Regulation of the Gene *Sex-lethal*: A Comparative Analysis of *Drosophila melanogaster* and *Drosophila subobscura*

Luiz O. F. Penalva,* Hiroshi Sakamoto,† Aurea Navarro-Sabaté,‡ Eiji Sakashita,† Begoña Granadino* Carmen Segarra‡ and Lucas Sánchez*

**Centro de Investigaciones Biológicas (C.S.I.C.), 28006 Madrid, Spain; †Department of Biology, Faculty of Science, Kobe University, Kobe 657, Japan; and ‡Departamento de Genética. Facultad de Biología, Universidad de Barcelona, 08071 Barcelona, Spain*

Manuscript received June 7, 1996
Accepted for publication September 13, 1996

ABSTRACT

The Drosophila gene *Sex-lethal* (*Sxl*) controls the processes of sex determination and dosage compensation. A *Drosophila subobscura* genomic fragment containing all the exons and the late and early promotors in the *Sxl* gene of *D. melanogaster* was isolated. Early *Sxl* expression in *D. subobscura* seems to be controlled at the transcriptional level, possibly by the *X*:A signal. In the region upstream of the early *Sxl* transcription initiation site are two conserved regions suggested to be involved in the early activation of *Sxl*. Late *Sxl* expression in *D. subobscura* produces four transcripts in adult females and males. In males, the transcripts have an additional exon which contains three translational stop codons so that a truncated, presumably nonfunctional Sxl protein is produced. The *Sxl* pre-mRNA of *D. subobscura* lacks the poly-U sequence presented at the polypirimidine tract of the 3′ splice site of the male-specific exon present in *D. melanogaster*. Introns 2 and 3 contain the Sxl-binding poly-U stretches, whose localization in intron 2 varies but in intron 3 is conserved. The Sxl protein is fully conserved at the amino acid level in both species.

IN *Drosophila melanogaster*, the gene *Sex-lethal* (*Sxl*) controls the processes of sex determination and dosage compensation (the products of the *X*-linked genes are present in equal amounts in females and males). This gene regulates the expression of two independent sets of genes (LUCCHESI and SKRIPSKY 1981). One set is formed by the sex determination genes; mutations in these genes affect sex determination while having no effect on dosage compensation (BAKER and WOLFNER 1988; STEINMANN-ZWICKY *et al.* 1990). The other set is formed by the *male-specific lethal* genes (*msl*'s); mutations in these genes affect dosage compensation while having no effect on sex determination (BAKER *et al.* 1994).

The gene *Sxl* produces two temporally distinct sets of transcripts corresponding to the function of the female-specific early and the non-sex-specific late promoters, respectively (SALZ *et al.* 1989). The early set is produced as a response to the *X*:A signal which controls *Sxl* expression at the transcriptional level (TORRES and SÁNCHEZ 1991; KEYES *et al.* 1992). Once the state of activity of *Sxl* is determined, an event that occurs at the blastoderm stage, the *X*:A signal is no longer needed and the activity of *Sxl* remains fixed (SÁNCHEZ and NÖTHIGER 1983; BACHILLER and SÁNCHEZ 1991). The identification of a set of genes involved in the initial step of *Sxl* activation indicates that a conventional genetic system is the basis of this signal (PARKHURST and ISH-HOROWICZ 1992; CLINE 1993; SÁNCHEZ *et al.* 1994). It has been shown in *D. melanogaster* that a fragment of 1400 bp upstream of the early *Sxl* transcription initiation site contains all the *cis*-acting elements responsible for interaction with the *X*:A signal (ESTES *et al.* 1995). This region contains E boxes that are the binding sites for the Scute-Daughterless (Sc-Da) heterodimers that act as activators of *Sxl*. It also contains one D box that acts as a binding site for the Dpn protein (HOSHIJIMA *et al.* 1995), another element of the *X*:A signal that acts as a repressor of *Sxl* (YOUNGER-SHEPHERD *et al.* 1992).

The late set of *Sxl* transcripts is formed by three male-specific and three female-specific transcripts. These appear slightly after blastoderm stage and persist throughout development. The male transcripts differ from the female by the inclusion of a male-specific exon that places a stop codon in the open reading frame. This gives rise to truncated, presumably nonfunctional, proteins. In females, this exon is spliced out and functional proteins are produced (BELL *et al.* 1988; BOPP *et al.* 1991). Therefore, the control of *Sxl* expression throughout development occurs by sex-specific splicing of its transcript. The capacity of *Sxl* to function as a stable switch is due to the positive autoregulatory function of its own product (CLINE 1984; BELL *et al.* 1991). Three functional domains in the *D. melanogaster* Sxl protein have been identified: two domains that interact with RNA (RNP domain) (BELL *et al.* 1988) plus a domain which is required for its own cooperativity (GN domain) (WANG and BELL 1994). The Sxl protein regulates its own RNA splicing by binding to poly-U stretches in introns 2 and 3, adjacent to the male-specific exon L3 (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993b).

Corresponding author: Lucas Sánchez, Centro de Investigaciones Biológicas, Velázques 144, 28006 Madrid, Spain.
E-mail: cibsrób@pinar2.csic.es

So far, early and late *Sxl* regulation have been studied in *in vitro* and *in vivo* experimental systems (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993b; WANG and BELL 1994; ESTES *et al.* 1995; HOSHIJIMA *et al.* 1995). A complementary approach to the study of *Sxl* regulation was undertaken in this investigation, namely its interspecific comparison. This work reports the cloning and characterization of the gene *Sxl* of *D. subobscura* and its comparison with the gene *Sxl* of *D. melanogaster*. These two species appear to have diverged ~30 mya (THROCKMORTON 1975). This evolutionary period is sufficiently distant for unconstrained sequences to have diverged extensively. The comparative analysis of the gene *Sxl* from both species should therefore allow the identification of putative *cis*-acting elements by sequence conservation (which might participate in the regulation of this gene) as well as the identification of Sxl protein functional domains.

MATERIALS AND METHODS

Fly strains: Flies were cultured on standard food at 25° (*D. melanogaster*) or 18° (*D. subobscura*).

Construction of a genomic library from *D. subobscura*: This was performed according to PIRROTTA (1986).

Cloning of the gene *Sxl* of *D. subobscura*: The *D. subobscura* genomic library was screened with a full-length female *Sxl* cDNA of *D. melanogaster* (SAMUELS *et al.* 1991). Hybridization conditions were: 63° for 18–20 hr, 5× SSC, 1% SDS, 2% powder milk and 100 mg/ml of salmon sperm DNA. Washes were repeated three times (20 min each) at 63° in 0.1× SSC and 0.1% SDS. Identification of positive clones, plaque purification, preparation of phage DNA, Southern blot analysis, identification of cross-hybridization fragments, subcloning of the restriction fragments into the plasmid pBluescript KS^+ and isolation of plasmid DNA were performed using the protocols described by MANIATIS *et al.* (1982).

***In situ* hybridization to polytene chromosomes:** The *ch cu* strain of *D. subobscura* (with the standard *X* chromosome arrangement) and the *Canton-S* strain of *D. melanogaster* were used to prepare salivary gland polytene chromosomes. Female third-instar larvae were chosen from uncrowded cultures raised at 17°. Conditions for larva dissection and polytene chromosome squashes were as described by MONTGOMERY *et al.* (1987). *D. subobscura* DNA was labeled by nick-translation using 16-biotin dUTP (Boehringer). Prehybridization, hybridization and detection conditions were as described by SEGARRA and AGUADÉ (1993). The location of the probes was determined using the cytological map of *D. subobscura* (KÜNZE-MÜHL and MÜLLER 1958) and the cytological and photographic maps of *D. melanogaster* produced by LEFEVRE (1976).

***In situ* hybridization to embryos of *D. subobscura*:** This was performed following the procedure described by TAUTZ and PFEIFLE (1989) with the modifications suggested by CUBAS *et al.* (1991). The probe was marked with digoxigenine (Boehringer) following the manufacturer's instructions. The stage of the embryos was identified according to CAMPOS-ORTEGA and HARTENSTEIN (1985).

Transcripts analysis: Poly-A^+ RNA preparation from adult males and females was performed following the procedure described in the Pharmacia RNA purification kit. Northern blots of poly-A^+ RNA were hybridized with different genomic fragments, previously subcloned in the plasmid pBluescript KS^+. These genomic fragments are shown in Figure 1.

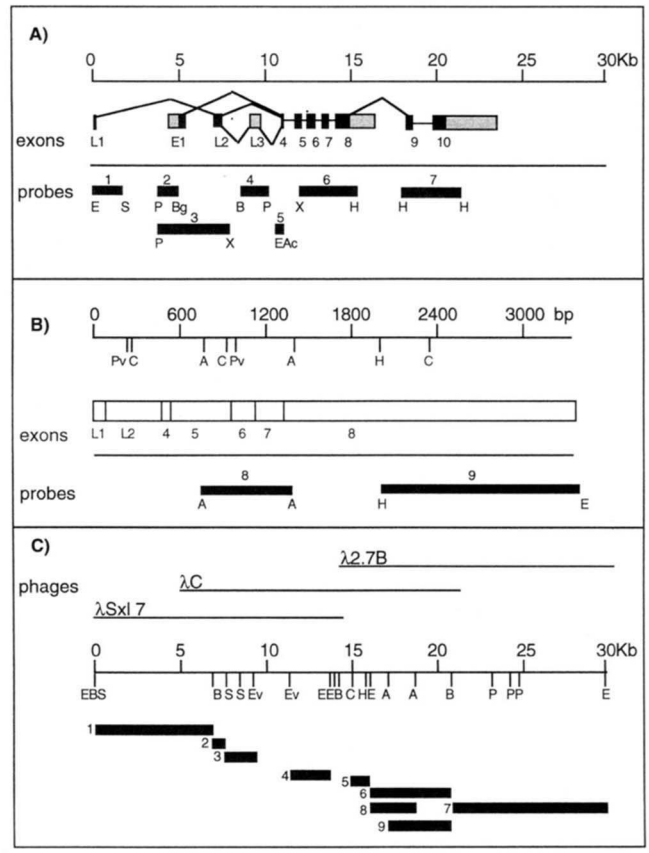

FIGURE 1.—The gene *Sxl* of *D. subobscura*. (A) Molecular map of the *D. melanogaster Sxl* gene and schematic representation indicating the splicing pattern of the early and late *Sxl* pre-mRNAs in females and males, as well as the position of the exons corresponding to the early and late *Sxl* RNAs (SAMUELS *et al.* 1991). E1 refers to the exon E1 that is specific to the early *Sxl* transcripts. L1, L2 and L3 refer to the exons 1, 2 and 3, respectively, that are specific for the late *Sxl* transcripts. L3 is the male-specific exon. (B) Molecular map of the full-length late female *Sxl* cDNA of *D. melanogaster* (SAMUELS *et al.* 1991), used for screening the *D. subobscura* genomic library, and schematic representation of the exons. The black horizontal bars in A and B, numbered from 1 to 9, indicate the position of the probes used for mapping the positions of the homologous exons in *D. subobscura*. Some of these *D. melanogaster* genomic fragments contain exons 9 and 10 or the germ line specific RNA. These are not present in the cDNA probe and correspond to the 3′ end of the transcription unit. Other *D. melanogaster* fragments contain sequences of the late exon L1, as well as sequences downstream and upstream of this exon. It also contains sequences upstream of the late promotor that constitute the 5′ end of the transcription unit. (C) Molecular map of the gene *Sxl* of *D. subobscura* and schematic representation of the three overlapping recombinant phages λ2.7B, λC and λSxl 7 isolated from the *D. subobscura* genomic library. The black horizontal bars indicate the position of the *D. subobscura Sxl* genomic fragments that hybridize with the *D. melanogaster* probes of A and B numbered from 1 to 9. A, *Ava*I; B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI, Ev, *Eco*RV; H, *Hind*III; P, *Pst*I; Pv, *Pvu*II, S, *Sal*I and X, *Xho*I.

RT-PCR analysis: To verify the existence of an additional exon in the male *Sxl* transcripts of *D. subobscura*, 10 μg of total RNA from both adult males and females were reverse-transcribed with AMV Promega. Amplifications were per-

formed by PCR following the protocol of FROHMAN *et al.* (1988). The primers were: 5′ GAAAGAGGAAGAGCACGCAG 3′ and 5′ TGTGCCTCCTCCCGTTTGTT 3′. Two amplification rounds were performed. The conditions of the first round were: 95° for 3 min, 62° for 2 min, 72° for 40 min—one cycle; 95° for 45 sec, 62° for 2 min, 72° for 1.5 min—25 cycles. Ten percent of the amplified product was amplified again in the second round of 15 cycles: 95° for 45 sec, 62° for 2 min, 72° for 1.5 min each cycle. RT-PCR products were analyzed by electrophoresis in 1.5% agarose, blotted into nylon membranes and hybridized with the *D. subobscura* genomic fragment homologous to the *Sxl* male-specific exon L3. The membranes were boiled to eliminate the previous probe and then hybridized with the *D. subobscura* genomic fragment homologous to the *D. melanogaster Sxl* genomic fragment containing exons 5, 6, 7 and part of 8, common to both sexes.

The RT-PCR amplified fragment from total RNA of adult females was cloned in the plasmid pBluescript KS$^+$, in the *Eco*RV site, and sequenced following the procedure mentioned below.

DNA sequencing: pBluescript KS$^+$ clones with small genomic subfragments were directly sequenced from both sides using the universal primers of the vector polylinker. Those pBluescript clones with larger genomic subfragments were sequenced after obtaining a set of nested deletions according to (HENIKOFF 1984). Sequencing was carried out by the dideoxynucleotide method (SANGER *et al.* 1977) using modified T7 polymerase from Pharmacia and 10 μCi of ^{35}S-dATP as a labeling precursor. Electrophoresis was performed with a LKB-Pharmacia Macrophor chamber using 6% acrylamide and 7.5 M urea, following the manufacturer's instructions. Other fragments were sequenced using an automatic 377 DNA sequencer (Applied Biosystem). The nucleotide sequence of the *D. melanogaster* genomic region spanning from 665 bp upstream of exon 2 to just downstream of exon 4 was determined by the dideoxysequencing method using Sequenase version 2.0 (U.S. Biochemicals). For this purpose, several lambda dash phage clones were isolated from a *D. melanogaster* genomic library. In all cases, the complete sequences of both strands of the cloned subfragments were determined. The sequence corresponding to the region spanning from ~4 kb upstream of exon 2 to exon 7 of the *Sxl* gene of *D. melanogaster* has been deposited with the database DDBJ (accession no. D84425). The sequence of the corresponding region, as well as the sequence of the female *Sxl* cDNA, of *D. subobscura* have been deposited with the database EMBL (accession no. X98370 and X98371)

Comparison of DNA sequences: Comparison of the DNA sequences of *D. melanogaster* and *D. subobscura* was performed using version 1.3 of the Ifasta program (PEARSON and LIPMAN 1988).

RESULTS

Isolation and molecular organization of the gene *Sxl* of *D. subobscura*: A genomic library of *D. subobscura* in λEMBL4 was synthesized (see MATERIALS AND METHODS) and screened with a full-length female *Sxl* cDNA of *D. melanogaster*. Three overlapping positive phages were isolated (λSxl 7, λC and λ2.7B) that together encompassed 26 kb of genomic DNA (Figure 1C).

The Sxl protein contains two RNA-binding domains (RNPs). It is possible, then, that the three isolated phages did not contain sequences of the *Sxl* homologous gene but rather sequences of another gene that codes for a protein which also contains this RNP motif. To provide supporting evidence that the isolated phages correspond to the *Sxl* homologous gene of *D. subobscura*, *in situ* hybridization of the three clones to salivary gland polytene chromosomes of *D. melanogaster* was performed. The three clones exclusively hybridized in cytogenetic band 6F where the gene *Sxl* of *D. melanogaster* maps (data not shown). These results indicate that the isolated phages contained sequences corresponding to the *Sxl* homologous gene of *D. subobscura*. *In situ* hybridization of the three phages to salivary gland polytene chromosomes of *D. subobscura* was also performed. These phages hybridized in cytogenetic band 10B of chromosome *A* (data not shown), which is homologous to the *X* chromosome of *D. melanogaster*.

In order to compare the organization between the gene *Sxl* of *D. melanogaster* and its homologue in *D. subobscura*, a set of Southern blots were performed with subclones from the three genomic phages λSxl 7, λC and λ2.7B. As probes, nine different fragments of the gene *Sxl* of *D. melanogaster* were used, which together contained the complete transcription unit of the *D. melanogaster Sxl* gene (Figure 1, A and B). All of the *D. melanogaster* probes hybridized with the 26-kb fragment of *D. subobscura* (Figure 1C). These results indicate that the whole transcription unit of the gene *Sxl* of *D. melanogaster* was contained within this 26 Kb fragment; the relative order of exons-introns being maintained. Thus, it is suggested that the whole transcription unit of the gene *Sxl* of *D. subobscura* was cloned.

Early expression of the gene *Sxl* of *D. subobscura*: To determine if the early regulation of *Sxl* in *D. subobscura* is like that in *D. melanogaster*, we probed embryos with a *D. subobscura* 0.2-kb genomic fragment that is homologous to exon E1 of the early *Sxl* transcripts of *D. melanogaster*. The only embryos that hybridized to this probe were at the syncytial blastoderm and cellular blastoderm stages (Figure 2). Hybridization to embryos at developmental stages preceding the syncytial blastoderm stage indicate the presence of maternal *Sxl* transcripts. Among 150 blastoderm embryos, 76 hybridized and 74 did not. These results indicate that early *Sxl* expression in *D. subobscura* takes place at the same developmental stages as in *D. melanogaster* and suggest that this early expression is regulated at the transcriptional level by the *X*:A signal. Embryos that showed positive hybridization would be females, where the *X*:A signal activates *Sxl*, and embryos showing no hybridization would be males. Moreover, as in *D. melanogaster*, no hybridization was observed in the pole cells, the precursors of the germ line.

The early *Sxl* transcripts of *D. melanogaster* contain the exon E1, which is not present in the late *Sxl* transcripts (see Figure 1A). We have sequenced and compared the exon E1 of *D. melanogaster* and *D. subobscura*. This exon shows a high degree of homology at both DNA (75%) and protein (80%) levels. This homology extends also to the 5′ UTR region (74.5% of homology), which con-

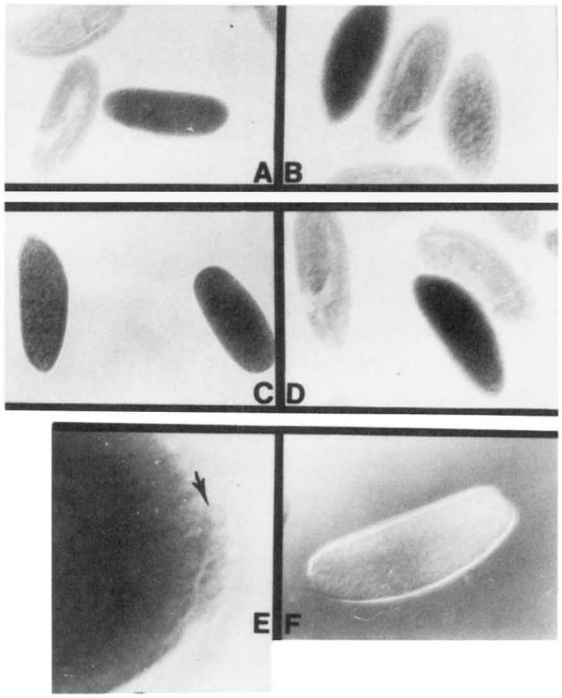

FIGURE 2.—Early expression of the gene *Sxl* of *D. subobscura*. *In situ* hybridization to *D. subobscura* embryos at different developmental stages. The probe used was a 0.2-kb genomic fragment of this species that contains the homologue to the early-specific exon E1 of *D. melanogaster*. (A–D) and (F) Examples shown of embryos at blastoderm or syncytial blastoderm stages that either hybridize (females) or do not hybridize (males) with the probe, plus embryos at postblatoderm stage that never hybridize. (E) Magnification is shown of the posterior region of an embryo at blastoderm stage that hybridizes but whose pole cells (marked with an arrow) do not show hybridization.

tains a sequence stretch (AAGTCCAACTT) that shows a high degree of homology with a sequence (AAGTGCAAGTT) found in the 5′ UTR of the *D. melanogaster* heat-shock genes (HOLMGREN *et al.* 1981). The sequence just 5′ of the translation initiation codon is not conserved in both species: CATT for *D. melanogaster* and AATG for *D. subobscura*. In both cases, the sequence shows low homology with the consensus sequence A/CAAA/C which is thought to be optimal for efficient translation initiation in *D. melanogaster* (CAVENER 1987). The protein domain encoded by this exon contains two more amino acids in *D. subobscura* than it does in *D. melanogaster*. The sequence of the 5′ splice site of this exon in *D. melanogaster* and in *D. subobscura* is AG GTATGT, which matches well with the consensus sequence AG GTRAGT. In both species, the highly conserved G residue at position +5 in the intron is conserved.

Sequencing of the genomic region located upstream of the early *Sxl* transcription initiation site. A comparative analysis of *D. melanogaster* and *D. subobscura*: A region of 1600 bp upstream of the early *Sxl* transcription initiation site was sequenced in *D. subobscura* and compared with the corresponding region of *D. melanogaster* (HOSHIJIMA *et al.* 1995; DDBJ accession no. D50435). Two conserved regions were found. Their location are shown in Figure 3. Region 1, which is ~700 bp upstream of the start site, contains three E boxes in *D. melanogaster* and four E boxes in *D. subobscura*. Region 2 which encompasses ~300 bp sequence near the transcription start site contains one E box and the D box, whose locations are conserved. In addition, the two regions contain other conserved sequences, surrounding the E and D boxes. Upstream of these two conserved regions, there were four E boxes in *D. subobscura* (−1000 to −1150 bp region) that are not present in the corresponding region of *D. melanogaster*.

Late expression of the gene *Sxl* of *D. subobscura*: To determine if the late regulation of *Sxl* is like that in *D. melanogaster*, we probed embryos with a *D. subobscura* 1.2-kb genomic fragment. This fragment contains the homologues to exons 5–8 of *D. melanogaster* that are common to the early and late *Sxl* transcripts. In this case, hybridization was observed in all of the embryos, independent of their developmental stages (data not shown). These results indicate that in *D. subobscura* the late *Sxl* transcription, common to both sexes, begins slightly later after the blastoderm stage, as in *D. melanogaster*.

The *Sxl* gene of *D. melanogaster* basically produces three transcripts in adult females (4.2, 3.3 and 1.9 kb) and three transcripts in adult males (4.4, 3.6 and 2.0 kb). There is another transcript of 3.3 kb that is expressed in the female germ line (BELL *et al.* 1988; SALZ *et al.* 1989). In order to characterize the transcripts from the gene *Sxl* of *D. subobscura*, Northern blots from both male and female adults were performed and subsequently hybridized with *D. subobscura* genomic fragments containing sequences homologous to exon L2 of *D. melanogaster*. Four transcripts of 5.3, 4.1, 3.5 and 2.7 kb were detected in adult females and four transcripts of 5.5, 4.3, 3.7 and 2.9 kb were detected in adult males (Figure 4A). Therefore, the *Sxl* transcripts of *D. subobscura* are larger than those of *D. melanogaster*. It is not known whether any of these transcripts correspond to the germ line specific RNA of *D. melanogaster*. The different size of these transcripts in both sexes parallels the situation in *D. melanogaster*, suggesting the presence of an additional exon in the male transcripts. This possibility was investigated by using RT-PCR to amplify cDNAs spanning the region between exons L2 and 8. The amplified cDNAs from males and females hybridized to the *D. subobscura* 1.2-kb fragment that contains sequences present in both sexes (Figure 4B). However, the male product was larger than the female and, unlike the female, it hybridized to the *D. subobscura* 1.8-kb probe (Figure 4C), which contains sequences homologous to the male-specific exon L3 of *D. melanogaster*. These results indicate that the *Sxl* transcripts of the males contained an additional exon and that the *Sxl* pre-mRNA follows a different splicing pattern in females and males. Therefore, in *D. subobscura* the late

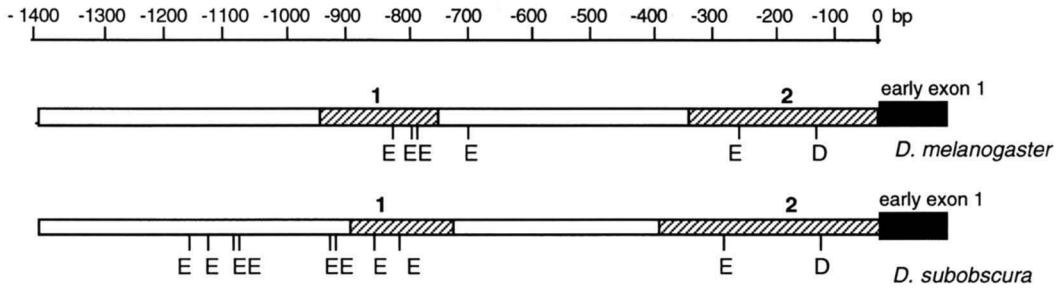

FIGURE 3.—Comparative analysis of the region upstream of the early *Sxl* promotor of *D. melanogaster* and its homologous region in *D. subobscura*. Schematic representation of the upstream region showing the two regions (cross-hatched) for which extensive sequence similarity between both species was found, as well as the locations of E and D boxes.

Sxl expression is controlled at the splicing level of its primary transcript; male transcripts containing an exon which is absent in female transcripts.

To test the specificity of the Sxl protein in *D. subobscura*, we performed inmunoblots of proteins from imaginal discs and brains of both male and female larvae with a monoclonal antibody that recognizes an epitope encoded by sequences downstream of the male-specific exon. In *D. melanogaster*, this antibody detects two Sxl proteins of 36 and 38 kD in females but not in males (BOPP *et al.* 1991). Two proteins of similar size were also detected in *D. subobscura* females but not in males (Figure 4D). These results would agree with the presence of translation stop codons in the male specific exon L3, so that in males a truncated Sxl protein would be manufactured.

The RT-PCR product of adult females containing the translated region of the *D. subobscura Sxl* gene was sequenced, as well as the genomic fragments containing the exons, and compared with the sequence corresponding to the female cDNA of *D. melanogaster* that is described in SAMUELS et al. (1991). The exon-intron structure and the two splicing acceptor sites of exon 5 are conserved (data not shown). At the nucleotide level, the average degree of homology of the coding region is very high, 86.1% in 1077 bp; 92.2% for the coding region of exon L2 (77 bp), 89.1% for exon 4 (46 bp), 83.8% for exon 5 (372 bp), 92.6% for exon 6 (108 bp), 88.4% for exon 7 (181 bp) and 83.3% for the coding region of exon 8 (293 bp).

The sequence GGAT just 5′ of the AUG translation initiation codon in exon L2 is conserved in both species and differs from the sequence found in early exon E1 (see above). The same GGAT sequence has been found just upstream of the translation start codon in exon L2 of *D. virilis* (BOPP *et al.* 1996). As in the case of exon E1, this sequence shows a poor match with the consensus sequence (A/CAAA/C) thought to be optimal for efficient translation in *D. melanogaster* (CAVENER 1987). The comparison of the predicted amino acid sequences of both Sxl proteins shows a very high degree of homology, 91.3% for 358 amino acids (Figure 5).

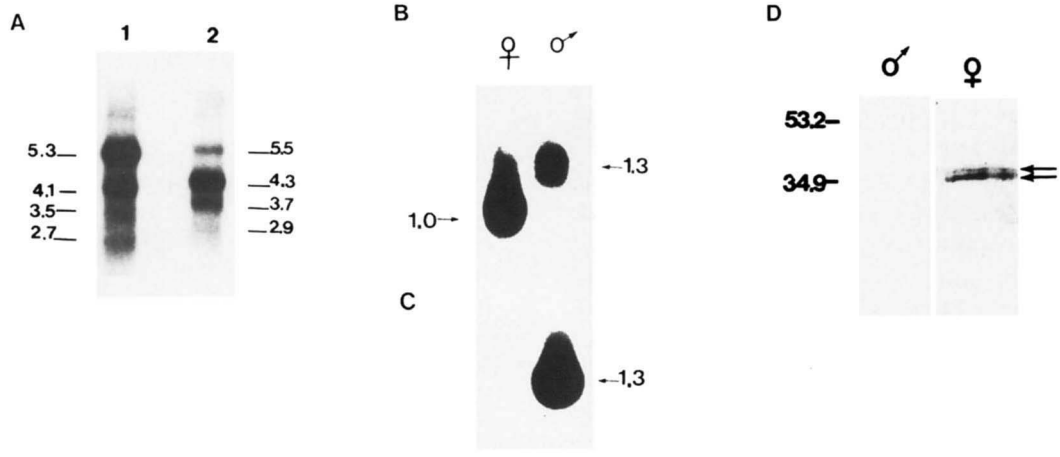

FIGURE 4.—Late expression of the gene *Sxl* of *D. subobscura*. (A) Northern blot of poly-A$^+$ RNA from adult females (lane 1) and males (lane 2). The probe used was the *D. subobscura* genomic fragment containing exon L2. (B) RT-PCR of poly-A$^+$ RNA of adult females and males. The probe used was the *D. subobscura* 1.2-kb genomic fragment that contains the exons 5–7 and part of exon 8 that are common to both sexes. The RT-PCR product had 1.3 kb in males and 1.0 kb in females. (C) RT-PCR of poly-A$^+$ RNA of adult females and males. The probe used was the *D. subobscura* 1.8-kb genomic fragment that contains the male-specific exon L3. (D) Analysis of the Sxl protein in *D. subobscura*. Immunoblots containing equivalent amounts of protein extracts from imaginal discs and brains of *D. subobscura* male and female larvae were probed with a monospecific antibody against the female-specific portion of the Sxl protein (BOPP *et al.* 1991). The arrows indicate the two bands corresponding to the 36- and 38-kD Sxl proteins.

```
mel   1 MYGNNNPGSNNNNGGYPPYGYNNKSSGGRGFGMSHSLPSGM(SRYAFSPQ)D   50
        :::::::::::::::::::::: :::::::::::::::::: :::::::: :
sub   1 MYGNNNPGSNNNNGGYPPYGYN-KSSGGRGFGMSHSLPSGM(SRYAFSPQ)D   49

mel   51 TEFSFPSSSSRRGYNDFPGCGGSG-G-NGGSANNLGGGNMCHLPPMASNN   98
         ::::::::::::::: ::: :: : : :::::::::: ::: : :: :::
sub   42 TEFSFPSSSSRRGYNEFPGGGGIGIGANGGSANNLGG-NMCNLLPMTSNN   98

mel   99 SLNNLCGLSLGSGGSDDLMN--DPRASNTNLIVNYLPQDMTDRELYALFR   146
         :: :::::::::::::: :   : : ::::::::::::::::::::::::
sub   91 SLSNLCGLSLGSGGSDDHMMMHDQRSSNTNLIVNYLPQDMTDRELYALFR   148

                                          RNP 1
mel  147 AIGPINTCRIMRDYKTGYSFGYAFVDFTSEMDSQRAIKVLNGITVRNKRL   196
         ::::::::::::::::::::::::::::::::::::::::::::::::::
sub  141 AIGPINTCRIMRDYKTGYSFGYAFVDFTSEMDSQRAIKVLNGITVRNKRL   198

mel  197 KVSYARPGGESIKDTNLYVTNLPRTITDDQLDTIFGKYGSIVQKNILRDK   246
         ::::::::::::::::::::::::::::::::::::::::::::::::::
sub  191 KVSYARPGGESIKDTNLYVTNLPRTITDDQLDTIFGKYGSIVQKNILRDK   248

                                          RNP 2
mel  247 LTGRPRGVAFVRYNKREEAQEAISALNNVIPEGGSQPLSVRLAEEHGKAK   296
         :::::::::::::::::::::::::::::::::::::::::: ::::::
sub  241 LTGRPRGVAFVRYNKREEAQEAISALNNVIPEGGSQPLSVRLAQEHGKAK   298

mel  297 AAHFMSQMGVVPANVPPPPPQPPPAHMAAAFNMMHRGRSIKSQQRFQNSHP   346
         ::::::: ::  :: ::::: :: :::   :: :::::::::::::  ::
sub  291 AAHFMSQIGVPSANAPPPPPPPP-HMAFN-NMVHRGRSIKSQQRFQKTHP   346

mel  347 YFDAKKFI 354
         :::: :::
sub  339 YFDAQKFI 354
```

FIGURE 5.—Comparative analysis of the predicted Sxl protein sequence of *D. melanogaster* and *D. subobscura*. The additional aminoacids when the most proximal acceptor splicing site of exon 5 is used are shown in parenthesis. The two regions that encode the RNP 1 and RNP 2 domains are underlined. The GN domain would correspond at least to the 38 first amino acids at the amino terminal end (WANG and BELL 1994).

Sequencing of the *Sxl* genomic region that is involved in the alternative splicing of *Sxl* pre-mRNA. A comparative analysis of *D. melanogaster* and *D. subobscura*: In the splicing of *D. melanogaster* late *Sxl* transcripts, the late exon L1 joins to exon L2, so that the early specific exon E1 functions as an intron. Moreover, the late *Sxl* transcripts show an alternate splicing pattern in both sexes: in females, exon L2 is joined to exon 4, so that exon L3 functions as an intron, whereas this exon L3 is incorporated in male transcripts (SALZ *et al.* 1989; SAMUELS *et al.* 1991; KEYES *et al.* 1992) (see Figure 1A). In order to identify *cis*-acting elements that might be involved in controlling the alternative *Sxl* pre-mRNA splicing, we sequenced the *D. subobscura* and *D. melanogaster* genomic region between early exon E1 and exon L2, as well as the male-specific exon L3 and their adjacent introns 2 and 3. It was not possible to make a perfect alignment between the sequences of both species because of the different size and the degree of divergence among these regions. For these reasons, only the comparative analysis of intron regions for which a certain degree of conservation was found is presented. The poly-T stretches in introns 2 and 3 were also compared. These poly-T sequences correspond to the poly-U stretches of the RNA that are the targets for the binding of the Sxl protein (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993a,b).

The gene *Sxl* of *D. virilis* has been cloned (BOPP *et al.* 1996). This gene shows sex-specific splicing, since the male *Sxl* transcripts contain an additional exon, L3. Only the complete sequence of this exon and the sequences of the regulated splice sites, as well as their comparative analysis with the corresponding regions of the gene *Sxl* of *D. melanogaster*, have been reported (BOPP *et al.* 1996). For this reason, the comparative analysis of the gene *Sxl* of *D. subobscura* and *D. virilis* is restricted to these regions.

Comparative analysis of the region between early specific exon E1 and exon L2: No poly-T stretches were found and the degree of homology was very low.

Comparative analysis of late intron 2: This intron is larger in *D. subobscura* than in *D. melanogaster*, and its size varies depending on which of the two exon L3 3′ splice sites are used in the two species. In the latter species the size is 2909 or 2927 bp if the first or the second 3′ splice sites, respectively, are used. The size

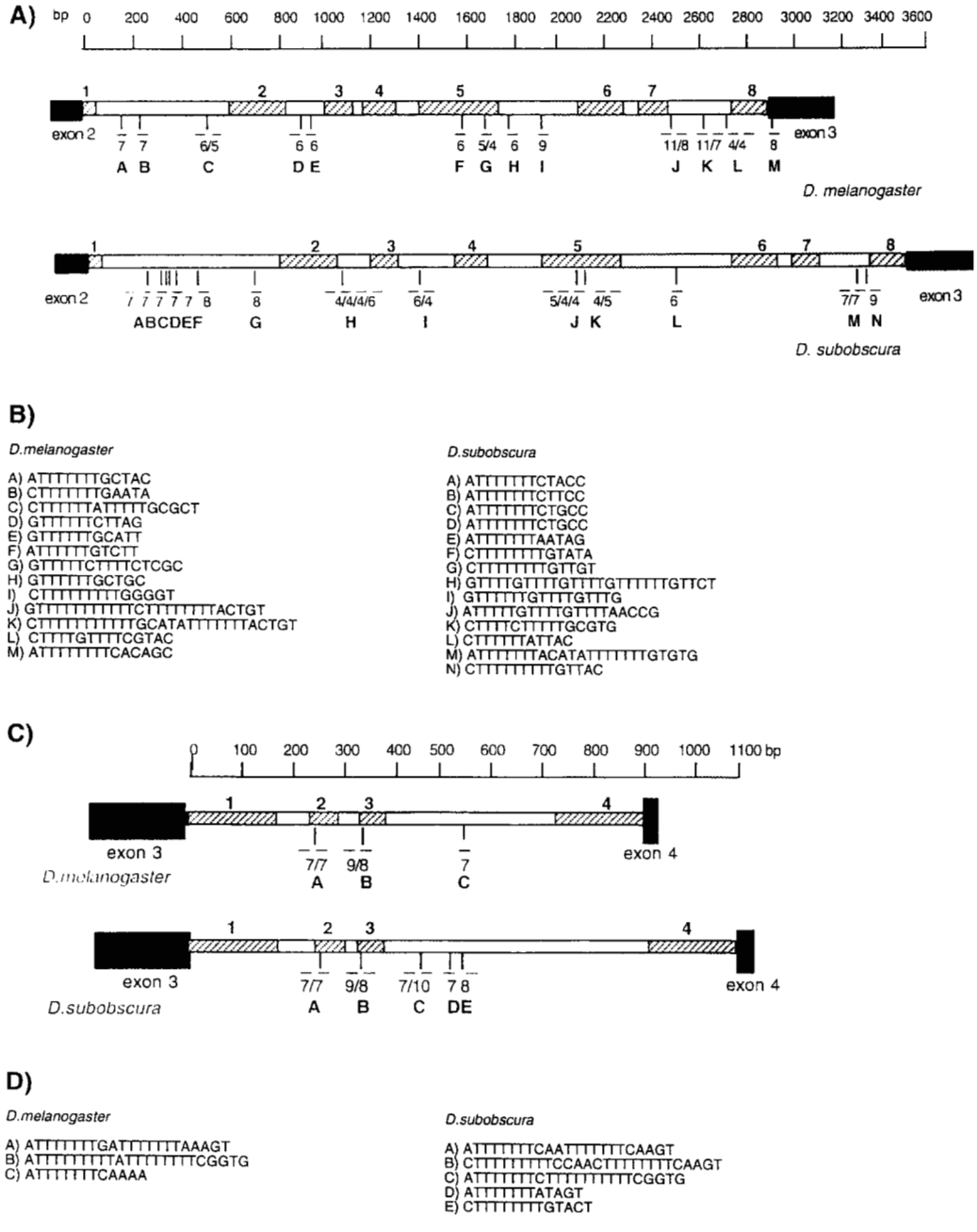

FIGURE 6.—Comparative analysis of late introns L2 and L3 of *D. melanogaster* and their homologues in *D. subobscura*. (A) Schematic representation of the intron 2 showing the eight regions (cross-hatched) for which extensive sequence similarity between both species was found, as well as the location of the poly-T sequences. The numbers refer to the number of T present in each sequence. (B) Nucleotide sequence of the poly-T stretches and their flanking bases. (C) Schematic representation of the intron 3 showing the four regions (cross-hatched) for which extensive sequence similarity between both species was found, as well as the location of the poly-T sequences. The numbers refer to the number of T present in each sequence. (D) Nucleotide sequence of the poly-T stretches and their flanking bases.

in the first species varies from 3559 to 3575 bp if the first or second 3′ splice sites, respectively, are used. Within the intron there are eight conserved regions which are distributed in the same relative order in the two species (Figure 6A). Some of these regions contain poly-T stretches, corresponding to the poly-Us of the RNA to which the Sxl protein binds. They are scattered along the intron and their locations vary between *D. melanogaster* and *D. subobscura* (Figure 6A). The size is similar but the flanking bases vary within and between the two species (Figure 6B). In *D. virilis*, poly-T stretches in intron 2 have also been found (BOPP *et al.* 1996). In *D. melanogaster* there is a T_8 stretch in the polypyrimidine tract of the male exon L3 3′ splice site. In *D. subobscura*, this T_8 stretch is modified: three T residues are replaced by three C residues. This T_8 stretch is also modified in *D. virilis*: two T residues are replaced by two C residues (BOPP *et al.* 1996).

Comparative analysis of the male specific exon L3: The male exon of *D. subobscura* has three instead of two translation stop codons and these stop codons are located at different positions than in *D. melanogaster* (Figure 7). The degree of homology between the male exon in the two species is low, agreeing with the fact that this exon functions as an intron in females and that its presence in male transcripts produces a truncated protein due to the translation stop codons. There are only two conserved regions, one in the middle and another at the 3′ end of this exon that extends into the beginning of intron 3. We also compared the male exon

```
5'ss exon 2

mel        CGAG GTAAATTTTtaAACTTg
sub        CGAG GTAAATTTTgtAACTTg
vir        CGAG GTAAATTTT--AACTTc
                5' splice site

3´ss exon 3

mel  aCGttgTaccGC-AGcTGAaTTTAAtcgTGTGTAG ACAT-attttTTTtCACAG C
sub  aCGatgTcctGC-AGcTGAaTTTAAtttTGTGTAG ACAT---ctcTTTcCACAG C
vir  tCGataTt-tGCCAGtTGA-TTTAA--tTGTGTAG ACATgatctcTTTtCACAG C
                                      3'splice site        3' splice site

5'ss exon 3

mel  ACgcAATGTT-tTTgAAtCGaGGACACc-TCcaAAgcCCT GTAAGTAAcaAgact
sub  ACaaAATGTT-cTTgAAtCGgGGACGCa-TC-gAAcgCCT GTAAGTAAcaAgcat
vir  ACaaAATGTTccTTaAA-CGaAGACGCcgTC--AAcgCCT GTAAGTAAagActag
                                            5' splice site

3´ss exon 4
                BP          Pp
mel  TtctatgatatCTAATTc---ATATTTTCTct-tccCcaaCgaaAtggAAtCGaTT--cATCgtaCAG tGGT
sub  Ttc--ccagtgCTAATAt---ATATTTTCTctatccCcctCatcAtcgAAtCGaTTgcaATCgtgCAG cGGT
vir  Taata----tgCTAATAtgcgATATTTTCTtctcttCaatCatcAatcAAaCG-TT-caATCaa-CAG cGGT
                                                                  3' splice site
```

FIGURE 7.—Comparative analysis of the regulated splice sites between *D. melanogaster, D. subobscura* and *D. virilis.*

of *D. subobscura* with that of *D. virilis* (BOPP *et al.* 1996). Aside from the homology at the 3′ splice sites and the presence of stop codons, the only other conserved sequence is at the 3′ end of this exon L3, which is also present in *D. melanogaster.* In the three species, there is a translation start codon (AUG) 32 nucleotides upstream of the exon L3-exon 4 splice junction. In *D. virilis,* this AUG initiates an open reading frame that extends into the downstream exons common to both sexes and codes for a Sxl protein of 35 kD in males (BOPP *et al.* 1996). However, in *D. subobscura* and *D. melanogaster,* there is no open reading frame after this AUG and males would only be expected to produce the truncated protein initiating from the AUG in exon L2.

Comparative analysis of late intron 3: Intron 3 is bigger in *D. subobscura* than in *D. melanogaster* (1108 *vs.* 919 bp) and contains four conserved regions (Figure 6C). The 5′ and 3′ splice sites and the adjacent sequences (regions 1 and 4 of Figure 6C) are conserved. The conserved sequence at the beginning of region 1 forms part of the conserved sequence (81% of homology in 58 bp) that starts at the 3′ end of exon L3 and extends to the beginning of intron 3, therefore including the 5′ splice site of the male specific exon L3. The other two conserved sequences (regions 2 and 3) are located approximately in the middle of the intron and also contain poly-T stretches. Their location within the intron is conserved (Figure 6C) and the flanking bases vary both within and between the species (Figure 6D). The intron 3 of *D. virilis* also has poly-T stretches (BOPP *et al.* 1996).

Comparative analysis of the regulated splice sites between *D. melanogaster*, *D. subobscura* and *D. virilis*: Figure 7 shows the comparative analysis of the regulated splice sites between *D. melanogaster, D. subobscura* and *D. virilis.* The 5′ splice site of exon L2 is fully conserved in the three species (AG GTAAAT) and differs from the consensus sequence (AG GTRAGT) (MOUNT *et al.* 1992) in that the highly conserved G residue at position +5 in the intron is replaced by an A residue in the three species. This contrasts with the 5′ splice site of the exon E1, which has a G residue at position +5 in both *D. melanogaster* and *D. subobscura* (in *D. virilis* the sequence of this exon has not been reported). The 5′ splice site of the male-specific exon L3 is fully conserved in the three species (CT GTAAGT) and differs from the consensus 5′ splice site sequence in that the AG residues are replaced by the CT residues. The two 3′ splice sites of the male-specific exon L3 are conserved in the three species: the first 3′ splice site is fully conserved and the second one differs only in that the residue at position −4 upstream of AG is a T in *D. melanogaster* and *D. virilis* and a C in *D. subobscura.* There is also a minor variation in the distance between the AG residues of the two 3′ splice sites: 17 bases in *D. virilis*, 14 bases in *D. subobscura* and 16 bases in *D. melanogaster.* The distance between the polypirimidine tract of the 3′ splice site of exon 4 and the AG residue of this splice site is approximately the same in the three species, though longer than the normal distance.

DISCUSSION

Organization of the gene *Sxl* of *D. subobscura*: In *D. subobscura,* four late *Sxl* transcripts have been identified

in adult flies of both sexes. The *Sxl* transcripts of this species are larger than those of *D. melanogaster* although the Sxl protein is of similar size in both species. These results indicate that the difference in the size of the *Sxl* transcripts is due to the larger size of the nontranslated regions in *D. subobscura*. Further, the gene *Sxl* of *D. subobscura* shows alternative splicing in females and males: the male transcripts differ from the female transcripts by an additional exon that contains three translation stop codons. Males therefore produce a truncated Sxl protein. This contrasts with the situation found in *D. virilis* where Sxl protein is produced in males, although their *Sxl* transcripts contain an additional exon L3 with translation stop codons. In the males of this species, an open reading frame exists downstream to the last stop codon that codes for a Sxl protein that is smaller than the Sxl protein of females. The function of this protein remains unknown (BOPP *et al.* 1996). In *D. pseudoobscura*, which belongs to the obscura group as *D. subobscura*, no Sxl protein has been found in males (BOPP *et al.* 1996).

The exon E1 of the early *Sxl* transcripts contains the translation initiation site and the leader sequence. The comparative analysis of this exon in the two species revealed high homology in the 5′ UTR, suggesting that this region might play an important role in the translation of early *Sxl* transcripts. This conserved sequence contains a stretch that is homologous to the sequence found in the 5′ UTR of heat-shock mRNAs (HOLMGREN *et al.* 1981). The *D. melanogaster* heat-shock genes *hsp22* and *hsp70* are examples of where the 5′ UTR of the mRNAs are involved in regulation of translation. Under conditions of heat-shock, the translation of all mRNAs is suppressed except for the heat-shock protein mRNAs, which are translated at a higher rate. However, when the 5′ UTR of these latter transcripts is deleted, their translation is prevented at high temperatures though not at 25° (KLEMENZ *et al.* 1985; MACGARRY and LINDQUIST 1985; HULTMARK *et al.* 1986). In the case of the early *Sxl* transcripts, it is suggested that the 5′ UTR might be involved in their efficient translation. Therefore, in females, the production of a certain amount of early Sxl protein can be assured since this protein is needed in all cells for early dosage compensation. It is also needed for establishing the positive autoregulation of *Sxl* through its involvement in female-specific splicing of the first *Sxl* pre-mRNA from the late constitutive promotor.

Control of the early *Sxl* expression: a comparison of *D. melanogaster* and *D. subobscura*: The study of early *Sxl* expression in *D. subobscura* shows that *Sxl* is initially expressed at the same developmental stage as in *D. melanogaster* and that this early expression seems to be controlled at the transcriptional level by the *X*:A signal in females. In *D. melanogaster* it has been shown that a fragment of 1400 bp upstream of the early *Sxl* transcription initiation site contains all the *cis*-acting elements required for the control of early *Sxl* expression by the *X*:A signal (ESTES *et al.* 1995). The study of the effect of different deletions within this fragment on early *Sxl* expression has revealed the existence of two regions (−200 to −400 bp, and −800 to −1400 bp) required for *Sxl* activation; the deletion of one region having a lesser effect than the deletion of both regions (ESTES *et al.* 1995). No deletion of the first 200 bp was analyzed. These two regions coincide with the conserved sequences (-1 to -372 bp and -759 to -946 bp) found in the interspecific comparative analysis. This agrees with the key role of these two regions in the activation of the early *Sxl* promotor. These regions contain E boxes, which are the binding sites for the Sc-Da heterodimers that act as activators of *Sxl* transcription. The region adjacent to the transcription initiation site also contains the single D box, which is the binding site for the Dpn protein (an element of the *X*:A signal) that acts as a repressor of *Sxl* transcription. In addition to the E and D boxes, the two regions contain other sequences that show a high degree of conservation, though they do no present homology with other known sequences involved in transcriptional control (Sequence Data Banks were searched). These conserved sequences might be the binding sites for other factors involved in the early activation of *Sxl*; for example, the Sis-a protein and/or other factors (SÁNCHEZ *et al.* 1994). Another possibility is that the conserved sequences are the binding sites for autosomal products (denominator elements of the *X*:A signal) that would suppress *Sxl* activation by preventing the binding of the Sc-Da complexes to the E boxes.

In *D. subobscura*, there are four E boxes within the −1000 to −1150 bp region that are not present in *D. melanogaster*. In addition, this region is not conserved in the two species. The involvement of these E boxes in the early activation of *Sxl* remains unknown. However, the fact that these E sequences are present suggests the possibility that they have a function. In this context, the five E boxes of *D. melanogaster*, compared with the nine E boxes of *D. subobscura*, might suggest the early *Sxl* promotor in this latter species to be weaker, and therefore require a greater quantity of activators and E boxes to reach a similar level of activation. An alternative, though not mutually exclusive explanation, is that the *X*:A signal in *D. subobscura* has a lower affinity for the *cis*-acting elements that control the early *Sxl* promotor. A higher number of these elements would then be required to ensure a similar level of *Sxl* activation in both species.

***cis*-acting elements involved in the splicing regulation of *Sxl* pre-mRNA: a comparison of *D. melanogaster* and *D. subobscura*:** *Splicing control of early Sxl pre-mRNA:* It is not known how splicing of the early *Sxl* pre-mRNA is controlled, though it is recognized that the Sxl protein does not take part in this process. It has been suggested that splicing of early *Sxl* pre-mRNA represents the default splice, so that the intrinsic structure of this early transcript determines its splicing pattern (HORABIN and SCHEDL 1996). The comparative analysis of *D. melanogaster* and *D. subobscura* revealed that the 5′ splice site of exon E1 is fully conserved in both species and shows greater homology with the consensus sequence than do

the 5′ splice site of exons L2 and L3. Thus, a possibility is that the 5′ splice site of exon E1 is stronger than the 5′ splice site of exons L2 and L3 for their union to exon 4. In addition, the exon E1 has a cap structure which favors this exon's participation in the splicing process (IZAURRAULDE *et al.* 1994; KATAOKA *et al.* 1994).

Splicing control of late Sxl pre-mRNA: The Sxl protein participates in the control of female-specific splicing of late *Sxl* pre-mRNA. The action of this protein occurs through its binding to the poly-U sequences located in introns 2 and 3. *In vitro* analyses have shown that the relative position of these poly-U's within the introns (WANG and BELL 1994), and their flanking bases influence the binding of the Sxl protein (SAKASHITA and SAKAMOTO 1994; SAMUELS *et al.* 1994; SINGH *et al.* 1995). In both species, the bases that flank the poly-U stretches show variation. This contradicts the central role of these flanking bases for Sxl binding *in vivo*. The *D. melanogaster Sxl* pre-mRNA contains a poly-U stretch in the 3′ splice site of the male-specific exon (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993b). This poly-U sequence, however, is absent in the *Sxl* pre-mRNA of *D. subobscura* and *D. virilis* (BOPP *et al.* 1996). This would support the idea that this sequence plays no central role in the control of splicing of *Sxl* pre-mRNA (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993a,b). This contrasts with the situation encountered in the *D. melanogaster tra* pre-mRNA splicing. This RNA has this poly-U stretch in the 3′ splice site of the non-sex-specific exon. This stretch seems to constitute the unique *cis*-acting element that plays a central role in splicing control of the *tra* pre-mRNA, since its alteration affects the sex-specific splicing of this RNA (SOSNOWSKI *et al.* 1989; INOUE *et al.* 1990; VALCÁRCEL *et al.* 1993). In fact, this is the only target where the Sxl protein binds, preventing the binding of the general splicing factor U2AF, which results in the female-specific splicing of this *tra* RNA (VALCÁRCEL *et al.* 1993). Therefore, the splicing mechanisms of the *tra* and *Sxl* RNAs seem to be different (VALCÁRCEL *et al.* 1993).

The comparison of *D. melanogaster* and *D. subobscura* has identified two conserved sequences, one at the 3′ end of intron 2 just upstream of the 3′ splice site of exon L3 and another at the 3′ end of this exon extending into the beginning of intron 3 and including the 5′ splice site of this exon L3. These sequences are also conserved in *D. virilis* (BOPP *et al.* 1996). These conserved sequences do not correspond to poly-U stretches that are the binding sites of the Sxl protein. It is suggested, then, that these sequences could be the binding targets for other factor(s) known to be required in splicing control of *Sxl* pre-mRNA. In this respect, three genes have been identified whose functions are required for proper female-specific *Sxl* RNA splicing: *fl(2)d* (GRANADINO *et al.* 1990–1992), *snf* (OLIVER *et al.* 1988, 1993; STEIMANN-ZWICKY 1988; SALZ 1992; ALBRECHT and SALZ 1993; BOPP *et al.* 1993; FLICKINGER and SALZ 1994) and *vir* (HILFIKER and NÖTHIGER 1991; HILFIKER *et al.* 1995). Mutations in these genes cause *Sxl* pre-mRNA to follow the male instead of the female-specific splicing pattern in female flies. The gene *snf* has been cloned and its product characterized (FLICKINGER and SALZ 1994). The Snf protein shows homology with the U1A snRNP of mammals. This protein, which is a component of the spliceosome, functions at an early point in the splicing process through its binding to the 5′ splice site. It has been proposed that the Snf protein might be involved in the recognition of the 5′ splice site of exon L3 (FLICKINGER and SALZ 1994). The *cis*-acting element of this recognition might be the conserved sequence mentioned above, where the Snf protein would bind. However, this sequence could also be the binding site of other factor(s) with which Snf would interact.

Results reported here and those previously reported (SAKAMOTO *et al.* 1992; HORABIN and SCHEDL 1993b; FLICKINGER and SALZ 1994; WANG and BELL 1994) would agree with the hypothesis of exon definition (ROBBERSON *et al.* 1990). This hypothesis states that the exon is the initial unit of spliceosome assembly: communication between recognition factors for the 3′ splice site and recognition factors for the 5′ splice site of the exon functions to concertedly recognize both ends of the exon during early assembly of the spliceosome. Following this hypothesis, the elimination of the male-specific exon L3 from late *Sxl* pre-mRNA would require that the function of the 3′ and 5′ splice sites of this exon be blocked. This blockage might be carried out by factors that bind to the conserved sequences (which are different from the poly-U stretches that bind the Sxl protein) located upstream and contiguous to the 3′ splice site and to the conserved sequence that includes the 5′ splice site. The binding of these factors might require the binding of the Sxl protein to the poly-U sequences in introns 2 and 3. In addition, the cooperativity between the Sxl proteins bound to introns 2 and 3 might be required for looping out the DNA region containing the male-specific exon L3. This brings into proximity exons L2 and 4, which will be joined.

Elimination of the early exon E1 from late Sxl pre-mRNA: The comparison of *D. melanogaster* and *D. subobscura* shows that the genomic region between the early exon E1 and exon L2 contains no conserved sequences and lacks the Sxl-binding poly-U stretches. This suggests that the Sxl protein does not directly participate in the elimination of early exon E1 from the late *Sxl* pre-mRNA. This elimination might be explained by the fact that early exon E1 lacks a 3′ splice site and potentially cannot be joined to late exon L1. However, it still contains a 5′ splice site that could appear to compete with the 5′ splice site of late exon L1 for joining exon L2, though this is something that does not appear to happen. A possibility is that the strength of this 5′ splice site is partly lost because the cap structure is no longer associated with this early exon E1. Alternatively, though not necessarily exclusive, the function of this 5′ splice site may be prevented by some factor(s).

The Sxl protein: a comparison of *D. melanogaster* and *D.*

subobscura: Both RNP and GN domains are fully conserved in the Sxl proteins of both species. This agrees with the functional role of these domains. Moreover, the Sxl protein is also highly conserved outside these domains. These other conserved regions might represent new domains of the Sxl protein required for its interaction with other proteins that take part in the splicing process. Alternatively, these conserved regions may not be new functional domains but structural domains whose conservation is required to allow the active conformation of the RNP and GN domains responsible for all interactions carried out by the Sxl protein.

The homologous *Sxl* gene of *Chrysomya rufifacies* has been cloned (MÜLLER-HOLTKAMP 1995). In this species, *Sxl* is transcribed in both sexes from the blastoderm stage throughout development and its pre-mRNA shows no sex differential splicing. Therefore, both males and females have the Sxl protein. This questions the role of *Sxl* in sex determination in this species. The homology between the Sxl proteins of *D. melanogaster* and *Ch. rufifacies* is restricted to the RNP domains (90% of homology) whereas the rest of the protein shows variation (18% of homology). In *D. melanogaster, D. subobscura* and *Ch. rufifacies* exon 5 has two 3′ splice sites. The use of the second splice site eliminates the same eight amino acids from the Sxl protein in these three species, so that these amino acids can behave as introns for certain Sxl proteins. The possibility of using different splice sites determines the production of distinct Sxl proteins with minor differences. These could be specific of certain cell types (BOPP *et al.* 1991; SAMUELS *et al.* 1991). The fact that the eight amino acids of exon 5 are conserved suggests that these amino acids may be important for the function of the Sxl protein that incorporates them.

The high degree of conservation of the Sxl protein of *D. melanogaster* and *D. subobscura* contrasts with the high degree of variation of the Tra protein of *D. hydei* and *D. virilis.* These proteins have an homology of 31 and 36%, respectively, when compared to the Tra protein of *D. melanogaster* (O'NEIL and BELOTE 1992). However, in spite of this low degree of conservation, the Tra protein from *D. virilis* confers the *tra* function when introduced into *D. melanogaster.* This can be explained by the fact that the functional domains of the Tra protein are made of small stretches of arginine-serine amino acids, which are conserved in the Tra proteins of these species.

We are grateful to R. DE ANDRÉS and D. MATEOS for technical assistance, to C. GODAY, R. ESTEBAN and G. GIOVINAZZO for help with the Western blot and to T. UEMURA for a Drosophila cDNA library. This work was supported by grant PB92-0006 from D.G.I.C.Y.T. (Spain) to L.S. and in part by grant PB94-923 from D.G.I.C.Y.T. (Spain) to M. Aguadé, and research grants from the Ministry of Education, Science and Culture of Japan, the Asahi Glass Foundation, the Inamori Foundation, and the Senri Life Science Foundation to H.S. L.O.F.P. was the recipient of a predoctoral fellowship from Ministerio de Educación y Ciencia-ICI, A.N.-S. is the recipient of a predoctoral fellowship from Comissió Interdepartamental de Recerca i Tecnologia, and E.S. is the recipient of a JSPS Research Fellowship for Young Scientists.

LITERATURE CITED

ALBRECHT, E. B., and H. K. SALZ, 1993 The *Drosophila* sex determination gene *snf* is utilized for the establishment of the female-specific splicing pattern of *Sex-lethal.* Genetics **134:** 801–807.

BACHILLER, D., and L. SÁNCHEZ, 1991 Production of XO clones in XX females of *Drosophila.* Genet. Res. **57:** 23–28.

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., M. GORMAN and I. MARÍN, 1994 Dosage compensation in *Drosophila.* Annu. Rev. Genet. **28:** 491–521.

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 similar to RNA binding proteins. Cell **55:** 1037–1046.

BELL, L. R., J. I. HORABIN, P. SCHEDL and T. W. CLINE, 1991 Positive autoregulation of *Sex-lethal* by alternative splicing maintains the female determined state in *Drosophila.* Cell **65:** 229–239.

BOPP, D., L. R. BELL, T. W. CLINE and P. SCHEDL, 1991 Developmental distribution of female-specific Sex-lethal proteins in *Drosophila melanogaster.* Genes Dev. **5:** 403–415.

BOPP, D., J. I. HORABIN, R. A. LERSH, T. W. CLINE and P. SCHEDL, 1993 Expression of *Sex-lethal* gene is controlled at multiple levels during the *Drosophila* oogenesis. Development **118:** 797–812.

BOPP, D., G. CALHOUN, J. I. HORABIN, M. SAMUELS and P. SCHEDL, 1996 Sex-specific control of *Sex-lethal* is a conserved mechanism for sex determination in the genus *Drosophila.* Development **122:** 971–982.

CAMPOS-ORTEGA, J. A., and V. HARTENSTEIN, 1985 *The Embryonic Development of Drosophila melanogaster.* Springer-Verlag, Berlin, Heidelberg.

CAVENER, D., 1987 Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. Nucleic Acids Res. **15:** 1353–1361.

CLINE, T. W., 1984 Autoregulatory functioning of a *Drosophila* gene product that establishes and maintains the sexually determined state. Genetics **107:** 231–277.

CLINE, T. W., 1993 Sex determination in *Drosophila*: how flies count to two? Trends Genet. **9:** 385–390.

CUBAS, P., J. CELIS, S. CAMPUZANO and J. MODOLELL, 1991 Proneural clusters of *achaete-scute* expresion and the generation of sensory organs in the *Drosophila* imaginal wing disc. Genes Dev. **5:** 996–1008.

ESTES, P., L. N. KEYES and P. SCHEDL, 1995 Multiple response elements in the *Sex-lethal* promoter ensure its female-specific expression pattern. Mol. Cell. Biol. **15:** 904–917.

FLICKINGER, T. W., and H. K. SALZ, 1994 The *Drosophila* sex determination gene *snf* encodes a nuclear protein with sequences and functional similarity to the mammalian U1A snRNP protein. Genes Dev. **8:** 914–925.

FROHMAN, M. A., M. K. DUSH and G. R. MARIN, 1988 Rapid production of full-lenght cDNAs from rare transcripts: amplification using a single gene-specific oligo-nucleotide primer. Proc. Natl. Acad. Sci. USA **85:** 8998–9002.

GRANADINO, B., S. CAMPUZANO and L. SÁNCHEZ, 1990 The *Drosophila melanogaster fl(2)d* gene is needed for the female-specific splicing of *Sex-lethal* RNA. EMBO J. **9:** 2597–2602.

GRANADINO, B., A. B. SAN JUÁN and L. SÁNCHEZ, 1991 The gene *fl(2)d* is required for various Sxl-controlled processes in *Drosophila* females. Roux's Arch. Dev. Biol. **200:** 172–176.

GRANADINO, B., A. B. SAN JUÁN, P. SANTAMARÍA and L. SÁNCHEZ, 1992 Evidence of a dual function in *fl(2)d* a gene needed for *Sex-lethal* expression in *Drosophila melanogaster.* Genetics **130:** 597–612.

HENIKOFF, S., 1984 Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:** 351–359.

HILFIKER, A., and R. NÖTHIGER, 1991 The temperature-sensitive mutation vir^{ts} *(virilizer)* identifies a new gene involved in sex determination of *Drosophila.* Roux's Arch. Dev. Biol. **200:** 240–248.

HILFIKER, A., H. AMREIN, A. DÜBENDORFER and R. NÖTHIGER, 1995 The gene *virilizer* is required for female-specific splicing controlled by *Sxl,* the master gene for sexual development in *Drosophila.* Development **121:** 4017–4026.

HOLMGREN, R., V. CORCES, R. MORIMOTO, R. BLACKMAN and M. MESELSON, 1981 Sequence homologies in the 5′ regions of four *Drosophila* heat-shock genes. Proc. Natl. Acad. Sci. USA **78:** 3775–3778.

HORABIN, J. I., and P. SCHEDL, 1993a Regulated splicing of the *Dro-*

sophila Sex-lethal male exon involves a blockage mechanism. Mol. Cell Biol. **13:** 1408–1414.

HORABIN, J. I., and P. SCHEDL, 1993b *Sex-lethal* autoregulation requires multiple cis-acting elements upstream and downstream of the male exon and appears to depende largely on controlling the use of the male exon 5′ splice site. Mol. Cell Biol. **13:** 7734–7746.

HORABIN, J. I., and P. SCHEDL, 1996 Splicing of the *Drosophila Sex-lethal* early transcripts involves exon skipping that is independent of Sex-lethal protein. RNA **2:** 1–10.

HOSHIJIMA, K., A. KOHYAMA, I. WATAKABE, K. INOUE, H. SAKAMOTO *et al.*, 1995 Transcriptional regulation of the *Sex-lethal* gene by helix-loop-helix proteins. Nucleic Acids Res. **23:** 3441–3448.

HULTMARK, D., R. KLEMENZ and W. GEHREING, 1986 Translational and transcriptional control elements in the untranslated leader of the heat-schock gene *hsp22*. Cell **44:** 429–438.

INOUE, K., K. HOSHIJIMA, H. SAKAMOTO and Y. SHIMURA, 1990 Binding of the *Drosophila Sex-lethal* gene product to the alternative splice site of *transformer* primary transcript. Nature **344:** 461–463.

IZAURRAULDE, E., J. LEWIS, C. MCGUIGAN, M. JANKOWSKA, DARZYNKIEWICZ *et al.*, 1994 A nuclear cap bindind protein complex involved in pre-mRNA splicing. Cell **78:** 657–668.

KATAOKA, N., M. OHNO, K. KANGAWA, Y. TOKORO and Y. SHIMURA, 1994 Cloning of a complementary DNA encoding an 80 kilodalton nuclear cap binding protein. Nucleic Acids Res. **22:** 3861–3865.

KEYES, L. N., T. W. CLINE and P. SCHEDL, 1992 The primary sex determination signal of *Drosophila* acts at the level of transcription. Cell **68:** 933–943.

KLEMENZ, K., D. HULTMARK and W. GEHRING, 1985 Selective translation of heat shock mRNA in *Drosophila melanogaster* depends on sequence information in the leader. EMBO J. **4:** 2053–2060.

KÜNZE-MÜHI., E., and E. MÜLLER, 1958 Weikre untersuchungen uber die chromosomale struktur und die naturlichen struckturtyper von *Drosophila subobscura*. Chromosoma **9:** 559–570.

LEFEVRE, G. J., 1976 *A Photographic Representation and Interpretation of the Polytene Chromosomes of D. melanogaster Salivary Glands*. Academic Press, New York.

LUCCHESI, J. C., and T. SKRIPSKY, 1981 The link between dosage compensation and sex differentiation in *Drosophila melanogaster*. Chromosoma **82:** 217–227.

MACGARRY, T., and S. LINDQUIST, 1985 The preferential translation of *Drosophila* hsp70 mRNA requires sequences in the untraslated leader. Cell **42:** 903–911.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

MONTGOMERY, E. B., B. CHARLESWORTH and C. H. LANGLEY, 1987 A test for the role of natural selection in the stabilization of transposable element copy number in a population of *Drosophila melanogaster*. Genet. Res. **49:** 31–41.

MOUNT, S. M., C. BURKS, G. HERTZ, G. D. STORMO, O. WHITE *et al.*, 1992 Splicing signals in *Drosophila*: intron size, information content and consensus sequences. Nucleic Acids Res. **20:** 4255–4262.

MÜLLER-HOLTKAMP, F., 1995 The *Sex-lethal* gene in *Chrysomya rufifacies* is highly conserved in sequence and exon-intron organization. J. Mol. Evol. **41:** 467–477.

OLIVER, B., N. PERRIMON and A. P. MAHOWALD, 1988 Genetic evidence that the *sans-fille* locus is involved in *Drosophila* sex determination. Genetics **120:** 159–171.

OLIVER, B., Y. J. KIM and B. S. BAKER, 1993 *Sex-lethal*, master and slave: A hierarchy of germ line sex determination in *Drosophila*. Development **119:** 897–908.

O'NEIL, M. T., and J. M. BELOTE, 1992 Interspecific comparison of *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. Genetics **131:** 113–128.

PARKHURST, S. M., and D. ISH-HOROWICZ, 1992 Common denominators for sex. Curr. Biol. **2:** 629–631.

PEARSON, W. R., and D. J. LIPMAN, 1988 Improved tools for biological sequence comparinson. Proc. Natl. Acad. Sci. USA **85:** 2444–2448.

PIRROTTA, V., 1986 Cloning *Drosophila* genes, pp. 83–110 in *Drosophila: a Practical Approach*, edited by D. B. ROBERTS. IRL Press, Oxford.

ROBBERSON, B. L., G. J. COTE and S. M. BERGET, 1990 Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell Biol. **10:** 84–94.

SAKAMOTO, H., K. INOUE, I. HIGUCHI, Y. ONO and Y. SHIMURA, 1992 Control of *Drosophila Sex-lethal* pre-mRNA splicing by its own female-specific product. Nucleic Acids Res. **20:** 5533–5540.

SAKASHITA, E., and H. SAKAMOTO, 1994 Characterization of RNA binding specificity of the *Drosophila* Sex-lethal protein by *in vitro* ligand selection. Nucleic Acids Res. **22:** 4082–4086.

SALZ, H. K., 1992 The genetic analysis of *snf*: a *Drosophila* sex determination gene requires for activation of *Sex-lethal* in both the germline and the soma. Genetics **130:** 547–554.

SALZ, H. K., E. M. MAINE, L. N. KEYES, M. E. SAMUELS, T. W. CLINE *et al.*, 1989 The *Drosophila* female-specific sex-determination gene, *Sex-lethal*, has stage-, tissue-, and sex-specific RNAs suggesting multiple modes of regulation. Genes Dev. **3:** 708–719.

SAMUELS, M. E., D. BOPP, R. A. COLVIN, R. F. ROSCIGNO, M. GARCÍA-BLANCO *et al.*, 1994 RNA binding by Sxl proteins *in vitro* and *in vivo*. Mol. Cell Biol. **14:** 4975–4990.

SAMUELS, M. E., P. SCHEDL and T. W. CLINE, 1991 The complex set of late transcriptes from the *Drosophila* sex determination gene *Sex-lethal* encodes multiple related polypeptides. Mol. Cell. Biol. **11:** 3584–3602.

SÁNCHEZ, L., and R. NÖTHIGER, 1983 Sex determination and dosage compensation in *Drosophila melanogaster*: production of male clones in XX females. EMBO J. **2:** 485–491.

SÁNCHEZ, L., B. GRANADINO and M. TORRES, 1994 Sex determination in *Drosophila melanogaster*: X-linked genes involved in the initial step of *Sex-lethal* activation. Dev. Genet. **15:** 251–264.

SANGER, F., S. NICKLEN and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463–5467.

SEGARRA, C., and M. AGUADÉ, 1993 Nucleotide divergence of the *rp49* gene region between *Drosophila melanogaster* and two species of the obscura group of *Drosophila*. J. Mol. Evol. **36:** 243–248.

SINGH, R., J. VALCÁRCEL and M. R. GREEN, 1995 Distinct binding and functions of higher eucaryotic polypirimidine tract-binding proteins. Science **268:** 1173–1176.

SOSNOWSKI, B. A., J. M. BELOTE and M. MCKEOWN, 1989 Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence dependent splice site blockage. Cell **58:** 449–459.

STEIMANN-ZWICKY, M., 1988 Sex determination in *Drosophila*: the X chromosomal gene *liz* is required for *Sxl* activity. EMBO J. **7:** 3889–3898.

STEINMANN-ZWICKY, M., H. AMREIN and R. NÖTHIGER, 1990 Genetic control of sex determination in *Drosophila*. Adv. Genet. **27:** 189–337.

TAUTZ, D., and C. PFEIFLE, 1989 A non-radiactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translation control of the segmentation gene *hunchback*. Chromosoma **98:** 81–85.

THROCKMORTON, L. H., 1975 *The Phylogeny, Ecology, and Geographic of Drosophila*. Plenum Press, New York.

TORRES, M., and L. SÁNCHEZ, 1991 The sisterless-b function of the *Drosophila* gene *scute* is restricted to the state when the X:A ratio signal determines the activity of *Sex-lethal*. Development **113:** 715–722.

VALCÁRCEL, J., R. SINGH, P. D. ZAMORE and M. R. GREEN, 1993 The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of *transformer* pre-mRNA. Nature **362:** 171–175.

WANG, J., and L. R. BELL, 1994 The Sex-lethal amino terminus mediates cooperative interactions in RNA binding and is essentialfor splicing regulation. Genes Dev. **8:** 2072–2085.

YOUNGER-SHEPHERD, S., H. VAESSIN, E. BIER, L. YEH JAN and Y. NUNG JAN, 1992 *deadpan*, an essential pan-neural gene encoding an HLH protein, acts as a denominator in *Drosophila* sex determination. Cell **70:** 911–922.

Communicating editor: T. SCHÜPBACH