The Y Chromosome of Drosophila melanogaster Contains a Distinctive Subclass of HeT-A-Related Repeats

Olga Danilevskaya,*† Ariel Lofsky,* Elena V. Kurenova† and Mary-Lou Pardue*

*Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and †Institute of Molecular Genetics, Russian Academy of Sciences, 123182 Moscow, Russia

Manuscript received November 12, 1992
Accepted for publication February 27, 1993

ABSTRACT

The HeT-A element is a transposable element with an apparent role in the structure of the telomeres of Drosophila melanogaster chromosomes. HeT-A transposition is the earliest event detected in healing of broken ends; HeT-A is also found on telomeres of unbroken chromosomes. Sequences with homology to HeT-A are never detected in euchromatic regions; however, clusters of HeT-A-related sequences occur in nontelomeric regions of the heterochromatic Y chromosome. Analysis of two of these Y-associated clusters shows them to be significantly different in structure from telomeric HeT-A elements, although the regions of shared sequence have >80% sequence identity in all cases. Telomeric HeT-A elements occur in chains, with the elements in the same orientation but variably truncated at their external ends and irregularly interspersed with unrelated sequences. In contrast, the nontelomeric Y elements are regular tandem repeats of parts of the HeT-A sequence joined to unrelated sequences which are not the same in the two clusters studied. The sequence structures suggest that the nontelomeric clusters on the Y chromosome do not arise by the same transposition mechanism that forms the telomeric clusters; instead the clusters on the Y may arise by a mechanism that is used more generally in the evolution of Y chromosomes. Although the telomeric and nontelomeric clusters appear to be formed differently, both are enriched in parts of the HeT-A sequence which may be important in the structure of heterochromatin.

HeT-A is a novel transposable element which appears to form part of the telomeres of Drosophila melanogaster chromosomes. This element transposes specifically to the broken ends of chromosomes to produce "healed" telomeres. HeT-A elements are always added in the same orientation with respect to the broken end. HeT-A transposition appears to be the first step in the reestablishment of the telomere; however, this reestablishment may be a complex process since telomeres of unbroken chromosomes contain multiple HeT-A elements as well as other sequences. Several HeT-A elements from telomeres of both "healed" and normal chromosomes have now been sequenced (Biessmann et al. 1992a). Elements from unbroken telomeres are variably truncated at the distal end and form a complex mosaic with other, less characterized, telomere-associated sequences (Valgeirsdottrir, Traverse and Pardue 1990; Karpen and Spradling 1992).

Some of the sequences that are associated with HeT-A elements at telomeres are also detected in the pericentric β heterochromatin by in situ hybridization to polytene chromosomes (Valgeirsdottrir, Traverse and Pardue 1990). In contrast, HeT-A sequences themselves are detected only in the terminal polytene band of each chromosome in such experiments. The high degree of replication of the euchromatic regions of polytene chromosomes makes it possible to detect <50 bp of homologous sequences within these regions by in situ hybridization. Thus, these experiments allow us to conclude that even partial HeT-A elements are absent in euchromatin. However, because some parts of the heterochromatin, including the Y chromosome, are underrepresented in polytene nuclei, these experiments do not allow strong conclusions about the presence of HeT-A sequences in nontelomeric heterochromatin.

The HeT-A sequences found in telomere regions are different from previously studied classes of repeated sequences. Typically, clustered repeated sequences consist of regular tandem repeats. In contrast, telomeric HeT-A clusters consist of irregular chains of elements. Within the chains, elements vary in size, being truncated from one end (and always the same end). The irregularity of the telomeric clusters is further increased by insertions which vary in size, sequence and position.

Previously it was reported that D. melanogaster telomeres contain typical tandem repeats; a clone of a 3.1-kb repeated sequence, pDm356, hybridized in situ to the telomeres of polytene chromosomes (Rubin 1977; Potter et al. 1979). More recent studies have shown that this in situ hybridization to telomeres was produced by cross-hybridization of a large segment within the 3.1-kb repeats that had a high level of sequence identity to the HeT-A elements in the telom-
ere regions. These 3.1-kb tandem repeats of pDm356 are actually located at internal positions on the Y chromosome (TRAVIS and PARDE 1989). We have concluded that the pDm356 clone is a member of a new subclass of sequences with HeT-A homology. This subclass is distinct from the subclass of telomeric HeT-A elements. It differs from the telomeric subclass in both structure (tandem repeats) and in location (nontelomeric sites on the Y chromosome).

We have now recovered three other clones of HeT-A-related sequences with tandem repeats. All appear to belong to the subclass identified by pDm356. One clone, AT-F, appears to be identical to the 356 repeat (TRAVIS and PARDE 1989). The second clone, Dm665, is different in sequence from pDm356 but also is found at internal sites on the Y chromosome. The 665 repeat is closely related to the repeat in the third clone, YAC 25-32, which has not been completely characterized (DANILEVSKAYA et al. 1991). In this report we compare two sets of these Y-associated non-telomeric sequences, 356 and 665, with a consensus sequence derived from HeT-A elements that are known to be on telomeres of other chromosomes. The Y-associated sequences show strong conservation (>80% identity) with sequences from portions of the telomeric HeT-A elements. Such homology indicates that the telomeric and the nontelomeric elements are derived from the same ancestral sequences. In spite of this relationship, the differences in structure between the two subclasses suggest that the origin of the clusters of Y-associated repeats is quite different from the transposition events which appear to generate the HeT-A element clusters at telomeres.

MATERIALS AND METHODS

Nomenclature: HeT-A elements make up one subfamily of HeT DNA. HeT DNA is a family of sequences found only in heterochromatin, identified by the ability to cross-hybridize with the 9.4 kb of D. melanogaster DNA in the clone, X-T-A (YOUNG et al. 1985; VALGEIRSDOTTIR, TRAVIS and PARDE 1990). HeT-A is the only subfamily of HeT DNA that has been studied at this time.

Sequence analyses: The HeT-A consensus was derived from comparison of 5 partial elements cloned from chromosomal telomeres. The alignment of these sequences has been discussed in our previous paper (BIESSMANN et al. 1992a). The sequence of the 665 repeat has been published (DANILEVSKAYA et al. 1991). We have sequenced the 356 repeat from clone p356-3B3 (POTTER et al. 1979) using the dideoxy chain-termination technique (SANGER, NICKLEN and COULSON 1977). Each strand was sequenced in both directions. Clone p356-3B3 contains a single repeat that appears to be identical to the multiple repeats in the clone we call AT-F (TRAVIS and PARDE 1989). The 356 and 665 repeats were cloned from convenient restriction sites occurring once per repeat. Since the ends of these tandem arrays are not known, the true start of each repeat is unknown. For comparison with HeT-A, each of the sequences has been circularly permuted so that the start of the HeT-A sequence identity is used as the start of the repeat. Database searches were done with the FASTA program (PEARSON and LIPMAN 1988). Other sequence analyses were done with programs from the University of Wisconsin Genetics Computer Group (DEVEREUX, HAEBERLI and SMITHIES 1984). The multiple alignment was made with the Multalin program (CORPET 1988).

Drosophila stocks: The X-Y translocation stocks were fertile reciprocal translocations between a B'Y chromosome and the proximal heterochromatin of an X chromosome (KENNISON 1981). Females of the stocks used had an attached-X chromosome and also carry one fragment of the X-Y translocation. To prepare DNA carrying the desired Y chromosome fragment, females carrying the Y chromosome fragment were selected by phenotype (B' or Y' depending on the Y fragment desired) and used for DNA preparation. Thus the only Y chromosome material was the selected fragment. The gt-l stock is the gt'w' stock of Bridges. Mutations are as described in LINDSLEY and ZIMM (1992).

Isolation of DNA: DNA was isolated from whole adult males and females as previously described (DANILEVSKAYA et al. 1991). For gt-l and Crimea stocks DNA was isolated from the heads of adult males and females as previously described (TRAVIS and PARDE 1989).

Restriction fragment analysis: DNA filters were prepared as described (TRAVIS and PARDE 1989). Hybridization was overnight at 60° in 4 X SET (1 X SET is 0.15 M NaCl, 0.03 M Tris-HCl, pH 7.0, 2 mM EDTA). A wash of 0.1 X SSC, 0.5% sodium dodecyl sulfate at 68°, used as a high stringency wash, was shown to prevent cross-hybridization between the 665 repeats, the 356 repeats, and telomere-associated repeats. Hybridization probes were made from gel-isolated DNA fragments labeled with [32P]dATP by random primer labeling (FEINBERG and VOGELSTEIN 1984).

RESULTS

Only a part of each Y-associated repeat contains homology with HeT-A sequence: The two Y-associated repeats, 665 and 356, are each found in tandem repeats. Although we do not know the number of repeats in any cluster, a cloned fragment with 5 copies of the 665 repeat has been analyzed (DANILEVSKAYA et al. 1991) while fragments with 3 and 4 repeats of 356 have been cloned (POTTER et al. 1979; TRAVIS and PARDE 1989). These clones give minimum estimates of the number of contiguous repeats in each cluster.

We have compared the sequence of each of the Y-associated repeats with the consensus sequence derived from five HeT-A elements that have been found in telomere regions. The HeT-A element has been defined by the sequences that have transposed onto two "healed" chromosome ends (BIESSMANN et al. 1992a). At this point no complete element has been isolated from a "healed" telomere; therefore the consensus describes only about 2 kb at the chromosome-proximal end of the HeT-A element and we have not been able to define the other end of the complete element. The chromosome-proximal 2 kb of HeT-A are unusual in that they contain no open reading frame; instead they have loosely conserved repeating motifs that suggest a structural function. It is interesting that each of the Y-associated repeats contains the repeated structure.
A sequences exist outside the 356 repeat, they must be limited to heterochromatin.

The 665 repeat contains approximately 650 bp of sequence with homology to the Stellate gene (Danilevskaya et al. 1991). The Stellate sequence is not a part of the HeT DNA family because Stellate is found at an euchromatic location, polytene region 12E on the X chromosome, as well as in the heterochromatin of both the X and Y chromosomes (Livak 1990; Balakireva et al. 1992; V. A. Gvozdev, personal communication). The 650 bp in the 665 repeat seem to have been pieced together from four fragments of Stellate sequence. Three of the fragments are from the Stellate coding region but they do not produce any large open reading frame in the 665 repeat. The 665 repeat also contains 94 bp with homology to the transposable element Copia.

The sequence identity between the Y-associated repeats and telomeric HeT-A elements is high and the sequence suggests a structural role: Within both the 356 and the 665 clusters the sequences appear to be evolving in concert. We do not have any measure of the degree of sequence divergence within a cluster. However we can detect cross-hybridization between repeats within a cluster at stringencies high enough to prevent cross-hybridization between sequences from different clusters. As discussed below, the between-cluster sequence identity is >80% for the HeT-A-related portions in all cases. Therefore the within-cluster sequence identity must be even higher.

If we consider only the segment of each repeat that shows HeT-A homology on the dotplot, the two Y-associated repeats can be easily aligned with the HeT-A consensus to produce long regions of strong sequence identity with some insertion/deletion gaps of moderate size. As noted above, the two repeats contain three partial HeT-A elements, one in the 665 repeat and two in the 356 repeat. If we ignore regions where sequences are gapped, the three Y-interior elements show >80% sequence identity with the telomere consensus. They also show about the same degree of sequence identity with each other (Table 1). The lowest level of sequence identity seen is the comparison of 356-2 with 665. These sequences overlap for only 90 nucleotides and show 78% identity. The percent of sequence identity and the numbers of gaps required for alignment seen for these four elements are very similar to those seen when telomere-associated HeT-A sequences were compared (Biessmann et al. 1992a).

The part of the telomere-associated HeT-A element that was defined in our earlier study (Biessmann et al. 1992a) has no evidence of open reading frames. This is very unusual for a transposable element; however, since we have characterized only one end of the HeT-A element, it seems possible that there may be coding regions in the chromosome-distal end. A likely can-
TABLE 1

<table>
<thead>
<tr>
<th>Sequences compared</th>
<th>No. identical bases/No. bases compared</th>
<th>Percent identity</th>
<th>No. gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>HeT-A/356-1</td>
<td>1607/1808</td>
<td>88</td>
<td>36</td>
</tr>
<tr>
<td>HeT-A/356-2</td>
<td>312/360</td>
<td>87</td>
<td>5</td>
</tr>
<tr>
<td>HeT-A/665</td>
<td>1051/1264</td>
<td>83</td>
<td>35</td>
</tr>
<tr>
<td>356-1/356-2</td>
<td>243/292</td>
<td>83</td>
<td>11</td>
</tr>
<tr>
<td>356-1/665</td>
<td>934/1157</td>
<td>81</td>
<td>35</td>
</tr>
<tr>
<td>356-2/665</td>
<td>70/90</td>
<td>78</td>
<td>4</td>
</tr>
</tbody>
</table>

Sequences in the alignment were compared by counting only those bases that were aligned with a base in the other sequence being compared. Percent identity = the fraction of identical sequences/the total number of sequences compared. No. gaps = number of gaps of one or more bases introduced to give the alignment used for the calculations.

...didate for the HeT-A coding region is the sequence which has been named the T element. This element contains two overlapping open reading frames with some similarity to those of other transposable elements (DANILEVSKAYA et al. 1992). T elements have been found only in telomeres and are always found in association with what we have defined as the HeT-A consensus but which we now know is only the chromosome-proximal 2 kb of the intact element (our unpublished results). In spite of the probability that the T element sequences were part of the complete HeT-A element, they were not included in the original definition of the HeT-A consensus (BIESSMANN et al. 1992a) because they had not been seen in the segments transposed to newly "healed" chromosomes. More recently BIESSMANN and colleagues (1992b) have detected sequences similar to the T element on a healed telomere, supporting our assumption that the T element is the HeT-A open reading frame and is actually part of the complete HeT-A element. It is interesting that, although these open reading frames appear to be part of the intact HeT-A element, they have not been detected in any of the tandemly repeated HeT-A-related sequences reported here.

Although the chromosome-proximal 2 kb of HeT-A sequence does not appear to be protein-coding, its conservation, both at telomeres and at other sites on the Y, argues that the sequence has a function. The sequence of the telomere-associated elements suggested a structural function. A residual short range repeat structure is detected throughout the sequence
in both the telomeric HeT-A elements and in the Y-associated repeats. It is readily visualized when two of the sequences are compared by dotplot (see Figure 2).

The most obvious difference between the sequence of 665 and the telomeric HeT-A elements lies at the end of the element that, on the healed chromosomes, forms the junction with the broken chromosome. On this end of the telomere elements, the DNA strand that runs 5' to 3' into the chromosome is terminated by an oligo(A) segment (3-28 residues in the elements studied). Because transposition of HeT-A is polar, this oligo(A) marks the inner, or proximal, end of the element. This structure strongly suggests that HeT-A transposes through an RNA intermediate and makes the junction by means of its poly(A) tail; however this has not been proven. HeT-A elements on unbroken telomeres also appear to be oriented with the oligo(A) end in the proximal (internal) position. The telomere elements are variably truncated at the end opposite the oligo(A), but all that have been studied, both from healed telomeres and from normal telomeres, have the oligo(A) end intact. Because the oligo(A) end was the end which all telomere elements had in common, we have adopted the convention of numbering from this end and, in order to present sequence in a 5' to 3' direction, used the sequence of the complement to the oligo(A)-containing strand. The first nucleotide is the one immediately 3' of the oligo(T) complement of the oligo(A). The first 130 nt of this end of HeT-A elements are exceptionally well conserved in all of the telomere elements sequenced (BIESSMANN et al. 1992a) but 685 shows no homology to the first 233 nucleotides of HeT-A. In contrast, 356, shows two regions of strong homology to the first part of the element, one in the first long region of homology and the other in the second short region of homology.

The Y-associated repeats are found in different clusters within the Y chromosome. The clusters appear to be evolving separately: Although the three HeT-A sequences described here have a high level of sequence identity, their sequences differ enough to suggest that they are evolving separately (Table 1). Under very stringent conditions we do not detect cross-hybridization between any of the three sequences. We have used such stringent conditions to map the different clusters of tandem repeats along the Y chromosome. Ideally such mapping would be carried out by in situ hybridization to metaphase chromosomes but we have not been able to achieve sufficiently stringent conditions in our in situ experiments to discriminate between these elements. Instead we have located the different repeat clusters by using a set of stocks with X-Y translocations produced by J. KENNISON (1981). Each of the translocations has a different breakpoint in the Y chromosome and the set allows the Y chromosome to be subdivided into several segments. In each translocation the Y chromosome fragment is joined to the heterochromatin flanking the X chromosome centromere. Since none of the Y-associated HeT-A sequences cross-hybridize with DNA from females under the conditions of these experiments, the X chromosome material is of no concern here.

The Y chromosome used to generate the translocations had been marked at the end of the long arm by Bar² and at the end of the short arm by yellow*. Both of these genes are derived from the X chromosome. When the marker genes were added to the Y chromosome, the telomeres of the Y chromosome were replaced with telomeres from the chromosome donating the marker genes. Therefore these translocation chromosomes can tell us nothing about the Y telomeres; however, the information about the interior regions of the Y appears to be reliable. Mapping with different translocation fragments is internally consistent (Figure 3) and fits well with restriction maps of Y chromosomes from several different stocks (Figure 4. Traverse and Pardue 1989).

The mapping experiment was performed with DNA from females carrying the fragment of the Y chromosome to be analyzed. This allowed us to analyze the Y fragments in the absence of an intact Y chromosome. Females showing expression of the gene marking the desired Y fragment were selected from the stock and used for DNA. Figure 3 shows the pattern of Southern blot hybridization to DNA from females carrying increasingly longer fragments of either the long or the short arm of the Y chromosome. Because the Y fragments are progressively longer in each population, each population provides some check on results with other populations. Any hybridization seen to a short fragment should also be seen with a longer fragment of the same arm. This redundant information is useful but there is a price. The most distal clusters on each arm can be mapped unambiguously but more proximal clusters may be masked unless they contain distinctive restriction fragments or are very abundant.

Hybridization to DNA from these translocation chromosomes (Figure 3A) localized the 665 cluster to a proximal region of the long arm of the Y, approximately to regions 10-13, on the map drawn by GATTI and PIMPINELLI (1983). This localization is consistent with an earlier report (DANILEVSKAYA et al. 1991) but, because the hybridization was done at higher stringency in the present experiment, several related clusters seen in the earlier experiment were not detected in this study. Even the high stringency conditions yield a doublet band of hybridized 665 probe, as well as several minor bands. The doublet band represents two clusters of repeats since each has been cloned separately (DANILEVSKAYA et al. 1991). The relation of the minor bands to the major bands has not been
FIGURE 3.—Analysis of X-Y translocations used to map clusters of 665 and 356 repeats on the Y chromosome. (A) Autoradiogram of Southern blot hybridization of DNA from females carrying different fragments of the Y chromosome, probed with the sequence of repeat 665. All DNA samples have been digested with BglII endonuclease and hybridized under the high stringency conditions given in MATERIALS AND METHODS. Only the two longest fragments of the long arm of the Y show bands of hybridized DNA. The longer fragment shows no evidence of having more 665 homologous sequence than the shorter fragment. Lanes 1–7 = DNA from females carrying the Y fragments shown in panel C. Lane 8 = DNA from females of the gt-1 stock. Lane δ = DNA from males of the gt-1 stock. (B) Autoradiogram of Southern blot hybridization of DNA from females carrying different fragments of the Y chromosome, probed with the sequence of repeat 356. All DNA samples were digested with EcoRI endonuclease and hybridized as in panel A. The repeat probe detects a complex set of bands. The complete set is present in the shortest fragment of both the long and short arm of the Y chromosome. The longer fragments of each arm show an apparently identical banding pattern, indicating that most, if not all, of the 356-homologous sequences are in the region covered by the small fragment of each arm. The 356 sequences must lie near the proximal end of the small fragments because the distal parts of these fragments are telomeres from the chromosomes donating the genes used to mark the Y chromosome (see panel C). The transferred sequences should also be present in the DNA of females lacking the Y fragments (lane 8) and this DNA does not hybridize with the 356 probe. The patterns of bands seen on these Y fragments are nearly indistinguishable from the patterns seen for Y chromosomes from the other D. melanogaster stocks studied with one exception; all of these fragments have one strong band of 356-homologous DNA (~10 kb) that has been seen only with derivatives of the BβY* and we believe to have arisen in the construction of this marked chromosome. Lanes marked as in A. λT-F = DNA from the λT-F clone which contains three repeats of the 3.1-kb 356 sequence. One repeat lacks an EcoRI site giving rise to the 6.2-kb band (TRaverse and PARdue 1989). The 1.6-kb band is a partial repeat generated in cloning. The other bands in this lane are partial digestion products. (C) Diagram showing the regions of the Y chromosome included in each of the fragments analyzed above. The schematic drawing of the Y chromosome is adapted from that of GATTI and PIMPINELLI (1983). The two hatched areas indicate the non-Y material translocated onto the chromosome to add the marker genes, B' and y+. The asterisk marks the centromere. Y' marks the long arm and Y marks the short arm. The thinner region indicates the nucleolus organizer. Lines above the drawing indicate the extent of the fragments in the DNA used in A and B. The numbers by each line correspond to the lanes in A and B. Dotted lines mark regions of uncertainty. The translocation stocks which produced the fragments are: 1 = T(X;Y)V24; 2 = T(X;Y)W27; 3 = T(X;Y)E15; 4 = T(X;Y)F12; 5 = T(X;Y)W19; 6 = T(X;Y)V8; 7 = T(X;Y)N12.
determined but some may represent end fragments from the clusters of repeats.

A similar filter probed with 356 sequences showed a more complex set of bands even under stringent hybridization conditions (Figure 3B). The prominent bands at 3.1 and 6.2 kb represent sequences that lie adjacent to each other on the fragment cloned in λT-F and thus appear to be part of the same cluster. The 6.2-kb fragment contains two of the 3.1-kb repeats which have lost the EcoRI restriction site between them. We do not know whether other bands are part of the same cluster. If not, these fragments must lie in closely adjacent clusters since the complete pattern is present on the smallest of the Y fragments.

The complete set of 356 bands is detected in the most distal fragment of both arms of the Y chromosome. No hybridization is seen to DNA from females without Y fragments, showing that all of the bands are Y-specific. They must therefore not be located in the new telomeres of this marked Y because these are derived from the chromosomes that donated the marker genes (indicated by hatched segments in the diagram in Figure 3C); instead the bands of 356 sequence appear to lie slightly proximal to the translocated segment on both arms. Longer fragments of the Y show nearly identical sets of bands, although the more heavily loaded lanes (5 and 6) reveal some very minor bands not seen in other lanes. The levels of Y-specific DNA in females from different stocks appear to be somewhat variable, suggesting that the extra Y chromosome fragments may not be maintained equally in the somatic tissues of females of different lines. However we see no clear increase in the amount of hybridizing material on the longer fragments. We conclude that most, if not all, of the 356 homologous sequences are clustered just proximal to the non-Y telomere segments that mark the two ends of the parent Y chromosome.

This translocation analysis shows a very symmetrical arrangement of 356 repeat clusters near the ends of the two arms of the Y chromosome. The pattern of bands hybridizing with 356 is nearly identical to the pattern seen with Y chromosomes from all of the other D. melanogaster stocks studied, although studies on intact Y chromosomes do not reveal whether the sequences are on one or both arms of the chromosome (TRAVERSE and PARDEE 1989). The history of the construction of the B'Yy+ (GATTI and PIMPINELLI 1983) suggests an explanation for the remarkable symmetry of the banding pattern. The first marker gene, y+, was translocated onto the long arm of the Y. Next an exchange between the long arm of one Y chromosome and the short arm of a second Y placed y+ on the short arm of the second Y. If the 356 clusters were close to the y+ fragment on the long arm, they might have been translocated with the y+ to the short arm. Because the chromosome receiving the y+ still had an intact long arm, it would now have 356 sequences on both ends, thus explaining the symmetrical pattern of bands. A final exchange placed B' on the long arm of the chromosome that had y+ on the short arm. We propose that this translocation occurred distal to the 356 cluster and thus left the doubly marked chromosome with a duplication of the 356 cluster.

The X-Y translocations do show one prominent band of hybridization not seen with any other Y chromosome (heavy band ~10 kb, Figure 3). This extra band is seen in each of the translocation chromosomes and thus seems to characterize the B'Yy+. The origin of the extra band on the B'Yy+ is not clear. It may have been present on the Y chromosome used to make the B'Yy+. Alternatively, perhaps there is a "hotspot" on the Y which accepted both of the marker translocations and had an associated sequence amplification event which generated the new band. Our mapping places the 356 repeat family very close to the telomere additions so the translocations might have affected adjacent material.

The HeT-A-homologous repeats in the interior regions of the Y chromosome appear to differ from telomere-associated HeT-A repeats not only in their origin but also in their maintenance: The pattern of restriction fragments seen with each of the different translocation stocks is highly conserved even though these stocks were made over 10 years ago (KENNISON 1981). In addition, the pattern is almost identical to that seen in similar experiments with D. melanogaster stocks, suggesting that the structure of this part of the Y chromosome is very stable (Figure 4A). (One exception would be the duplication of the entire pattern that we suggest was generated in constructing the B'Yy+.) The only notable difference between the pattern in the translocation stocks and the pattern seen for Y chromosomes in other stocks is the band which we suggest may have arisen in conjunction with the addition of the first marker gene.

In contrast to the conservation of the restriction patterns of the Y-interior sequences, D. melanogaster stocks show marked changes in the restriction fragments hybridizing with a telomere-specific HeT-A sequence. In the stocks compared in Figure 4B only one of the major bands of HeT-A hybridization is conserved in all three stocks. Note that, with this telomere probe, the patterns of male and female DNA are very similar, as expected since the telomere probe is detecting sequences on chromosomes present in both sexes. We assume that bands seen only in male DNA are from the telomere region of the Y and bands that are heavier in female DNA are from the X chromosome telomeres. The differences in the pattern of restriction fragments seen between stocks could be due to mutation within restriction sites, to transposition (which could result in either insertions or deletions), or to exchange between telomeres. At the
moment there is clear evidence for transposition in telomere regions (Bießmann et al. 1992b) but it seems likely that the other mechanisms may also contribute to the very fluid state of these telomere regions. The internal regions of the Y chromosome seem to be much less fluid than the telomere sequences, in spite of their repeated nature.

**DISCUSSION**

The location of HeT-A-related repeats in interior regions of the Y chromosome is in apparent contradiction to the telomere localization of HeT-A elements: One of the remarkable features of the HeT-A element is the specificity of its transposition. Several transposition events have now been detected and in each case the element has moved to the broken end of a chromosome. Multiple elements have also been found at the telomeres of unbroken chromosomes, suggesting that transposition can also occur to normal tips; however such events have not been documented. Furthermore, we have never seen any traces of HeT-A sequences in the euchromatic arms of any D. melanogaster stock. Why then do we find so much HeT-A-homologous sequence in regions of the heterochromatic Y chromosome that are topologically equivalent to the euchromatic arms of other chromosomes (regions we are calling “Y-internal”)? (We assume that the Y chromosome also has telomere-associated HeT-A elements but they are not included in this study.)

To resolve some of the questions raised by the Y-associated repeats, we have determined the sequences of two repeats that appear to be representative of two very abundant clusters on the Y chromosome and compared them with a telomeric consensus sequence. The differences in the bands of telomeric sequences detected in the three stocks contrast sharply with the conservation of the Y sequence pattern (A) seen for the same stocks. Open squares mark a few of the major bands of hybridization that differ between stocks. The marks are placed between the lanes of male and female DNA for each stock since the bands are present in both sexes. The few bands present only in male DNA may represent the telomere of the Y chromosome.

**Figure 4.**—Comparison of the conservation of genomic organization of telomeric HeT-A sequences with those in internal clusters on the Y chromosome. (A) Autoradiogram of DNA probed with the 356 Y-specific repeat. The DNA was cut with EcoRI and hybridized at the same high stringency used for Figure 3. The banding patterns of this Y-specific sequence are very similar in the three stocks. The only notable difference, the decrease in the 6.2 band and the increase in the 3.1 band in Crimea males, is probably due to the change in a single restriction site. (B) Autoradiogram of DNA probed with the sequence of the telomeric HeT-A element A1 (Bießmann et al. 1992a) HindIII and hybridized at the same high stringency used for Figure 3. The differences in the bands of telomeric sequences detected in the three stocks contrast sharply with the conservation of the Y sequence pattern (A) seen for the same stocks. Open squares mark a few of the major bands of hybridization that differ between stocks. The marks are placed between the lanes of male and female DNA for each stock since the bands are present in both sexes. The few bands present only in male DNA may represent the telomere of the Y chromosome.
which should tolerate repeated DNA (Wyman and Wertzian 1987). Even with such libraries, the 356 and 665 repeats (plus one YAC that has not yet been studied as thoroughly, Danilevskaya et al. 1991) are the only tandem repeats that have been found in our clone searches. They are also the only clones with HeT-A homology that show clear evidence of nontelomeric locations. Thus they represent a distinct subclass. We have not found tandem repeats containing the T element sequences, the presumed coding regions of HeT-A elements. This raises the possibility that only some parts of the HeT-A sequence can enter into the process that generates the Y-internal repeats.

It may be significant that the part of the telomeric HeT-A element that is included in the 356 and 665 repeats is the presumed structural region, containing short internal repetitive motifs which are easily detected in dotplots. If this region contributes to the heterochromatic structure of telomere regions, it might play the same role in the heterochromatic Y chromosome. We know little about the non-HeT-A sequences in each repeat. The Stellate sequences on the Y are thought to regulate production of a testis-specific protein from a gene on the X chromosome (Livar 1990). It is not clear whether the portions of Stellate sequence in the 665 repeat have any role in this regulation.

HeT-A transposition to interior sites may differ in mechanism from transposition to telomeric sites:

The interior locations of the HeT-A-related repeats discussed in this report are quite different from those of the telomere-specific HeT-A elements. We believe that this difference reflects a very different mechanism of origin for the repeats. The structure of telomeric HeT-A elements suggests strongly that these elements transpose through a polyadenylated RNA intermediate which attaches to chromosome ends by an oligo(A) segment. Neither 665 nor 356 have structures which are consistent with such a mechanism. 665 seems to lack the minimal structure for retrotransposition since it lacks the oligo(A) and about 200 nucleotides from this highly conserved end of HeT-A. In contrast, each 356 repeat has two oligo(A) ends and so should have required two adjacent events if it originated by retrotransposition.

The sequence homology between the Y-internal repeats and HeT-A indicates that they are closely related and it seems possible that the Y-internal repeats are derived from telomeric HeT-A elements. Before discussing how this might have happened, it is useful to consider the structure of HeT-A elements at telomeres. The initial study of "healed" telomeres showed that, with time after a healing event, populations began to show both longer and shorter telomeres on the healed ends (Biessmann et al. 1990), suggesting that HeT-A addition is a first step in a longer process. More recent studies have shown that the HeT-A elements on the healed ends become gradually shorter (Biessmann et al. 1992b). Because of this general shortening, the transposed sequences that have been used to define the HeT-A element have been incomplete and we do not know the structure, or the size, of the distal end of the intact element. Longer telomeres on "healed" chromosomes have been produced by occasional transposition of a second HeT-A element onto the first. The second transposition appears to involve the same mechanism as the initial event. As with the original transposition, the attachment appears to be independent of the sequence on the end of the chromosome (in this case, the chromosome end is the truncated end of the preexisting HeT-A element). The new junction is made by an oligo(A) sequence at the 3' end of one DNA strand on the proximal end of the second element. The prominent 3' oligo(A) strongly suggests that the HeT-A element moves via an RNA intermediate. The net effect of gradual truncation of the external end combined with sporadic transpositions of additional elements in the same polar orientation is to yield a linked chain of variably truncated HeT-A elements. The DNA in the lambda T-A clone, which appears to come from a normal X chromosome telomere, shows such a chain (Valgeirsdottir, Traverse and Pardue 1990; Biessmann et al. 1992a). We propose that such chains, with insertions of other sequences which may occur secondarily, are general components of the telomeres.

Both the 356 and the 665 repeat could have arisen from fragments of a telomeric HeT-A chain (Figure 5). The 356 fragment would include parts of two elements from the chain. The 665 fragment would be derived from within a single element. In either case the fragment might be transposed to the Y chromosome either before or after undergoing amplification into a chain of tandem repeats. The mechanism involved is not clear but such a mechanism appears to have been used repeatedly in the evolution of the Y chromosome for many kinds of sequences. The Y chromosome of Drosophila hydei contains tandem 0.2 kb repeats of sequence with homology to a segment coding for 26s ribosomal RNA (Huisjen and Hennig 1987). The Y chromosome of D. simulans contains a large block of 2.4-kb repeats of sequence homologous to the nontranscribed spacer of the genes for ribosomal RNA (Lohe and Roberts 1990). We suggest a similar transposition and amplification of fragments from telomeric heterochromatin. At least some of these amplified segments may be important in Y chromosome structure. McKee, Habera and Vrana (1992) have recently suggested that the Drosophila simulans repeats are involved in the meiotic pairing of X and Y chromosomes.

It would be interesting to know whether the amplification of the repeats occurs after integration in the Y chromosome or during the transposition, perhaps
HeT-A fragment transposed and was then amplified. The 665 repeat contains sequences with HeT-A portions of the 356 and 665 repeats. Alternatively, the non-euchromatic and one heterochromatic region on the cluster of Stellate sequences in the same segment of homology to the Stellate locus and there is another HeT-A sequence in each repeat might lie adjacent to the target site to produce the tandem repeats after insertion of the HeT-A fragment. Along with some adjacent Stellate sequence to yield the 665 clusters. However, Stellate sequences are found in at least two other places in the genome, in one euchromatic and one heterochromatic region on the X chromosome (Livak 1990; Balakireva et al. 1992; V.A. Gvozdev, personal communication). Thus we cannot eliminate the possibility that the Stellate sequences in the 665 repeat did not originate from the Y chromosome sequences. If any of the non-HeT-A sequences in the repeats occur as insertions into telomeric HeT-A chains, they could have been included in the original fragment transposed to the Y.

Multiple sites of HeT-A-related sequences may correlate with the general heterochromatic nature of the Y chromosome: The Y chromosome of D. melanogaster is heterochromatic over its entire length and, like much of the heterochromatin, does not polytenize in salivary gland nuclei. Therefore Y chromosomes are much more difficult to study than other Drosophila chromosomes. However the combination of genetic studies with analysis of mitotic chromosomes is building a picture of the structure of this unusual chromosome. Recently eight different simple sequence satellite DNAs have been shown to map in a complex pattern of clusters scattered along the length of the D. melanogaster Y chromosome (Bonaccorsi and Lohe 1991). These satellite sequences are localized in the pericentric heterochromatin of other chromosomes; the wide distribution of satellite sequences might correlate with the chromatin structure of the Y chromosome.

The HeT-A-related sequences, like the satellites, appear to be specific for heterochromatin. Since the segment of the HeT-A sequence amplified in these clusters is the segment with a potential structural function, these sequences might be in part responsible for the heterochromatic nature of the Y chromosome. In addition to the clusters mapped in this study, slightly lower stringency hybridization detects additional clusters of HeT-A-related sequences along the length of the Y chromosome (Danilevskaya et al. 1991).

The patterns of Y chromosome HeT-A-related repeats are strongly conserved in different D. melanogaster stocks: Both the 356 and the 665 probes hybridize at very high stringency with complex patterns of restriction fragments. We have not determined interrelationship of all of these fragments but can draw a few conclusions. The 3.1- and 6.2-kb bands in the 356 family appear to represent adjacent repeats that differ by the loss of an EcoR1 cleavage site in the 6.2-kb band. The two heavy bands in 665 appear to be from adjacent clusters of repeats since they have been separated by cloning in YACs. Surprisingly, the complex structure of the clusters of these repeats is almost completely conserved in the different D. melanogaster stocks that we have studied. The stocks include Oregon R, gt-1, Canton S and Crimea (this study; Traverse and Pardue 1989). These stocks have been separated long enough to have quite different distributions of mobile elements in their euchromatic regions (e.g., Dawid et al. 1981). By comparison, the HeT-A-related clusters appear relatively stable. The most notable change in the hybridization pattern is seen for the Crimea Y chromosome which shows less of the 6.2-kb band with the 356 probe but more of the 3.1-kb band, suggesting that more of the repeats contain the EcoR1 cleavage site. The X-Y translocation chromosomes have been carried in stocks for over 10 years yet they, too, maintain the general patterns. The only significant difference from the pattern of normal Y chromosomes is the extra band of hybridization that we suggest may have been generated in a translocation event in the creation of the B*Yy' Y chromosome. We also suggest that the short arm of the B*Yy' contains an exact duplication of the 356 cluster from the long arm.
The most obvious exception to the apparent stability of Y chromosome organization was seen with our earlier study of the 356 repeat, which was analyzed using the \( \mu T-F \) clone. In that study the 356 repeat was mapped on the Y chromosome by using two deleted Y chromosomes, S10 and S12 (Traverse and Pardue 1989). The conclusion was that the 356 sequence was found only on the short arm of the Y, predominantly in the pericentric region, and that part of the 356 sequence had been lost from the short arm of the deleted chromosomes. The experiments reported here have caused us to reevaluate the earlier experiment. Our more recent results suggest that one and probably both of the deleted chromosomes used in those earlier experiments had undergone more than a simple deletion and that the conclusions drawn from them may not be valid for other Y chromosomes (unpublished results). In contrast, the X-Y translocations used in this work give consistent and therefore presumably reliable mapping. The restriction maps of each of the translocations are also consistent with the map of the intact Y chromosome in the stocks studied.

It is worth noting that the X-Y translocations used here are each carried as the only Y chromosome in the stock and are therefore constantly selected for function. On the other hand the deleted Y chromosomes (S10 and S12) used in our earlier study were carried in stocks with an attached XY to provide the necessary fertility factors. Thus the deleted Y chromosomes were not essential and may have been able to undergo sequence rearrangement more freely. It is possible that the more constant nature of the restriction pattern of the X-Y translocations used here is a result of the strong selection for function.

In contrast to the Y-associated repeats, the \( \mu T-A \) elements on the telomeres of these stocks show significant stock-specific differences in restriction patterns. There is mounting evidence that telomeres in yeast and humans are dynamic structures (Zakian 1989). Our studies suggest that Drosophila telomeres are also fluid.

**Does the generation of the Y-associated repeats reflect general mechanisms in the evolution of Y chromosomes?** The two Y-associated repeats analyzed here appear to represent two separate events in which part of the \( \mu T-A \) element has been transposed and amplified. In neither case does the \( \mu T-A \) fragment suggest typical retroposition. It is interesting to compare these studies with studies on the Y chromosome in Drosophila miranda. In D. miranda an autosome has become attached to the Y chromosome sometime within the last 5 million years. The attached autosome is gradually acquiring the characteristics of the Y chromosome, displaying changes in chromosome structure as well as loss of gene activity. The chromosome is now referred to as the neo-Y chromosome. In contrast, the homolog of the autosomal part of the neo-Y appears to be evolving into an X chromosome (designated X2). Some genes on the X2 are now capable of dosage compensation, the enhanced transcription in males that compensates for the inactivity of the homologous gene on the neo-Y.

Steinemann and Steinemann (1992) have analyzed the region around the four larval cuticle protein (Lep) genes from the X2 and compared this to the homologous region from the neo-Y chromosome. The striking difference between the regions of the two chromosomes is the large number of insertion sequences found on the neo-Y but not on the X2. Two of the insertions have sequences similar to known retroposons; however, the sequences of most of the insertions yield no clues to transposition mechanism. This larger group of insertions might represent intermediate steps in the formation of the Y-associated clusters that we have found in D. melanogaster (e.g., a transposition that has not yet been amplified).

One of the D. miranda neo-Y insertion sequences was used as a probe for in situ hybridization to D. miranda polytene chromosomes. It showed heavy hybridization over the formerly autosomal sequences that are now on the neo-Y and only partially polytenized in salivary gland nuclei. In contrast, the homolog, X2, shows approximately the same density of sites of hybridization as do the autosomes and X chromosome. Thus in the relatively short evolutionary time since the autosome was translocated onto the Y chromosome, the sequence of the homologs has evolved very differently; the sequence on the neo-Y has acquired many more copies of this insertion sequence whereas its homolog, the X2, has not.

Because the autosomes that now form the neo-Y and the X2 in D. miranda were homologs in the ancestral stock, their current differences reflect, not the starting DNA sequence, but the new character as an X or Y chromosome. The increased level of transposed sequences is consistent with what we now know of the structure of the Y chromosome in D. melanogaster, suggesting that the autosomal sequences on the neo-Y are in the process of acquiring similar structure. Why should the autosomal sequence acquire these repeated sequences when attached to the Y chromosome when the free autosome does not? None of the studies has given a clear answer to this question; however, there is evidence that, whatever the mechanism may be, it does not apply to all types of repeated sequences. Long repeats of mono- and dinucleotides show a behavior that is diametrically opposed to that of the more complex repeats found on Y chromosomes. The mono- and dinucleotide repeats actually become increased on the autosome that becomes the X2, so that levels approach the higher levels seen on all other Drosophila X chromosomes. In contrast, the neo-Y continues to show the level of these repeats.
typical of autosomes (LOWENHAUPT, RICH and PARDOE 1989). Thus it appears that sequence, rather than simply the repeated nature of the sequence, is a factor in the accumulation of the repeats on the Y chromosome.

CONCLUSION

Studies on polytene chromosomes have given abundant evidence that sequences of transposable elements are found in heterochromatic regions. It is usually assumed that these sequences represent elements that have jumped into heterochromatin only to be inactivated by the chromatin environment and trapped. Fragments of transposable elements recovered from heterochromatin are thought to result from accidental degradation of this excess baggage of trapped elements.

Our results suggest a different view. For at least one class of transposable element, HeT-A, the association with heterochromatin may not be accidental. Instead, sequences of this element may have a role in forming heterochromatin. HeT-A differs from known transposable elements in having >2 kb of conserved noncoding sequence. This noncoding sequence has a short-range repetitive nature that could influence chromatin structure. In this study we have characterized fragments of the HeT-A element recovered from nontelomeric heterochromatin. These fragments do not appear to result from accidental degradation of the element. Instead they show preferential amplification of the presumed structural DNA segment. Restriction fragment analysis of Y chromosomes from distantly related D. melanogaster stocks shows that these amplified regions are remarkably stable components of the Y chromosome in this species. Taken together, these observations suggest that these repeats may contribute to the structure of the heterochromatic Y chromosome.

This work has been supported by grants from the American Cancer Society and the National Science Foundation. Sequence analysis has been done through the Whitaker College Computing Facility of the Massachusetts Institute of Technology.

LITERATURE CITED


SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequenc-
Y Chromosome HeT-A-Related Repeats


Communicating editor: P Cherbas