Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns

Rodney N. Nagoshi[1,2] and Bruce S. Baker[1]

[1]Department of Biological Sciences, Stanford University, Stanford, California 94305-5020 USA

Sex-specific alternative RNA splicing of the *doublesex* (dsx) pre-mRNA results in sex-specific polypeptides that regulate both male and female somatic sexual differentiation in *Drosophila melanogaster*. We have molecularly characterized a class of *dsx* mutations that act in *cis* to disrupt the regulation of *dsx* RNA processing, causing the *dsx* pre-mRNA to be spliced in the male-specific pattern regardless of the chromosomal sex of the fly. These *dsx* mutations are associated with rearrangements in the female-specific exon just 3′ to the female-specific splice acceptor. The mutations do not affect the female-specific splice sites or intron that are identical to wild-type sequences. These results indicate that sequences in the female-specific exon are important for the regulation of sex-specific RNA splicing, perhaps by acting as sites of interaction with *trans*-acting regulators. Furthermore, the data suggest that female-specific regulation of *dsx* RNA processing occurs by promoting the usage of the female splice acceptor site, rather than by repressing the usage of the alternative male-specific splice acceptor.

[*Key Words*: RNA splicing; sex determination; *doublesex*; *Drosophila*]

Received September 12, 1989; accepted October 13, 1989.

RNA splicing occurs through the interaction of *trans*-acting factors, such as the individual components of spliceosomes, with conserved *cis*-acting sequences found in pre-mRNAs. These sequences include the splice donor, the splice acceptor, and the branchpoint, all of which are required for correct splicing in both in vitro and in vivo systems (for review, see Green 1986; Padgett et al. 1986; Sharp et al. 1987a,b). In a growing number of genes, however, RNA splicing is complicated by the fact that a single pre-mRNA can be differentially spliced (Breitbart et al. 1987). The usage of alternative RNA splice sites provides a mechanism by which a single gene can produce multiple, structurally related proteins whose expression can be regulated. Well-characterized examples of regulated splicing include the vertebrate calcitonin–cGRP gene (Amara et al. 1982; Sabate et al. 1985; Leff et al. 1987), the mammalian fibronectin gene (Kornblihtt et al. 1985; Scharzbauer et al. 1987; Barone et al. 1989), the IgM heavy-chain gene (Alt et al. 1980; Early et al. 1980; Rogers et al. 1980; Galli et al. 1987, 1988; Peterson and Perry 1989), and the adenovirus E3 transcription unit (Nevins and Chen-Kiang 1981; Adami and Nevins 1988). Alternative RNA processing can also be used as an 'on-off switch,' in which one splicing from encodes a protein product and an alternative form does not (Boggs et al. 1987; Bell et al. 1988; for review, see Bingham et al. 1988).

How the choice is made between differentially used splice sites is not well understood. A priori expectations are that *cis*-acting regulatory elements are required to confer differences between the alternatively used RNA splice sites to create asymmetry in their availability to the cellular splicing machinery. A number of *cis*-acting regions have been proposed to be involved in splice site choice; these appear to differ depending on the gene studied. For example, in the calcitonin–cGRP gene, differential poly(A) site selection and exon usage lead to alternative mature transcripts. This regulated RNA processing depends on sequences at the differentially used 3′ splice acceptor but appears to be independent of sequences at the poly(A) sites (Leff et al. 1987; Adami and Nevins 1988). This observation leads to the suggestion that *trans*-acting regulatory factors act to promote the usage of one splice acceptor over another which, in turn, determines the poly(A) site to be used. In comparison, two different models have been proposed for the regulation of differential RNA processing at the IgM gene, which also exhibits alternative 3′ exon usage and polyadenylation. Galli et al. (1987, 1988) suggested that the regulation of IgM occurs at the level of poly(A) site

[2]Current address: Department of Biology, University of Iowa, Iowa City, Iowa 52242 USA.

choice which, in turn, determines the splicing pathway that is taken. An alternative model was proposed by Peterson and Perry (1989). They presented evidence that the IgM RNA splice pattern is determined by competition between the use of a nonconsensus IgM splice donor and a nearby poly(A) site. The choice between these mutually exclusive sites is dependent on cell-specific, *trans*-acting factors. Additional sites required in *cis* for splice site selection have been implicated in studies on other loci, including the branch site (Noble et al. 1987; Fu et al. 1988) and sequences within the alternatively spliced exons (Somasekhar and Mertz, 1985; Mardon et al. 1987; Barone et al. 1989; Hampson et al. 1989; Laski and Rubin 1989; Streuli and Saito 1989). This diversity of regions involved in regulating splice choice suggests that multiple mechanisms exist for controlling RNA processing.

A complete understanding of how RNA processing is regulated will require the characterization of both *cis*- and *trans*-acting factors that regulate RNA processing site choice. One of the best systems for studies of this kind is the regulatory pathway that controls somatic sexual development in *Drosophila*. Extensive genetic and molecular studies have demonstrated that the sex-specific regulation of this developmental pathway is controlled primarily, if not solely, by sex-specific RNA splicing (for review, see Baker 1989; Hodgkins 1989). Particularly advantageous is that at least some of the loci required in *trans* for the sex-specific splicing of these genes have been identified.

In *Drosophila*, the choice between male or female somatic development is determined by the assessment of the X/A ratio, the number of X chromosomes relative to the number of sets of autosomes. The X/A ratio determines somatic sexual differentiation by ultimately controlling the expression of the bifunctional *doublesex (dsx)* gene (Baker and Ridge 1980). In chromosomally female flies (X/A = 1 : 1), a hierarchy of regulatory genes act in a response to the X/A ratio to cause *dsx* to be expressed in a female-specific manner. The *dsx* female product represses female differentiation. In contrast, in male flies (X/A = 1 : 2) *dsx* is expressed in a male-specific manner and the *dsx* male product acts to repress female differentiation. In this regard *dsx* is unique among the loci involved in somatic sexual differentiation in that it is the only gene whose active function is required to regulate the development of both sexes (Baker and Ridge 1980). Loss-of-function mutations of *dsx* cause intersexual development in both chromosomally male and female flies.

The basis of the bifunctionality of the *dsx* gene derives from sex-specific RNA processing (Burtis and Baker 1989). The *dsx* gene produces both male- and female-specific mRNAs that are derived from a common pre-mRNA and have the same first three 5′ exons (Fig. 1). They differ in their 3′ exons as a consequence of the use of sex-specific splice acceptor and poly(A) sites. One female RNA of 3.5 kb and two male-specific transcripts of 2.9 and 3.9 kb are detected by Northern blot analysis. The two male RNAs arise from differential poly(A) site usage in the last male-specific, noncoding exon. This male- and female-specific alternative RNA processing results in sex-specific polypeptides that differ in their carboxy-terminal domains.

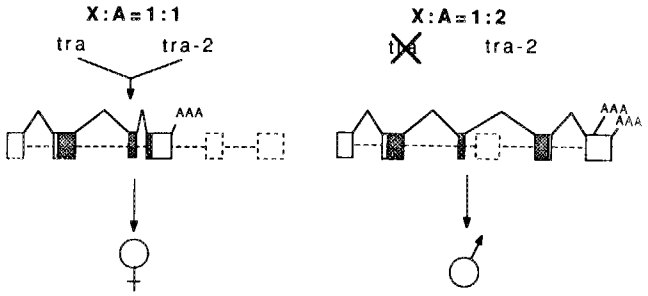

Figure 1. Model for the regulation of *dsx*. The *dsx* gene displays regulated RNA splicing and polyadenylation, which results in sex-specific polypeptides (Burtis and Baker 1989). In chromosomally female flies, where the ratio of the number of X chromosomes to sets of autosomes (X/A) is 1 : 1, the *tra* and *tra-2* genes act to cause female-specific RNA processing of the *dsx* gene (Nagoshi et al. 1988). In chromosomally male flies (X/A = 1 : 2), the *tra* gene is inactive. In the absence of female-specific regulation, the *dsx* gene is expressed in the default male mode. RNAs are not drawn to scale. (Open boxes) Exons; (boxes outlined by broken lines) unspliced exons; (shaded boxes) open reading frames; (AAA) polyadenylation sites; (diagonal lines) splice pattern; (broken line) introns.

Male-specific RNA splicing and polyadenylation of the *dsx* pre-mRNA is thought to be the default state of *dsx* RNA processing, because it occurs in the absence of activity of the genes known to regulate sexual differentiation (Baker et al. 1987; Nagoshi et al. 1988). In contrast, female-specific RNA processing at *dsx* is dependent on the activity of the *transformer (tra)* and *transformer-2 (tra-2)* genes (Fig. 2), as loss-of-function mutations of *tra* or *tra-2* in chromosomally female flies results in the male *dsx* RNA splicing pattern (Nagoshi et al. 1988). Although it has not been established that the *tra* and *tra-2* products interact directly with the *dsx* pre-mRNA to impose female-specific RNA splicing, a direct interaction is suggested by the fact that the predicted *tra-2* polypeptide contains a 90-amino-acid domain found in many proteins known to bind RNA (Amrein et al. 1988; Goralski et al. 1989). In addition, the predicted *tra-2* (Goralski et al. 1989; Mattox et al. in prep) and *tra* (Belote et al. 1989) proteins also contain stretches of arginine- and serine-rich regions that are characteristic of some proteins that are involved in, or that may regulate, RNA processing (Chou et al. 1987; Bingham et al. 1988; Query et al. 1989).

We are interested in identifying and characterizing *cis*-acting *dsx* sequences that are important for sex-specific RNA processing. These are potential sites of interaction for *trans*-acting regulatory factors such as the *tra* and *tra-2* products. In this paper we describe a class of *cis*-acting mutations at *dsx* that alter the *dsx* sex-specific RNA splicing pattern. We describe the phenotype, location, and structure of these lesions and discuss their potential role in the regulation of *dsx* RNA processing.

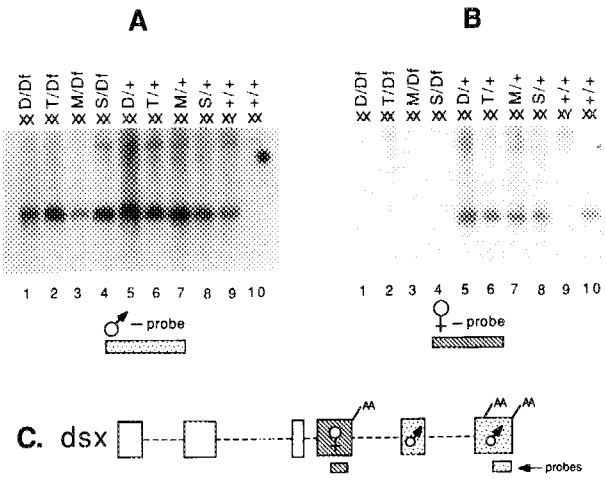

Figure 2. Expression of the sex-specific *dsx* RNAs in flies carrying the *dsx* dominant mutations. Northern blots displaying the sex-specific *dsx* RNA patterns found in flies carrying the *dsx* dominant mutations are shown. The two autoradiographs in each set represent separate blots derived from the same RNA preparations and prepared in parallel. (*A*) Northern blot hybridized with a male-specific probe showing the expression of the male-specific *dsx* transcript in flies hemizygous (lanes *1–4*) or heterozygous (lanes 5–8) for a *dsx* dominant allele. D, T, M, and S represent the dsx^D, dsx^T, dsx^M, and dsx^S dominant alleles, respectively. Df designates the dsx^{M+R15} deletion (Baker and Wolfner 1988) that is deleted for the entire *dsx* locus. Wild-type RNAs (+/+) are from Canton S flies. (*B*) Northern analysis of the same blot hybridized to a probe specific for the female-specific *dsx* transcript (*C*) Diagram of the *dsx* gene (not drawn to scale) describing the location of the sex-specific exons and the origin of the sex-specific probes used. Hatching indicates female-specific sequences; stippled boxes designate male-specific sequences.

Results

dsx dominant mutations

Genetic studies have identified a group of four dominant mutations at *dsx* that are distinct from the recessive loss-of-function *dsx* mutations in their phenotypic effects (Duncan and Kaufman 1975; Baker and Ridge 1980; Nothiger et al. 1980, 1987). The *dsx* dominant mutations are designated dsx^D, dsx^M, dsx^T, and dsx^S and are independent in origin but have nearly identical mutant phenotypes (summarized in Table 1). The *dsx* dominant mutations have two effects on *dsx* gene expression: (1) They eliminate the female-specific *dsx* function, and (2) they cause constitutive expression of the male *dsx* function. Thus, chromosomally female (XX) flies hemizygous for the *dsx* dominant mutations [$dsx^{dom}/Df(dsx)$] develop somatically as males. However, when the *dsx* dominant alleles are heterozygous with a wild-type *dsx* allele, they transform XX flies into intersexes that are indistinguishable from those derived from homozygous recessive *dsx*-null mutations (see below). In contrast, no mutant phenotype is found in chromosomally male (XY) flies carrying a *dsx* dominant mutation, either hemizygous or over a wild-type *dsx* allele. This is consistent with the dominant *dsx* alleles expressing the male-specific *dsx* function that is normally active in male flies.

dsx *dominant mutations alter sex-specific RNA splicing*

To determine how the *dsx* dominant mutations disrupt *dsx* expression, we first examined the effects of these mutations at the transcript level. In particular, because previous studies demonstrated that sex-specific regulation at *dsx* is mediated through the control of RNA processing, we examined whether the *dsx* dominant mutations disrupt the regulation of *dsx* RNA splicing.

The *dsx* dominant mutations alter the types of *dsx* transcripts produced in chromosomally female flies. This is illustrated by Northern blot analysis using single-stranded DNA probes specific for either male (Fig. 2A) or female *dsx* transcripts (Fig. 2B). The wild-type pattern is shown in lane 9 for males and lane 10 for females. Lanes 1–4 contain RNA from chromosomally female flies that carry a *dsx* dominant allele over a deletion of the *dsx* locus. Each of the four dominant alleles produces the male-specific *dsx* transcripts (Fig. 2A) and not the female-specific *dsx* transcript (Fig. 2B), even though these flies are chromosomally female. In chromosomally female flies heterozygous for a *dsx* dominant and a dsx^+ allele (lanes 5–8), both male- and female-specific transcripts are present—the female transcript from the dsx^+ allele and the male transcript from the *dsx* dominant chromosome.

These results demonstrate that the *dsx* dominant mutations are *cis*-acting lesions that not only eliminate the

Table 1. dsx *dominant mutant phenotypes*

	Somatic phenotype	*dsx* function ♂	*dsx* function ♀
XX; $\frac{dsx^{dom}}{Df(dsx)}$[b]	male	+	–
XX; $\frac{dsx^{dom}}{+}$	intersex	+	+
XX; $\frac{dsx^-}{dsx^-}$	intersex	–	–
XY; $\frac{dsx^{dom}}{Df(dsx)}$	male	+	–
XY; $\frac{dsx^{dom}}{+}$	male	+	–
XY; $\frac{dsx^-}{dsx^-}$	intersex	–	–

The somatic phenotypes of the *dsx* dominant mutations are described. The presence (+) or absence (–) of *dsx* function is extrapolated from the sexual phenotype of the fly.
[a]All four dominant mutations gave similar phenotypes.
[b]*Df(dsx)* is dsx^{M+R15} (B. Baker et al., in prep.), a deletion for the entire *dsx* locus.

female transcript but also cause the male *dsx* transcript to be expressed in both chromosomally female and male flies. This indicates that the two phenotypes displayed by the *dsx* dominant mutations, the absence of female *dsx* function, and the constitutive expression of male *dsx* function are the consequences of an abnormal pattern of *dsx* RNA processing. The ectopic expression of the male *dsx* transcript in chromosomally female flies argues against the dominant lesions affecting the stability or accumulation of the female-specific *dsx* RNA. Rather, it appears that the pre-mRNA from the *dsx* dominant alleles is incapable of female-specific splicing and polyadenylation, instead displaying the default (male) RNA processing pattern.

The *dsx* RNA splicing pattern in flies hemizygous for a *dsx* dominant allele is similar to that found in flies lacking *tra* or *tra-2* activity. In both cases, the absence of the female-specific *dsx* RNA is associated with the presence of the male-specific *dsx* transcript. A simple explanation for these results is that the *dsx* dominant mutations are *cis*-acting lesions that prevent *tra* and *tra-2* , or factors regulated by these genes, from interacting with the *dsx* pre-mRNA. This would prevent female-specific regulation of *dsx* RNA processing, thereby causing constitutive usage of the default male splice acceptor site.

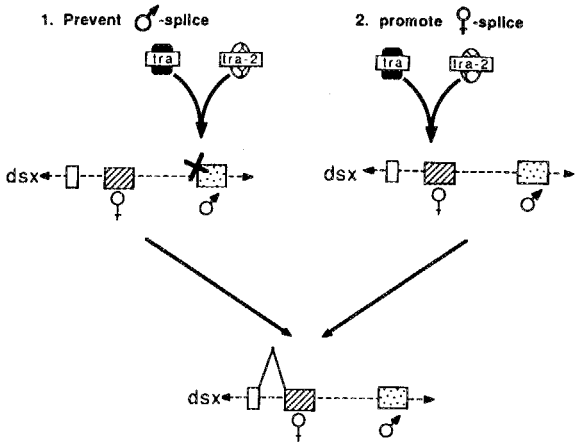

Figure 3. Models for the regulation of *dsx* RNA splicing by *tra* and *tra-2*. The portion of the *dsx* gene that contains the common splice donor, the female-specific exon, and the first male-specific exon is shown. In model 1 the *tra* and *tra-2* products act to block the male splice in chromosomally female flies. This will allow the female splice acceptor to be used. In model 2, the *tra* and *tra-2* products act positively on the female splice acceptor to promote the occurrence of female-specific splicing. Diagonal lines indicate female-specific sequences; stippled boxes designate male-specific sequences.

Mapping of the dsx *dominant mutations*

Because the *dsx* dominant lesions act in *cis* to disrupt sex-specific RNA processing, it is likely that they are located in regions that are important for sex-specific regulation. The mapping of these lesions should provide insight into the mechanism by which *dsx* RNA processing is controlled. Two topics, in particular, can be addressed. First, because the *dsx* pre-mRNA undergoes both sex-specific splicing and polyadenylation, the regulation of *dsx* RNA processing could be dictated by the choice of splice acceptor, poly(A) site, or a combination of the two. The clustering of the *dsx* dominant lesions in the vicinity of one of these RNA processing sites would argue for its importance in the regulation of *dsx* sex-specific expression. Second, the regulation of *dsx* RNA processing by *trans*-acting factors, such as *tra* and *tra-2*, could occur by either positive or negative mechanisms (Fig. 3). In chromosomally female flies, the wild-type *tra* and *tra-2* products could have a negative regulatory role by acting at the male splice acceptor [or poly(A) site] to prevent male-specific RNA processing from occurring. This would then allow the alternative female splice acceptor to be used. On the other hand, *tra* and *tra-2* could act positively to promote the usage of the female-specific processing sites instead of the default male-specific sites. The location of the *dsx* dominant lesions relative to the male- and female-specific RNA processing sites could help distinguish between these possibilities.

The initial localization of the *dsx* dominant mutations was done by restriction fragment length polymorphism (RFLP) analysis, using a combination of both Southern blotting of genomic DNAs and the DNA cloning and characterization of each of the four *dsx* dominant alleles (Baker and Wolfner 1988; this paper). All four alleles displayed restriction pattern alterations from the wild-type map within an ~3-kb interval containing the female-specific splice acceptor and polyadenylation sites. Across the 40-kb *dsx* locus, this was the only region where all four dominant alleles differed in their restriction pattern from the wild-type Canton S pattern. The correlation in four independent mutations between the location of a rearrangement in a discrete region of the *dsx* gene and the *dsx* dominant mutant phenotype provides strong support for the inference that these rearrangements are the mutations.

Sequence analysis of the mutant DNAs precisely localized the structural lesions associated with each of the dominant mutations. All four dominant alleles are associated with rearrangements with breakpoints in the vicinity of the female-specific splice acceptor (Fig. 4). The dsx^D, dsx^T, and dsx^M alleles are each associated with an insertion of middle repetitive elements that lie 48, 127, and 175 bp, respectively, 3′ of the female-specific splice acceptor site and within the female-specific exon. The dsx^S allele carries a 448-bp deletion associated with the addition of five bases: CCAAA. The 5′ breakpoint of the deletion is located 108 bp, 3′ of the female-specific splice acceptor. These rearrangements are completely localized within the female-specific exon. In contrast, the female-specific splice sites and the entire female-specific intron were found to be identical to the wild-type DNA sequence for all four dominant mutations.

Unlike the close proximity of these rearrangements to the female splice acceptor, all are located >600 bp upstream from the female canonical AATAAA polyadenylation signal. The most distant rearrangement, the dsx^D

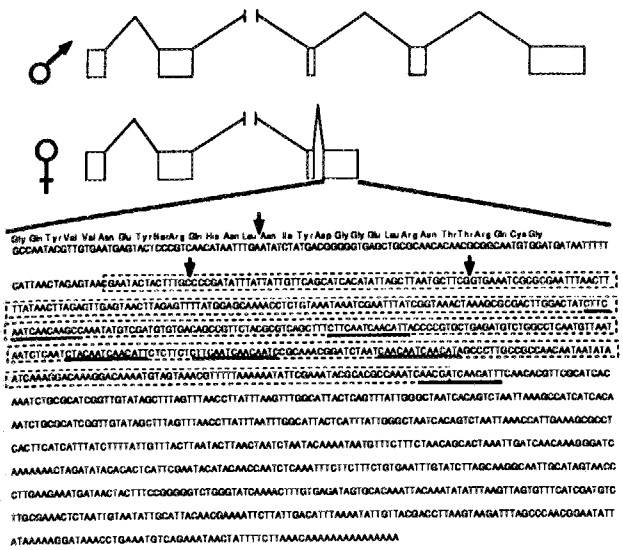

Figure 4. Location of the *dsx* dominant mutations. The map of the *dsx* gene is shown (not drawn to scale). Below this, the DNA sequence of the 5′ portion of the female-specific exon is displayed. All four *dsx* dominant alleles are associated with rearrangements in this sequenced region. Arrows indicate the location of insertions that are associated with the following dominant alleles dsx^M, dsx^T, and dsx^D, listed in order from the female-specific splice acceptor. The broken box indicates the sequences deleted in the dsx^S allele. This deletion is also associated with a CCAAA insertion at the deletion breakpoint (not shown). The six 13-mer repeats are delineated by the underlined regions.

insertion, is found 1.2 kb away from the female AATAAA hexanucleotide. This hexanucleotide sequence is the most 5′ regulatory element known to be required in *cis* for pre-mRNA cleavage–polyadenylation (for review, see Manley 1988); thus, it seems unlikely that the *dsx* dominant-associated rearrangements are directly affecting RNA cleavage and polyadenylation. The male-specific splice acceptors and poly(A) sites are even more distant from these rearrangements. The most 3′ *dsx* dominant rearrangement, the dsx^S deletion, is >3 kb from the nearest male-specific splice acceptor and >8 kb from the male polyadenylation site.

The close proximity of the *dsx* dominant lesions to the female-specific splice acceptor site strongly suggests that these mutations have their primary effect on the usage of the female-specific splice acceptor. Because these mutations do not disrupt the consensus splicing elements in the female-specific splice sites or intron, the evidence suggests that the integrity of sequences in the female-specific exon are required for female-specific regulation of *dsx* RNA splicing. An interesting structural feature of this exon is the presence of six copies of a 13-mer sequence (Burtis and Baker 1989) interspersed in the region deleted by the dsx^S deletion (Fig. 4). A portion of this repeat is found in the adjacent 114-bp female-specific intron (Burtis and Baker 1989). The possibility that this repeat may be involved in the interaction of the *dsx* pre-mRNA with *trans*-acting regulators is currently being tested.

Characterization of the dsx *dominant-associated insertions*

The insertions associated with dsx^D, dsx^M, and dsx^T are described in Table 2. Both the dsx^D and dsx^T insertions are associated with a 20- to 25-residue stretch of poly(A) at one of their ends, a characteristic of certain families of middle repetitive retroposons in *Drosophila melanogaster* (Di Nocera et al. 1983; Pittler and Davis 1987; Di Nocera 1988). The dsx^M element shares substantial sequence homology with one such family of repeats, the F elements (Table 2B; Di Nocera et al. 1983), in the region immediately adjacent to the poly(A) stretch. However, the dsx^M element may represent a different subgroup of this family because no homology could be found with that of the consensus F element in the sequence of the end opposite the poly(A) stretch. No sequence homologies could be found for the dsx^T element with the F family or with the other sequences contained in the National Institutes of Health (NIH) GenBank (release 58.0) and EMBL Nucleotide Sequence Data Library (release 17.0). Thus, it may represent a new family of RNA-mediated transposable elements. The dsx^D insertion was found to have substantial sequence similarity with the roo transposable element (Table 2B; Scherer et al. 1982; Garfinkel et al. 1983). It is 7.8 kb in size and is present in ~100 copies per genome, as estimated from genomic Southern blots.

Male and female dsx *products are antagonistic*

On the basis of genetic studies that indicate that the male- and female-specific *dsx* products act to repress female and male sexual differentiation, respectively, the simultaneous expression of both products should repress sexual differentiation altogether. The expected phenotype of this situation is difficult to predict; one might expect the absence of sexually dimorphic tissues or that these tissues would develop but be sexually undifferentiated. Instead, the *dsx* dominant mutation over dsx^+ genotype results in an intersexual phenotype that is indistinguishable from that resulting from *dsx* null mutations (Table 1; Baker and Ridge 1980; Nothiger et al. 1987). This suggests that the male and female *dsx* products are mutually antagonistic; their simultaneous expression in a single cell is functionally equivalent to the absence of *dsx* products. From these observations, each of the *dsx* sex-specific products can be seen to have two activities: (1) When only one of the products is expressed, it represses differentiation of the opposite sex; and (2) when both *dsx* products are expressed, they inactivate each other. The reason for this mutual antagonism is not known.

The data in Figure 2 indicate that the antagonism between the activities of the *dsx* male and female products does not occur at the level of RNA synthesis or pro-

Table 2. *Sequence analysis of the* dsx *dominant lesions*

Allele	Proximal end	Distal end	Repeated sequence at site of insertion	Insert size (kb)
	A. *Sequence of the ends of the* dsx *dominant insertion elements*			
dsx^D	ACAAGTGTGTAACTTGGTGCTTAATATAATTA----------	----------TCTAAATCCCTAATAAGAAGACTTT	GCTTC	7.7
dsx^M	$(A)_{23}$TGAAACGAAAATAAATAACTTAGAGA----------	----------TCAGTAAGGTATCCGGGATTTCGCTGTAT	ATAA	3.9
dsx^T	GCCGAATTCCCTATAGTGAGTCGTAATCATGTC----------	----------AGAGTTCAATAAATATAAAAAAATT$(A)_{25}$	TTTGCC	4.5
	B. *Sequence comparison of the* dsx^M *and* dsx^D *insertions with known transposable elements*			
dsx^M	$(A)_{23}$TGAAACGAAAATAAATAACTTAGAGAGAACATATTCTTGATTGTTATTCAAATTTTTAGAATTGATATTGTTCCTA			
F element[a]	$(A)_{23}$TGAAACGAAAATAAATAACTTAGA - AGAACATATTCTTGATTGTTATTCAAATTTTTAGAATTGATATTGTTCATA			
dsx^D	ACAAGTGTGTACTTGTGCTTA - ATATAATTA----------	----------TTTTGAGTTGCTCATTTCAGAAGAATAATCCCTAAAA		
roo^b	ACAAGTGTGTACTTGTGCTTATATA - AATTT----------	----------TTTTGAGTTGCTCATTTCAGAAGAATAAACCCTAAAA		

[a]F-element sequence is from Di Nocera et al. (1983).
[b]From Scherer et al. (1982).

cessing. In chromosomally female flies carrying both a *dsx* dominant mutation with a wild-type *dsx* allele, both male- and female-specific *dsx* transcripts are expressed (Fig. 2, lanes 5–8); the male product from the *dsx* dominant allele and the female product from the dsx^+ gene. This indicates that the presence of the female *dsx* product does not repress the production of the male *dsx* RNA nor does the male *dsx* product inhibit female-specific *dsx* RNA processing.

Discussion

In this study we examined a unique class of *dsx* mutations, the *dsx* dominant alleles, which alter two aspects of *dsx* gene expression: (1) They eliminate the female-specific *dsx* function, and (2) they cause constitutive expression of the male-specific *dsx* function. Both of these mutant phenotypes are the consequences of alterations in the pattern of *dsx* RNA processing such that only male-specific splicing and polyadenylation occurs in both chromosomally male and female flies. To explain these results, we propose the following mechanism for the action of the *dsx* dominant mutations and the potential role of the *tra* and *tra-2* genes.

The male-specific splicing pattern is a default state for *dsx* RNA processing, it occurs in the absence of sex-specific regulatory functions requiring only the general housekeeping splicing machinery of the cell (Baker et al. 1987; Nagoshi et al. 1988). In the absence of sex-specific regulation, the male splice acceptor is preferentially used over the female acceptor, perhaps reflecting the closer similarity of the male splice acceptor to the *Drosophila* consensus splice acceptor sequence as compared to that shown by the female splice acceptor (Burtis and Baker 1989). We propose that female-specific *trans*-acting regulatory factors, particularly the *tra* and *tra-2* gene products, act to overcome this preferential usage of the male acceptor, thereby allowing the female splice to occur in chromosomally female flies. The *dsx* dominant mutations prevent this regulation from occurring, perhaps by disrupting regions in the *dsx* RNA required for binding to the *tra* and *tra-2* products. Without the intervention of these *trans*-acting factors, the *dsx* primary transcript is processed in the default pattern, resulting in male-specific splicing and polyadenylation. This mechanism explains how a single lesion in the *dsx* gene, the *dsx* dominant mutations, can both prevent female-specific RNA processing and cause constitutive production of the male *dsx* transcripts.

On the basis of the above model, there are two possible ways by which the regulation of female-specific splicing could occur. *Trans*-acting regulators, such as *tra* and *tra-2*, could act in a positive fashion to promote the usage of the nonconsensus female-specific splice site. Alternatively, the female-specific regulators could act negatively by preventing the usage of the default male splice acceptor. The proximity of the dominant lesions to the female splice acceptor is most consistent with *tra* and *tra-2* promoting the usage of the female splice site. Furthermore, the locations of the dominant lesions implicate exon sequences within 200 bp 3′ of the female-specific splice acceptor to be required for the regulation of female-specific splicing.

There is precedent for the involvement of exon sequences in splice acceptor site selection (Somasekhar and Mertz 1985; Reed and Maniatis 1986). Reed and Maniatis (1986) demonstrated with in vitro studies that if two identical splice acceptors are placed in tandem array downstream from a splice donor, the nearest splice acceptor is preferentially used. However, this preference can be shifted to the more distant splice acceptor by altering exon sequences immediately 3′ of the first splice acceptor. In the human fibronectin gene, sequences within a differentially spliced exon were found to be required in vivo for regulated RNA splicing (Barone et al. 1989). The deletion or reversal of orientation of an 81-bp

region in the middle of this exon prevented the usage of the adjacent splice acceptor site. A similar phenomenon was described for the human leukocyte common antigen (LCA) RNA in in vivo experiments (Streuli and Saito 1989). The LCA pre-mRNA undergoes cell-type-specific usage of splice acceptor sites, the regulation of which is dependent on the physical integrity of several discrete regions in the differentially spliced exon.

Reed and Maniatis (1986) suggested that exon sequences could participate directly in splice site recognition by determining the affinity of adjacent splice sites for splicing components. This could occur through RNA secondary structure, perhaps by making the splice site conformationally more accessible, or alternatively, the exon sequences themselves might bind directly to factors involved in splice site selection. The latter possibility seems likely in the current case, because of the finding that two genes that may directly control sex-specific RNA processing at *dsx*, *tra* (Belote et al. 1989), and *tra-2* (Amrein et al. 1988; Goralski et al. 1989) encode for polypeptides with structural similarities to known RNA-binding proteins.

An alternative possibility is that the spacing between the sex-specific splice acceptors and polyadenylation sites is critical for sex-specific regulation. There could be, for example, a functional interaction between female-specific splicing and polyadenylation such that variations in the spacing of the splice acceptor from the polyadenylation site would have dramatic effects on splice site selection. Alternatively, there may be competition between the usage of the male and female splice acceptors (or polyadenylation sites) whose resolution is, in part, dependent on the physical distance between the sites. A potential example of this type of mechanism is the IgM heavy-chain gene which, like *dsx*, displays both alternative splice acceptor and alternative poly(A) site selection (Alt et al. 1980; Early et al. 1980; Rogers et al. 1980). Poly(A) site selection was reported to be dependent on the spacing between the alternative poly(A) sites (Galli et al. 1987, 1988). Another study indicated that changes in the relative locations of the differential splice donor to the poly(A) sites can affect RNA splice site selection (Peterson and Perry 1986, 1989). The occurrence of a similar mechanism for the regulation of *dsx* seems unlikely in view of the fact that either insertions of as much as 7.8 kb (dsx^D) or a deletion of 0.5 kb (dsx^S) within the region that separates the female splice acceptor from both the male splice acceptor and female polyadenylation site result in identical patterns of *dsx* pre-mRNA splicing.

In summary, we have identified *cis*-acting sequences in the *dsx* gene that have profound effects on the regulation of sex-specific splicing of the *dsx* RNA. These are likely to be sites of interaction with the *trans*-acting regulators of *dsx* RNA processing; two likely candidates for such factors are the products of the *tra* and *tra-2* genes. The delineation of how these *cis*- and *trans*-acting factors interact to regulate sex-specific splicing at *dsx* should have general relevance toward understanding how RNA processing is regulated.

Methods

Fly strains

The genetic phenotypes of the dsx^D (Duncan and Kaufman 1975) and dsx^M (Baker and Ridge 1980; Nothiger et al. 1980) alleles have been described previously. The dsx^T allele was isolated in W. Gehring's laboratory in Basel. All three are spontaneous in origin. The dsx^S allele was induced by X-ray mutagenesis in an inbred background for which the parental chromosome is available (K. Brandt-Rosquist, pers. comm., as cited in Baker and Wolfner 1988). The deletion, dsx^{M+R15}, removes the entire *dsx* gene and is described in B. Baker et al. (in prep.). Flies were raised on a standard cornmeal, molasses, yeast, agar medium, containing propionic acid as a mold inhibitor and supplemented with live yeast.

RNA preparation

RNAs for Northern blot analyses were obtained as follows: Adult flies (10–50) were homogenized in a Dounce homogenizer in 10 ml of 1 : 1 homogenization buffer/phenol [homogenization buffer: 0.15 M NaCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl (pH 8.0), 0.5% SDS]. The supernatant was extracted twice with 1 : 1 phenol/chloroform and ethanol-precipitated. The nucleic acid precipitate was resuspended in 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA and the concentration determined by OD_{260}.

RNA blotting and hybridization

Northern analysis was done as described in Nagoshi et al. (1988). Total RNA (10–20 μg per lane) was loaded in a formaldehyde gel and blotted to Nytran (Schleicher & Schuell), as described in Maniatis et al. (1982). Single-stranded DNA probes were derived from *dsx* genomic and cDNA sequences subcloned into the M13 vector, using primed synthesis from a 17-mer M13 primer. Hybridizations were done in 5× SSPE (or 5× SSC), 5× Denhardt's solution, 50 μg/ml salmon sperm DNA, and 0.2% SDS at 42°C.

DNA sequencing

Genomic DNA fragments were subcloned into either M13 (Yanisch-Perron et al. 1985) or Bluescript (Strategene) vectors. Single- and double-stranded templates were prepared by standard methods and sequenced by the method of Sanger et al. (1977). Computer resources for sequences were provided by the Bionet National Computer Resources for Molecular Biology, which is funded by the Biomedical Research Technology Program, Division of Research Resources, National Institutes of Health.

Acknowledgments

We thank W. Gehring for the dsx^T stock. B. Brandt-Rosquist for dsx^S and parental stocks, G. Bohm for preparation of *Drosophila* and bacterial media, and B. Oliver, M. Kuroda, T. Goralski, L. Ryner, M. Gorman, and W. Mattox for helpful comments and discussion on the manuscript. This work was supported by grants from the National Institutes of Health to R.N.N. and B.S.B.

References

Adami, G. and J.R. Nevins. 1988. Splice site selection dominates over poly(A) site choice in RNA production from com-

plex adenovirus transcription units. *EMBO J.* **7:** 2107–2116.

Alt, F.W., A.L.M. Bothwell, M. Knapp, E. Siden, E. Mather, M. Koshland, and D. Baltimore. 1980. Synthesis of secreted and membrane-bound immunoglobulin Mu heavy chains as directed by mRNAs that differ at their 3′ ends. *Cell* **20:** 293–301.

Amara, G.G., V. Jonas, M.G. Rosenfeld, E.S. Ong, and R.M. Evans. 1982. Alternative RNA processing in calcitonin gene expression generates mRNAs encoding different polypeptide products. *Nature* **298:** 240–244.

Amrein, H. M. Gorman, and R. Nothiger. 1988. The sex-determining gene tra-2 of Drosophila encodes a putative RNA binding protein. *Cell* **55:** 1025–1035.

Baker, B.S., and K.A. Ridge. 1980. Sex and the single cell: I. On the action of major loci affecting sex determination in *D. melanogaster*. *Genetics* **94:** 383–423.

Baker, B.S. R.N. Nagoshi, and K.C. Burtis. 1987. Molecular genetic aspects of sex determination of Drosophila. *Bioessays* **6:** 66–70.

Baker, B.S. and M.F. Wolfner. 1988. A molecular analysis of *doublesex*, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. *Genes Dev.* **2:** 477–489.

Baker, B.S. 1989. Sex in flies: The splice of life. *Nature* **340:** 521–524.

Barone, M.V. C. Henchcliffe, F.E. Baralle, and G. Paolella. 1989. Cell type specific trans-acting factors are involved in alternative splicing of human fibronectin pre-mRNA. *EMBO J.* **8:** 1079–1085.

Bell, L.R., E.M. Maine, P. Schedl, and T.W. Cline. 1988. Sex-lethal, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* **55:** 1037–1046.

Belote, J.M., M. McKeown, R.T. Boggs, R. Ohkawa, and B.A. Sosnowski. 1989. Molecular genetics of *transformer*, a genetic switch controlling sexual differentiation in *Drosophila*. *Dev. Genet.* **10:** 143–154.

Bingham, P.M., T-B. Chou, I. Mims, and Z. Zachar. 1988. On/off regulation of gene expression at the level of splicing. *Trends Genet.* **4:** 134–138.

Boggs, R.T., P. Gregor, S. Idriss, J.M. Belote, and M. McKeown. 1987. Regulation of sexual differentiation in *D. melanogaster* via alternative processing of RNA from the transformer gene. *Cell* **50:** 739–747.

Breitbart, R.E., A. Andreadis, and B. Nadal-Ginard. 1987. Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes. *Annu. Rev. Biochem.* **56:** 467–495.

Burtis, K.C. and B.S. Baker. 1989. *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* **56:** 997–1010.

Chou, T.-B., Z. Zachar, and P.M. Bingham. 1987. Developmental expression of a regulatory gene is programmed at the level of splicing. *Cell* **6:** 4095–4104.

Di Nocera, P.P., M.E. Digan, and I.B. Dawid. 1983. A family of oligo-adenylate-terminated transposable sequences in *Drosophila melanogaster*. *J. Mol. Bol.* **168:** 715–727.

Di Nocera, P.P. 1988. Close relationship between non-viral retroposons in *Drosophila melanogaster*. *Nucleic Acids Res.* **16:** 4041–4052.

Duncan, I., and T. Kaufman. 1975. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: Mapping of the proximal portions of the right arms. *Genetics* **80:** 733–752.

Early, P., J. Rogers, M. Davis, K. Calame, M. Bond, R. Wall, and L. Hood. 1980. Two mRNAs can be produced from a single immunoglobulin gene by alternative RNA processing pathways. *Cell* **20:** 313–319

Fu, X.-Y, J.D. Colgan, and J.L. Manley. 1988. Multiple cis-acting sequence elements are required for efficient splicing of simian virus 40 small-t antigen pre-mRNA. *Mol. Cell. Biol.* **8:** 3582–3590.

Galli, G., J. Guise, P.W. Tucker, and J.R. Nevins. 1988. Poly(A) site choice rather than splice site choice governs the regulated production of IgM heavy-chain RNAs. *Proc. Natl. Acad. Sci.* **85:** 2439–2443.

Galli, G., J. Guise, M.A. McDevitt,, P.W. Tucker, and J.R. Nevins. 1987. Relative position and strengths of poly(A) sites as well as transcription termination are critical to membrane versus secreted μ-chain expression during B-cell development. *Genes Dev.* **1:** 471–481.

Garfinkel, M.D., R.E. Pruitt, and E.M. Meyerowitz. 1983. DNA sequences, gene regulation and modular protein evolution in the Drosophila 68C glue gene cluster. *J. Mol. Biol.* **168:** 765–789.

Goralski, T.J., J-E. Edström, and B.S. Baker. 1989. The sex determination locus transformer of Drosophila encodes a polypeptide with similarity to RNA binding proteins. *Cell* **56:** 1011–1018.

Green, M.R. 1986. Pre-mRNA splicing. *Annu. Rev. Genet.* **20:** 671–708.

Hampson, R.K., L. La Follette, and R.M. Rottman. 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. *Mol. Cell. Biol.* **9:** 1604–1610.

Hodgkins, J. 1989. Drosophila sex determination: A cascade of regulated splicing. *Cell* **56:** 905–906.

Kornblihtt, A.R., K. Umezawa, and F.E. Baralle. 1985. Primary structure of human fibronectin: Differential splicing may generate at least 10 polypeptides from a single gene. *EMBO J.* **4:** 1755–1759.

Laski, F.A. and G.M. Rubin. 1989. Analysis of the *cis*-acting requirements for germ-line-specific splicing of the P-element ORF2–ORF3 intron. *Genes Dev.* **3:** 720–728.

Leff, S.E., R.M. Evans, and M.G. Rosenfeld. 1987. Splice commitment dictates neuron-specific alternative RNA processing in calcitonin/CGRP gene expression. *Cell* **48:** 517–524.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: A laboratory manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Manley, J.L. 1988. Polyadenylation of mRNA precursors. *Biochim. Biophys. Acta.* **950:** 1-12.

Mardon, H.J., G. Sebastio, and F.E. Baralle. 1987. A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Res.* **15:** 7725–7733.

Mattox, W., M.J. Palmer, and B.S. Baker. 1990. Alternative splicing of the sex determination gene *tra-2* is sex-specific in the germ-line but not in the soma. *Cell* (in prep.).

Nagoshi, R.N., M. McKeown, K.C. Burtis, J.M. Belote, and B.S. Baker. 1988. The control of alternative splicing at genes regulating sexual differentiation in D. melanogaster. *Cell* **53:** 229–236.

Nevins, J.R. and S. Chen-Kiang. 1981. Processing of adenovirus nuclear RNA to mRNA. *Adv. Virus Res.* **26:** 1–35.

Noble, J.C.S., Z-Q. Pan, C. Prives, and J.L. Manley. 1987. Splicing of SV40 early pre-mRNA to large T and small t mRNAs utilizes different patterns of lariat branch sites. *Cell* **50:** 227–236.

Nothiger, R., M. Roost, and T. Schubach. 1980. *'Masculinizer'* is an allele of *'doublesex'*. *Dros. Inf. Serv.* **55:** 118.

Nothiger, R. M. Leuthold, N. Andersen, P. Gerschwiler, A.

Gruter, W. Keller, C. Leist, M. Roost, and H. Schmid. 1987. Genetic and developmental analysis of the sex-determining gene *doublesex (dsx)* of *Drosophila melanogaster*. *Genet. Res.* **50:** 113–123.

Padgett, R.A., P.J. Grabowski, M.M. Konarska, S. Seiler, and P.A. Sharp. 1986. Splicing of messenger RNA precursors. *Annu. Rev. Biochem.* **55:** 1119–1150.

Peterson, M.L. and R.P. Perry. 1986. Regulated production of μ_m and μ_s mRNA requires linkage of the poly(A) addition sites and is dependent on the length of the μ_s–μ_m intron. *Proc. Natl. Acad. Sci.* **83:** 8883–8887.

———. 1989. The regulated production of μ_s and μ_m mRNA is dependent on the relative efficiencies of μ_s poly(A) site usage and the CM4- to M1-splice. *Mol. Cell. Biol.* **9:** 726–738.

Pittler, S.J. and R.L. Davis. 1987. A new family of the polydeoxyadenylated class of *Drosophila* transposable elements identified by a representative member at the *dunce* locus. *Mol. Gen. Genet.* **208:** 325–328.

Query, C.C., R.C. Bentley, and J.D. Keene. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell* **57:** 89–101.

Reed, R. and T. Maniatis. 1986. A role for exon sequences and splice site proximity in splice-site selection. *Cell* **46:** 681–690.

Rogers, J., P. Early, C. Carter, K. Calame, M. Bond, L. Hook., and R. Wall. 1980. Two mRNAs with different 3′ ends encode membrane-bound and secreted forms of immunoglobulin μ chain. *Cell* **20:** 303–312.

Sabate, M.I., L.S. Stolarsky, J.M. Polak, S.R. Bloom, I.M. Varndell, M.A. Ghatei, R.M. Evans, and M.G. Rosenfeld. 1985. Regulation of neuroendocrine gene expression by alternative RNA processing. *J. Biol. Chem.* **260:** 2589–2592.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Scharzbauer, J.E., R.S. Patel, D. Fonda, and R.O. Hynes. 1987. Multiple sites of alternative splicing of the rat fibronectin gene transcript. *EMBO J.* **6:** 2573–2580.

Scherer, G., C. Tschudi, J. Perera, H. Delius, and V. Pirrotta. 1982. B104, a new dispersed gene family in *Drosophila melanogaster* and its analogies with retroviruses. *J. Mol. Biol.* **157:** 435–451.

Sharp. P.A. 1987a. Splicing of messenger RNA precursors. In 'The Harvey Lectures,' Series 81, pp. 1–31. Alan R. Liss, New York.

———. 1987b. Splicing of messenger RNA precursors. *Science* **235:** 766–771.

Somasekhar, M.B. and J.E. Mertz. 1985. Exon mutations that affect the choice of splice sites used in processing SV40 late transcripts. *Nucleic Acids Res.* **13:** 5591–5609.

Streuli, M. and H. Saito. 1989. Regulation of tissue-specific alternative splicing: Exon-specific cis-elements govern the splicing of leukocyte common antigen pre-mRNA. *EMBO J.* **8:** 787–796.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33:** 103–119.