The Molecular Structure of TE146 and Its Derivatives in Drosophila melanogaster

R. Lovering,1 N. Harden2 and M. Ashburner

Department of Genetics, University of Cambridge, Cambridge, England

Manuscript received December 7, 1990
Accepted for publication February 22, 1991

ABSTRACT

TE146 is a giant transposon of Drosophila melanogaster. It carries two copies of the white and roughest genes, normally found on the X chromosome. The structure of this transposon has been studied at the molecular level. TE146 may transpose to new chromosome positions, excise and be lost from the genome or undergo internal rearrangements. The termini of TE146 are foldback DNA elements (FB); the transposon also carries two internal FB elements. Loss or internal rearrangement of TE146 involves recombination between different FB elements. These events have been mapped molecularly, by taking advantage of the fact that the FB sequences are composed largely of a regular 155-bp repeat sequence that is cut by the restriction enzyme TaqI, and are shown to be nonrandom. We suggest that these FB-FB exchange events occur by mitotic sister-chromatid exchange in the premeiotic germ line.

A family of very large transposons was discovered in Drosophila melanogaster by Ising and Ramel over 25 years ago (RAMEL 1966; ISING and RAMEL 1976). This family originated by the transposition of two genes, white and roughest from a Base (Muller-3) X chromosome to chromosome arm 2R, and was first detected because the transposon acted as a dominant suppressor of white (ISING 1964). The unusual property of this suppressor was first evident when its genetic position jumped from a wild-type sequence chromosome arm 2R to a Cy balancer chromosome. Since then, this TE has been hopping around the genome of D. melanogaster and over 200 different transpositions have now been mapped (ISING and BLOCK 1984; G. ISING, personal communication; D. GUBB and J. ROOTE, unpublished data). These TEs show two other properties, each indicative of their unstable nature. First, they may excise from a site and be lost from the genome. Because the TE carries a functional allele of white (either w or a reversion of this allele to an almost wild-type form), losses are very easily detected as long as the TE is kept on a genotypically w- background. Second, the TE can undergo internal duplications or deletions. This can be seen by the changing reaction of the white genes carried by the TE to the zeste1 mutation, since the yellow eye-color of z1 requires two contiguous copies of w+ for its expression (Pirrotta 1991, for a recent review). Originally, the TE carried a single copy of white and gave a red eye-color on a z1 w- background. Sometimes, its white genes become suppressible by z1. These new forms of the TE have duplicated their white gene (and, usually, their roughest gene, too). Such duplicated forms of the TE can subsequently lose one copy of white and revert to a form that is red-eyed (with z1). In fact, this process of duplication, deletion, subsequent duplication and then deletion can be followed ad infinitum, with the TE remaining at the same insertion site (GUBB et al. 1985, 1986). The first clue to the molecular nature of the sequences that cause the TE to be transposable and unstable came with the discovery that the termini of the element are foldback sequences (GOLDBERG, PARO and GEHRING 1982). POTTER (1982a) had characterized a family of foldback DNA that was repetitive in the genome of all strains of D. melanogaster, usually being found at about 20–30 different chromosomal locations. Sequences of this family (FB) were found at the junction of the TE with sequences at its insertion site. They were also found to be associated with transposons of independent origin, e.g., w and wN (GOLDBERG, PARO and GEHRING 1982; LEVIS, COLLINS and RUBIN 1982, PARO, GOLDBERG and GEHRING 1983) and with the unstable mutation wDZ (LEVIS, COLLINS and RUBIN 1982; LEVIS and RUBIN 1982; ZACHAR and BINGHAM 1982). FB elements are often associated with other DNA sequences (TRUETT, JONES and POTTER 1981). One class of FB elements is associated with a particular sequence with coding potential, known as the NOF sequence. The composite FB-NOF element was first found at an end of TE28 by GOLDBERG, PARO and GEHRING (1982) and then at the unstable w' mutation (LEVIS, COLLINS...
and Rubin 1982). FB-NOF sequences are normally rare in the genome (0–2 copies) although many strains carrying Ising's transposon, as well as Basc and wozl strains, may have 10–20 copies (Harden and Ashburner 1990; Lovering 1988). Two FB-NOF elements have been sequenced (Templeton and Potter 1989; Harden and Ashburner 1990) and the 4-kb NOF sequence has the potential to code for a 120-kD polypeptide. Although circumstantial, the evidence is that the NOF sequence encodes a function required for FB mediated transposition.

The structure of FB elements suggests that recombination between different FBs may be the mechanism for their excision and transposition. In two cases there is evidence that this is indeed so. The unstable FB-NOF associated w allele frequently mutates to stable, male-viable, w' derivatives. Many of these are 14-kb deletions that result from exchange between the FB-NOF element within w' and a FB element 14-kb more distal (Collins and Rubin 1984). Several "complete" losses of one TE, TE146(Z), inserted at 35B1.2 within the no-oellii (noc) gene (Gubb et al. 1985a), have been analyzed molecularly by Chia et al. (1985a). These losses of the TE revert the strong mutant noc allele associated with the insertion to noc'. Their important feature is that they are all imprecise excisions of the TE and leave between 3 and 10 kb of DNA, which is largely, if not solely, FB DNA, at the insertion site. They must have resulted from an exchange between the FB elements that flank TE146(Z).

Despite the evidence for FB/FB mediated exchange being important for the instability of the TE, the precise structures of the FB elements, and more especially of their junctions with non-FB sequences, are not known. In this paper we describe these junctions and provide further evidence for the nature of FB mediated recombination.

MATERIALS AND METHODS

Stocks: The original stock of TE146(Z)/CyO was from G. Ising. The SR series of spontaneous derivatives of TE146(Z) was recovered as red-eyed w' w185; TE146(Z)/CyO flies from zeste-eyed parents. They have been characterized genetically by Gubb et al. (1986).

Probes: noc sequences: TE146(Z) is inserted at coordinate 1075 kb of the Adh region chromosome walk, within DNA covered by the phage λob9.04 (Chia et al. 1985b). Two clones, described by Chia et al. (1985a), were used to study the junctions between noc and TE146(Z) sequences. One (nocD) is a 0.7-kb HindIII-SalI fragment from coordinates 1075 to 1082. The TE is inserted 50 bp 5' of the SalI site. This probe is, in practice, specific for sequences immediately proximal to the insertion site. For sequences immediately proximal to this site the clone nocP was used. This is the 0.65-kb SalI-EcoRI fragment that, in wild-type DNA, is immediately proximal to nocD.

white sequences: The distal end of the white sequences carried by TE146(Z) is included in the phage λm11B.2 of Lewis, Bingham and Rubin (1982). This phage wholly includes the 5-kb HindIII-BamHI fragment D of pΔ of Goldberg, Paro and Gehring (1982), identified as being close to the distal end of white sequences carried by TE28 (Figure 1a). Two fragments of λm11B.2 (provided by K. O'Hare) were subcloned in pBR322, the 2.6-kb SalI-EcoRI fragment that corresponds to coordinates 15.5 to –18.2 kb of the white walk (Lewis, Bingham and Rubin 1982) (we will call this w:SR26) and the 450-bp HindIII-EcoRI fragment that includes sequences from –15.5 to 16 kb of the w:HR0.45 clone. In addition, as a probe to the white gene sequence itself, we used the 11-kb KpnI-EcoRI fragment carrying DNA from coordinates 4 to +7 (see Figure 5b). This clone was made by K. Moses from λm2.1 of Lewis, Bingham and Rubin (1982) (provided by K. O'Hare).

rst sequences: The rst allele has not been characterized at the molecular level. However, Paro, Goldberg and Gehring (1983) isolated several clones from TEs which they mapped to sequences proximal to rst. One of these, T2, a 0.6-kb subclone in pBR329 from the phage Xob9.04 derived from a TE98 library, was the gift of M. Goldberg (see Figure 1b). A genomically unique 1-kb SalI-EcoRI fragment of T2 (in pBR325) was used as a probe for sequences normally proximal to the rst gene (rst:SR0.8). This fragment is the same as Paro, Goldberg and Gehring's T2B, TE98 carries a NEB element between its rst gene and the adjacent FB DNA, within the EcoRI-SalI fragment. rst:SR0.8 does not include these NEB sequences.

NOF sequences: The NOF probe has been described by Harden and Ashburner (1990), it is the EcoRI-SalI fragment of pA1 of Goldberg, Paro and Gehring (1982) and was the gift of M. Goldberg.

Molecular techniques: Routine techniques for the isolation and analysis of DNA are described in Maniatis, Fritsch and Sambrook (1982). The extraction of DNA from single flies used the method described by Jowett (1986). For the partial digestion of DNA with TaqI the enzyme concentration was 0.3 unit/μg DNA at 65° for 15–30 min. Denaturing conditions were used for the Southern transfers, as described by Reed and Mann (1985).

A show hybridization to polytene chromosomes was done with either 32P-labeled probes as described in Ashburner (1989).

RESULTS

A summary of the gross organization of TE146(Z) is shown in Figure 2. The basic features of this structure, that this TE includes a tandem repeat of w' and rst' genes bounded by elements of Potter's FB family (Truett, Jones and Potter 1981), have been determined by genetic (Gubb et al. 1985, 1986) and previous molecular (Chia et al. 1985a) studies. We identify eight different DNA boundaries of sequences that are contiguous in TE146(Z) but not in wild-type DNA. To help the presentation of the data these are labeled A to H, and the adjacent FB sequences will be called limbs A to H. An FB element consists of two limbs in opposite orientation with a variable amount of non-FB DNA separating them. The eight boundaries have been mapped, by probing DNA extracted from flies of an appropriate genotype after digestion with restriction enzymes, electrophoresis and transfer to filters.

The insertion site of TE146 in the noc gene: Chia
Transposon structure in Drosophila

FIGURE 1.—(a) Restriction map of the DNA distal to the white gene of D. melanogaster (from LEVIS, BINGHAM and RUBIN 1982) showing the positions of the various clones used in this work and the distal limit of the \(w^{11E4}\) deletion (mapped by ZACHAR and BINGHAM 1982). Fragment D is that isolated from \(TE28\) by GOLDBERG, PARO and GEHRINC (1982).

(b) Restriction maps of the region proximal to the roughest gene in \(z^{1} w^{11E4}\) and \(TE98\) showing the position of the \(rst\):SR0.8 probe and its relationship to the \(T2\) and \(\lambda 98/2\) clones.

et al. (1985a,b) showed that \(TE146(Z)\) is inserted in sequences of the \(noc\) gene, as expected since this insertion results in a strong mutant \(noc\) phenotype (GUBB et al. 1985). The insertion site is within the 0.7-kb HindIII-SalI fragment at coordinate -108.2 to -107.5 of the \(Adh\) region chromosome walk. Junctions A + B are represented by a 11.4-kb \(PstI\) fragment seen when DNA from \(TE146(Z)\) is probed with \(noc:D\). Junctions G + H are represented by a 9.0-kb \(PstI\) fragment seen when this DNA is probed with \(noc:P\) (CHIA et al. 1985a).

The orientation of the white genes within \(TE146\):

Genetically, \(TE146(Z)\) is known to carry two functional white genes that are in tandem orientation with respect to each other (GUBB et al. 1986). These genetic data also strongly suggested that the most distal gene carried by the TE was white (rather than \(rst\)), that is to say the map of the TE could be represented as \(noc^{n} white \; rst\); white with \(noc^{n}\) and \(noc^{p}\) representing the distal and proximal regions of the \(noc\) gene relative to the TE’s insertion site. To verify this structure at the molecular level we used spontaneous derivatives of \(TE146(Z)\) that carry only a single white gene. As suggested by GUBB et al. (1986), and as will be proven below, some of these \(SR\) derivatives have lost the distal copy of white and some the proximal copy. When DNA from \(TE146(Z)\) is digested with \(PstI\) and probed with \(w:HR0.45\) two fragments from \(TE\) DNA hybridize, of 11.4 and 7.0 kb (Figure 3, tracks 1–3). DNA from the spontaneous red-eyed derivative \(SR5\) lacks the 11.4-kb fragment while DNA from another derivative, \(SR35\), lacks the 7.0-kb fragment. These data suggest that \(SR5\) has lost one of the copies of white and \(SR35\) the other.

The 11.4-kb \(PstI\) fragment comes from junctions A + B, since it is also hybridized by the \(noc:D\) probe, but not by the \(noc:P\) probe. The 7.0-kb \(Pst\) fragment is from junction F since it fails to hybridize to either of the two \(noc\) probes (Figure 3). \(SR5\) retains junction F (the 7.0-kb fragment) but not junctions A + B, the 11.4-kb fragment being replaced by a novel fragment of 17-kb (Figure 3, track 4). We conclude that \(SR5\)
FIGURE 2.—A summary of the molecular map of the noc region (coordinates are from the EcoRI site 5' to Adh) showing the insertion site of TE146(Z) (CHIA et al. 1985b). TE146(Z) is shown inserted at coordinate −107.5. The positions of the two noc probes used to map sites from this insertion site are indicated. Some restriction enzyme sites within the body of the TE are indicated. Some of these had also been mapped by CHIA et al. (1985a) but data for them all are given in this paper. The letters A to H indicate the eight junctions between FE (as wriggly lines) and non-FB (as straight lines) DNA. The lengths of FB sequences are not to scale, but their total lengths are indicated (in kb). The positions of probes to white, roughest and NOF sequences within the TE are shown below the diagram. Abbreviations for restriction enzyme sites used are: H, HindIII; P, PstI; R, EcoRI and S, SalI.

FIGURE 3.—Genomic DNA from z′ w1164; TE146(Z)/CyO and from two spontaneous red-eyed derivatives of this stock, SR5 and SR35, was digested with PstI and then hybridized in turn with w:HR0.45 (tracks 1–3), noc:D (tracks 4–6), noc:P (tracks 7–9) and rst:SR0.8 (tracks 10–12). With the white gene probe the 7.5- and 3.8-kb fragments are from the z′ w1164 chromosome; with the noc gene probes the 4.6-kb fragment is from the CyO chromosome; with the rst probe the 5.5- and 2.0-kb fragments are from the X chromosome. (We would expect the intensity of hybridization of the w:HR0.45 probe to be similar to the 7.0- and 11.4-kb fragments in track 2 (as in Figure 7a). Presumably, the higher molecular weight fragment was poorly transferred to the filter here).

has lost the distal copy of white. SR35 retains junctions A + B (the 11.4-kb fragment) but has lost junction F (the 7.0-kb fragment that is hybridized by w:HR0.45). We conclude that this derivative has lost the proximal white gene. These conclusions were confirmed by probing the same filter with noc:P (Figure 3, tracks 7–9). This hybridizes to a 9.0-kb PstI fragment in TE146(Z) and SR5 (junctions G + H) but to a novel, 7.0-kb, fragment in SR35.

The interpretation of these data is that a white gene is close to junctions A + B but distant from junctions G + H, i.e., that the structure summarized in Figure 2, deduced from genetic data, is correct.

The position of rst sequences: Although the rst gene has not been characterized molecularly, a probe to DNA sequences proximal to rst was available from the experiments of PARO, GOLDBERG and GEHRING (1983). This probe, rst:SR0.8, was used with the same filters of PstI digested DNA. With DNA from TE146(Z) three PstI fragments hybridize, of 10.8, 9.0 and 2.5 kb. Both SR5 and SR35 lack one of these (figure 3, tracks 10 and 12). The 10.8-kb fragment is absent from SR5 and is not replaced by any novel fragment; both the 10.8-kb and 9.0-kb fragments are absent from SR35 and are replaced by a novel 7.0-kb fragment. Which of these fragments represents the fusion with noc sequences? Only the proximal noc probe, noc:P hybridizes to these bands, in fact to the 9.0-kb band of TE146(Z) and SR5 and to the 7.0-kb fusion band of SR35. Therefore, the 9.0-kb PstI fragment represents junction G + H and the 10.8-kb fragment represents junction C. These data confirm the model of TE146(Z)’s structure shown in Figure 2.

The position of NFB sequences in TE146(Z): By in situ hybridization to polytene chromosomes TE146(Z)
carries two sites of NEB sequence (McGill 1985). NEB is a transposable element of the retroviral class that was mapped between the rst and FB sequences of TE98 by Paro, Goldberg and Gehring (1983). A comparison of the restriction maps of TE98 and of junctions G + H of TE146(Z) shows them to be very similar (see below). This suggests that a NEB element is located between rst and FB sequences at junction G.

When PstI digested DNA of TE146(Z) was probed with rst:SR0.8 there is a 2.5-kb fragment that hybridizes (Figure 3, track 11). This fragment results from the insertion of NEB into a 2.0-kb PstI region of the wild-type DNA. Its size is unchanged in both SR5 and SR35 (Figure 3, tracks 10 and 12), but note that in these its intensity relative to the 2.0-kb fragment is reduced, as expected since these SRs should have one and not two rst genes. These data suggest that both of the rst genes of TE146(Z) have a NEB element inserted at a similar site. Interestingly, the Basc chromosome, from which this family of TEs was derived, does not have a NEB sequence at the homologous position proximal to rst, a conclusion drawn from the PstI fragment sizes of this chromosome and confirmed by digestion with EcoRI and XhoI (LOVERING 1988).

Are the two white junctions in TE146(Z) the same? Ising's family of TEs originated carrying one copy of the white and roughest genes. TE146(Z), and many other members of this family have duplicated these, by an unknown mechanism (see DISCUSSION). In TE146(Z), we have already shown that the distal limit of the proximal copy of the transposed white sequences (i.e., junction F) is about 16.4-kb distal to the site of copia insertion in w (Harden and Ashburner 1990; see also Goldberg, Paro and Gehring 1982). Is junction B the same? To study this, DNA was digested with DraI and probed with w:SR2.6, a wild-type sequence that crosses the sequenced FB/white junction. With w:11E4 (a deletion of w that does not extend this far distal, see Figure 1a) and TE146(Z)w2, derivatives of TE146(Z) that have lost their w genes (Chia et al. 1985a), four hybridizing DraI fragments are seen (Figure 4a, tracks 2 and 3; their origin is indicated on the DraI map, Figure 4b). With TE146(Z) itself there is an extra 0.46-kb band (track 1); this must represent the FB/w junction. Since only one band is seen, this suggests that this junction is the same for both w genes. Note that the intensity of hybridization of the probe to this band is about equal to that of the 0.56-kb fragment from the w:11E4 chromosome.

Five different SR derivatives, each containing only one w gene, have been analyzed in a similar manner (Figure 4a, tracks 4–8). All show the same fragments as TE146(Z), but the 0.46-kb band hybridizes less intensively than the 0.56-kb band (Figure 4a, track 1). It so happens that all of the SRs used in this experi-

![Figure 4](image-url)
fragment that spans the copia insertion site in \( w^c \). These data suggest that the \( w \) genes of \( TE146(Z) \) still carry a small (ca. 250 bp) insertion in \( w \), presumably a copia LTR (see Carbonare and Gehring 1985).

The structure of \( TE146(Z) \): The data discussed so far are summarized in Figure 6. We have determined the gross structure of junctions A to H by probing suitable restriction digests with a variety of clones whose positions in their corresponding wild-type genes are known. These data have been confirmed by digests with two other enzymes (HindIII and SalI) (Lovering 1988). The region of greatest uncertainty is that of limbs D and E; these will be discussed below.

The structure of half-losses of \( TE146(Z) \): One of the first indications that \( TE246(Z) \) carried two copies of \( w^+ \) was the phenotype of \( z^1 w^-; TE246(Z)/+ \) flies. These are zeste-colored, and not red (Gubb et al. 1985). However, a stock of \( z^1 w^{1144}; TE146(Z)/CyO \) frequently gives red-eyed derivatives and many of these have been selected (Gubb et al. 1986). The majority, though not all, of these SRs have a single copy of \( white \) by a number of criteria: unlike the original form of \( TE146 \) they have three, and not six, polytene chromosome bands; they show only one, and not two, \( in \text{ situ} \) hybridization sites with a \( white \) gene probe and they produce only half as much red-pigment as does \( TE146(Z) \) (Gubb et al. 1986).

Two of these SRs have been used in the analysis of the junctions between \( TE146 \) and \( noc \) sequences, \( SR5 \) and \( SR35 \). In \( SR5 \) junctions A and B are changed in size but junctions G and H remain as in the original form of this \( TE \). In \( SR35 \) junctions A and B are unchanged but junctions G and H are novel. We concluded (above) that \( SR5 \) had lost the distal copy of \( white \) and retained the proximal and that \( SR35 \) had lost the proximal copy and retained the distal. A further 15 SRs have been studied in the same way: six are similar to \( SR5 \) and eight to \( SR35 \), although the precise sizes of the novel junction fragments may differ between SRs (Figure 7 and Table 1). One, \( SR47 \), differs from these in that both the distal and proximal junctions (i.e., both \( A + B \) and \( G + H \)) are novel. This may have been a consequence of the primary event which generated this derivative or of some event occurring subsequently. Two of the SRs, \( SR33 \) and \( SR64 \), have polymorphic fragments representing junctions \( A + B \) and \( E + F \), respectively (Figures 7a and b). These are true polymorphisms since DNA from single flies of either of these strains show only one of the two forms. These polymorphisms are evidence for secondary changes occurring to junctions.

**NOF in \( TE146(Z) \) half-losses:** An interesting correlation is seen when the half-loss derivatives of \( TE146(Z) \) are probed for NOF sequences. All of the distal half-losses retain a 7.0-kb \( PstI \) fragment that hybridizes with NOF (Figure 7e) (except the unusual \( SR47 \)). None of the proximal losses retain this fragment. The first conclusion drawn from these data is that loss of the distal half of the element occurs by an event involving sequences distal to NOF, i.e., distal to
Figure 6.—Restriction enzyme map of the junctions of FB and non-FB DNA in TE146(Z). FB DNA is shown, to scale, as wriggly lines, except for limb E which is estimated to be >1 kb. The nature of sequences between limbs D and E is unknown. The positions of the probes used to deduce these maps are indicated by the solid bars.

Figure 7.—Genomic DNA from TE146(Z) and several different spontaneous half-losses (SRs) was digested with PstI and the filter hybridized in turn with (a) w:HR0.45, (b) noc:D, (c) noc:P, (d) rst:SR0.8 and (e) NOF probes. Fragments common to all tracks, and those whose sizes are not indicated alongside the TE146(Z) track, are from the 1A1174 or Cy0 chromosomes.

junction E. The interpretation of the proximal half-losses is not as straightforward.

As can be seen from Figure 7e and Table 1 two of the proximal half-losses (SR26 and SR35) show only the small 3.7-kb PstI fragment hybridizing with NOF. This fragment is present in TE146(Z) and all of its derivatives. It is from a site at 66AB (Figure 8a and HARRISON 1989). The conclusion is that SR26 and SR35 lack NOF sequences within the TE. The events which generated these derivatives presumably involved sequences distal to NOF, i.e., limbs C, D or E.

All of the other proximal half-losses possess novel PstI fragments hybridizing with NOF and, in many cases, more than one such fragment. These originate in two quite different ways. In SR24 and SR44 the 10.5-kb PstI fragment also hybridizes to the noc:P probe. It is presumably a consequence of a novel junction between the FB-NOF sequences and noc DNA, i.e., a novel junction “G + H.” The other NOF fragments in these SRs do not hybridize to noc probes:
in fact they are not from the TE at all but represent amplified NOF sequences that appear to have been generated coincidentally with the SR. For three derivatives this has been confirmed by in situ hybridization to polytene chromosomes with a NOF probe: SR2 and SR51 do not show a NOF in situ site in region 35 but have a site at 94A (Figure 8b). The third derivative is SR48, a deletion of the proximal-half of TE146(Z) extending across the adjacent osp and Adh loci (GUBB et al. 1986). This chromosome has no NOF sequence in region 35 but does have one in region 37. None of these sites have been seen to be occupied by NOF in TE146(Z) nor have NOF hybridizing restriction fragments of an appropriate size ever been seen in this stock. These SRs, therefore, have been generated by an event involving sequences distal to FB-NOF. This interpretation is confirmed by probing these half-losses with rst:SR0.8. TE146(Z) shows two PstI fragments, of 10.8 and 9.0 kb, that hybridize with this probe. The 9.0-kb fragment extends from the PstI site near the proximal copy of rst to that in noc DNA proximal to the TE's insertion site. This fragment is retained in all of the distal half-losses and is lost from all of the proximal half-losses. The 10.8-kb PstI fragment extends from the PstI site near the distal copy of rst to a PstI site that maps between limbs D and E. This fragment is only retained by SR24 and SR44, the two losses of the proximal half of the TE that retain NOF DNA. In SR38 and SR51 this 10.8-kb fragment is replaced by one of 11.0 kb, in SR2, SR26 and SR35 it is replaced by a 7.0-kb PstI fragment. These novel fragments indicate the sizes of the DNA remaining between rst and proximal-noc sequences, a fact confirmed by their hybridization with the noc:P probe (Table 1).

The complete loss of TE146(Z), i.e., to a white-eyed derivative, involves exchange between the flanking FB elements, that is between junctions A + B and G + H. These losses leave a few kilobases of FB DNA within the noc gene (CHIA et al. 1985a). The data from the half-losses indicate that these originate by a similar class of event, that is exchange between flanking FB sequences (Figure 9).

**TABLE 1**

Sizes of PstI fragments of TE146 and its derivatives detected by the five different probes used in these experiments

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>(w_{HR0.45})</th>
<th>noc:D</th>
<th>noc:P</th>
<th>NOF</th>
<th>rst:SR0.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>3.8, 7.5</td>
<td>4.6</td>
<td>4.6</td>
<td>3.7</td>
<td>2.0, 5.5</td>
</tr>
<tr>
<td>TE146(Z)</td>
<td>7.0, 11.4</td>
<td>11.4</td>
<td>9.0</td>
<td>7.0</td>
<td>2.5, 9.0, 10.8</td>
</tr>
<tr>
<td>Distal half-losses:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR3</td>
<td>7.0</td>
<td>17</td>
<td>9.0</td>
<td>7.0</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR33</td>
<td>7.0</td>
<td>11.35</td>
<td>9.0</td>
<td>7.0, 8.0</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR46</td>
<td>7.0</td>
<td>12</td>
<td>9.0</td>
<td>7.0</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR47</td>
<td>8.4</td>
<td>8.4</td>
<td>4.6</td>
<td>8.0</td>
<td>2.5, 4.6</td>
</tr>
<tr>
<td>SR63</td>
<td>7.0</td>
<td>13</td>
<td>9.0</td>
<td>7.0, 11.5</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR64</td>
<td>7.0, 8.8</td>
<td>13</td>
<td>9.0</td>
<td>7.0, 8.8</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR68</td>
<td>7.0</td>
<td>12</td>
<td>9.0</td>
<td>7.0, 8.0</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>SR100</td>
<td>7.0</td>
<td>14</td>
<td>9.0</td>
<td>7.0, &gt;20</td>
<td>2.5, 9.0</td>
</tr>
<tr>
<td>Proximal half-losses:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR2</td>
<td>11.4</td>
<td>11.4</td>
<td>7.0</td>
<td>11.5</td>
<td>2.5, 7.0</td>
</tr>
<tr>
<td>SR24</td>
<td>11.4</td>
<td>11.4</td>
<td>10.5, 6.5</td>
<td>10.5, 6.5</td>
<td>2.4, 10.8</td>
</tr>
<tr>
<td>SR26</td>
<td>11.4</td>
<td>11.4</td>
<td>7.0</td>
<td>—</td>
<td>2.5, 7.0</td>
</tr>
<tr>
<td>SR35</td>
<td>11.4</td>
<td>11.4</td>
<td>7.0</td>
<td>—</td>
<td>2.5, 7.0</td>
</tr>
<tr>
<td>SR38</td>
<td>11.4</td>
<td>11.4</td>
<td>11.0</td>
<td>8.0</td>
<td>2.5, 11.0</td>
</tr>
<tr>
<td>SR44</td>
<td>11.4</td>
<td>11.4</td>
<td>10.5</td>
<td>10.5, 8.0</td>
<td>2.5, 10.8</td>
</tr>
<tr>
<td>SR51</td>
<td>11.4</td>
<td>11.4</td>
<td>11.0</td>
<td>11.5, 8.0</td>
<td>2.5, 11.0</td>
</tr>
</tbody>
</table>

No hybridization is indicated by —. Fragments from the X or CyO chromosomes are only listed in the "+" row. Numbers in italics represent novel fragments.
Transposon structure in Drosophila

**Figure 8.** (a) *In situ* hybridization to the polytene chromosomes of *z' w1118; TE146(Z)/CyO* with a tritiated NOF probe, showing two sites of hybridization, at 35B within the TE and at 66AB on chromosome arm 3L. (b) *In situ* hybridization to the polytene chromosomes of *z' w1118; TE146(Z) SR2/CyO* with a tritiated NOF probe showing hybridization to 94A on chromosome arm 3R but not to the site of the TE (large arrow).

*FB* sequences are only partially digested with this enzyme then a ladder of fragments, with a 155-bp periodicity, is seen. The method used here was simply to partially digest DNA from *TE146(Z)* and its derivatives with *TaqI* and then hybridize the DNA with a unique probe from a flanking region, *i.e.*, *noc:D* or *noc:P*. Then, the size of a labeled fragment will represent its distance from the probe, *i.e.*, from the junction of *noc* and *FB* sequences. DNA from both distal and proximal half-losses were used to study the internal junctions. If the *TaqI* sites are interrupted within the repetitive region of *FB* DNA then the position and size of sequences that lack *TaqI* sites will be seen from the break in the *TaqI* ladder (see LEVIS, COLLINS and RUBIN 1982).

Figure 10 is a more detailed view of the external side of an *FB* element [after TRUETT, JONES and POTTER (1981) and POTTER (1982a)]. Two features are important for our analysis: the first is the cluster of three *TaqI* sites between bases 134 and 154, the second is the *HinfI* site at base 53.

**Junctions A + B:** When DNA from *TE146(Z)* is completely digested with *EcoRI* and *HindIII* and then probed with *noc:D* there is an 8.5-kb hybridizing fragment. Partial *TaqI* digestion of this fragment gives an additional 7.0-kb band and a ladder of *TaqI* fragments. This ladder is interrupted between 3.4 and 3.85 kb (Figure 11, track 2). With a *HinfI* partial digestion of the *HindIII* fragment there is a 7.05-kb band. These fragments from partial digestion, the 7.0-kb *TaqI* band and the 7.05-kb *HinfI* band, originate by cutting the *HindIII* fragment at the end representing junction B (see Figure 11, track 3). (The smaller fragments from junction A have run off this gel.) More extensive *TaqI* digestion simply leads to a *TaqI* ladder down to a size of about 1.2 kb; this ladder is interrupted between 3.4 and 3.85 kb due to an absence of *TaqI* sites this distance from the probe (Figure 11, track 1). Below 1.2 kb the ladder becomes irregular and the stoichiometry between the bands changes. This results from the irregular distribution of *TaqI* sites at the end of an *FB* element and, in particular, the cluster of three *TaqI* sites about 0.95 kb from the probe (Figure 10). Our interpretation of

[Diagram showing restriction maps and hybridization patterns]

**Figure 9.** A summary of the mapping of *SR* derivatives of *TE146(Z).* The top line is a partial *PstI* (P) restriction map of *TE146(Z)* (not to scale) showing the sizes of the fragments seen with the various probes (which are shown as bars). The structure of seven distal half-losses and of seven proximal half-losses are drawn below, with the DNA deleted indicated as lying between the brackets. *FB* DNA is indicated by wriggly lines.

365
these data, that is to say a restriction map of junctions A + B, is given in Figure 12a.

**Junctions G + H:** The filters used in the analysis of junctions A + B were washed free of label and then rehybridized with *noc*/*P* so as to map the proximal junctions of *TE146(Z)*. The data are shown in Figure 11, tracks 4, 5 and 6. The initial products of partial *TagI* or *Hinfl* digestion are 7.7- to 7.75-kb bands. Further *TagI* digestion yields a *TagI* ladder down to about 1.1 kb and then a few smaller bands, including an intense band at 0.85 kb (due to the *TagI* cluster at junction H). The *TagI* ladder shows an asymmetric interruption, between 6.7 and 7.3 kb. These data are interpreted in Figure 12b.

**Junctions C, D and E:** Five of the *SRs* (*SR2, SR26, SR35, SR38* and *SR51*) were interpreted as being the consequence of exchange between limbs C or D and G or H. If so, the pattern of *TagI* partial digestion products, seen after hybridization with the *noc*/*P* probe, should be identical, at least at the proximal end of limb H. If the events generating these half-losses had been exchanges between limbs C and G, then limb H should be identical to that seen in *TE146(Z)*. On the other hand, had the exchange events been between limbs D and H then the lengths of the *TagI* ladder at junction H may differ from that seen in the parental element. Although the small *TagI* fragments at junction H of *SR2, SR26* and *SR35* are the same size as those in *TE146(Z)* (i.e., 0.75, 0.85 and 1.1 kb) the 155-bp ladder is not. It extends only 4.0 kb, not 6.7 kb as in *TE146(Z)*. After a 0.45-kb region without *TagI* sites the ladder restarts, representing the 155-bp repeat of limb C (Figures 13 and 14).

Junction D was analysed with *EcoRI-Hinfl* digested DNA from *SR64*, a loss of the distal copies of *white* and *roughest* (Figure 13, tracks 3 and 4). The pattern of small *TagI* fragments (0.85-1.2 kb), at the distal end of limb A is identical in *TE146(Z)* and *SR64*. The 155-bp ladder, presumably the consequence of an exchange between limbs C and D, is uniform to 3.5 kb from the *noc*/*D* probes. From 3.5 to 3.9 kb there are three irregularly spaced *TagI* sites, followed by 0.4-kb lacking *TagI* sites. Thereafter the pattern of sites is regular for 2.15 kb; these are the sites within limb D. The higher molecular weight *TagI* sites are not regular and must lie in non-FB sequences between limbs D and E. The complexity of this region is also evident from its pattern of *Hinfl* sites (Figure 13, track 4) and has been confirmed by digestion with other enzymes (see Figure 14a). The resolution of the high molecular weight *TagI* fragments is insufficient to allow any detailed analysis of the FB sequences of limb E. Figure 14a is an interpretation of these data.

**Junction F:** Junction F has been cloned and sequenced (HARDEN and ASHBURNER 1990). The sequence of junction F indicates the presence of nine *TagI* sites within the NOF sequence that lies between...
the HindIII site and FB DNA (Figure 14c). These generate a number of TaqI partial restriction fragments between 6.9 and 9.4 kb in size which have only been poorly resolved by electrophoresis (see Figure 15). Nevertheless the map of this junction deduced from TaqI mapping is wholly consistent with its DNA sequence.

In summary, TaqI restriction enzyme mapping, using probes adjacent to the insertion site of TE146, have allowed a detailed analysis of all of the FB/non-FB DNA junctions within this element. This allows a more detailed analysis of the events that led to loss of parts of this element when the SR derivatives originated.

**Recombination generating the SR derivatives of TE146(Z):** The frequent occurrence of spontaneous red-eyed derivatives of TE146(Z) by the loss of one copy of white suggests a rôle of FB elements in their origin (GUBB et al. 1986). If so, then the SRs should have novel junctions between FB and non-FB sequences. At a gross level we have already shown this to be true (see above). Moreover, the precise organization of FB repeats should also be novel in these derivatives, since the FB limb itself will have been generated by a recombination event between two FB limbs. These features of the SRs have been studied by TaqI partial restriction enzyme mapping.

**Proximal half-losses:** The proximal half-losses of TE146(Z) are of two types, as seen from the size of their fusion fragments with noc DNA and by the fact that only some retain NOF sequences (see above). We have concluded (above) that SR2, a proximal half-loss that lacks NOF DNA, resulted from an exchange between limbs D and H (see Figure 14a). The restriction patterns of SR26 and SR35 are very similar (if not identical) to that of SR2 (Figure 15). The two proximal half-losses that retain NOF sequences have very different restriction patterns when probed with
368 R. Lovering, N. Harden and M. Ashburner

**DISCUSSION**

The outstanding molecular feature of ISING's family of transposons is their intimate association with foldback DNA sequences (GOLDBERG, PARO and GEHRING 1982). It is the FB DNA that gives the TE its genetic instability. In principle, we can imagine that FB sequences bounding any other DNA sequence could become unstable and form a transposon; just so happens that a favorable genetic background, and a philosophy of "treasuring one's exceptions," lead to the identification of this TE family in which the white and roughest genes were mobilized. Three important questions follow from the molecular characterization of these TEs: (1) what is the mechanism of recombination between FB elements that leads to loss or transposition of the TE? (2) how is the propensity for recombination seen by FB elements that bound a transposon controlled? and (3) given the existence of 20–50 copies of the FB sequence in the "average" genome of D. melanogaster, how is promiscuous recombination between them, which would result in great genomic instability, avoided? Clearly, these three questions are related, and answering any one would throw light on the other two. As a first step towards an analysis of these problems we have made a detailed molecular characterization of one particular member of ISING's TE family, TE146(Z), and of many of its spontaneous derivatives.

There is a strong similarity between the instability of TE146(Z) and that found by GABAY and LAUGHNAN (1970) for Dp(1;1)MNB-8, an unusual derivative of Dp(1;1)B. Dp(1;1)MNB-8 was highly unstable in the male premeiotic germ-line, giving rise to at least three different classes of product (GABAY and LAUGHNAN 1970). Like some FB associated mutations (i.e., w^S, see BINGHAM and ZACHAR 1989), this duplication showed a tendency to stabilize in stock. Moreover, the genetic basis of this duplication's instability was not separable from the duplication itself (LAUGHNAN, GABAY and MONTGOMERY 1971).

**TE146(Z)** is an unusual, but by no means unique, member of the ISING family of transposons since it is duplicated for both its white and roughest genes (GUBB et al. 1985, 1986). Three different classes of event can readily be seen to affect TE146(Z): complete excision, "half-exciscion," where one copy of w^S, w^D, or w^y is lost and the other remains, and transposition (GUBB...
Transposon structure in Drosophila

369

FIGURE 14.—(a) Interpretation of the restriction map of limbs D and E and of the fusion between limbs A and C in SR64. The proximal EcoRI site is within NOF DNA; it is 918 bp 5′ to the junction between FB and NOF sequences (see Templeton and Potter 1989). The FB sequences of limb E have not been mapped, their representation is schematic. (b) Interpretation of the restriction map of limb C and of the fusion between limbs D and H seen in SRs that have lost the proximal half of TE146. (c) A restriction map of the fusion of limbs F and H as deduced from restriction enzyme mapping of SR44. The HindIII site is the central such site within NOF sequences (at 1030 bp of the sequence of Harden and Ashburner 1990); the NOF/FB junction is at position 3412. The positions of the TaqI sites within the NOF DNA distal to limb F are from Harden (1989). FB limbs are represented as wriggly lines, the gaps between them represent DNA lacking TaqI sites.


A hallmark of FB sequences is their length heterogeneity (Trueett, Jones and Potter 1981). Not only are the limbs of FB elements heterogeneous in length, both between and within elements (see Potter 1982b), but also there is great heterogeneity in the length and nature of the sequences that lie between the limbs of an element. Such heterogeneity is evident from the different FB elements associated with TE146. The lengths of the FB limbs vary between 0.6 and 6.0 kb, and the lengths of the loop sequences that separate the limbs from 0.5 to 4.0 kb. The longest loop is that between junctions E and F, this loop is the NOF sequence, first found to be associated with FB and TE elements by Goldberg, Paro and Gehring (1982) and Paro, Goldberg and Gehring (1983). The NOF sequence is also found associated with the FB element of the w6 mutation (Collins and Rubin 1982). Two NOF sequences have been determined, including the one from TE146(Z) (Templeton and Potter 1989; Harden and Ashburner 1990). The sequence data indicate that NOF may code for a large polypeptide, which, it has been suggested, is required for recombination between or within FB elements.

A genetic analysis of spontaneous derivatives of TE146(Z) that were no longer phenotypically zeste when on a z1 background (i.e., the SR derivatives) lead to the suggestion that most resulted from the loss of either the distal or proximal copy of w6 rst+. We have shown by a molecular analysis that this is indeed so. With one exception (SR47) all of these “half-loss” SRs change one noc/FB junction but leave the other conserved; of 14 SRs, seven have lost the distal copy of w6 rst+ and seven the proximal. It is clear, however (see Figure 9), that the different FB limbs were not
involved with equal probability in the exchanges which generated these losses. For example, all of the losses of the distal half of TE146(Z) retain the FB-NOF element and have novel FB elements at junction A whose structures suggest that they resulted from an exchange between limbs A and C. In no case was there an exchange involving limb B, or one involving either limb of the FB-NOF element. Most (5/7) of the proximal half-losses have lost FB-NOF, and the structure of their junction with noc sequences indicates their origin by exchange between limbs D and H. Only SR24 and SR44 have originated by an exchange involving the FB-NOF element, in fact of its proximal limb (F) with limb H. In no case were limbs E or G involved.

There is an additional nonrandomness in the exchanges that generated a particular class of derivative, the similarity in restriction enzyme sites between derivatives of the same class suggests that there are preferred regions within FB limbs where exchanges occur. This is most clearly seen by the similarity of junction D in SR2, SR26 and SR35 (see Figure 15). These observations all confirm the suggestion (Potter 1982b) that heterogeneity in length, and precise structure, of FB limbs occurs as a consequence of exchanges between limbs. The repetitive nature of the FB sequence (Potter 1982a) is, of course, an ideal substrate for exchange between sequences that are not precisely aligned.

We have confirmed and extended the data that implicate FB-FB exchange events in the genetic instability of ISINC's transposon. All of the derivatives we have characterized have FB limbs of novel fine structure, in comparison with those of the parental TE146(Z). In particular, the evidence that all of the half-losses have an altered distal or proximal noc/FB junction indicates that half-loss does not occur by exchange between the duplicated copies of w and rst. Were it to do so, then both the distal and proximal junctions would remain unchanged in structure. It is clear, however, that these events are not normal meiotic exchanges. Although the instability of TE146(Z) and other members of this family is germ-line specific it is due to a premeiotic rather than meiotic process (Gubb et al. 1985; G. Ising, personal
communication). Moreover it occurs with a similar frequency in males and females (see Ashburner 1989).

The obvious alternative models are that the exchanges between FB elements occur either as sister- or as iso-chromatid events (Petersen and Laughnan 1963, 1964; Gabay and Laughnan 1970; Green and Lefevre 1979). With respect to the half-loss class of spontaneous-red derivatives of TE146(Z) discussed in this paper, no distinction can be made between these two types of exchange. There is, however, another class of SR derivative that differs from that of the half-loss SRs in several respects: these derivatives remain cytologically large insertions and their w+ gene(s) cannot be suppressed by z1 in any genotype (Gubb et al. 1986). We originally thought that these SRs (e.g., SR36) possess two copies of white, from the evidence of in situ hybridization. Higher resolution in situ analysis with biotinylated probes shows that these anomalous SRs have three white genes (D. Gubb and J. Treneer, unpublished data) (in retrospect this is even evident from the [3H]thymidine in situ analysis, see Figure 3a of Gubb et al. 1986). Tripletions can result from duplications by sister-chromatid, but not by iso-chromatid, exchange. The simplest interpretation of these data is that loss, half-loss and triplication of TE146(Z) occur by premeiotic sister-chromatid exchange that is limited to the FB limbs of the transposon.

This work was supported by an MRC Programme Grant to M.A. and, in part, by an SERC Studentship to R.L. and by a Commonwealth Scholarship to N.H; R.L. also thanks the Cambridge Philosophical Society and Churchill College, Cambridge for financial support. We thank David Gubb, Trevor Lockett and John Roote and our other colleagues in Cambridge for their advice and help. David Gubb, John Roote, Gunar Ising, Kevin Moses, Kevin O'Hare and Michael Goldberg provided stocks and clones that made this work possible, and David Gubb and John Roote gave us unpublished data. We thank Vincent Pirrotta for a preprint of his review on the zeste gene.

LITERATURE CITED


Ising, G., 1964 A recessive lethal in chromosome 2, which in single dose has an effect on the eye color of white animals. Drosophila Inform. Serv. 39: 84.


Transposon structure in Drosophila


Communicating editor: A. CHOVNICK