Restriction-Map Variation at the Zeste-tko Region in Natural Populations of Drosophila melanogaster

M. Aguadé,* N. Miyashita,† and C. H. Langley†

*Departament de Genetica, Facultat de Biologia, Universitat de Barcelona; and †Laboratory of Molecular Genetics, Natural Institute of Environmental Health Sciences

Restriction-map variation in 64 X chromosome lines extracted from three different natural populations of Drosophila melanogaster was investigated with seven six-nucleotide-recognizing enzymes for a 20-kb region including the zeste and tko genes. Ten restriction-site and four length polymorphisms (two insertions and two deletions) were detected. Contrary to the predicted lower level of variation for genes on the X chromosome, the level of variation attributable to nucleotide substitution (estimated heterozygosity/nucleotide = 0.004) was similar to that previously reported for autosomal loci. The amount of insertion/deletion variation in the studied region was within the range observed in autosomal regions and thus not explainable by a simple selection model against the effects of insertional mutations. A general lack of linkage disequilibrium between polymorphic sites was observed.

Introduction

The advent of recombinant DNA technology has made possible the cloning and characterization of particular regions of the genome and, with it, the possibility to analyze DNA sequence variation in natural populations. A new approach can therefore be used to address the old standing questions in population genetics. According to population genetics theory, genes located on the X chromosome are expected to show levels of variation lower than those shown by genes on autosomes, for two reasons: (1) the effect of random genetic drift in reducing levels of variation should be stronger because of the smaller effective population size for genes on the X chromosome and (2) the effect of selection against recessive deleterious variants on the X chromosome should be stronger than that on the autosomes because of hemizygosity in the heterogametic sex.

In natural populations of Drosophila melanogaster the level of variation detected in the white locus region (Langley and Aquadro 1987; Miyashita and Langley 1988) was similar to that reported elsewhere for autosomal genes (Langley et al. 1982; Leigh-Brown 1983; Aquadro et al. 1986). This departure from expectation may not be representative of genes on the X chromosome but may be peculiar to the white region. To investigate this question further, several other regions of the X chromosome have been surveyed for the same sample of 64 isogenic X-chromosome lines of D. melanogaster that were cited by Miyashita and Langley (1988). The region of interest for the present study is a 20-kb fragment including the zeste and tko genes (fig. 1); it has

1. Key words: restriction polymorphism, insertion/deletion variation, X chromosome.

Address for correspondence and reprints: Dr. M. Aguadé, Departament de Genetica, Facultat de Biologia, Universitat de Barcelona, Barcelona, Spain.

© 1989 by The University of Chicago. All rights reserved.
0737-4963/89/0602-0002$02.00
FIG. 1.—Summary restriction map of the region encompassing the zeste and tko genes for 64 X chromosome lines of Drosophila melanogaster. Monomorphic sites are below the line (see text for PvuII), and polymorphic sites are above the line (H = BamHI; R = EcoRI; D = HindIII; P = PstI; V = PvuII; C = SacI; and S = Sall). Outside the 20-kb probed region only the nearest monomorphic site for each enzyme is shown. Approximate size for insertion/deletions is given in table 1. Also shown is a schematic representation of the zeste and tko transcripts as well as the phage clone used as probe (CS1014), and the fragments used for germ line transformation by different authors. (1) pSXR2 fragment used by Royden et al. (1987); (2) EcoRI 5.8-kb fragment used by Gunaratne et al. (1986), and (3) BamHI 3.9-kb fragment used by Mariani et al. (1985).
been chosen both for its location on the X chromosome, in bands 3A2-3 (Judd et al. 1972), and for its high density of transcriptional units, two to four transcription units covering 10 contiguous kilobases of the total 20 kb studied.

The *zeste* gene is a regulatory gene that affects expression of at least three other loci, *bithorax (BX-C)*, *decapentaplegic (dpd)*, and *white (w)*, in a manner that depends on the homologous pairing of the chromosomes bearing the target gene (transvection; Lewis 1954; Judd 1988). The *zeste* gene recently has been cloned and characterized (Mariani et al. 1985; Gunaratne et al. 1986). Germ-line transformation experiments have shown that at least some aspects of normal *zeste* expression reside in a 5.8-kb EcoRI fragment (Gunaratne et al. 1986) (fig. 1). Other transformation experiments show that a 3.9-kb BamHI fragment (fig. 1) is sufficient to obtain the phenotype associated with the mutant *z*~*mut*~ (Pirrotta et al. 1987). The gene consists of three exons separated by two small introns, coding for a putative protein of 555 amino acids. Although mutant alleles of *zeste* have been studied intensively, the wild-type function of *zeste* is not yet clear.

The *tko* (technical knockout) mutation is a behavioral mutation of *D. melanogaster* which causes “bang sensitivity” and which has been mapped genetically 0.006 map units distal to the *zeste* gene (Judd et al. 1972). Germ-line transformation has shown that a 3.1-kb fragment of genomic DNA distal to the *zeste* gene (fig. 1) is sufficient to give a wild-type phenotype (Royden et al. 1987). This fragment produces only one complete transcript, which codes for a putative protein of 140 amino acids.

In the *tko-zeste* region two additional transcripts, 3.9- and 4.2-kb long, have been identified (Mariani et al. 1985; Gunaratne et al. 1986) which have been shown to lie between both genes (Royden et al. 1987). Although these transcripts have not been well characterized and their function has not yet been identified, their existence contributes further to the density of transcriptional units in the studied region.

In the present paper we report the study of naturally occurring variation at the *tko-zeste* region, as evidenced by restriction-map analysis in a sample of 64 X-chromosome isogenic lines extracted from three different populations.

**Material and Methods**

**Fly Stocks**

Sixty-four isogenic X-chromosome lines extracted from three different populations (20 from North Carolina, 27 from Texas, and 17 from Fukuoka, Japan) have been used in the present study. Lines were constructed as described in Miyashita et al. (1986).

**Restriction-Map Analysis**

Seven hexanucleotide-recognizing restriction enzymes were used in the present study: *BamHI, EcoRI, HindIII, PstI, PvuII, Sall*, and *SacI*. Genomic DNA was CsCl purified according to Bingham et al. (1981). Seven single digestions and two double digestions (*EcoRI and BamHI, HindIII and SacI*) have been performed. Digestion, electrophoresis, blotting, and probing were carried out as described elsewhere (Miyashita and Langley 1988). The λ phage clone CS1014 of Gunaratne et al. (1986) was used for probing in the present analysis.

**Results**

Figure 1 gives the restriction map of a 20-kb region encompassing the *zeste* and *tko* genes, showing all restriction sites scored in the present study (except for *PvuII*)
and the location of insertion/deletions. Coordinates (in kilobases) in the restriction map are those of Gunaratne et al. (1986). The resolution of the restriction map is relatively coarse, since not all possible double digestions were carried out. Only one PvuII restriction site has been mapped because nine of the fragments generated by this enzyme, ranging in size from 0.4 kb to 2.6 kb, showed no variation; only the tenth fragment (0.9 kb long) showed size variation, and this polymorphism can be easily attributed to the acquisition of the restriction site shown in figure 1 within the 3.9-kb fragment sequenced by Pirrotta et al. (1987). Ten restriction-site polymorphisms (of 55 sites scored) and four length polymorphisms (two insertions and two deletions) have been detected in the 64 X-chromosome lines studied. Restriction-site polymorphisms are distributed along the whole probed region, while length polymorphisms are restricted to the centromere-proximal (3' to zeste) region. There is evidence of a polymorphism 5' of the probed region (between coordinates -16.2 and -12.6), but it could not be unambiguously attributed to the acquisition of a new restriction site or to an insertion/deletion.

Table 1 shows the 27 different haplotypes observed and their frequency distribution. Only two haplotypes (1 and 3) are present in the three geographical samples; two (4 and 5) are present in both North American samples; another two (11 and 12) are present both in Texas and in Fukuoka; and the rest, either uniquely or multiply represented, are present in only one of the three sampled areas.

Frequencies of individual restriction-site variants range from 0.031, for sites present twice in the sample, to 0.406. Three of the four length polymorphisms are present only once in the sample (in fact, they are unique to North Carolina). Deletion del.1 is the only length variant multiply represented. Each occurrence could be molecularly unique and independent in origin, since they are identical only in approximate size and location. No insertion/deletion variation has been observed in the Japanese population.

From restriction-site polymorphisms three different estimates of variation attributable to nucleotide substitutions have been calculated, for each population as well as for the pooled data. Individual population estimates were very similar to pooled estimates. Although the assumptions underlying each particular method are different—selective neutrality and stochastic stationarity for Hudson's (1981) heterozygosity per nucleotide, no assumption for Nei et al.'s nucleotide diversity (Nei and Li 1979; Nei and Tajima 1981), and linkage equilibrium for Engels' (1981) heterozygosity per nucleotide—the estimates obtained are very similar (0.0040 ± 0.0013, 0.0044 ± 0.0026, and 0.0043 ± 0.0003, respectively).

The amount of insertion/deletion variation in the zeste-tko region has been measured both as the frequency of lines with large insertion/deletions (relative to the most common restriction-map haplotype) per kilobase surveyed (frequency = 0.0055) and as the average number of gap nucleotides, per nucleotide site, between two randomly chosen haplotypes [average no. of gap nucleotides = 0.0239; Nei 1987, eq. (10.18)].

Since the number of polymorphic sites, as well as the allele and two-locus haplotype frequencies, do (in some cases) differ in the three studied populations (table 1), the possible existence of linkage disequilibrium between sites has been analyzed separately for each population. Both restriction-site and length polymorphisms in which the rarer variant is not unique have been considered. A total of 36, 28, and 10 pairwise comparisons have been done for the samples from North Carolina, Texas, and Fukuoka, respectively, of which only 2, 2, and 5, respectively, have evidenced significant gametic associations (data not shown). Only one of the comparisons, that between
Table 1

Frequency Distribution of Haplotypes

<table>
<thead>
<tr>
<th>Haplotype Number</th>
<th>PstI</th>
<th>SacI</th>
<th>EcoRI</th>
<th>PvuII</th>
<th>PstI</th>
<th>HindIII</th>
<th>PstI</th>
<th>HindIII</th>
<th>SacI</th>
<th>del.1</th>
<th>ins.1</th>
<th>ins.2</th>
<th>del.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10.9</td>
<td>-2.6</td>
<td>0</td>
<td>5.55</td>
<td>6.2</td>
<td>6.6</td>
<td>9.65</td>
<td>9.75</td>
<td>11.4</td>
<td>0.6</td>
<td>8.7</td>
<td>6.7</td>
<td>0.1</td>
<td>NC</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TX</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>JPN</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Frequency:</td>
<td>0.14</td>
<td>0.03</td>
<td>0.03</td>
<td>0.22</td>
<td>0.23</td>
<td>0.20</td>
<td>0.28</td>
<td>0.03</td>
<td>0.41</td>
<td>0.11</td>
<td>0.08</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

**NOTE.**—Plus signs (+) and minus signs (−) indicate, respectively, the presence or absence of restriction-site and insertion/deletion variation. NC = North Carolina; TX = Texas; and JPN = Japan. Location of restriction-site polymorphisms is given according to the coordinates of Gunaratne et al. (1986). The approximate size of insertions (ins) and deletions (del) is expressed in kilobases. In the lower part of the table, the frequency of the less common variant is given for each polymorphism, both for the pooled data and for each position.
sites PstI (+6.2) and HindIII (+9.75), consistently shows a significant linkage disequilibrium in the same direction in the three populations.

Discussion

The action of both random genetic drift and natural selection against recessive deleterious alleles in reducing standing levels of variation is expected to be stronger for genes on the X chromosome than for those located on the autosomes. Standard population genetics theory expects that genes on heterochromosomes will have a smaller effective population size than will those on autosomes, owing to hemizygosity of the heterogametic sex. This hemizygosity is also the cause of increased selection pressure on recessive deleterious alleles of genes on the X chromosome. Theoretical studies of recessive lethals (Haldane 1927) and more recent studies of null-activity alleles at enzyme loci (Langley et al. 1981) have shown a reduction in the relative amount of variation on the X chromosome. A similar result should be expected for variation at the DNA level if a large portion of that variation were indeed mildly deleterious (Ohta 1972; Kimura 1983). Unfortunately there is no statistical procedure to rigorously test for any reduction in restriction polymorphism on the X chromosome. The similarity in the estimates obtained in previous studies provides no evidence for heterogeneity among the chromosomes in restriction-site polymorphism in *Drosophila melanogaster*. Neither Langley and Aquadro (1987) nor Miyashita and Langley (1988) found a reduction in the level of DNA sequence variation in the *white* locus region. The data on the *zeste-tko* region reported here show the same trend as those previous analyses; in fact, the estimated heterozygosity per nucleotide, 0.004, is similar to that reported for autosomal loci [0.006 for *Adh* (Langley et al. 1982; Aquadro et al. 1986) and 0.002 for *hsp70* (Leigh-Brown 1983)].

Most studies of variation at the DNA level in *D. melanogaster* have shown that large insertions (known or presumed to be transposable elements) are not uncommon in natural populations. But, unlike restriction-site polymorphisms, these large insertions show a skewed frequency distribution toward low individual frequencies. This observation has been considered an indication of their deleterious effect on fitness. This kind of polymorphism should show an even greater reduction in the amount of variation on the X chromosome than is seen in restriction-site polymorphisms. When the number of *copia*-like elements in autosomes and X chromosomes has been compared by in situ hybridization (Montgomery et al. 1987), two of the three elements studied failed to show the expected reduction. Because these results could not be explained by simple selection against the effects of insertional mutations, an alternative model of generation of dominant lethal rearrangements by unequal exchanges was proposed (Montgomery et al. 1987; Langley et al., accepted). In the data presented here there is similarly no evidence for a reduction in the amount of length polymorphisms on the X chromosome. In fact, the insertion/deletion variation (0.006, measured as the frequency of lines with large insertion/deletions per kilobase surveyed) in the *zeste-tko* region is within the range observed in autosomal regions [0.005 and 0.017 for *Adh* (Langley et al. 1982; Aquadro et al. 1986), 0.006 for *hsp70* (Leigh-Brown 1983), and 0.005 for *rosy* (Aquadro et al. 1988)].

In the present data there is no evidence for any heterogeneity in the distribution of restriction-site polymorphisms in the *zeste-tko* region studied. It should be pointed out that the only polymorphism observed within the coding region of the genes (site *PvuII* at +3.4) corresponds to a change already reported [position 2788 of Pirrotta et al. (1987), which causes a silent change in the putative polypeptide]. Length poly-
morphisms, on the other hand, have been found exclusively in the flanking regions, in agreement with previous studies.

Miyashita and Langley's (1988) analysis of variation at the white locus region has shown that linkage equilibrium seems to be reached when pairs under comparison are separated by ≥2 kb. Accordingly, a general lack of linkage disequilibrium would be expected in the zeste-tko region if levels of recombination in that region were comparable to those observed in the white locus region. In the studied region there is in fact only one case of consistent linkage disequilibrium among the three sampled populations. The observation of a larger amount of linkage disequilibrium in the Japanese population may be attributable to the small sample size, or this may be reflecting a younger evolutionary history for this population.

Even if the analysis of the zeste-tko region does not support the expectation of reduced levels of variation on the X chromosome as compared with that on the autosomes, no general assessment can yet be made concerning that question. Analysis of variation at a larger sample of loci, in both the autosomes and X chromosome, are needed to draw any general conclusions.

Acknowledgment

We would like to thank M. Goldberg for providing the CS1014 clone.

LITERATURE CITED


MASATOSHI NEI, reviewing editor

Received June 13, 1988; revision received September 9, 1988