Cloning and molecular characterization of the trithorax locus of *Drosophila melanogaster*  
(homoeotic gene/developmental genetics/insertional mutations)

BRIAN A. MOZER* AND IGOR B. DAWID
Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892

Contributed by Igor B. Dawid, February 10, 1989

ABSTRACT The trithorax (trx) locus of *Drosophila melanogaster* affects segment determination primarily in the thoracic region. Mutant flies show transformations of the third and, to a lesser extent, first thoracic segment toward the second thoracic segment; abdominal transformations also occur. Prior genetic evidence suggested that these effects are based on interactions between trx and genes of the bithorax complex and Antennapedia complex. Further, interactions between the maternal effect locus female sterile homoeotic (fsh) and trx have been observed. To aid in a molecular analysis of trx function, we have cloned the locus by a P-element transposon tagging approach. Five insertion mutations have been mapped within a region of about 10 kilobases; one of these mutations reverted coincident with the loss of the insertion. Transcription mapping suggests that two RNAs of about 12 and 15 kilobases are the major transcripts of the trx locus and that the transcription unit comprises a region of about 25 kilobases. Transcripts from the trx locus are distributed uniformly in early embryos, but at 14–16 hr after fertilization the ventral nerve cord contains a higher concentration of trx RNA than other regions of the embryo.

A genetic analysis of segmental determination in *Drosophila* has identified several genes involved in this process. These genes in concert allow groups of cells of the blastoderm to recognize positional information present in the embryo and to manifest this information by differentiating into the various tissues that make up the segments of the larva and ultimately the adult. One class of homoeotic genes, many of which are present in the bithorax complex (BX-C) and Antennapedia complex (ANT-C), specifies the segmental identity of discrete body domains. In addition to cross-regulatory interactions between the genes within the two complexes (1–3), other loci have been identified that appear to be involved in controlling the expression of BX-C and ANT-C genes. Homoeotic loci of the latter class are required in most if not all segments, and many have a significant maternal component (4, 5, 41). Some of these loci share as their loss-of-function phenotype the transformation of most or all segments toward the terminal abdominal segment (6). For two well-studied representatives of this class, Polycomb (Pc) and extra sex combs (esc), this phenotype is dependent on the presence of the BX-C; therefore, it has been postulated that Pc and esc are involved in the negative regulation of BX-C genes (1, 7). This proposal is supported by the visualization of the expression of Ultrabithorax (*Ubx*), one of the BX-C genes, at ectopic sites in *Pc* (8, 9) and *esc* (10) mutant embryos. Several genes have been identified whose phenotype may be interpreted as having a positive role in *Ubx* regulation. These include certain alleles of the fushi tarazu (*fz*) gene (11), Regulator of postbithorax (*Rg-pbx*), an allele of hunchback (hb) (12), and the trithorax (*trx*) gene.

The first *trx* allele was discovered by Lewis (13) and originally named Regulator of bithorax (*Rg-bx*) because of its interaction with the BX-C. Most *trx* alleles are zygotic lethals; larvae homozygous for certain mutant alleles exhibit patchy homoeotic transformations of abdominal segments toward more anterior identities. Flies with various *trx* allele combinations, including homozygous viable *trx* alleles, show transformations of the first and third thoracic segment toward the second thoracic segment, and also partial transformations of abdominal segments toward a more anterior identity (14–18). Many aspects of the mutant phenotype are enhanced when the fly also carries a mutation or deletion of the BX-C; conversely, some of these effects are reduced by extra copies of the BX-C (17, 18). From such effects it can be deduced, in a formal sense, that *trx* behaves like an activator of the BX-C. In addition to interactions with the BX-C, *trx* also interacts with the ANT-C (16, 19). The viable mutation *trx* exhibits a maternal effect on expressivity and penetrance (14), but the effect of removal of maternal *trx* function on the embryonic phenotype of *trx* is minimal, as examined with germ-line clones (20). Thus, any maternal contribution to *trx* function is comparatively slight. In addition to its function in the embryo, *trx* activity is required in the imaginal cells to specify normal segment identity in the adult (16). Clones homozygous for lethal alleles of *trx* induced during larval development by somatic recombination transform to second thoracic identity in most of the structures of the adult head, first leg, third leg, haltere, and genitalia. In addition *trx* clones in the posterior compartment of the wing transform to anterior and thus resemble engrailed clones (16). These observations suggest that the requirement for *trx* activity is not restricted to any particular segment and that *trx* may also play a role in regulating genes that define compartment boundaries in the wing. Thus it appears that the *trx* locus interacts with many of the major loci known to be involved in specification of segment identity. Further, *trx* interacts strongly with the maternal effect locus female sterile homoeotic (*fsh*), which has been cloned (21). To provide a basis for analyzing the molecular mechanism of *trx* function, we have cloned the gene and examined the expression and distribution of its transcripts during development.

MATERIALS AND METHODS

*Drosophila* Strains. The *trx* strain has been generated by G. Karpen (personal communication and ref. 22). The dysgenic allele *P2* and revertant *R4* were a gift of P. Ingham (Oxford University). The ethyl methanesulfonate-induced allele *trx* (23) and the dysgenic alleles B16, B17, and B18 were gifts of J. Kennison (National Institutes of Health).

Abbreviations: BX-C, bithorax complex; ANT-C, Antennapedia complex.

*Present address: Division of Biology, California Institute of Technology, Pasadena, CA 91125.*
Construction of Genomic Libraries. Total genomic DNA from the trx88B strain (provided by G. Karpen, Carnegie Institution, Baltimore) was digested to completion with EcoRI and cloned into the EcoRI site of λgt10 (Promega). Total genomic DNA from the trx2P strain was partially digested with Sau3A and cloned into the BamHI site of EMBL3 (24). Wild-type DNA was isolated by "chromosomal walking" from the Canton S library of Maniatis et al. (25). Restriction fragments were subcloned into the Bluescript vector (Stratagene).

RNA Blot Analysis. Total RNA was isolated by the guanidinium thiocyanate/phenol procedure of Sargent et al. (26). Poly(A) RNA was fractionated by oligo(dT)-cellulose chromatography essentially as described (27). RNAs were separated by electrophoresis through 1% agarose gels containing 5 mM methylmercuric hydroxide in 50 mM boric acid/5 mM sodium borate/10 mM sodium sulfate (28) and transferred by electroblotting to Nytran (Schleicher & Schuell) in 12 mM Tris/6 mM sodium acetate/0.6 mM EDTA, pH 7.8.

Hybridization. Nitrocellulose "plate lifts" and Southern blots were hybridized as described by Digan et al. (21). RNA blots using DNA probes were hybridized and washed according to Church and Gilbert (29). Blots using RNA probes were hybridized in 50% (vol/vol) formamide/5× SSPE/1% SDS with herring sperm DNA (100 μg/ml) as carrier at 65°C (1× SSPE is 0.18 M NaCl/10 mM Na2HPO4, pH 7.4/1 mM EDTA, pH 7.4). Filters were washed at 65°C once in 5× SSPE/1% SDS, once in 1× SSPE/0.1% SDS, and three times in 0.2× SSPE/0.1% SDS.

In Situ Hybridization. In situ hybridization was performed as described in Ingham et al. (30). RNA probes for blotting or in situ hybridization were synthesized in vitro using T7 or T3 RNA polymerase in the presence of the appropriate radiolabeled ribonucleotide triphosphate.

RESULTS

Isolation of DNA from the trx Locus. Entry into the trx locus was obtained by cloning the breakpoint fragments from a strain bearing a P-element transposon inserted into the trx locus. The trxPr7, ry+1888 strain, hereafter called trxrib7, carries the transposon P[rib7, ry+1] that contains a Drosophila ribosomal transcription unit and the rosy marker gene at position 88B and fails to complementtrx mutations (G. Karpen, personal communication and ref. 22). A 4.4-kilobase (kb) EcoRI fragment was isolated from a trx88B genomic library by homology to rosy and P-element sequences, and flanking sequences were used to isolate two phases (clones 12 and 20) from a wild-type genomic library. Clone 20 overlaps the proximal end of a chromosome walk of Parkhurst et al. (31), orienting the two clones along the chromosome. An additional 50 kb of genomic DNA was isolated proximal to clone 12.

The restriction map of the trx region is shown in Fig. 1; it is drawn in reverse orientation to the genetic map in accordance with transcriptional polarity (see below). Various restriction fragments that together cover the entire 50-kb region were used as probes in genomic Southern blot analysis of several trx mutants. Among the mutant strains tested five were found to be associated with insertions (indicated in Fig. 1), and a sixth mutant appears to represent a deletion (see below). Examples of blots supporting the mutant mapping are shown in Fig. 2.

trx88B carries an insertion of a 21-kb P-element transposon (rib7) into the 1.35-kb EcoRI-HindIII fragment at position 34.5 to 33 on the DNA map. A blot of EcoRI-digested DNA from ry506 (the parental strain) and from trx88B/ry506 heterozygotes, hybridized with a probe from the rosy gene, shows two additional bands of 11.8 kb and a 4.4 kb in the mutant (Fig. 2). The 11.8-kb fragment is an internal fragment from the transposon including part of rosy, whereas the 4.4-kb fragment contains the remaining portion of the rosy gene and the proximal breakpoint of the transposon.

The dysgenic allele P2 is associated with an insertion into the same 1.35-kb EcoRI-HindIII fragment (position 34.5 to 33) that is affected in the rib7 chromosome. The inserted fragment was cloned from a P2 genomic library, and the insertion found to be a 0.9-kb P element. HindIII-digested genomic DNA from the mutant strain P2 and from a revertant (R4) generated by hybrid dysgenesis (P. Ingham, personal

![Diagram](image-url)

**Fig. 1.** The trx locus. The top lines show the four Φ phage inserts covering the locus. Next is a scale in kb and a restriction map of genomic DNA. The five mapped insertion mutations are indicated, with the lengths of inserts and points of insertion drawn approximately to scale. Immediately below the map are the locations of three probes that were used for RNA blotting as illustrated in Fig. 3. The bottom of the figure shows the position of cDNA clones and the polarity of transcription. Symbols for the restriction sites shown are: B, BglII; H, HindIII; S, SalI; Xb, XbaI; X, XhoI. A restriction site polymorphism between wild-type strains is indicated by parenthesis.
communication) was hybridized with a probe extending from position 47 to position 29.5 (Fig. 2). Because the HindIII site at position 34.5 is polymorphic in wild-type Drosophila strains (data not shown), Southern blots detect either a 4-kb plus a 5-kb fragment, or a 9-kb fragment in this region. The polymorphic site is present in the TM6 balancer chromosome but absent in both the mutant (P2) and the revertant (R4) chromosomes. The insertion of a 0.9-kb P element with its two HindIII sites into the 9-kb HindIII fragment generated a 4.5-kb doublet band in the mutant. Reversion is accompanied by the loss of the doublet band and restoration of the 9-kb wild-type band. The change observed in the P2 mutation and its reversion are the strongest evidence that we have cloned the trx locus.

The dysgenic alleles trx^B16 and trx^B18 (J. Kennison, personal communication) show insertions of 1.5 and 2.5 kb, respectively, into the 3-kb EcoRI fragment at position 32 to 29 (Fig. 2). The dysgenic allele trx^B17 (J. Kennison, personal communication) is associated with a 1-kb insertion into the distal region of the 5.4-kb EcoRI fragment at position 26.5 to 21 (Fig. 2). The trx allele E3 showed changes in restriction pattern consistent with a deletion of about 900 base pairs at position 16 (data not shown).

trx Transcripts. Transcription from the region that includes the breakpoints of the insertion mutations and surrounding arrays was examined by RNA blot analysis, using poly(A)^+ RNA from various developmental stages. Fig. 3 shows examples with three probes that are indicated below the genomic map of Fig. 1. Probe 2, derived from the 1.35-kb EcoRI-HindIII fragment that spans the trx^B16 and trx^B2 breakpoints (position 34.5 to 33), detects two major transcripts of 12 and 15 kb in RNA from embryos and pupae (Fig. 3). The 12-kb mRNA is the predominant species found in embryos; it was detected in RNA from 0- to 3-hr embryos and increased in 3- to 6-hr embryos. Additional transcripts of 10 kb and 3 kb were detected at 0-3 hr only; the 10-kb RNA is quite rare but was seen repeatedly. The 15-kb RNA could not be detected in 0- to 3-hr RNA, was rare but detectable in 3- to 6-hr RNA (Fig. 3), and became equal in abundance with 12-kb RNA between 6 and 12 hr (data not shown). In larvae (data not shown) and in 8- to 9-day pupae (Fig. 3, lane MP), the most abundant trx transcript is the 15-kb RNA, but the 12-kb RNA is also present at lower levels. The 15-kb RNA is present in poly(A)^+ RNA from adult males and females; in addition, a 10-kb transcript is also seen in females (data not shown). The fact that the 10-kb transcript is female specific and present in 0- to 3-hr embryos suggests that it is maternally derived. Because probe 2 was an RNA probe, its use allowed us to determine the polarity of transcription of the 12- and 15-kb RNAs, which is distal (5') to proximal (3') (Fig. 1).

Probe 2 is the most distal fragment that detected the 12-kb and 15-kb transcripts. The 4-kb region adjacent to probe 2, from positions 33 to 29 and containing the B16 and B18 breakpoints, appears to constitute an intron common to both RNAs since these sequences did not detect any RNA (see also below). The region from positions 29 to 24 contains repetitive sequences that obscured the blots. The 15 kb of DNA from positions 24 to 9 hybridized to both the 12- and 15-kb transcripts, the most proximal probe that detected both RNAs being the 2-kb Xba I-Xho I fragment at positions 13 to 9. Probes derived from a 10-kb region proximal (or 5' in terms of the trx transcription unit) to this fragment failed to detect the 12- and 15-kb RNAs but hybridized to several smaller transcripts that are not known to be related to trx. Likewise, sequences distal to probe 2 hybridized to smaller RNAs (Fig. 3, probe 1). Thus, probes derived from the region of DNA to which trx mutations have been mapped detected two predominant transcripts of 12 and 15 kb, which we consider as the major transcripts of this gene.

Characterization of cdNA Clones for the Major trx Transcripts. Several cdNA clones were obtained from embryonic libraries as indicated at the bottom of Fig. 1. Clone A2 appears to correspond to the RNA visualized by probe 1 (Fig. 3) and is likely to be unrelated to trx. The other four cdNAs, and several additional examples that represent portions of the
same regions, were selected with genomic sequences that hybridize to the 12- and 15-kb RNAs and are, therefore, expected to represent the major trx transcripts. We cannot assign any of the cDNAs to either the 12- or 15-kb RNA. The mapping of cDNAs was done by cross hybridization and restriction mapping, and, therefore, exon boundaries have been assigned only approximately. Further, only the major introns have been identified, and additional introns probably occur. Nevertheless, the analysis of cDNA clones has confirmed some general conclusions derived from RNA blotting, including the approximate ends of the transcription unit and the presence of a large intron between positions 24 and 33. An alternatively spliced RNA represented by cDNA 9 includes an exon within this region. The additional exon in cDNA 9 is slightly less than 1 kb long and cannot by itself account for the size difference between the 12- and 15-kb RNAs, which remains unexplained at present.

The CAG/CAA repeat, also called M, opa, or strep, a sequence found in many homeotic genes of the BX-C and ANT-C and elsewhere (32), is present in the trx gene at positions 14–15 and 25–26 (Fig. 1). The repeated sequence is included in cDNAs 9 and 6C12, showing that it is transcribed into RNA.

Localization of trx Transcripts in the Embryo. The distribution of the trx transcripts in the developing embryo was examined by in situ hybridization. A single-copy portion of cDNA clone 15 that contains sequences complementary to the 12- and 15-kb transcripts was used to generate RNA probes. In all cases, sense-strand probes that do not detect the trx RNAs by blot analysis showed only background labeling over the embryo sections (data not shown). trx transcripts were detected initially in a uniform distribution throughout the cytoplasm of embryos undergoing the nuclear divisions prior to cellularization (Fig. 4 A and B). The presence of grains over sections of embryos from precellular blastoderm stages may represent maternal RNA. Transcripts were detected in all cells of the cellular blastoderm and gastrula, suggesting that there is no early germ-layer specificity of trx expression. A similar homogeneous pattern of expression persisted in embryos undergoing germ-band elongation (Fig. 4 C and D). Furthermore, there was no indication of a nonuniform distribution of grains along the anterior–posterior or dorsal–ventral axis in cross sections of embryos from any of the stages mentioned above.

Many of the homeotic genes of the BX-C and ANT-C are preferentially expressed in certain tissues during later stages of embryogenesis; this also appears to be true for trx expression. Fig. 4 E and F shows in situ hybridization of a trx probe to a sagittal section of a late stage embryo; prominent labeling is apparent in the ventral nerve cord at levels above that seen in other tissues.

**DISCUSSION**

A DNA region from 88B was identified as the trx gene of *Drosophila* by mapping of five insertion mutations. Reversion of one of these mutations, trx^P^, results in the excision of the inserted P element. Analysis of RNA products suggests that the length of the trx transcription unit is about 25 kb. This region, which includes the breakpoints of all mapped insertions, is transcribed in a distal-to-proximal direction to give rise to RNAs of 12 and 15 kb which we consider as the major zygotic trx transcripts. Distinct transcripts were mapped to

---

**Fig. 4.** Expression of trx transcripts in the developing embryo. (A and B) Dark-field and bright-field views, respectively, of a sagittal section of a syncytial blastoderm stage embryo hybridized with a ^3^H-labeled probe from the trx^rib/ttx^P^ breakpoint fragment. Uniform labeling of the entire embryo is seen. (C and D) Dark-field and bright-field views, respectively, of a sagittal section of an embryo undergoing germ-band extension hybridized with a ^3^H-labeled probe corresponding to the 900-base-pair 3′-terminal SalI–EcoRI fragment from cDNA 15. Hybridization signal is seen throughout the cells of the germ band, while far fewer grains are present over the interior yolk. (E and F) Dark-field and bright-field views, respectively, of a sagittal section of a 14- to 18-hr-old embryo hybridized with the same probe as in B. Anterior is to the left, dorsal is down. There is prominent labeling over the ventral cord (VC) and weaker labeling over other tissues.
genomic regions both distal and proximal to this region; their relationship to trx, if any, is unknown.

The expression of the *trx* gene is developmentally regulated. The 15-kb RNA is the major transcript in larvae, pupae, and adults, whereas the 12-kb transcript predominates in early embryogenesis. Both transcripts are correlated with the zygotic lethal function for *trx* (15). A rare 10-kb mRNA, present in adult females and 0- to 3-hr embryos, may correspond to the maternal component of *trx* function (14, 18). RNA gel blotting and analysis of cDNA clones derived from portions of the embryonic forms of the 12- and/or 15-kb transcripts suggest that a common 5′ exon is joined to a long region of common 3′ exons. The breakpoints for two of the *trx* insertions (B16 and B18) fall within a long intron. Alternate splicing of exons between the common 5′ and 3′ exon regions, as seen in cDNA clones 9 and 15, may account for some of the difference between the 12-kb and 15-kb transcripts. Whereas the CAG/CAl repeat is present in two positions in the *trx* gene (see Results), no homeobox homology could be detected by Southern blot hybridization.

Examination of the spatial distribution of *trx* transcripts in the developing embryo suggests that the zygotic mRNAs are ubiquitously expressed until late in embryogenesis. Although we cannot rule out that one of the two transcripts is spatially restricted, the observed distribution is consistent with the requirements for *trx* in most or all of the segments as defined by clonal analysis (16) and the distribution of homeotic transformations seen in *trx* homozygotes (14, 19). Several genes that function to specify cell fates in early development are ubiquitously expressed in the embryo. The distribution of the maternal RNA of the T testis segmental transformations are minor in the embryo as compared to the adult. Duncan and Lewis (17) and Capdevila and Garcia-Bellido (18) have proposed that the *trx* gene (specifically the *Rg-bx* allele studied by these authors) initiates BX-C gene expression by establishing a gradient along the anterior–posterior axis of the embryo. The finding that differential expression of the genes of the BX-C and ANT-C occurs in larval cells in the absence of *trx* suggests that this gene functions in the maintenance rather than the initiation of BX-C and ANT-C expression (20). In addition, the spatial distribution of *trx* RNAs in the embryo does not provide any support for a gradient distribution of the protein product.

Genetic interactions between the maternal effect mutation female sterile hemoctric (*fsh*) and *trx* suggest that zygotic *trx* function is dependent on maternal *fsh* gene product (21, 39, 40). The pleiotropic effects of these two mutations and the ubiquitous expression of the *trx* transcripts, as shown here, and of the maternal transcripts from *fsh* (42) in the early embryo suggest that these genes may have additional functions unrelated to the determination of segment identity. Alternatively the distribution of the transcripts may not reflect the ultimate distribution of the protein products or their function which might, for example, be mediated by interactions with other molecules. Further characterization of the *trx* gene products should lead to a resolution of these questions.

We thank Gary Karpen, Phil Ingham, and Jim Kennison for mutant strains without which this work could not have been done; L. Kauvar for cDNA libraries; Der-Hwa Huang, Alexander Mazo, Susan Parkhurst, and especially Susan Haynes and Jim Kennison for information, advice, and varied help throughout.