots, A_m, B_1, B_2, \ldots, B_n)\), and a polymorphic chromosome, C, with an arbitrary number, \(k\), of chromosome arrangements \((C_1, C_2, \ldots, C_n, C_{n+1}, \ldots, C_k)\). Such a system allows for \(mnk\) different gametic combinations (haplotypes). A full description of the \((mnk - 1)\) independent haplotype frequencies requires the same number of parameters. The appropriate description for studying the dynamics of the system includes three types of parameters: the allelic frequencies, the locus-inversion disequilibria, and the disequilibria between the two loci within each of the chromosome arrangements. There are \((m + n + k - 3)\) independent allelic frequencies. Let \(p_j\) \((q_j)\) denote the population frequency of the allele \(A_j\) \((B_j)\), and \(u_i\) \((u_j)\) that of chromosome arrangement \(C_i\) \((C_j)\). There are \(mnk\) disequilibrium parameters describing the association between locus A and chromosome C of the form:

\[
D_{A_jC_i} = f(A_j, C_i) - p_j u_i, \quad (2)
\]

yet only \((m-1)(k-1)\) are independent. Similarly, there are \((n-1)(k-1)\) independent disequilibrium parameters between locus B and chromosome C. Finally, any chromosome arrangement, \(C_i\), may be considered as a separate subpopulation, and the disequilibria between the two loci calculated within this subpopulation. There can be \((m-1)(n-1)\) independent disequilibrium parameters \((D_{ABjC_l})\) describing the association between the two loci within each of the \(k\) chromosome arrangements.

An alternative description of the \(mnk\) haplotypes may be obtained using a different set of parameters: the allelic frequencies, the pairwise gametic disequilibria \(D_{AB}, D_{AC}\) and \(D_{BC}\), and the three-locus disequilibria \(D_{ABC}\) (see Bennet, 1954; Thompson & Baur, 1984; Robinson et al. 1991). For each allele combination, \(A_jB_k\), a disequilibrium parameter may be defined in the total population, \(D_{A_jB_k} = f(A_j, B_k) - p_j q_k\). There will be \(mn\) such disequilibrium parameters, yet only \((m-1)(n-1)\) will be independent. This total disequilibrium may be shown to be equal to:

\[
D_{A_jB_k} = \sum_{i=1}^{k} u_i D_{A_jC_i} + \sum_{i < j} \frac{D_{A_jC_i} D_{B_kC_j}}{u_i u_j}, \quad (3)
\]

where the following simplifying notation is used:

\[
D_{A_jC_i} = u_j D_{A_jC_i} - u_i D_{A_jC_i} \quad \text{and} \quad D_{B_kC_j} = u_j D_{B_kC_j} - u_i D_{B_kC_j}. \quad (4)
\]

Equation 3 shows that the total disequilibrium between two loci can be partitioned into \(k\) within arrangements components (as many as different arrangements, first term in the right hand of the equation) and \(k(k-1)/2\) between arrangements components (as many as different heterokaryotypes, second term in the right hand). This formula is a generalization of that provided by Ruiz et al. (1991) for the simple case of two chromosome arrangements. In addition, each haplotype has a three-locus disequilibrium parameter associated with it, \(D_{ABC}\) (Thompson & Baur, 1984). Thus, there are \(mnk\) such parameters, yet only \((m-1)\)
(n-1)(k-1) are independent. The three-locus disequilibrium corresponding to a given haplotype, say A, B, C, may be written as follows:

\[ D_{ABC}(t) = u_1 D_{AB} + u_2 D_{AC} - u_1 u_2 D_{BC} \]

(9)

Therefore, both the two-locus total disequilibria, \( D_{AB} \), and the three-locus disequilibrium, \( D_{ABC} \), can be expressed in terms of their descriptive parameters, i.e. they do not provide new information about the system.

The range of these disequilibrium measures varies greatly with allelic frequencies. Thus, they are inappropriate for comparisons among populations, generations and/or pairs of loci with different allelic frequencies. The standardized measure of disequilibrium, \( D' \) (Lewontin, 1964), is very useful for comparative purposes because its range is frequency independent (Lewontin, 1988). For the total disequilibrium, \( D_{ABC}(t) \), it can be computed as either \( D_{AB}(t)/D_{max} \) when \( D_{AB}(t) > 0 \) (where \( D_{max} = \min(p(1-q), (1-p)q) \)), or \( D_{AB}(t)/D_{min} \) when \( D_{AB}(t) < 0 \) (where \( D_{min} = \max(-p, -q) \)). That is, \( D'_{AB} \) stands for the amount of actual disequilibrium relative to the maximum disequilibrium of the same sign that might be attained by the population under study.

(ii) Per generation rate of decay of the gametic disequilibrium: two chromosome arrangements

We derive in this section the per generation rate of change of the gametic disequilibrium in a system with two chromosome arrangements, \( C_1 \) and \( C_2 \) (frequencies \( u_1 \) and \( u_2 \), respectively). Hereafter the sub-indices of the alleles are dropped out to simplify the notation, i.e. \( A \) and \( B \) stand now for any given pair of alleles at the two loci (\( A_1 \) and \( B_1 \)). The subindices of the arrangements are also dropped out from locus-inversion disequilibria, \( D'_{AC} \) and \( D'_{BC} \), and the disequilibria within arrangements, \( D'_{ABC} \).

The per generation rate of decay of the gametic disequilibrium is therefore:

\[ D_{ABC}(t+1) = u_1 D_{ABC}(t) - u_1 u_2 D_{AB}(t) + D_{AC}(t) D_{BC}(t) - u_1 D_{AC}(t) D_{BC}(t) \]

(10)

and the disequilibrium within arrangement \( C_i \) may be expressed in a similar way. Using (3), (6) and (7), the total disequilibrium \( D_{AB} \) in generation \( t+1 \) may be derived:

\[ D_{AB}(t+1) = u_1 D_{ABC}(t) - u_1 u_2 D_{AB}(t) + D_{AC}(t) D_{BC}(t) - u_1 D_{AC}(t) D_{BC}(t) + u_2 D_{ABC}(t) - u_1 u_2 D_{AB}(t) + D_{AC}(t) D_{BC}(t) - u_1 D_{AC}(t) D_{BC}(t) \]

(11)

(iii) Disequilibrium dynamics with two chromosome arrangements

As pointed out before, in our model, the only parameters that can change from one generation (\( t \)) to the next (\( t+1 \)) are the two disequilibria within, \( D_{ABC}(t) \) and \( D_{ACBD}(t) \), and the two locus-inversion disequilibria that build together the component of the total disequilibrium, \( D_{BTW} = (D_{AC} D_{BC})/(u_1 u_2) \). Using matrix algebra and eqns 6 and 7, it can be proved that the general recursions for \( D_{BTW}(t) \), \( D_{ABC}(t) \), and \( D_{ACBD}(t) \) are:

\[ D_{BTW}(t) = D_{BTW}(0) A_1^t + D_{BTW}(0) A_2^t + \ldots + D_{BTW}(0) A_1^t + D_{BTW}(0) A_2^t \]

where the \( A_i \) are the eigenvalues of the system (see the Appendix for notation and derivation). Obviously, the total disequilibrium between \( A \) and \( B \), \( D_{AB}(t) \), can also be expressed in these terms:

\[ D_{AB}(t) = \lambda_1^t D_{BTW}(0) + \lambda_2^t D_{BTW}(0) + \lambda_3^t D_{BTW}(0) + \lambda_4^t D_{BTW}(0) + \lambda_5^t D_{BTW}(0) + \lambda_6^t D_{BTW}(0) \]

(12)

It can be shown that \( 0 < \lambda_1 < \lambda_2 < \lambda_3 < 1 \) unless \( r_a = r_y = 0 \), in which case \( \lambda_1 = 1 \). Therefore, it is obvious from eqn 9 that the system will always tend to an equilibrium point in which \( D_{AB}(t) = D_{BTW}(t) = D_{ABC}(t) = D_{ACBD}(t) = 0 \), except when \( r_a = r_y = 0 \). In the latter case, when \( t \) tends to \( \infty \), \( D_{ABC}(t) = D_{ACBD}(t) = 0 \) and \( D_{AB}(t) = D_{BTW}(t) \), yet this will never happen as long as double crossovers and gene conversion events present \( r_{12} \) from being zero. The asymptotic rate per generation of approach to equilibrium is given by \( \min[1 - \lambda_1, 1 - \lambda_2, 1 - \lambda_3] \); comparing these rates with that in a system without inversions (\( r_1 \) or \( r_2 \) it is obvious that the approach to equilibrium is always delayed in a system with two chromosome arrangements, just as expected from their recombination-reducing effect.

A most interesting feature of this system is that the total disequilibrium and the within arrangements disequilibria do not converge to equilibrium in a monotonically decreasing way even through the locus-inversion disequilibria do. The generation at which the maximum or minimum value of disequilibrium is attained (\( t_{max} \)) can be ascertained by equating to zero
These oscillations are the product of their dynamical transient, they can last many generations. For example, if equilibrium. Even though these oscillations are extreme and intermediate values (for example, positive disequilibria were given all possible combinations of recombination parameters. In these simulations, production of new within disequilibria allows new oscillations, but their magnitude is negligible because the first partial derivatives of eqns. 9 and 10. It can be proved in this way that these disequilibria can have local maxima and minima for \( t > 0 \), i.e. they can increase and decrease during their convergence to equilibrium. Even though these oscillations are transient, they can last many generations. For example, if \( r_x = r_y = 0.2, r_A = r_B = 0.25 \) and \( D_{ABC}(0) = -0.25 \) and \( D_{AB}(0) = 0 \), then \( t_{\text{max}} \) for \( D_{ABC} \) is 72 and \( t_{\text{max}} \) for \( D_{AB} \) is 34.

A deterministic simulation study was carried out to explore the full space of possible initial conditions and recombination parameters. In these simulations, locus-inversion disequilibria and within arrangements disequilibria were given all possible combinations of extreme and intermediate values (for example, positive maximum values for \( D_{AC} \) and \( D_{BC} \); a negative maximum value for \( D_{AB} \) and a positive maximum for \( D_{ABC} \)) and similarly with the allelic and inversion frequencies (for example, \( p_1 = q_1 = 0.5, u_1 = 0.9 \) and \( u_2 = 0.1 \)). Recombination parameters covered the three cases depicted in Fig. 1 and included also a survey of variation in the magnitude of recombination. Every set of recombination parameters was combined with every set of initial values of frequencies and disequilibria.

As analytically shown from eqns 9 and 10, the total and within arrangements disequilibria can experience oscillations during their convergence to equilibrium. These oscillations are the product of their dynamical interactions and, therefore, are useful to their study. The necessary conditions for these oscillations to take place are the following.

1. The within disequilibria can increase in two different ways. First, when \( D_{AB} = D_{AC} = D_{BC} = 0.25, D_{ABC} = 0 \), \( r_A = r_B = 0.2, r_x = r_y = 0.2 \) (Fig. 1 c). In this case, as can be seen from eqn 7, the increase is a consequence of double crossing-over in the heterokaryotypes (Fig. 2). The disequilibria generated in this manner always have the same sign as the between component of the total disequilibrium; it is generated in the same amount in both arrangements,

\[
[D = 001). \]

Gametic disequilibria and chromosome inversions

Fig. 2. Two arrangements. Temporary increase of within disequilibria (initial conditions: \( D_{AB} = D_{AC} = D_{BC} = 0.25, D_{ABC} = 0 \), \( r_A = r_B = 0.2, r_x = r_y = 0.2 \). [D and \( D' \) are shown.]

but may decay at a different rate in each one. Second, when there is a difference in disequilibrium between the two arrangements, double crossing-over can also produce disequilibrium within an arrangement from the disequilibrium within the other because of the transfer of non-equilibrium gametes (haplotypes) between arrangements.

2. The basic condition for the total disequilibrium to increase is that its components have different signs and decay at different rates. This will happen when:

(2.1) In a given generation, \( D_{AB(TW)} \) and \( D_{AB(CT)} \) have different signs and different absolute values. Then \( D_{AB} \) will increase if the rate of decay of the within component whose sign is imposed on \( D_{AB} \) is large enough for this disequilibrium to disappear and the other disequilibrium to become dominant. This will occur more easily when the two within disequilibria decay at different rates, and this will happen when (see eqn 9) (a) recombination frequencies are different in the two homokaryotypes, or (b) the two chromosome arrangements have different frequencies (Fig. 3 a).

(2.2) The sign of the between component is different from the sign of the within components and their values and rates of decay are such that they determine alternatively the sign of \( D_{AB} \) (Fig. 3 b).

Given that in this case the total disequilibrium includes three components, there can be at most a major local maximum and a major local minimum in the approach to the equilibrium. Of course, the production of new within disequilibria allows new oscillations, but their magnitude is negligible because
the new within disequilibria are always small (about $10^{-4}$, see Fig. 2).

These simulations are just aimed to explore the behaviour of our model in the whole space of possible initial conditions. However, situations with special biological significance can be studied. An example is represented in Fig. 3c. If an inversion, after increasing in frequency, remains associated with the gamete in which it originated, a total disequilibrium can be generated of inverse sign to that in the original population, giving rise to the kind of dynamics we have just described.

(iv) Three chromosome arrangements. Per generation rate of decay of gametic disequilibrium and disequilibrium dynamics

We will consider now the case of three chromosome arrangements, $C_1$, $C_2$ and $C_3$. We derive the per generation rate of decay of the various disequilibria under the same conditions as above. The simulations were carried out using the same criteria as above. The associations of allele $A$ and the three chromosome arrangements in generation $t + 1$ are:

$$D_{AC}^{(t+1)} = D_{AC}^{(t)} - r_A D_{AC}^{(t)} - r_{BC} D_{BC}^{(t)},$$
$$D_{AC}^{(t+1)} = D_{AC}^{(t)} + r_A D_{AC}^{(t)} - r_{A} D_{AC}^{(t)},$$
$$D_{AC}^{(t+1)} = D_{AC}^{(t)} + r_{A} D_{AC}^{(t)} + r_{BC} D_{BC}^{(t)} - r_{A} D_{AC}^{(t)},$$

where the notation follows that in (4). Analogous expressions can be derived in locus $B$. We can see that the locus-inversion disequilibria do not decrease monotonically in this case. They are not independent because $D_{AC} + D_{AC} + D_{AC} = D_{BC} + D_{BC} + D_{BC} = 0$. These disequilibria, therefore, may experience oscillations due to their different signs and rates of decay (Fig. 4).

Fig. 4. Three arrangements. Temporary increase in between disequilibria (initial conditions: $p_1 = p_2 = 0.5, u_1 = 0.34, u_2 = u_3 = 0.33, D_{AB} = 0.09, D_{AC} = 0.17, D_{BC} = 0.005,$ $D_{ABC} = 0, D_{ABC} = 0.0044, D_{ABC} = -0.0272, r_1 = r_2 = r_3 = 0.3, r_A = r_B = r_B = 0.05, r_{AB} = r_{BC} = 0.1$).

The disequilibrium between alleles $A$ and $B$ within chromosome arrangement $C_i$ in generation $t + 1$ is:

$$D_{ABC}^{(t+1)} = D_{ABC}^{(t)}(1 - u_A r_{1} - u_2 r_{12} - u_3 r_{13} - u_2 r_{2} - u_3 r_{23} - u_3 r_{3}) + D_{ABC}^{(t)}(r_A D_{AC}^{(t)} + r_B D_{BC}^{(t)})$$

Again, analogous expressions can be derived for disequilibria within arrangements $C_2$ and $C_3$. The production of within disequilibria is rather more complicated in this system than in the one with two chromosome arrangements. Within disequilibria can undergo two oscillations (Fig. 5).

Finally, the total disequilibrium between locus $A$ and $B$ in generation $t + 1$ is:

$$D_{AB}^{(t+1)} = D_{AB}^{(t)}(1 - u_4 r_1 - u_4 r_2 - u_4 r_3)$$

The conditions for $D_{AB}$ to increase are fully analogous to those of the system with two arrangements. The only difference is that there can be more oscillations because of the presence of three within and three between components which may

https://doi.org/10.1017/S0016672300033486 Published online by Cambridge University Press
Gametic disequilibria and chromosome inversions

Have different signs and may decay at different rates (Fig. 6).

4. Discussion

When two multiallelic loci are linked to a system of polymorphic inversions, the total disequilibrium between any two alleles can be partitioned into two types of components: (1) those that are functions of the disequilibria within chromosome arrangements; and (2) those that depend on the locus-inversion disequilibria (the between components). This decomposition is based on the recombination-redistribution effect of inversions and is necessary for two main reasons. On one side, it has a clear practical application (see Betrán et al., 1995), allowing one to disentangle the source of particular allelic associations observed in natural populations. On the other side, since the disequilibria within arrangements and the locus-inversion disequilibria decay at different rates, it is the appropriate starting point to study the dynamics of these systems.

We have used for this dynamical analysis a relatively simple model in which the only evolutionary factor involved was recombination. We focused our attention on such a model because any special feature of its dynamics may be attributed entirely to the presence of inversions that reduce and redistribute recombination. The effect of other evolutionary forces, such as selection, migration, mutation or genetic drift, is currently under investigation (Navarro et al., unpublished). One of the features of our model is no interference, which is a realistic assumption only when distances separating loci are long enough (Foss et al., 1993). However, even when interference has an effect in our system, it will not be a qualitative but a quantitative effect: the percentage of haplotype (gamete) transference between arrangements will tend to be increased when compared with the percentage of independent transference of the two loci.

The disequilibrium between two loci not linked to chromosome inversions decays according to the known expression: $D(t) = D(0)(1 - r)^t$ (Crow & Kimura, 1970). The addition of two or more arrangements complicates the situation considerably, and the formulae for the rate of decay of the various disequilibria become complex functions of the various recombination parameters and the frequencies of the chromosome arrangements.

When only two chromosome arrangements are considered, locus-inversion disequilibria decay monotonically at a rate which depends only on the recombination rate between each locus and the inversion (eqn 7). If this recombination is very low, as when the locus is located within the inversion, the formula may be approximated by $D(t) = D(0)e^{-rt}$ (Ishii & Charlesworth, 1977). When a third arrangement is present, however, locus-inversion disequilibria change at rates which are functions of the various recombination rates and the arrangement frequencies (eqn 11). Consequently, they do not always decay but may suffer transient increases due to the flow of alleles between arrangements (Fig. 5). This flow can, given the approximate recombination rates, accelerate the rate of decay of disequilibria. If those disequilibria are maintained by selection, the arrangement mainly facilitating flux (usually the intermediate arrangement) will be selectively eliminated.

The disequilibrium within a particular chromosome arrangement does not decay monotonically but shows a complex rate of change whether two (eqn 7) or three arrangements (eqn 12) are considered. It may even increase temporarily due to the 'flow' of disequilibrium from the between components or from the disequilibria within other chromosome arrangements (Figs. 3 and 6). This 'flow' is produced by double crossing-over and would not take place if either $r_A = r_A' = r_B = r_B' = r_B'' = 0$ or $r_A = r_A' = r_A'' = 0$. The conversion of the disequilibrium between into disequilibrium within was already noted by Charlesworth (1974, see his table 3). If we assume that recombination is completely suppressed in heterokatyotypes, then the disequilibrium within arrangement $C_5$ decays monotonically at a rate $(1 - u_i r_i)$. This shows that, other things being equal, within disequilibria will have a longer life in the less frequent arrangements. Moreover, the lower the frequency of an arrangement the smaller will be its effective population size ($N_{eq} = u_i 2N$; Nei & Li, 1980) and the easier it will be for genetic drift to generate disequilibria within it. Therefore, within disequilibria should be more frequently found in the less frequent arrangements (setting aside problems of statistical power for testing the disequilibria).

The disequilibrium within and the locus-inversion disequilibria have very different dynamics. Our results show that, as a general rule, when $(r_A, r_B) = (r_A')$, the within disequilibrium will always decay at a faster rate...
than the locus-inversion disequilibria. This tendency will decrease with an increasing number of arrangements, but it will still be true when using realistic values for recombination parameters, that is, when recombination rates in heterokaryotypes are as small as $10^{-4}$ or $10^{-5}$ per loci per generation. This is consistent with the more frequent finding of significant disequilibria of the latter type (see Nei and Li, 1980; Sperlich & Pfriem, 1986; Zapata & Alvarez, 1987, 1992; Krimbas & Powell, 1992).

The dynamics of the total disequilibrium between two loci is even more complex because this disequilibrium is made up of several components each decaying at a different rate (eqns 9 and 10). This disequilibrium may also increase temporarily and even show oscillations lasting up to hundreds of generations (Figs. 3 and 6). It should be recalled though that total disequilibrium is not generated de novo, it is only released by recombination from the disequilibrium between and/or within concealed in the initial conditions. If recombination between the two loci and the inversions is assumed to be zero, then the total disequilibrium changes at the rate:

$$D_{AB}(t+1) = D_{AB}(t) - u_1^2 r_1 D_{ABC_1} - u_2^2 r_2 D_{ABC_2}.$$  (14)

This expression shows that even in this simplified situation, the total disequilibrium will exhibit a complex pattern of decay. Its dynamics cannot be described using a single parameter, an effective recombination rate, such as the average frequency of recombination. Instead of that, it is necessary to obtain information about all possible recombination parameters (as in Betran et al. 1995) to describe the system fully.

The dynamics of disequilibria in a system with chromosome inversions parallel that in a subdivided population with migration (Nei & Li, 1973; Li & Nei, 1974; Feldman & Christiansen, 1975; Tachida, 1994): chromosome arrangements play the role of sub-populations and recombination that of migration. Covariance in genic frequencies between subpopulations generates disequilibria within populations when migration is present. Likewise, when locus-inversion disequilibria are present, double crossing-over and/or conversion generates disequilibria within chromosome arrangements. Migration may also transfer disequilibrium from one subpopulation to another, and again recombination may produce a similar effect. Moreover, disequilibrium in a subdivided population exhibits an oscillatory behaviour similar to that of the total disequilibrium in our system. Of course, when no selection, mutation or drift are involved, both systems will tend towards complete linkage equilibrium. There are, however, important differences which make polymorphic inversions a more complex situation. Firstly, the magnitudes of the implied parameters are very different; migration will generally be greater than recombination in heterokaryotypes. Secondly, while disequilibrium decays at the same rate in all subpopulations, it may change at a different rate within each arrangement; migration will act as double crossing-over or double gene conversion but will not allow single-locus interchanges, i.e. there is no independent locus-inversion recombination. Moreover, any single arrangement added to the system necessarily adds its own (and different) recombination parameters to it, so that a single parameter describing between arrangements interchange, as migration, is not possible. Finally, recombination implies necessarily reciprocal interchange and migration does not.

We thank A. Barbadilla and J. M. Ranz for valuable discussion and criticism. We are also grateful to referees who provided very useful suggestions. Work was supported by a FPI fellowship from the DGU (Generalitat de Catalunya, Spain) to A.N. and grant no. PB93-0844 from the DGICYT (Ministerio de Educación y Ciencia, Spain) to A.R.

Appendix

The recurrence equations are obtained from (7) and (8). In matrix notation those equations may be written as:

$$D(t+1) = RD(t),$$

where $D(t)$ is the $(3 \times 1)$ vector with elements $D_{AB(C)}$, $D_{ABC_1}$ and $D_{ABC_2}$; and $R$ is the following $(3 \times 3)$ matrix:

$$R = \begin{pmatrix}
(1-r_A)(1-r_B) & 0 & 0 \\
-r_A r_B & 1-r_1 u_1-r_A u_2-r_B u_2+r_A r_B u_2 & r_A r_B u_1 \\
-r_A r_B u_1 & 1-r_2 u_2-r_A u_1-r_B u_1+r_A r_B u_1 \\
\end{pmatrix}.$$  

The eigenvalues for $R$ are $\lambda_1 = (1-r_A)(1-r_B)$, $\lambda_2 = (2+g-b)/2$, $\lambda_3 = (2+g+b)/2$, and the corresponding eigenvectors are:

$$V_1 = \begin{pmatrix} a \\ e \\ f \end{pmatrix}, \quad V_2 = \begin{pmatrix} 0 \\ c-b \\ d \end{pmatrix}, \quad V_3 = \begin{pmatrix} 0 \\ c+b \\ d \end{pmatrix},$$
where

\[\begin{align*}
a &= (u_1 u_2 (r_1 r_2 + r_3 (r_4 r_5 + r_6 - r_7 - r_8) - r_4 r_6 (r_1 + r_2)) / (r_4 r_6)) \\
b &= ((u_1 (r_4 r_6 - r_3 r_5 + r_7 + r_8) - u_2 (r_1 - r_2 r_4 + r_5 r_6 + r_7 + r_8)) + 4 u_1 u_2 r_4 r_6 r_8^{1/2}) \\
c &= (r_4 + r_6 - r_3 r_5 (u_1 - u_2) - r_1 u_1 + r_2 u_2) \\
d &= 2 r_4 r_6 u_1 \\
e &= u_4 (r_2 r_3) \\
f &= u_4 (r_1 r_2) \\
g &= (r_4 - r_6 + r_3 r_5 - r_1 - r_2 u_2).
\end{align*}\]

The general recursions are:

\[\begin{align*}
D_{BTW}(t) &= D_{BTW}(0) \lambda_1^t \\
D_{ABCD}(t) &= k \lambda_1^t + m_1 \lambda_2^t + m_2 \lambda_3^t \\
D_{ABC}(t) &= \beta \lambda_1^t + \gamma \lambda_2^t + \delta \lambda_3^t,
\end{align*}\]

where

\[\begin{align*}
k &= (D_{BTW}(0) e_t) / a \\
 j &= (D_{BW}(0) f_t) / a \\
m_1 &= ((b - c) D_{ABCD}(0) a d - D_{ABCD}(0) a(b + c) - D_{BTW}(0) a d + D_{BTW}(0) f(b + c)) / (2 a d) \\
m_2 &= ((b + c) D_{ABCD}(0) a d + D_{ABCD}(0) a(b - c) - D_{BTW}(0) a d - D_{BTW}(0) f(b - c)) / (2 a d) \\
m_3 &= (D_{ABCD}(0) a b c + D_{ABCD}(0) a d + D_{BTW}(0) (d e - f c - f b)) / (2 a b d) \\
n_2 &= (D_{ABCD}(0) a b c + D_{ABCD}(0) a d + D_{BTW}(0) (d e - f c - f b)) / (2 a b d).
\end{align*}\]

References


MATHEMATICA v2.2 (1993). Wolfram Research Inc.


disequilibrium analysis at three linked loci with application to HLA data. *Genetics* **129**, 231–248.