Interspecific Comparison of the *transformer* Gene of Drosophila Reveals an Unusually High Degree of Evolutionary Divergence

Michael T. O'Neil and John M. Belote[1]

Department of Biology, Syracuse University, Syracuse, New York 13244

Manuscript received November 21, 1991
Accepted for publication January 25, 1992

ABSTRACT

The *transformer* (*tra*) gene of *Drosophila melanogaster* occupies an intermediate position in the regulatory pathway controlling all aspects of somatic sexual differentiation. The female-specific expression of this gene's function is regulated by the *Sex lethal* (*Sxl*) gene, through a mechanism involving sex-specific alternative splicing of *tra* pre-mRNA. The *tra* gene encodes a protein that is thought to act in conjunction with the *transformer-2* (*tra-2*) gene product to control the sex-specific processing of *doublesex* (*dsx*) pre-mRNA. The bifunctional *dsx* gene carries out opposite functions in the two sexes, repressing female differentiation in males and repressing male differentiation in females. Here we report the results from an evolutionary approach to investigate *tra* regulation and function, by isolating the *tra*-homologous genes from selected Drosophila species, and then using the interpecific DNA sequence comparisons to help identify regions of functional significance. The *tra*-homologous genes from two Sophophoran subgenus species, *Drosophila simulans* and *Drosophila erecta*, and two Drosophila subgenus species, *Drosophila hydei* and *Drosophila virilis*, were cloned, sequenced and compared to the *D. melanogaster tra* gene. This comparison reveals an unusually high degree of evolutionary divergence among the *tra* coding sequences. These studies also highlight a highly conserved sequence within intron one that probably defines a *cis*-acting regulator of the sex-specific alternative splicing event.

THE *transformer* (*tra*) gene of *Drosophila melanogaster* is one of a set of major regulatory loci controlling somatic sexual differentiation (reviewed by BAKER 1989; STEINMANN-ZWICKY, AMREIN and NÖTHIGER 1990). Like most of the other genes that have been identified in this regulatory pathway, the *tra* gene is expressed in a sex-specific manner. In chromosomally male individuals (*i.e.*, *XY*), the *tra* gene is functionally "off," resulting in the fly following the default pathway of male differentiation. In chromosomally female flies (*i.e.*, *XX*), the *tra* gene is functionally "on," where it acts to direct proper female differentiation of all sexually dimorphic somatic tissues. Loss-of-function mutations of the *tra* gene are, therefore, female-specific in their effects, with *XX; tra*$^-$ individuals being transformed to maleness, and *XY; tra*$^-$ being unaffected.

Molecular genetic studies of the *tra* gene have given us insight into how this sex-specific expression is achieved (BUTLER *et al.* 1986; MCKEOWN, BELOTE and BAKER 1987; BELOTE *et al.* 1989). Although the *tra* gene is transcribed from the same promoter in both sexes, the pre-mRNA is subjected to sex-specific alternative splicing of the first intron (BOGGS *et al.* 1987). In males, the *tra* pre-mRNA is spliced in such a way that the mature transcript contains no long open reading frame (ORF), and can encode only a truncated, inactive *tra* polypeptide. In females, about half of the *tra* pre-mRNA that is made is spliced in a different manner, to yield an mRNA with an ORF encoding a polypeptide of 197 amino acids, representing the functional *tra* gene product. The alternative splicing of *tra* pre-mRNA is regulated by the functioning of the *Sex lethal* (*Sxl*) gene (NAGOSHI *et al.* 1988; BELL *et al.* 1991). This gene is also expressed in a female-specific manner, and it encodes an RNA binding protein that is thought to act directly on *tra* pre-mRNA to influence its splicing (BELL *et al.* 1988; INOUE *et al.* 1990). It does this by inhibiting the use of the non-sex-specific 3′ splice site and allowing the use of the downstream female-specific 3′ splice site (SOSNOWSKI, BELOTE and MCKEOWN 1989). In males, no *Sxl* protein is produced, and *tra* pre-mRNA is spliced using the preferred non-sex-specific 3′ splice site by default.

Genetic and molecular analyses of *tra* suggest that its gene product acts in conjunction with the *transformer-2* (*tra-2*) gene product to control the sex-specific expression of the *doublesex* (*dsx*) gene (BAKER and RIDGE 1980; NAGOSHI *et al.* 1988; MCKEOWN, BELOTE and BOGGS 1988). The *dsx* gene is a bifunctional locus that has active, but opposite, functions in males and females (BAKER and RIDGE 1980; NÖTHIGER *et al.* 1987; BAKER and WOLFNER 1988). In males, *dsx* acts

[1] To whom reprint requests should be sent.

to repress female-specific differentiation of the sexually dimorphic somatic tissues, thereby leading to normal male development, and in females it acts to repress male differentiation of the soma, resulting in normal female development. The sex-specific expression of *dsx* is achieved by sex-specific alternative processing of *dsx* pre-mRNA, yielding different *dsx* gene products in the two sexes (BURTIS and BAKER 1989). It has been proposed that the *tra* and *tra-2* gene products affect *dsx* pre-mRNA splicing by activating the female-specific 3′ splice site in the *dsx* transcript (BURTIS and BAKER 1988; NAGOSHI and BAKER 1990; HOSHIJIMA *et al.* 1991; HEDLEY and MANIATIS 1991; RYNER and BAKER 1991).

Among the important questions concerning the *tra* gene's expression and function are (1) how is the sex-specific splicing of *tra* pre-mRNA regulated by factors such as the *Sxl* gene product? and (2) what is the exact function of the *tra* protein? With respect to the first question, one relevant topic of inquiry deals with the *cis*-acting sequences that are responsible for the sex-specific control of intron 1 splicing. Work to date has focused on using site-directed mutagenesis procedures to manipulate the *tra* gene *in vitro*, and then examining the effects of those changes on the sex-specific splicing of the pre-mRNA in either transgenic flies (SOSNOWSKI, BELOTE and MCKEOWN 1989), or in transgenic cell lines (INOUE *et al.*, 1990). As an alternative approach, we have examined the *tra* gene in other Drosophila species, in order to identify conserved sequence elements within the regulated intron that might identify important regulatory sequences controlling the alternative splicing event. With respect to the second question stated above, it was hoped that the interspecific comparison of the *tra* gene's coding region would reveal conserved domains of the protein and help focus future studies of the *tra* protein on regions of functional significance. Because the species chosen for our study have been estimated to have diverged approximately 60 million years ago, sequence conservation should be limited to functionally relevant regions of the gene.

MATERIALS AND METHODS

Fly strains: *Drosophila hydei* flies were obtained from a stock provided by DAVID SULLIVAN, Syracuse University, *Drosophila erecta* flies from a stock provided by DONAL HICKEY, University of Ottawa, and *Drosophila virilis* from a stock provided by LARRY SALKOFF, Washington University School of Medicine. Flies were maintained at 22° on cornmeal-molasses-yeast-agar media containing propionic acid as a mold inhibitor. Mutations of *Drosophila melanogaster* are described in LINDSLEY and GRELL (1968) or in LINDSLEY and ZIMM (1985, 1986, 1987, 1988, 1990).

Genomic libraries: The *Drosophila simulans* and *D. hydei* libraries were constructed by ligating size-fractionated genomic DNA that had been partially digested with *Mbo*I into the *Bam*HI site of the bacteriophage λEMBL4 vector (FRISCHAUFF *et al.* 1983). The recombinant DNA was packaged into phage particles using the Gigapack Bacteriophage λ Packaging Kit from Stratagene according to the supplier's protocol, and plated on *Escherichia coli* strain Q359. An additional *D. hydei* library, similar to the one described above, but amplified, was provided to us by DAVID SULLIVAN, Syracuse University.

The *D. erecta* and *D. virilis* libraries were amplified libraries constructed by DONAL HICKEY, University of Ottawa, and RONALD BLACKMAN, Harvard University, respectively, using the bacteriophage λEMBL 3 vector (FRISCHAUFF *et al.* 1983).

Screening of the recombinant λ phage libraries was done essentially as described in MANIATIS, FRITSCH and SAMBROOK (1982). For each library, 15 plates containing 10^4 plaque-forming units (pfu)/plate were prepared, and plaques lifted onto nitrocellulose filters (Millipore) and prepared for hybridization according to the method of BENTON and DAVIS (1977). Probes were made by random-primed labelling of agarose gel-purified restriction fragments of cloned *D. melanogaster* transformer gene sequences using the Boehringer Mannheim Random Primed Labelling Kit. The *tra* clones used as probes in this study included pIBItra3.9R, a 3.9-kb *Eco*RI fragment, derived from a recombinant λ phage isolated from the Canton-S genomic library of MANIATIS *et al.* (1978), cloned into the plasmid vector pIBI76 (International Biotechnologies, Inc.), and two cDNA clones, pIBI6.5.1 and pIBItra3.1P containing inserts derived from recombinant phage isolated from embryonic and pupal cDNA libraries, respectively, of POOLE *et al.* (1985), and subcloned into the pIBI76 vector. pIBI6.5.1 represents a full-length female-specific *tra* transcript and pIBItra3.1P represents a full length non-sex-specific transcript. See MCKEOWN, BELOTE and BAKER (1987) and BOGGS *et al.* (1987) for the molecular map and sequences of the *D. melanogaster tra* gene region and its transcripts. Hybridization was carried out overnight at 42° in 5 × SSPE, 0.1% sodium dodecyl sulfate (SDS), 1 × Denhardt's solution, 0.1 mg/ml sheared salmon sperm DNA, and 50% formamide (for the normal stringency hybridizations), or either 29% or 39% formamide (for the reduced stringency hybridizations). Filters were washed at 50° in three changes of wash buffer (0.1 × SSPE, 0.1% SDS for normal stringency hybridizations, or 0.5 × SSPE, 0.1% SDS for reduced stringency hybridizations). Identification of positive clones, plaque purification, preparation of phage DNA, DNA fragment isolation, Southern blot analysis, identification of cross-hybridizing fragments, subcloning of the restriction fragments into the plasmid vector pIBI76 and isolation of plasmid DNA was carried out using the protocols described in MANIATIS, FRITSCH and SAMBROOK (1982).

Southern blot analysis of fly DNA: Drosophila genomic DNA was isolated as described in BENDER, SPIERER and HOGNESS (1983). Restriction digestion, agarose gel electrophoresis and transfer of the DNA to GeneScreen Plus membrane (NEN, Du Pont) was carried out as described in MANIATIS, FRITSCH and SAMBROOK (1982). Hybridization and wash conditions were the same as those described above for screening of the libraries.

DNA sequencing: For each of the cloned *tra*-homologous genes, a series of nested deletions were generated by *Exo*III digestion according to the method of HENIKOFF (1987) using the Erase-A-Base kit from Promega. Plasmid DNA was purified by the ammonium acetate method of MORELLE (1989). Double stranded DNA templates were sequenced by the dideoxynucleotide chain termination method of SANGER, NICKLEN and COULSON (1977) using [^{35}S]dATP (New England Nuclear). Primers for the sequencing reactions

were either the T7 or SP6 promoter oligonucleotides purchased from Promega. Reaction products were electrophoresed according to the electrolyte gradient method of SHEEN and SEED (1988). DNA sequence data was managed using the programs of DNA STAR (Madison, Wisconsin). Comparisions of nucleotide sequences were performed by the algorithm of WILBUR and LIPMAN (1983).

***In situ* hybridization:** Salivary glands from late third instar larvae were dissected in 0.7% saline, incubated for 10 sec in 45% acetic acid, and placed in a drop of 1 part lactic acid, 2 parts H_2O, and 1 part acetic acid for 5 min. A siliconized coverslip was then applied and the glands were squashed. After freezing the slide in liquid nitrogen, the coverslip was removed and the chromosomes dehydrated through an ethanol series. Hybridization of the chromosomes was carried out according to the procedure of PLILEY, FARMER and JEFFERY (1986). Probes were prepared for *in situ* hybridization by random primed labelling of gel purified restriction fragments. The modified deoxynucleotide used was either biotinylated dUTP (Bethesda Research Laboratories) or digoxigenin-substituted dUTP (Boehringer Mannheim). Signal detection was done using a Detek-I-hrp Kit (ENZO) following the procedure of ASHBURNER (1989), for the biotinylated probes, or using the Genius Kit (Boehringer Mannheim) detection procedure for the digoxigenin-labeled probes (TAUTZ and PFEIFLE 1989).

***P* element-mediated germline transformation:** *D. melanogaster* embryos were collected within 30 min of egg deposition, dechorionated manually and injected with DNA as described in SPRADLING (1986). The injected DNA was cesium banded plasmid DNA consisting of a 2.8 *Pst*I fragment containing the *D. virilis tra*-homologous gene region inserted into the pW8 transformation vector of KLEMENZ, WEBER and GEHRING (1987) which carries the *white*$^+$ (w^+) eye color gene as a marker. Recipient flies were from the *w*; $ry^{506}P[ry^+\Delta 2\text{-}3](99B)$ strain of ROBERTSON *et al.* (1988). Since this strain carries a genomic source of transposase, no helper plasmid was used. G_0 survivors were mated to *w; TM3,ri p*^{p} *sep Sb bx*^{34e} *e*^{s}*/TM6B, Hu Tb e ca* flies, and progeny scored for pigmented eyes to identify transformants. Such transformants were then crossed to *w; Df(3L)st*^{j7}*, Ki roe p*^{p}*/TM6B, Hu Tb e ca* mates and the *Df(3L) st*^{j7}*, Ki roe p*^{p}*/Balancer* progeny with pigmented eyes were mated to *w; th st tra cp in ri p*^{p}*/TM6B, Hu Tb e ca* or *w; th st tra cp in ri p*^{p}*/TM3,ri p*^{p} *sep Sb bx*^{34e} *e*^{s}. From this and subsequent crosses, flies were generated that were hemizygous or homozygous for the *tra* mutation and carried one or two copies of the *virilis tra*-homologous gene. In all of these crosses, males carried the B^s *Y* chromosome that allowed us to distinguish *XX* and *XY* individuals on the basis of the dominant phenotype associated with B^S. Flies of interest were mounted between coverslips according to the method of SZABAD (1978) and examined under bright field optics using a Zeiss Axioplan microscope.

Transcript analysis: Preparation of poly(A$^+$) RNA and Northern blot analyses were carried out as described in MCKEOWN, BELOTE and BAKER (1987). DNA probes were made by random primed labelling of gel purified restriction fragments as described above. The procedure used for the generation of strand-specific RNA probes (*i.e.*, riboprobes) was a modification of MELTON *et al.* (1984). Template DNA was linearized by restriction digestion and T7 RNA polymerase was used to create the run-off transcripts from the T7 promoter of plasmid vector pIBI76. The [α-^{32}P]CTP (NEN, Du Pont)-labeled transcripts were electrophoresed through a 5% native polyacrylimide gel in 0.5 × TBE buffer (89 mM Tris-base, 89 mM boric acid, 2 mM Na_2EDTA). After electrophoresis, the gel was exposed to Kodak X-OMAT AR film for 15 min at room temperature. After development, the film was aligned with the gel with the positive exposure denoting the position of the riboprobe transcript. This region was then cut out of the gel, placed in a 1.5-ml Eppendorf tube containing 0.38 ml 2 M NH_4OAc, 0.1% SDS and rocked for 2 hr at 37°. The supernatant was then precipitated with 10 μg yeast tRNA and 2.5 × volume absolute ethanol for 15 min at −70°. The RNA pellet was resuspended in 25 μl of 0.2 M NaPIPES, pH 6.4, 2 mM NaCl, 5 mM Na_2EDTA.

For the amplification of reverse-transcribed RNA by polymerase chain reaction poly(A$^+$) RNA was prepared as follows. Flies (0.5–1.0 g) were ground to a powder in liquid nitrogen and added to 4.0 ml of 5 M guanidine isothiocyanate, 10 mM EDTA, 50 mM HEPES, 5% β- mercaptoethanol and homogenized in a Dounce tissue solubilizer. After brief centrifugation at 4° the supernatant was transferred to 1 ml of 5.7 M CsCl in HE buffer (10 mM HEPES, 1 mM EDTA). Dry Sarkosyl was added to 4% and the samples heated to 60° for 5 min. This solution was then layered over 1 ml of 5.7 M CsCl and centrifuged at 20° for 36 hr at 36,000 rpm in an AH650 rotor (Du Pont) to pellet the RNA. The RNA pellet was resuspended in 8 M urea in HE and then extracted several times with an equal mixture of phenol:chloroform. The RNA was then precipitated with a 0.1 × volume of 2.5 M NaOAc and a 2 × volume of absolute ethanol at −20°. The RNA was resuspended in diethylpyrocarbonate-treated H_2O, and poly(A$^+$) RNA was selected by passage over an oligo-dT-cellulose column as described in MANIATIS, FRITSCH and SAMBROOK (1982).

Polymerase chain reaction amplification of cDNA (cDNA-PCR) was carried out as described by KAWASAKI and WANG (1989). cDNA was synthesized in a 20-μl total volume reaction mixture of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 1 mM each dNTP, 1 unit/μl RNasin (BRL), 100 pmol random hexamers (Pharmacia), 0.5 μg poly(A$^+$) RNA isolated from *D. virilis* adult males or females, and 200 units MuLV reverse transcriptase (BRL). Samples were incubated 10 min at room temperature, then 1 hr at 42°, and reactions stopped by heating to 95° for 5 min. Polymerase chain reaction amplification was carried out using the GenAmp DNA Amplification Reagent Kit (Perkin Elmer Cetus). Reaction mixtures consisted of 10 mM each dNTP, 01 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 1 μl of cDNA template, prepared as above, 0.5 unit AmpliTaq polymerase (Perkin Elmer Cetus), and 20 μM of each primer, in a total volume of 20 μl. The primers used for the amplification of *D. virilis* cDNAs were: primer Dvir2–5′ GACGCACATAGCTTCTGG 3′, and primer Dvir5′-5′ ATGGACGCCGACAGCAG 3′. Primers were synthesized by the Syracuse University DNA/Protein Core Facility using a Coder 300 DNA Synthesizer (Du Pont). The polymerase chain reaction was carried out using a Techne thermal cycler for 50 cycles, each cycle consisting of 1 min at 94°, 30 sec at 48°, and 2 min at 72°. PCR products were electrophoresed through 2% agarose gels, transferred to Gene Screen membrane filters (Du Pont), and probed with random primed ^{32}P-labeled restriction fragments of the *D. virilis tra*-homologous gene. Autoradiograms were developed after about 1 or 2 weeks of exposure.

RESULTS

Isolation of *tra*-homologous genes from *D. simulans, D. erecta, D. hydei* and *D. virilis*: As a preliminary step toward the isolation of *tra*-homologous

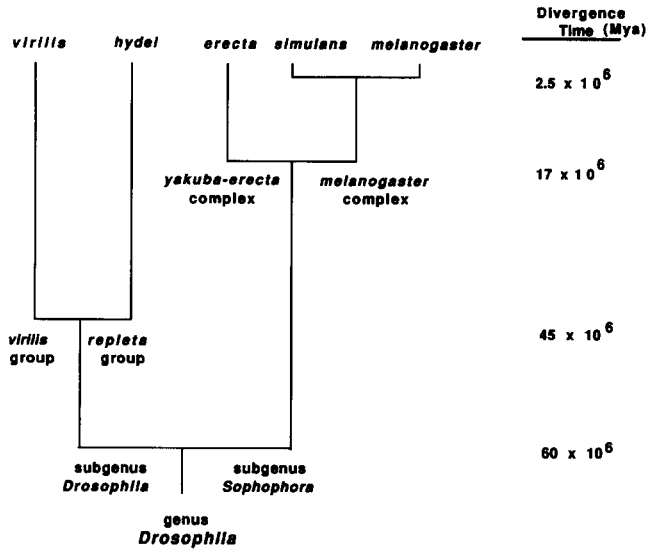

FIGURE 1.—Phylogenetic relationships of the Drosophila species used in this study. Estimates of divergence times from THROCKMORTON (1975), BEVERLEY and WILSON (1984), and LACHAISE *et al.* (1988). Mya, millions of years ago.

genes, we carried out genomic Southern blot analyses on DNA isolated from four Drosophila species representing various degrees of divergence from *D. melanogaster*. The Drosophila species selected for our study included two species from the *melanogaster* subgroup of the Sophophora subgenus, *D. simulans* and *D. erecta*, and two species from the Drosophila subgenus: *D. hydei* of the *repleta* radiation, and *D. virilis* of the *virilis* radiation. Figure 1 shows the phylogenetic relationships, and the estimated divergence times, among these species. Southern hybridizations were carried out under either normal or reduced stringency hybridization conditions, using as probes the cloned *tra* gene region from *D. melanogaster* (see MATERIALS AND METHODS for details). Hybridization of the *tra* cDNA probe pIBItra3.1P under normal stringency conditions allowed the detection of a single cross-hybridizing region in the Sophophoran species, while no hybridization signal was detected in the more distantly related species (data not shown). Reducing the stringency of the hybridization conditions, by lowering the amount of formamide in the hybridization solution from 50% to 29%, allowed us to detect cross-hybridizing regions in all of the Drosophila species. For all of the species there were more than one cross-hybridizing region detected under these conditions, but the one strongest hybridization signal was presumed to be the best candidate for the *tra*-homologous gene region, with the weaker cross-hybridizing regions possibly being *tra* "cognates," DNA regions that share sequence similarity with *tra* but that do not represent the *tra*-homologous gene.

Having thus defined the hybridization conditions that should allow the isolation of *tra*-homologous genes from the other Drosophila species, we next screened genomic λ phage libraries of these four species under the appropriate conditions, using the *tra* region genomic probe pIBItra3.9R. For each library screened, the recombinant λ phage that were isolated by molecular cloning were then restriction mapped and aligned. The phage DNAs were then subjected to Southern blot analysis using a *melanogaster tra* cDNA clone as probe to further delimit the *tra* cross-hybridizing regions. For each species, two or more overlapping recombinant phage were isolated. Figure 2 shows the restriction maps of representative phage corresponding to the *tra*-homologous region of each species.

DNA sequences of the *tra*-homologous genes of the Sophophoran species, *D. simulans* and *D. erecta*: Because the *simulans* and *erecta tra*-homologous sequences were isolated by screening under normal stringency conditions, and because only a single genomic region is detected under those same conditions in the Southern blot analysis, it is assumed that the recombinant phage that were isolated in these screens contain the *tra*-homologous genes. Southern blot analysis of the recombinant λ phage clones λDs24, from *D. simulans*, and λ De108.3, from *D. erecta*, using the *D. melanogaster tra* cDNA clone pIBItraP3.1 as a probe, revealed that the *tra*-homologous genes were contained within a 2.0-kb *Eco*RI/*Bam*HI fragment of the *simulans* phage, and within a 3.0-kb *Bam*HI fragment of the *erecta* phage. These fragments were therefore subcloned into plasmid vectors and a portion, corresponding to the *tra* transcribed region, of each was sequenced. Figure 3 shows the nucleotide sequence comparisons of the Sophophoran species examined.

The structure of the *tra* gene has been extensively characterized in *D. melanogaster* (MCKEOWN, BELOTE and BAKER 1987; BOGGS *et al.* 1987). The *tra* transcription unit is relatively small, about 1100 base pairs in length, with two introns. The first intron exhibits sex-specific alternative RNA splicing. One type of splice is seen in both sexes and involves the removal of a 73-nucleotide long intron, from the 5′ splice site to the non-sex-specific 3′ splice site shown in Figure 3. The RNA resulting from this splice has no long ORF, and is assumed to be nonfunctional due to the production of a truncated *tra* protein. In females, about half of the transcripts undergo an alternative splice in which a different 3′ splice site is used (labeled female-specific 3′ splice site in Figure 3). The RNA resulting from this splice has a long ORF encoding the functional *tra* protein. Thus, this sex-specific splicing event provides the basis for the female-limited function of the tra gene.

The nucleotide sequences of *D. simulans* and *D. erecta* can be aligned with that of *D. melanogaster* to give the same gene structure. The initiation and ter-

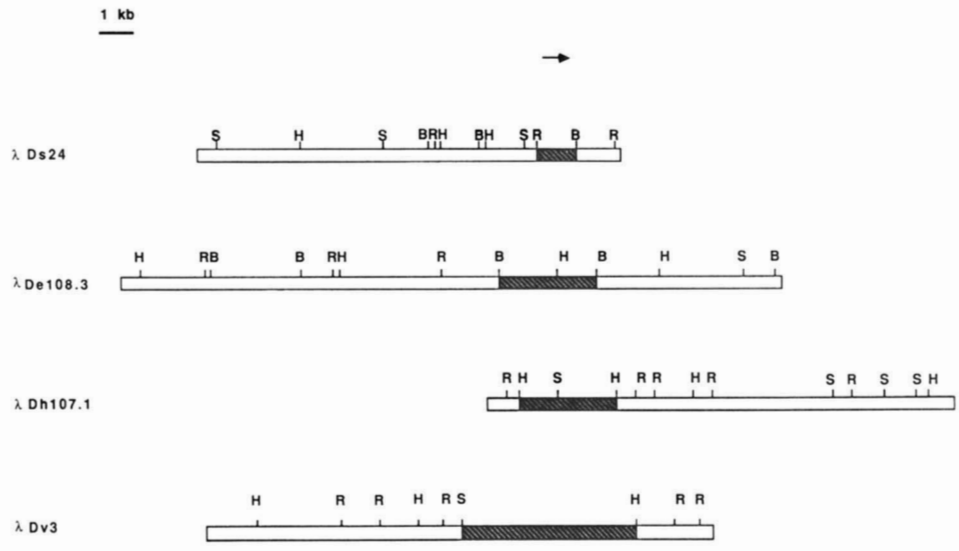

FIGURE 2.—Restriction maps of representative recombinant λ phage containing the *tra*-homologous regions of *D. simulans* (λ Ds24), *D. erecta* (λ De108.3), *D. hydei* (λ Dh107.1), and *D. virilis* (λ Dv3). Bacteriophage vector arms are not shown. Cross-hatched regions represent restriction fragments that hybridize to the *D. melanogaster tra* cDNA pIBItra3.1P probe. The arrow shows the inferred position and direction of transcription of the *tra*-homologous genes, as inferred from using strand-specific *tra* probes on Northern blots, and from subsequent DNA sequence analysis. B = *Bam*HI; H = *Hin*dIII; R = *Eco*RI; S = *Sal*I.

mination codons are in similar positions, although the termination codon is UAA in *D. erecta*, and UGA in *D. simulans* and *D. melanogaster*. All three species have two in-frame methionine codons at the beginning of the ORF. It is not known which of these is used as the initiation codon *in vivo*. The exon/intron arrangement has been preserved: the splice junction sequences are conserved, and they occur in the same reading frame positions in all three species. While the occurrence of the alternative splice of intron 1 has not been directly demonstrated in these other species, the conservation of the 5′ and 3′ splice junctions, and the conserved female specific ORF, strongly suggest that the expression of *tra* in these species is very similar to that in *D. melanogaster*. Northern blot analysis of *tra* expression in *D. erecta* shows that there is a sex-specific pattern of *tra* transcripts indistinguishable from that seen in *D. melanogaster* (data not shown). Additional Northern blot analyses using strand specific probes, also show that the arrangement of the flanking transcription units is similar to that in *D. melanogaster*; *i.e.*, the 5′ flanking gene is transcribed from the same strand as *tra* and has its 3′ end near the start of the *tra* transcription unit, and the 3′ neighboring gene is transcribed from the opposite strand as *tra* and has its 3′ end very near the 3′ end of the *tra* gene (data not shown). In *D. melanogaster*, the 3′ neighboring gene, *l(3)73Ah*, corresponding to a complementation group defined by a late pupal lethal mutation, has been shown to actually overlap the *tra* transcription unit by about 75 bases at the 3′ ends of both genes (BOGGS *et al.*, 1987). Given the degree of sequence conservation in the 3′ untranslated region of *tra* (*e.g.*, 59 continuous bases identical among all three species, the longest stretch of nucleotide sequence identity in the entire sequenced region), it is likely that this overlapping gene organization is present in these other species, as well.

Table 1 shows the nucleotide sequence conservation among these three species. As expected from the known phylogenetic relationships, the degree of similarity is higher when *melanogaster* and *simulans* are compared (93.6% identity) than when either of these species is compared to *erecta* (86.0% and 85.6% identity). The higher degree of conservation of sequences in intron 1 than in intron 2 is not unexpected, since intron 1 exhibits sex-specific alternative splicing while the splicing of intron 2 is presumably not regulated. In *D. melanogaster* it has been shown that the sex-specific splicing of intron 1 is controlled by sequences contained within the intron itself (SOSNOWSKI, BELOTE and MCKEOWN 1989), and these regulatory sequences are likely conserved. The close relationship between these members of the melanogaster subgroup, however, precludes using these comparisons to accurately delimit any well defined region of the intron that might be implicated in this regulation, since not enough time has passed for the nonimportant sequences to have diverged.

One obvious difference between the coding region of *tra* in *D. melanogaster* and that of the other two species is the presence of a 39-base pair sequence in the central exon of *D. melanogaster* that is missing in *D. simulans* and *D. erecta*. Given the phylogenetic relationships of these three species, the most parsimonious explanation is that there has been a duplication of this sequence in the lineage leading to *D. melanogaster* rather than two independent deletions

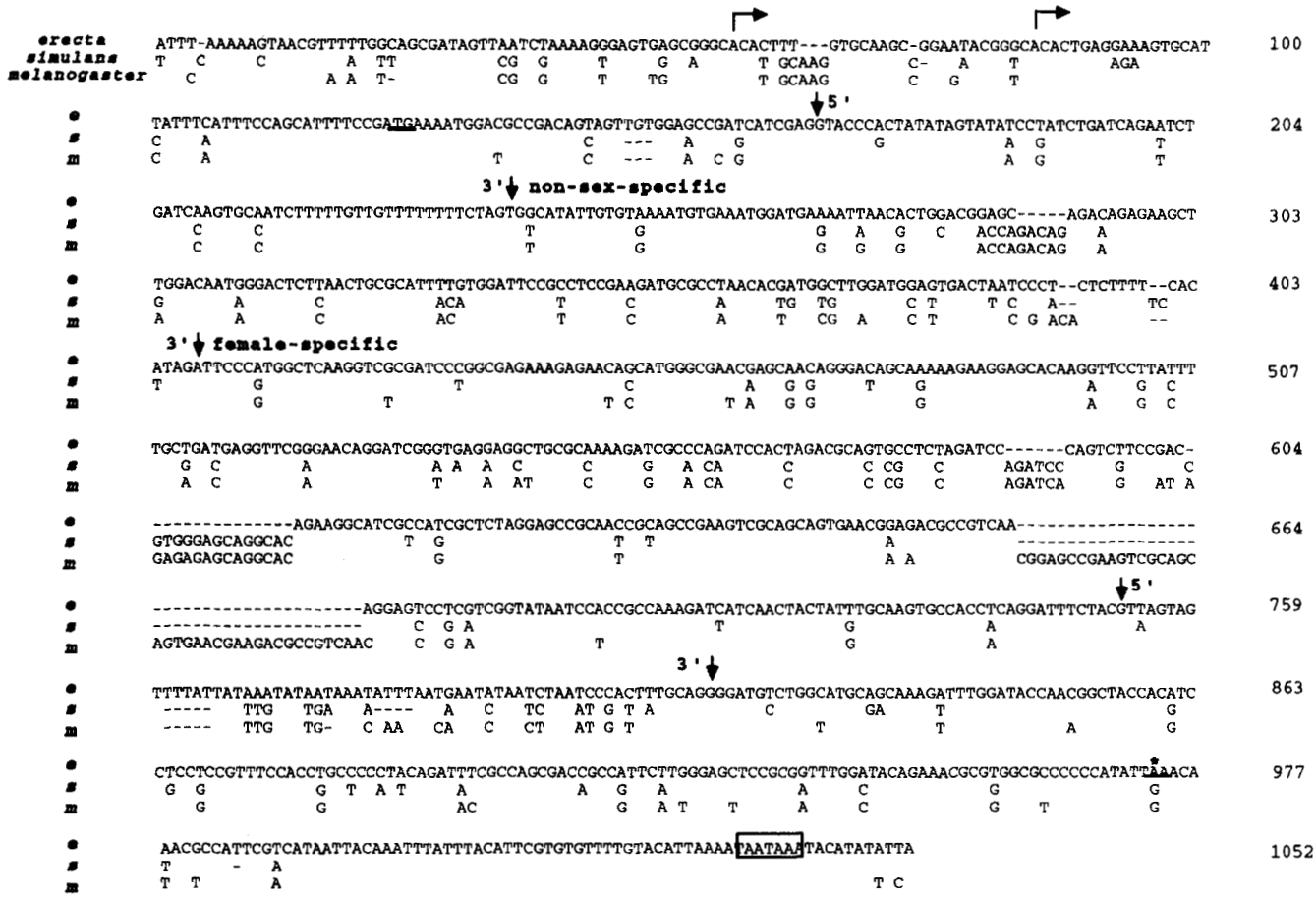

FIGURE 3.—Comparison of the DNA sequences of the *tra* genes from the Sophophoran subgenus species *D. erecta*, *D. simulans* and *D. melanogaster*. The complete nucleotide sequence is shown only for *D. erecta*. For *D. melanogaster* [taken from BOGGS *et al.* (1987)] and *D. simulans*, only those bases that differ from the *D. erecta* sequence are shown. Dashes represent deletions in the sequence relative to the other species. The numbers refer to the nucleotides in the *D. erecta* sequence. The two rightward facing arrows show the sites where transcription is initiated in *D. melanogaster* (BOGGS *et al.* 1987; MCKEOWN, BELOTE and BOGGS 1988). The other arrows show the 5′ and 3′ splice sites for the two introns. The first intron shows sex-specific alternative splicing, with the more distal 3′ splice site being used only in females (BOGGS *et al.* 1987). The underlined sequences represent the translation start and translation stop (with asterisk) sites. The boxed sequence is the consensus polyadenylation signal. These sequences have been aligned using the algorithm of WILBUR and LIPMAN (1983).

occurring in the *D. simulans* and *D. erecta* lineages. This is supported by the fact that in *D. melanogaster*, these extra 39 bases exist as an almost perfect tandem repeat (37/39 identity) of the adjacent sequence. This duplication occurs in the very arginine-serine-rich region of the *tra* gene (see below, Figure 8), and its presence in either one or two copies is presumably not critical for the functioning of the *tra* protein.

Other smaller deletions/insertions also are seen when the coding regions of these three species are aligned (Figures 3 and 8). For example, *D. erecta* has an extra three nucleotides, encoding a cysteine, in exon 1 that are not seen in *D. melanogaster* and *D. simulans*. In exon 2 there are two sites where an extra six nucleotides, encoding R S, and an extra 15 nucleotides, encoding R E/G S R H, are present in *D. melanogaster* and *D. simulans*, but not in *D. erecta*. These differences among the coding regions yield different size *tra* proteins in the three species: 197, 184 and 178 amino acids, for *D. melanogaster*, *D. simulans* and *D. erecta*, respectively.

Because of the relatively close relationship among these three species, information about what regions of the *tra* gene are conserved and what regions are not conserved is limited. To better examine this, comparisons of these Sophophoran species with the Drosophila group species, *D. hydei* and *D. virilis*, was carried out.

Characterization of the *tra*-homologous genes of *D. hydei* and *D. virilis*: The reduced stringency conditions that were used for the isolation of the *tra*-homologous regions of *D. hydei* and *D. virilis* allow cross-hybridization of more than one genomic region, although in each of these cases there is one predominant cross-hybridization signal that is noticeably stronger than the others. Characterization of the three recombinant phage isolated from the *virilis* library showed that they all corresponded to the same genomic region. Restriction mapping of these clones yielded a map that matched our expectations based on the pattern of the most strongly hybridizing restric-

TABLE 1

Interspecific comparisons of the *transformer* gene

Gene region		Percentage				
		melanogaster	*simulans*	*erecta*	*hydei*	*virilis*
Entire gene[a,b]	*melanogaster*	—	93.6	86.0	62.5	63.3
	simulans	—	—	85.6	58.7	62.1
	erecta	—	—	—	57.9	62.1
	hydei	—	—	—	—	73.5
Coding region	*melanogaster*	—	93.1	87.7	60.6	63.5
	simulans	—	—	88.0	59.9	61.2
	erecta	—	—	—	59.8	65.0
	hydei	—	—	—	—	72.5
Intron 1	*melanogaster*	—	94.3	86.0	54.7	54.5
	simulans	—	—	84.4	52.0	55.6
	erecta	—	—	—	56.7	57.6
	hydei	—	—	—	—	73.3
Intron 2	*melanogaster*	—	92.7	67.8	i.h.[c]	i.h.
	simulans	—	—	i.h.	i.h.	i.h.
	erecta	—	—	—	i.h.	62.5
	hydei	—	—	—	—	i.h.
Protein sequence[d]	*melanogaster*	—	86.3	77.3	31.3	35.8
	simulans	—	—	81.1	31.8	34.1
	erecta	—	—	—	31.7	42.0
	hydei	—	—	—	—	62.6

[a] All DNA sequence comparisons were done using the algorithm of WILBUR and LIPMAN (1983). Percentages refer to the nucleotide sequence similarity index.

[b] The entire gene includes coding regions, the two introns, 100 base pairs of sequence 3′ to the end of the coding region, and 100–150 base pairs of DNA 5′ to the start of the coding region.

[c] i.h. indicates that there was insufficient homology for a meaningful comparison to be made.

[d] Protein sequence comparisons were done using the algorithm of NEEDLEMAN and WUNSCH (1970). Percentages refer to the amount of amino acid identity in the aligned sequences.

tion fragments in reduced stringency genomic Southern blots. This suggests that these recombinant phage carry DNA corresponding to the *tra*-homologous region and that they do not represent *tra* "cognate" loci. The recombinant phage isolated from the *hydei* library could be grouped into five classes based on their restriction maps and each class cross-hybridized to the *tra* probe with differing strengths. When the stringency conditions were increased by raising the formamide concentration to 39%, only the most strongly hybridizing class of clones still hybridized to the probe. The class of clones that cross-hybridized most strongly to *tra*, which included clone λ Dh107.1 (see Figure 2), had a restriction map that matched the pattern seen for the most strongly hybridizing fragments seen in genomic Southern blots, and is presumed to be the *tra*-homologous region. The other classes of weakly cross-hybridizing clones could be related to the *tra* cognate loci that have been characterized in *D. melanogaster* and that show only limited sequence similarity to *tra* DNA (J. R. EBERLE and J. M. BELOTE, unpublished).

Chromosome localization of the *tra*-homologous genes of *D. virilis* and *D. hydei*: Since the putative *tra*-homologous genes from *D. hydei* and *D. virilis* were isolated using reduced stringency hybridization conditions, and since such conditions permit cross-hybridization to more than one gene region in genomic Southern blots, it is conceivable that the primary candidates for the *tra*-homologous genes were not really the true *tra* homologs. To provide supporting evidence for our assumptions that the λ Dh107.1 clone and the λ Dv3 clone represent the *tra*-homologous regions from *D. hydei* and *D. virilis*, respectively, we mapped their chromosomal positions by *in situ* hybridization to salivary gland chromosomes. The chromosomal position of the *tra* gene in *D. melanogaster* is 73A9.10 on the left arm of chromosome three (BUTLER *et al.* 1986; MCKEOWN, BELOTE and BAKER 1987). If chromosome linkage arrangements are conserved among these three species, as is suggested by several studies involving *in situ* hybridization of cloned genes to polytene chromosomes of related species (BROCK and ROBERTS 1983; LOUKAS and KAFATOS 1986; WHITING *et al.* 1989; TONZETICH, HAYASHI and GRIGLIATTI 1990), then we would expect the *tra*-homologous gene of *D. virilis* to map to chromosome three (ALEXANDER 1976) and that of *D. hydei* to map to chromosome four [using BERENDES' (1963) chromosome numbering system].

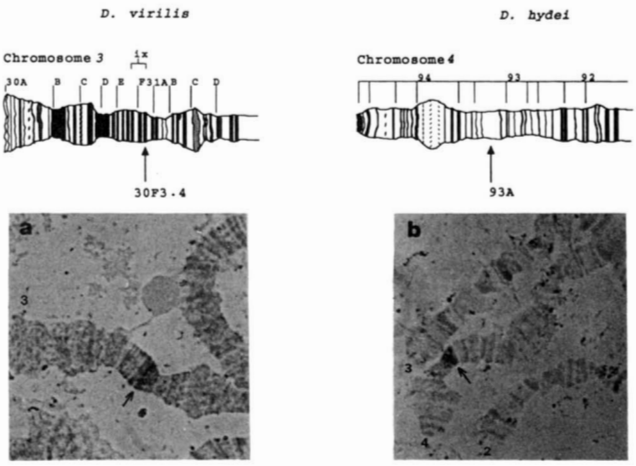

FIGURE 4.—*In situ* hybridization of *tra*-homologous genomic clones to third instar larval salivary gland chromosomes from (a) *D. virilis* and (b) *D. hydei*. Sketches above the photomicrographs show the map positions as determined by the hybridization signal. The *D. virilis* map, showing the tip of chromosome *3*, is drawn from the photomicrographs of GUBENKO and EVGEN'EV (1984), and the *D. hydei* map, showing the tip of chromosome *4*, is redrawn from BERENDES (1963) and ANANIEV AND BARSKY (1982). The approximate map position of the *D. virilis intersex* gene, *ix*, as determined by recombinational mapping experiments (LEBEDEFF 1939; GUBENKO and EVGEN'EV 1984), is also indicated.

When a 2.8-kb *Pst*1 fragment containing the *tra* cross-hybridizing sequences from *D. virilis* was used as probe to *D. virilis* salivary gland chromosomes, hybridization was detected at a single site, band 30F3.4 of chromosome *3* on the map of GUBENKO and EVGEN'EV (1984) (Figure 4a). This site is well separated from the site of hybridization of the probe 521 used by WHITING *et al.* (1989) in their comparisons of chromosome homologies between *D. melanogaster* and *D. virilis*. In that study, they found that this probe, which hybridizes to chromosome region 73DE in *D. melanogaster* hybridizes to region 32D on chromosome *3* of *D. virilis*. Similarly, NEUFELD, CARTHEW and RUBIN (1991) in their interspecific study of the *sina* locus mapped this gene to 73D in *D. melanogaster* and 32C-D in *D. virilis*. Given that the *tra* gene of *D. melanogaster* is at 73A9.10, only about two letter divisions (corresponding to about six dark bands) away from 73C-D, our observation that the *tra*-homologous gene of *D. virilis* at 30F is at least 20 dark bands away from the sites of the 521 and *sina* loci at 32C-D demonstrates that during the evolutionary histories of these two species there have been chromosome rearrangements that have changed the colinearity of the chromosome banding maps in that region. It is interesting to note that the site of the *tra*-homologous gene coincides with the published map position of the *intersex* (*ix*) locus of *D. virilis* (LEBEDEFF 1934, 1938, 1939; GUBENKO and EVGEN'EV 1984). Mutant alleles of this gene have a phenotype similar to *tra* mutants in *D. melanogaster*: chromosomally female *D. virilis* individuals homozygous for *ix* mutations are shifted phenotypically toward maleness while chromosomally male flies are unaffected. In fact, the strong alleles of *ix* transform *XX* flies into phenotypically normal-looking males, just like *tra* does in *D. melanogaster*. Given these similarities in phenotype and map position, it is a reasonable possibility that the *ix* gene of *D. virilis* represents the *D. melanogaster tra*-homologous gene.

Hybridization of a probe containing the putative *D. hydei tra*-homologous gene (*i.e.*, p3.0H1.7S, a subclone of the 3.0-kb *Hin*dIII fragment from λ Dh107.1) to *D. hydei* salivary gland chromosomes resulted in a single hybridization signal at position 93A on chromosome *4* (Figure 4b). As with the *D. virilis tra*-homologous clone, the chromosome linkage of the *D. hydei* clone is consistent with our expectations based on the assumption that it represents the *tra*-homologous gene.

While these experiments do not prove that these clones represent the *tra*-homologous genes, they do provide good supporting evidence that this is the case, and taken together with the results presented below, it is almost certain that these genes from *D. hydei* and *D. virilis* represent the *tra*-homologous genes.

***P* element-mediated germline transformation of *D. virilis tra* DNA into *D. melanogaster*:** As an additional test to confirm that the *tra*-cross-hybridizing clones from the Drosophila subgroup species represent the *tra*-homologous gene, we used *P* element germline transformation methods to introduce the putative *tra* gene from *D. virilis* into *D. melanogaster*, and then tested the transduced gene for its ability to provide tra^+ function in *tra* mutant flies. The 2.8-kb *Pst*I fragment containing the *tra* cross-hybridizing region from *D. virilis* was subcloned into the *P* element-transformation vector pW8 (KLEMENZ, WEBER and GEHRING 1987) and introduced into the *D. melanogaster* genome by germline transformation. One transgenic line was obtained, which was then crossed to the appropriate stocks to construct chromosomally female flies that were homozygous for the *tra* mutant allele. This allele is a deletion of the *tra* coding region (MCKEOWN, BOGGS and BELOTE 1988). These transgenic flies should be phenotypically male unless the transduced *virilis tra* DNA is able to supply tra^+ activity, and thus shift the flies' phenotypes towards femaleness. As shown in Figure 5, the *D. virilis tra*-homologous transgene is able to supply tra^+ function. This is most evident in the genital region of the adult, where there is a marked reduction in the development of male genital structures, such as parts of the penis apparatus and the claspers, and the appearance of female-specific structures, such as vaginal plates and thorn bristles. The shift toward femaleness is also apparent in the altered pigmentation patterns of the abdominal tergites, as well as in the structure of the anal plates. Additional evidence of a shift toward

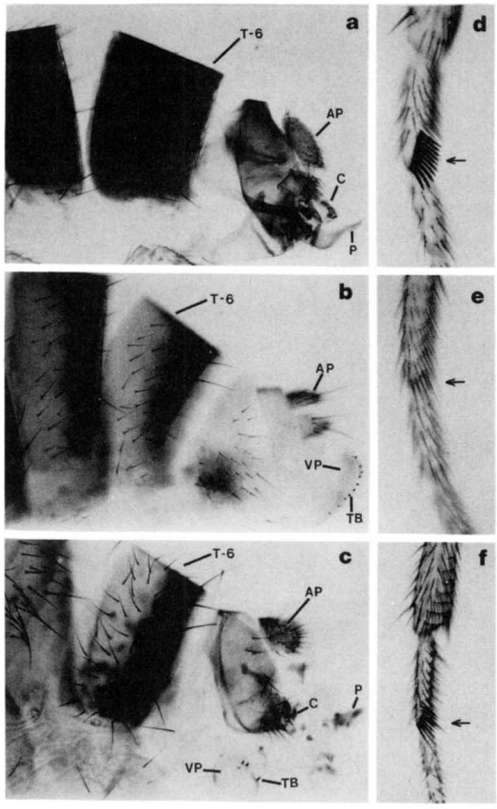

FIGURE 5.—Partial rescue of the *tra* mutant phenotype in transgenic *D. melanogaster* carrying the *tra* homologous gene from *D. virilis*. Terminalia (a) and basitarsus of the foreleg (d) of an adult *XX*; *tra/tra* fly showing normal male sexual phenotype. Terminalia (b) and basitarsus of the foreleg (e) of an adult *XX*; tra^+/tra^+ fly showing normal female sexual phenotype. Terminalia (c) and basitarsus of the foreleg (f) of a transgenic *XX*; tra/*tra* fly that is homozygous for an *X* chromosome carrying a transduced copy of *pW8Dv2.8P*, a *P* element vector containing the *tra* homologous gene region from *D. virilis*. AP, anal plates; C, clasper; P, penis; TB, thorn bristles; T-6, sixth abdominal tergite; VP, vaginal plate. The arrows in d, e, and f point to the sex comb regions.

female development is seen in the sex-comb region of the foreleg, where the transgenic flies have bristles resembling what is seen in females, rather than the thick, blunt bristles characteristic of the male sex comb. While the shift toward femaleness is not complete in this transgenic line, it has been observed that even the *D. melanogaster tra* gene is frequently unable to completely rescue the *tra* mutation when transduced into the genome by *P* element-mediated germline transformation. In fact, in most of the reported experiments where transgenic flies carrying transduced tra^+ genes have been produced, only a minority of the lines showed complete phenotypic rescue of tra mutants (BUTLER *et al.* 1986; MCKEOWN, BELOTE and BAKER 1987; FENG, SCHIFF and CAVENER 1991; J. BELOTE and M. MCKEOWN, unpublished observations). This lack of complete rescue in these experiments is likely due to the *tra* transgene being especially sensitive to chromosomal position effects. Regardless of the reason for the failure of complete female development of *XX; tra/tra* flies in our transgenic line, the observation that there is substantial female differentiation shows that the *tra*-homologous gene from *D. virilis* is able to function quite well in *D. melanogaster*. The fact that chromosomal males carrying this transduced *virilis tra* gene are normal and fertile further suggests that the sex-specific alternative splicing of the first intron is being properly carried out, since other studies on the *D. melanogaster tra* gene show that the loss of the non-sex-specific splice site most frequently leads to default splicing of the *tra* pre-mRNA at the female-specific 3′ splice site in *XY* individuals, leading to male sterility and/or sex transformation toward femaleness (SOSNOWSKI, BELOTE and MCKEOWN 1989; B. A. SOSNOWSKI, R. T. BOGGS and M. MCKEOWN, personal communication).

DNA sequences of the *tra*-homologous genes of the Drosophila species, *D. hydei* and *D. virilis*: Southern blot analysis of the recombinant λ phage clones λDh107.1, from *D. hydei*, and λDv3, from *D. virilis*, using the *D.melanogaster tra* cDNA clone pIBItraP3.1 (BOGGS *et al.* 1987) as a probe indicated that the *tra*-homologous genes were located within a 3.0-kb *Hin*dIII fragment of the *hydei* phage and within a 2.8-kb *Pst*I fragment of the *virilis* phage. These fragments were therefore subcloned into plasmid vectors and a portion, corresponding to the *tra* transcribed region, of each was sequenced. Figure 6 shows the nucleotide sequences of these two species, and Table 1 indicates the degree of similarity among all of the species examined. The alignment of the sequences with each other, and with that of *D. melanogaster*, suggests that the approximate size and structure of the *tra* gene is fairly well conserved among all of the species examined. The protein-coding regions and the positions of the introns were deduced by comparison with the *melanogaster* sequence. Unlike the case with the Sophophoran species, there is only one methionine codon at the beginning of the ORF in these two Drosophila species and it corresponds to the second ATG seen in the Sophophoran species. The exon/intron structure is similar in all species examined: the splice junctions of both introns are conserved at the 5′ and 3′ ends, and the introns interrupt the *tra* reading frame in similar ways. Intron 1 occurs between the G and the A of a conserved GAT aspartate codon, and intron 2 occurs just after a conserved tyrosine codon TAC, in all five species.

Northern blot analysis of *D. hydei* poly(A^+) RNA shows a sex-specific pattern of *tra* transcripts similar to what is seen in *D. melanogaster* (data not shown). In order to confirm that there is sex-specific alternative splicing of the first intron in flies from the Drosophila subgenus, PCR primers flanking intron 1 were synthesized, and a cDNA-PCR analysis was carried out on RNA isolated from male and female *D. virilis*

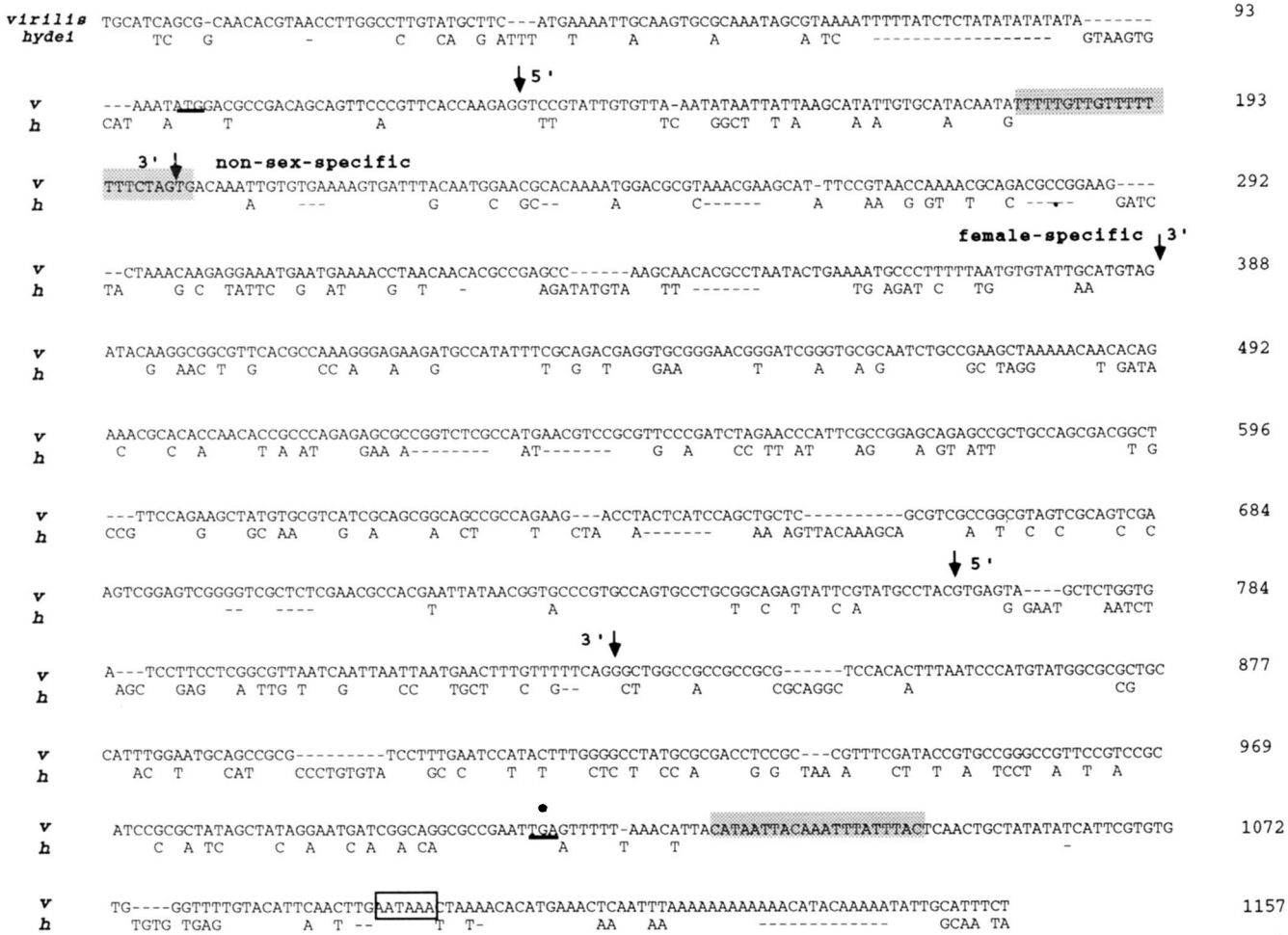

FIGURE 6.—Comparison of the DNA sequences of the *tra* genes from the Drosophila subgenus species *D. virilis* and *D. hydei*. These sequences have been aligned using the algorithm of WILBUR and LIPMAN (1983). The complete nucleotide sequence is shown only for *D. virilis*. For *D. hydei* only those bases that differ from the *D. virilis* sequence are given. Dashes represent deletions in the sequence relative to the other species. The numbers refer to the nucleotides in the *D. virilis* sequence. The arrows show the 5′ and 3′ splice sites of the two introns, as inferred from comparisons with the more completely characterized *D. melanogaster tra* gene. The underlined sequences represent the translation start and translation stop (with asterisk) sites. The boxed sequence is the consensus polyadenylation signal. The shaded sequences are the only two regions greater than ten base pairs that show perfect matches with the sequence of the Sophophoran species' *tra* genes (see Figure 3).

adults. As shown in Figure 7, *D. virilis* shows a sex-specific pattern of *tra* transcripts consistent with their having an alternative splicing event at intron 1 similar to what is operative in *D. melanogaster*. While the exact splice sites have not been determined for the *D. virilis* gene, comparison of the sequence with that of *D. melanogaster* strongly suggests that the non-sex-specific and female specific splice sites are as shown in Figure 6. Northern blot experiments, using strand specific probes also show that the sizes and orientations of the two transcription units flanking *tra* are similar to what is seen in *D. melanogaster* (data not shown). Specifically, the 5′ flanking gene is transcribed from the same strand as *tra* and specifies a major transcript of about 1.6 kb, and the gene 3′ to *tra* is transcribed from the opposite strand and lies very close to, and possibly overlapping, the 3′ end of the *tra* transcription unit. This gene specifies a transcript of about 1.4 kb.

The most striking aspect of the *tra* interspecific sequence comparison is the high degree of divergence shown by this gene (see Table 1). This is especially apparent when the Drosophila subgenus species are compared with the Sophophorans. Because of the high degree of divergence between these two groups of species, an unambiguous alignment of the sequences is impossible. At the level of nucleotide sequence comparisons, the *tra* gene region similarity among these species is only about 60% identity, too low for the DNA sequences of all five species to be reliably aligned. When the noncoding regions are discounted, the sequence similarity is still only about 60%. In fact, the two longest stretches of nucleotide sequence identity are in intron 1, 23 base pairs corresponding to the sequence immediately upstream and flanking the non-sex-specific 3′ splice site (see below), and in the 3′ untranslated region, 21 base pairs surrounding the

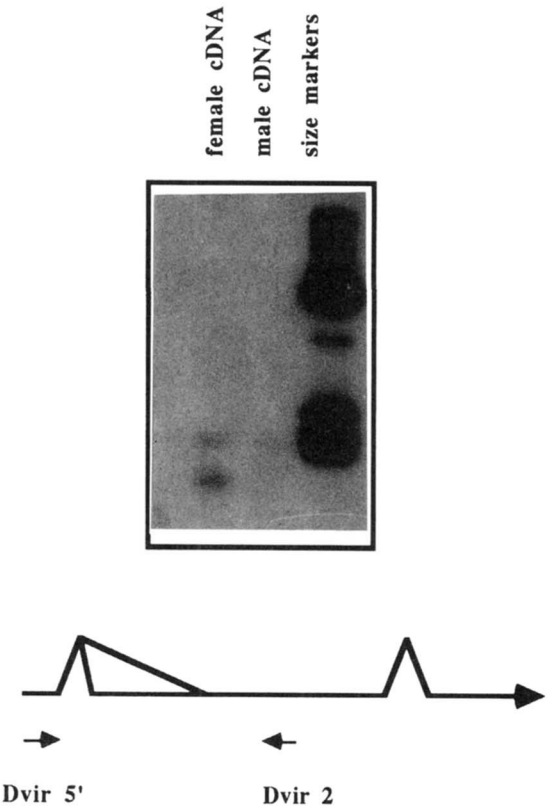

FIGURE 7.—Sex-specific expression of the *D. virilis tra*-homologous gene as shown by cDNA-PCR analysis. The locations of the primers are indicated by the arrows below the diagram of the *tra* transcript, and represent sequences located at positions 98–114 (Dvir 5′) and 599–516 (Dvir 2) on the sequence shown in Figure 6. The non-sex-specific and female-specific cDNA-PCR products are expected to be 449 base pairs and 261 base pairs in length, respectively.

sequence ATTTATTTA, a sequence which has been suggested to play a role in RNA turnover (SHAW and KAMEN 1986) (Figure 6). While the comparison of *D. hydei* and *D. virilis* shows a few stretches of sequence identity in the 5′ upstream region, comparison with the Sophophoran species fails to reveal any obvious conserved sequence elements in this region of the gene. One remarkable feature of these sequence comparisons is the high level of deletion/insertion changes that are observed when the protein coding sequences are compared. For example, when *D. hydei* and *D. virilis* are aligned, there are ten regions within the coding regions in which a deletion or insertion is postulated to have occurred. This suggests that there is an unusual degree of flexibility in the protein's structure that still allows function, since if there was not this flexibility, such deletions or insertions would be selected against.

Given the high degree of divergence at the level of the DNA, it is not surprising to see a correlative high level of divergence when the protein sequences are compared (Table 1 and Figure 8). When the *tra* proteins of all five species are aligned, there are only a few limited stretches of amino acid sequence identities. The longest stretches of conservation among all five species are two regions of six contiguous amino acids: MDADSS in the first exon and PYFADE in the second exon. The functional significance of these two conserved regions is not known. A search of the GenBank database for proteins with such regions failed to turn up any meaningful homologies. In addition to these regions, the *tra* genes in all species were found to contain extensive regions of arginine-serine, or serine-arginine, dipeptides interspersed throughout the central exon. While the high arginine-serine content of the *tra* gene is conserved, the number and spacings of these dipeptides is quite variable among the species examined. It has been observed that several proteins that play roles in pre-mRNA splicing, *e.g.*, the gene products of the *su(w^a)* and *tra-2* genes of *D. melanogaster*, and the 70K U1 snRNP and SF2/ASF splicing factors of humans, exhibit similar arginine-serine rich domains (BINGHAM *et al.* 1988; AMREIN, GORMAN and NÖTHIGER 1988; GORALSKI, EDSTRÖM and BAKER 1989; QUERY, BENTLEY and KEENE 1989; GE, ZUO and MANLEY 1991; KRAINER *et al.* 1991). Since the *tra* gene product is thought to interact with the *tra-2* gene product to influence the female-specific splicing of the *doublesex* gene's pre-mRNA, this common feature of these gene products may relate to a general role in pre-mRNA splicing. In fact, recent experiments by LI and BINGHAM (1991) have suggested that such arginine-serine domains help target proteins to a subnuclear compartment involved in RNA splicing.

Another interesting feature revealed by the interspecies comparisons of *tra* is that the carboxy-terminal one-third of all five proteins are very rich in proline residues (16/67 = 24% for *D. melanogaster*, 17/67 = 25% for *D. simulans* and *D. erecta*, 26/74 = 35% for *D. hydei*, and 20/68 = 29% for *D. virilis*), while the primary sequences of this part of the protein are not all that well conserved. Inspection of Figure 8 reveals that throughout most of the coding sequence, the only parts of the *tra* protein that can be aligned among all five species are short and scattered regions consisting of only a few contiguous amino acids. This lack of conservation among the different species' *tra* proteins represents one of the most extreme cases of rapid protein evolution, and is made more remarkable by our observation that the *D. virilis tra*-homologous gene can function reasonably well when expressed in *D. melanogaster* transgenic flies.

Conservation of a presumptive regulatory sequence within the alternatively spliced first intron: As mentioned above, the longest stretch of nucleotide sequence conservation occurs within intron 1, flanking the regulated non-sex-specific 3′ splice site (Figure 9). This conserved sequence is contained within the

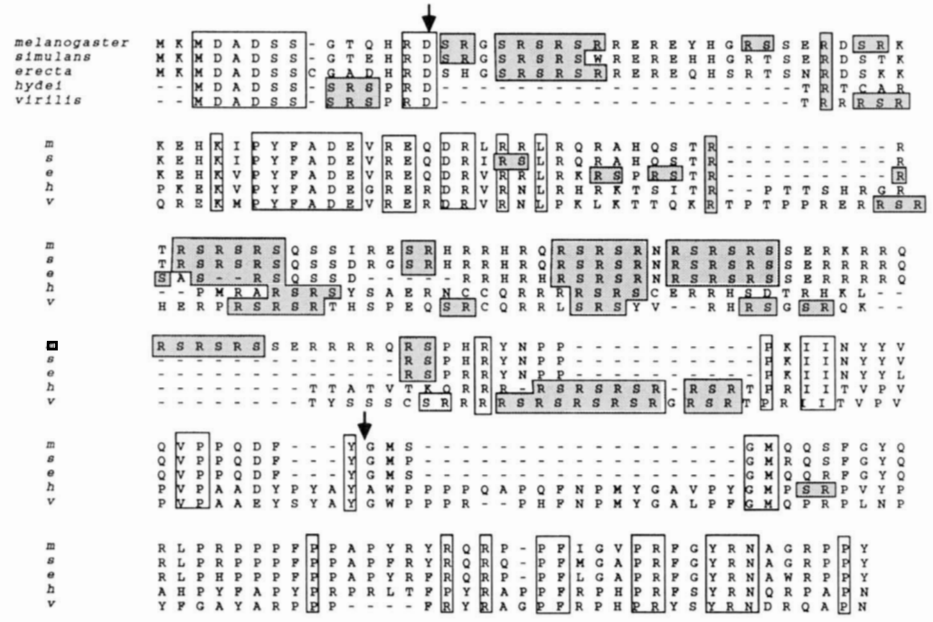

FIGURE 8.—The predicted amino acid sequences for the *tra* proteins of *D. melanogaster*, *D. simulans*, *D. erecta*, *D. hydei* and *D. virilis* . The dashes represent deletions in the sequence relative to the other species. Nonshaded boxes indicate amino acid identity among the five species. Shaded boxes highlight the arginine-serine, or serine-arginine, dipeptides found in all of the species. For some regions, especially in the central exon, the alignment shown is only one of many possible alignments consistent with the interspecific comparisons. The arrows indicate the sites of the two introns. It should be noted that it is not known whether translation is initiated at the first or the second methionine codon in the Sophophoran species.

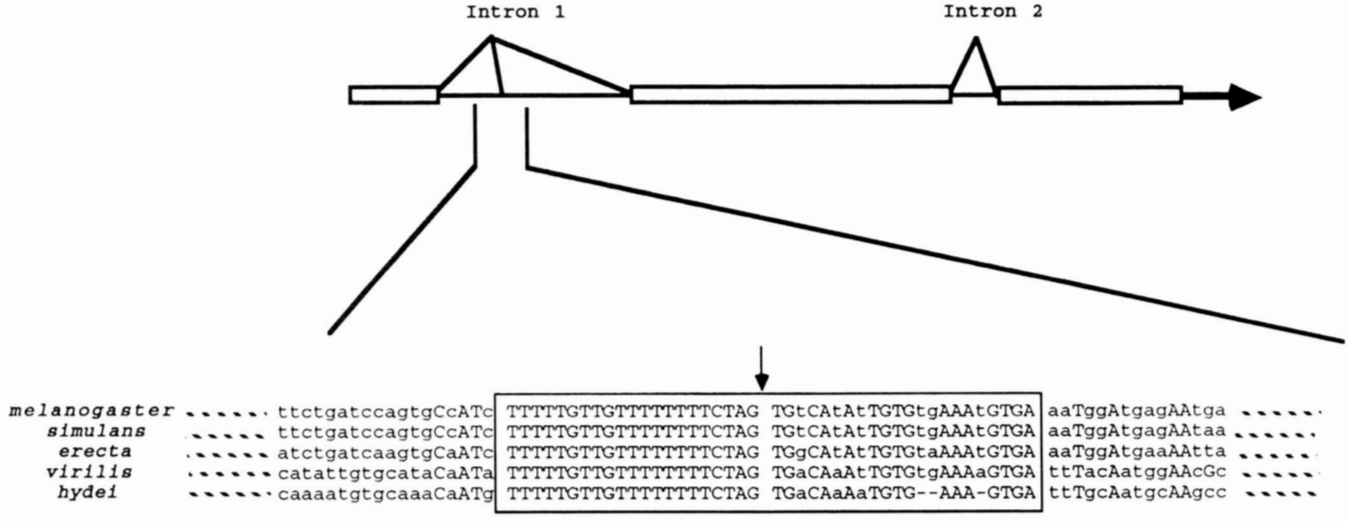

FIGURE 9.—Comparison of the non-sex-specific 3′ splice sites of the alternatively spliced intron 1 of *D. melanogaster*, *D. simulans*, *D. erecta*, *D. hydei* and *D. virilis*. The arrow indicates the position of the 3′ splice site. The upper case letters represent nucleotides that are conserved among all five species. The boxed sequence represents the region of *tra* that shows the highest degree of conservation (38/43 = 88% identity) as compared to the rest of the gene.

region of the *tra* gene that has been implicated in the sex-specific regulation of the alternative splice event that controls the female-specific expression of *tra* function in *D. melanogaster* (SOSNOWSKI, BELOTE and MCKEOWN 1989; INOUE *et al.* 1990). While most of intron 1 shows very little conservation, similar to the nonconserved, unregulated intron 2, there is conservation of the 21 bases immediately upstream, and the 18/22 bases immediately downstream, of the non-sex-specific 3′ splice site across all species examined.

One point that needs to be emphasized here is that the conserved intron sequences do not necessarily identify only those sequences of *tra* that interact with *Sxl* to control the female-specific splicing event (*i.e.*, by interfering with the use of the non-sex-specific 3′ splice site). While the sequences that interact with *Sxl* should be among the conserved sequences, it is also critical for normal sexual differentiation and fertility that the non-sex-specific splice site be highly favored over the female-specific splice site in the absence of

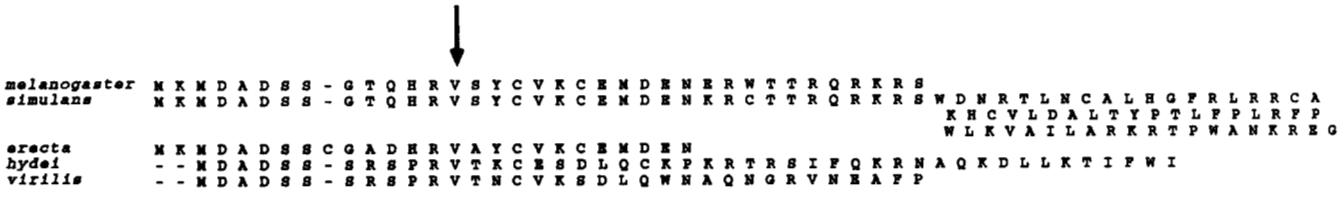

FIGURE 10.—Comparison of the predicted amino acid sequences potentially encoded by the open reading frames of the non-sex-specific *tra* transcripts of *D. melanogaster*, *D. simulans*, *D. erecta*, *D. hydei* and *D. virilis*. The dashes represent deletions in the sequence relative to the other species. The arrow denotes the site of the non-sex-specific intron.

regulation by *Sxl* (*i.e.*, in males). It has been shown that even partial use of the female-specific splice site in males leads to partial sex transformation, with an accompanying sterility phenotype (SOSNOWSKI, BELOTE and MCKEOWN 1989). Therefore, it is important that the female splice site virtually never be favored over the non-sex-specific splice site in males. Any change in the sequence that reduces the ability of the non-sex-specific splice site to be used in preference to the downstream female-specific splice site would likely be selected against. Thus, the conserved sequences highlighted in Figure 9 should include not only (1) sequences that interact with *Sxl* protein to cause blockage of the non-sex-specific 3′ splice site in females, but also (2) sequences that ensure that this non-sex-specific 3′ splice site is highly favored over the downstream female-specific 3′ splice site in the absence of *Sxl* function.

Interspecific comparison of the non-sex-specific transcript's ORF: In *D. melanogaster*, a deletion of the *tra* gene has no detectable phenotypic effect in *XY* males, suggesting that the non-sex-specific transcript has no necessary function in males (MCKEOWN, BELOTE and BOGGS 1988). Additional studies in this species have further suggested that the non- sex specific transcript may not be required in females either, and that it might represent a truly nonfunctional transcript. This conclusion is based on experiments involving *XX* female flies that were deleted for their endogenous *tra* loci but that carried minigene constructs consisting of the female specific *tra* cDNA sequence driven by the *hsp70* heat shock promoter (MCKEOWN, BELOTE and BOGGS 1988). Such flies constitutively express the female-specific *tra* transcript, but are missing the non-sex-specific *tra* transcript. The observation that these flies develop as females with normal sexual morphology demonstrates that the non-sex-specific transcript plays little if any role in causing female differentiation (MCKEOWN, BELOTE and BOGGS 1988). However, the observation that these transgenic flies were sterile left open the possibility that the non-sex-specific gene product is responsible for some subtle function necessary for female fertility. An alternative explanation for the above result is that the female-specific minigene construct is not being expressed in the proper developmental or tissue-specific pattern for normal fertility to occur, perhaps due to the exogenous promoter, and that the lack of the non-sex-specific *tra* gene product in these flies is unrelated to their sterile phenotype. This latter explanation is supported by the observation that forced expression of the non-sex-specific *tra* transcript, from a non-sex-specific *hsp70-tra* cDNA minigene, in transgenic flies carrying the female-specific minigene, did not rescue the female sterility phenotype (MCKEOWN, BELOTE and BOGGS 1988).

Interspecific comparisons of the *tra* gene sequences provides us with an independent approach for addressing the question of whether or not the non-sex-specific transcript encodes a necessary function in either sex. If this transcript is not functional in either sex, but is simply a "by-product" of this gene's particular mode of regulation (sex-specific alternative splicing), then one would not expect to see any significant conservation of the polypeptide potentially encoded by the non-sex-specific transcript's ORF, at least in that portion of the reading frame that is not shared with the female specific transcript's ORF. If, on the other hand, this transcript does encode a functional protein that plays a subtle, but real, role in some aspect of male or female development, then the non-sex-specific ORF might show some conserved features. Figure 10 shows the potential protein products encoded by the non-sex-specific transcripts of the five species examined in this study. As can be readily seen, there is extemely rapid divergence in the putative protein coding sequence starting in the second exon. (Conservation of exon one is expected since that is shared by the functional female-specific message.) For example, there is a large difference in the size and composition of the putative protein products of the closely related sibling species *D. melanogaster* and *D. simulans* (36 amino acids for *D. melanogaster vs.* 94 amino acids for *D. simulans*). Moreover, when the two subgenera are compared, it is seen that other than exon one, the amino acid sequences cannot be aligned across all species. The few conserved positions behind the splice site in *D. virilis* are most likely a consequence of the conservation of the non-sex-specific 3′ splice junction regulatory region discussed in the previous section and shown in Figure 9, and do not reflect conserved protein coding capacity. The simplest inter-

TABLE 2

Amino acid sequence identity between genes of *D. melanogaster* and their homologous genes in species of the Drosophila subgenus

Gene	Species compared with *D. melanogaster*	Amino acid identity (%)	References[a]
transformer	*D. hydei*	31	1
transformer	*D. virilis*	36	1
period	*D. virilis*	~55	2
sevenless	*D. virilis*	63	3
E74	*D. virilis*	65	4
hunchback	*D. virilis*	~80	5
engrailed	*D. virilis*	~80	6
l(3)Ai	*D. virilis*	~80	7
Adh	*D. hydei*	82	8
Ultrabithorax	*D. funebris*	85	9
mastermind	*D. virilis*	88	10
Rh4	*D. virilis*	94	11
sina	*D. virilis*	97	11
hsp82	*D. virilis*	97	12

[a] The references are (1) this report; (2) COLOT, HALL and ROSBASH (1988); (3) MICHAEL, BOWTELL and RUBIN (1990); (4) JONES, DALTON and TOWNLEY (1991); (5) TREIER, PFEIFLE and TAUTZ (1989); (6) KASSIS *et al.* (1986); (7) K. SAVILLE and J. M. BELOTE, unpublished; (8) SULLIVAN, ATKINSON and STARMER (1990); (9) WILDE and AKAM (1987); (10) NEWFIELD, SMOLLER and YEDVOBNICK (1991); (11) NEUFELD, CARTHEW and RUBIN (1991); (12) BLACKMAN and MESELSON (1986).

pretation of the extreme lack of conservation of the non-sex-specific ORF is that the non-sex-specific transcript is either not translated, or if it is, it encodes a nonfunctional polypeptide.

DISCUSSION

The most surprising finding to come out of this study was the unusually high degree of divergence shown by the *tra* coding region, when compared with other genes that have been examined by interspecific sequence comparisons among these Drosophila species. For example, published values of amino acid sequence identity between genes of *D. melanogaster* and their homologs from species of the Drosophila subgenus (*e.g.*, *D. virilis*, *D. hydei* and *D. funebris*) range from about 55% for the *period* gene (COLOT, HALL and ROSBASH 1988; THACKERAY and KYRIACOU 1990) to 97% for the *hsp82* (BLACKMAN and MESELSON 1986) and *sina* (NEUFELD, CARTHEW and RUBIN 1991) genes, with typical values in the range of 80–90% (see Table 2). In contrast, the *tra* gene coding region shows only 31–36% amino acid identity when the genes of these two subgenera are compared.

While there are examples of genes that exhibit rapidly evolving domains interspersed within more highly conserved domains, for example, the *period* (COLOT, HALL and ROSBASH 1988; THACKERAY and KYRIACOU 1990), *sevenless* (MICHAEL, BOWTELL and RUBIN 1990), *mastermind* (NEWFIELD, SMOLLER and YEDVOBNICK 1991), *engrailed* (KASSIS *et al.* 1986), and *hunchback* (TREIER, PFEIFLE and TAUTZ 1989) loci, in these cases the rapidly evolving sequences make up a small (*i.e.*, <30%) part of the overall coding region, and they frequently are regions of relatively simple sequence motifs. Moreover, within the conserved portions of these genes there are reasonably long stretches (*e.g.*, 20–90, or more, in a row) of identical amino acids. The *tra* gene is remarkable in that the rapidly evolving sequences constitute most of the coding region, and the longest contiguous stretch of identity is only nine amino acids, when the *D. melanogaster* and *D. virilis* genes are aligned, or six amino acids, when *D. melanogaster* and *D. hydei* are compared.

Because of the rapidly evolving nature of the *tra* coding region, and since transformation studies show that the *D. virilis* gene can function in transgenic *D. melanogaster*, this interspecific comparison is particularly useful in delimiting domains of the *tra* gene that are of functional significance. Molecular genetic analyses in *D. melanogaster* have suggested that the *tra* gene product acts in concert with the *tra-2* gene product, perhaps involving a physical interaction between these two proteins, in order to activate the female-specific splicing of *dsx* pre-mRNA, thereby causing female differentiation. In the experiments reported here, the *D. virilis tra* transgene is able to act in a reasonably normal manner to direct female development, implying that this transgene encodes a *tra* protein that is able to act together with the *D. melanogaster tra-2* protein to direct the female-specific splicing of the *melanogaster dsx* pre-mRNA. Thus, the rather limited sequence conservation of the *tra* coding region highlights the elements of the protein that are sufficient for these functions, and this information should provide guidance for further studies aimed at a more detailed dissection of the *tra* protein's function.

While the present study was confined to comparisons within the genus Drosophila, future studies of a more practical nature might involve the isolation of the *tra*-homologous genes from other insect species (*e.g.*, certain insect pests), with the long-range goal of using them as agents for the manipulation of sexual development, or in schemes for the generation of unisexual laboratory populations. The results presented here suggest that strategies involving cross-hybridization might be problematic for the isolation of this particular gene, due to its unusually high evolutionary divergence. However, this interspecific sequence comparison does highlight a few sequences within the *tra* gene that are conserved, and that can therefore be used for designing primers that might facilitate the isolation of the *tra* homologous genes from other, more distantly related dipterans, by PCR amplification methods.

We are indebted to KAREN KACZMAN DANIEL and MARY MILLER

for their excellent technical assistance, and to SHANNON IRVING for assistance with the photography. We thank DONAL HICKEY, CYNTHIA SMITH and DAVID SULLIVAN for providing some of the fly strains and/or genomic libraries used in this study. We especially thank DAVID SULLIVAN, SHOZO YOKOYAMA and TOM STARMER for their helpful comments. This work was supported by U.S. Public Health Service grant GM36579 from the National Institutes of Health.

LITERATURE CITED

ALEXANDER, M. L., 1976 The genetics of *Drosophila virilis*, pp. 1365–1419 in *The Genetics and Biology of Drosophila*, Vol. 1C, edited by M. ASHBURNER and E. NOVITSKI. Academic Press, New York.

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

ANANIEV, E. V., and V. E. BARSKY, 1982 A photographic map of *Drosophila hydei* polytene chromosomes. Chromosoma **87:** 239–245.

ASHBURNER, M., 1989 *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

BAKER, B. S., 1989 Sex in flies: the splice of life. Nature **340:** 521-524.

BAKER, B. S., and K. RIDGE, 1980 Sex and the single cell: On the action of the major loci affecting sex determination in *Drosophila melanogaster*. Genetics **94:** 383–423.

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.

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.

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:** 229–240.

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.

BENDER, W., P. SPIERER and D. S. HOGNESS, 1983 Chromosome walking and jumping to isolate DNA from the *Ace* and *rosy* loci and the *bithorax* complex in *Drosophila melanogaster*. J. Mol. Biol. **168:** 17–33.

BENTON, W. D., and R. W. DAVIS, 1977 Screening λgt recombinant clones by hybridization to single plaques *in situ*. Science **196:** 180–182.

BERENDES, H. D., 1963 The salivary gland chromosomes of *Drosophila hydei* Sturtevant. Chromosoma **14:** 195–206.

BEVERLEY, S. M., and A. C. WILSON, 1984 Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. J. Mol. Evol. **21:** 1–13.

BLACKMAN, R. K., and M. MESELSON, 1986 Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the Drosophila *hsp82* gene. J. Mol. Biol. **188:** 499–415.

BOGGS, R. T, P. GREGOR, S. IDRISS, J. M. BELOTE and M. MCKEOWN, 1987 Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the *transformer* gene. Cell **50:** 739–747.

BROCK, H. W., and D. B. ROBERTS, 1983 Location of the LSP-1 genes in *Drosophila* species by *in situ* hybridization. Genetics **103:** 75–92.

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.

BUTLER, B., V. PIRROTTA, I. IRMINGER-FINGER and R. NÖTHIGER, 1986 The sex determining gene *tra* of Drosophila: molecular cloning and transformation studies. EMBO J. **5:** 3607–3613.

COLOT, H. V., J. C. HALL and M. ROSBASH, 1988 Interspecific comparison of the *period* gene of Drosophila reveals large blocks of non-conserved coding DNA. EMBO J. **7:** 3929–3937.

FENG, Y., N. M. SCHIFF and D. R. CAVENER, 1991 Organ-specific patterns of gene expression in the reproductive tract of Drosophila are regulated by the sex-determination genes. Dev. Biol. **146:** 451–460.

FRISCHAUF, A.-M., H. LEHRACH, A. POUSTKA and N. MURRAY, 1983 Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170:** 827–842.

GE, H., P. ZUO and J. L. MANLEY, 1991 Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. Cell **66:** 373- 382.

GORALSKI, T. J., J.-E. EDSTRÖM and B. S. BAKER, 1989 The sex determination locus *transformer-2* of Drosophila encodes a polypeptide with similarity to RNA binding proteins. Cell **56:** 1011–1018.

GUBENKO, I. S., and M. B. EVGEN'EV, 1984 Cytological and linkage maps of *Drosophila virilis* chromosomes. Genetica **65:** 127–139.

HEDLEY, M. L., and T. MANIATIS, 1991 Sex-specific splicing and polyadenylation of *dsx* pre-mRNA requires a sequence that binds specifically to *transformer-2* protein. Cell **65:** 579- 588.

HENIKOFF, S., 1987 Unidirectional digestion with exonuclease III in DNA sequence analysis. Methods Enzymol. **155:** 156–165.

HOSHIJIMA, K., K. INOUE, I. HIGUCHI, H. SAKAMOTO and Y. SHIMURA, 1991 Control of *doublesex* alternative splicing by *transformer* and *transformer-2* in Drosophila. Science **252:** 833–836.

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

JONES, C. W., M. W. DALTON and L. H. TOWNLEY, 1991 Interspecific comparisons of the structure and regulation of the Drosophila ecdysone-inducible gene *E74* . Genetics **127:** 535–543.

KASSIS, J. A., S. J. POOLE, S. J. WRIGHT and P. O'FARRELL, 1986 Sequence conservation in the protein coding and intron regions of the *engrailed* transcription unit. EMBO J. **5:** 3583–3589.

KAWASAKI, E. R., and A. M. WANG, 1989 Detection of gene expression, pp. 89–97 in *PCR Technology*, edited by H. A. ERLICH. Stockton Press, New York.

KLEMENZ, R., U. WEBER and W. J. GEHRING, 1987 The white gene as a marker in a new P-element vector for gene transfer in Drosophila. Nucleic Acids Res. **15:** 3947–3959.

KRAINER, A. R., A. MAYEDA, D. KOZAK and G BINNS, 1991 Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. Cell **66:** 383–394.

LACHAISE, D., M. L. CARIOU, J. R. DAVID, F. LEMEUNIER, L. TSACAS and M. ASHBURNER, 1988 Historical biogeography of the *Drosophila melanogaster* species subgroup. Evol. Biol. **22:** 159–225.

LEBEDEFF, G. A., 1934 Genetics of hermaphroditism in *Drosophila virilis*. Proc. Natl. Acad. Sci. USA **20:** 613–616.

LEBEDEFF, G. A., 1938 Intersexuality in *Drosophila virilis* and its bearing on sex determination. Proc. Natl. Acad. Sci. USA **24:** 165–172.

LEBEDEFF, G. A., 1939 A study of intersexuality in *Drosophila virilis*. Genetics **24:** 553–586.

LI, H., and P. M. BINGHAM, 1991 Arginine/serine rich domains

of the *su(w^a)* and *tra* RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. Cell **67:** 335–342.

LINDSLEY, D. L., and E. H. GRELL, 1968 *Genetic Variations of Drosophila melanogaster*. Carnegie Inst. Wash. Publ. 627.

LINDSLEY, D. L., and G. ZIMM, 1985 The genome of *Drosophila melanogaster*. Drosophila Inform. Serv. **62.**

LINDSLEY, D. L., and G. ZIMM, 1986 The genome of *Drosophila melanogaster*. Drosophila Inform. Serv. **64.**

LINDSLEY, D. L., and G. ZIMM, 1987 The genome of *Drosophila melanogaster*. Drosophila Inform. Serv. **65.**

LINDSLEY, D. L., and G. ZIMM, 1988 The genome of *Drosophila melanogaster*. Drosophila Inform. Serv. **67.**

LINDSLEY, D. L., and G. ZIMM, 1990 The genome of *Drosophila melanogaster*. Drosophila Inform. Serv. **68.**

LOUKAS, M., and F. C. KAFATOS, 1986 The actin loci in the genus *Drosophila*: establishment of chromosomal homologies among distantly related species by *in situ* hybridization. Chromosoma **94:** 297–308.

MANIATIS, T., E. F. FRITSCH and J. SAMBROOK, 1982 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

MANIATIS, T., R. C. HARDISON, E. LACY, J. LAUER, C. O'CONNELL, D. QUON, G. K. SIM and A. ESTRATIADIS, 1978 The isolation of structural genes from libraries of eukaryotic DNA. Cell **15:** 687–701.

MCKEOWN, M., J. M. BELOTE and B. S. BAKER, 1987 A molecular analysis of *transformer*, a gene in *Drosophila melanogaster* that controls female sexual differentiation. Cell **48:** 489–499.

MCKEOWN, M., J. M. BELOTE and R. T. BOGGS, 1988 Ectopic expression of the female *transformer* gene product leads to female differentiation of chromosomally male Drosophila. Cell **53:** 887–895.

MELTON, D. A., P. A. KREIG, M. R. REBAGLIATI, T. MANIATIS, K. ZINN and M. GREEN, 1984 Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12:** 7035–7056.

MICHAEL W. M., D. D. L. BOWTELL and G. M. RUBIN, 1990 Comparison of the *sevenless* genes of *Drosophila virilis* and *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **87:** 5351–5353.

MORELLE, G., 1989 A plasmid extraction procedure on a miniprep scale. Focus **11:** 7–8.

NAGOSHI, R. N., and B. S. BAKER, 1990 Regulation of sex-specific RNA splicing at the *Drosophila doublesex* gene: *cis*-acting mutations in exon sequences alter sex-specific RNA splicing patterns. Genes Dev. **4:** 89–97.

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.

NEUFELD, T. P., R. W. CARTHEW and G. M. RUBIN, 1991 Evolution of gene position: chromosomal arrangement and sequence comparison of the *Drosophila melanogaster* and *Drosophila virilis sina* and *Rh4* genes. Proc. Natl. Acad. Sci. USA **88:** 10203–10207

NEWFIELD, S. J., D. A. SMOLLER and B. YEDVOBNICK, 1991 Interspecific comparison of the unusually repetitive *Drosophila* locus *mastermind*. J. Mol. Evol. **32:** 415–420.

NÖTHIGER, R., M. LEUTHOLD, N. ANDERSEN, P. GERSCHWILER, A. GRÜTER, W. KELLER, C. LEIST, M. ROOST and H. SCHMID, 1987 Genetic and developmental analysis of the sex-determining gene "double sex" (*dsx*) of *Drosophila melanogaster*. Genet. Res. **50:** 113–123.

PLILEY, R. J. L. FARMER and D. E. JEFFERY, 1986 *In situ* hybridization of biotinylated DNA probes to polytene salivary chromosomes of Drosophila species. Drosophila Inform. Serv. **63:** 147–149.

POOLE, S. J., L. KAUVAR, B. DREES and T. KORNBERG, 1985 The *engrailed* locus of Drosophila: structural analysis of an embryonic transcript. Cell **40:** 37–43.

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.

ROBERTSON, H. M., C. R. PRESTON, R. W. PHILLIS, D. M. JOHNSON-SCHLITZ, W. K. BENZ and W. R. ENGELS, 1988 A stable genomic source of *P* element transposase in *Drosophila melanogaster*. Genetics **118:** 461–470.

RYNER, L. C., and B. S. BAKER, 1991 Regulation of *doublesex* pre-mRNA processing occurs by 3′ splice site activation. Genes Dev. **5:** 2071–2085.

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

SHAW, G., and R. KAMEN, 1986 A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46:** 659–667.

SHEEN, J.-Y., and B. SEED, 1988 Electrolyte gradient gels for DNA sequencing. Biotechniques **6:** 942–944.

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.

SPRADLING, A. C., 1986 P element-mediated transformation, pp. 175–198 in *Drosophila: A Practical Approach*, edited by D. B. ROBERTS. IRL Press, Oxford.

STEINMANN-ZWICKY, M., H. AMREIN and R. NÖTHIGER, 1990 Genetic control of sex determination in *Drosophila*. Adv. Genet. **27:** 189–237.

SULLIVAN, D. T., P. W. ATKINSON and W. T. STARMER, 1990 Molecular evolution of the alcohol dehydrogenase genes of the genus *Drosophila*. Evol. Biol. **24:** 107–147.

SZABAD, J., 1978 Quick preparation of Drosophila for microscopic analysis. Drosophila Inform. Serv. **53:** 215.

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

THACKERAY, J. R., and C. P. KYRIACOU, 1990 Molecular evolution in the *Drosophila yakuba period* locus. J. Mol. Evol. **31:** 389–401.

THROCKMORTON, L. H., 1975 The phylogeny, ecology, and geography of *Drosophila*, pp. 421–469 in *Handbook of Genetics*, Vol. 3, edited by R. C. KING. Plenum Press, New York.

TONZETICH, J., S. HAYASHI and T. A. GRIGLIATTI, 1990 Conservatism of sites of tRNA loci among the linkage groups of several *Drosophila* species. J. Mol. Evol. **30:** 182–188.

TREIER, M., C. PFEIFLE and D. TAUTZ, 1989 Comparison of the gap segmentation gene *hunchback* between *Drosophila melanogaster* and *Drosophila virilis* reveals novel modes of evolutionary change. EMBO J. **8:** 1517–1525.

WHITING, J. H., JR., M. D. PLILEY, J. L. FARMER and D. E. JEFFERY, 1989 *In situ* hybridization analysis of chromosomal homologies in *Drosophila melanogaster* and *Drosophila virilis*. Genetics **122:** 99–109.

WILBUR, W. J., and D. J. LIPMAN, 1983 Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80:** 726–730.

WILDE, C. D., and M. AKAM, 1987 Conserved sequence elements in the 5′ region of the *Ultrabithorax* transcription unit. EMBO J. **6:** 1393–1401.

Communicating editor: R. E. DENELL