Evolution of the Ribosomal DNA Spacers of Drosophila melanogaster: Different Patterns of Variation on X and Y Chromosomes

Scott M. Williams,*1 Glenn R. Furnier,1,2 Eric Fugø* and Curtis Strobeck*

*Department of Zoology and 1Department of Forest Science, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

Manuscript received June 23, 1986
Revised copy accepted February 27, 1987

ABSTRACT

Length variation of the ribosomal gene spacers of Drosophila melanogaster was studied. Analysis of 47 X chromosomal and 47 Y chromosomal linked rDNA arrays collected from five continents indicates that the arrays on the two chromosomes differ qualitatively. The Y-linked arrays from around the world share little or no similarity for either their overall length or the organization of their spacers. Most of the X-linked arrays do, however, share a major length spacer of 5.1 kb. In addition, those X-linked arrays that have a major 5.1-kb band have similar spacer organization as demonstrated by genomic DNA digestions with several restriction enzymes. These data strongly support the hypothesis that spacer length patterns on only X-linked genes are maintained primarily by natural selection.

It is now evident that multigene families are common in eukaryotic organisms and that the members of these families are not evolving independently of each other (BRITTEN and KOHNE 1968; HOOD, CAMPBELL and ELGIN 1975; OHTA 1980; ZIMMER et al. 1980; DOVER 1982; DOVER et al. 1982; ARNHEIM 1983). Although gene family members are more similar to each other than would be expected if they were evolving independently, the extent of their similarity is not completely known for any family. Recent studies have in fact indicated that some nucleotide sequences within an individual are different among gene family members (GONZALEZ et al. 1985). Even though the mechanism(s) of homogenization have not yet been determined, it is clear that in some species members of a gene family on the same chromosome are much more similar to each other than their counterparts on other chromosomes (TARTOF and DAWID 1976; YAGURA, YAGURA and MURAMATSU 1979; ARNHEIM et al. 1980, 1982; Dvorak et al. 1984; WILLIAMS, DESALLE and STROBECK 1985; METZENBERG et al. 1985; CARLSON and HOGENESS 1985). Systems with these differences among chromosomes are particularly well suited to the study of multigene family evolution because they allow an examination of the levels of homogenization both within and among chromosomal lineages. Because different mechanisms will account for homogenization at these two levels (KIMURA and OHTA 1979; OHTA and DOVER 1983; WILLIAMS and STROBECK 1985), studying such a system should give insight into the underlying mechanisms of homogenization.

One gene family that has been studied in detail is the ribosomal gene family (WELLAUER et al. 1976; BUONGIORNO-NARDELLI et al. 1977; KUNZ et al. 1981; COEN, THODAY and DOVER 1982; COEN, STRACHEN and DOVER 1982; BONCINELLI et al. 1983). It is particularly amenable to evolutionary and structural study because one part of the family, the nontranscribed spacer (NTS), is extremely variable in length in several species (WELLAUER et al. 1976; BUONGIORNO-NARDELLI et al. 1977; COEN, THODAY and DOVER 1982; COEN, STRACHEN and DOVER 1982; BONCINELLI et al. 1983). Because the length variability results from the presence of different numbers of internally repeated nucleotide sequences (WELLAUER et al. 1976; KUNZ et al. 1981; DOVER et al. 1982; GHIKARAIISHI et al. 1983; SAGHAI-MAROOF et al. 1984; WILLIAMS, DESALLE and STROBECK 1985; YANG-YEN et al. 1985) and can be generated by unequal exchange (FEDOROFF 1979; TRECO, BROWNELL and ARNHEIM 1982; COEN, THODAY and DOVER 1982; DOVER et al. 1982; WILLIAMS and STROBECK 1985; EICKSON and SCHMICKEL 1985), this trait is at least as, and probably more, variable than nucleotide sequence. Therefore, NTS length variation is an extremely sensitive measure of variation within and among ribosomal gene arrays. In addition, the amount of chromosome variation within this gene family is species dependent.

In Drosophila melanogaster, the rDNA family arrays occur on both the X and Y chromosomes (RITossa 1976). Molecular characterization of these arrays has demonstrated that both X and Y chromosomes have diagnostic markers for nucleotide sequence (YAGURA, YAGURA and MURAMATSU 1979), NTS length (INDIK and TARTOF 1980), and insertion sequences in the 28 S gene (WELLAUER, DAWID and TARTOF 1978). Non-
theless, the restriction maps do not differ substantially between the two chromosomal arrays (TARTOF and DAWID 1976). These studies have only examined one or a few chromosomes from a single geographic region (COEN, THODAY and DOVER 1982) or laboratory stocks for comparisons (TARTOF and DAWID 1976; BONCINELLI et al. 1983) and do not offer an overview of the type and extent of variation in this gene family in nature. We have extended the previous results to include an examination of X and Y chromosomal linked rDNA spacer lengths from 12 wild strains collected around the world. Because five continents were represented in this study, the results should be a good representation of the type and level of variation found in this species in nature. The variation we found suggests that X chromosomal and Y chromosomal rDNA arrays are evolving via different mechanisms, with selection playing an important role in the evolution of X-linked rDNA spacers but not Y-linked ones.

MATERIALS AND METHODS

Fly stocks: Fly stocks were constructed which had only a single X chromosome and a single Y chromosome extracted from nature. Three or four stocks were constructed from each of 12 isofemale lines collected from five continents. The isofemale lines were kindly provided by C. AQUADRO. These stocks were derived in the following manner: single males from an isofemale line were crossed to virgin females with the compound chromosome -O/(e1) m4/y/Fm7 Y*y+y' v f B (described in Drosophila Information Service 1978). The female progeny from this cross were mated with their fathers. The resulting stocks contained only the X and Y chromosomes from the father.

The 12 isofemale lines were collected from eight populations: Argentina (AR 4 and AR 6), Australia (AU 2 and AU 14), Benin (Benin IIc), Central Africa (CA 7), North Carolina (NC 1 and NC 2), France (FR V3-1), Taiwan (TA 20), and Vietnam (VI 13-1 and VI 15-1). Either three or four X chromosomes and three or four Y chromosomes were isolated from each line for NTS length analysis (Table 1). In total, 47 X chromosome derived and 47 Y chromosome derived rDNA spacers were examined.

DNA extraction and manipulation: For DNA extraction males from the single chromosome lines were crossed to B(2)Y/Df(1) bb 1158Y virgin females [described in LINDSLEY and GREL (1968)]. B+ males and females were collected for the extractions. Because the Df(1) bb 1158 Y chromosome has no rDNA, the only rDNA in B- progeny is derived from the paternal X in females or paternal Y in males. DNA was extracted from 50-150 mg of the collected flies by the method of ISH-HOROWICZ et al. (1979).

For total length variation studies, the DNA was digested with HaIII (Figure 1) according to the manufacturer’s specifications (Bethesda Research Laboratories) and electrophoresed on 0.6% agarose gels. The DNA was then transferred to nitrocellulose, hybridized to radioactive probe, washed, and autoradiographed as described elsewhere (WILLIAMS, DiSALLE and STROBECHEK 1985). The blots were hybridized to the probe pDmr 103HH2 kindly provided by G. DOVER. For mapping of the NTS length variation, digests were done with both HaIII and either Alul or Ddel (Figure 1). In these cases, the DNA was electrophoresed on 1.2% agarose gels.

RESULTS

Ribosomal spacer length variation within lines and geographic regions: The NTS length differences within each isofemale line are most pronounced between X- and Y-linked arrays. In no case are the patterns of spacers from the same isofemale line identical for the two chromosomes. In fact, they rarely share spacers of the same length, and when they do the observed length is generally due to different internal organization of the spacers on the X and Y chromosomal arrays (see below). Both chromosomes also carry a large number of length variants (5-20 bands) within a single array.

In contrast to the large differences between X and Y chromosome-linked rDNA spacers, each isofemale line carried only a few variant patterns for each chromosome (Figure 2, Table 1). Only twice did a line have two distinct Y-linked patterns (Table 1). This observation is of note because it demonstrates that these lines were generated from females that had
FIGURE 2.—Variation of NTS lengths within an isofemale line. Autoradiograph of DNA digested with HaeIII from an isofemale line collected in Australia (AU 2). Headings A, B, C, and D represent four independent isolates derived from single males of the original isofemale line. Note that the Y chromosomal NTS from isolate A differs from the other Y chromosomal spacers and that the X-linked spacers from A and B differ from each other and the X-linked spacers from C and D. The length is denoted in kilobases.

It does suggest that the X-linked spacers are more variable within a population than the Y-linked spacers. This is consistent with results from more intensive intrapopulation studies in *D. melanogaster* (P. Cluster, personal communication).

**Variation among geographic regions:** Not only do X-linked NTSs differ from Y-linked NTSs from the same region, but the two arrays differ from each other with respect to the amount and type of variation among geographic regions (Figure 3). Y chromosomal rDNA NTSs differ substantially among geographic regions. This is true for samples from the same continent as well as for those from different continents. The patterns of NTS lengths from the various strains are consistent with independent evolution of the spacers on each of the Y chromosomes, because they share few if any bands. In fact, the Y-linked NTSs from different populations appear to be as different from each other as Y-derived NTSs are from X-derived NTSs.

The X-linked NTSs are very different from the Y-linked NTSs in their level of variation. Unlike the Y-derived NTSs, most X-linked NTSs (9 of 12) share a common band of 5.1 kb (11 of the 12 populations studied are represented in Figure 3a), an observation consistent with earlier findings (Coen, Thoday and Dover 1982). This NTS length is only present in 2 of the 12 strains on Y-linked arrays (Figure 3b). These two lines (AR 4 and AR 6) represent only one population (Table 1) and the Y chromosomes are identical. The presence or absence of the 5.1-kb fragment on the Y chromosomes was confirmed by direct comparison to X chromosomes having this fragment (data not shown). An earlier study also found 5.1-kb NTSs on Y chromosomes, but only one population was examined (Coen, Thoday and Dover 1982). The divergence among Y chromosome spacers contrasts sharply with the more homogeneous X-linked spacers.

One way of determining the extent of similarity among X chromosome arrays and among Y chromosome arrays is by using a shared fragment analysis. Comparisons among arrays have been done using only the most prominent bands in each array (Table 2). By using only major lengths of the most common array types, most ambiguity that would be created with the less intense bands is eliminated, although some information is necessarily lost. As the major bands represent most of the copies within an array, however, this measure should estimate array differences well. The mean proportion of shared fragment among Y arrays using this approach is 0.04 (including the Y chromosomes from the Argentina and Australia populations which are identical). The mean proportion of shared fragments among X arrays is 0.20. These two values are significantly different ($t = 4.25 P < 0.001$), but

```plaintext
5.1
-5.1
```
TABLE 2

Proportion shared fragments: above diagonal—Y chromosomes below diagonal—X chromosomes

<table>
<thead>
<tr>
<th>Straina</th>
<th>AR 4</th>
<th>AR 6</th>
<th>AU 2</th>
<th>AU 14</th>
<th>BE</th>
<th>CA</th>
<th>NC 1</th>
<th>NC 2</th>
<th>FR</th>
<th>TA</th>
<th>VI 13-1</th>
<th>VI 15-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 4</td>
<td>—</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AR 6</td>
<td>0.20</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AU 2</td>
<td>0.14</td>
<td>0.25</td>
<td>—</td>
<td>1.00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AU 14</td>
<td>0.20</td>
<td>0.25</td>
<td>0.14</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BE</td>
<td>0.25</td>
<td>1.00</td>
<td>0.25</td>
<td>0.25</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.20</td>
<td>0</td>
</tr>
<tr>
<td>NC 1</td>
<td>0.25</td>
<td>1.00</td>
<td>0.25</td>
<td>0.25</td>
<td>1.00</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>NC 2</td>
<td>0.14</td>
<td>0.25</td>
<td>0.14</td>
<td>0.25</td>
<td>0.25</td>
<td>0.14</td>
<td>0.25</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FR</td>
<td>0.25</td>
<td>1.00</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>1.00</td>
<td>0.25</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>TA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.33</td>
</tr>
<tr>
<td>VI 13-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.33</td>
<td>—</td>
</tr>
<tr>
<td>VI 15-1</td>
<td>0.25</td>
<td>0.20</td>
<td>0.17</td>
<td>0.17</td>
<td>0.25</td>
<td>0.20</td>
<td>0.25</td>
<td>0.25</td>
<td>0</td>
<td>0</td>
<td>0.35</td>
<td>0.40</td>
</tr>
</tbody>
</table>

*a Abbreviations as in Table 1.

Figure 3.—Autoradiographs of X-linked and Y-linked rDNA spacers from 11 isofemale lines. Each lane represents the NTS length pattern of a single chromosome array from each of the isofemale lines. The flies were isolated and DNA extracted and treated as described in the text and digested with *HaeIII*. (a) X-linked arrays—note that spacers from AR 4, AR 6, AU 2, AU 14, Benin IIC (BE), NC 1, NC 2 and France V3-1 (FR) all have 5.1-kb lengths. (b) Y-linked arrays—only spacers from AR 4 and AR 6 have 5.1-kb long spacers and these two arrays are identical.

are primarily indicative of the shared 5.1-kb fragment among X chromosomes.

Mapping the spacer length variation: The source of the length variation within a chromosomal array was determined by two double digests (*HaeIII*/AluI and *HaeIII*/Ddel). This procedure allowed us to determine which of the several internally repeated sequences was responsible for most of the spacer length variation in the different arrays. Digests using AluI divided the spacers into three parts: the 5' end of the spacer to the 240 bp AluI repeats, the 240-bp AluI repeats, and the 3' end of the spacer which extends from the AluI repeats to the transcribed part of the rDNA family (Figure 1). This digest showed that both X and Y chromosomes share a common band of 1.0 kb (Figure 4), which is the 3' end of the spacer (Figure 1). The universality of this 1.0-kb fragment is not surprising since this region contains the transcription initiation site and the sequence and length of this region have been shown to be important in the control of transcription (Kohorn and Rae 1982). The arrays on the two chromosomes differ elsewhere with respect
FIGURE 4.—*HaeIII/AluI* mapping of the NTS lengths. The autoradiographs are of DNA double digested with the two enzymes *HaeIII* and *AluI* and run on 1.2% agarose gels. The DNA is identical to that of Figure 3 (a) X-linked arrays—AR 4, AR 6, AU 2, AU 14, Benin HC (BE), NC 1, NC 2 and France V3-1 (FR) all have 1.9-kb fragments, (b) Y-linked arrays—no Y-linked spacers have 1.9-kb fragments.

to internal spacer organization. The fragments derived from the 5′ end of the spacer differ substantially (Figure 4), indicating that the number of internal repeats in this region is different on the two sex chromosomes. Not only do these data demonstrate that more of the X-linked spacer variation is attributable to changes in the 5′ region than in its Y-linked counterpart, but it also shows that X chromosomes from different populations share more common bands than the Y chromosomes. Specifically, the X-linked arrays that have 5.1-kb NTSs all have a 1.9-kb fragment at their 5′ ends, and this band length is in most spacers in a single array. This strongly suggests that the X-linked 5.1-kb spacers share the same organization with the 5′ ends being 1.9 kb long and a specific number of *AluI* repeats in the middle of the spacer. The Y-derived 5.1-kb spacers do not share this organization (Figure 4, lanes AR 4, AR 6). In addition, these double digests indicate that both X and Y chromosomal arrays have restriction site polymorphism. Specifically, Y-linked arrays from Benin, Central Africa and Vietnam and X-linked arrays from Central Africa, Taiwan and Vietnam appear to have *AluI* sites not present in the other arrays (Figure 4).

The *HaeIII/DdeI* double digests confirm the interpretation that more X-linked than Y-linked variation is caused by a variable number of *DdeI* internal repeats at the 5′ end of the spacer. The digestion with *DdeI* creates a similar pattern for all X-derived arrays, with a 0.9-kb band at the 5′ end of the spacer (Figure 5a). The Y-linked spacers, however, show more variation in the number of small repeats at the 5′ end of the NTS (Figure 5b). This supports the argument that the X chromosome-linked 5.1-kb spacers share the same organization with a set number of the small repeating elements at the 5′ end of the spacer, a set number of 340-bp *DdeI* repeats, and a set number of the 240-bp *DdeI* (or *AluI*) repeats. As compared to this conservative organization for the X-linked 5.1-kb spacers, the Y-linked arrays are extremely variable in their length and makeup. Overall, as with the NTS lengths shown by the *HaeIII* digests, the X chromo-
somes are more similar to each other than the \( Y \) chromosomes.

**DISCUSSION**

The data presented here demonstrate that X-linked rDNA arrays commonly have \( \text{HaeIII} \) 5.1-kb long spacer fragments, whereas \( Y \)-linked arrays rarely do. The observation that all X-linked arrays, but not \( Y \)-linked arrays, with 5.1-kb bands have 1.9-kb \( \text{AluI} \) bands also supports the argument that each type of array is unique. These differences between sex chromosomes are particularly informative in the sense that they show that the kind and level of variation on the two sex chromosomes is qualitatively different.

The double digests suggest that the 5.1-kb bands on X-linked rDNA arrays are caused by a conserved organization, with all the internally repeating sequences in the same multiplicity in several widely dispersed populations. The probability of these repeats being in the same multiplicity and adding up to the same overall length in several geographic subpopulations is extremely low if only molecular mechanisms are responsible for this pattern. In particular, a theoretical study has demonstrated that even with only a single internal repeat type, spacer lengths will diverge due to unequal sister chromatid exchange if no constraints are placed on the system (Williams and Strobeck 1985). It would be even less likely that three independent repeats would converge without an even more intense set of constraints. Divergence of NTS lengths caused by unequal homologous exchange would occur similarly. Therefore, even if recombination within the rDNA arrays is much more frequent between different X chromosomes than between \( X \) and \( Y \) chromosomes, the \( X-X \) exchanges still would likely create new length variants. The frequency of recombination between rDNA arrays is, however, similar between different \( X \) chromosomes and \( X \) and \( Y \) chromosomes (S. M. Williams, J. A. Kennison and C. Strobeck, unpublished data), indicating that the observed \( X \) chromosome patterns are not caused by this mechanism. In addition, homologous exchange between chromosome lineages that have diverged due to unequal sister chromatid ex-
change will likely result in extremely variable X-linked rDNA arrays and not similar ones from around the world as reported in this paper. For these reasons, we argue that X-linked NTS lengths must be maintained by natural selection. This can be done by having upper and lower limits to the numbers of internally repeated sequences or by having an optimal value for each of the repeats. We also consider biased gene conversion within the X-linked spacers, which could result in the same type of pattern we have observed (DOVER 1982), to be included within the realm of natural selection.

Selection seems to be the predominant force maintaining the 5.1-kb X-linked spacers, but it is not the only mechanism operating, as evidenced by the number of spacers of different lengths on an array (Figure 3). As mentioned above, NTS length variants are generated by molecular mechanisms such as unequal exchange among internal repeats. This process is ongoing and will continue to generate new length variants (WILLIAMS and STROBECK 1985), suggesting that the rDNA arrays are in constant flux. The predominance of the X-linked 5.1-kb NTS implies that there exists a balance between two opposing forces, the molecular mechanisms that generate new length variants and selection that eliminates X-linked arrays with too few 5.1-kb spacers. Therefore, NTSs on X chromosomes are held at 5.1 kb in spite of unequal exchange, not because of it. We emphasize that while unequal exchange creates new length variants it also homogenizes nucleotide sequence. The number of NTS lengths per chromosomal array suggests the frequency and therefore the importance of unequal exchange in sequence homogenization.

Selection on the spacer region is plausible because it has been demonstrated in Xenopus that some of these internal repeats act as enhancers for rDNA transcription (BUSBY and REEDER 1983; REEDER and ROAN 1985; LABHART and REEDER 1984; REEDER 1984; DE WINTER and MOSS 1986). Therefore, lengths of the spacers can affect the efficiency of transcription both at the level of an rDNA copy and for the entire rDNA system as a whole (REEDER, ROAN and DUNAWAY 1983). And it is well known that reduction in the number of transcribed rDNA genes has a large phenotypic effect on several species of Drosophila (RITOSA 1976; FRANZ and KUNZ 1981; DE SALLE, SLIGHTOM and ZIMMER 1986), demonstrating the importance of proper transcription for a normal phenotype.

Although the argument for selection on X-linked NTSs is strong, no such analogous holds for the Y-linked spacers. In fact, the amount of divergence among Y-linked spacers from the different regions suggests that they are evolving independently. This observation indicates that the similarities of the X-linked spacers cannot be caused by migration among the populations studied, strongly supporting the argument for natural selection. It remains to be explained why only X-linked spacers are under selective constraints. One possibility is that the X-linked ribosomal genes are preferentially transcribed. In such a system only the X-linked array would be under selection and the Y-linked genes would be able to diverge without ill effect to the individual.

We would like to thank DTS. M. MAHANAY, K. MORGAN, C. AQUADRO, A. TEMPLETON, and G. DOVER for reading and commenting on earlier drafts of this manuscript. We would also like to thank C. AQUADRO for supplying the fly stocks from around the world and G. DOVER for providing the rDNA spacer probe, pDmr 103HH2. M. DRYGAS helped with the figures. This work was supported by an Alberta Heritage Fund for Medical Research Fellowship (to S.M.W.), a Natural Science and Engineering Research Council (NSERC) of Canada Fellowship (to G.R.F.), and an NSERC grant (to C.S.).


Communicating editor: M. T. CLEG. 1


Communicating editor: M. T. CLEG.