Recombination Rate Predicts Inversion Size in Diptera

Mario Cáceres, Antonio Barbadilla and Alfredo Ruiz

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

Manuscript received November 15, 1998
Accepted for publication May 3, 1999

ABSTRACT

Most species of the Drosophila genus and other Diptera are polymorphic for paracentric inversions. A common observation is that successful inversions are of intermediate size. We test here the hypothesis that the selected property is the recombination length of inversions, not their physical length. If so, physical length of successful inversions should be negatively correlated with recombination rate across species. This prediction was tested by a comprehensive statistical analysis of inversion size and recombination map length in 12 Diptera species for which appropriate data are available. We found that (1) there is a wide variation in recombination map length among species; (2) physical length of successful inversions varies greatly among species and is inversely correlated with the species recombination map length; and (3) neither the among-species variation in inversion length nor the correlation are observed in unsuccessful inversions. The clear differences between successful and unsuccessful inversions point to natural selection as the most likely explanation for our results. Presumably the selective advantage of an inversion increases with its length, but so does its detrimental effect on fertility due to double crossovers. Our analysis provides the strongest and most extensive evidence in favor of the notion that the adaptive value of inversions stems from their effect on recombination.

OVER 60 years of research on Drosophila inversion polymorphism have yielded a vast amount of empirical information. However, besides the notion that some kind of balancing selection is responsible for its maintenance in natural populations (Sperlich and Pfriem 1986; Krimbas and Powell 1992; Powell 1997), very few genus-wide generalizations have been made to settle which property, if any, is selected for in evolutionarily successful inversions. A consistent relationship has been found between size and frequency of inversions in different species: rare or endemic (unsuccessful) inversions are usually small, while polymorphic or fixed (successful) inversions are predominantly medium-sized (Olvera *et al.* 1979; Brehm and Krimbas 1991; Krimbas 1992; Cáceres *et al.* 1997). This suggests that natural selection discriminates among inversions of different sizes, likely favoring those of intermediate physical length (Krimbas and Powell 1992; Powell 1997).

Because the most conspicuous effect of paracentric inversions is the substantial reduction of recombination within the inverted chromosomal segment in heterozygous individuals (Sturtevant and Beadle 1936; Navarro *et al.* 1997), it appears that natural selection operates upon polymorphic inversions through their effect on recombination (Kojima and Schaffer 1964; Dobzhansky 1970; Charlesworth and Charlesworth 1973; Charlesworth 1974; Álvarez and Zapata 1997). Accordingly, selection would act on recombination length of inversions, not physical length, and the observed relationship between evolutionary success and physical length of inversions would merely be a reflection of the actual correlation between physical and recombination length. Interspecific comparisons are crucial to test this hypothesis. Drosophila species seem to differ both in physical length of inversions (Cáceres *et al.* 1997) and in their recombination map length (True *et al.* 1996). If recombination length of inversions, rather than physical length, were the property selected for, differences in physical length of inversions among species would be accounted for by differences in recombination map length among species.

On the basis of the observed intermediate size of successful inversions, a simple selective model can be considered, by which natural selection favors a constant optimal recombination length (γ) of inversions across species (a less constrained and more realistic model would assume an optimal interval, but the conclusions would not differ qualitatively from those of this simpler model). This would be the net result of the selective advantage gained by the reduction of recombination and the detrimental effect on fertility of the inversions (see discussion). If species vary in their genome recombination map length (G), selection for optimal recombination length of inversions will result in variable optimal physical lengths (L) among species. Let L_i and G_i be, respectively, the optimal physical length of inversions and the recombination map length of species i, then

$$L_i = \gamma \frac{1}{G_i}.$$

Corresponding author: Mario Cáceres, Departament de Genètica i de Microbiologia, Facultat de Ciències, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain.
E-mail: mariocs@cc.uab.es

Taking physical length of inversions in relative units, as a proportion of the total physical length of the genome, makes our model independent of the genome length of each species, which could vary among them. Thus, this model predicts a positive linear correlation across species between the relative physical length of selectively favored or successful inversions (that should be distributed around the species optimal physical length) and $1/G_i$. Conversely, the physical length of unsuccessful inversions, *i.e.*, those with zero or low selective advantage, is expected to be unrelated to the species recombination map length. We can test this model via its predictions, with unsuccessful inversions being used as the control data set.

In this article we carry out a comprehensive statistical analysis of inversion size and recombination map length in Diptera species for which appropriate data are available. The results corroborate the above predictions. We have found (1) that recombination map length varies greatly across species; (2) significant differences in length among species in successful inversions, but not in unsuccessful inversions; and (3) contrasting results, depending on the class of inversion considered, for the relationship between the two variables. In successful inversions, physical length of inversions and species recombination map length are strongly correlated, but this is not so in unsuccessful inversions.

MATERIALS AND METHODS

An exhaustive search of relevant data was made in the extant literature. We studied Diptera species with (1) detailed cytological maps of the salivary gland chromosomes and unambiguous descriptions of chromosomal inversions and (2) recombination (linkage) maps of at least one chromosome. Ten Drosophila species (*Drosophila ananassae*, *D. buzzatii*, *D. funebris*, *D. hydei*, *D. mediopunctata*, *D. melanogaster*, *D. persimilis*, *D. prosaltans*, *D. pseudoobscura*, and *D. subobscura*) and two Anopheles species (*A. gambiae* and *A. stephensi*) met our requirements.

Physical length of inversions was measured as the distance between the two breakpoints relative to the total physical length of the genome (the euchromatic portion represented in the polytene chromosome maps) in percentage. For each inversion, the length was obtained by comparison with the arrangement from which it originated (Olvera *et al.* 1979), assuming that the major determinant of the fate of an inversion is its interaction with the parental chromosome (because this will be the one with which any new chromosomal arrangement will be combined as heterozygote most frequently). In *D. buzzatii*, *D. pseudoobscura*, and *D. subobscura* the data were directly taken from previous studies (Olvera *et al.* 1979; Krimbas 1992; Cáceres *et al.* 1997). Because they are scarce and have distinct genetic dynamics, neither pericentric nor sex-ratio inversions were included in the data set. According to their evolutionary success, paracentric inversions were classified as "successful" and "unsuccessful," and both groups were analyzed separately. An *a priori* criterion, based on the geographical distribution and frequency of the different chromosomal arrangements, was adopted to maximize the proportion of selectively favored inversions in the successful class. Only those inversions described previously as common and widespread (Olvera *et al.* 1979; Moore and Taylor 1986; Krimbas 1992; Lemeunier and Aulard 1992) or that have been reported as present in a significant fraction of the species distribution, namely at least 25% of the sampled localities (Cáceres *et al.* 1997), were considered successful. All other inversions, which are rare or restricted to a few localities, were considered unsuccessful. The 207 inversions used in the study are listed in Table 1.

Recombination map data available for the 12 species are given in Table 2. When necessary, the published linkage map of each chromosome was corrected with the widely used Kosambi mapping function (Crow 1990). Then, its length was multiplied by $(n + 1)/(n - 1)$ to account for the different number of markers (n) per chromosome (Chakravarti *et al.* 1991). To estimate the total recombination map length of a given species, G_i, the lengths of all chromosomes in the haploid set were summed. When no data were available for a given chromosome, its recombination length was inferred from the recombination lengths of other chromosome(s) for which data were available by assuming proportionality with its physical length (in the polytene chromosome maps). To test this assumption of proportionality, we computed the correlation between the relative physical length and the relative recombination length of chromosomes. We used the recombination data of species with at least two mapped chromosomes (8 out of 12 species; Table 2). Because the variables take relative values, only $a - 1$ of the a pairs of values of a species are independent, and we omitted one data pair (chosen at random) of each of the 8 species. The Pearson r value was 0.76 (d.f. 13; $P = 0.0010$), which supports our estimation procedure. In Drosophila, recombination is limited to females, while in *Anopheles gambiae* and *A. stephensi*, males and females have similar recombination rates (Parvez *et al.* 1985; Zheng *et al.* 1996). Accordingly, to make recombination values directly comparable between different chromosomes and different genera, the estimated G_i were multiplied by $\frac{1}{2}$ for Drosophila autosomal inversions and by $\frac{2}{3}$ for X-linked inversions (Begun and Aquadro 1992). However, the different recombination value for the X chromosome with regard to the autosomes, as well as the absence of X-chromosome inversions in 9 of the 12 species, renders the statistical analysis of inversion length and recombination unbalanced and makes it intractable. Therefore, X-linked inversions (4 successful and 14 unsuccessful) were omitted from the correlation analysis, although it should be noted that their behavior is very similar to that of autosomal inversions (see Figure 1).

RESULTS

While the two Anopheles species have a relatively small recombination map length, there is remarkable variation in recombination map length among Drosophila species, from 285.4 cM in *D. prosaltans* to 1007.6 cM in *D. subobscura* (Table 3). Mean length of successful inversions also varies greatly among species, from 3.64% in *D. mediopunctata* to 11.45% in *D. prosaltans*, and the differences are statistically significant as shown by the ANOVA ($F = 5.28$; d.f. 11, 70; $P < 0.0001$) or the nonparametric Kruskal-Wallis test ($H = 27.92$; d.f. 11; $P = 0.0033$). To test the effect of species recombination map length on physical length of successful inversions, we performed an analysis of variance and regression, where the variation among species in physical length of inversions is partitioned into linear and nonlinear components (Table 4). The F-test showed a very significant

TABLE 1

Inversions used in the study classified according to their evolutionary success

Species	Successful inversions	Unsuccessful inversions	Reference[a]
D. ananassae	*2LA, 3LA, 3RA*	*2LB, 2LC, 2RA, 3RD*	(1, 2)
D. buzzatii	*2j, 2z^3, 2q^7, 4s*	*2y^3, 2c^9, 2d^9, 2e^9, 2g^9, 2h^9, 2i^9, 2j^9, 2r^9, 2s^9, 3j^2, 5c^2*	(3, 4)
D. funebris	*II-1, II-2, II-3, III-1, IV-1*		(5)
D. hydei	*2a^2*		(6)
D. mediopunctata	*2 AB, 2 AC*		(7)
D. melanogaster	*(2L)NS, (2L)t, (2R)NS, (3L)M, (3L)P, (3R)C, (3R)K, (3R)M, (3R)Mo, (3R)P*	*(1)12A;18D, (1)16D;18D, (2L)A, (2L)W, (2L)22A;26B, (2R)NC, (2R)O, (2R)49B;56A, (3L)L, (3L)Y, (3L)62D;68A*	(8)
D. persimilis	*CO, KL, MD, RD, SE, WT*	*HU, MA, MR, NA, TP, TU, VI, WA, WE*	(9–13)
D. prosaltans	*PXLa, PXLd, PIILa, PIIRa*	*PXLb, PXLc, PXLe, PXLf, PXLg, PXRa, PIIRb, PIIRc, PIIIa, PIIIb*	(14, 15)
D. pseudoobscura	*AR, CH, CU, EP, HI, OA, OL, PP, SC, TA, TL*	*AF, AM, BE, CC, EB, FC, HY, IZ, MA, MI, OZ, PA, PI, PO, SA, SB, SJ, SO, TE, TH, UR, VA, ZI*	(16)
D. subobscura	*A$_1$, A$_2$, J$_1$, J$_3$, J$_4$, U$_1$, U$_2$, U$_3$, U$_4$, U$_6$, U$_7$, U$_8$, E$_1$, E$_2$, E$_3$, E$_4$, E$_8$, E$_9$, E$_{12}$, O$_1$, O$_2$, O$_3$, O$_4$, O$_5$, O$_6$, O$_7$, O$_8$, O$_{22}$*	*A$_3$, A$_4$, A$_5$, A$_6$, A$_7$, A$_8$, J$_5$, U$_5$, U$_9$, U$_{10}$, U$_{11}$, U$_{12}$, E$_5$, E$_{13}$, E$_{14}$, E$_{15}$, E$_{16}$, E$_{17}$, E$_{18}$, E$_{19}$, E$_{20}$, O$_9$, O$_{10}$, O$_{11}$, O$_{12}$, O$_{13}$, O$_{14}$, O$_{15}$, O$_{16}$, O$_{17}$, O$_{18}$, O$_{19}$, O$_{20}$, O$_{21}$, O$_{23}$, O$_{24}$, O$_a$*	(17)
A. gambiae	*2La, 2Rb, 2Rc, 2Rd, 2Rj, 2Rk, 2Ru*	*2Rt*	(18, 19)
A. stephensi	*2Rb*	*2Rc, 2Rd, 2Re, 2Rf, 2Lc, 2Ld, 3Ra, 3Rb, 3Rc, 3Lb, 3Lc, 3Ld, 3Le, 3Lf, 3Lg, 3Lh, 3Li, 3Lj*	(20, 21)

[a] References: (1) Futch (1996); (2) Hinton and Downs (1975); (3) Ruiz *et al.* (1984); (4) Barker *et al.* (1985); (5) Dubinin and Tiniakov (1946); (6) Wasserman (1962); (7) Kastritsis (1966); (8) Lemeunier and Aulard (1992); (9) Dobzhanski and Epling (1944); (10) Dobzhanski (1948); (11) Spiess (1950); (12) Spiess (1965); (13) Beckenbach (1986); (14) Bicudo (1973); (15) Bicudo *et al.* (1978); (16) Olvera *et al.* (1979); (17) Krimbas (1992); (18) Coluzzi and Sabatini (1967); (19) Coluzzi *et al.* (1979); (20) Coluzzi *et al.* (1973); (21) Mahmood and Sakai (1984).

correlation between physical length of inversions and $1/G_i$. However, the variances within species were not homogeneous, even after log transformation. Thus, we also performed a resampling test as follows. A random sample with replacement is drawn of the inversion length data (78 inversions in our case). Let N_i be the number of inversions analyzed in species *i*. N_1, N_2, . . . , N_{12} sampled values are assigned randomly to species 1, 2, . . . , 12. To the whole random sample, the same analysis of variance and regression shown in Table 4 is applied, and the *F* value computed. The *F* distribution of 10,000 replicates obtained with this procedure is compared with the *F* empirical value, and its significance is estimated. This analysis corroborated the parametric significant probabilities (Table 4). As shown in Figure 1A, the average physical length of successful inversions decreases as species recombination map length increases.

Clearly contrasting results were obtained for unsuccessful inversions. First, they are smaller than successful inversions (mean length 4.72% *vs.* 6.40%) as shown by a *t*-test ($t = 4.10$; d.f. 205; $P < 0.0001$) or a sign test ($P = 0.0391$). Second, the weighted mean variance in inversion length within species (computed dividing each inversion length value by the species average of its class to take out the length differences) is larger for the unsuccessful inversions than for the successful ones ($F = 2.33$; d.f. 70, 116; $P = 0.0001$). Third, no significant differences in mean length of unsuccessful inversions among species are found (ANOVA: $F = 1.31$; d.f. 8, 116; $P = 0.25$; Kruskal-Wallis test: $H = 13.16$; d.f. 8; $P = 0.11$). Finally, the correlation between physical length of inversions and species recombination map length is not significant (Table 4). Therefore, there is no relationship between physical length of unsuccessful inversions and recombination map length (Figure 1B).

So far, we have considered the mean values of species as independent data points. This implicitly assumes that polymorphic inversions are not inherited across species but arise *de novo* in each species (see discussion). It could be argued, nevertheless, that closely related species share their trait values because of common ancestry (Felsenstein 1985; Harvey and Pagel 1991). Figure 2 shows the phylogenetic relationships of the 12 species. To test for a phylogenetic clustering of recombination values, we performed a nested ANOVA at four taxo-

TABLE 2

Recombination data for the 12 species of Diptera

Species	Chromosome	Number of markers (n)	Linkage map (cM)	Kosambi's correction	$(n + 1)/(n - 1)$ correction	Reference[a]
D. ananassae	X	31	106.9	107.6	114.8	(1)
	2	27	113.1	114.1	122.9	(1)
	3	26	103.5	104.1	112.4	(1)
D. buzzatii	X	14	108.9	108.9	125.7	(2)
	2	11	138.5	138.5	166.2	(2)
D. funebris	X	13	158.0	172.2	200.9	(3)
D. hydei	X	25	115.7	117.6	127.4	(4)
D. mediopunctata	X	6	89.0	94.4	132.2	(5)
D. melanogaster	X	868	73.1	73.1	73.3	(6)
	2	624	110.0	110.0	110.4	(6)
	3	573	110.9	110.9	111.3	(6)
D. persimilis	X	11	164.5	191.6	229.9	(7)
D. prosaltans	X	14	100.0	105.5	121.7	(8)
	2	7	59.0	61.3	81.7	(8)
	3	5	52.0	54.7	82.0	(8)
D. pseudoobscura	X	34	228.2	228.2	242.0	(9)
	2	11	101.3	113.8	136.6	(10)
	3	12	68.0	69.4	82.0	(10)
	4	7	69.2	72.4	96.5	(10)
D. subobscura	A	14	150.0	167.6	193.4	(11)
	O	17	228.3	226.1	254.4	(12)
	E	6	107.4	108.4	151.8	(12)
A. gambiae	X	46	48.9	48.9	51.1	(13)
	2	57	72.4	72.4	75.0	(13)
	3	28	93.7	93.7	100.6	(13)
A. stephensi	2	7	98.4	128.4	171.2	(14)
	3	6	77.2	89.1	124.7	(14)

[a] References: (1) Tobari (1993); (2) Schafer *et al.* (1993); (3) Perje (1955); (4) Hess (1976); (5) Sampaio *et al.* (1996); (6) Lindsley and Zimm (1992); (7) Beers (1937); (8) Spassky *et al.* (1950); (9) Orr (1995); (10) Anderson (1990); (11) Spurway (1945); (12) Loukas *et al.* (1979); (13) Zheng *et al.* (1996); (14) Parvez *et al.* (1985).

nomic levels: genus, subgenus, group, and species (Bell 1989; Harvey and Pagel 1991). Most of the variation was due to the "between groups within subgenera" level, which explained 68% of the total variation (Table 5). A nested ANOVA was also applied to the physical length of inversions. For successful inversions, we found again that the group level explained most of the variation in inversion length (50%). On the other hand, for unsuccessful inversions, nearly all the length variation (~96%) was found within species and none in the other taxonomic levels (Table 5). To account for the possible phylogenetic effect on the observed correlation between physical length of successful inversions and recombination map length, we repeated the same analysis of variance and regression of Table 4 on the eight species groups. Because little additional variance remains at higher levels for both variables, the groups can be considered as statistically independent. As in Table 4, the F value for the regression with $1/G_i$ (19.92) was highly significant either by the parametric (d.f. 1, 6; $P = 0.0043$) or the resampling test ($P = 0.0060$). We also applied Felsenstein's (1985) independent contrasts method to our data (using the CAIC program of Purvis and Rambaut 1995). The method requires the prior knowledge of the phylogeny and the branching times of species. We used the time estimates available in the literature either for our species or for closely related species (Figure 2; Russo *et al.* 1995; Powell 1997). The correlation between physical length of successful inversions and $1/G_i$ for the independent contrasts was 0.47, which is marginally significant (d.f. 9; $P = 0.07$).

DISCUSSION

Our analysis shows that the physical length of evolutionarily successful inversions differs among species and that there is a negative correlation between inversion length and species recombination map length (Figure 1A) that explains a sizable part of inversion length variance among species (61%). For both variables, species recombination map length and physical length of successful inversions, the group level accounts for a significant proportion of the variance (Table 5). Hence, two alternative explanations are possible. First, the correla-

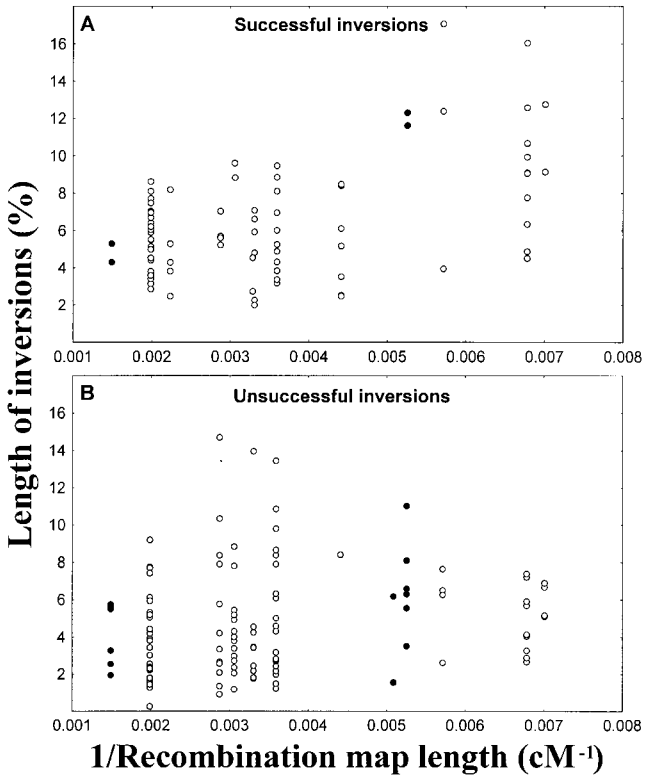

Figure 1.—Scatterplot of physical length of inversions (in percentage of the total physical length of the genome) *vs.* the inverse of the species recombination map length. (A) Successful inversions ($N = 82$) and (B) unsuccessful inversions ($N = 125$). X-linked inversions (4 successful and 14 unsuccessful) are shown as solid circles and display similar behavior as autosomal inversions (open circles).

tion could have a purely phylogenetic origin, *i.e.*, species could share both characters due to common ancestry. In this case, it should disappear when appropriate comparative methods that take into account the phylogeny are used. Second, there could be a real causal relationship between recombination map length and inversion length, which would result in a pseudophylogenetic correlation as a consequence of the phylogenetic clustering of recombination values. In this second case, the correlation could diminish but should not disappear entirely when comparative methods are used.

Two different arguments allow us to exclude a phylogenetic explanation for the observed correlation. First, when we obviate the phylogenetic effect, considering just the eight independent species groups (Bell 1989; Harvey and Pagel 1991), the same highly significant correlation is still found. Likewise, despite the reduction of statistical power, the independent contrasts method (Felsenstein 1985) also yields a marginally significant correlation. Second, the phylogenetic explanation assumes that the species trait values are inherited from common ancestors. Polymorphic inversions are nevertheless very recent in evolutionary terms, and none of our inversions are shared between closely related species. One can still imagine that closely related species could share a mutational propensity to produce inversions within a given length range. However, this possibility can be safely excluded because then we would expect to observe the same effect on both successful and unsuccessful inversions, and this was certainly not the case.

The comparison between successful and unsuccessful inversions is critical for the interpretation of our data.

TABLE 3

Total recombination map length and mean length of inversions for the 12 species of Diptera

Species	Recombination map (cM)	Successful inversions			Unsuccessful inversions		
		$\bar{L}$ (%)	SD	N	$\bar{L}$ (%)	SD	N
D. ananassae	350.1	11.14	6.64	3	5.76	2.17	4
D. buzzatii	696.5	5.90	0.79	4	5.35	4.21	12
D. funebris	895.6	4.82	2.14	5	—	—	—
D. hydei	655.2	9.61	—	1	—	—	—
D. mediopunctata	609.6	3.64	1.28	2	—	—	—
D. melanogaster	294.9	9.09	3.52	10	4.64	1.95	11
D. persimilis	605.1	4.79	2.20	6	4.21	3.80	9
D. prosaltans	285.4	11.45	1.60	4	6.49	2.02	10
D. pseudoobscura	557.1	5.84	2.22	11	4.96	3.41	23
D. subobscura	1007.6	5.40	1.62	28	4.07	2.10	37
A. gambiae	226.7	5.24	2.55	7	8.40	—	1
A. stephensi	327.1	8.84	—	1	4.18	1.87	18
Total		6.40	3.06	82	4.72	2.77	125

$\bar{L}$ (%), mean physical length of inversions expressed as the percentage of the total physical length of the genome; SD, standard deviation; N, number of inversions for each species.

TABLE 4

Analysis of variance and regression to test the effect of species recombination map length (G_i) on physical length of inversions

	Successful inversions				Unsuccessful inversions			
Source of variation	d.f.	MS	*F*	*P*[a]	d.f.	MS	*F*	*P*[a]
Among species	11	25.19	4.05	0.0002 (0.0011)	8	6.32	0.79	0.61 (0.53)
Regression $1/G_i$	1	167.84	15.36	0.0029 (0.0054)	1	16.38	3.36	0.11 (0.13)
Deviations	10	10.92	1.76	0.08 (0.08)	7	4.88	0.61	0.75 (0.67)
Within species	66	6.22			102	7.95		

X-linked inversions (4 successful and 14 unsuccessful) were excluded from this analysis. MS, mean square.
[a] The probability was assessed by the parametric *F*-test and by a resampling test with 10,000 replicas (shown in parentheses).

Inversions that have survived and flourished, becoming common and widespread in natural populations (successful inversions), are a highly selected subset of all occurring inversions: those with the highest selective advantage. On the other hand, unsuccessful inversions are a mixture comprising unique (recently arisen) inversions and also some inversions currently endemic or restricted to a small portion of the species distribution. They are scattered over the entire range of recombination values and taxons, and, although some of them might have been partially sieved by natural selection, it is clear that their selective advantage cannot be large. As previously noted (Olvera *et al.* 1979; Cáceres *et al.* 1997), unsuccessful inversions are smaller than successful inversions and have a larger length variance within species. Furthermore, their length does not differ across species and does not show either phylogenetic clustering around the group level or correlation with recombination map length. These clear contrasting results rule out any mutational cause for the length distribution of successful inversions among species and point to natural selection as the only possible explanation.

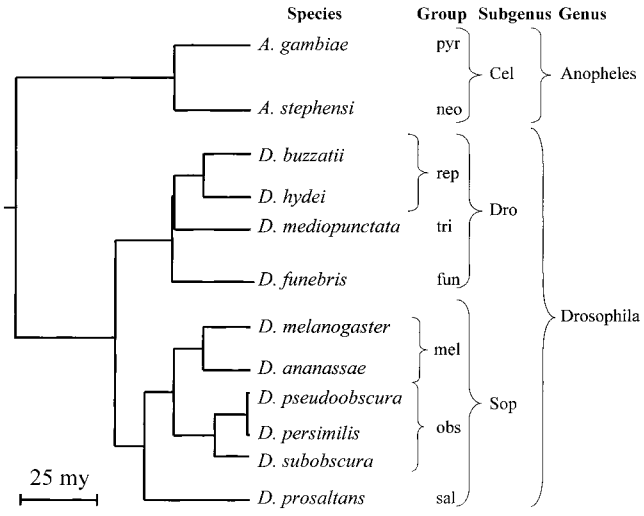

Figure 2.—Phylogenetic tree and branching times of the 12 species of Diptera used in the study. To test for a phylogenetic clustering of the variables, four taxonomic levels were considered (genus, subgenus, group, and species). Abbreviations: pyr, pyretophorus; neo, neocellia; rep, repleta; tri, tripunctata; fun, funebris; mel, melanogaster; obs, obscura; sal, saltans; Cel, Cellia; Dro, Drosophila; Sop, Sophophora.

To explain the correlation between physical length of successful inversions and species recombination map length, we must assume then that natural selection is acting either (i) directly through recombination length of inversions, or (ii) indirectly through a trait correlated with recombination. Perhaps species with similar recombination values within the same taxon also share the same selective pressure on inversion length because of an unknown cause independent of recombination. This explanation does not seem quite parsimonious, and, given the results of the comparative analyses and the expected connection of inversion length with recombination, we believe that recombination-mediated selection is indeed responsible.

Two opposite selective forces seem to be acting on inversion length. On one side, three alternative but nonmutually exclusive theories predict a positive relationship between the selective advantage of an inversion and its length, *i.e.*, that long inversions should be favored. Under the genic selection model (Nei *et al.* 1967), the longer the inversion, the greater its selective advantage, provided that it contains few or no deleterious alleles (Santos 1986). The coadaptation theory proposes that inversions reduce recombination in the heterokaryotypes, allowing the capture of favorable allelic combinations (Charlesworth and Charlesworth 1973; Charlesworth 1974) and the building up of coadapted gene complexes (Kojima and Schaffer 1964; Dobzhansky 1970; Álvarez and Zapata 1997). Accordingly, the probability of "catching" two or more genes with epistatic effects on fitness increases with the size of the inversion, and the selective advantage gained

TABLE 5

Nested analysis of variance to test for a phylogenetic clustering of species recombination map length and physical length of successful and unsuccessful inversions

Source of variation	Recombination map length ($1/G_i$)				Physical length of inversions							
					Successful inversions				Unsuccessful inversions			
	d.f.	MS $\times 10^6$	F_S	Variance component (%)	d.f.	MS	F_S	Variance component (%)	d.f.	MS	F_S	Variance component (%)
Between genera	1	0.10	0.01	0.00	1	3.22	0.23	0.00	1	1.47	0.24	0.00
Between subgenera	1	8.40	2.03	18.14	1	14.20	0.30	0.00	1	5.98	0.60	0.00
Between groups	5	4.14	8.04*	68.02	5	46.98	7.39*	50.21	3	10.00	2.24	4.15
Between species	4	0.51		13.84	4	6.36	1.02	0.16	3	4.47	0.56	0.00
Within species	—	—	—	—	66	6.22		49.63	102	7.95		95.85

* $P < 0.05$.

by the inversion increases with recombination distance between them (Charlesworth and Charlesworth 1973). Finally, stabilizing selection on a quantitative trait should favor those modifiers that reduce recombination, *e.g.*, inversions, because they reduce the genetic variance of the trait and increase the mean fitness of the population (Mather 1943; Charlesworth 1993). If the modifier reduces the map length of a chromosome, its selective advantage would be proportional to the reduction of recombination (Charlesworth 1993). In other words, if the modifier is an inversion, its selective advantage would be proportional to its recombination length.

On the other side, inversion length also has negative consequences on fertility. Because of the ordered oogenesis of females in Drosophila and other Diptera, the unbalanced chromosomes resulting from single crossovers within the inverted region of heterokaryotypes are always set into the polar bodies and no inviable zygotes are formed (Sturtevant and Beadle 1936; Carson 1946). In Drosophila, there is no chiasma formation in males, and, in Anopheles spermatogenesis, the bridge between the two nuclei at anaphase I prevents their separation and their development into sperm (White 1973). The real problem arises with double crossovers, which produce one-fourth of unbalanced gametes because four-strand double crossovers yield only unbalanced gametes (Sturtevant and Beadle 1936; Roberts 1976). This fertility effect selects against long inversions because of their increased probability of double crossovers (Navarro *et al.* 1997). The operation of the two opposite factors is evident from Figure 1A. Large inversions are totally absent in species with long recombination maps, while there is a relative paucity of small inversions in species with short recombination maps.

According to the previous discussion, selection favors larger inversions in species with a low recombination rate than in species with a high recombination rate. One might predict, consequently, that species with larger recombination maps should accumulate more inversions than species with shorter recombination maps, because the smaller inversions block only a relatively small part of the chromosome and leave space for new inversions to settle. Our results agree with this prediction. It seems that there is a positive correlation between the number of successful inversions per species and the recombination map length ($r = 0.47$), although, because we have only 12 species, the correlation was not statistically significant ($P = 0.12$). There are nevertheless differences among species in the presence or absence of inversions that do not fit into this explanation. For instance, *D. simulans* and *D. mauritiana*, with a recombination map similar to that of their close relative *D. melanogaster*, exhibit no inversion polymorphism. Likewise, *D. virilis*, with no known polymorphic inversions, has one of the largest recombination maps of the Drosophila genus. Thus other factors, such as the age of

the species or differences in the molecular mechanisms that generate inversions, *e.g.*, the smaller level of middle repetitive DNA in *D. simulans* relative to *D. melanogaster* (Dowsett and Young 1982), should also be important determinants of the number of polymorphic inversions per species. In the *D. buzzatii* species complex, a significant negative correlation between length of polymorphic inversions and number of polymorphic inversions per species was observed (Cáceres *et al.* 1997). However, in our genus-wide data set such correlation was not significant ($F = 1.14$; d.f. 1, 10; $P = 0.31$), which suggests that recombination, rather than number of inversions, is the main determinant of inversion size in different species.

The significant correlation observed between physical length of successful inversions and recombination map length is striking if one considers that the several sources of error underlying the diverse data used would tend to hamper our ability to detect a trend. Although we have estimated each species' recombination map length as accurately as possible, our estimates are inevitably approximate in some cases. Moreover, our model assumes that the distribution of crossovers along the chromosome is uniform, which is not always the case (True *et al.* 1996). Relative inversion length was also measured approximately, because cytological distances are only rough estimates of real physical distances and to determine the exact location of the breakpoints is usually difficult. Finally, in some cases, inversions are found associated on the same chromosome (overlapping or in strong linkage disequilibrium), with more complicated recombination-reducing effects, depending on the chromosomal arrangements combined in each population, than just the recombination length of the inversion (Krimbas 1992). Thus, that the trend's signal could emerge so strongly is remarkable and suggests a consistent and general phenomenon. The population genetics of polymorphic inversions has long been one of the most powerful and mysterious examples of natural selection in action. Our statistical analysis of the available empirical evidence corroborates the notion that inversions are selected by their effects on recombination and adds further evidence to the importance of recombination differences among species in their evolutionary dynamics.

We appreciate very much the comments on the manuscript of A. Berry, B. Charlesworth, J. Coyne, R. C. Lewontin, A. Navarro, J. M. Ranz, and two anonymous referees. A. Berry considerably improved the article readability. We thank E. Betrán, M. Coluzzi, A. Fontdevila, A. M. de Ron, J. Rozas, C. Segarra, V. P. Sharma, S. K. Subbarao, and C. Zapata for providing us articles with linkage maps and/or descriptions of chromosomal inversions. Work was supported by a FI fellowship from the Comissionat per a Universitats i Recerca (Generalitat de Catalunya, Spain) awarded to M.C. and grant PB95-0607 from the Dirección General de Investigación Científica y Técnica (Ministerio de Educación y Cultura, Spain) awarded to A.R.

LITERATURE CITED

Álvarez, G., and C. Zapata, 1997 Conditions for protected inversion polymorphism under supergene selection. Genetics **146:** 717–722.

Anderson, W. W., 1990 Linkage map of *Drosophila pseudoobscura*, pp. 3188–3189 in *Genetic Maps: Locus Maps of Complex Genomes, Vol. 3, Lower Eukaryotes*, edited by S. J. O'Brien. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Barker, J. S. F., F. M. Sene, P. D. East and M. A. Q. R. Pereira, 1985 Allozyme and chromosomal polymorphism of *Drosophila buzzatii* in Brazil and Argentina. Genetica **67:** 161–170.

Beckenbach, A. T., 1986 The third chromosome inversion polymorphisms in northern populations of *Drosophila persimilis*. Can. J. Genet. Cytol. **28:** 401–408.

Beers, C. V., 1937 Linkage groups in *Drosophila pseudoobscura*, race B. Genetics **22:** 577–586.

Begun, D. J., and C. F. Aquadro, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. Nature **356:** 519–520.

Bell, G., 1989 A comparative method. Am. Nat. **133:** 553–571.

Bicudo, H. E. M. C., 1973 Chromosomal polymorphism in the saltans group of Drosophila. I. The *saltans* subgroup. Genetica **44:** 520–552.

Bicudo, H. E. M. C., M. K. Hosaki, J. Machado and M. C. N. Marques, 1978 Chromosomal polymorphism in the saltans group of Drosophila. II. Further study on *D. prosaltans*. Genetica **48:** 5–15.

Brehm, A., and C. B. Krimbas, 1991 Inversion polymorphism in *Drosophila obscura*. J. Hered. **82:** 110–117.

Cáceres, M., A. Barbadilla and A. Ruiz, 1997 Inversion length and breakpoint distribution in the *Drosophila buzzatii* species complex: is inversion length a selected trait? Evolution **51:** 1149–1155.

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

Chakravarti, A., L. K. Lasher and J. E. Reefer, 1991 A maximum likelihood method for estimating genome length using genetic linkage data. Genetics **128:** 175–182.

Charlesworth, B., 1974 Inversion polymorphism in a two-locus genetic system. Genet. Res. **23:** 259–280.

Charlesworth, B., 1993 Directional selection and the evolution of sex and recombination. Genet. Res. **61:** 205–224.

Charlesworth, B., and D. Charlesworth, 1973 Selection of new inversions in multilocus genetic systems. Genet. Res. **21:** 167–183.

Coluzzi, M., and A. Sabatini, 1967 Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. Parassitologia **9:** 73–88.

Coluzzi, M., M. Di Deco and G. Cancrini, 1973 Chromosomal inversions in *Anopheles stephensi*. Parassitologia **15:** 129–136.

Coluzzi, M., A. Sabatini, V. Petrarca and M. Di Deco, 1979 Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. Trans. R. Soc. Trop. Med. Hyg. **73:** 483–497.

Crow, J. F., 1990 Mapping functions. Genetics **125:** 669–671.

Dobzhansky, Th., 1948 Genetics of natural populations. XVI. Altitudinal and seasonal changes produced by natural selection in certain populations of *Drosophila pseudoobscura* and *Drosophila persimilis*. Genetics **33:** 158–176.

Dobzhansky, Th., 1970 *Genetics of the Evolutionary Process*. Columbia University Press, New York.

Dobzhansky, Th., and C. Epling, 1944 Contributions to the genetics, taxonomy, and ecology of *Drosophila pseudoobscura* and its relatives. Carnegie Inst. Washington Publ. **554:** 1–183.

Dowsett, A. P., and M. W. Young, 1982 Differing levels of dispersed repetitive DNA among closely related species of Drosophila. Proc. Natl. Acad. Sci. USA **79:** 4570–4574.

Dubinin, N. P., and G. G. Tiniakov, 1946 Structural chromosome variability in urban and rural populations of *Drosophila funebris*. Am. Nat. **80:** 393–396.

Felsenstein, J., 1985 Phylogenies and the comparative method. Am. Nat. **125:** 1–15.

Futch, D. G., 1966 A study of speciation in South Pacific populations of *Drosophila ananassae*. Univ. Texas Publ. **6615:** 79–120.

Harvey, P. H., and M. D. Pagel, 1991 *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.

Hess, O., 1976 Genetics of *Drosophila hydei* Sturtevant, pp. 1343–1363 in *The Genetics and Biology of Drosophila*, Vol. 1c, edited by M. Ashburner and E. Novitski. Academic Press, New York.

Hinton, C. W., and J. E. Downs, 1975 The mitotic, polytene, and meiotic chromosomes of *Drosophila ananassae*. J. Hered. **66:** 353–361.

Kastritsis, C. D., 1966 Cytological studies on some species of the tripunctata group of Drosophila. Univ. Texas Publ. **6615:** 413–473.

Kojima, K., and H. E. Schaffer, 1964 Accumulation of epistatic gene complexes. Evolution **18:** 127–131.

Krimbas, C. B., 1992 The inversion polymorphism of *Drosophila subobscura*, pp. 127–220 in *Drosophila Inversion Polymorphism*, edited by C. B. Krimbas and J. R. Powell. CRC Press, Boca Raton, FL.

Krimbas, C. B., and J. R. Powell, 1992 Introduction, pp. 1–52 in *Drosophila Inversion Polymorphism*, edited by C. B. Krimbas and J. R. Powell. CRC Press, Boca Raton, FL.

Lemeunier, F., and S. Aulard, 1992 Inversion polymorphism in *Drosophila melanogaster*, pp. 339–405 in *Drosophila Inversion Polymorphism*, edited by C. B. Krimbas and J. R. Powell. CRC Press, Boca Raton, FL.

Lindsley, D. L., and G. G. Zimm, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.

Loukas, M., C. B. Krimbas, P. Mavragani-Tsipidou and C. D. Kastritsis, 1979 Genetics of *Drosophila subobscura* populations. VIII. Allozyme loci and their chromosome maps. J. Hered. **70:** 17–26.

Mahmood, F., and R. K. Sakai, 1984 Inversion polymorphisms in natural populations of *Anopheles stephensi*. Can. J. Genet. Cytol. **26:** 538–546.

Mather, K., 1943 Polygenic inheritance and natural selection. Biol. Rev. **18:** 32–64.

Moore, B. C., and C. E. Taylor, 1986 Drosophila of southern California. III. Gene arrangements of *Drosophila persimilis*. J. Hered. **77:** 313–323.

Navarro, A., E. Betrán, A. Barbadilla and A. Ruiz, 1997 Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics **146:** 695–709.

Nei, M., K. Kojima and H. E. Schaffer, 1967 Frequency changes of new inversions in populations under mutation-selection equilibria. Genetics **57:** 741–750.

Olvera, O., J. R. Powell, M. E. De La Rosa, V. M. Salceda, M. I. Gaso *et al.*, 1979 Population genetics of Mexican Drosophila. VI. Cytogenetic aspects of the inversion polymorphism in *Drosophila pseudoobscura*. Evolution **33:** 381–395.

Orr, H. A., 1995 A new linkage map of the *Drosophila pseudoobscura* X chromosome. Dros. Inf. Serv. **76:** 127–128.

Parvez, S. D., K. Akhtar and R. K. Sakai, 1985 Two new mutations and a linkage map of *Anopheles stephensi*. J. Hered. **76:** 205–207.

Perje, A. M., 1955 Genetic and cytological studies of *Drosophila funebris*. Some sex-linked mutations and their standard order. Acta Zool. **36:** 51–66.

Powell, J. R., 1997 *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University Press, New York.

Purvis, A., and A. Rambaut, 1995 Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. Comput. Appl. Biosci. **11:** 247–251.

Roberts, P. A., 1976 The genetics of chromosomal aberration, pp. 67–184 in *The Genetics and Biology of Drosophila*, Vol. 1a, edited by M. Ashburner and E. Novitski. Academic Press, London.

Russo, C. A. M., N. Takezaki and M. Nei, 1995 Molecular phylogeny and divergence times of Drosophila species. Mol. Biol. Evol. **12:** 391–404.

Ruiz, A., H. Naveira and A. Fontdevila, 1984 La historia evolutiva de *Drosophila buzzatii*. IV. Aspectos citogenéticos de su polimorfismo cromosómico. Genét. Ibér. **36:** 13–35.

Sampaio, M. C., F. R. Varandas, S. C. Vaz and A. B. Carvalho, 1996 Mapeamento do cromossomo X de *Drosophila mediopunctata*. Rev. Brasil. Genét. **19** (Suppl.): 268.

Santos, M., 1986 The role of genic selection in the establishment of inversion polymorphism in *Drosophila subobscura*. Genetica **69:** 35–45.

Schafer, D. J., D. K. Fredline, W. R. Knibb, M. M. Green and J. S. F. Barker, 1993 Genetics and linkage mapping of *Drosophila buzzatii*. J. Hered. **84:** 188–194.

Spassky, B., S. Zimmering and Th. Dobzhansky, 1950 Comparative genetics of *Drosophila prosaltans*. Heredity **4:** 189–200.

Sperlich, D., and P. Pfriem, 1986 Chromosomal polymorphism in natural experimental populations, pp. 257–309 in *The Genetics and Biology of Drosophila*, Vol. 3e, edited by M. Ashburner, H. L. Carson and J. N. Thompson, Jr. Academic Press, London.

Spiess, E. B., 1950 Experimental populations of *Drosophila persimilis* from an altitudinal transect of the Sierra Nevada. Evolution **4:** 14–33.

Spiess, E. B., 1965 A discovery and rediscovery of third chromosome arrangements in *Drosophila persimilis*. Am. Nat. **99:** 423–425.

Spurway, H., 1945 The genetics and cytology of *Drosophila subobscura*. I. Element A. Sex-linked mutants and their standard order. J. Genet. **46:** 268–286.

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.

Tobari, Y. N., 1993 Linkage maps, pp. 49–51 in *Drosophila ananassae—Genetical and Biological Aspects*, edited by Y. N. Tobari. Japan Scientific Societies Press, Tokyo.

True, J. R., J. M. Mercer and C. C. Laurie, 1996 Differences in crossover frequency and distribution among three sibling species of Drosophila. Genetics **142:** 507–523.

Wasserman, M., 1962 Cytological studies of the repleta group of the genus Drosophila. IV. The hydei subgroup. Univ. Texas Publ. **6205:** 73–84.

White, M. J. D., 1973 *Animal Cytology and Evolution*. Cambridge University Press, London.

Zheng, L., M. Q. Benedict, A. J. Cornell, F. H. Collins and F. C. Kafatos, 1996 An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. Genetics **143:** 941–952.

Communicating editor: R. R. Hudson