Key aspects of the primary sex determination mechanism are conserved across the genus *Drosophila*

James W. Erickson¹,² and Thomas W. Cline¹,*

¹Department of Molecular and Cell Biology, Division of Genetics, University of California, Berkeley, CA 94720-3204, USA
²Department of Biological Sciences, Columbia University, New York, NY 10027, USA

*Author for correspondence (e-mail: sxlcline@uclink.berkeley.edu)

Accepted 29 May; published on WWW 21 July 1998

SUMMARY

In *D. melanogaster*, a set of ‘X:A numerator genes’, which includes *sisterlessA* (*sisA*), determines sex by controlling the transcription of *Sex-lethal* (*Sxl*). We characterized *sisA* from *D. pseudoobscura* and *D. virilis* and studied the timing of *sisA* and *Sxl* expression with single cell-cycle resolution in *D. virilis*, both to guide structure-function studies of *sisA* and to help understand sex determination evolution. We found that *D. virilis* *sisA* shares 58% amino acid identity with its *melanogaster* ortholog. The identities confirm *sisA* as an atypical bZIP transcription factor. Although *virilis* *sisA* can substitute for *melanogaster* *sisA*, the protein is not fully functional in a heterologous context. The putative *sisA* regulatory sequence CAGGTAG is a potential ‘numerator box,’ since it is shared with the other strong X:A numerator gene, *sisB*, and its target, *SxlPv*. Temporal and spatial features of *sisA* and *SxlPv* expression are strikingly conserved, including rapid onset and cessation of transcription in somatic nuclei, early cessation of *sisA* transcription in budding pole cells and persistent high-level *sisA* expression in yolk nuclei. Expression of *sisA* and *Sxl* is as tightly coupled in *virilis* as it is in *melanogaster*. Taken together, these data indicate that the same primary sex determination mechanism exists throughout the genus *Drosophila*.

Key words: *D. virilis*, *D. pseudoobscura*, *D. melanogaster*, *sisA*, *Sxl*, pole cells, yolk nuclei, X:A ratio, sex determination, *Sciara*.

INTRODUCTION

Although sexual dimorphism is a ubiquitous feature of metazoans, the genetic programs that determine sex ratios differ remarkably among species (see Graves, 1996; Hodgkin, 1992) and can differ even among members of the same species (Hilfiker-Kleiner et al., 1994; Kallman, 1968; Miura et al., 1996). This rapid diversification of developmental mechanisms makes sex determination an appealing subject for comparative molecular analysis aimed at understanding how genetic pathways evolve, especially since sex determination mechanisms are particularly amenable to experimental manipulation. Interspecific comparisons are also useful for understanding structure-function relationships for sex determination genes and their products.

For these reasons, we have compared the sex-determination gene *sisterlessA* (*sisA*) among three species that span the breadth of the *Drosophila* radiation: *virilis*, *pseudoobscura* and *melanogaster*. The common ancestor of these species is estimated to have lived 40-60 Myr ago (Powell and DeSalle, 1995; Russo et al., 1995). We have also included a comparison of the temporal and spatial aspects of the expression pattern for *Sex-lethal* (*Sxl*), the regulatory target of *sisA*. *Sxl* is the master regulatory gene that heads a gene hierarchy controlling all aspects of sexual development (reviewed in Cline and Meyer, 1996).

Sex in *D. melanogaster* is determined by the effect on *Sxl* expression of the zygotic dose of at least four ‘X:A numerator genes,’ one of the most influential being *sisA* (reviewed in Cline and Meyer, 1996). For chromosomal females (2X2A), the double dose of these X-linked genes launches the female developmental program by transiently activating a *Sxl* ‘establishment’ promoter, *Pe*. By contrast, chromosomal males (1X2A) have only a single copy of these four genes and therefore fail to generate a level of numerator gene products sufficient to activate *SxlPv*. The male developmental program follows by default in the absence of *Sxl* protein.

The pulse of *Sxl* protein generated in females by the action of the numerator elements on *SxlPv* engages a positive feedback loop on *Sxl* pre-mRNA splicing after *SxlPv* shuts off and a ‘maintenance’ promoter, *SxlPm*, turns on. Primary transcripts from *SxlPm*, unlike those from *SxlPv*, are generated in both sexes but are processed into functional mRNA only in the presence of *Sxl*, an RNA-binding protein. By imposing the productive splicing mode for *Sxl* pre-mRNA, *Sxl* protein ensures its own continued synthesis throughout development. Appropriate sexual differentiation and X-chromosome dosage compensation follow as a consequence of the effects of *Sxl* on...
a variety of switch genes whose sex-specific functions are more specialized. Although males activate \( Sxl_{pm} \), they lack the early pulse of \( X:\lambda \) protein made in response to a 2X dose of numerator genes. Consequently, males splice \( Sxl_{pm} \)-derived RNA in a non-productive fashion and downstream switch gene targets are expressed in their male mode by default.

In \( D. melanogaster \), \( sisA \) encodes a transcription factor that resembles proteins of the basic-leucine zipper (bZIP) class in a number of key respects, but which lacks other features shared by most bZIP proteins (Erickson and Cline, 1993). By determining whether these defining similarities were evolutionarily conserved, we sought to ascertain their significance and thereby assess the validity of the bZIP classification for \( sisA \). In addition, we tested the functional consequences of aspects of the amino acid sequence that were not conserved.

Unlike the other \( X:\lambda \) numerator genes, \( sisA \) is not expressed in somatic cells after the blastoderm stage, nor anywhere else after mid-embryogenesis (Erickson and Cline, 1993). It is, however, highly expressed in yolk nuclei, suggesting that any non-sex-specific function it might have is likely to be related to yolk utilization. We wondered whether \( sisA \) in other species would also exhibit strong expression in yolk nuclei. If \( sisA \) was available for recruitment as a very early acting sex-determination signal element because of a prior involvement in yolk utilization very early in development, one would expect \( sisA \) expression in yolk nuclei to be at least as conserved evolutionarily as its expression in somatic nuclei.

In \( D. melanogaster \), \( sisA \) is among the earliest genes expressed. Its transcription begins in nuclear cycle 8 before any of the rapidly dividing syncytial nuclei of the young embryo have reached the egg cortex (Erickson and Cline, 1993). Cycle 8 is before the germline (i.e., the pole cells) segregates from the soma, and it is considerably before somatic cells first form in cycle 14 and most embryonic gene expression begins (reviewed in Foe et al., 1993; Pritchard and Schubiger, 1996). A striking feature of \( sisA \) expression is its abrupt cessation in those nuclei that bud off to form pole cells. This shutting off of \( sisA \) is the earliest indication of fundamental differences between sex determination in the soma and germline of \( D. melanogaster \). Would such an early onset of \( sisA \) transcription in the embryo and such an abrupt shut off in the precursors of the germline be characteristic of other species as well?

The early onset of transcription for \( sisA \) in \( D. melanogaster \) allows for an early onset of transcription for its regulatory target, \( Sxl \). Moreover, cessation of \( sisA \) expression is paralleled by the subsequent silence of \( Sxl_{pm} \). Transcription of \( Sxl \) in female somatic nuclei begins in nuclear cycle 12, only four short cell cycles after \( sisA \) and still prior to the stage at which somatic cells first form (Erickson and Cline, 1993). Would other species determine their sex as early as \( melanogaster \) and would they display such a close coupling between regulator and target?

**MATERIALS AND METHODS**

**Isolation of \( D. viridis \) and \( D. pseudoobscura \) genomic clones**

\( D. viridis \) \( sisA \) was recovered from a \( \lambda \) EMBL3 genomic library provided by G. Rubin (UC Berkeley). Plaque lifts were hybridized with a full-length \( D. melanogaster \) \( sisA \) cDNA probe in 6x SSC, 0.2% SDS, 5x Denhardt’s solution, 100 μg/ml ssDNA at 50°C. Washes were in 1x SSC at 55°C. Hybridizing phage were re-screened with a \( D. melanogaster \) \( l(1)10Bb \) cDNA probe in 6x SSC, 0.25% non-fat dry milk at 65°C. Washes were in 0.2x SSC, 0.1% SDS at 65°C. \( D. pseudoobscura \) \( sisA \) was recovered from a \( \lambda \) EMBL3 genomic library provided by A. Berry and R. Lewontin (Harvard University). The library was screened with a probe encoding the last 87 residues of \( melanogaster \) \( sisA \). Conditions for the primary and counter-screen were as described for \( l(1)10Bb \). Other techniques were standard (Sambrook et al., 1989).

**Sequences and sequence analysis**

GenBank accession numbers for \( sisA \) from \( D. melanogaster \), \( D. viridis \) and \( D. pseudoobscura \) are AF046044, AF045585 and AF045586, respectively, and AF045587 for \( l(1)10Bb \) of \( melanogaster \). A portion of the \( D. viridis \) \( Sxl_{pm} \) region (Genbank #AF046045) was amplified from genomic DNA using primers CGCTTACTAGCTCTTAATCG and GACAGCGACGAGTGTGTCAG. P. Gergen (SUNY Stony Brook) generously shared unpublished \( 5' \) runt sequence. Sequence analyses were performed with BIBI and GCG software. Percentage identity was calculated excluding residues opposite gaps from consideration.

**P-element-mediated germline transformation and genetic analysis**

\( Ncol \) and \( BglII \) sites were introduced into \( D. viridis \) \( sisA \) at the start of translation and in the 3' UTR by PCR amplification with the following primers: \( Dv4Nco \) CAACTGACCATGGAACAGAG and \( Dv3Bgl \) GTAATGTAAGCACCAGCCGC. The \( D. viridis \) \( coding sequence \) was inserted into the \( melanogaster \) gene at the \( Ncol \) and \( BamHI \) sites. This replaces the \( melanogaster \) \( coding region \) and 31 bp of \( 3' \) sequence with the \( viridis \) \( coding region \) and 14 bp of \( 3' \) sequence. A 6.4 kb \( HaeIII \) fragment containing the chimeric gene and \( l(1)10Bb \) was cloned into the \( P \)-element transformation vector Casper. The plasmid carries the same regulatory information as \( P[A10, w^{+}\text{c}c, sisA^*, l(1)10Bb^*] \) described in Erickson and Cline (1993), and germline transformants were obtained by the same techniques. Alleles and balancer chromosomes are listed at http://flybase.bio.indiana.edu.

**Embryo in situ hybridization**

\( D. viridis \) embryos were collected at 25°C on our standard cornmeal/molasses/agar food smeared with live yeast paste. Embryo fixation, preparation and hybridization with digoxigenin-labeled RNA probes, followed Lehmann and Tautz (1994) except that proteinase K concentration and incubation time were reduced for \( viridis \) embryos, and antibody incubations were in 0.1 M maleate, 0.15 M NaCl, 2% Boehringer Mannheim Blocking Reagent, 0.1% Tween 20, pH=7.5. Embryos were treated with DAPI (0.5 μg/ml) to stain nuclei. \( D. viridis \) \( sisA \) probes contained the coding region and 13 bases of 3' sequence. \( D. viridis \) \( Sxl \) exon E1 was cloned by PCR amplification of genomic DNA using mismatch primers virE1Bam5' TitaGATCCCTTTAA-ATAGCTACGCAC and virE1Xba3' GGAGAA TGCAtTtCT. The product was cloned into the \( BamHI \) and \( XbaI \) sites of pBluescript KS(+). The 252 bp insert includes the 198 bp UTR by PCR amplification with \( BglII \) and \( NcoI \) sites were introduced into \( runt \) by PCR amplification with \( GATCTACGCAC \) and \( virE1Xba3' \) GGAGAA TGCAtTtCT. The product was cloned into the \( BamHI \) and \( XbaI \) sites of pBluescript KS(+). The 252 bp insert includes the 198 bp UTR by PCR amplification with \( BglII \) and \( NcoI \) was identified as focused dots of DAPI staining.
emphasizes the largest continuous blocks of conserved sequence. Identities are shown only in regions with extended similarity. A conserved heptamer is tentatively labeled ‘numerator box’ due to its occurrence in functionally related genes (see Fig. 2).

For mRNAs, the Sxl exon E1 probe clearly discriminates between Sxl_{D. melanogaster}-derived and Sxl_{D. pseudoobscura}-derived species, since only Sxl_{D. melanogaster} mRNA contains this exon; however, discrimination between nascent transcripts is not as clear, since E1 is included in the first intron of Sxl_{D. pseudoobscura} transcripts (Bopp et al., 1996). Hence probe E1 should detect nascent transcripts from Sxl_{D. pseudoobscura} that have not yet made their exon 1-2 splice. All cycle 12 embryos that stained with probe E1 exhibited two nuclear dots in at least some nuclei and were therefore female. In cycle 13 and early cycle 14, females could be identified by the presence of two intense nuclear dots per nucleus, but also by a cytoplasmic signal. While similarly staged male embryos did show two intense nuclear dots in at least some nuclei and were therefore female. In cycle 12 and early cycle 14, females could be identified by the presence of two intense nuclear dots per nucleus, but also by a cytoplasmic signal. While similarly staged male embryos did show very faint single dots in their nuclei, they exhibited no cytoplasmic signal. While similarly staged male embryos did show very faint single dots in their nuclei, they exhibited no cytoplasmic Sxl mRNA. In late cycle 14 females, cytoplasmic Sxl mRNA levels declined and intense nuclear dots disappeared. At this point, both sexes displayed very faint nuclear dots that are not apparent in photographs. These faint dots seem likely to represent Sxl_{D. pseudoobscura}-derived primary transcripts, although we cannot exclude the possibility that they reflect a low basal level of Sxl_{D. melanogaster} expression in both sexes.

**RESULTS**

A three-species comparison of *sisA* gene and protein structure

To clone *sisA* from *D. virilis* and *D. pseudoobscura*, we took advantage of the close linkage discovered in *D. melanogaster* between *sisA* and *l(1)10Bb* (Erickson and Cline, 1993), a small gene that is highly conserved from plants to animals (Hla et al., 1995). We correctly assumed that synteny would be maintained in *virilis* for two genes only 2 kb apart in *melanogaster*. Consequently, we first screened *virilis* and *pseudoobscura* genomic libraries with a *melanogaster* *sisA* cDNA probe, then determined which phages were likely to carry the *sisA* ortholog by screening positives with the *melanogaster* *l(1)10Bb* probe. The *sisA* regions were sequenced and compared. Because the structure and much of the sequence of *sisA* was highly conserved, identification of the orthologs was unambiguous. For all three species, the *sisA*-coding region is small and there are no introns.

Outside the coding region, nucleotide sequence similarity was apparent only in a few upstream regions (Fig. 1). Given the rate of sequence divergence in the genus *Drosophila*, these conserved regions are likely to be important for regulation (Hartl and Lozovskaya, 1995). We wondered whether any of these putative regulatory sequences might be shared with other X:A numerator elements or with their target, Sxl. The heptamer CAGGGTA was most promising in this regard. Within *sisA*, this sequence was found on the sense strand at a corresponding position less than 200 bp upstream of the transcription start site in all three species (Figs 1, 2). In *melanogaster* and *pseudoobscura*, but not *virilis*, another copy was found on the antisense strand closer to the start site.

Among the sex determination genes, the expression of *sisB* most closely matches that of *sisA*. Within the 543 bases of upstream sequence published for *D. melanogaster* *sisB*, the heptamer appears twice on the antisense strand at positions comparable to those conserved in *sisA* (Fig. 2). The gene *runt* is a weaker numerator element whose expression occurs slightly later and in a more spatially restricted domain than that of *sisA* or *sisB*. Within the 440 bases of 5′ sequence available for *runt* in *melanogaster*, *pseudoobscura* and *virilis*, there are no occurrences of this heptamer or its complement.

Two copies of the heptamer occur in conserved positions near the transcription start site in *Sxl* that is activated by the numerator elements (Fig. 2). Approximately 80 base pairs

![Fig. 1](image-url) Conserved nucleotide sequences in the *sisA* promoter region. The 5′ region of *sisA* is shown for *D. melanogaster* (*mel*), *D. pseudoobscura* (*pse*) and *D. virilis* (*vir*). (+1) identifies the presumptive transcription start sites, and (AUG) the translation starts. The alignment emphasizes the largest continuous blocks of conserved sequence. Identities are shown only in regions with extended similarity. A conserved heptamer is tentatively labeled ‘numerator box’ due to its occurrence in functionally related genes (see Fig. 2).
upstream of the SxlPe start, the heptamer is found on the antisense strand in D. melanogaster and D. subobscura and on the sense strand in D. virilis. Approximately 50 bases further upstream, the heptamer (allowing one mismatch in melanogaster) is repeated on the sense strand in all three species.

The deduced amino acid sequence of sIS A is compared in Fig. 3 for the three species. The three proteins (Dv-sIS A, Dp-sIS A and Dm-sIS A) are similar in size, charge and hydrophilicity. Dm-sIS A shares 58% amino acid identity with Dv-sIS A and 55% with Dp-sIS A. It appears that SIS A has two structural domains: a conserved bipartite N-terminal region of unknown function and a C-terminal dimerization and DNA-binding region containing an unusual basic leucine zipper (bZIP) motif. These domains are separated by a non-conserved ‘spacer’ of varying size.

The most conserved part of sIS A is the bZIP motif. The features that sIS A shares with other bZIP family members are preserved, but so are the features that are so unusual for a protein of this class. Among these atypical features is the absence of characteristic single or paired alanines located three and/or four residues downstream of the invariant asparagine in the basic region (Hurst, 1995). Also atypical is the pattern of amino acids in the ‘a’ position of the dimerization helix. The ‘a’ and ‘d’ residues serve as the dimerization interface and are, in most bZIP proteins, large hydrophobic amino acids. ‘d’ is usually leucine, consistent with the L•L•I•L sequence found for sIS A. ‘a’ generally contains other hydrophobic residues, but in sIS A, the invariant sequence of ‘a’ residues is A•R•E•G. This arrangement, combined with the lack of helix-stabilizing residues in the ‘e’ and ‘g’ positions, suggests that sIS A is unlikely to form stable homodimers (Krylov et al., 1994; Vinson et al., 1993), in agreement with results of two-hybrid analyses (Liu and Belote, 1995). The positions in the zipper domain that diverge most, ‘c’ and ‘f’, are predicted to lie on the outside surface of the helix and thus not to participate in dimer formation.

D. virilis sisA protein functions suboptimally in D. melanogaster

To determine if the D. virilis sisA protein is capable of providing sisA function in D. melanogaster, we expressed Dv-sisA under the control of melanogaster regulatory sequences and assayed its function in transgenic animals. The melanogaster sisA protein-coding region was replaced with that for virilis within a P-element construct previously shown to provide full sisA function (see Materials and Methods). Data in Table 1 document that the sisA protein sequence divergence between these two species has impaired but not eliminated the ability of the protein to function in the heterologous situation. A single copy of a Dv-sisA expressing chimeric transgene fully rescued females homozygous for the hypomorphic, female-specific lethal allele sisA 1; however, the chimeric transgene was not fully effective at rescuing hemizygous sisA 1 females, a more stringent test of sisA function (Cline, 1988). Consistent...
with the results for females, the chimeric transgenes also displayed sisA+ activity in males (not shown). This was manifested as a male-specific lethal effect of the transgene in combination with duplications of sisB+ that was somewhat weaker than the male-lethal effect previously reported for the DM-sisA parental transgene.

**Conservation of sisA expression**

In *D. melanogaster*, the early onset of sisA transcription, the extinction of sisA transcription in pole cells and the abrupt cessation of sisA transcription during nuclear cycle 14 are all reflections of this gene’s participation in sex determination through its regulation of Sxl. The persistence of sisA transcription in *melanogaster* yolk nuclei suggested that the gene might also have functions that are not sex-specific. We found that all four of these unusual features have been conserved over the >40 Myr separating *D. melanogaster* and *D. virilis*.

Assaying by in situ hybridization, we first detected *D. virilis* sisA transcripts in nuclear cycle 8 (Figs 4, 5). At this point, the nuclei are migrating outwards to the cortex. Although differences were noted among nuclei in the time of onset, all nuclei had begun to transcribe sisA within one cell cycle. Initially, sisA transcripts were found primarily within nuclei. Unambiguous cytoplasmic signals arose only later during cycle 10. Other *Drosophila* genes that turn on at a comparably early stage display similar behavior (Pritchard and Schubiger, 1996). Nuclei near the posterior pole expressed lower levels of sisA mRNA than nuclei elsewhere. All somatic and yolk nuclei continued to express sisA in the subsequent cleavage cycles. In contrast, expression was extinguished in the germline precursors, the pole cells, during cycle 9. Transcripts were readily apparent in prepole cell nuclei as the pole buds formed, but disappeared midway through bud formation (Fig. 5E,F).

Transcription of sisA was maintained during the subsequent cycles of the syncytial blastoderm (Fig. 4). For somatic nuclei, the level of sisA mRNA peaked in late cycle 12 or early cycle 13. By this point, the sisA regulatory target, SxlP, has become active (see below). In some embryos, focused dots of sisA hybridization could be seen in the nuclei. Such dots represent nascent transcripts (Pritchard and Schubiger, 1996; Rothe et al., 1992; Shermoen and O’Farrell, 1991). Although hybridization dots could not be seen in all embryos, or even in all the nuclei of any single embryo, embryos displaying nuclear dots fell into two classes: individuals with only a single dot per nucleus and embryos with two (Fig. 5C,D). Presumably, this reflects an X-chromosome location for sisA in *D. virilis* with females carrying two alleles and males only one.

After cycle 13, nuclear dots disappeared, signaling the cessation of sisA transcription. The cytoplasmic hybridization signals begin to disappear from the periphery of the embryo first, but by early to mid cycle 14, the last traces of sisA mRNA had disappeared from all surface nuclei. As had been the case cycles of the syncytial blastoderm (Fig. 4). For somatic nuclei, the level of sisA mRNA peaked in late cycle 12 or early cycle 13. By this point, the sisA regulatory target, SxlP, has become active (see below). In some embryos, focused dots of sisA hybridization could be seen in the nuclei. Such dots represent nascent transcripts (Pritchard and Schubiger, 1996; Rothe et al., 1992; Shermoen and O’Farrell, 1991). Although hybridization dots could not be seen in all embryos, or even in all the nuclei of any single embryo, embryos displaying nuclear dots fell into two classes: individuals with only a single dot per nucleus and embryos with two (Fig. 5C,D). Presumably, this reflects an X-chromosome location for sisA in *D. virilis* with females carrying two alleles and males only one.

<table>
<thead>
<tr>
<th>sisA transgene</th>
<th>Relative viability of sisA1/sisA1 females*</th>
<th>Relative viability of Df(sisA)/sisA1 females†</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>P(sisA-MBL) 45</td>
<td>103%</td>
<td>90%</td>
</tr>
<tr>
<td>P(sisA-VIR) 5</td>
<td>84%</td>
<td>58%</td>
</tr>
<tr>
<td>P(sisA-VIR) 6</td>
<td>92%</td>
<td>74%</td>
</tr>
<tr>
<td>P(sisA-VIR) 22</td>
<td>104%</td>
<td>28%</td>
</tr>
</tbody>
</table>

*Crosses @25°C: y w sisA1/Binsinscy Δ δ × Δ δ y w sisA1/Y; P(w+mc, sisA)+.
†Crosses @25°C: y w Df(1)N71, sisA1/FM7c Δ δ × Δ δ y w sisA1/Y; P(w+mc, sisA).
‡y w sisA1/Balancer; P(w+mc, sisA)+ females served as viability references.

**Table 1. D. virilis sisA functions suboptimally in D. melanogaster**

![Fig. 4. Time course of sisA expression in D. virilis. Transcripts were detected by in situ hybridization with a sisA probe. Precellular embryos are from cycles 5, 8, 9, 10, 12, 13 and 14. Early gastrula (gast) and germ-band-extended (gbe) embryos are shown.](image)
with *D. melanogaster*, no subsequent somatic or germ-cell transcription of sisA was observed in *virilis* embryos.

In *D. virilis* as in *D. melanogaster*, sisA transcripts persisted in the yolk nuclei until the demise of those nuclei 10 to 12 hours after fertilization (Fig. 4). The primary yolk nuclei derive from 70-100 nuclei that fail to migrate with the majority of nuclei toward the periphery of the embryo during cycles 8 and 9. Instead they fall back into the central yolk mass and eventually become polyploid (Foe and Alberts, 1983; Foe et al., 1993). During cycle 14, sisA transcript levels in these nuclei increased abruptly and remained high thereafter. During normal embryogenesis, some nuclei drop from the periphery into the surface of the yolk mass to become ‘secondary’ yolk nuclei, perhaps as a consequence of errors in replication or division (Frenz and Glover, 1996; Sullivan et al., 1993). Although not highly polyploid, these nuclei also express high levels of sisA. Probable secondary yolk nuclei can be seen in Fig. 4 (cycle 14) near the periphery and at the anterior pole of the embryo. Their behavior suggests that persistence of sisA transcription is likely to be a consequence of environmental differences between the central yolk mass and the yolk-free periphery, rather than polyploidy per se.

**The timing of Sxl{Pe} expression is conserved in *D. virilis***

The observation that the timing of sisA expression is essentially identical between *D. virilis* and *D. melanogaster* leads to the question of whether the timing of expression of Sxl{Pe}, the target of sisA, is equally conserved. It had been shown previously that Sxl is expressed in a sex-specific manner at the cellular blastoderm stage in both *D. virilis* and *D. subobscura* (Bopp et al., 1996; Penalva et al., 1996), and for *subobscura* this expression had been shown to be due to Sxl{Pe}; however, the temporal resolution of that analysis was relatively crude. Consequently, we performed in situ hybridization experiments with an exon E1 probe for Sxl{Pe} products, analyzing the young embryos cell cycle by cell cycle. We found that the temporal regulation of Sxl{Pe}, like that of sisA, is virtually identical between these two distant *Drosophila* species.

The accumulation of *D. virilis* Sxl{Pe} transcript across the whole embryo is shown in Fig. 6, while Fig. 7 illustrates the time course of ‘nuclear dot’ development reflecting the synthesis of Sxl{Pe} transcripts. A Sxl{Pe} signal was first detected as light dots of staining in interphase nuclei of cycle 12 embryos (Fig. 7A). The fact that Sxl{Pe} activation is female-specific in *virilis* was established by the observation that as the dots darkened, they ultimately occurred only in pairs, reflecting the XX (female) karyotype of the Sxl-positive embryos. Nevertheless, Sxl{Pe} was not necessarily activated simultaneously on both X chromosomes within a single nucleus. Individual early cycle 12 embryos exhibited mosaic patterns of nuclear dots in which some nuclei expressed Sxl from both X chromosomes, some from only one X, and others from neither (Fig. 7A). Slightly later in cycle 12, and in cycle 13 and 14 embryos, the nuclear dots became much more intense and these intense dots were always found in pairs (Fig. 7B-D). Cytoplasmic Sxl transcripts first appeared in late cycle 12 females (Fig. 6). They reached maximum levels shortly after the peak in nuclear dot staining was reached – late cycle 13 to early cycle 14. In the nuclei of cycle 13 and 14 embryos that exhibited no cytoplasmic accumulation of exon-E1-containing mRNA, only a single faint dot of staining was visible (Fig. 7F). This dot is likely to represent unspliced Sxl{Pe} transcripts in males (see Materials and Methods).
As cycle 14 continued, shut off of SxlPe was signaled by the disappearance of clear-cut nuclear dots. Degradation of transcripts made prior to this point was evidenced by the rapid loss of cytoplasmic staining thereafter. The shut-off of SxlPe and degradation of Sxl mRNA appear to be non-uniform, since a series of weakly defined stripes became apparent in late blastoderm and early gastrula embryos (Fig. 6), just as had been observed for D. melanogaster (Barbash and Cline, 1995; Erickson and Cline, 1993). Loss of nuclear dots occurred first in the lighter staining areas (not shown). As was true for melanogaster, transcripts derived from SxlPe in D. virilis were not evident in pole cells or in embryos at later stages, regardless of sex (not shown).

**DISCUSSION**

The present study adds the strong X:A numerator element sisterlessA to the list of Drosophila sex determination genes for which interspecific comparisons have been made. That list now includes three genes of the primary sex determination signal (sisA, sisB, runt), the target of the primary signal (Sxl), three regulatory genes downstream of Sxl that are specifically involved in somatic sexual differentiation (tra, tra2 and dsx), and two regulatory targets of Sxl that control dosage compensation (msl-1 and msl-2) (Bone and Kuroda, 1996; Bopp et al., 1996; Botella et al., 1996; Chandler et al., 1997; Marin et al., 1996; O’Neil and Belote, 1992; Penalva et al., 1996; Pepling and Gergen, 1995; Raymond et al., 1998).

**A potential 'numerator box': promoter sequence similarities among numerator elements sisA and sisB and their target, Sxl**

One aim of this study was to use DNA sequence conservation to identify potential regulatory regions that allow genes to function as X:A numerator elements. The heptomeric sequence CAGGTAG, which is conserved in all three sisA genes examined here, is a promising candidate. It appears in pairs less than 200 bp upstream of the transcription start site not only for sisA, but also for sisB and Sxl (Fig. 2). These three genes have in common an unusually early onset of expression that allows the embryo to establish X-chromosome dosage compensation by the time general transcription begins. Perhaps this sequence provides for such an early start of transcription.

**Conservation of the unusual bZIP functional domain of sisA protein**

The protein-coding sequences of sisA display a commonly observed pattern: highly conserved regions separated by completely diverged sequences (Fig. 3). Conservation of the bZIP domain confirms the validity of this motif assignment, despite the occurrence of non-canonical residues in key

---

**Fig. 6.** Time course of SxlPe expression in D. virilis. SxlPe-derived transcripts were detected by in situ hybridization with an exon E1 probe. The embryo in interphase of cycle 11 (11i) has not stained. The embryo in early interphase 12 (12ei) is a female that had just initiated SxlPe expression and exhibits nuclear staining only (see also Fig. 7A). The later interphase 12 (12i); interphase 13 (13i); anaphase 13 (13a); and early cycle 14 (14) embryos are females exhibiting nuclear and cytoplasmic Sxl staining. The gastrulating embryo (gast) is female. The cycle 14 female represents the darkest staining observed and is similiar in stage to the cycle 14 male.
positions. The N-terminal region of the protein was also found to be conserved, notwithstanding the fact that the N-terminal 15 amino acids can be deleted with only modest effects on in vivo function (Erickson and Cline, 1993). An intriguing feature of the melanogaster protein was a run of six glycine-serine pairs hypothesized to serve as a flexible spacer linking two functional domains. Although this run is not conserved, it occurs in a region that is notable for its variable length and lack of conservation, consistent with such a spacer role.

The level of amino acid sequence conservation for sisA protein within the genus (55-68% identity) is comparable to that seen with the Drosophila sex determination genes sisB and tra-2, is less than that for runt and Sxl (80-90%), and is considerably greater than that for tra (36%). Thus sisA, like most other sex-signal genes of Drosophila, does not display the remarkably rapid evolution observed for the mammalian sex-signal gene, SRY (Pamilo and O’Neill, 1997).

**Heterospecific functioning of sex-determination genes**

Genomic fragments from *D. virilis* that included *tra* or *tra-2* provided function to *D. melanogaster*, but the level of activity was below that of endogenous alleles (Chandler et al., 1997; O’Neill and Belote, 1992). Similarly, sisB from *D. subobscura* provided sex-determination activity to melanogaster (Botella et al., 1996), but it is not clear how close to wild-type that activity might have been. Because these heterospecific transgenes carried both regulatory and structural information from the foreign species, the basis for their suboptimal performance remains unclear. In our study of the functioning in *D. melanogaster* of sisA from *D. virilis*, only the protein-coding sequence was foreign. Nevertheless, the results resembled those with *tra* and *tra2*: the foreign protein provided some function, but less than the endogenous gene. This result implies that the sequence divergence among the sisA proteins is functionally significant.

**The temporal and spatial pattern of sisA expression is conserved**

We show here that the many distinctive features of the sisA expression pattern described for *D. melanogaster* are conserved in *D. virilis*. Not only were sisA transcripts first detected in cycle 8 embryos from both species, all nuclei expressed sisA by the end of that cell cycle. In contrast, Pritchard and Schubiger (1996), studying two other early expressed *Drosophila* genes *ftz* and *Kr*, found transcriptional activation to be more gradual, with different nuclei in the same embryo beginning transcription during different cell cycles. Whether this reflects fundamental differences in the mechanisms by which these early genes are activated remains to be determined.

In part because so little is known about the role of the yolk nuclei and the factors that control expression of their genes, our observation that persistent sisA expression in yolk nuclei is a conserved feature of this gene does not yet allow us to distinguish between the possibility that sisA has an important function in yolk utilization, and the alternative possibility that persistent yolk expression is simply a consequence of the way the gene is activated for sex determination. Ner and Travers (1994) noted that the chromatin environment of the early *Drosophila* embryo is similar to that of the yolk nuclei in being rich in HMG-d protein and deficient in histone H1. Hence it is possible that factors that allow sisA to be transcribed during the nuclear multiplication stage might lead to its continued expression in the yolk simply because the chromatin environment of those nuclei remains unchanged. Arguing against such an ‘environmental’ explanation, however, is the fact that sisB, which begins transcription as early as sisA, does not exhibit persistent high-level expression in the yolk nuclei (Erickson and Cline, 1993). If sisA expression in yolk nuclei is functionally significant, it will be necessary to examine species beyond the genus *Drosophila* to determine whether this might be the ancestral function predating sisA’s recruitment as a sex-determination signal element.

The behavior of sisA in the pole cells, the precursors to the germline, has implications for how those cells remain transcriptionally inactive prior to germ band extension. It had been suggested that this quiescence might be due to pole cells being derived from transcriptionally inactive nuclei that are kept inert until the germ band extends (Ner and Travers, 1994; Williamson and Lehmann, 1996). This explanation is inconsistent with the behavior of sisA, which establishes that prepole cell nuclei, and even budding pole cells, transcribe genes that will be silent after bud formation is complete. Since pole cell nuclei remain in mitotic synchrony with the other embryonic nuclei for some time after bud formation is complete (Foe et al., 1993), the sisA shut-off seems not to be a consequence of some obvious change in the mitotic cycle. There is evidence that *Drosophila* and *C. elegans* both keep
early germ cells transcriptionally silent by preventing phosphorylation of the carboxy-terminal domain of RNA polymerase II (Seydoux and Dunn, 1997). The behavior of \textit{sisA} suggests that active Pol II is either inactivated or excluded from the nuclei as pole cells form. \textit{sisA} might serve as a useful molecular marker in a search for genes regulating such events.

**Correlation between \textit{sisA} expression and \textit{Sxl} \textit{Pe} activation, and the conservation of X-chromosome counting mechanisms**

The fact that \textit{sisA} and \textit{Sxl} \textit{Pe} exhibit precisely the same temporal and spatial expression patterns in two species that span the breadth of the \textit{Drosophila} radiation argues that the same primary mechanism for determining sex exists throughout this genus. In both species, \textit{Sxl} \textit{Pe} is first activated in cycle 12 diplo-X embryos and shuts off during cycle 14. The pattern of \textit{Sxl} RNA hybridization dots in early cycle 12 embryos shows that the activation of \textit{Sxl} \textit{Pe} must be a stochastic process that occurs independently not only in each nucleus, but even on each X chromosome. Such behavior favors the hypothesis that the rate-limiting step in \textit{Sxl} activation is assembly of an active transcription complex on the promoter itself, rather than formation of active heterodimeric transcription factors in the nucleoplasm. There seems to be a difference between the initial activation of \textit{Sxl} \textit{Pe} and its reactivation after mitosis: nascent transcript dots were not apparent on anaphase and telophase chromosomes at the ends of cycles 12 and 13, but when they reappeared in interphase diplo-X embryos in cycles 13 and 14, they were always seen as two dots per nucleus. The temporal pattern of \textit{Sxl} \textit{Pe} activity is consistent with a direct effect of \textit{sisA} protein on \textit{Sxl} \textit{Pe} in both \textit{virilis} and \textit{melanogaster}, since in both species \textit{Sxl} \textit{Pe} activity begins when the \textit{sisA} mRNA level is near a very transient maximum and ceases shortly after \textit{sisA} mRNA disappears.

Although the initial events in sex determination for \textit{D. virilis} and \textit{D. melanogaster} appear identical, primary events in sex determination for other Diptera are clearly very different (Hilfiker-Kleiner et al., 1994; Meise et al., 1998; Puchalla, 1994; Saccone et al., 1998; Traut, 1994). Is the \textit{Drosophila} mechanism a relatively recent evolutionary innovation, or might it represent instead an ancient mechanism for Dipteran sex determination that has been lost in some lineages? In this connection, it is interesting to note that the \textit{Sciaraeidae} and \textit{Cecidomyiidae}, Dipterans that diverged from the \textit{Drosophila} lineage some 200 Myr ago (see Sommer et al., 1992), may possess an X chromosome counting mechanism for sex determination that operates as early as that in \textit{Drosophila} (de Saint Phalle and Sullivan, 1996; Gerbi, 1986; Metz, 1938; Stuart and Hatchett, 1991).

Similarities between sex determination mechanisms in genera as diverse as \textit{Sciara} and \textit{Drosophila} may be purely coincidental, but those among species within the genus \textit{Drosophila} clearly are not. Advancing to the next level of sophistication in understanding the extent of similarities among \textit{Drosophila} species – for example, is \textit{sisA} as important a numerator element in \textit{virilis} as in \textit{melanogaster}? – is likely to require bona fide genetic analysis in these other species. Recent advances in transgenic techniques (Lohe and Hartl, 1996; Lozovskaya et al., 1996) make such an approach less formidable than it seemed not long ago.

We thank S. Feist, L. Subrahmanyan, J. Walker and M. Bell for help with library screening and sequencing, Y. Wei for assistance with transgenes, A. Berry, R. Lewontin and G. Rubin for genomic libraries, B. Meyer and L. Megna for comments on the manuscript, and B. de Saint Phalle for discussions about \textit{Sciara}. Research support was from US NIH grant GM23468 (T. W. C.), ACS grant RPG-97-079-01-DB (J. W. E.), and a Searle Community Trust Award (J. W. E.).

**REFERENCES**


for leucine zipper stability and dimerization specificity: e and g interhelical interactions. *EMBO J.* 13, 2849-2861.


