Volume 13 Number 2 1985 Nucleic Acids Research

Distinct characteristics of loop sequences of two *Drosophila* foldback transposable elements

Heidi L.Brierley* and S.Steven Potter

Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06457, USA

Received 10 September 1984; Revised and Accepted 18 December 1984

ABSTRACT

A few foldback (FB) transposable elements have, between their long terminal inverted repeats, central loop sequences which have been shown to be different from FB inverted repeat sequences. We have investigated loop sequences from two such FB elements by analyzing their genomic distribution and sequence conservation and, in particular, by determining if they are normally associated with FB elements. One of these FB loop sequences seems to be present in a few conserved copies found adjacent to FB inverted repeat sequences, suggesting that it represents an integral component of some FB elements. The other loop sequence is less well-conserved and not usually associated with FB inverted repeats. This sequence is a member of another family of transposable elements, the HB family, and was found inserted in an FB element only by chance. We compare the complete DNA sequences of two HB elements and examine the ends of four HB elements.

INTRODUCTION

Transposable elements, units of DNA which can move to new locations in the genome, have been studied in many procaryotic and eucaryotic systems (for reviews see 1, 2). In Drosophila melanogaster several families of these elements have been characterized (3) including the foldback (FB) family (4). Foldback elements are among the most interesting transposable elements both because of the many activities associated with them and because of their unusual structures. They can, of course, insert into genes and inactivate them. These mutations are particularly interesting because they revert at a high frequency by precise excision of both the element and one copy of the target site duplication (5). Also, study of the w^{DZL} mutation (6,7) and derivatives of the w^{c} mutatiion (8) suggest that FB elements can generate rearrangements of adjacent DNA . Furthermore, two FB elements can cooperate to move large blocks of DNA around the genome (9).

Structurally, the FB elements have long inverted terminal repeats and sometimes a central loop between the repeats. The lengths of these repeats and loops vary from element to element. We have previously shown that the

inverted repeats carry a peculiar organization of sequences with highly conserved complex sequences at the termini which gradually change into simple 31 base-pair (bp) tandem repeats as one moves toward the center of the element (10, 11).

The FB loops are perhaps even more unusual. Restriction mapping data suggest that often a loop structure seen by electron microscopy is simply caused by the conjunction of a longer inverted repeat with a shorter one in the same FB element (10). The excess tandem repeats from the longer end constitute the loop structure. Very few FB elements have been demonstrated to carry loop sequences that are different from FB inverted repeat sequences (11, 7).

We have investigated the nature of the loop sequences contained in two such FB elements. One of these elements is the FB which Levis et al (7) associated with white crimson (w^c), an unstable allele of the white locus; we designate this element FBw^c. Levis et al (7) had previously shown that the loop of FBw^c was not homologous to FB inverted repeats or to the loop of FB4. The other FB element reported on is FB4; its DNA sequence revealed an interesting structure within the loop (11). There is an open reading frame with characteristic gene punctuation signals in a region flanked by short inverted repeats. The presence of this structure raised questions concerning the conservation of this potential gene region and its distribution in the genome. Is this structure an integral part of certain FB elements or is it an independent transposable element? We investigated each of the two FB loop sequences with respect to its genomic distribution and its association with FB inverted repeat sequences to determine if the sequence was usually between FB inverted repeats or if it was found there only by chance.

Our data suggest that the FBw^c loop sequences are indeed integral components of some FB elements and not themselves independent elements. Some interesting observations concerning the structure and evolution of FB inverted repeat sequences also came from this study. Our investigation of the FB4 loop sequence, however, indicates that it represents a member of an independent family of transposable elements which we designate HB elements. Many other homologous members have been found without any associated FB inverted repeat sequences. FB4, therefore, is an unusual composite structure with one transposable element (an HB) inserted into another (an FB). In this report we compare the complete sequences of two HB elements and analyze the ends of four HB elements. Furthermore, we describe an example in which internal HB sequence is repeated in a flanking region near an HB element.

MATERIALS AND METHODS

Materials

Enzymes were purchased from New England Biolabs. Isotopes were from New England Nuclear. Nitrocellulose filters were from Schleicher and Schuell. Agarose was from Sigma, and acrylamide was from Biorad.

DNA Preparation and Filter Hybridization

The construction of the recombinant pBR322-*Drosophila melanogaster* DNA library and the preparation of the filters for screening this library have been previously described (12, 13). Plasmid DNA was prepared by the rapid boil method of Holmes and Quigley (14) with previously described modifications (15). Restriction site mapping was done by size analysis of segments generated by single and double digests. DNA segments were prepared by using low-melting agarose as previously described (10). Transfer of DNA from agarose gels to nitrocellulose filters was done in the manner of Southern (16). Sometimes one gel was used to make two filters by placing additional nitrocellulose and blotter paper below the gel and eliminating the 20X SSC supply (17).

DNA segments were nick translated in the manner of Rigby et al (18) using [α-^{32}P]dNTP from NEN. The labelled DNA was either passed over a Sephadex G50 column or ethanol precipitated in the presence of 0.2 M sodium acetate. Filters were presoaked in 1X Denhardt's solution (19) (1X = 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin, 0.02% ficoll) and 4X SSCP (1X = 0.015 M sodium citrate, 0.12 M NaCl, 0.02 M sodium phosphate (pH 7)) for at least 4 hours at 65°C. The nick-translated DNA was allowed to hybridize to the filters in 1X Denhardt's, 4X SSCP, and 0.1 mg/ml sonicated salmon sperm DNA for 15-24 hours at 65°C with agitation. The filters were washed extensively in 1X Denhardt's and 4X SSCP, with multiple wash changes, at 65°C with gentle agitation. The second to last wash was done at 55°C, and the final wash was done in 2X SSC (1X = 0.015 M sodium citrate, 0.15 M NaCl) at room temperature.

DNA Sequence Analysis

DNA segments were 3' end labelled with [α-^{32}P]dNTP by using the large fragment of *E. coli* DNA Pol I (20) in the manner of Levis et al (21). DNA sequencing was done by the procedure of Maxam and Gilbert (22); the chemical reactions for G, T+C, C, and A>C were used. 75 cm long gels of 4-20% acrylamide were used. The Stanford University SEQ computer program was used in the sequence analysis.

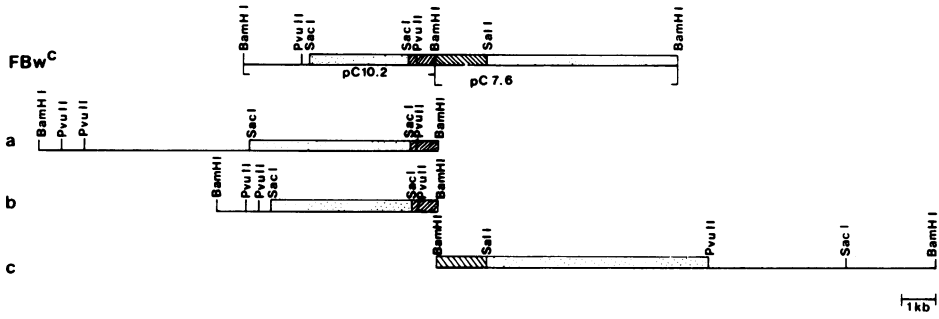

Figure 1. Restriction maps of FBw^{c} clones (pC10.2 and pC7.6) and homologous clones. Homologous regions are shaded similarly. The striped regions are loop segments, and the stippled regions contain FB inverted repeat sequences.

RESULTS

FBw^{c} Loop Sequences

An FB element with a loop which is not homologous to either the loop of FB4 or the FB inverted repeats has been associated with the white crimson^{c} allele by Levis et al (7). If the loop sequences of this element, FBw^{c}, are integral components of FB elements, one would predict that they would generally be found within FB elements rather than independent of them. Two segments from the FBw^{c} loop, the 0.8 kilobase-pair (kb) Sac I-Bam HI segment from pC10.2 and the 1.5 kb Bam HI-Sal I segment from pC7.6 (top of figure 1), were used to screen a pBR322-*D. melanogaster* DNA library for homologous clones. This library was constructed previously using partially Bam HI-digested DNA from wild type Oregon R flies (12). The results of this screen indicated that in the *Drosophila* DNA from which this library was made sequences homologous to the FBw^{c} loop were present approximately two times per haploid genome (6 positive clones/42,500 screened clones). DNA from hybridizing clones from this screen was digested, electrophoresed through agarose, and Southern blotted. Each of the two FBw^{c} loop segments and the 1.4 kb Pvu II-Hind III segment containing the right inverted repeat from FB4 (top of figure 5) were nick translated and separately hybridized with the Southern blot filters.

The resulting maps are shown in figure 1. The segments with homology to the FBw^{c} loop segments were bordered by the same restriction sites and of the same size as the probe segments. Each of the clones with homology to one of the loop segments also contained adjacent homology to FB inverted repeat sequences. Thus, at least within the limits of these experiments, these FBw^{c} loop sequences were conserved and always associated with FB inverted repeat

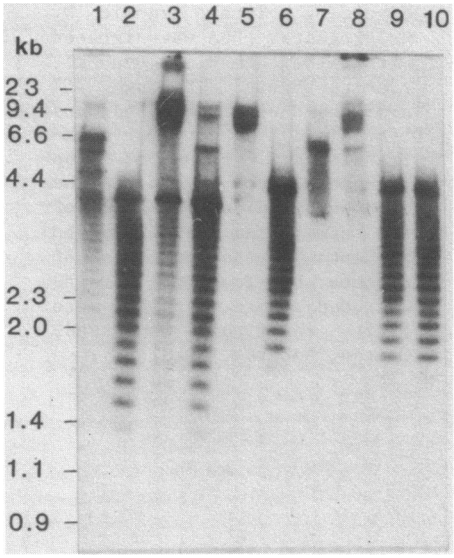

Figure 2. Restriction digests of an FBw^{c}-homologous clone, "b" in figure 1, were electrophoresed through 1.2% agarose, Southern blotted, and hybridized with nick-translated FB inverted repeat sequences from FB4. A different digest was run in each lane; the enzymes used were: Bam HI + Sac I (lane 1), Pvu II + Sac I (lane 2), Sac I (lane 3), Sac I + Sal I (lane 4), Sal I (lane 5), Pvu II + Sal I (lane 6), Bam HI + Sal I (lane 7), Bam HI (lane 8), Bam HI + Pvu II (lane 9), Pvu II (lane 10).

sequences. The restriction sites in the flanking DNA vary among the clones, indicating different chromosomal positions. Each clone contains only one half of an element because Bam HI segments were cloned and the FBw^{c} loop contained a Bam HI site.

In autoradiographs from the hybridization of the FB inverted repeat probe to Southern blots of the FBw^{c}-homologous clones, we observed ladder patterns in certain lanes. An example is given in figure 2; the other ladders were similar in appearance. Stained with ethidium bromide, the gels from which these Southern blots were made did not show bands corresponding to those in the ladders. However, we did notice that in those lanes which showed ladders on the autoradiographs the DNA segments which contained the FB inverted repeat sequences gave bands which were of relatively weak intensity for the segments' sizes. The ladders appeared in Sac I and Pvu II digests of two clones, pC10.2 and "b" in figure 1, which contain the left sides of FBw^{c}-homologous elements. The third clone containing the left side of one of these elements, "a" in figure 1, gave ladders in only the Sac I digests. The two clones, pC7.6 and "c" in figure 1, which contain the right sides of FBw^{c}-homologous elements did not give ladders. It has been previously shown by DNA sequencing that the FB inverted repeats have a hierarchical structure in which five slightly different types of 31 bp repeats are arranged such that every fifth repeat is of the same type (11). This gives these sequences a periodicity of 155 bp. The spacing of the bands in the ladders of the

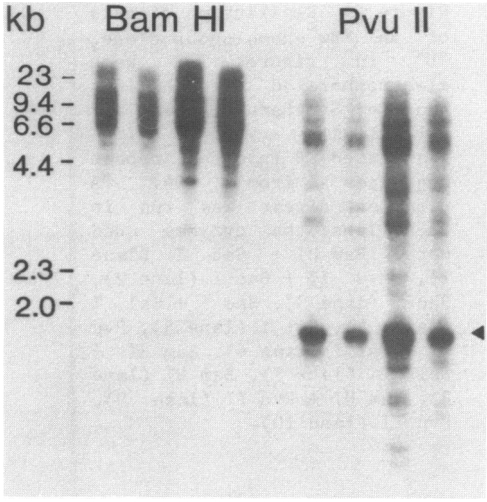

Figure 3. Total genomic D. melanogaster DNA was digested with either Bam HI or Pvu II, electrophoresed through .8% agarose, Southern blotted, and hybridized with nick-translated 1.6 kb Pvu II loop segment from FB4. The arrowhead indicates bands of the same size as the probe segment. In the two sets of digests DNA from four strains was used; these strains were (left to right) Samarkand, Canton, Amherst, Swedish.

FBw^{c}-homologous clones is approximately equal to this 155 bp repeat length. In an additional set of experiments DNA from each of the five clones was digested separately with Sac I and with Pvu II. The digests were electrophoresed through agarose, Southern blotted, and hybridized with the FB inverted repeat probe. The results of these experiments confirmed those of the earlier experiments. Since the agents responsible for the ladders seemed possibly to be different contaminants of Sac I and Pvu II, a similar set of experiments was done using Sac II and Pvu I. This set of experiments yielded neither ladders nor DNA bands near 155 bp in size.

FB4 Loop Sequences

The loop of FB4 includes an open reading frame which might code for a transposition factor. The 1.6 kb Pvu II segment from this loop (top of figure 5) contains this potential gene region. In order to study the genomic organization of the FB4 loop sequence, the Pvu II segment was nick translated and hybridized to Southern blots of digested, total genomic D. melanogaster DNA. In one case the DNA was digested with Pvu II; in another it was digested with Bam HI, which does not cut within FB4. The results (figure 3) showed: 1) homologous sequences in genomic Pvu II segments of several different sizes in addition to the size of the probe segment, indicating that the probe sequence was not strictly conserved, and 2) homologous sequences in genomic Bam HI segments in a range of sizes, indicating that the probe sequence was present in a number of different sites in the genome. The patterns of hybridization varied only slightly among the four strains of

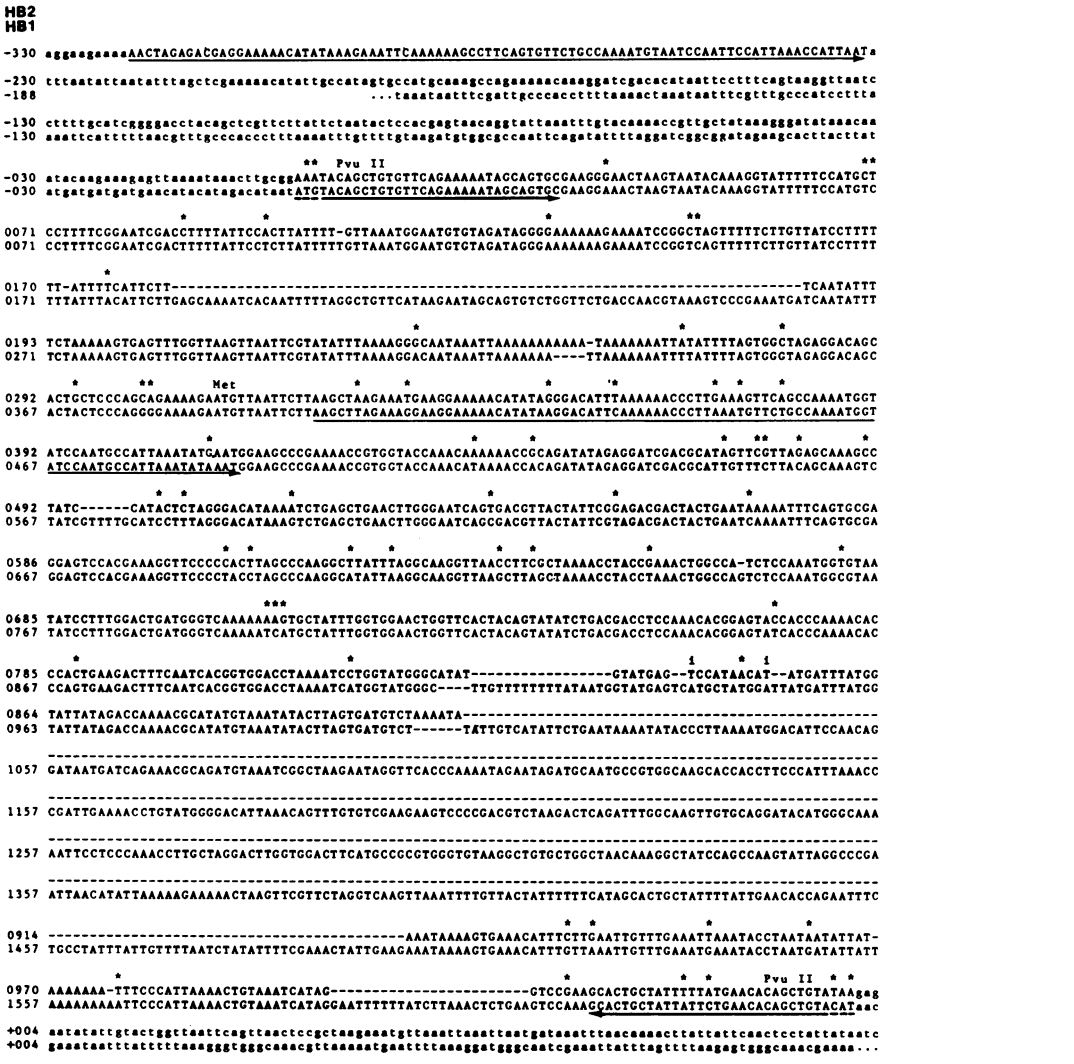

Figure 4. The DNA sequences of elements HB2 and HB1 are compared. The HB sequences are capitalized. Where a base differs in the two elements this is indicated by an asterisk (*); where a base is absent from one element this is indicated by a dash (-). The ten bases marked off by the i's are inverted in HB2 relative to those in HB1. The bold shorter arrows underline the HB terminal inverted repeats; the precise outside ends are not known. The thin longer arrows underline sequences which are contained within the elements and are directly repeated externally to the left of HB2. Some flanking sequences are given in lower case, including 100 bp of each of FB4's inverted repeats outside of HB1. The Pvu II sites near the ends of the HB elements are indicated, as well as the start codon for the open reading frame in HB1. The first translation termination signals start at base 360 in HB2 and at base 831 in HB1.

D. melanogaster that were tested (figure 3).

The FB4 Pvu II loop segment was also used to select an homologous clone from a previously constructed pBR322-D. melanogaster DNA library (12). On this selected clone the primary homology was localized to a 1.0 kb Pvu II segment, instead of a 1.6 kb one as in FB4. The restriction maps differed outside this Pvu II segment also, and this second clone did not contain homology to FB inverted repeat sequences.

The sequence of the entire FB4 element had been determined previously (11) so, in order to more closely analyze the relationship of the two homologous clones, the 1.0 kb Pvu II segment from the second was sequenced. Much of the two Pvu II segments is homologous. Short inverted repeats flank the homologous regions within each of the Pvu II segments. These homologous elements with short inverted repeats have been designated HB1 (inside FB4) and HB2. A comparison of the sequences of HB1 and HB2 is given in figure 4. The primary difference is that one large section (493 bp) and several smaller ones present in the HB1 sequence are deleted from the HB2 sequence. There are also many individual base differences between the two sequences. One such change introduces a stop signal in HB2 only 48 bases after the start signal; the open reading frame in HB 1 continues for an additional 396 bases.

To further investigate the nature of the association of HB sequences with FB inverted repeat sequences and the possible relationship of HB sequences with the FBw^{c} loop sequences, we used the 1.0 kb Pvu II segment from HB2 (figure 5) to thoroughly screen the pBR322-D. melanogaster DNA library. The results of this screen indicated that sequences with homology to this segment were present in approximately 20 copies per haploid genome

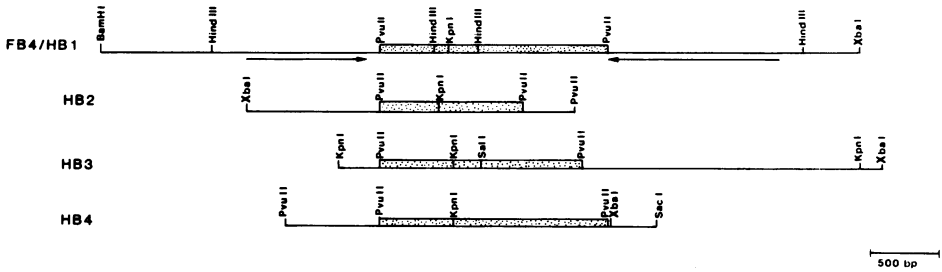

Figure 5. Restriction maps of regions containing HB elements. The element HB1 is within the loop of the element FB4. Homologous Pvu II segments are shaded; the Pvu II sites which bound them are near the ends of the HB elements. The arrows indicate the locations of FB4's inverted repeats. Hind III sites are only indicated in the top map.

(40 positive clones/42,500 screened clones). DNA from randomly selected hybridizing clones was prepared, digested, electrophoresed through agarose, and Southern blotted. Quadruplicate filters were used in four separate hybridizations with four nick-translated segments. These probes were the 1.0 kb Pvu II segment from HB2 (figure 5), the 1.4 kb Pvu II-Hind III segment containing the right inverted repeat from FB4 (figure 5), the 0.8 kb Sac I-Bam HI segment from pC10.2 (figure 1), and the 1.5 kb Bam HI-Sal I segment from pC7.6 (figure 1); these last two segments were from the FBw^{c} loop.

The Pvu II segment hybridized to 23 selected clones, but the FB inverted repeat segment hybridized to none of these 23 new clones. This indicated that HB sequences were generally not associated with FB inverted repeat sequences. The segments from the FBw^{c} loop did not hybridize to any of the selected clones. This suggested that the loop sequences from FB4 and FBw^{c} were not related.

DNA from several of these new homologous clones was digested, electrophoresed through agarose, and Southern blotted. These clones were restriction mapped, and the Southern blot filters were used in a hybridization to localize the regions of primary homology to the 1.0 kb Pvu II segment. On many of the analyzed clones the regions of homology are Pvu II segments with approximate sizes of 1.0, 1.4, and 1.6 kb. We have sequenced parts of two of these new clones to better understand the ends of the homologous regions. It should be noted that on a few clones the HB homologous Pvu II segments were larger than 1.6 Kb, but restriction mappings and Southern blot hybridizations showed that these HB elements were no longer bordered by Pvu II restriction sites and these clones were not studied further. Restriction maps of the regions containing four HB elements with homologous Pvu II segments are given in figure 5. Elements HB3 and HB4 were shown by sequencing to have short terminal inverted repeats. The sequences of the ends of four HB elements are listed in figure 6. These ends are not

```
          outside                                    inside
HB1  left  GACATAATATGTACAGCTGTGTTCAGAAAAATAGCAGTGC|GA
     right TAAAGCAATACATGTCGACACAAGTCTTATTATCGTCACG|AA

HB2  left  ACTTGCGGAAATACAGCTGTGTTCAGAAAAATAGCAGTGC|GA
     right TATAAGAGAATATGTCGACACAAGTATTTTTATCGTCACG|AA

HB3  left  TTTAAAGAATATTCAGCTGTGTTCAATGAAATAGCAGTGC|GA
     right CACGTACATATATGTCGACACAAGTATTTTTGTCGTCACG|AA

HB4  left  TCGTACATATCTACAGCTGTGTTCAGAA AATAGCAGTGC|GA
     right GATCTATATACATGTCGACACAATTCTTTTTATCGTCACG|AA

           35344335774878888888888785776887888888888 44
```

Figure 6. DNA sequences of the ends of four HB elements. Both left and right end sequences are from the same strand. The vertical lines mark the inside ends of the HB inverted repeats. The precise outside ends have not been defined. The bottom row indicates the number of bases out of the eight that agree at each position.

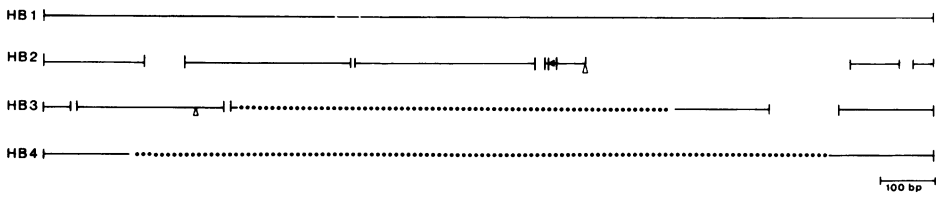

Figure 7. HB2 and those portions of HB3 and HB4 which have been sequenced are compared with HB1. Horizontal lines in the bottom three maps represent sequences which are also present in the corresponding locations in HB1. Dotted sections have not been sequenced. A gap between two short vertical lines represents a deletion of sequences which are present in the corresponding location in HB1. An open triangle indicates an insertion of sequences not present in HB1; the length of the base of the triangle indicates the length of the insertion. The line segment between two short vertical lines and containing an arrowhead represents an inversion in HB2 relative to HB1. Sequence differences of 5 bp or more in length are indicated.

fully conserved. The innermost base of the inverted repeats is apparent, but the outermost base is not obvious. The inverted repeats may be 29-32 bp in length, and there may or may not be direct target site repeats. Those portions of HB3 and HB4 that were sequenced are diagrammed in figure 7 which compares four HB elements. The data presented in this figure indicate that different HB elements have different deletions when compared to HB1. The HB elements also have many base changes with respect to one another.

In both Southern blot analysis and sequence data we noticed an interesting secondary homology on the clone containing HB2. Approximately 250 bp outside of the left end of the element there is an imperfect direct repeat of about 90 bp of sequence found inside the element (figure 4). This external repeat has 78% homology with the internal copy and 83% homology with the corresponding copy inside HB1. The two internal sequences themselves share 91% homology in this region.

DISCUSSION

Many FB elements contain only FB inverted repeat sequences without any detectable different loop sequences, as shown by electron microscopy and restriction mapping (10). Such FB elements may be deletion products of other larger FB elements, and certain FB loops may contain a gene for an FB-specific transposase which also acts on incomplete FB elements. A situation similar to this has been described for P elements by O'Hare and Rubin (23). To better understand the loop sequences found in FB elements we

have investigated two different loop sequences from FB elements as to their genomic representation and their association with FB inverted repeat sequences.

We isolated three clones with homology to the FBw^c loop (figure 1). In these clones the restriction sites external to the elements and, therefore, the elements' genomic locations vary from those of FBw^c. This is as expected since the library we screened was made with DNA from flies without the w^c phenotype. We estimated that FBw^c loop sequences were present approximately twice per haploid genome in non-w^c flies. For the w^c strain Levis et al (7) reported four, or possibly five, sites with FBw^c loop sequences while Paro et al (9) reported six sites.

The sizes of the segments homologous to the FBw^c loop segments and the patterns of the restriction sites in the central loop regions imply that the FBw^c loop sequences are conserved. This apparent conservation and the fact that these sequences were found only in association with FB inverted repeats indicate that the FBw^c loop sequences may be functionally important to FB elements. We suggest that the FBw^c loop may code for functions involved in the regulation of FB transposition. Paro et al (9) have associated both the FB inverted repeats and the FBw^c loop sequences from FB-NOF with several (historically related) white locus transposons. It should be noted that the FB inverted repeats on the FBw^c-homologous clones may not extend the full length of the segments shown to contain them in figure 1. Levis et al (7) report that those of FBw^c do not. The central loop region may exceed the homologous segments indicated in figure 1, and the outer ends of the inverted repeats may be well inside the restriction sites which bound the segments homologous to the FB inverted repeat sequences. On the left side, however, the outer ends of the repeats may reach the Sac I sites since Sac I sites are often present at such junctions.

Ladder patterns of hybridization of FB inverted repeat sequences to FBw^c-homologous clones were observed in specific cases. Only the three clones containing the left sides of the elements showed these patterns, and only when they were digested with Sac I or, for two out of the three, with Pvu II. The right side clones and Bam HI- or Sal I-digests of the left side clones did not give ladders. These interesting results suggest that the two sides of an FB element can be quite distinct, with particular sequence patterns evenly distributed on one side and completely absent from the other side. Figure 2 shows hybridization patterns from one of the left side clones, "b" in figure 1; ladders are present in the lanes where Sac I or Pvu

II was included and are absent from the other lanes. The DNA of the ladder bands which hybridized to FB inverted repeat DNA must have come from the primary segment containing the FB inverted repeat. On the gels from which the Southern blots were made the small amount of DNA in each ladder band was not visible, but the absence of the total of this DNA from the the primary FB inverted repeat-containing band was observable. This FB inverted repeat DNA band was less intense than expected when compared with the other DNA bands in the same lane. Evidently partial digestion of some of the FB inverted repeat sequences occurred, and because of the tandem repeat construction of these sequences, this generated ladders of restriction segments.

The specificity of this phenomenon indicates that different agents, with different recognition sites, contaminating the enzymes Sac I and Pvu II were responsible since one clone ("a" in figure 1) was sensitive to one of these agents and not to the other. The enzymes Sac I and Pvu II themselves do not seem to be responsible, since conditions of 100-fold over digestion failed to reduce the ladders to monomer size. We tested the obvious possible contaminants Sac II and Pvu I, but the results implied that these were not responsible for the ladders. Perhaps other as yet uncharacterized contaminating enzymes were involved.

FB inverted repeat sequences have been found to have a complex structure which some distance in from the outside ends is made up of tandem 31 bp direct repeats. Slightly different types of these direct repeats are organized in sets of five such that a 155 bp repeat length may be established (11). Taq I sites with this 155 bp spacing have been noted in FB inverted repeats (10), and in the FBw^{c}-homologous clones the ladder bands have approximately this spacing.

The specificity observed for the ladders also indicates that there are differences among FB inverted repeats, even between two FB inverted repeats on the same element (FBw^{c}). The left inverted repeat of FBw^{c} shows the ladder pattern in Sac I or Pvu II digests, whereas the right one does not. This implies that the inverted repeats are generated individually and do not often exchange information by recombining with each other. This contrasts with the previous study of an FB4 rearrangement which indicated that "flip-flop" recombination between inverted repeats of a single element could occur (15). However, FB4, with its composite construction, might represent a special case.

When the potential gene region was discovered on the loop of FB4, one proposal was that this might be an FB transposase gene. Another proposal was

that the FB4 loop segment might be an independent transposable element since the loop segment had its own short terminal inverted repeats, distinct from the FB inverted repeats. The genomic Southern blots were done to begin to study the distribution and conservation of this loop segment.

The pattern of the hybridization of the Pvu II segment from the FB4 loop (HB1) to the Bam HI-digested, total genomic D. melanogaster DNA, showing multiple bands (figure 3), indicated that homologous DNA was present in a number of different sites in the genome. Hybridization to the Pvu II-digested DNA indicated that the Pvu II segment containing the potential gene was not strictly conserved since, although the band corresponding to the size of the probe is darkest, there are several additional bands (figure 3). Some of these additional homologous Pvu II segments could have been due to simple base changes affecting the Pvu II recognition sites; the loss of a site would result in a larger segment extending to the next flanking site. Significant divergence affecting the potential gene could have produced other additional bands; in the case of the Pvu II segment from HB2, sequence analysis indicated that the region of the potential gene was actually altered considerably.

Three lines of evidence suggest that the sequences homologous to the FB4 loop segment represent independent transposable elements: 1) the sequences have their own short terminal inverted repeats, 2) the sequences generally are found independent of FB terminal inverted repeats, and 3) the sequences are moderately repetitive in copy number. We have designated this group of repetitive elements the HB family. In the hybridization of the Pvu II segment to the genomic Southern blot (figure 3) the patterns of homologous segments are similar across the four strains. This similarity suggests that the repetitive sequences predate the isolation of these strains and that the elements have transposed infrequently since then. Alternatively, these elements entered these strains separately and had specific target sites. This seems less likely; the DNA sequences surrounding the analyzed elements are not obviously related.

Although two FB elements apparently can mobilize other segments of DNA between them (9), there is no evidence that a single FB element can "pick up" DNA between its inverted repeats. The best explanation for the structure of the loop in FB4 is the insertion of HB1. The very low frequency of HB transposition suggested by the cross-strain similarities could be due to any of several possibilities. If the potential gene in the sequence of HB1 codes for an HB transposase, very little of this transposase may be made since

other HB elements contain deletions and multiple base changes affecting this region. Even the potential gene in HB1 may contain sequence changes such that it no longer encodes a functional transposase, or the gene may code not for a transposase but for a repressor. Base changes in the inverted repeats of the HB elements may have caused them to be no longer recognized by transposition factors. Also, the presence of other unfavorable conditions might account for the positional stability of the HB elements.

Two other families of Drosophila transposable elements with short terminal inverted repeats are the hobo (24) and P elements (23). The hobo element which inserted near a glue protein gene had 12 bp perfect inverted repeats and generated an 8 bp target site duplication (24). P elements have 31 bp perfect inverted repeats, the sequence of which is conserved from one element to the next, and these elements generate 8 bp direct repeats of target site DNA (23). In the sequences of the ends of four HB elements listed in figure 6, the inner end of the inverted repeats is well defined and, therefore, is marked by the vertical lines. The outer end is less apparent due to base changes in the inverted repeats and the lack of obvious direct repeats, or possibly due to variability in the lengths of the elements. The inverted repeats may be 29-32 bp in length, as judged by the homology among the 8 end sequences. If the inverted repeats are taken to be 30 bp long, then there could be 2 bp direct repeats just outside of them. Longer direct repeats are not found consistently on the four elements. These elements may have inserted into their sites long enough ago that the sequence changes since then have made any direct repeats once present no longer apparent.

The size heterogeneity of the HB elements is similar to that of hobo (24) and P elements (23). However, the numerous sequence changes and multiple deletions from one HB to another (for example see figure 4) indicate a situation different from that described for the P elements. Each P element differs from the full length element only by a single deletion and very few other changes (23). The different size classes of homologous Pvu II segments in the HB elements are generated by different sets of internal deletions. The homologous segments which are not bordered by Pvu II sites are probably due to both simple base changes in the Pvu II recognition sites and larger rearrangements. HB elements may cause instabilities and rearrangements which involve the adjacent sequences as well as the element. This is suggested by the structure of the clone containing HB2; a direct repeat of part of the HB sequence is found a short distance outside of the element.

If the sequence in the middle of FB4 is a transposable element independent of FB elements, then the gene may still code for a transposase, or other function, for these HB elements. There may be a transposase encoded elsewhere in the genome which acts on FB elements, or perchance the FBw^c loop sequence or another yet undiscovered FB loop sequence encodes such a function.

Many P transposable elements are apparently related to one another by being different deletion products of larger complete P elements (23). If FB elements were related in a similar manner, one would expect two nonhomologous FB loops to have come from a loop containing both sequences. With this model in mind, we screened the clones with homology to the FB4 loop for homology to the FBw^c loop. No clone with homology to both the FB4 loop segment and either of the loop segments from FBw^c was found. In light of the apparent nature of the FB4 loop sequence, this result is not surprising. If the FB4 loop sequence is an independent element, then it would not be expected to have a common origin with the FBw^c loop sequence.

In summary, the FBw^c loop appears to represent an integral and conserved component of several FB transposable elements and, therefore, might be functionally important for the FB family. The FB4 loop, however, represents a transposable element of the distinct HB family by chance found inserted into an FB transposable element.

ACKNOWLEDGMENTS:

We thank Mary Collins and Gerald M. Rubin for providing the pC10.2 and pC7.6 clones. We thank Dan Weinstein for assisting in the initial stages of this work.

This work was supported by grant number GM 30816 from the National Institutes of Health and by PHS grant number CA 00852 from the National Cancer Institute, DHHS.

*Present address: Department of Biological Sciences, Stanford University, Stanford, CA 94305, USA

LITERATURE CITED

1. Calos,M.P. and Miller,J.M. (1980) Cell 20, 579-595.
2. Shapiro,J.A. and Cordell,B. (1982) Biol. Cell 43, 31-54.
3. Spradling,A.C. and Rubin,G.M. (1981) Annu. Rev. Genet. 15, 219-264.
4. Potter,S., Truett,M., Phillips,M., and Maher,A. (1980) Cell 20, 639-647.
5. Collins,M. and Rubin,G.M. (1983) Nature (London) 303, 259-260.
6. Bingham,P.M. (1981) Cold Spring Harbor Symp. Quant. Biol. 45, 519-525.
7. Levis,R., Collins,M., and Rubin,G.M. (1982) Cell 30, 551-565.
8. CollinsmM. and Rubin,G.M. (1984) Nature (London) 308, 323-327.
9. Paro,R., Goldberg,M.L., and Gehring,W.J. (1983) EMBO J. 2, 853-860.
10. Truett,M.A., Jones,R.A., and Potter,S.S. (1981) Cell 24, 753-763.

11. Potter,S.S. (1982) Nature (London) 297, 201-204.
12. Potter,S.S., Brorein,W.J.,Jr., Dunsmuir,P., and Rubin,G.M. (1979) Cell 17, 415-427.
13. Grunstein,M. and Hogness,D.S. (1975) Proc. Nat. Acad. Sci. USA 72, 3961-3965.
14. Holmes,D.S. and Quigley,M. (1981) Anal. Biochem. 114, 193-197.
15. Potter,S. (1982) Mol. Gen. Genet. 188, 107-110.
16. Southern,E.M. (1975) J. Mol. Biol. 98, 503-517.
17. Smith,G.E. and Summers,M.D. (1980) Anal. Biochem. 109, 123-129.
18. Rigby,P.W.J., Dieckmann,M., Rhodes,C., and Berg,P. (1977) J. Mol. Biol. 113, 237-251.
19. Denhardt,D.T. (1966) Biochem. Biophys. Res. Commun. 23, 641-646.
20. Klenow,H. and Henningsen,I. (1970) Proc. Nat. Acad. Sci. USA 65, 168-175.
21. Levis,R., Dunsmuir,P., and Rubin,G.M. (1980) Cell 21, 581-588.
22. Maxam,A.M. and Gilbert,W. (1980) in Methods Enzymol., Grossman,L. and Moldave,K. Eds., Vol. 65, pp. 499-560, Academic Press, New York.
23. O'Hare,K. and Rubin,G.M. (1983) Cell 34, 25-35.
24. McGinnis,W., Shermoen,A.W., and Beckendorf,S.K. (1983) Cell 34, 75-84.