

Mark L. Siegal · Bruce S. Baker

Functional conservation and divergence of *intersex*, a gene required for female differentiation in *Drosophila melanogaster*

Received: 19 July 2004 / Accepted: 29 September 2004 / Published online: 12 November 2004
© Springer-Verlag 2004

Abstract In *Drosophila melanogaster*, somatic sexual differentiation is regulated by a well characterized genetic hierarchy, by which the ratio of X chromosomes to autosomes (X:A) ultimately directs the deployment of sex-specific transcription factors encoded by *doublesex* (*dsx*) and *fruitless* (*fru*). In other dipterans, the X:A ratio is not the primary sex-determination signal. Correspondingly, the *Drosophila* hierarchy is not fully conserved. In all non-drosophilid fly species examined, *Sex-lethal* (*Sxl*), the master switch at the top of the *Drosophila* hierarchy, does not control somatic sex. This rapid divergence contrasts with the apparently deep conservation of *dsx*, which in *Drosophila* controls virtually all aspects of somatic sex except for male courtship behavior (which is controlled by *fru*). Sex-specific *dsx* mRNAs have been reported in Diptera, Lepidoptera and Hymenoptera, and *dsx* homologs in nematodes and mammals are required for aspects of male differentiation. Thus, it seems that the bottom of the hierarchy is rather ancient, especially compared with the top. To test this, we cloned insect and vertebrate homologs of the *Drosophila* gene *intersex* (*ix*), which functions together with *dsx* at the bottom of the hierarchy in females. When expressed in *D. melanogaster* females mutant at the endogenous *ix* gene, dipteran and lepidopteran *ix* homologs restore proper sexual differentiation, substantiating the hypothesis that *ix*, like *dsx*, is broadly conserved. When the mouse homolog is expressed it produces a dominant-negative phenotype suggesting partial functional divergence. Our results raise the possibility that a func-

tional association between *ix*- and *dsx*-related gene products existed before the origin of the bifunctional *dsx* gene used in insect sex determination.

Keywords Sex determination (genetics) · Evolution · Intersex protein · *Drosophila*

Introduction

An unexpected and remarkable finding of research in diverse model organisms is the deep homology and functional conservation of key developmental and regulatory genes (Carroll et al. 2001). What is known about the great variety of sex-determining systems in animals seems to contradict this pattern (Bull 1983). Even closely related species differ in their primary chromosomal sex-determination signals. One of the best studied animal lineages in this regard is the insect order Diptera, which includes a number of devastating crop and livestock pests (including medflies and blowflies) and disease vectors (including tsetse flies and mosquitoes), as well as the geneticist's darling, the vinegar fly *Drosophila melanogaster* (for phylogenetic relationships among the fly taxa discussed in this paper, see McAlpine and Wood 1989). The cyclorrhaphan ("higher") Diptera are particularly well studied, and include species with male heterogamety (XY males, XX females), others with female heterogamety (XX males, XY females), and still others with homomorphic (i. e., cytologically indistinguishable) sex chromosomes that differ only in which alleles of a sex-determining locus they carry. Moreover, even similar cytology might mask underlying differences in mechanism: in *D. melanogaster* an XXY fly is female, because she has two X chromosomes, whereas in the medfly *Ceratitidis capitata* an XXY fly is male, because he has a Y chromosome (Willhoft and Franz 1996). Some species, such as the housefly *Musca domestica*, even have natural strains with different primary sex-determining loci, some of which convert autosomes into new sex chromosomes by virtue of their presence (Bull 1983). The inconstancy of sex chromo-

Edited by D. Tautz

M. L. Siegal · B. S. Baker
Department of Biological Sciences, Stanford University,
Stanford, CA, 94305-5020, USA

Present address:

M. L. Siegal (✉)
Department of Biology, New York University,
100 Washington Square East,
New York, NY, 10003, USA
e-mail: mark.siegal@nyu.edu

somes is perhaps most dramatically demonstrated by the humpbacked fly *Megaselia scalaris*, in which a male-determining factor transposes on chromosome to chromosome at a low rate, creating a new Y chromosome each time it so moves (Traut 1994). For comparison with the rapid divergence of sex-determining mechanisms, consider that all cyclorrhaphan dipterans undergo embryonic segmentation in the same, stereotypical “long germband” mode (Sander 1976), and that the cascade of regulatory genes executing segmentation in *Drosophila* appears to be conserved among cyclorrhaphans (Stauber et al. 1999). The anomalously rapid divergence of sex-determination mechanisms presents both a fascinating puzzle and a unique opportunity to study the evolution of a key developmental process.

Sex determination in *D. melanogaster* is a model for genetic hierarchies in development. The sequence of steps between the initial assessment of the number of X chromosomes and the deployment of sex-specific transcription factors is extremely well characterized (Cline and Meyer 1996; Fig. 1a), and enormous progress is being made on understanding how this pathway is integrated with major patterning hierarchies to achieve sex-specific differentiation (Christiansen et al. 2002). This wealth of information makes possible a detailed comparative analysis of sex determination in other fly species. For example, the rapid evolution of primary sex-determination signals is reflected in the functional divergence within Diptera of the gene *Sex-lethal* (*Sxl*). *Sxl* occupies a fundamental position at the top of the sex hierarchy in *Drosophila*, interpreting the primary sex-determination signal—the ratio of X chromosomes to sets of autosomes. When the dosage of so-called numerator genes on the X chromosome is sufficient, as it is in XX but not XY embryos, early expression of *Sxl* is established and continued presence of SXL protein is maintained by an autoregulatory splicing loop. SXL thereby initiates a cascade of regulatory interactions that effect female differentiation in chromosomal females (XX; Cline and Meyer 1996). However, in other cyclorrhaphan dipterans examined—the tephritid *C. capitata*, the muscid *M. domestica*, the calliphorid *Chrysomya rufifacies*, and the phorid *M. scalaris*—the homolog of *Sxl* is not expressed sex-specifically (Saccone et al. 1998; Meise et al. 1998; Müller-Holtkamp 1995; Sievert et al. 1997, 2000). Furthermore, ectopic expression in *D. melanogaster* males of either *C. capitata* or *M. domestica* SXL does not activate female differentiation (Saccone et al. 1998; Meise et al. 1998). Therefore, in non-drosophilids *Sxl* is almost certainly not involved in somatic sex determination.

At the other extreme, there is evidence that the *Drosophila* gene *doublesex* (*dsx*)—which functions at the bottom of the sex-determination hierarchy, producing alternatively spliced mRNA isoforms that encode sex-specific transcription factors—is not only functionally conserved in Diptera (Sievert et al. 1997; Kuhn et al. 2000) and Lepidoptera (Suzuki et al. 2003), but is homologous to *mab-3* and *mab-23*, *Caenorhabditis elegans* genes required for aspects of male sexual

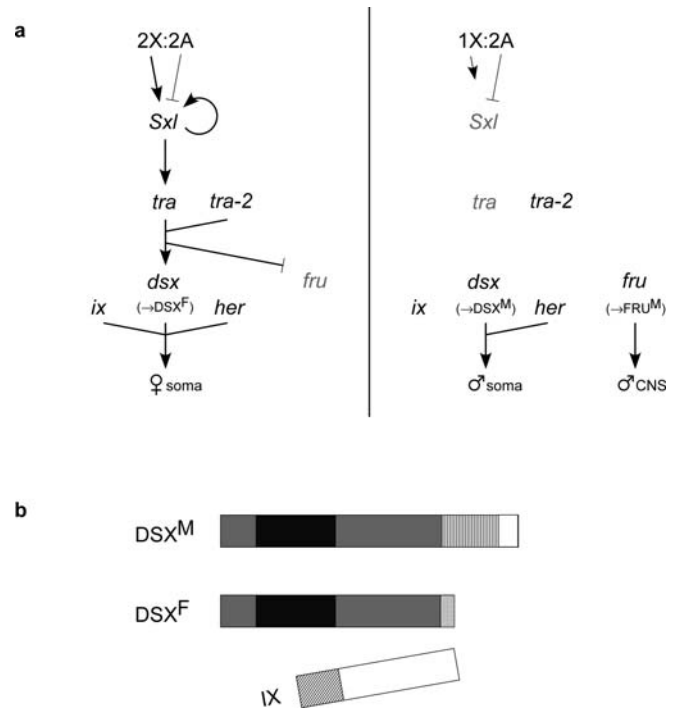


Fig. 1 Sex-determination hierarchy of *Drosophila melanogaster* (a) and schematics of Doublesex (DSX) and Intersex (IX) proteins (b). **a** The primary sex-determination signal in *D. melanogaster* is the ratio of X chromosomes to sets of autosomes. In females (2X:2A), the dosage of numerator genes on the X chromosome is sufficient to activate the early expression of *Sex-lethal* (*Sxl*). Once established, SXL production is maintained by an autoregulatory splicing loop. In males (1X:2A), the dosage of numerator genes is not sufficient to overcome the repressive influence of an autosomal denominator, and therefore SXL is not produced. SXL causes transcripts of *transformer* (*tra*) to be spliced such that an active TRA protein is made. TRA acts with the protein product of *tra-2* to cause transcripts of *dsx* and *fruitless* (*fru*) to be spliced female-specifically, such that a female-specific DSX protein is made (DSX^F) and no female-specific FRU protein is made (certain transcripts of *fru* are essential and not sex-specific, so this figure only considers *fru* expression from the P1 promoter, which produces the sex-specifically processed transcripts). DSX^F acts with the protein product of *ix*, which is transcribed sex-nonspecifically, and the protein product of *hermaphrodite* (*her*), which encodes a zinc-finger transcription factor, to promote female differentiation and repress male differentiation of the soma. In males, the absence of active TRA leads to default, male-specific splicing of *dsx* and *fru* transcripts, such that a male-specific DSX protein is made (DSX^M) and male-specific FRU isoforms are made (FRU^M). DSX^M promotes male differentiation and represses female differentiation of the soma, interacting with HER in some aspects. FRU^M creates the potential for male courtship behavior by promoting male-specific differentiation of the central nervous system. In the figure, genes that produce no active protein in the given sex are shaded gray. **b** The DSX^M and DSX^F proteins are identical over much of their lengths, differing only in their carboxy termini. The common region (gray) contains the DNA-binding DM domain (black). The DSX^M carboxy terminus contains a proline- and serine-rich putative transactivation domain (vertical stripes). The DSX^F carboxy terminus (stippled) lacks such a domain. IX, which forms a specific physical interaction with DSX^F, contains no recognizable DNA-binding domain, but does contain a proline-, glycine-, glutamine- and serine-rich region (diagonal stripes) similar to known transcriptional activators. Unique sequences with unspecified function are in white

differentiation (Raymond et al. 1998; Lints and Emmons 2002), as well as to *Dmrt1*, a mouse gene required for testis development and whose human ortholog, *DMRT1*, is also testis-specifically expressed and is present with other *dsx*-related genes in a region of chromosome 9 associated with an XY sex-reversal condition (Raymond et al. 1999, 2000). Thus, it appears that the bottom of the *Drosophila* sex hierarchy is rather ancient, whereas the top is rapidly diverging.

An intermediate level of divergence is shown by the *transformer* (*tra*) gene, which occupies an intermediate stage in the *Drosophila* sex-determination hierarchy, transmitting the signal from *Sxl* to *dsx*. In addition to affecting splicing of its own pre-mRNA, SXL causes the pre-mRNA of *tra* to be spliced into a female-specific form, which is translated into an active TRA protein; the absence of SXL in males leads to default splicing of *tra* transcripts so that no active TRA is made. Active TRA leads to female-specific splicing of *dsx* transcripts so that a female-specific protein (DSX^F) is made; without TRA there is default splicing of *dsx* transcripts so that a male-specific protein (DSX^M) is made. TRA is a member of the SR family of splicing factors, and *tra* is also one of the most rapidly diverging genes in *Drosophila* at the amino acid sequence level (O'Neil and Belote 1992; Kulathinal et al. 2003). This makes it challenging to identify orthologs of *tra* in non-drosophilid lineages. Nevertheless, the *tra* ortholog of the medfly *C. capitata* has been identified by synteny, and has been shown to act in sex determination. Indeed, *C. capitata* TRA appears to regulate the splicing of its own pre-mRNA; this autoregulatory loop, which resembles that of *Sxl* in *Drosophila*, is blocked in male embryos by the action of a Y-chromosome locus, called *M*, that is yet to be identified molecularly (Pane et al. 2002). *tra* may be acting at the top of the sex-determination hierarchy in the honeybee, *Apis mellifera*, as well. The honeybee, a typical hymenopteran, has diploid females and haploid males, which develop from unfertilized eggs. However, diploids are only female when they are heterozygous at the *complementary sex determiner* (*csd*) locus; fertilized eggs carrying two copies of the same *csd* allele develop as males. It has recently been reported that *csd* encodes an SR protein, which might be the honeybee ortholog of *tra*, and that sex-specific *dsx* transcripts exist in this species as well (Beye et al. 2003). Thus, at the very least, it can be said that SR-protein-mediated generation of sex-specific *dsx* transcripts appears to have much broader taxonomic distribution than any of the mechanisms employed to establish sex-specific activity of these SR proteins themselves (e.g., *Sxl* of *D. melanogaster*, *M* of *C. capitata*, allelic complementarity of *A. mellifera*). Due to the aforementioned rapid sequence divergence, it will be difficult to establish that *csd* is indeed a *tra* ortholog. More dense taxonomic sampling will certainly be necessary, especially as it has been reported that *dsx* transcripts in the silkworm *Bombyx mori*, which as a lepidopteran is more closely related to *Drosophila* than is the honeybee, are sex-specifically spliced but appear to lack features characteristic of TRA regulation (Suzuki et al. 2001).

Wilkins (1995) forwarded a hypothesis, later expanded upon by himself and others (Pomiankowski et al. 2004; Wilkins 2002; Marín and Baker 1998; Schütt and Nöthiger 2000; Shearman 2002), that sex-determination pathways have evolved from the bottom up, and therefore that the later a gene acts in a pathway the more likely it is to be conserved among distantly related species. According to Wilkins's original hypothesis, the addition of a new genetic function at the top of a pathway would be driven by natural selection to maintain an optimal sex ratio. That is, during episodes in which a population's sex ratio becomes skewed toward having too many males or too many females, there would be strong selective pressure to shift the sex ratio toward the under-represented sex, and such a change would most likely happen at the top of the pathway, where a single new (or newly recruited) factor could bias all downstream activities.

The available molecular data on the conservation and divergence of *dsx*, *tra* and *Sxl* support the Wilkins hypothesis, but a number of components of the *D. melanogaster* sex-determination system remain to be investigated in other species. Furthermore, the role of *dsx* homologs in sexual differentiation in diverse taxa seems to suggest broad functional conservation, but there are clear functional distinctions among the *dsx* homologs. For example, as mentioned above, *D. melanogaster dsx* controls nearly all aspects of somatic sex in both males and females, whereas *C. elegans mab-3* and *mab-23* are responsible for partially overlapping subsets of the male (and not female) differentiation program, and mouse *Dmrt1* is required for proper testis differentiation but not for establishing a male fate for the undifferentiated, bipotential gonad.

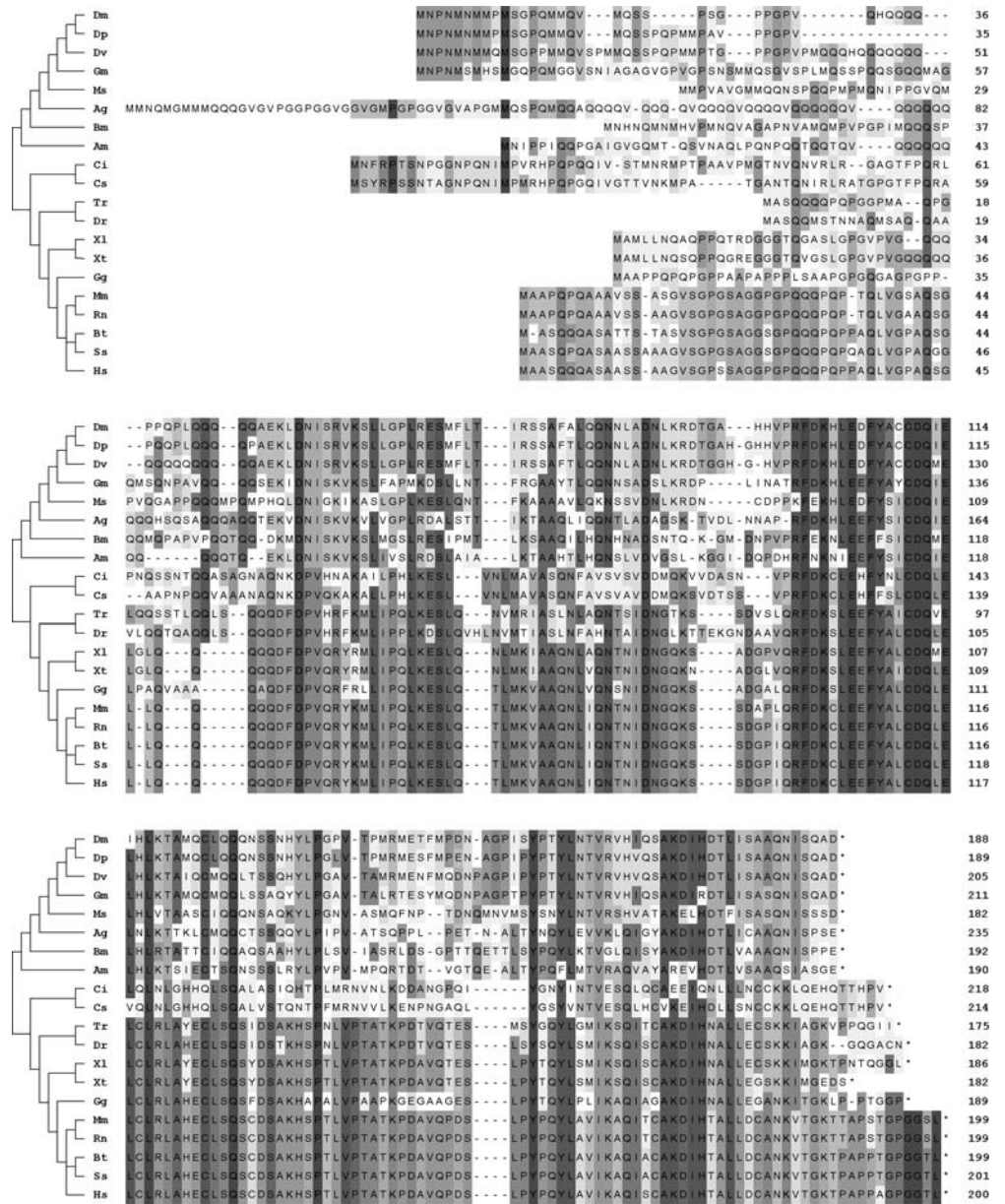
To test further the Wilkins hypothesis of bottom-up evolution of sex-determination pathways, and with the longer term goal of understanding better the functional history of *dsx*-related genes, we chose to identify and characterize homologs of the *D. melanogaster intersex* (*ix*) gene. As shown in Fig. 1a, *ix* acts in concert with *dsx* at the bottom of the sex-determination hierarchy in females. The DSX^M and DSX^F proteins are identical over much of their length (see Fig. 1b). They have identical DNA-binding domains, but differ in their carboxy termini due to TRA-regulated alternative splicing of the *dsx* pre-mRNA. The DSX^M protein is longer, and contains a proline- and serine-rich tail resembling a transactivation domain (Burtis and Baker 1989; Seipel et al. 1992). The DSX^F protein lacks such a domain. In contrast, IX contains no recognizable DNA-binding domain, but does contain a proline-, glycine-, glutamine- and serine-rich region, which resembles known transcriptional activation domains (Garrett-Engle et al. 2002). It was hypothesized therefore that IX and DSX^F might act together in a complex to achieve the same kind of functionality as DSX^M alone. That is, IX might provide DSX^F with a domain corresponding to—but functionally complementary to—the male-specific DSX tail, allowing DSX to activate or repress target genes in a female-specific manner. This hypothesis was strongly supported by Garrett-Engle et al.

(2002), who showed that IX does indeed specifically interact with DSX^F, forming a complex that binds the regulatory region of the *Yolk protein 1* and *2* (*Yp*) genes, which are known DSX targets. A role for IX in direct modulation of transcription is also supported by the recent finding (Sato et al. 2003) that a putative mammalian IX homolog is a subunit of the Mediator complex (see Discussion).

We present here the identification, cloning and characterization of *ix* homologs from diverse metazoan species. As described above, functional comparisons of *Sxl*, *tra* and *dsx* homologs across species have taken advantage of the knowledge that in *D. melanogaster* each of these genes produces sex-specific transcripts and proteins. *D. melanogaster ix*, on the other hand, is not sex-specifically transcribed (Garrett-Engele et al. 2002) and likely produces functional protein in both sexes as

well (Waterbury et al. 1999). Therefore, we chose to test functional conservation by using transgenic assays of homolog function in *D. melanogaster*. While such assays test biochemical conservation of the IX protein, rather than conservation of its biological role, they comprise a critical first step in understanding the evolution of the *ix* gene. We show that *ix* protein does indeed appear to be fairly broadly conserved, more akin to the conservation seen for *dsx* than that seen for *Sxl*. We discuss these results in the context of hypotheses concerning the evolution of sex determination and other developmental pathways, and also with respect to the evolution of the *dsx*-related gene family.

Fig. 2 Amino acid alignment of Intersex (IX) homologs. On the left are the known phylogenetic relationships among the species whose IX homologs are shown (*Dm Drosophila melanogaster*, *Dp D. pseudoobscura*, *Dv D. virilis*, *Gm Glossina morsitans morsitans*, *Ms Megalotia scalaris*, *Ag Anopheles gambiae*, *Bm Bombyx mori*, *Am A. mellifera*, *Ci Ciona intestinalis*, *Cs C. savignyi*, *Tr Takifugu rubripes*, *Dr Danio rerio*, *Xl Xenopus laevis*, *Xt X. tropicalis*, *Gg Gallus gallus*, *Mm Mus musculus*, *Rn Rattus norvegicus*, *Bt Bos taurus*, *Ss Sus scrofa*, *Hs Homo sapiens*). At each position in the alignment, amino acids are shaded by percent identity. Asterisks indicate stop codons



Materials and methods

Identification of *intersex* homologs

Database searches

DNA sequences appearing to encode proteins with significant sequence identity to *D. melanogaster* IX (GenBank AAN37397.1) were identified by translated BLAST (tblastn) searches of available databases. These putative *ix*-homologous sequences were then individually inspected and, when possible, full-length coding sequences were assembled. The first search was of the GenBank non-redundant (nr) database (Release 140, searched 3/8/04), which yielded putative *ix* homologs from *Anopheles gambiae* (represented by cDNA BX066894), *Mus musculus* (AK005112.1), *Rattus norvegicus* (XM_214868.2), *Homo sapiens* (AK000411.1/hIntersex—see Sato et al. 2003), *Gallus gallus* (BX935654.1), and *Danio rerio* (AF506210.1). A subsequent search of 46 metazoan genomes available through NCBI yielded putative homologs from *D. pseudoobscura* (contig AADE01000173.1), *A. mellifera* (contig AADG02020501.1), *Ciona intestinalis* (contig AABS01000063.1), *Ciona savignyi* (contig AACT01000191.1), and *Takifugu rubripes* (contig CAAB01000368.1). The *C. intestinalis* full-length coding sequence was not readily apparent from the genomic sequence alone, so cDNA sequence was also acquired, by searching the *Ciona* genome project EST collection at <http://ghost.zool.kyoto-u.ac.jp/> (databases seqC posted 9/30/02 and seqN posted 9/27/02). This cDNA sequence (represented by clone cicl100o08) aided in identifying the *C. savignyi* coding sequence as well. Similarly, identification of the *Takifugu rubripes* coding sequence was aided by aligning the genomic sequence with the coding sequence of *Danio rerio*. Additional putative *ix* homologs were identified by searches of the GenBank EST divisions. A tblastn search using *D. melanogaster* IX sequence as the query yielded the coding sequence from *Glossina morsitans morsitans* (from overlapping sequences BX557985.1 and BX553813.1). A tblastn search using the inferred protein sequence encoded by the putative *Mus musculus ix* homolog as the query yielded the coding sequences from *Sus scrofa* (BX673104.1) and *Bos taurus* (CB447336.1). A tblastn search using the inferred protein sequence encoded by the putative *Takifugu rubripes ix* homolog as the query yielded sequences from *Salmo salar*, *Oryzias latipes* and *Oncorhynchus mykiss*, but none of these three sequences appeared to contain full-length coding sequence and so are not shown in Fig. 2. Searches of the *Xenopus tropicalis* and *Xenopus laevis* EST databases at <http://www.sanger.ac.uk> (databases posted 3/8/04) yielded the coding sequences of putative *ix* homologs from these two species (from CF782265.1 and CB201403.2, respectively).

Additional putative homologs

The putative *ix* homolog of *D. virilis* was cloned by hybridization of a *D. melanogaster ix* probe to a *D. virilis* genomic DNA library. The probe was made by incorporating [α -³²P]-labeled dATP and dCTP (New England Nuclear, Boston, Mass., USA) into random-hexamer-primed extension products synthesized with the Klenow fragment of DNA polymerase I (Life Technologies, Rockville, Md., USA) using as template an *EcoRI/ClaI* restriction fragment containing full-length *ix* coding sequence from cDNA clone G2 (Garrett-Engle et al. 2002). The *D. virilis* library, constructed in vector λ EMBL3, by Ron K. Blackman, was obtained from Thomas Kaufman (Thummel 1993). Hybridization to lifts of approximately 10⁵ plaques was performed overnight at 42°C in buffer containing 2× SSC, 0.1% sodium dodecyl sulfate (SDS) and 50 μg/ml sonicated salmon sperm DNA. Filters were washed at 55°C, first in 2× SSC/1% SDS, then in 1× SSC/0.1% SDS, then in 0.2× SSC/0.1% SDS, and finally in 0.1× SSC/0.1% SDS, after which they were exposed to film. Two phage clones (named ix10 and ix5) were identified as positive plaques. Restriction enzyme mapping indicated the two clones were overlapping, and Southern hybridization of the *ix* probe to restriction digested clone DNA indicated that a terminal 1.7-kb *EcoRI-SalI* fragment (*SalI* in vector arm) of ix5 and an internal 2.2-kb *EcoRI-SalI* fragment of ix10 contained *D. virilis ix*. These fragments were subcloned into the plasmid vector pBluescript II KS + (Stratagene, La Jolla, Calif., USA). DNA sequencing of these clones confirmed that they did indeed overlap and contain *D. virilis ix*. Alignment of this sequence with *D. melanogaster ix* identified the extent of the *D. virilis* coding sequence (GenBank accession number AY648338).

The putative *ix* homolog of *M. scalaris* was identified by reverse transcription of *M. scalaris* RNA, followed by PCR amplification using degenerate oligonucleotides designed to match presumed conserved regions of the *ix* coding sequence. A wildtype *M. scalaris* stock was provided by Dixie-Ann Persaud (University of Central Florida), and maintained at room temperature on standard *Drosophila* medium, occasionally supplemented with freeze-dried liver (Carolina Biological Supply Company, Burlington, N.C., USA). RNA was isolated from 50 adult *M. scalaris* by TRIzol reagent (Life Technologies), followed by RNeasy (Qiagen, Valencia, Calif., USA). First-strand cDNA was synthesized using the polyT-containing adapter primer AP (3' RACE System, Life Technologies). The cDNA was amplified using Taq polymerase (Life Technologies), RACE System universal adapter primer AUAP and the degenerate oligonucleotide ixDU4 (5'-GTICCMGITYGAIAARMAYYTIGARGANTT), with 40 cycles of 94°C for 1 min, 40°C for 1 min and 72°C for 2 min. Diluted amplification reactions were then reamplified with ixDU4 and the nested primer ixDL (5'-ATIARIGTRTCRTGIATRTCYYTNGC), with the same cycling conditions. Amplification products were cloned into the TOPO-TA pCR II vector (Invitrogen,

Carlsbad, Calif., USA), and the DNA sequence of three clones was obtained. These sequences indicated that the *ix*-homologous sequence had indeed been identified. This sequence was used to obtain the rest of the *M. scalaris* coding sequence. First, 5' RACE was performed (5' RACE System, Life Technologies) as follows: first-strand cDNA was synthesized using primer MsL1 (5'-CGTAGCAA-CATGTGATCTGACAGT), dC-tailed, amplified with primer MsL3 (5'-TGCATAGAAGCTACATTTCCCTGGT) and RACE System adapter primer AAP (35 cycles of 94°C for 1 min, 52°C for 1 min and 72°C for 2 min), and reamplified with nested primer MsL2 (5'-GCAGCTGT-TACCAAATGCAATTCA) and AUAP (35 cycles of 94°C for 1 min, 53°C for 1 min and 72°C for 2 min). Amplification products were cloned into the pCRII vector and sequenced. One clone contained sequence extending 5' to the putative start codon of *M. scalaris ix*. Then, 3' RACE was performed using the already obtained polyT-primed cDNA, as follows: cDNA was amplified using primers MsU1 (5'-ATGCATGGAATGATGCCTGT) and AUAP (30 cycles of 94°C for 1 min, 53°C for 1 min and 72°C for 1 min), then reamplified with nested primer MsU2 (5'-CCCCAAAATTTGAAAAGCACT) and AUAP (30 cycles of 94°C for 1 min, 52°C for 1 min and 72°C for 1 min). Amplification products were cloned into the pCRII vector and sequenced. One clone gave sequence extending 3' to the putative stop codon of *M. scalaris ix*, thereby completing the full-length coding sequence (GenBank accession number AY648339).

The putative *ix* homolog of *Bombyx mori* was first identified by tblastn search of the *B. mori* EST database, SilkBase (<http://www.ab.a.u-tokyo.ac.jp/silkbase/>), using *D. melanogaster IX* sequence as the query. This search yielded a sequence (tesS0692) containing a stretch of 60 nucleotides that, when translated, showed high identity to the query. This nucleotide sequence was used to obtain the full-length coding sequence of *B. mori ix*. *B. mori* fertilized eggs and powdered food medium were obtained from Mulberry Farms (Fallbrook, Calif., USA). When available, fresh young mulberry leaves were also used to feed silkworm larvae. First, RNA was isolated from 30 recently hatched *B. mori* larvae and poly-T-primed cDNA was made as above for *M. scalaris*. The cDNA was amplified using primer BmB (5'-AACCTGCAC-CAGTGCCTCAA) and AUAP (30 cycles of 94°C for 1 min, 53°C for 1 min and 72°C for 2 min), then reamplified with nested primer BmC (5'-TGGGATCCT-TACGCGAATCT) and AUAP (same cycling conditions). Amplification products were cloned into the pCRII vector and sequenced. The rest of the *B. mori ix* coding sequence was obtained by inverse PCR. Genomic DNA was isolated from a single late-stage larva (weighing approximately 100 mg), using the protocol developed by the Berkeley *Drosophila* Genome Project for *Drosophila* inverse PCR (<http://www.fruitfly.org/about/methods/inverse.pcr.html>). This protocol was also followed for restriction digestion of the genomic DNA (complete digestion with *SspI*) and ligation of digested DNA in dilute solution to promote intramolecular ligation. Digested, ligated DNA was

amplified with primers BmA (5'-ATGGGTCCTGGTA-CAGGCAT) and BmC (30 cycles of 94°C for 1 min, 53°C for 1 min and 72°C for 3 min). A 1.8-kb product was obtained, cloned into pCRII, and sequenced, completing the coding sequence (GenBank accession number AY648340). Note that the SilkBase EST collection now includes a longer cDNA match to *ix*, ovS311C09f, which was not present at the time of the original search.

Amino acid sequence alignment

The inferred amino acid sequences encoded by the putative full-length coding sequences obtained as detailed above were aligned using the GCG program Pileup (Wisconsin Package Version 10.3, Accelrys, San Diego, Calif., USA), then adjusted manually and shaded by percent identity in MacVector (version 7.2, Accelrys).

Drosophila melanogaster stocks and transformation

Mutations and chromosomes not referenced are described elsewhere (FlyBase Consortium 2003; <http://flybase.org/>). The *ix*-mutant genotype used in this study was *ix*³/*Df*(2*R*) *en-B*. The *ix*³ allele is a strong hypomorph, and the *Df*(2*R*) *en-B* completely eliminates the *ix* locus (Garrett-Engel et al. 2002). Crosses were carried out at 25°C, and flies were maintained on standard cornmeal-dextrose-agar medium.

The *ix* coding sequences of *D. melanogaster*, *D. virilis*, *M. scalaris*, *B. mori* and *M. musculus* were cloned separately into the *P*-element transformation vector pUASp (Rørth 1998), as detailed below. *P*-element-mediated germline transformation into strain *w*¹¹¹⁸ was performed using the transposase source *pπ25.7wc* (Karess and Rubin 1984) at a concentration of 150 μg/ml and *ix*-containing pUASp constructs at a concentration of 450 μg/ml, following standard techniques (Rubin and Spradling 1982; Spradling and Rubin 1982). Transformed flies were identified by eye color, and lines with single *P*-element insertions were established by crosses with *w*⁻ stocks containing dominantly marked balancer chromosomes. Additional insertions were generated by re-mobilizing existing insertions with the stably integrated genomic source of *P* transposase, $\Delta 2-3$ (Robertson et al. 1988). In most cases, third-chromosome insertions of the UAS constructs were used for subsequent testing. When second-chromosome insertions were used, they were first recombined onto the *ix*³-bearing chromosome.

The pUASp constructs for each species' *ix* gene were made by amplifying the coding sequence using an upper primer containing an *EcoRI* site at its 5' end and a lower primer containing a *NotI* site at its 5' end. Following the *EcoRI* site, the upper primer contained the *Drosophila* consensus translation initiation sequence (CAAC; Cavener 1987), followed immediately by the coding sequence of the homolog, beginning with the start codon. Similarly, the lower primer continued from the *NotI* site into the last few codons of the gene, beginning with the stop codon. The

primers for each species (indicated by first two letters in primer name) were: DmU-RI (5'-AAGAATTCAACATGAATCCCAACATGAACATGATG), DmL-NotI (5'-AAGCGGCCGCTATCAATCAGCCTGCGAAATGTT), DvU-RI (5'-AAGAATTCAACATGAATCCAAATATGAATATGATG), DvL-NotI (5'-AAGCGGCCGCTATCAATCGGCCTGAGATATATT), MsU-RI (5'-AAGAATTCAACATGATGCCTGTGGCTGTAGGGATGATG), MsL-NotI (5'-AAGCGGCCGCTATCAGTCACTCGAGGAAATATTTTG), BmU-RI (5'-AAGAATTCAACATGAATCACAATCAAATGAATATG), BmL-NotI (5'-AAGCGGCCGCTATCATTTCGGGCGGTGATATATTCTG), MmU-RI (5'-AAGAATTCAACATGGCTGCGCCCAACCACAGGCT), MmL-NotI (5'-AAGCGGCCGCTATCAGAGGCTGCCGC-CAGGGCCAGT). Coding sequences were amplified with the appropriate pair of primers, using the high-fidelity eLONGase polymerase mix (Life Technologies). Amplification templates for *M. scalaris* and *B. mori* were the polyT-primed cDNAs described above. Corresponding templates for *D. melanogaster*, *D. virilis* and *Mus musculus* were made by the same protocol as was used for *Megaselia scalaris*, except that *D. virilis* RNA was isolated from 30 adult flies and *M. musculus* RNA was isolated from one adult female liver (kindly provided and processed by Hank Bayle from a wildtype laboratory mouse destined for sacrifice). Reactions for *D. melanogaster* and *M. scalaris* templates were performed with 35 cycles of 94°C for 30 s, 53°C for 30 s and 68°C for 1 min. Reactions for *D. virilis* and *B. mori* templates were performed with 40 cycles of 94°C for 30 s, 40°C for 30 s and 68°C for 1 min. The reaction for the *M. musculus* template was performed with 35 cycles of 94°C for 30 s, 59°C for 30 s and 68°C for 1 min. Amplification products were digested with *EcoRI* and *NotI*, then cloned into *EcoRI*- and *NotI*-digested pBluescript II KS+. Clones were sequenced to identify ones without replication errors. The *D. virilis* sequence had five differences from the previously obtained genomic sequence—four synonymous differences and one insertion of a glutamine codon. As multiple clones of this cDNA were sequenced, we conclude that these differences are not replication errors but are polymorphisms between the *D. virilis* strain used to make the genomic library (Texmelucan) and the strain (Pasadena) used to make the cDNA (see GenBank accession number AY649563 for the cDNA sequence). For the selected clones from *D. melanogaster*, *D. virilis*, *M. scalaris* and *B. mori*, the coding sequence cassette was isolated as a *KpnI-NotI* fragment and inserted into *KpnI*- and *NotI*-digested pUASp, yielding the corresponding species-specific *pP{UASp-ix.S}* constructs (in the text, each construct is identified with the *ix* sequence's species of origin, e.g., *pP{UASp-Dme\ix.S}*; the '.S' suffix serves as a unique identifier for the constructs described in this paper). Because the *M. musculus* coding sequence contains a *KpnI* site, it was isolated as an *EcoRI-NotI* cassette, which then replaced the corresponding fragment of *pP{UASp-Dme\ix.S}*.

Cuticle preparations

Adult flies were preserved in a 3:1 mixture of 70% ethanol and glycerol. To make cuticle preparations, preserved flies were rinsed thoroughly in water, dissected to separate abdomens from thoraces, incubated in 10% NaOH for 1 h at 60°C, rinsed thoroughly in water, rinsed thoroughly in 70% ethanol, dehydrated in 100% isopropanol for 5 min, transferred to Gary's Magic Mountant (Struhl 1981), and mounted on a glass slide under a coverslip.

Results

Putative homologs of *D. melanogaster ix* were identified either by DNA hybridization (phage library screening or degenerate-oligonucleotide RT-PCR) or by database searching, as described in [Materials and methods](#). An alignment of 20 inferred full-length amino acid sequences encoded by the homologs is shown in [Fig. 2](#). All identified homologs are from metazoan species; no plant, fungal, eubacterial or archeal genomes searched contained a sequence with significant similarity to *ix*. Each genome searched contained at most one unique sequence with apparent homology to *ix*. The possibility remains that duplications of *ix* exist in the genomes that have not yet been completely sequenced, but in the absence of the complete sequences of these genomes, as well as those of species representing key lineages missing from the available databases (such as molluscs and non-insect arthropods), we provisionally assume that the *ix* sequences reported here are orthologs. Among the Metazoa, the only lineage lacking *ix*-homologous sequences, and for which complete genome sequence data are available, is the nematode lineage including *C. elegans*, *C. briggsae* and *Brugia malayi*.

Several notable features emerge from the amino acid alignment in [Fig. 2](#). The homologs appear to have a conserved organization, roughly divided into the two regions depicted in [Fig. 1b](#). The amino termini of the proteins (corresponding to positions 1–46 in the *D. melanogaster* sequence) do not align well, because they are low complexity sequences. However, the amino acid composition of this region of the proteins—rich in glutamine, proline, glycine and serine residues—is conserved. The similarity of this region to known transcription regulators (Garrett-Engel et al. 2002) is thus a conserved feature of the IX protein. Carboxy-terminal to this region are several stretches of high amino acid identity. The most striking such block, corresponding to positions 98–117 in the *D. melanogaster* sequence, contains a high proportion of polar amino acids, as well as two absolutely conserved phenylalanine residues. This block does not match any domain of known function, and appears to exist only in these IX homologs. Another conserved feature of the IX homologs is their short length. The *D. melanogaster* IX protein is only 188 amino acids in length. The *A. gambiae* IX protein, due to extended glutamine repeats in its amino-

terminal region, is the longest homolog, yet it is only 234 amino acids in length.

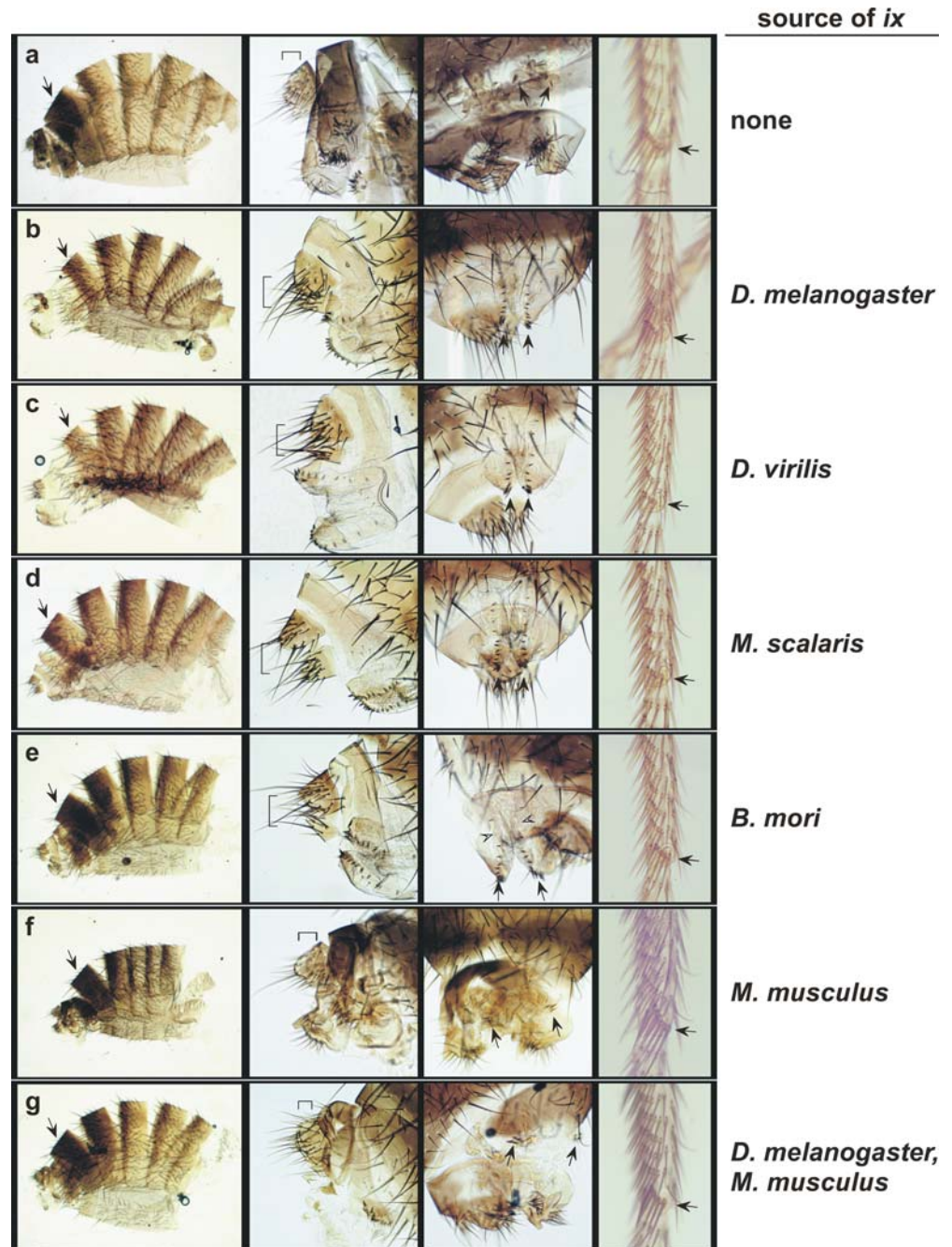
Transgenic tests of functional conservation

To test for functional conservation of the *ix* coding sequences, we expressed a subset of them individually in transgenic *D. melanogaster* females that were mutant at the endogenous *ix* gene. The *ix* promoter has not yet been defined, so we decided to employ the GAL4-UAS binary system (Brand and Perrimon 1993) for driving expression of transgenic homolog cDNAs. At the outset of the

experiment it was unclear what timing and spatial pattern of *ix* expression would be required for full function, but the GAL4-UAS system gave us the flexibility to try various GAL4 drivers. For the experiments described here, we used the *tubulin 1 α* (α *Tub84B*) promoter driving GAL4 ($P\{w^{+mC}=tubP-GAL4\}LL7$, abbreviated as $P\{tubP-GAL4\}$; Lee and Luo 1999), which gives essentially constitutive expression throughout development. The *ix* cDNAs were cloned downstream of the UAS in pUASp (Rørth 1998), enabling their activation by GAL4.

The *ix* homologs used for these experiments, chosen to represent a range of species divergence times and primary sex-determining mechanisms, derive from *D. melanogaster*-

Fig. 3 Cuticular phenotypes of *Drosophila melanogaster* females expressing cDNAs corresponding to *ix* homologs. See text for full genotypes. **a** Mutant females (*Df(2R)en-B/ix³*) have male-like pigmentation of tergite 6 (arrow, panel 1, lateral view of abdomen), laterally shifted anal plates (bracket, panel 2, lateral view), few, disorganized vaginal teeth (arrows, panel 3, ventral view), and last transverse row bristles (LTRB) that are thickened, darkened, increased in number, and rotated 45° relative to the other transverse rows (arrow, panel 4). **b–d** Mutant females expressing *D. melanogaster* (**b**), *D. virilis* (**c**), or *Megaselia scalaris* (**d**) *ix* cDNA appear wildtype, with female-like pigmentation (panel 1), dorsal and ventral anal plates (panel 2), ordered rows of vaginal teeth (panel 3), and LTRB that are similar in appearance, number and orientation to the other transverse row bristles (panel 4). **e** Mutant females expressing *Bombyx mori* *ix* cDNA appear slightly intersexual, with male-like pigmentation (panel 1), dorsal and ventral anal plates (panel 2), ordered rows of vaginal teeth (panel 3) with some missing bristles (open arrowheads), and LTRB that are slightly thicker and darker than the other transverse rows and also slightly rotated upward (panel 4). **f–g** Females homozygous (**f**) or heterozygous (**g**) for *ix* and expressing *Mus musculus* *ix* cDNA are intersexual, yet differ from mutant females not expressing *M. musculus* *ix* cDNA in that their LTRB are curved upward past a 45° angle (panel 4)



ter (positive control); the distantly related *Drosophila* species, *D. virilis*; the basal cyclorrhaphan fly, *M. scalaris*; the lepidopteran, *B. mori*; and the mammal, *M. musculus*. The corresponding UASp constructs (see [Materials and methods](#)) are referred to as $pP\{UASp-Dme\backslash ix.S\}$, $pP\{UASp-Dvir\backslash ix.S\}$, $pP\{UASp-Msca\backslash ix.S\}$, $pP\{UASp-Bmor\backslash ix.S\}$ and $pP\{UASp-Mmus\backslash ix.S\}$.

D. melanogaster females mutant for *ix* display characteristic morphological defects and are sterile (Chase and Baker 1995; Garrett-Engel et al. 2002). Wildtype females lack the dark pigmentation on abdominal tergites 5 and 6 that is characteristic of wildtype males. Female *ix* mutants have male-like pigmentation, especially on tergite 6 (Fig. 3a panel 1). Wildtype females have one dorsal and one ventral anal plate, whereas wildtype males have one left and one right anal plate. In *ix* females, the ventral plate develops incompletely and the dorsal plate is split such that each side is shifted laterally toward a male-like fate, although the plates commonly remain fused on the dorsal side (Fig. 3a panel 2). There are a number of sex-specific genital bristles that distinguish wildtype females and males. Among the most prominent are the gonopod thorn bristles (commonly known as “vaginal teeth”), which form two parallel rows on either side of the vaginal opening. In *ix* females, there is a dramatic reduction in the number of vaginal teeth, and they are no longer present in ordered rows (Fig. 3a panel 3). Bristles on the foreleg are also sexually dimorphic. In wildtype females, the last (most distal) transverse row of bristles on the basitarsus of the foreleg (LTRB) typically has 5–6 bristles that are similar in appearance and orientation to the bristles of the other transverse rows on that leg segment. In wildtype males, these bristles form a sex comb, which consists of approximately ten bristles that are thicker and more darkly pigmented than the transverse row bristles and are oriented perpendicular to the other rows. Female *ix* mutants have an intermediate phenotype: there are approximately 7–8 LTRB bristles, they are thicker and more darkly pigmented than the corresponding female bristles but not quite as thick and dark as the corresponding male bristles, and the row is set at a 45° angle with respect to the other rows (Fig. 3a panel 4).

The efficacy of the GAL4-UAS approach is evident in Fig. 3b, which shows females carrying the same *ix* mutations as in Fig. 3a (a strong hypomorph over a deficiency, see [Materials and methods](#)) but also expressing wildtype *D. melanogaster ix* by virtue of carrying the *tubulin*-promoter GAL4 driver and the UASp-*ix* construct (four independent third-chromosome UASp insertions tested; full genotype: $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/P\{UASp-Dme\backslash ix.S\}$). The transgenically expressed cDNA fully rescues the mutants, restoring fertility as well as proper female abdominal pigmentation, anal plate orientation, vaginal teeth number and placement, and LTRB bristle number and appearance.

Expressing the *D. virilis* cDNA (three independent UASp insertions tested; full genotype $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/P\{UASp-Dvir\backslash ix.S\}$ or $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/TM6B, Tb^1, P$

$\{UASp-Dvir\backslash ix.S\}$) also fully rescues the mutants, restoring fertility as well as proper female abdominal pigmentation, anal plate orientation, vaginal teeth number and placement, and LTRB bristle number and appearance (Fig. 3c). The same is true of the *M. scalaris* cDNA (four independent UASp insertions tested; full genotype: $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/P\{UASp-Msca\backslash ix.S\}$ or $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/TM6B, Tb^1, P\{UASp-Msca\backslash ix.S\}$; Fig. 3d).

The *ix* cDNA from the silkworm *B. mori* does not fully restore wildtype female development. Four independent insertions of the UAS construct were tested, none of which completely rescued the *ix*-mutant females ($w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/P\{UASp-Bmor\backslash ix.S\}$ or $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/TM6B, Tb^1, P\{UASp-Bmor\backslash ix.S\}$). Moreover, wildtype female development was not fully restored even when the UAS construct was present in two copies ($w^{1118}/w^{1118}; P\{UASp-Bmor\backslash ix.S\}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/TM6B, Tb^1, P\{UASp-Bmor\backslash ix.S\}$). Despite expression of the *B. mori ix*, fertility was severely reduced, although a small number of progeny of these double-UAS females was recovered. These females also manifest some morphological defects. Pigmentation of tergite 6 is male-like (Fig. 3e panel 1). The dorsal and ventral anal plates, however, are present and appear normal (Fig. 3e panel 2). The vaginal teeth are present in two neat rows but there are gaps where bristles would be expected to be present (Fig. 3e panel 3). Likewise, the LTRB appear mostly female-like, but are slightly darker and slightly rotated upward (Fig. 3e panel 4). This nearly complete rescue of the *ix*-mutant females suggests partial functional divergence between the *D. melanogaster* and *B. mori IX* proteins.

Expressing the *ix* cDNA from the mouse *M. musculus* does not restore wildtype female development at all. Four insertions of the UAS construct were tested, and none yielded a phenotype resembling a wildtype female ($w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/P\{UASp-Mmus\backslash ix.S\}$ or $w^{1118}/w^{1118}; Df(2R)en-B/ix^3; P\{UASp-Mmus\backslash ix.S\}; P\{tubP-GAL4\}/+$ or $w^{1118}/w^{1118}; P\{UASp-Mmus\backslash ix.S\}; Df(2R)en-B/ix^3; P\{tubP-GAL4\}/+$). The pigmentation of tergite 6 was male-like (Fig. 3f panel 1), the anal plates were shifted laterally (Fig. 3f panel 2), the vaginal teeth were reduced in number and not present in rows (Fig. 3f panel 3) and the LTRB were darker, thicker and rotated upwards. Interestingly, expressing the mouse cDNA in females carrying one wildtype copy of the endogenous *ix* gene (any of the above genotypes, except at the endogenous *ix* locus: $Df(2R)en-B/+$ or $+/ix^3$) yielded a strong intersexual transformation. Indeed, *D. melanogaster* females heterozygous for endogenous *ix* and expressing mouse *ix* are indistinguishable from females homozygous mutant for *ix* and expressing mouse *ix* (Fig. 3g). This suggests that mouse *ix* acts as a dominant-negative in *D. melanogaster*, a conclusion supported by the observation that females homozygous wildtype for *ix* and expressing mouse *ix* have much less severe morphological defects (not shown).

IX has been shown to dimerize both with itself and with DSX^F (Garrett-Engle et al. 2002), so it is possible that mouse IX retains the capacity for one or both of these interactions, yet is incapable of having the proper activating or repressing effect on target genes. One indication that mouse IX is misregulating target genes comes from closer comparison of the LTRB on females expressing mouse *ix* (Fig. 3f, g) with those on *ix*-mutant females not expressing mouse *ix* (Fig. 3a). As noted above, the canonical intersexual LTRB phenotype is a straight row of bristles, rotated 45° to the proximal-distal axis of the leg. The corresponding bristle row in the two genotypes expressing mouse *ix* is not straight. Indeed it is curved upward beyond a 45° angle. This observation is inconsistent with mouse IX acting solely to interfere with endogenous IX or DSX^F function, because then it would at most yield a phenotype corresponding to complete loss of *ix* or *dsx* function (i.e., a straight row of bristles at a 45° angle). Instead, the bristles are transformed more toward a sex-comb fate, suggesting that mouse IX might be misregulating one or more target genes in this differentiating tissue, shifting its fate to be slightly more male than intersexual. As *dsx* target genes become elucidated, this possibility can be tested more directly.

Discussion

Because of the sheer variety of sex-determination mechanisms, it is tempting to consider their evolution as an intellectual curiosity, rather than as a model for the evolution of other developmental pathways. Indeed, there are reasons to suppose that sex-determination pathways are under different constraints than other developmental pathways, the most notable being that in many systems, mutating a key sex-determination gene will yield a viable, fertile individual of the “default” sex, thereby causing little or no reproductive disadvantage (Marín and Baker 1998). Nonetheless, there are also important features shared by sex-determination and other pathways that merit attention. As has been noted (Wilkins 1995, 2002; Marín and Baker 1998), particularly relevant are considerations of the topology of the regulatory hierarchy. That is, the place at which a new factor may be added depends upon how the existing factors are connected to one another and to the ultimate phenotype. This truism clearly holds for sex-determination and other developmental pathways alike, but what is striking is that the core of Wilkins’s “bottom-up” hypothesis of sex-determination evolution—conservation of later-acting regulatory factors despite divergence of early-acting factors—is also the core of prominent views concerning the evolution of embryonic patterning, specifically the “hourglass” metaphor of Raff (1996) and the related, but more narrowly defined, “egg-timer” metaphor of Duboule (1994), both of which seek to explain the relative fixity of animal body plans despite the evolutionary lability of early development. Thus, sex determination might indeed serve as a valuable model for the evolution of other hierarchies, especially because its rapid

divergence enables high-resolution comparative analysis using more closely related species than would be meaningful when studying, for example, body-plan divergence.

We have presented here evidence that the IX protein, which in *D. melanogaster* physically interacts with the DSX^F protein, is rather broadly conserved in function. Homologs deriving from species well beyond the genus *Drosophila* are capable of functionally replacing the *D. melanogaster* gene. In our transgenic assays, expression of the *ix* cDNA of the basal cyclorrhaphan fly, *M. scalaris*, restores proper sexual development in otherwise mutant females, to an extent indistinguishable from that seen with expression of the *ix* cDNA of *D. melanogaster* itself. More distant homologs show partial divergence. Expression of the *B. mori ix* cDNA only partially restores female differentiation in *ix*-mutant female *D. melanogaster*. One possible explanation of the partial rescue is that the IX and DSX^F proteins co-evolve, such that their interaction is imperfect when they are taken from highly divergent species. It will be interesting to explore this possibility using genetic and biochemical assays of the interaction of IX and DSX^F pairs taken from species spanning a range of divergence times. It is also intriguing that the sexual phenotype most sensitive to the replacement of *D. melanogaster ix* with *B. mori ix* is the pigmentation of tergite 6. Abdominal pigmentation varies considerably among *Drosophila* species, and the sexually dimorphic *D. melanogaster* pattern has arisen multiple times independently (Gompel and Carroll 2003). One might therefore expect that such a recently evolved link in the sexual differentiation program would best reveal differences between distantly related homologous regulatory proteins, because there has not been long-term, constant selection to maintain this particular interaction between the regulator (s) and target(s).

The distant history of IX, its interacting regulatory proteins, and its targets remains largely unknown, but our analysis of mouse IX suggests at least some conservation of function among quite distantly related metazoans. Whereas the *Mus musculus ix* cDNA does not restore female differentiation when expressed in *ix*-mutant female *D. melanogaster*, there is evidence of conservation of its biochemical function, in that *D. melanogaster* females heterozygous for *ix* are wildtype in their sexual development, but are transformed to intersexes by expression of the mouse *ix*. The mouse IX thus appears to retain some ability to interact with one or more components of the endogenous sex-determination hierarchy—probably IX or DSX^F—even though it is not able to regulate target genes properly. A recent study has shown that the human IX homolog is a component of the Mediator transcriptional coactivation complex (Sato et al. 2003). It is not yet known what developmental processes require IX in mammals (mouse *ix* is transcribed in both soma and gonads of both male and female adults; data not shown), but it is notable that the best-scoring IX-interacting protein in a high-throughput yeast two-hybrid assay of the *D. melanogaster* proteome (Giot et al. 2003) is Nut2, a conserved component of the fly and mammalian Mediator

complexes. Thus, even though the biological function of IX might turn out to be different in divergent animal lineages, it might have a common mode of action, linking specific DNA-binding proteins to the general transcriptional machinery. More complete elucidation of IX-interacting proteins in different lineages will be needed to test this speculation.

Good candidates in vertebrates for potential IX interaction partners are the members of the DSX-related DM family of proteins (at least eight family members exist in humans, seven in mice), most of which are largely uncharacterized functionally (Volff et al. 2003). *Dmrt* genes can be divided into two classes, those containing a DSX^M-like proline- and serine-rich domain at the carboxy terminus and those resembling DSX^F in that they lack such a domain (Ottolenghi et al. 2002). Members of the latter class make better candidates for IX-binding partners, as they likely lack the ability to activate transcription on their own. In the tilapia fish, *Oreochromis niloticus*, the *Dmrt1* ortholog, which contains a proline- and serine-rich domain, is expressed exclusively in the testis, whereas another DM family protein, *tDMO*, lacks the proline- and serine-rich domain and is expressed exclusively in the ovary (Guan et al. 2000). The tilapia experiments were performed only on adults and no distinction was made between germline and somatic expression. Nevertheless, this finding raises the enticing possibility of a rather ancient role for *dsx* homologs in both male and female sexual differentiation. However, no human or mouse *dsx* homolog has yet been shown to have a role in female differentiation. A possible candidate for such a role is *Dmrt7*, which is expressed in the early gonad of XX but not XY mice, but is expressed later in development in both ovaries and testes (Kim et al. 2003). Thus, it remains unclear when—and how many independent times—a *dsx*-related gene acquired a role in female differentiation (Ottolenghi et al. 2002; Kim et al. 2003). Obviously, it is also unclear when (and how many times) an *ix* homolog acquired a role in female differentiation. Even if IX homologs in vertebrates were to interact with one or more DM family members, they would not necessarily be functioning in sexual development. Several mouse *Dmrt* genes are indeed expressed sex-differentially in the developing gonad (Kim et al. 2003), but as yet only *Dmrt1* has been shown to be required for proper sexual differentiation. Furthermore, it has been reported that *Dmrt2* is not required for sexual differentiation in the mouse, and its zebrafish ortholog, *terra*, is involved in somite development, so the biological functions of DM proteins might be quite diverse (Volff et al. 2003).

The potential diversity of biological processes in which IX homologs participate is at least as great as that in which DSX homologs participate, because it is possible that IX homologs have interaction partners beyond the DM protein family. Therefore, it will be important to know the phenotypic consequences of loss of *ix* function in *Mus musculus*, and, for that matter, in species such as *Megaselia scalaris* and *B. mori*, because the ability of a homolog to replace *D. melanogaster ix* does not guarantee

that the homolog is functioning in sex determination (and only in sex determination) in the donor species. Such further biochemical and genetic investigations into IX function will provide a more direct test of the Wilkins (1995) hypothesis of bottom-up evolution of the sex-determination pathway. More specifically, they hold the promise of illuminating the evolutionary diversification of *dsx*-related genes, as well as the evolutionary origin of the *Drosophila* sex-determination hierarchy.

Acknowledgements The authors thank Dixie-Ann Persaud, Hank Bayle, Thomas Kaufman, Robin Cooper, Douglas Harrison, and Liqun Luo and his laboratory members for reagents; members of the Baker lab for helpful advice; Cindy Smith for maintaining our mulberry tree; Peter Ray for donating mulberry leaves from his tree; Guennet Bohm for laboratory assistance; Dmitri Petrov and two anonymous reviewers for helpful comments; and Rachel Drysdale, Kathy Matthews and Bill Gelbart for guidance in naming transgenic constructs. This work was supported by NSF/Sloan Foundation and NIH NRSA postdoctoral fellowships awarded to M.L.S. and by an NIH grant to B.S.B.

References

- Beye M, Hasselmann M, Fondrk MK, Page RE Jr, Omholt SW (2003) The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* 114:419–429
- Brand AH, Perrimon N (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118:401–415
- Bull JJ (1983) Evolution of sex determining mechanisms. Benjamin/Cummings, Menlo Park
- Burtis KC, Baker BS (1989) *Drosophila doublesex* gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* 56:997–1010
- Carroll SB, Grenier JK, Weatherbee SD (2001) From DNA to diversity: molecular genetics and the evolution of animal design. Blackwell, Malden
- Cavener DR (1987) Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res* 15:1353–1361
- Chase BA, Baker BS (1995) A genetic analysis of *intersex*, a gene regulating sexual differentiation in *Drosophila melanogaster* females. *Genetics* 139:1649–1661
- Christiansen AE, Keisman EL, Ahmad SM, Baker BS (2002) Sex comes in from the cold: the integration of sex and pattern. *Trends Genet* 18:510–516
- Cline TW, Meyer BJ (1996) Vive la différence: males vs females in flies vs worms. *Annu Rev Genet* 30:637–702
- Duboule D (1994) Temporal colinearity and the phylotypic progression: a basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development Suppl*:135–142
- FlyBase Consortium (2003) The FlyBase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Res* 31:172–175
- Garrett-Engle CM, Siegal ML, Manoli DS, Williams BC, Li H, Baker BS (2002) *intersex*, a gene required for female sexual development in *Drosophila*, is expressed in both sexes and functions together with *doublesex* to regulate terminal differentiation. *Development* 129:4661–4675

- Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N, Carroll S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL Jr, White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna MP, Chant J, Rothberg JM (2003) A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736
- Gompel N, Carroll SB (2003) Genetic mechanisms and constraints governing the evolution of correlated traits in drosophilid flies. *Nature* 424:931–935
- Guan G, Kobayashi T, Nagahama Y (2000) Sexually dimorphic expression of two types of DM (*Doublesex/Mab-3*)-domain genes in a teleost fish, the Tilapia (*Oreochromis niloticus*). *Biochem Biophys Res Commun* 272:662–666
- Karess RE, Rubin GM (1984) Analysis of P transposable element functions in *Drosophila*. *Cell* 38:135–146
- Kim S, Kettlewell JR, Anderson RC, Bardwell VJ, Zarkower D (2003) Sexually dimorphic expression of multiple *doublesex*-related genes in the embryonic mouse gonad. *Gene Expr Patterns* 3:77–82
- Kuhn S, Sievert V, Traut W (2000) The sex-determining gene *doublesex* in the fly *Megaselia scalaris*: conserved structure and sex-specific splicing. *Genome* 43:1011–1020
- Kulathinal RJ, Skwarek L, Morton RA, Singh RS (2003) Rapid evolution of the sex-determining gene, *transformer*: structural diversity and rate heterogeneity among sibling species of *Drosophila*. *Mol Biol Evol* 20:441–452
- Lee T, Luo L (1999) Mosaic analysis with a repressible cell marker for studies of gene function in neuronal morphogenesis. *Neuron* 22:451–461
- Lints R, Emmons SW (2002) Regulation of sex-specific differentiation and mating behavior in *C. elegans* by a new member of the DM domain transcription factor family. *Genes Dev* 16:2390–2402
- Marin I, Baker BS (1998) The evolutionary dynamics of sex determination. *Science* 281:1990–1994
- McAlpine JF, Wood DM (1989) Manual of Nearctic Diptera, vol. 3. Research Branch, Agriculture Canada
- Meise M, Hilfiker-Kleiner D, Dubendorfer A, Brunner C, Nothiger R, Bopp D (1998) *Sex-lethal*, the master sex-determining gene in *Drosophila*, is not sex-specifically regulated in *Musca domestica*. *Development* 125:1487–1494
- Müller-Holtkamp F (1995) The *Sex-lethal* gene homologue in *Chrysomya rufifacies* is highly conserved in sequence and exon-intron organization. *J Mol Evol* 41:467–477
- O’Neil MT, Belote JM (1992) Interspecific comparison of the *transformer* gene of *Drosophila* reveals an unusually high degree of evolutionary divergence. *Genetics* 131:113–128
- Ottolenghi C, Fellous M, Barbieri M, McElreavey K (2002) Novel paralogy relations among human chromosomes support a link between the phylogeny of *doublesex*-related genes and the evolution of sex determination. *Genomics* 79:333–343
- Pane A, Salvemini M, Delli Bovi P, Polito C, Saccone G (2002) The *transformer* gene in *Ceratitidis capitata* provides a genetic basis for selecting and remembering the sexual fate. *Development* 129:3715–3725
- Pomiankowski A, Nöthiger R, Wilkins A (2004) The evolution of the *Drosophila* sex-determination pathway. *Genetics* 166:1761–1773
- Raff RA (1996) The shape of life: genes, development, and the evolution of animal design. University of Chicago Press, Chicago, Ill.
- Raymond CS, Shamu CE, Shen MM, Seifert KJ, Hirsch B, Hodgkin J, Zarkower D (1998) Evidence for evolutionary conservation of sex-determining genes. *Nature* 391:691–695
- Raymond CS, Parker ED, Kettlewell JR, Brown LG, Page DC, Kusz K, Jaruzelska J, Reinberg Y, Flejter WL, Bardwell VJ, Hirsch B, Zarkower D (1999) A region of human chromosome 9p required for testis development contains two genes related to known sexual regulators. *Hum Mol Genet* 8:989–996
- Raymond CS, Murphy MW, O’Sullivan MG, Bardwell VJ, Zarkower D (2000) *Dmrt1*, a gene related to worm and fly sexual regulators, is required for mammalian testis differentiation. *Genes Dev* 14:2587–2595
- Robertson HM, Preston CR, Phillis RW, Johnson-Schlitz DM, Benz WK, Engels WR (1988) A stable genomic source of P element transposase in *Drosophila melanogaster*. *Genetics* 118:461–470
- Rørth P (1998) Gal4 in the *Drosophila* female germline. *Mech Dev* 78:113–118
- Rubin GM, Spradling AC (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science* 218:348–353
- Saccone G, Peluso I, Artiaco D, Giordano E, Bopp D, Polito LC (1998) The *Ceratitidis capitata* homologue of the *Drosophila* sex-determining gene *Sex-lethal* is structurally conserved, but not sex-specifically regulated. *Development* 125:1495–1500
- Sander K (1976) Specification of the basic body pattern in insect embryogenesis. *Adv Insect Physiol* 12:125–238
- Sato S, Tomomori-Sato C, Banks CAS, Parmely TJ, Sorokina I, Brower CS, Conaway RC, Weliky Conaway J (2003) A mammalian homolog of *Drosophila melanogaster* transcriptional coactivator Intersex is a subunit of the mammalian Mediator complex. *J Biol Chem* 278:49671–49674
- Schütt C, Nöthiger R (2000) Structure, function and evolution of sex-determining systems in Dipteran insects. *Development* 127:667–677
- Seipel K, Georgiev O, Schaffner W (1992) Different activation domains stimulate transcription from remote (‘enhancer’) and proximal (‘promoter’) positions. *Embo J* 11:4961–4968
- Shearman DCA (2002) The evolution of sex determination systems in dipteran insects other than *Drosophila*. *Genetica* 116:25–43
- Sievert V, Kuhn S, Traut W (1997) Expression of the sex determining cascade genes *Sex-lethal* and *doublesex* in the phorid fly *Megaselia scalaris*. *Genome* 40:211–214
- Sievert V, Kuhn S, Paululat A, Traut W (2000) Sequence conservation and expression of the *Sex-lethal* homologue in the fly *Megaselia scalaris*. *Genome* 43:382–390
- Spradling AC, Rubin GM (1982) Transposition of cloned P elements into *Drosophila* germ line chromosomes. *Science* 218:341–347
- Staubert M, Jäckle H, Schmidt-Ott U (1999) The anterior determinant *bicoid* of *Drosophila* is a derived *Hox* class 3 gene. *Proc Natl Acad Sci USA* 96:3786–3789
- Struhl G (1981) Anterior and posterior compartments in the proboscis of *Drosophila*. *Dev Biol* 84:372–385
- Suzuki MG, Ohbayashi F, Mita K, Shimada T (2001) The mechanism of sex-specific splicing at the *doublesex* gene is different between *Drosophila melanogaster* and *Bombyx mori*. *Insect Biochem Mol Biol* 31:1201–1211
- Suzuki MG, Funaguma S, Kanda T, Tamura T, Shimada T (2003) Analysis of the biological functions of a *doublesex* homologue in *Bombyx mori*. *Dev Genes Evol* 213:345–354
- Thummel C (1993) Compilation of *Drosophila* cDNA and genomic libraries. *Dros Info Serv* 72:180–183
- Traut W (1994) Sex determination in the fly *Megaselia scalaris*, a model system for primary steps of sex chromosome evolution. *Genetics* 136:1097–1104
- Volff J-N, Zarkower D, Bardwell VJ, Schartl M (2003) Evolutionary dynamics of the DM domain gene family in metazoans. *J Mol Evol* 57(Suppl1):S241–S249
- Waterbury JA, Jackson LL, Schedl P (1999) Analysis of the *Doublesex* female protein in *Drosophila melanogaster*. Role in sexual differentiation and behavior and dependence on Intersex. *Genetics* 152:1653–1667
- Wilkins AS (1995) Moving up the hierarchy: a hypothesis on the evolution of a genetic sex determination pathway. *BioEssays* 17:71–77
- Wilkins AS (2002) The evolution of developmental pathways. Sinauer, Sunderland, Mass.
- Willhoelt U, Franz G (1996) Identification of the sex-determining region of the *Ceratitidis capitata* Y chromosome by deletion mapping. *Genetics* 144:737–745