Recombination and Gene Flux Caused by Gene Conversion and Crossing Over in Inversion Heterokaryotypes

Arcadio Navarro, Esther Betrán, Antonio Barbadilla and Alfredo Ruiz

Departament de Genètica i de Microbiologia, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain

Manuscript received July 11, 1996
Accepted for publication February 21, 1997

ABSTRACT

A theoretical analysis of the effects of inversions on recombination and gene flux between arrangements caused by gene conversion and crossing over was carried out. Two different mathematical models of recombination were used: the Poisson model (without interference) and the Counting model (with interference). The main results are as follows. (1) Recombination and gene flux are highly site-dependent both inside and outside the inverted regions. (2) Crossing over overwhelms gene conversion as a cause of gene flux in large inversions, while conversion becomes relatively significant in short inversions and in regions around the breakpoints. (3) Under the Counting model the recombination rate between two markers depends strongly on the position of the markers along the inverted segment. Two equally spaced markers in the central part of the inverted segment have less recombination than if they are in a more extreme position. (4) Inversions affect recombination rates in the uninverted regions of the chromosome. Recombination increases in the distal segment and decreases in the proximal segment. These results provide an explanation for a number of observations reported in the literature. Because inversions are ubiquitous in the evolutionary history of many Drosophila species, the effects of inversions on recombination are expected to influence DNA variation patterns.

A considerable effort has been devoted during the last decade to describe and interpret patterns of DNA variation in the genus Drosophila (KREITMAN and WAYNE 1994). Perhaps the main conclusion drawn from these studies is that recombination affects levels of polymorphism, explaining about one-quarter of variance among genes in nucleotide diversity (BERRY *et al.* 1991; BEGUN and AQUADRO 1991, 1992; AGUADÉ and LANGLEY 1994; MORIYAMA and POWELL 1996). Genes located in regions of the genome with low levels of recombination have low levels of polymorphism, due either to hitchhiking with favorable mutations (KAPLAN *et al.* 1989; AQUADRO and BEGUN 1993; AQUADRO *et al.* 1994) or to deleterious background selection (CHARLESWORTH *et al.* 1993; CHARLESWORTH 1994; HUDSON 1994; HUDSON and KAPLAN 1995). About three quarters of all Drosophila species, including *D. melanogaster*, are polymorphic for paracentric inversions (SPERLICH and PFRIEM 1986; KRIMBAS and POWELL 1992), which reduce recombination in the heterokaryotypes and move genes from one chromosomal region to another in the homokaryotypes (STURTEVANT and BEADLE 1936). Thus, inversions are an important factor to be considered when interpreting patterns of DNA variation both within and between species.

Inversions reduce recombination within the inverted region of heterokaryotypic females for two reasons: (1) chiasmata are partially inhibited by the inversion loop (ROBERTS 1976; COYNE *et al.* 1991, 1993) and further, (2) when crossovers do take place, they most often give rise to unbalanced meiotic products (STURTEVANT and BEADLE 1936; ROBERTS 1976). The recombination reduction, however, is not complete because viable recombinant gametes may arise by double crossing over (STURTEVANT and BEADLE 1936; SPURWAY and PHILIP 1952; NOVITSKI and BRAVER 1954; LEVINE 1956) and by gene conversion (CHOVNICK 1973). The relative importance of double crossing over and gene conversion as causes of recombination in inversion heterozygotes is unclear. Typically crossing over overwhelms gene conversion as a recombination force, but this might not hold in inversion heterokaryotypes if the rate of gene conversion were much higher than that of double crossing over. That gene conversion can be the important recombination source in this case is suggested by the thorough experimental analysis of recombination at the *rosy* locus of *D. melanogaster* (CHOVNICK 1973) and by recent population DNA variability surveys of the *rp49* locus of *D. subobscura* (ROZAS and AGUADÉ 1993, 1994) and the *amylase* gene region of *D. pseudoobscura* (POPADIĆ and ANDERSON 1995).

Inversions also affect recombination in the noninverted segments of the heterokaryotypic chromosome. Outside the inverted region, recombination rates are usually very low near the breakpoints and increase gradually as one moves away from them. In some cases, recombination may be even increased, relative to the homokaryotypes, in regions of the chromosome far

Corresponding author: Arcadio Navarro, Departament de Genètica i de Microbiologia, Facultat de Ciències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain.
E-mail: arcadi@cc.uab.es

away from the inversion (GRELL 1962; LUCCHESI and SUZUKI 1968). An asymmetrical reduction of recombination has often been observed between the proximal and distal regions of inversion heterokaryotypes. Sometimes the reduction is higher in the distal segment (*e.g.*, ROBERTS 1962), whereas in others the reduction is stronger in the proximal segment (CARSON 1953 and references therein). These observations have always been interpreted in terms of the effect of chromosome inversions on chromosome pairing and chiasma formation (ROBERTS 1976). However, they might be caused, at least in part, by the simultaneous occurrence of crossovers in the inverted and proximal regions of the chromosome that produce unbalanced gametes (PTASHNE 1960; ROBERTS 1976). This possibility has never been explored in the literature.

So far, no detailed theoretical analysis of the effect of inversions upon recombination inside and outside the inverted region has been carried out. This paper aims to help fill this vacuum. When dealing with inversion heterokaryotypes, a useful distinction may be made between recombination and flux. Recombination implies at least two markers or sites and refers to the generation of gametes with allelic combinations not present in the parental chromosomes. Flux is defined as the probability of a site being transferred from one chromosome arrangement to another during meiosis in heterokaryotypic females. Thus, flux can be considered a special case of recombination: that taking place between a specific marker or site and the inversion itself (the other marker).

Using two different mathematical models of recombination (with and without interference) we derive the expected rates of recombination and gene flux caused by crossing over and conversion inside and outside the inverted segment of heterokaryotypes. Specifically, the objectives of our analysis are as follows: (1) to assess the relative importance of the two recombination factors, (2) to describe the pattern of recombination and flux site by site along the chromosome, (3) to compare the rates of recombination and flux between homo- and heterokaryotypes, and (4) to compare the two recombination models as potential explanations of the available recombination data.

These two models of recombination are the Poisson model (HALDANE 1919) and the Counting model (FOSS *et al.* 1993). The former model ignores the well established fact of chiasma interference (MATHER 1938; see KARLIN and LIBERMAN 1994 for a recent review). In inversion heterokaryotypes, interference is expected to be especially relevant because single crossovers do not generate recombinant gametes and double crossing over becomes the important event. To ascertain the effect of interference, we have used the Counting model. Among all the models of interference so far proposed, this is the model providing the best fit to the vast amount of recombination data gathered from Drosophila and Neurospora (FOSS *et al.* 1993; MCPEEK and SPEED 1995; ZHAO *et al.* 1995). Thus, although some of its predictions concerning gene conversion are not fulfilled in Saccharomyces, it still is the best choice for Drosophila species (FOSS and STAHL 1995).

MODELS AND PARAMETER VALUES

In a meiotic tetrad, potential or intermediate recombination events (C events, see Table 1 for a list of symbols) can be resolved either as conversions with associated crossover ($C_\times$ events) or as conversions without associated crossover (C_0 events) (MORTIMER and FOGEL 1974; FOSS *et al.* 1993; FOSS and STAHL 1995). Let m be the average number of C_0's for each $C_\times$. If λ is the average number of C events in a chromosomal region, the average number of $C_\times$'s will be $\lambda/(m+1)$. In our analysis, C events are always assumed to be independently (Poisson) distributed along the chromosome and the specific distributions of $C_\times$'s and C_0's determine the particular recombination model. We consider a heterokaryotype for two arrangements that differ by a single paracentric inversion. The presence of the inversion delimits three segments in the involved chromosomal arm, proximal (P), inverted (I) and distal (D); each with its own λ (Figure 1). Also, any two markers A and B inside a given region divide it into three different segments whose lengths are α, $(\beta - \alpha)$ and $(1 - \beta)$, where α and β are the genetic distances of the two markers from one of the inversion breakpoints (the nearest breakpoint when dealing with uninverted regions) relative to the length of that region, $0 \leq \alpha \leq \beta \leq 1$ (Figure 1). λ and m are assumed to be independent of the presence of chromosomal inversions and, thus, for any given region, they are equal in homo- and heterokaryotypes. Also, we assume that there is no chromatide interference.

Poisson model: This model (HALDANE 1919) assumes that recombination intermediates (C events) resolve randomly either as $C_\times$ events, with probability $1/(m+1)$, or as C_0 events, with probability $m/(m+1)$. Accordingly, both crossovers and gene conversion events follow a Poisson distribution with parameters $\lambda/(m+1)$ and λ, respectively.

Counting model: This model (FOSS *et al.* 1993) conceives crossing-over interference as the outcome of a "machine" that counts as follows. If a random C resolves as a $C_\times$, then the next m C's must resolve as C_0 events. After the m C_0's, the next C must resolve as a $C_\times$, and so on. Thus, the machine generates a rigid sequence of $C_\times$ and C_0 events along the chromosome: $\cdots C_\times (C_0)_m C_\times (C_0)_m C_\times \cdots$. To make the whole process stationary, the $m+1$ possible $C_\times (C_0)_m$ patterns (*i.e.*, $C_\times C_0 \cdots C_0$, $C_0 C_\times \cdots C_0$, $\cdots$, $C_0 C_0 \cdots C_\times$) must be considered equally probable. Therefore, interference results from the rigidity of the sequence that makes it dependent on the genetic distance rather than

TABLE 1

Meaning of symbols

Symbol	Meaning
C	Potential recombination event
C_x	Crossover event (a conversion event is always associated with it)
C_0	Conversion event without associated crossover
m	Number of C_0's relative to the number of C_x's. In the interference model, number of C_0's between two consecutive C_x's
k, n	Number of C events in the inverted (k) and proximal (n) segments of a tetrad
λ	Average number of C events in a given chromosome or region
λ_I, λ_P, λ_D	λ in the inverted, proximal and distal regions, respectively
ρ_{ij}	Proportion of balanced and recombinant gametes produced by heterokaryotypic tetrads with i C_x's in the inverted and j C_x's in the proximal segments
P_v	Probability of viable gametes produced by an heterokaryotypic female
P_i	Probability of a number, i, of C_x events inside the inverted region
Pp_j	Probability of a number, j, of C_x events in the proximal region
α, β	Genetic distances from markers A and B, relative to that of the chromosome region in which they lay, to (1) an arbitrary breakpoint when inside the inverted region or (2) to the nearest breakpoint when outside the inverted region
$R_{Cx}(\alpha, \beta)$	Recombination between α and β caused by crossing over inside or outside the inverted region
R_C	Recombination between α and β caused by conversion
$\Phi_{Cx}(\beta)$	Gene flux caused by crossing over per site and per generation inside and outside the inverted region
Φ_C	Gene flux caused by conversion per site and per generation
$\Phi_{Total}(\beta)$	Total gene flux per site and per generation
q_P, q_D	Number of C_0's between the proximal (distal) inversion breakpoint and the nearest C_x event to the distal (proximal) direction of the breakpoint
$\bar{L}_{CxT}(m, \lambda_i)$	Average relative length of a double crossing-over tract given m and λ_i
L_{CT}	Average length of a conversion tract (in bp)
L_I, L_P, L_D	Length of the inverted, proximal and distal chromosome regions (in bp)

the physical distance. Note also that chiasma interference affects only crossing over (C_x events) but not gene conversion, which remains unchanged. That is, the probability of conversion points is Poisson distributed with parameter λ.

Parameter values: Although the formulae we derive in this paper are general, we will illustrate and discuss them using data from Drosophila because most of the evidence on inversions and recombination comes from this genus. For *D. melanogaster*, HILLIKER and CHOVNICK (1981) and HILLIKER *et al.* (1991) have estimated $m \approx 4$ (*i.e.*, four C_0 events are expected for each C_x event) at the *rosy* locus. Furthermore, using classical recombination data for many Drosophila loci, FOSS *et al.* (1993) and MCPEEK and SPEED (1995) have found that a Counting model with $m = 4$ allows for the best fit to the observations. The average length of a gene conversion tract has been estimated to be 352 bp in the *rosy* locus of *D. melanogaster* (HILLIKER *et al.* 1994) and 122 bp in the *rp49* locus of *D. subobscura* (BETRÁN *et al.* 1996). The former estimate is based in many more data and will be used here.

The *D. melanogaster* genome has an estimated size of 165 Mb, two-thirds of which (~120 Mb) are euchromatic (MERRIAM *et al.* 1991; HARTL and LOZOVSKAYA 1994). However, other species have larger genome sizes. For instance, the genome of *D. virilis* is ~313 Mb long, with 150 Mb of euchromatin (HARTL and

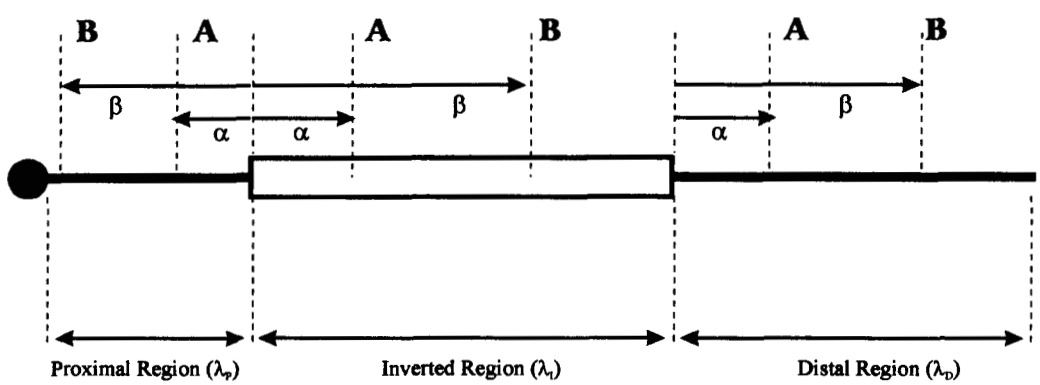

FIGURE 1.—Regions under study. Every one has its own λ (average number of C events). A and B are two markers and α and β are their genetic distances from the breakpoints (relative to that of the region).

LOZOVSKAYA 1995). A typical *D. melanogaster* chromosomal element (excluding the dot chromosome) has between 24 and 30 Mb. An average inversion captures one-third of a chromosomal element if the two breakpoints arise independently (FEDERER *et al.* 1967). Therefore, we may consider the length of a typical paracentric inversion as between 8 and 10 Mb.

The genetic length of the five chromosomal elements of *D. melanogaster* (excluding again the dot chromosome) ranges from 47 to 73 cM, the average length being 56 cM (LINDSLEY and ZIMM 1992), but chromosomes can be up to twice as long in other Drosophila species. For example, *D. virilis X* chromosome measures 171 cM (GUBENKO and EVGEN'EV 1984). This means that ~1.1–3.4 C_x's are expected per meiosis and per chromosomal element. Given that the average inversion captures one third of the chromosome, a reasonable range of values for the average number of C_x's in the inverted region ($\lambda/(m+1)$) is between 0.4 and 1.1. Accordingly, if $m = 4$ (see above), the corresponding number of C's within the inversion (λ_I) will be between two and six.

RESULTS

Probabilities of unbalanced and recombinant gametes in heterokaryotypes: Let P_i be the probability of i crossovers in the inverted segment; Pp_j, the probability of j crossovers in the proximal zone; and P_{ij}, the probability of i and j crossovers in the inverted and proximal segments, respectively. We neglect the small probabilities of more than two crossovers in any of the two regions, therefore $i, j = 0, 1$ or 2. Under the Poisson model, P_{ij} can be easily derived using the Poisson distribution as follows:

$$P_{ij} = P_i \cdot Pp_j = \frac{e^{-\frac{\lambda_I}{m+1}}\left(\frac{\lambda_I}{m+1}\right)^i}{i!}\,\frac{e^{-\frac{\lambda_P}{m+1}}\left(\frac{\lambda_P}{m+1}\right)^j}{j!}. \quad (1)$$

Under the Counting model, however, the rigidity of the $\cdots C_x(C_0)_m\cdots$ sequence makes Pp_j dependent on the number of crossovers in the inverted segment. The expression for P_{ij} under this model is derived in APPENDIX A.

Single crossovers within the inversion loop produce acentric fragments that are lost and dicentric chromosomes that form a bridge at anaphase I. Because of the ordered oogenesis in Drosophila females, unbalanced chromosomes are always set into the polar bodies and no inviable zygotes are produced (STURTEVANT and BEADLE 1936; CARSON 1946). On the other hand, double crossovers within the inversion loop produce 1/4 of unbalanced gametes because four-strand double crossovers (probability 1/4) yield 100% unbalanced gametes (STURTEVANT and BEADLE 1936; ROBERTS 1976). As shown in Figure 2a, unbalanced gametes are also generated in proportion 1/4 when two crossovers, one within the inversion loop and another in the proximal region, take place simultaneously (PTASHNE 1960; ROBERTS 1976). Other combinations of crossovers in the inverted and proximal segments of an heterokaryotype for a paracentric inversion also produce unbalanced gametes in variable proportion (Figure 2, b–d; Table 2). Let ρ_{ij} be the proportion of balanced gametes when there are i C_x's in the inverted segment and j C_x's in the proximal segment (Table 2). The average proportion of balanced gametes when there are i C_x's in the inverted zone (ρ_i) is given, for the Poisson model, by the following expression:

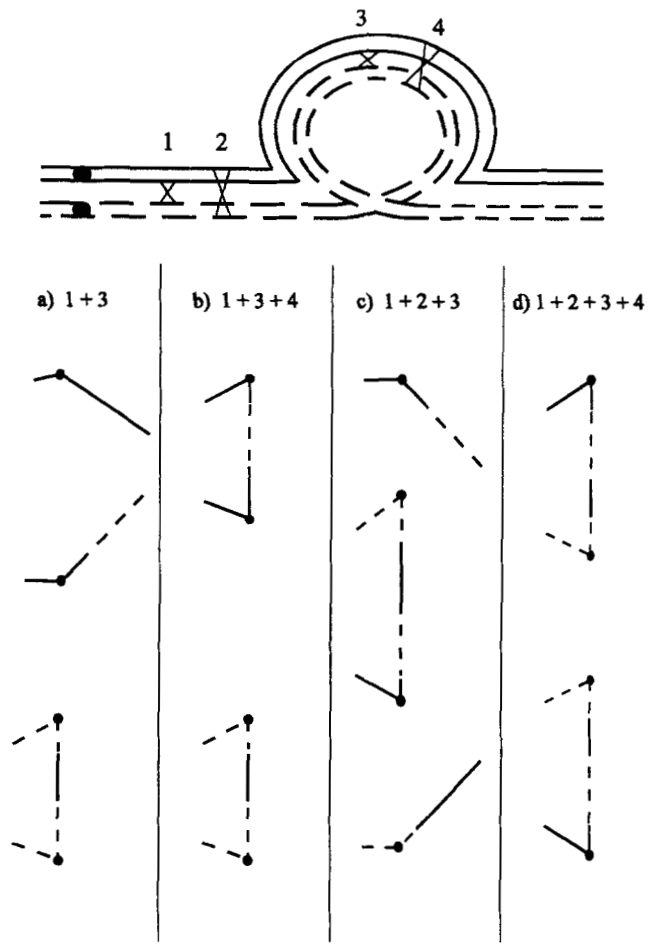

FIGURE 2.—Four examples of Anafase II configurations produced by a sample of four different combinations of crossovers involving the proximal and inverted segment. Anafase II bridges in cases a, b and d will produce a fraction of unbalanced gametes ($^1/_2$, 1 and 1, respectively).

$$\rho_i = \frac{\sum_{j=0}^{2}\left(P(j C_x\text{'s Prox} \mid i C_x\text{'s In})\,\rho_{ij}\right)}{\sum_{j=0}^{2} P(j C_x\text{'s Prox} \mid i C_x\text{'s In})}. \quad (2)$$

Combining the probabilities of zero, one or two crossovers in the inverted and proximal regions, P_{ij}, with their correspondent ρ_{ij}'s, we obtain the following proportion of viable gametes:

TABLE 2

Expected gametes produced by meioses with different numbers of crossovers in a Drosophila female heterokaryotypic for a paracentric inversion

Crossovers		Balanced gametes			Unbalanced gametes
Inverted segment	Proximal segment	Parental	Recombinant	Total (ρ_{ij})	
1	0	1	0	1	0
	1	3/4	0	3/4	1/4
	2	7/8	0	7/8	1/8
2	0	3/8	3/8	3/4	1/4
	1	5/16	5/16	5/8	3/8
	2	11/32	11/32	11/16	5/16

Parental and recombinant are referred to the inverted segment.

$$P_v = \frac{\Sigma_{i=0}^{2}\Sigma_{j=0}^{2} P_{ij}\rho_{ij}}{\Sigma_{i=0}^{2}\Sigma_{j=0}^{2} P_{ij}}. \quad (3)$$

As shown by expression (3), natural selection is expected to act against inversions due to the semisterility of heterokaryotypes, the maximum selection coefficient against the heterokaryotype being $s = 1 - P_v$. The selection expected against paracentric inversions is mild but increases with both the genetic length of the inversion and the length of the proximal region, *i.e.*, its distance to the centromere (Figure 3).

Recombination and flux inside the inverted region caused by double crossing over: *Poisson model:* Given two independent crossovers in the inverted region of a given gamete, the probability that the two markers A and B recombine, *i.e.*, the probability that one of the two C_x's lies inside the A-B region and the other one outside it, is

$$R_{Cx}((\alpha, \beta)|2C_x\text{'s In}) = 2(\beta - \alpha)(1 - (\beta - \alpha)), \quad (4)$$

where α and β are the genetic distances of the two markers from one of the inversion breakpoints relative to the length of the inversion (Figure 1). Therefore, the recombination probability of the two markers is derived multiplying expression (4) by P_2, ρ_2 and $^1/_2$ and dividing by the probability of viable gametes, P_v:

$$R_{Cx}(\alpha, \beta) = \frac{\rho_2}{2P_v}(\beta - \alpha) \times (1 - (\beta - \alpha))e^{-(\lambda_I/(m+1))}\left(\frac{\lambda_I}{m+1}\right)^2. \quad (5)$$

It follows from (5) that, for a fixed distance between A and B, $R_{Cx}(\alpha, \beta)$ is constant and does not depend on the position of the two markers.

The flux caused by double crossing over for any site B in the inverted region, $\Phi_{Cx}(\beta)$, is a particular case of recombination and can be readily derived from (5) making $\alpha = 0$. Figure 4 shows a graphical representation of $\Phi_{Cx}(\beta)$ as a function of β for realistic values of parameter λ_I. Flux reaches its maximum at the center of the inversion and decreases, following a parabola, as we move away from the center; it is null at the breakpoints.

Counting model: Given k C events and two C_x events inside the inverted region, the probability of recombination between A and B, $R_{Cx}(\alpha, \beta)$, is obtained by adding up the probabilities of all possible combinations of C_x's and C_0's that allow for a single C_x to be inside the A-B segment. The result (APPENDIX B, expressions A7 and A8) is not simple because the rigidity of the $\cdots C_x(C_0)_m\cdots$ sequence must be taken into account.

While the Poisson model predicts constant recombination values along the entire inverted segment, the Counting model predicts that recombination rates will depend strongly on λ_I and on the position of the markers along the inverted segment (Figure 5a). The reason for this is that double crossover tracts are longer under the Counting model (see APPENDIX C, for mathematical proof). Long double crossover tracts make two markers very difficult to separate when both of them lie in a central position and, thus, their recombination rate can be smaller than when they are in more extreme locations.

The rate of flux for each site B, $\Phi_{Cx}(\beta)$, can be derived, as above, replacing $\alpha = 0$ in expression (A8). When $\Phi_{Cx}(\beta)$ is compared with the function previously obtained for the Poisson model, it proves to be more sensitive to changes in λ_I (Figure 4). The Counting model predicts greater flux than the Poisson model for $\lambda_I > 4$ (approximately) and smaller flux than the Poisson model for $\lambda_I < 4$. The explanation of this fact is again related to the higher average length of double crossover tracts under the Counting model (APPENDIX C). When an inversion is short, it is very difficult for two C_x's to be inside it because $(m + 2)$ C's must also be present, so that flux is always small in short inversions. On the contrary, in long inversions the rigidity of the model increases both the probability of two C_x's in the inversion and the length of the transferred tract.

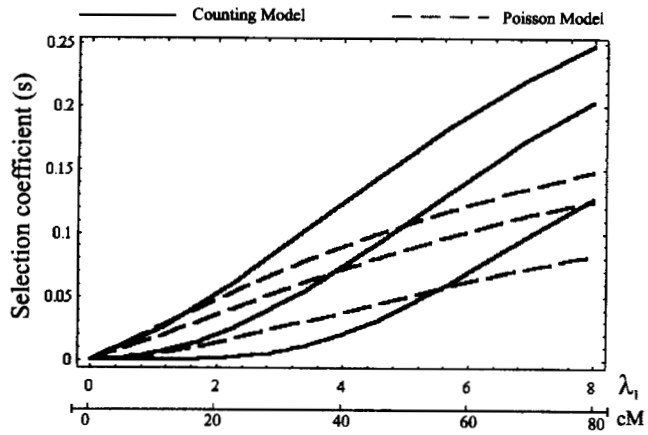

FIGURE 3.—Maximum selection coefficients caused by the production of unbalanced gametes in paracentric inversions of different genetics lengths (λ_P = 0, 2 and 4 from the lower to the higher line). - - -, Poisson model; —, Counting model, interference pattern m = 4.

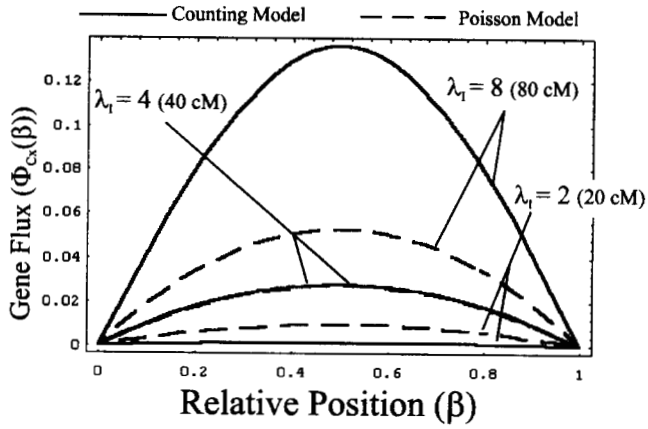

FIGURE 4.—Exact flux probabilities due to double crossing over for any position β inside the inverted region. - - -, Poisson model; —, Counting model, interference pattern m = 4. (λ_I = 5).

Flux and recombination inside the inverted region caused by gene conversion: *Poisson model:* Given a conversion event, the longer its tract length, the larger its recombination effect. Let L_{CT} and L_I be the expected length in bp of a conversion tract and of the inverted segment, respectively. Assuming unbiased conversion and no postmeiotic segregation, every C event in a tetrad generates 1/4 "converted" gametes (see DISCUSSION) and an approximate expression for the flux per nucleotide per generation caused by gene conversion (Φ_C) is

$$\Phi_C = \frac{\lambda_I L_{CT}}{4 L_I}. \qquad (6)$$

Because $L_{CT} \ll L_I$, the edge effect of inversions is negligible, and expression (6) is valid for virtually every site within the inversion. If two sites are farther away than the average conversion tract length, simultaneous conversion of the two sites can be neglected because of its extremely low probability. Then, the conversion-caused recombination between any two sites is the addition of the conversion probabilities of the two sites. That is, (6) must be multiplied by 2 to obtain the recombination probability, R_C.

Counting model: Interference affects gene conversion rates because the unbalanced gametes produced by crossing over do not contain a random number of C events. The average number of C events that occur in the balanced gametes (effective average number of C events), is

$$\lambda_I^{eff} = \frac{1}{P_v}\left[\mathrm{Conv}_0 + \rho_1 \mathrm{Conv}_1 + \rho_2 \mathrm{Conv}_2\right], \qquad (7)$$

where Conv_i is the average number of C events given i C_x's inside the inversion,

$$\mathrm{Conv}_0 = \sum_{k=0}^{m} \frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{(m+1)-k}{m+1} k,$$

$$\mathrm{Conv}_1 = \sum_{k=1}^{m} \frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{k}{m+1}(k-1) + \sum_{k=m+1}^{2m+1} \frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{2(m+1)-k}{m+1}(k-1),$$

$$\mathrm{Conv}_2 = \sum_{k=m+2}^{2m+2} \frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{k-(m+1)}{m+1}\left(\frac{k}{4} + \frac{k-1}{2}\right). \qquad (8)$$

When comparing λ_I and λ_I^{eff} along a realistic range of λ_I values ($2 < \lambda_I < 8$), the order of magnitude of the predicted Φ_C never changes, so that, for our purposes, (6) is approximate enough.

Total flux inside the inverted region and equivalence points: *Poisson model:* A single parameter, $\Phi_{Total}(\beta)$, jointly describing the total flux caused by gene conversion and double crossing over can be defined for every position in the inverted region adding up $\Phi_{Cx}(\beta)$ and Φ_C. Note that while the flux caused by conversion events depends both on the physical and on the genetic length of the inversion, the flux caused by double crossing over depends exclusively on the genetic length of the inverted region. The β values at which conversion equals crossing over as a cause of flux (equivalence points, β_{eq}) can be obtained by solving the equation $\Phi_{Cx}(\beta) = \Phi_C$ for β. There are two β_{eq} values. Let us focus on the one closest to the proximal breakpoint. Higher β_{eq} values imply an increasing relative importance of conversion over crossing over as a cause of gene flux inside the inverted region. In general, crossing over is the dominant force in the central region, while conversion becomes important near the breakpoints (Figure 6a). Figure 6b shows β_{eq} for several values of λ_I. Only for very tiny inversions ($\lambda_I < 0.1$, which means < 1 cM when m = 4) would gene conversion be the main cause of flux along the entire inverted segment.

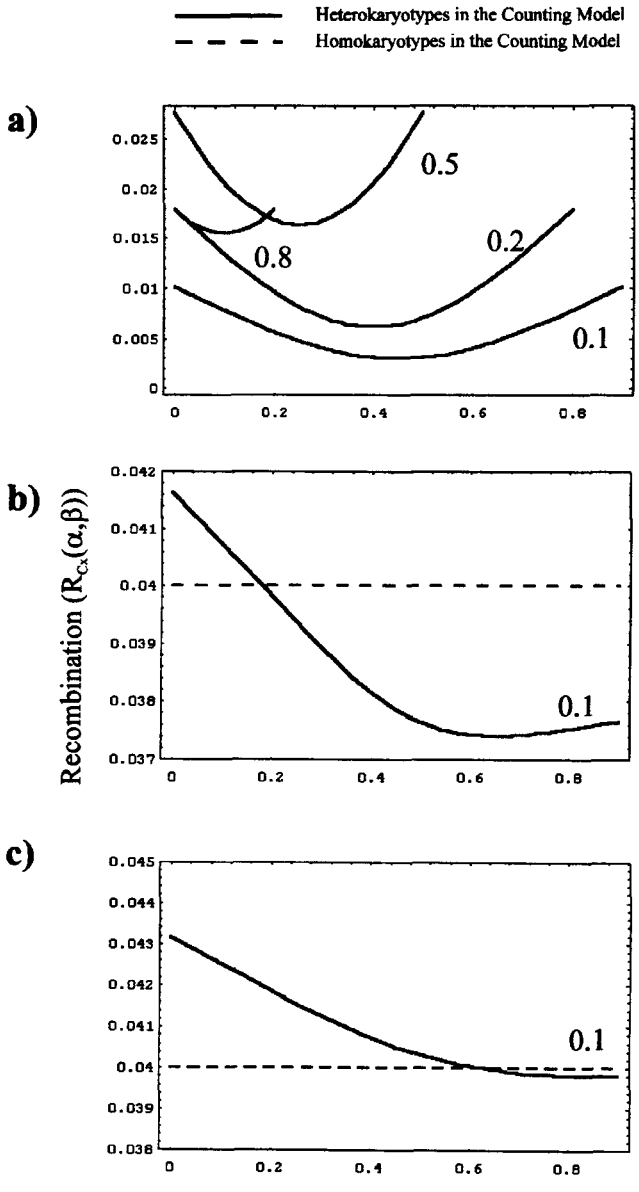

FIGURE 5.—Recombination probabilities ($R_{Cx}(\alpha, \beta)$) between two markers, A and B, at fixed distances from each other and at different distances from the inversion breakpoints, which are referred to as distances of A to the left breakpoint. Counting model (m = 4, λ_I = 4, λ_P = 4, λ_D = 4). ---, Counting model, homokaryotypes. —, Counting model, heterokaryotypes. (a) Inverted segment [$(\beta - \alpha)$ = 0.1, 0.2, 0.5, 0.8]. (b) Proximal segment [$(\beta - \alpha)$ = 0.1]. (c) Distal segment [$(\beta - \alpha)$ = 0.1].

Counting model: In this case, developing an analytical expression for the equivalence points (β_{eq}) is hardly possible. However, by means of a numerical approach it is possible to find values of β_{eq} for given values of the parameters (Figure 6b). When comparing these values with those obtained for the Poisson model, a major change has taken place in the relationship between gene conversion and crossing over. In the Counting model, the range of values of λ_I for which gene conversion is important as a cause of flux increases by more than an order of magnitude (approximately, $\lambda_I < 1.5$ in the Counting model *vs.* $\lambda_I < 0.1$ in the Poisson model). On the other hand, in long inversions, when double crossing over is the dominant force, the Counting model predicts smaller values of β_{eq} than the Poisson model, that is, the Counting model proves again to be more sensitive to changes in λ_I values.

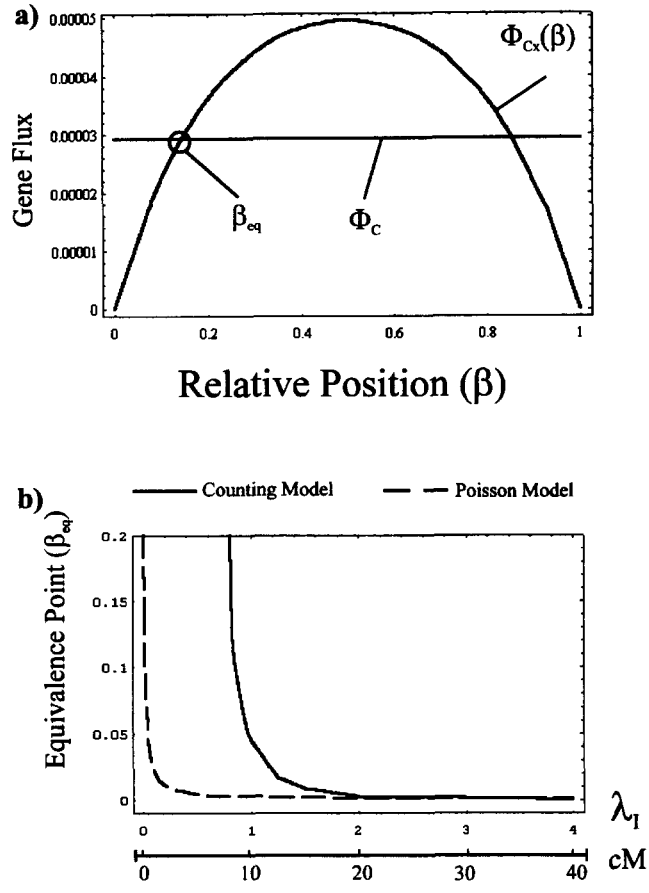

FIGURE 6.—Equivalence points. (a) Counting model (m = 4, λ_I = 1, L_{CT} = 352 bp, L_I = 3 Mb, λ_P = 4). The double crossing-over caused flux (parabolic line) and the gene conversion-caused flux (straight line) equal each other at $\beta_{eq} \approx$ 0.17. (b) β_{eq} positions under different λ_{In} values (L_{CT} = 352 bp, L_I = 8 Mb, λ_P = 4). ---, Poisson model; —, Counting model (m = 4).

Recombination and flux outside the inverted region caused by crossing over: *Poisson model:* The recombination frequency between any two markers A and B in the distal region (Figure 1) is easily obtained by multiplying the Poisson probability of one or more C_x events between the two markers by $^1/_2$, the proportion of recombinant gametes when one or more C_x events take place in a tetrad.

$$R_{Cx}(\alpha, \beta) = {}^1/_2(1 - e^{-(\beta-\alpha)(\lambda_D)/(m+1)}), \quad (9)$$

where $\lambda_D/(m + 1)$ is the average number of C_x events in the distal segment under study, and α and β are the genetic distances, relative to that of the noninverted segment, from markers A and B, respectively, to the

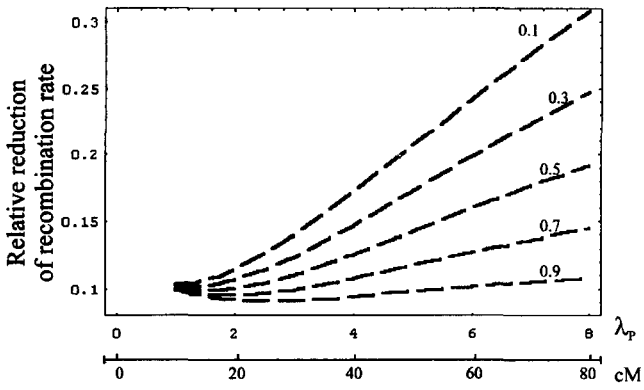

FIGURE 7.—Reduction of recombination in heterokaryotypes (relative to that of homokaryotypes) between two markers A and B in the proximal segment at different distances from each other [$(\beta - \alpha) = 0.1, 0.3, 0.5, 0.7, 0.9$], under different values of λ_P ($\lambda_I = 4$) and under the Poisson model.

nearest breakpoint. This expression holds also for homokaryotypes, because it is only dependent on the distance between the two markers. The flux caused by crossing over, $\Phi_{Cx}(\beta)$, is the Poisson probability of one or more C_x events along β and can be derived from (9) making $\alpha = 0$.

The situation is more complex in the proximal segment because some combinations of crossovers in the proximal and inverted segments increase the production of unbalanced gametes (Table 2). Neglecting the probability of more than two C_x's either in the inverted or the proximal region, the recombination between A and B in the proximal region can be approximated by

$$R_{Cx}(\alpha, \beta) = \frac{\Sigma_{i=0}^{2} (P_{i1}(\beta - \alpha)\rho_{i1}) + \Sigma_{i=0}^{2} \times (P_{i2}(1 - (1 - (\beta - \alpha))^2)\rho_{i2})}{2 \Sigma_{i=0}^{2} \Sigma_{j=0}^{2} (P_{ij}\rho_{ij})}. \quad (10)$$

The increased proportion of unbalanced gametes produces a considerable decrease of recombination rates in the proximal zone. (Figure 7).

Counting model: The recombination between two markers A and B at distances α and β from an arbitrary reference point when no inversion is present, that is, in any region of an homokaryotype, is given by the probability of one or more C_x events between the two markers multiplied by $^1/_2$, the number of recombinant gametes (Figure 5, b and c):

$$R_{Cx}(\alpha, \beta) = \frac{1}{2} \times \left(1 - \sum_{j=0}^{m} \frac{e^{-\lambda(\beta-\alpha)}(\lambda(\beta - \alpha))^j}{j!} \frac{(m + 1) - j}{(m + 1)}\right), \quad (11)$$

where j is the number of C events between A and B. In heterokaryotypes the situation is more complicated because the Counting model predicts that the crossovers inside the inverted region are correlated with those outside it. Given that some special crossovers in the inverted and proximal regions produce unbalanced gametes, their corresponding recombination events in the distal segment will also be included in unbalanced gametes. Taking this relationship into account, the recombination between two markers outside the inverted region, $R_{Cx}(\alpha, \beta)$, is derived in APPENDIX D for both the distal and proximal regions (expressions A13 and A16). Again, an expression for gene flux outside the inverted region under the Counting model, $\Phi_{Cx}(\beta)$, can be derived making $\alpha = 0$ in those expressions.

Figure 5, (b and c) shows a comparison of the recombination probability outside the inverted region of homo- and heterokaryotypes. The Counting model predicts that in the distal region recombination rates will be, most of the time, higher in heterokaryotypes than in homokaryotypes (Figure 5c). On the other hand, in the proximal segment recombination rates are, as average, smaller in heterokaryotypes (although not as small as under the Poisson model, Figure 7). In other words, inversions induce a polarity in recombination rates by reducing them in the proximal segment and, under the Counting model, increasing them in the distal segment.

Flux and recombination outside the inverted region caused by gene conversion: *Poisson and Counting models:* Let L_{NIn} be the length of a noninverted region measured in bp. Using the same arguments as above, we get an approximate expression for the flux caused by gene conversion under the Poisson model:

$$\Phi_C = \frac{\lambda_{NIn} L_{CT}}{4 L_{NIn}}, \quad (12)$$

where λ_{NIn} is the average number of C events in any noninverted segment. Again, an approximate expression for the conversion-caused recombination, R_C, can be obtained multiplying Φ_C by 2. For the same reasons as above, (12) also approximates gene conversion in the Counting model.

Total flux outside the inverted region and equivalence points: *Poisson model:* The total flux, $\Phi_{Total}(\beta)$, caused by gene conversion and by crossing over can also be defined here adding up $\Phi_{Cx}(\beta)$ and Φ_C. It is also possible to find the equivalence points (β_{eq}) that, of course, are different in the proximal and distal regions (expressions not shown). However, a general conclusion can be drawn: outside the inverted region β_{eq} is always very small (of the order of 10^{-4}) because crossing over is always the dominant force.

Counting model: A numerical approach is also needed at this point to calculate some β_{eq} values. The higher recombination rates predicted by the Counting model in regions close to the inversion make β_{eq} values only slightly smaller when compared with those predicted for the Poisson model.

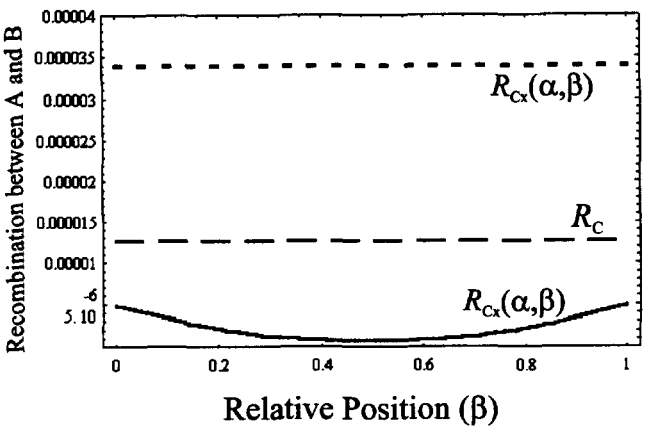

FIGURE 8.—Predictions of recombination rates in inversion *In(3R)* P_{18} of *D. melanogaster* between two markers A and B at a fixed distance ($\beta - \alpha$) and at different distances from the inversion breakpoints. Inversion is 17.9 cM ($\lambda_I \approx 1.79$, when $m = 4$; $\lambda_P = 0$) and 15,560 kb long. Locus *rosy* lies in the central region and the maximum possible distance, relative to that of the inversion, between two mutant markers inside *rosy* is 10^{-3} (CHOVNICK 1973), so ($\beta - \alpha$) $\approx 10^{-3}$. —, Recombination caused by double crossing-over, $R_{Cx}(\alpha, \beta)$, under the counting model ($m = 4$). ---, $R_{Cx}(\alpha, \beta)$ under the Poisson model. – – –, Recombination caused by conversion, R_C, under both models. Under the Counting model, the rates of recombination caused by gene conversion are an order of magnitude higher than those caused by double crossing over.

DISCUSSION

Recombination inside the inverted region: Under the Poisson model, recombination between two markers inside the inverted region of heterokaryotypes depends only on the distance between these markers. In contrast, under the Counting model, the position of markers becomes a key factor. Because of the greater average length of double crossover tracts when interference is present (whatever the model used to study interference), two markers in the center of an inversion will recombine less than if they were in a more lateral position within the inversion (Figure 5). This is consistent with CHOVNICK's (1973) finding of a very low percentage of crossovers between two markers in the *rosy* locus of *D. melanogaster,* which lies in the center of the 17.9-cM long inversion *In(3R)* P_{18}. Using markers inside *rosy,* at a maximum distance of $\sim 10^{-2}$ cM from each other, C_x and C_0 events were distinguished and scored. Chovnick's results show a clear predominance of gene conversion as a cause of recombination between close markers in the inverted region of heterokaryotypes. As shown in Figure 8, those results are not explainable under the Poisson model unless a strong inhibition of crossing over is postulated. The Counting model, on the other hand, accounts for Chovnick's data because it predicts that the rate of recombination caused by gene conversion is one order of magnitude higher than that caused by double crossing over.

Gene flux inside the inverted region: Three general results arise from our study. (1) Flux caused by crossing over is null at the breakpoints. (2) The central region of inversions presents, in general, higher flux rates because of the flux caused by double crossing over, which dominates there. Conversion only becomes important in regions close to the breakpoints. The length of these regions of conversion dominance is inversely proportional to the genetic length of the inversion. (3) In very short inversions, where conversion becomes the main recombination force, flux rates will be more evenly distributed along the inverted segment. Long inversions, on the contrary, will present huge flux differences between the central region and the regions close to the breakpoints (Figure 4). These results agree with the pattern of gene flux found in different regions of inversion *In(3L)Payne* of *D. melanogaster* (WESLEY and EANES 1994; HASSON and EANES 1996).

All three results are common to the two models of recombination used. Yet, the two models show marked differences in their flux predictions (Figure 4), which, again, make the Counting model more compatible with observations. Under the Poisson model, only in very short inversions ($\lambda_I < 0.1$; *i.e.,* inversions shorter than 1 cM) will conversion be of some importance as a cause of gene flux. On the other hand, under the Counting model the range of λ_I values for which conversion is the dominant force all over the inverted region is ~15 times larger, and the interval of conversion dominance around the breakpoints embraces a longer segment. This makes interference an important factor to explain the DNA variability observed in *rp49* (ROZAS and AGUADÉ 1994) and *Amy* (POPADIĆ and ANDERSON 1995; POPADIĆ *et al.* 1995) in which conversion has been found to be the only cause of flux between arrangements in regions close to inversion breakpoints. For instance, gene *rp49* lies in chromosome O of *D. subobscura,* very close to one of the breakpoints of inversion O_3. A sample of 34 sequences of this gene was obtained by ROZAS and AGUADÉ (1994), 17 from arrangement O_{ST} and 17 from O_{3+4}. Using their data, BETRÁN *et al.* (1997) estimated that whereas 19 conversion tracts had been interchanged between the two arrangements in the sample during the last 2.78×10^7 generations, no trace was found of flux caused by double crossing over.

The above discussion clearly illustrates two conclusions. The first one is that gene conversion and crossing over have different roles when considering either recombination or gene flux. In general, crossing over is expected to be the dominant force but conversion will overcome it in two cases: (1) as a cause of gene flux between arrangements around breakpoints, as found by ROZAS and AGUADÉ (1994); and (2) with interference, as a cause of recombination between closely linked markers in the inverted region of heterokaryotypes, as found by CHOVNICK (1973). The last tendency will be stronger when dealing with markers lying in the middle of the inverted region, because double crossovers cannot easily separate them.

The second conclusion is that interference explains better the empirical data. Under the Poisson model, claims to an almost complete inhibition of crossing over in inversion heterokaryotypes have to be made to explain data such as those of CHOVNICK (1973) or ROZAS and AGUADÉ (1994). Furthermore, to make these studies consistent with some other data sets in which nonnegligible rates of double crossing over have been detected inside inverted regions (see KRIMBAS and POWELL 1992 for a review), this postulated inhibition must change from one inversion to the other. Under the Counting model, such claims are not necessary. In this case, double crossing over is unlikely in short inversions ($\lambda_I < 2$, that is <20 cM), so that the expected flux rate caused by double crossing over is much smaller under the Counting model than under the Poisson model. In addition, long inversions (approximately, $\lambda_I > 4$, that is >40 cM) will have higher flux rates under the Counting model than under the Poisson model because double crossovers are highly probable in long inversions. Therefore, the Counting model fits the observations reported in the literature better, because it allows for both conversion domination in short inversions or in intervals close to the breakpoints and high double crossing over rates in long inversions. In general, these conclusions apply not only to the Counting model but to any model of interference because they depend on the fact that positive interference always induces a greater distance between neighboring crossovers.

Recombination and gene flux outside the inverted region: As expected, under both models recombination in the noninverted regions of heterokaryotypes is dominated everywhere by crossing over, and gene flux is lower for sites located at shorter distance from the breakpoints. Also, under both models, there is a strong reduction of recombination between any two markers in the proximal zone (Figures 5b and 7). This reduction, which extends even to markers that are segregating independently of the inversions, is caused by the production of unbalanced gametes by given combinations of crossovers in the proximal and inverted segments. On the contrary, crossovers in the distal segment have no effect on fertility. Thus, inversions create a polarity in the chromosome arm where they occur, which can account at least for part of the reduction of recombination in the proximal zone described in the literature (CARSON 1953 and the references therein). However, there are still differences between the Counting and the Poisson models. Because under the Counting model recombination events inside the inverted region are related with events outside it, inversions have an unexpected and remarkable effect on recombination rates: they can increase them in regions close to the breakpoints of heterokaryotypes, mainly in the distal region (Figure 5c). This result, which adds a further recombination-redistributing effect of inversions, may also be useful to explain the increment in recombination rates outside the inversion found by several authors (GRELL 1962; the references in LUCCHESI and SUZUKI 1968 and KRIMBAS and POWELL 1992).

Assumptions and consequences of our models: Two main assumptions underlie both models. First, we have assumed that the presence of inversions does not alter the number of recombination events, C's, C_0's and C_x's, that occur in a given region. This assumption is necessary if one wishes to study, as we do here, the effects of recombination *per se*. However, several authors have shown the existence of crossing over inhibition in the breakpoints and in areas nearby (NOVITSKI and BRAVER 1954; RUIZ and ALBEROLA 1983; COYNE *et al.* 1991; COYNE *et al.* 1993. See also KRIMBAS and POWELL 1992 for a review). This inhibition seems to have a "mechanical" basis because it decreases as the distance from the breakpoint increases. In any case, the inhibition affects both C_x's and C_0's (CHOVNICK *et al.* 1971; HILLIKER and CHOVNICK 1981). Therefore, it will not alter qualitatively our conclusions regarding the relative importance of crossing over and conversion as sources of recombination.

Our second assumption is that gene conversion events come exclusively from complete chromatid conversions, *i.e.*, that every heteroduplex is repaired, so that there is no postmeiotic segregation. Although postmeiotic segregation has been described in *D. melanogaster* in some special cases (as in repair-defective mutants, see CARPENTER 1982; LESLIE and WATT 1986), its effect on conversion rates seems to be too complex (LAMB and HELMI 1982) and small (LESLIE and WATT 1984, 1986) to be worth including in our models.

The presence of inversions enforces a dramatic modification of recombination rates along the whole chromosome. As we have shown, inversions do more than reduce recombination: they redistribute it in several ways, generating complex patterns of recombination and gene flux along the chromosome. Our equations provide a site by site description of those patterns that is crucial to study both the dynamics of gametic disequilibria (ISHII and CHARLESWORTH 1977; NAVARRO *et al.* 1996) and the patterns of molecular variability in inversion systems (KREITMAN and WAYNE 1994; SCHAEFFER 1994). The two-locus 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 disequilibria. For neutral variation, the rate of decay of the locus-inversion disequilibria depends only on gene flux. Therefore, we expect these disequilibria to be stronger and nucleotide divergence between arrangements to be higher in low flux regions (see Figure 4). Other issues concerning molecular variability, such as the extent to which nucle-

otide polymorphism is affected by inversions, or how the presence of different gene arrangements affect hitchhiking and background selection, are beyond the scope of this paper. Our results are a first step for further dynamical studies aimed at answering such specific questions about nucleotide variability within a chromosomal context.

We thank A. BERRY, A. CABALLERO, M. CÁCERES, A. LLOPART, J. V. MAÑANAS, J. M. RANZ, J. ROZAS and two anonymous referees for valuable discussion and criticism. Work was supported by a FPI fellowship from the Generalitat de Catalunya, Spain to A.N. and grants PB93-0844 and PB95-0607 from the Ministerio de Educación y Ciencia, Spain to A.R.

LITERATURE CITED

AGUADÉ, M., and C. H. LANGLEY, 1994 Polymorphism and divergence in regions of low recombination in *Drosophila*, pp. 67–76 in *Non-Neutral Evolution*, edited by B. GOLDING. Chapman & Hall, New York.

AQUADRO, C. F., and D. J. BEGUN, 1993 Evidence for an implication of genetic hitchhiking in the *Drosophila* genome, pp. 159–178 in *Mechanisms of Molecular Evolution*, edited by N. TAKAHATA and A. G. CLARK. Sinauer Associates Inc., Sunderland, MA.

AQUADRO, C. F., D. J. BEGUN and E. C. KINDAHL, 1994 Selection, recombination and DNA polymorphism in *Drosophila*, pp. 46–56 in *Non-Neutral Evolution*, edited by B. GOLDING. Chapman & Hall, New York.

BEGUN, D. J., and C. F. AQUADRO, 1991 Molecular population genetics of the distal portion of the X chromosome in *Drosophila:* evidence for genetic hitchhiking of the *yellow-achaete* region. Genetics **129:** 1147–1158.

BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Genet. Res. Camb. **53:** 7–15.

BERRY, A. J., J. W. AJIOKA and M. KREITMAN, 1991 Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. Genetics **129:** 1085–1098.

BETRÁN, E., J. ROZAS, A. NAVARRO and A. BARBADILLA, 1997 The estimation of the number and the length distribution of gene conversion tracts from population DNA sequence data. Genetics **146:** 000–000.

BURNHAM, C. R. (Editor), 1980 *Discussions in Cytogenetics*. St. Paul, MN.

CARPENTER, A. T. C., 1982 Mismatch repair, gene conversion, and crossing over in two recombination defective mutants of *Drosophila melanogaster*. Genetics **69:** 179–209.

CARSON, H. L., 1946 The selective elimination of inversion dicentric chromatids during meiosis in the eggs of *Sciara impatiens*. Genetics **31:** 95–113.

CARSON, H. L., 1953 The effect of inversions on crossing over in *Drosophila robusta*. Genetics **38:** 168–186.

CHARLESWORTH, B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. Camb. **63:** 213–227.

CHARLESWORTH, B., M. T. MORGAN and D. CHARLESWORTH, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.

CHOVNICK, A., 1973 Gene conversion and transfer of genetic information within the inverted region of inversion heterozygotes. Genetics **75:** 123–131.

COYNE, J. A., S. AULARD and A. BERRY, 1991 Lack of underdominance in a naturally occurring pericentric inversion in *Drosophila melanogaster* and its implications for chromosome evolution. Genetics **129:** 791–802.

COYNE, J. A., W. MEYERS, A. CRITTENDEN and P. SNIEGOWSKI, 1993 The fertility effects of pericentric inversions in *Drosophila melanogaster*. Genetics **134:** 487–496.

FEDERER, W. T., G. D. STEEL and B. WALLACE, 1967 Mathematical models for chromosomal inversions. Genetics **55:** 783–795.

FOSS, E. J. and F. W. STAHL, 1995 A test of a counting model for chiasma interference. Genetics **139:** 1201–1209.

FOSS, E., E. LANDE, F. W. STAHL and C. M. STAINBERG, 1993 Chiasma interference as a function of genetic distance. Genetics **133:** 681–691.

GRELL, R. F., 1962 A new model for secondary nondisjunction: the role of distributive pairing. Genetics **47:** 1737–1754.

GUBENKO, I. S., and M. B. EVGEN'EV, 1984 Cytological and linkage maps of *Drosophila virilis* chromosomes. Genetica **65:** 127–139.

HALDANE, J. B. S., 1919 The combination of linkage values and the calculation of distances between the loci of linked factors. J. Genet. **8:** 299–309.

HARTL, D. L., and E. R. LOZOVSKAYA, 1994 Genome evolution: between the nucleosome and the chromosome, pp. 579–592 in *Molecular Ecology and Evolution: Approaches and Implications* edited by D. SCHIERWATER, B. STREIT, G. P. WAGNER and R. DE SALLE. Birkhäuser, Basel.

HARTL, D. L. and E. R. LOZOVSKAYA, 1995 *The Drosophila Genome Map: A Practical Guide*. Springer-Verlag, Heildelberg, Germany.

HASSON, E., and W. F. EANES, 1996 Contrasting histories of three gene regions associated with *In(3L)Payne* of *Drosophila melanogaster*. Genetics **144:** 1565–1575.

HILLIKER, A. J. and A. CHOVNICK, 1981 Further observations on intragenic recombination in *Drosophila melanogaster*. Genet. Res. Camb. **38:** 281–296.

HILLIKER, A. J., S. H. CLARK and A. CHOVNICK, 1991 The effect of DNA sequence polymorphisms on intragenic recombination in the *rosy* locus of *Drosophila melanogaster*. Genetics **129:** 779–781.

HILLIKER, A. J., G. HARAUZ, A. G. REAUME, M. GRAY, S. H. CLARK *et al.*, 1994 Meiotic gene conversion tract length distribution within the *rosy* locus of *Drosophila melanogaster*. Genetics **137:** 1019–1026.

HUDSON, R. R., 1994 How can the low levels of Drosophila sequence variation in regions of the genome with low levels of recombination be explained. Proc. Natl. Acad. Sci. USA **91:** 6815–6818.

ISHII, K., and B. CHARLESWORTH, 1977 Associations between allozyme loci and gene arrangements due to hitchhiking effects of new inversions. Genet. Res. Camb. **30:** 93–106.

KARLIN, S., and U. LIBERMAN, 1994 Theoretical recombination processes incorporating interference effects. Theor. Pop. Biol. **46:** 198–231.

KREITMAN, M., and M. L. WAYNE, 1994 Organisation of genetic variation at the molecular level: lessons from *Drosophila*, pp. 157–183 in *Molecular Ecology and Evolution: Approaches and Implications*, edited by D. SCHIERWATER, B. STREIT, G. P. WAGNER and R. DE SALLE. Birkhäuser, Basel.

KRIMBAS, C. B., and J. R. POWELL, 1992 *Drosophila Inversion Polymorphism*. CRC Press, Inc. Boca Raton, FL.

LAMB, B. C., and S. HELMI, 1982 The extent to which gene conversion can change allele frequencies in populations. Genet. Res. Camb. **39:** 199–217.

LESLIE, J. F., and W. B. WATT, 1984 Postmeiotic segregation as a source of mosaics in diploid organisms. Genet. Res. Camb. **44:** 225–229.

LESLIE, J. F., and W. B. WATT, 1986 Some evolutionary consequences of the molecular recombination process. Trends Genet. **23:** 288–291.

LEVINE, R. P., 1956 Crossing over and inversions in coadapted systems. Am. Nat. **90:** 41–45.

LINDSLEY, D. L., and G. G. ZIMM, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego, CA.

LUCCHESI, J. C., and D. T. SUZUKI, 1968 The interchromosomal control of recombination. Annu. Rev. Genet. **2:** 53–86.

MATHER, K., 1938 Crossing-over. Biol. Rev. **13:** 252–292.

MERRIAM, J., M. ASHBURNER, D. L. HARTL and F. C. KAFATOS, 1991 Toward cloning and mapping the genome of Drosophila. Science **254:** 221–225.

MCPEEK, M. S., and T. P. SPEED, 1995 Modeling interference in genetic recombination. Genetics **139:** 1031–1044.

MORIYAMA, E. N., and J. R. POWELL, 1996 Intraspecific nuclear DNA variation in Drosophila. Mol. Biol. Evol. **13:** 261–277.

MORTIMER, R. K., and S. FOGEL, 1974 Genetical interference and gene conversion, pp. 263–275 in *Mechanisms in Recombination*, edited by R. F. GRELL. Plenum, New York.

NAVARRO, A., E. BETRÁN, C. ZAPATA and A. RUIZ, 1996 Dynamics of gametic disequilibria between loci linked to chromosome inversions: the recombination redistributing effect of inversions. Genet. Res. Camb. **67:** 67–76.

NOVITSKI, E., and G. BRAVER, 1954 An analysis of crossing over within a heterozygous inversion in *Drosophila melanogaster*. Genetics **39:** 197–209.

POPADIĆ, A., and W. ANDERSON, 1995 Evidence of gene conversion in the Amylase multigene family of *Drosophila pseudoobscura*. Mol. Biol. Evol. **12:** 564–572.

POPADIĆ, A., D. POPADIĆ and W. ANDERSON, 1995 Interchromosomal exchange of genetic information between gene arrangements on the third chromosome of *Drosophila pseudoobscura*. Mol. Biol. Evol. **12:** 938–943.

ROBERTS, P. A., 1962 Interchromosomal effects and the relation between crossing over and non-disjunction. Genetics **47:** 1691–1709.

ROBERTS, P. A., 1976 The genetics of chromosome aberration, pp. 67–184 in *The Genetics and Biology of Drosophila*, Vol. 1a, edited by M. ASHBURNER and E. NOVITSKI. Academic Press, London.

ROZAS, J., and M. AGUADÉ, 1993 Transfer of genetic information in the rp49 region of *Drosophila subobscura* between different chromosomal gene arrangements. Proc. Natl. Acad. Sci. USA **90:** 8083–8087.

ROZAS, J., and M. AGUADÉ, 1994 Gene conversion is involved in the transfer of genetic information between naturally occurring inversions of *Drosophila*. Proc. Natl. Acad. Sci. USA **91:** 11517–11521.

RUIZ, A., and J. ALBEROLA, 1983 Lack of evidence of embryonic mortality in the progeny of *Drosophila buzzatii* females heterozygous for included inversions. Dros. Inf. Syst. **59:** 106–107.

RUIZ, A., H. NAVEIRA and A. FONTDEVILA 1984 La historia evolutiva de *Drosophila buzzatii*. IV. Aspectos citogenéticos de su polimosfismo cromosómico. Genét. Ibér. **36:** 13–35.

SCHAEFFER, S. W., 1994 Molecular population genetics in *Drosophila pseudoobscura:* three future directions, pp. 29–45 in *Non-Neutral Evolution,* edited by B. GOLDING. Chapman & Hall, New York.

SINNOT, E. W., L. C. DUNN and T. DOBZHANSKY, 1962 *Principles of Genetics,* Ed. 5. McGraw-Hill, New York.

SPERLICH, D., and P. PFRIEM, 1986 Chromosomal polymorphism in natural and experimental populations, pp. 257–309 in *The Genetics and Biology of Drosophila,* Vol. 3e. Academic Press, London.

SPURWAY, H., and U. PHILIP, 1952 Genetics and cytology of *Drosophila subobscura* VII. Abnormal gene arrangements in element A. J. Genet. **51:** 198–215.

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

VAN VALEN, L., and R. LEVINS, 1968 The origins of inversion polymorphisms. Am. Nat. **102:** 5–23.

WESLEY, C. S., and W. F. EANES, 1994 Isolation and analysis of the breakpoint sequences of chromosome inversion *In(3L) Payne* in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **91:** 3132–3136.

ZHAO, H., T. P. SPEED and M. S. MCPEEK, 1995 Statistical analysis of crossover interference using the chi-square model. Genetics **139:** 1045–1056.

Communicating editor: D. CHARLESWORTH

APPENDIX A

Probabilities of i C_x's in the inverted segment and j C_x's in the proximal segment under the counting model: *Derivation of P_i:* The rigidity of the $\cdots C_x(C_0)_m \cdots$ sequence implies that there are m, and only m, C_0's between any two consecutive C_x's. So at least $(m + 2)$ C events inside the inverted region are required for a double crossover to take place. Let k be the number of C events inside the inversion, then $k \geq 6$ is needed when $m = 4$. In addition, the λ_I values that we are considering (see Parameter values) allow us to ignore, in general, cases in which there are more than $2(m + 1)$ C's (that is, cases in which $k > 10$ when $m = 4$), because of their very small probability. Therefore, the range of k values to consider is $m + 2 \leq k \leq 2(m + 1)$. For this range, the probability of 2 C_x's when there are kC's in the inverted region is

$$P(2C_x\text{'s In}|k\ C\text{'s}) = \frac{k - (m + 1)}{m + 1}, \qquad \text{(A1)}$$

and the probability of a number k of C events inside the inverted region of a given tetrad is the Poisson probability of k events with parameter λ_I:

$$P(kC\text{'s}) = \frac{e^{-\lambda_I}\lambda_I^k}{k!}. \qquad \text{(A2)}$$

Therefore,

$$P_2 = P(kC\text{'s})\,P(2C_x\text{'s In}|k\ C\text{'s}) = \sum_{k=m+2}^{2m+2}\left(\frac{e^{-\lambda_I}\lambda_I^k}{k!}\,\frac{k - (m + 1)}{m + 1}\right). \qquad \text{(A3)}$$

Analogous reasoning allows us to obtain P_0 and P_1 under the counting model (Table A1).

Derivation of Pp_j: Under the counting model, the rigidity of the $\cdots C_x(C_0)_m C_x \cdots$ sequence makes Pp_j dependent on the number of C_x events inside the inversion. That is, the events inside the inverted region influence the events outside it as follows:

1. The number, q_P, of C_0's between the proximal inversion breakpoint and the nearest C_x event in the distal direction define $m + 1$ possible different patterns inside the inverted region ($0 \leq q_P \leq m$).
2. Different numbers of C_x events inside the inverted region (i) imply different probabilities of q_P. The probability of q_P and i C_x's in the inverted region is

$$P(q_P \cap 0C_x\text{'s In}) = \sum_{k=0}^{q_P}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\,\frac{1}{m + 1}$$

$$P(q_P \cap 1C_x\text{'s In}) = \sum_{k=q_P+1}^{m+1+q_P}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\,\frac{1}{m + 1}$$

$$P(q_P \cap 2C_x\text{'s In}) = \sum_{k=m+2+q_P}^{2m+2+q_P}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\,\frac{1}{m + 1}. \qquad \text{(A4)}$$

3. Different values of q_P imply different probabilities of zero, one and two C_x's in the proximal zone (Pp_j values):

$$P(0C_x\text{'s Prox}|q_P) = \sum_{n=0}^{m-q_P}\frac{e^{-\lambda_P}\lambda_P^n}{n!}$$

$$P(1C_x\text{'s Prox}|q_P) = \sum_{n=(m+1)-q_P}^{(2m+1)-q_P}\frac{e^{-\lambda_P}\lambda_P^n}{n!}$$

$$P(2C_x\text{'s Prox}|q_P) = \sum_{n=(2m+2)-q_P}^{(3m+2)-q_P}\frac{e^{-\lambda_P}\lambda_P^n}{n!}, \qquad \text{(A5)}$$

where n is the number of C events in the proximal segment.

Combining (A4) and (A5) we can easily obtain P_{ij}:

$$P_{ij} = P(iC_x\text{'s In} \cap jC_x\text{'s Prox}) = \sum_{q_P=0}^{m} [P(q_P \cap iC_x\text{'s In})P(jC_x\text{'s Prox}|q_P)]. \quad \text{(A6)}$$

APPENDIX B

Recombination inside the inverted region under the counting model: Two markers A and B inside the inverted region divide it into three different segments, A, A-B, and B, whose lengths are α, $(\beta - \alpha)$ and $(1 - \beta)$. To know the probability that A and B recombine we must follow three steps.

1. Given k C's and two C_x's inside the inverted region, the number, q_P, of C_0's between the proximal inversion breakpoint and the nearest distal C_x event can take $k - (m + 2) + 1$ different values $[0 \leq q_P \leq k - (m + 2)]$. These values are equally probable because of the stationarity of the model.
2. For every value of q_P the number, s, of C's in segment A must be $0 \leq s \leq q_P + 1 + m$. This range of values allows for zero or one C_x events in segment A. If the range was larger, recombination would be impossible because the two C_x's would be in segment A.

Every s value has a probability of $(\alpha)^s$ and there are $\binom{k}{s}$ possible ways in which s C's from a group of k C's can be grouped. Of course, s C's in segment A imply $(k - s)$ C's in the remainder of the inverted region.

3. The possible number of C events in segment A-B, r, will have different ranges depending on the presence or absence of a C_x event in segment A. If $s \leq q_P$, there is no C_x in A, so that one of the two C_x's must be in segment A-B and the other one in segment B for recombination to take place. In this case $(q_P - s + 1 \leq r \leq q_P - s + 1 + m)$. If $s > q_P$, there is a C_x event in segment A, so that the other C_x must be in segment A-B. In this case $(m + 1 - s + q_P + 1 \leq r \leq k - s)$. Every r value has a probability of $(\beta - \alpha)^r(1 - \beta)^{k-s-r}$ and there are $\binom{k-s}{r}$ possible ways in which r C's from a group of $k - s$ C's can be grouped.

Adding up the probabilities of the events that will cause recombination, as described in 1–3, we obtain an expression for the probability of recombination between markers A and B given k C's and two C_x's in the inverted region of a given gamete.

$$R_{Cx}((\alpha, \beta)|kC\text{'s} \cap 2C_x\text{'s In}) = \tfrac{1}{2}\sum_{q_P=0}^{k-(m+2)}\left[\frac{1}{k-(m+2)+1}\left\{\sum_{s=0}^{q_P}\left(\binom{k}{s}\alpha^s \sum_{r=q_P-s+1}^{q_P-s+1+m}\left(\binom{k-s}{r}(\beta-\alpha)^r(1-\beta)^{k-s-r}\right)\right)\right.\right.$$
$$\left.\left. + \sum_{s=q_P+1}^{q_P+1+m}\left(\binom{k}{s}\alpha^s \sum_{r=(m+1)-(s-q_P-1)}^{k-s}\left(\binom{k-s}{r}\times(\beta-\alpha)^r(1-\beta)^{k-s-r}\right)\right)\right\}\right]. \quad \text{(A7)}$$

Multiplying this expression by the probabilities of k C's and two C_x's and by the probability of viable gametes with two C_x's (ρ_2) and dividing by the proportion of viable gametes, we derive the recombination rate of any two markers in a given position of the inverted region:

$$R_{Cx}(\alpha, \beta) = \frac{\rho_2}{P_v}\sum_{k=m+2}^{2m+2} P(k\ C\text{'s})P(2C_x\text{'s In}|k\ C\text{'s}) \times R_{Cx}((\alpha, \beta)|k\ C\text{'s} \cap 2C_x\text{'s In}). \quad \text{(A8)}$$

APPENDIX C

The average crossover tract length in the counting model: Let us consider the average length of the double crossover tract, relative to that of the inversion. When there are k randomly distributed C's in the inverted region, it is divided into $k + 1$ intervals whose average length, always relative to that of the inversion, will be $1/(k + 1)$. There are always m C's between two consecutive C_x's, thus there are always $m + 1$ intervals

TABLE A1

Probabilities of zero, one and two C_x events in the inverted region

Probability of i C_x's	Poisson model	Counting model
P_0	$e^{-\frac{\lambda_I}{m+1}}$	$\sum_{k=0}^{m}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\frac{(m+1)-k}{m+1}$
P_1	$e^{-\frac{\lambda_I}{m+1}}\frac{\lambda_I}{m+1}$	$\sum_{k=1}^{m}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\frac{k}{m+1} + \sum_{k=m+1}^{2m+1}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\frac{2(m+1)-k}{m+1}$
P_2	$\frac{1}{2}e^{-\frac{\lambda_I}{m+1}}\left(\frac{\lambda_I}{m+1}\right)^2$	$\sum_{k=m+2}^{2m+2}\frac{e^{-\lambda_I}\lambda_I^k}{k!}\frac{k-(m+1)}{m+1}$

between two C_x's and the mean length of a double crossover tract ($\bar{L}_{CxT}$) given m and k is

$$\bar{L}_{CxT}(m, k) = \frac{m+1}{k+1}. \quad \text{(A9)}$$

The average length of a double crossing-over tract is calculated by weighting the lengths obtained in (A9) by the Poisson probability of k C's inside the inverted region and by the probability of two C_x's when there are k C's. For every value of λ_I and m, this length is

$$\bar{L}_{CxT}(m, \lambda_I) = \frac{\sum_{k=m+2}^{2m+2} [P(kC\text{'s}) P(2C_x\text{'s In} | kC\text{'s}) \bar{L}_{CxT}(m, k)]}{\sum_{k=m+2}^{2m+2} [P(2C_x\text{'s In}) P(2C_x\text{'s In} | kC\text{'s})]}$$

$$= \frac{\sum_{k=m+2}^{2m+2} \left[\left(\frac{\lambda_I^k}{k!} \right) \left(\frac{k-(m+1)}{k+1} \right) \right]}{\sum_{k=m+2}^{2m+2} \left[\left(\frac{\lambda_I^k}{k!} \right) \left(\frac{k-(m+1)}{m+1} \right) \right]}. \quad \text{(A10)}$$

This lengths are larger under the counting model than under the noninterference model, in which the average double crossover tract is $^1/_3$ the length of the inversion [we get this value from formula (A9) by making $m = 0$ and $k = 2$].

APPENDIX D

Recombination outside the inverted region under the counting model: Just as in APPENDIX B, two markers A and B in any noninverted region divide it into three different segments, A, A-B and B, whose lengths are α, $(\beta - \alpha)$ and $(1 - \beta)$. A reasoning completely analogous to that of APPENDIX B allows us to obtain expressions for the recombination in the proximal and distal segments.

Proximal segment: Different values of q_P imply different recombination rates between two markers A and B in the proximal region because they imply different probabilities of one or two C_x's in that zone. Let us first concentrate in the case of 1 crossover in the proximal zone.

When there is a single C_x event in the proximal zone it must lie in the A-B segment for recombination to be possible. Also, n, the number of C's in the proximal segment, must be $m + 1 - q_P \le n \le 2m + 1 - q_P$. For every q_P value, there must be s C's in segment A ($0 \le s \le m - q_P$). Again, every s value has a probability of $(\alpha)^s$ and there are $\binom{n}{s}$ possible ways in which s C's from a group of n C's can be grouped. Of course, s C's in segment A imply $(n - s)$ C's in the remainder of the proximal region. Finally, the number of C events in segment A-B, r must be $m + 1 - q_P - s \le r \le n - s$. Every r value has a probability of $(\beta - \alpha)^r (1 - \beta)^{k-s-r}$ and there are $\binom{n-s}{r}$ possible ways in which r C's from a group of $n - s$ C's can be grouped. These considerations lead us to

$$P(1C_x\text{'s Prox} \cap R_{Cx}(\alpha, \beta) | q_P) = {}^1/_2 \sum_{n=m+1-q_P}^{2m+1-q_P} \left[\frac{e^{-\lambda_P} \lambda_P^n}{n!} \sum_{s=0}^{m-q_P} \left(\binom{n}{s} \alpha^s \sum_{r=m+1-q_P-s}^{n-s} \left(\binom{n-s}{r} \times (\beta - \alpha)^r (1 - \beta)^{n-s-r} \right) \right) \right]. \quad \text{(A11)}$$

When there are two C_x's in the proximal segment, the number, n, of C's in the proximal zone given q_P must lie in the range $2m + 2 - q_P \le n \le 3m + 2 - q_P$. Also, the number, s, of C's in segment A must be $0 \le s \le 2m + 1 - q_P$. This range of values allows for zero or one C_x events in segment A. The possible number of C events in segment A-B, r, will have different ranges depending on the presence of absence of a C_x event in segment A. If $s \le m - q_P$, there is no C_x in A, so that at least one of the two C_x's must be in segment *A-B* for recombination to take place. In this case r must be $m + 1 - q_P - s \le r \le n - s$. If $s > m - q_P$, there is a C_x in segment A, so that the other C_x must be in segment A-B for recombination to be possible. In this case r must be $2m + 2 - s - q_P \le r \le n - s$. Adding up all these probabilities, we obtain

$$P(2C_x\text{'s Prox} \cap R_{Cx}(\alpha, \beta) | q_P) = {}^1/_2 \sum_{n=2m+2-q_P}^{3m+2-q_P} \left[\frac{e^{-\lambda_P} \lambda_P^n}{n!} \left\{ \sum_{s=0}^{m-q_P} \left(\binom{n}{s} \alpha^s \sum_{r=m+1-q_P-s}^{n-s} \left(\binom{n-s}{r} (\beta - \alpha)^r (1 - \beta)^{n-s-r} \right) \right) + \sum_{s=m+1-q_P}^{2m+1-q_P} \left(\binom{n}{s} \alpha^s \sum_{r=2m+2-q_P-s}^{n-s} \left(\binom{n-s}{r} \times (\beta - \alpha)^r (1 - \beta)^{n-s-r} \right) \right) \right\} \right], \quad \text{(A12)}$$

which allows us to derive

$$R_{Cx}(\alpha, \beta) = \left(\sum_{q_P=0}^{m} \sum_{i=0}^{2} [P(q_P \cap i\ C_x\text{'s In}) \times (P(1C_x\text{'s Prox} \cap R_{Cx}(\alpha, \beta) | q_P) \rho_{i1} + P(2C_x\text{'s Prox} \cap R_{Cx}(\alpha, \beta) | q_P) \rho_{i2})] \right) \Big/ \sum_{i=0}^{2} \sum_{j=2}^{2} (P_{ij} \rho_{ij}). \quad \text{(A13)}$$

Distal segment: In this case we can define q_D as the number of C_0's between the distal inversion breakpoint and the nearest C_x event in the proximal direction. Just as above, different values of q_D imply different recombination rates between two markers, A and B, in the distal region. In addition, in this case we can establish a relationship between q_D and q_P that will allow us to calculate

the average viability given q_P, k, m, and i, the number of C_x's in the inverted segment.

$$q_P = k + (m - q_D) \quad \text{when} \quad i = 0$$
$$q_P = k - (q_D + 1) \quad \text{when} \quad i = 1$$
$$q_P = k - (m + 2 + q_D) \quad \text{when} \quad i = 2$$

From these expressions we can readily obtain the joint probabilities of any given q_P and i values and of balanced gametes.

$$P(q_D \cap 0C_x\text{'s In} \cap \text{Balan}) = \sum_{k=0}^{q_D} \left[\frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{1}{m+1} \sum_{j=0}^{2} (P(jC_x\text{'s Prox} | q_P)\rho_{0j}) \right],$$

$$P(q_D \cap 1C_x\text{'s In} \cap \text{Balan}) = \sum_{k=q_D+1}^{m+1+q_D} \left[\frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{1}{m+1} \sum_{j=0}^{2} (P(jC_x\text{'s Prox} | q_P)\rho_{1j}) \right],$$

$$P(q_D \cap 2C_x\text{'s In} \cap \text{Balan}) = \sum_{k=m+2+q_D}^{2m+2+q_D} \left[\frac{e^{-\lambda_I}\lambda_I^k}{k!} \frac{1}{m+1} \sum_{j=0}^{2} (P(jC_x\text{'s Prox} | q_P)\rho_{2j}) \right]. \quad \text{(A14)}$$

For every value of q_D, the probability of recombination between two markers in the distal region is as follows:

$$R_{Cx}((\alpha, \beta) | q_D) = \tfrac{1}{2} \sum_{n=0}^{\infty} \frac{e^{-\lambda_D\alpha}(\lambda_D\alpha)^n}{n!} \times \left(1 - \sum_{j=0}^{\eta-1} \frac{e^{-\lambda_D(\beta-\alpha)}(\lambda_D(\beta-\alpha))^j}{j!}\right), \quad \text{(A15)}$$

where λ_D is the average number of C's in the distal segment; $\eta = \{(1 + \text{Int}[(n + q_D)/(m + 1)])(m + 1) - (n + q_D)\}$; Int$[x]$ is the greatest integer less than or equal to x; and α and β are the distances from the markers to the nearest breakpoint ($0 < \alpha < \beta < 1$) relative to the length of the distal segment.

Finally, combining (A14) and (A15), we can easily obtain a formula for the recombination rate between two markers in the distal region:

$$R_{Cx}(\alpha, \beta) = \frac{\sum_{i=0}^{2} \sum_{q_D=0}^{m} [P(q_D \cap iC_x\text{'s In} \cap \text{Balan}) R_{Cx}((\alpha, \beta) | q_D)]}{\sum_{i=0}^{2} \sum_{j=0}^{2} (P_{ij}\rho_{ij})}. \quad \text{(A16)}$$