



## A male-specific *doublesex* isoform reveals an evolutionary pathway of sexual development via distinct alternative splicing mechanisms

Caihong Han<sup>1,3</sup>, Qionglin Peng<sup>1,3</sup>, Xiangbin Su<sup>1</sup>, Limin Xing<sup>1</sup>, Xiaoxiao Ji<sup>1</sup> & Yufeng Pan<sup>1,2</sup>  

The *doublesex/mab-3* related transcription factor (*Dmrt*) genes regulate sexual development in metazoans. Studies of the *doublesex* (*dsx*) gene in insects, in particular *Drosophila melanogaster*, reveal that alternative splicing of *dsx* generates sex-specific Dsx isoforms underlying sexual differentiation. Such a splicing-based mechanism underlying sex-specific *Dmrt* function is thought to be evolved from a transcription-based mechanism used in non-insect species, but how such transition occurs during evolution is not known. Here we identified a male-specific *dsx* transcript (*dsx*<sup>M2</sup>) through intron retention (IR), in addition to previously identified *dsx*<sup>M</sup> and *dsx*<sup>F</sup> transcripts through alternative polyadenylation (APA) with mutually exclusive exons. We found that Dsx<sup>M2</sup> had similarly masculinizing function as Dsx<sup>M</sup>. We also found that the IR-based mechanism generating sex-specific *dsx* transcripts was conserved from flies to cockroaches. Further analysis of these *dsx* transcripts suggested an evolutionary pathway from sexually monomorphic to sex-specific *dsx* via the sequential use of IR-based and APA-based alternative splicing.

<sup>1</sup>The Key Laboratory of Developmental Genes and Human Disease, School of Life Science and Technology, Southeast University, Nanjing 210096, China. <sup>2</sup>Co-innovation Center of Neuroregeneration, Nantong University, Nantong 226019, China. <sup>3</sup>These authors contributed equally: Caihong Han, Qionglin Peng. ✉email: [pany@seu.edu.cn](mailto:pany@seu.edu.cn)

Sexually dimorphic traits and behaviors are mediated by sex-specific expression of regulatory genes. Substantial studies reveal that the *doublesex/mab-3* related transcription factor (*Dmrt*) genes have been found in all studied animal models and humans for sexual differentiation<sup>1–3</sup>. *Dmrt* transcription factors share a common DM (Dsx/Mab-3) domain which is a zinc-finger DNA binding motif, but show little sequence conservation in other parts<sup>1</sup>. In humans, *Dmrt1* is linked to disorders of sex development and testicular germ cell tumors<sup>4,5</sup>. *Dmrt1* is preferentially expressed in the male gonad and required postnatally to maintain gonadal sex in mammals and other vertebrates<sup>2,6–8</sup>. In the nematode worm *Caenorhabditis elegans*, *mab-3* is mainly expressed in male tails and necessary for proper morphogenesis and differentiation of copulatory structures<sup>9</sup>. In the fruit fly *Drosophila melanogaster*, sex-specific Dsx isoforms, Dsx<sup>M</sup> in males and Dsx<sup>F</sup> in females, control sexual differentiation and behaviors in both sexes, and are generated through alternative polyadenylation (APA) with mutually exclusive exons<sup>10–14</sup>.

Despite conserved roles of *Dmrt* genes in sexual differentiation in metazoans, regulatory mechanisms underlying their sex-specific function are diversified<sup>1–3,15,16</sup>. The alternative splicing-based regulation of *Dmrt* genes has only been found in insects and often mediated by the female-specific expression of *transformer* (*tra*) with few exceptions<sup>10,16–19</sup>. In vertebrates and nematodes, *Dmrt* genes are not sex-specifically spliced, but transcribed in a predominantly male-specific manner for the development of male-specific traits<sup>2</sup>. It has been proposed that the insect-specific mechanism based on alternative splicing of *dsx* evolved from a more ancient mechanism based on male-specific transcription<sup>16,17</sup>, but how such transition occurs during evolution is rarely known.

The canonical insect *tra-dsx* pathway generates sex-specific Dsx isoforms by the selective use of distinct 3' exons in the two sexes<sup>16,17</sup>. In *D. melanogaster* females, Tra and Tra-2 bind to the female-specific exon of *dsx* and direct female-specific splicing of *dsx<sup>F</sup>*, while in males the absence of Tra function permits the default splicing of *dsx<sup>M</sup>* including both common and male-specific exons<sup>10,20–22</sup>. Dsx<sup>M</sup> and Dsx<sup>F</sup> have common DM domain and two oligomerization domains, thus may bind to the same sets of target genes; however, they have been found to oppositely regulate target gene expression to establish male- or female-specific differentiation, possibly through their sex-specific C-terminus<sup>23–26</sup>.

In this study, we report that intron retention (IR) generates a male-specific *dsx* transcript and propose an evolutionary pathway from sexually monomorphic to sex-specific splicing of *dsx*. We identify another male-specific *dsx* transcript (*dsx<sup>M2</sup>*), in which the intron linking the last common exon and the female-specific exon is retained. The IR-generated Dsx<sup>M2</sup> has a masculinizing role like Dsx<sup>M</sup> and is crucial for male courtship robustness in *D. melanogaster*. We further show that such mechanism generating sex-specific *dsx* transcripts is deeply conserved in insects from fruit flies to cockroaches. The male-specific intron retention depends on the presence of a weak splicing acceptor sequence inside the intron, as well as the regulation of Tra. Comparison of the structures of *dsx* transcripts, in light of their difference of conservation, suggests an evolutionary pathway from sexually monomorphic to sex-specific *dsx* via the sequential use of IR-based and APA-based alternative splicing.

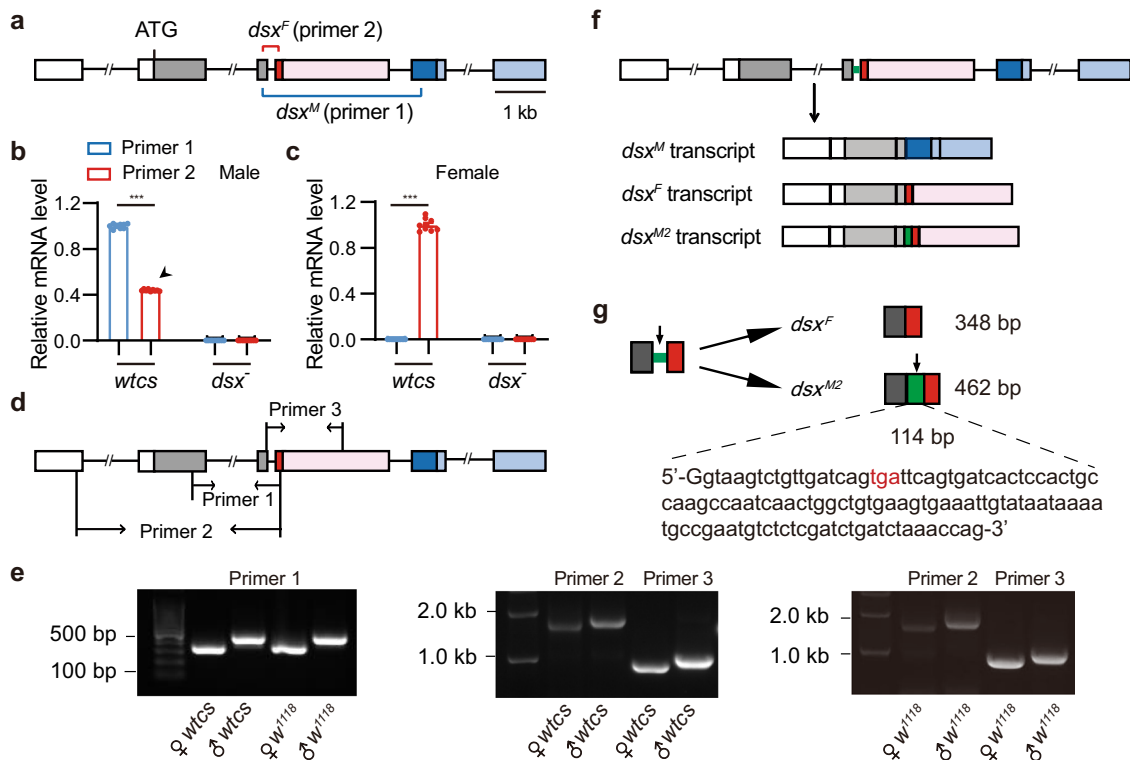
## Results

**Identification of a novel *dsx* transcript *dsx<sup>M2</sup>* through intron retention.** It has been determined that the *dsx* pre-mRNA is sex-specifically spliced to yield the male-specific *dsx<sup>M</sup>* and female-specific *dsx<sup>F</sup>* mRNAs. In an attempt to determine relative

expression of *dsx* transcripts in the two sexes in wild-type *Canton-S* (*w<sup>tcs</sup>*) and *dsx* mutant (*dsx<sup>683-7058</sup>/dsx<sup>1649-9625</sup>*) flies, we performed quantitative PCR (qPCR) experiments using two pairs of primers targeting the male-specific or female-specific locus in *dsx*, respectively (Fig. 1a). We found that *dsx<sup>M</sup>* was indeed expressed exclusively in *w<sup>tcs</sup>* males, but not in *w<sup>tcs</sup>* females or *dsx* mutant flies; however, comparable levels of *dsx* expression were detected in both *w<sup>tcs</sup>* males and females, but not in *dsx* mutant flies, by using the primer pair targeting *dsx<sup>F</sup>* (Fig. 1b, c). These results suggest the presence of *dsx<sup>F</sup>* or a novel *dsx* transcript, in addition to *dsx<sup>M</sup>*, in male flies.

To identify the *dsx* transcript detected in male flies, we next performed reverse transcription PCR (RT-PCR) experiments using three different primer pairs targeting different sites of the *dsx<sup>F</sup>* transcript (Fig. 1d), and found that all PCR products from male flies were larger than those from females (Fig. 1e). The same results were also obtained in another broadly used wild-type *w<sup>1118</sup>* flies (Fig. 1e). By sequencing these PCR products and sequence alignment, we found that the novel *dsx* transcript in males is quite similar to *dsx<sup>F</sup>*, with only 114 bp intron not being spliced out and retained between the last common exon and the female-specific exon (Fig. 1f, g). Thus, we identified a novel male-specific *dsx* transcript, hereafter referred to as *dsx<sup>M2</sup>*, in addition to the previously identified male-specific *dsx<sup>M</sup>* and female-specific *dsx<sup>F</sup>* transcripts (Supplementary Fig. 1a). Indeed, previous RNA-seq results already suggested the possibility of the 114 bp intron retention in males but not females<sup>27,28</sup> (Supplementary Fig. 1b and Supplementary Table 1). The predicted Dsx<sup>M2</sup> protein contains the common N-terminus of Dsx proteins and a specific C-terminus with only six amino acids resulting from the early stop codon inside the retained intron (Supplementary Fig. 2).

***dsx<sup>M2</sup>* functions like *dsx<sup>M</sup>* but has limited roles.** The above results identified a male-specific *dsx<sup>M2</sup>* transcript, but how it is generated from intron retention and whether it plays any role in sexual development or behavior is not known. To investigate whether and where *dsx<sup>M2</sup>* expresses, we tried to generate a polyclonal antibody against the eight amino acids in the C-terminus of the predicted Dsx<sup>M2</sup> protein but failed to detect any signal through immunostaining and western blot experiments. We next performed qPCR experiments to quantify the relative expression levels of *dsx<sup>M</sup>* and *dsx<sup>M2</sup>* in various tissues. We found that *dsx<sup>M2</sup>* was expressed in the fly head, thorax, and forelegs, but the level of *dsx<sup>M2</sup>* mRNA was generally lower than the level of *dsx<sup>M</sup>* mRNA (Supplementary Fig. 3), which is also consistent with the result using the whole fly body (Fig. 1b). To investigate whether *dsx<sup>M2</sup>* plays a role in sexual development and/or behavior, we generated a *UAS-dsx<sup>M2</sup>RNAi* construct targeting the retained intron (Fig. 2a), combined with previously used *UAS-dsx<sup>M</sup>RNAi* construct targeting the male-specific exon<sup>14</sup>, and validated their efficiency with three different *GAL4* drivers, the *actin-GAL4*, the pan-neuronal *R57C10-GAL4*<sup>29</sup>, and *dsx<sup>GAL4</sup>*, using qPCR (Fig. 2b, c). Both *dsx<sup>M</sup>* and *dsx<sup>M2</sup>* RNAi lines knocked down the corresponding *dsx* mRNA efficiently though not in the same level (Fig. 2b, c), and did not reduce the level of the other transcript (Supplementary Fig. 4). We next knocked down *dsx<sup>M</sup>* or *dsx<sup>M2</sup>* in all *dsx*-expressing cells [*dsx<sup>GAL4(Δ2)</sup>*] and found that males with *dsx<sup>M</sup>* knocked down were intersexual and displayed little courtship, while males with *dsx<sup>M2</sup>* knocked down had regular male appearance but reduced courtship and mating success with virgin female targets (Fig. 2d). These results suggest that *dsx<sup>M2</sup>* is involved in regulating male courtship but not, if any, development of sexually dimorphic traits. To further confirm whether *dsx<sup>M2</sup>* functions in the nervous system to mediate male courtship, we knocked down *dsx<sup>M</sup>* or *dsx<sup>M2</sup>* pan-neuronally using



**Fig. 1** Identification of a male-specific *dsx<sup>M2</sup>* transcript through intron retention. **a** Schematic structure of the *dsx* gene. Boxes: exons; white and gray boxes: common 5'-UTR and coding sequences respectively in both sexes; red and pink: coding and 3'-UTR sequences of the female-specific exon; dark and light blue: coding and 3'-UTR sequences of male-specific exons. Primers against sex-specific sequences are indicated. **b**, **c** Relative mRNA expression levels of presumed *dsx<sup>M</sup>* cycle at 60% humidity transcript (**b**, primer 1) and *dsx<sup>F</sup>* transcript (**c**, primer 2) in *wtcs* and *dsx* mutant (*dsx<sup>683-7058/dsx<sup>1649-9625</sup></sup>*) males and females.  $n = 9$  based on three replicates for each.  $***p < 0.001$ , Mann-Whitney U test. Note that a *dsx* transcript (arrowhead) is detected in males using the primer against the female-specific transcript (**b**). **d**, **e** Reverse transcription PCR (RT-PCR), using three pairs of primers as indicated (**d**), and sequencing identify the male-specific *dsx<sup>M2</sup>* transcript, in addition to the female-specific *dsx<sup>F</sup>* transcript, in both *wtcs* and *w<sup>1118</sup>* flies (**e**). For primer 1: 462bp and 348bp for male and female products, respectively; for primer 2: 1833bp and 1719bp for male and female products, respectively; for primer 3: 1069bp and 955bp for male and female products, respectively. **f** Illustration of the male-specific *dsx<sup>M2</sup>* transcript, in addition to previously identified *dsx<sup>M</sup>* and *dsx<sup>F</sup>* transcripts. **g** The *dsx<sup>M2</sup>* transcript differs from the *dsx<sup>F</sup>* transcript with the 114 bp intron (in green) retained. Red letters (tga) indicate the stop codon inside the retained intron.

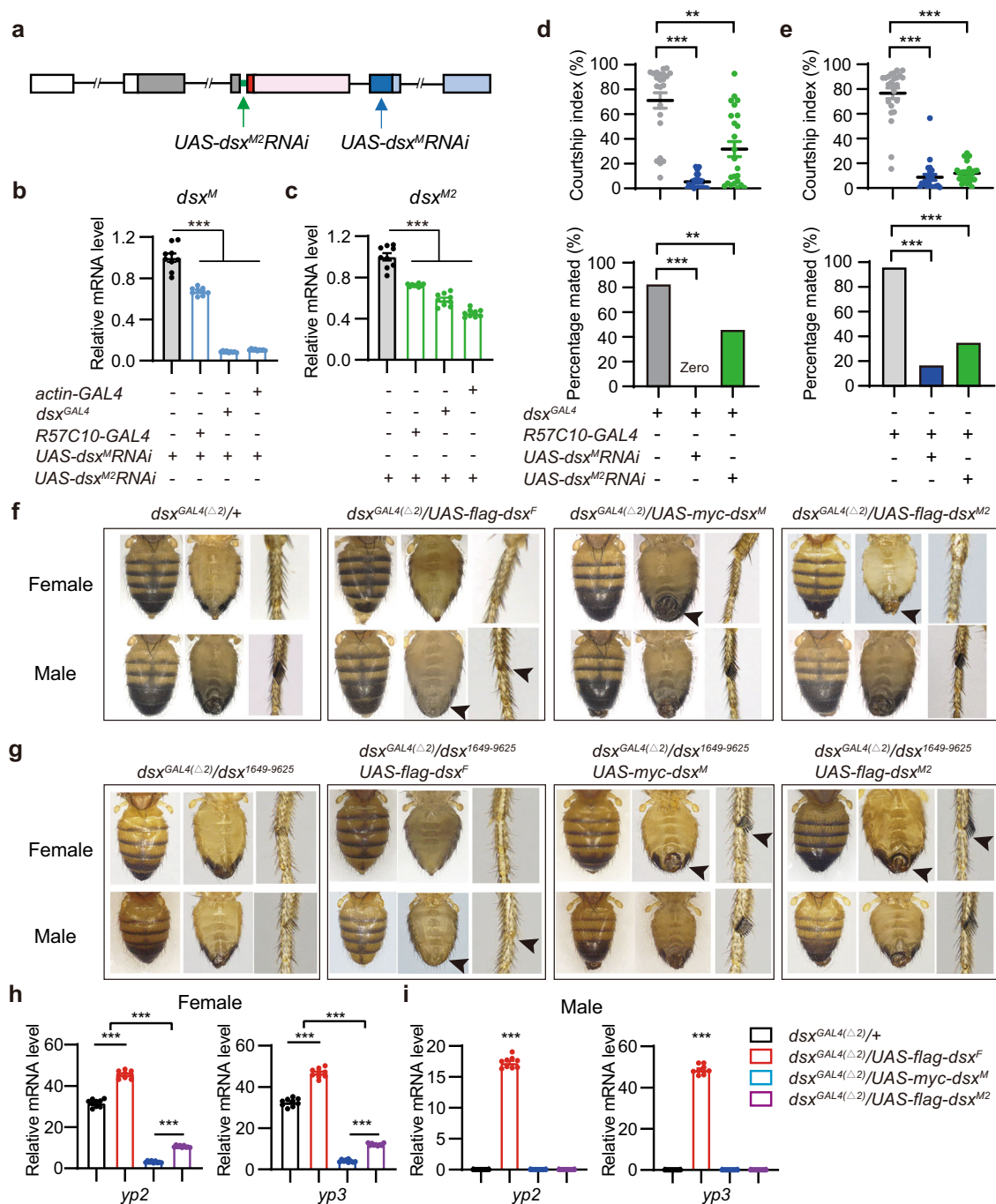
the *R57C10-GAL4* driver. Males with *dsx<sup>M</sup>* or *dsx<sup>M2</sup>* knocked down pan-neuronally had regular male appearance but much reduced courtship levels and mating success (Fig. 2e). These results indicate a crucial role of *dsx<sup>M2</sup>*, like *dsx<sup>M</sup>*, in regulating male courtship intensity.

To further investigate *dsx<sup>M2</sup>* function, we generated constructs overexpressing each *dsx* isoform (*UAS-flag-dsx<sup>F</sup>*, *UAS-myc-dsx<sup>M</sup>* and *UAS-flag-dsx<sup>M2</sup>*) and validated their efficiency using qPCR experiments and immunostaining with anti-Flag and anti-Myc antibodies (Supplementary Fig. 5). We then overexpressed these *dsx* isoforms in all *dsx*-expressing cells and found that overexpressing *Dsx<sup>F</sup>* feminized male development (genitals and sex combs), while overexpressing *Dsx<sup>M</sup>* or *Dsx<sup>M2</sup>* masculinized female development (genitals) (Fig. 2f). To avoid potential interference of overexpressed and indigenous *Dsx* isoforms (e.g., *Dsx<sup>M2</sup>* and *Dsx<sup>F</sup>*), we next overexpressed these *dsx* isoforms in a *dsx* mutant background [*dsx<sup>GAL4(Δ2)</sup>/dsx<sup>1649-9625</sup>*] in which both sexes were intersexual, and found that expressing *Dsx<sup>F</sup>* induced female differentiation in both sexes, while expressing *Dsx<sup>M</sup>* or *Dsx<sup>M2</sup>* strongly masculinized genital and sex comb development in both sexes (Fig. 2g). These results indicate that *Dsx<sup>M2</sup>* has a potential masculinizing role like *Dsx<sup>M</sup>*.

As transcription factors, *Dsx* proteins regulate sexual differentiation through their target genes, of which three genes encoding the female-specific Yolk Proteins (YPs) have been

intensively studied under control of *Dsx*<sup>10,24,30</sup>. To compare the transcriptional regulation of different *Dsx* isoforms on target genes, we overexpressed the three *Dsx* isoforms driven by the *dsx<sup>GAL4(Δ2)</sup>* and tested relative expression changes of *yp2* and *yp3*. Compared to control females, overexpressing *Dsx<sup>F</sup>* significantly increased *yp2* and *yp3* expression, while overexpressing *Dsx<sup>M</sup>* or *Dsx<sup>M2</sup>* significantly reduced their expression in females (Fig. 2h). Note that *yp2* and *yp3* expression were more severely reduced in females expressing *Dsx<sup>M</sup>* than those expressing *Dsx<sup>M2</sup>*, suggesting that *Dsx<sup>M2</sup>* had a weaker role of transcriptional inhibition compared to *Dsx<sup>M</sup>* (Fig. 2h). The expression of *yp2* or *yp3* was undetectable in control males, but significantly increased in males overexpressing *Dsx<sup>F</sup>*, which further confirmed the role of *Dsx<sup>F</sup>* in promoting *yp2* and *yp3* transcription (Fig. 2i). Taken together, the knockdown experiments indicate a crucial role of *dsx<sup>M2</sup>* in the nervous system for male courtship robustness but not sexual development, while the overexpression experiments indicate that *Dsx<sup>M2</sup>* has a potentially masculinizing role like *Dsx<sup>M</sup>*.

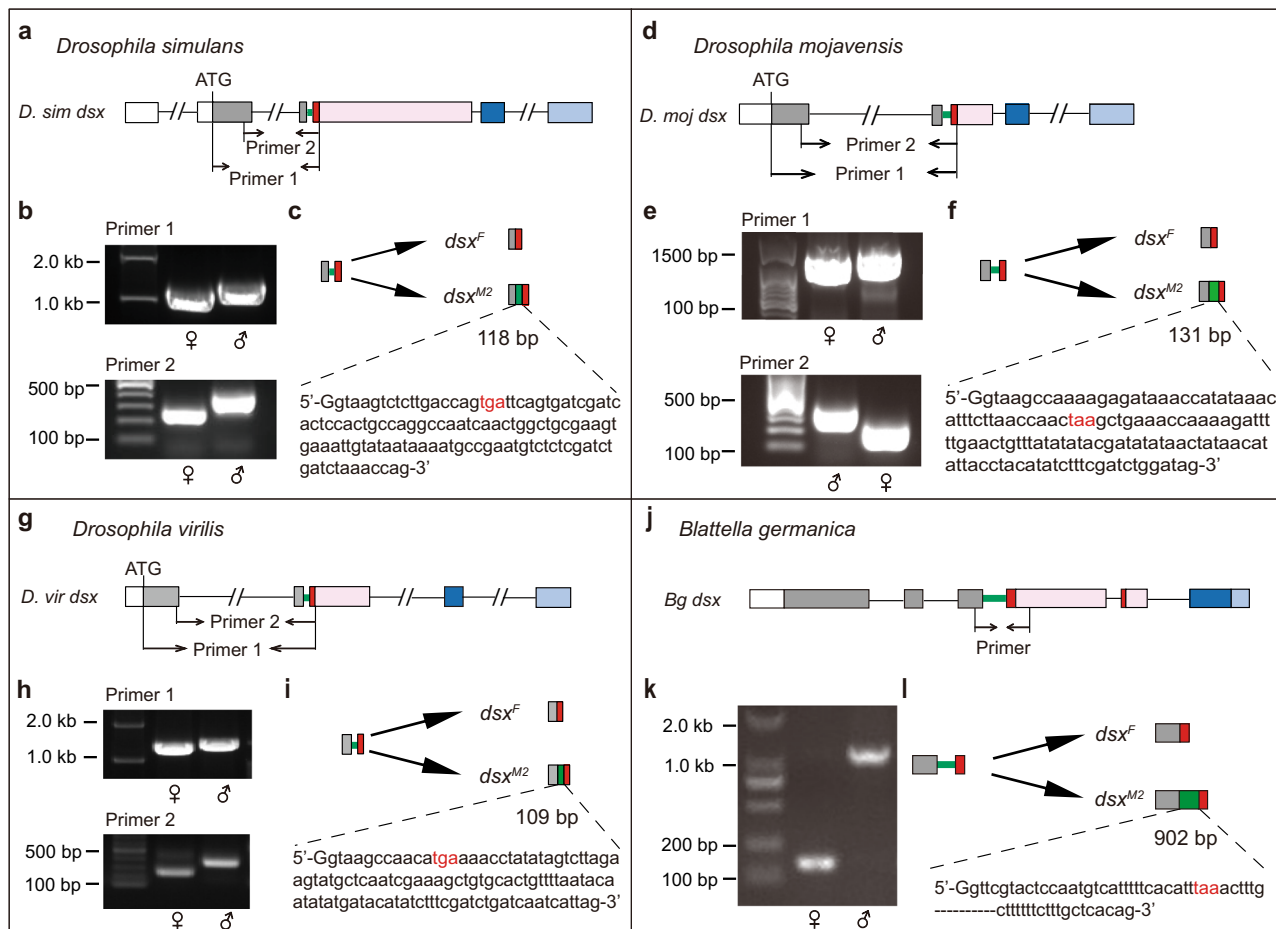
**Intron retention is a common mechanism to generate sex-specific *dsx* isoforms.** As we identified the male-specific *dsx<sup>M2</sup>* transcript through intron retention, which is a potent mechanism underlying sexual differentiation, we asked if such an alternative splicing form also existed in other animal species. We tested three



**Fig. 2** *dsx<sup>M2</sup>* has a potentially masculinizing function like *dsx<sup>M</sup>*. **a** RNAi targeting *dsx<sup>M</sup>* and *dsx<sup>M2</sup>* as indicated by blue and green arrows, respectively. **b, c** Relative mRNA expression levels of *dsx<sup>M</sup>* (**b**) and *dsx<sup>M2</sup>* (**c**) in control and RNAi-mediated males.  $n = 9$  based on three replicates for each.  $***p < 0.001$ , Mann-Whitney U test. **d, e** Knocking down *dsx<sup>M</sup>* or *dsx<sup>M2</sup>* in *dsx*-expressing cells (**d**) or pan-neuronally (**e**) impairs male courtship and mating success with females.  $n = 23, 24, 24, 24, 24$  for *dsx<sup>GAL4</sup>/+*, *dsx<sup>GAL4</sup>/UAS-dsx<sup>M</sup>RNAi*, *dsx<sup>GAL4</sup>/UAS-dsx<sup>M2</sup>RNAi*, *R57C10-GAL4/+*, *R57C10-GAL4/UAS-dsx<sup>M</sup>RNAi* and *R57C10-GAL4/UAS-dsx<sup>M2</sup>RNAi* respectively. For courtship index,  $**p = 0.0012$  and  $***p < 0.001$ , Kruskal-Wallis test with Dunn's multiple comparisons test; for percentage mated,  $**p = 0.0087$  and  $***p < 0.001$ , Chi-square test. **f, g** Overexpression of Dsx<sup>F</sup> feminized, while overexpression of Dsx<sup>M</sup> or Dsx<sup>M2</sup> masculinized *dsx*-expressing cells, including external genitalia and sex comb (arrowhead), under wild-type background [**f**, *dsx<sup>GAL4(Δ2)/+</sup>*] or mutant background [**g**, *dsx<sup>GAL4(Δ2)/dsx<sup>1649-9625</sup></sup>*] for *dsx*. **h, i** Relative mRNA expression levels of *dsx* target genes, *yp2* and *yp3*, in females (**h**) and males (**i**). Expression levels of *yp2* and *yp3* were increased with Dsx<sup>F</sup> overexpression and decreased with Dsx<sup>M</sup> or Dsx<sup>M2</sup> overexpression.  $n = 9$  based on three replicates for each.  $***p < 0.001$ , Mann-Whitney U test. Error bars indicate SEM.

other *Drosophila* species including the closely related *D. simulans* and two other distant species *D. mojavensis* and *D. virilis*, by performing RT-PCR experiments using multiple primer pairs targeting sequences including the potentially retained intron and

a portion of the female-specific exon, followed by sequencing. We found retention of the 118 bp intron in *D. simulans* males (Fig. 3a–c), 131 bp intron in *D. mojavensis* males (Fig. 3d–f) and 109 bp intron in *D. virilis* males (Fig. 3g–i), which generated



**Fig. 3 Intron retention-induced sex-specific splicing is widely conserved.** Retention of the intron that separates the last common exon and the female-specific exon generates male-specific  $dsx^{M2}$  transcripts in *D. simulans* (a–c), *D. mojavensis* (d–f), *D. virilis* (g–i) and *B. germanica* (j–l). The *dsx* gene structures are illustrated in a, d, g and j. White and gray boxes indicate common 5'-UTR and coding sequences respectively in both sexes. Red and pink boxes indicate coding and 3'-UTR sequences of the female-specific exon. Dark and light blue boxes indicate coding and 3'-UTR sequences of male-specific exons. The thick green line indicates the intron spliced out in the  $dsx^F$  transcript and retained in the  $dsx^{M2}$  transcript. Primers for RT-PCR experiments (b, e, h and k) are indicated respectively (a, d, g and j). Full sequences of the retained intron of  $dsx^{M2}$  are indicated (c, f and i) except for the 902bp one in *B. germanica*  $dsx^{M2}$  (l). The capital letter "G" from the upstream exon is added in front of the intron sequence to ensure that the stop codon highlighted in red is aligned within the open reading frame.

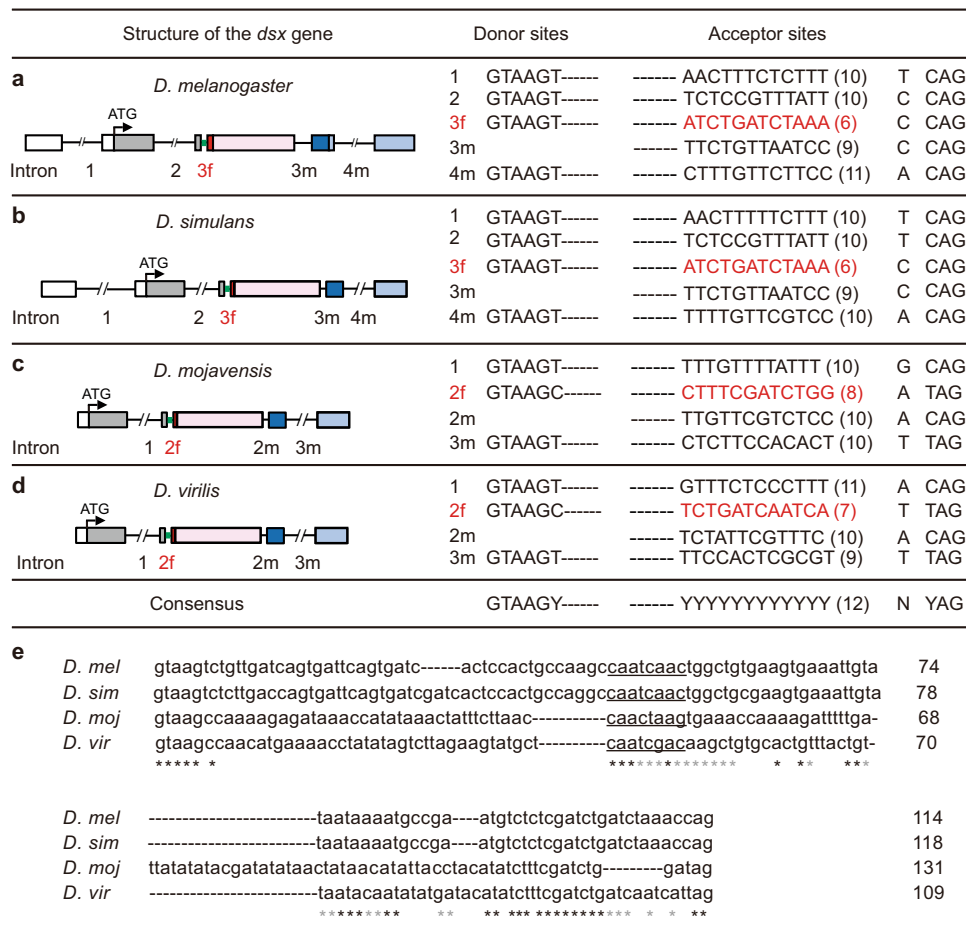
$dsx^{M2}$  transcripts in the same way as in *D. melanogaster* males. All predicted Dsx<sup>M2</sup> proteins have conserved DM domain including a zinc-finger DNA binding domain and oligomerization domains, and few amino acids in the C-terminus due to stop codons in the retained intron (Supplementary Fig. 6). Protein sequence alignment of Dsx isoforms in these *Drosophila* species revealed that sex-specific C-terminus of Dsx<sup>F</sup> proteins were almost identical, while those of Dsx<sup>M</sup> and Dsx<sup>M2</sup> proteins were less conserved (Supplementary Fig. 7), which is consistent with previous findings on the different levels of conservation of Dsx<sup>M</sup> and Dsx<sup>F</sup> isoforms across insect species<sup>31</sup>.

To further test whether intron retention could be a conserved way to generate alternative splicing of *dsx*, we performed RT-PCR experiments targeting the *dsx* gene in a hemimetabolous insect, the German cockroach *Blattella germanica*, as it is not only a far distant insect species apart from *D. melanogaster* (>300 million years), but also with the *dsx* gene structure being previously studied<sup>17,32</sup>. We designed a pair of primer targeting the last common exon and the first female-specific exon respectively (Fig. 3j) and obtained sex-specific fragments by PCR (Fig. 3k). Through sequencing, we revealed the existence of the male-specific  $dsx^{M2}$  transcript that retains the 902 bp intron linking the last common exon and the first female-specific exon (Fig. 3l), in

addition to previously identified *dsx* transcripts<sup>17</sup>. Taken together, these results indicate that intron retention is a common mechanism to generate sex-specific *dsx* isoforms in distant insect species.

### ***dsx* intron retention is jointly regulated by a weak splice acceptor and transformer.**

To further investigate factors affecting intron retention and generating sex-specific *dsx* isoforms, we compared sequences of all introns of *dsx* in the four *Drosophila* species. We did not compare the acceptor sequences in *Blattella germanica*, as we have insufficient knowledge about the consensus acceptor sequence in this species. We found that all introns of *dsx* in *D. melanogaster* and *D. simulans* start with the donor sequence GTAAGT and end with acceptor sequences (T/C)<sub>n</sub>NCAG (N indicates A, C, G or T) (Fig. 4a, b). Previous studies already suggest such consensus acceptor sequence<sup>10</sup>, where the number of pyrimidines (T/C) upstream of the last four nucleotides (NCAG) is at least 9. Indeed, the common (intron 1 and 2) and male-specific (intron 3m and 4m) acceptors have 9–11 pyrimidines upstream of NCAG. As for the retained intron (spliced out in female, 3f), the acceptor has only 6 pyrimidines upstream of NCAG (Fig. 4a, b). We next compared donor and acceptor



**Fig. 4 Retained introns share conserved sequence properties.** **a-d** Comparison of donor and acceptor sequences of spliced and retained introns. The structure of the *dsx* gene including introns (black lines), the retained intron for *dsx<sup>M2</sup>* (thick green line), and exons (boxes), as well as splicing donor and acceptor sequences for each intron in *D. melanogaster* (**a**), *D. simulans* (**b**), *D. mojavensis* (**c**) and *D. virilis* (**d**). Introns were labeled with numbers followed by “f” (female-specific splicing) or “m” (male-specific splicing), respectively. As for the consensus sequences, Y indicates pyrimidine, and N indicates either A, G, T or C. The number of pyrimidines in the 12 upstream nucleotides of the acceptor sequence NYAG is indicated in the brackets for each intron. **e** Comparison of retained intron sequences of *dsx<sup>M2</sup>* in four *Drosophila* species. Underlined letters indicate potential Tra binding sites. Black asterisks indicate perfect matches of amino acids among four species, and gray asterisks indicate three matches out of four species.

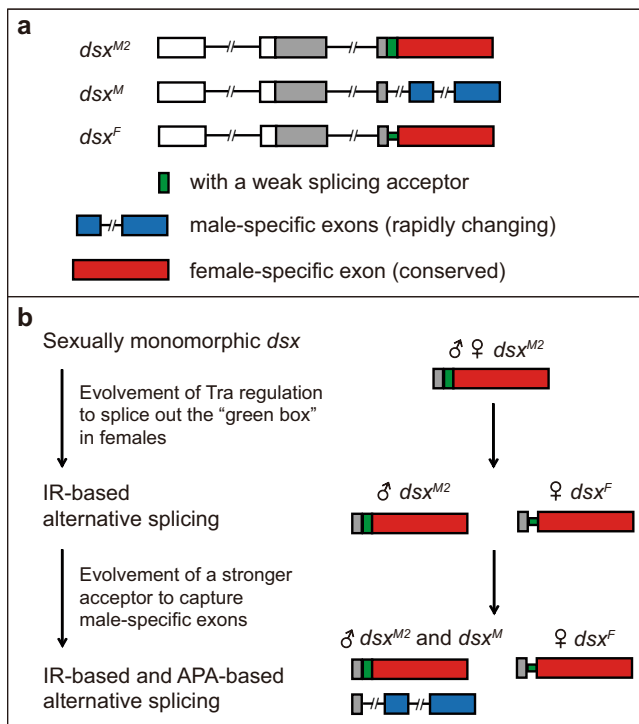
sequences of introns of the *dsx* gene in *D. mojavensis* and *D. virilis*. We found similar results as in *D. melanogaster* and *D. simulans*: the common and male-specific acceptors have 9–11 pyrimidines, while the female-specific acceptors have 7 or 8 pyrimidines upstream of NCAG (Fig. 4c, d). These results indicate that the presence of a weak splice acceptor (6–8 pyrimidines) may lead to intron retention of the *dsx* gene in males of four *Drosophila* species.

We next compared full sequences of retained introns in the four *Drosophila* species. Surprisingly, these intron sequences, despite many differences, share intensive identity in two regions: one in the 3' end of the intron overlapping the region of the weak acceptor site that may contribute to intron retention as above mentioned, and the other in the middle of these introns containing the core sequence (CAATCAAC) of the *tra* binding sequence, TC(T/A)(T/A)CAATCAACA, that occurs six times in the female-specific exon (Fig. 4e). The conservation of the above core sequence, which is a potential *tra* binding site within the retained introns, suggests its potential role in *tra*-mediated alternative splicing. Indeed, we observed intron retention in females with *tra* knocked down in all cells (*actin-GAL4/+; UAS-traRNAi/+*) (Supplementary Fig. 8). Together these results suggest that the weak splice acceptor and *tra* regulation jointly

determine the splicing modes between *dsx<sup>M2</sup>* (intron retained) and *dsx<sup>F</sup>* (intron spliced out).

**A proposed evolutionary pathway of *dsx* alternative splicing.**

The identification of the *dsx<sup>M2</sup>* transcript promoted us to analyze how the two forms of alternative splicing (IR and APA) were evolved to generate the modern day *dsx* transcripts (Fig. 5a). Substantial studies already elucidated an evolutionary pathway from sexually monomorphic *dsx* to sex-specific *dsx* transcripts upon the evolution of *tra-dsx* regulation in females<sup>16,17</sup>. A simple assumption is that sex-specific *dsx* transcripts come from a sexually monomorphic transcript *dsx<sup>M2</sup>*, or *dsx<sup>M</sup>*, or both. Due to the nature of close relationship between *dsx<sup>M2</sup>* and *dsx<sup>F</sup>* (whether an intron with a weak splicing acceptor is spliced out or not), and the more conservation of the female-specific exon than the male-specific exons in insect species<sup>31,33</sup>, we propose that *dsx<sup>M2</sup>* could serve as an ancient monomorphic transcript in both sexes, followed by two sequential events evolving alternative splicing: the IR-based mechanism generating *dsx<sup>M2</sup>* in males and *dsx<sup>F</sup>* in females with the evolution of *Tra* regulation in females, and later the APA-based mechanism generating the rapid-changing *dsx<sup>M</sup>* with the evolution of a stronger splicing acceptor to capture male-specific exons (Fig. 5b).



**Fig. 5 An evolutionary pathway of *dsx* alternative splicing underlying sexual development.** **a** Structures of the modern day *dsx* transcripts, *dsx*<sup>M2</sup>, *dsx*<sup>M</sup>, and *dsx*<sup>F</sup> in *D. melanogaster*. **b** A proposed evolutionary pathway from sexually monomorphic *dsx*<sup>M2</sup> to sex-specific *dsx* transcripts sequentially via the IR-based and APA-based mechanisms.

## Discussion

Previous studies have shown that alternative splicing of *dsx* generates sex-specific isoforms (Dsx<sup>M</sup> and Dsx<sup>F</sup>) underlying sexual differentiation in *D. melanogaster* and many other insects<sup>1,2,10,16</sup>. Such splicing mechanism involves the selective use of sex-specific 3' exons, which is under regulation of Tra and Tra-2<sup>10,20–22</sup>. Our results identified the male-specific Dsx<sup>M2</sup> isoform through intron retention, which is a common mechanism generating sex-specific Dsx isoforms in insects across 300 million years.

Recent studies on the alternative splicing of *dsx* using a variety of insect species revealed that the *tra*-mediated female-specific splicing of *dsx* could be evolved from sexually monomorphic isoforms with male-biased expression<sup>16,17</sup>. Indeed, vertebrates, nematodes and crustaceans use male-biased transcription of *Dmrt* genes to direct male-specific function<sup>1,2,34</sup>. In this regard, the evolution of *dsx*<sup>F</sup> may come from *dsx*<sup>M</sup> and/or *dsx*<sup>M2</sup> under Tra regulation, possibly through the functional gain of Tra binding sites in the female-specific exon of *dsx*. Indeed, knocking down Tra function in females induced intron retention-based splicing of *dsx*<sup>M2</sup> instead of *dsx*<sup>F</sup>. The coding sequence of the female-specific exon, which is also included as 3'-UTR of *dsx*<sup>M2</sup>, is more conserved than the male-specific exon of *dsx*<sup>M</sup> that rapidly changes not only across insect orders but also across closely related genera within each insect order<sup>31,33</sup>, excluding the possibility of *dsx*<sup>M</sup> as the origin of sexually monomorphic *dsx* transcript. The identification of the intron retention-based *dsx*<sup>M2</sup> transcript fits well into a position as the ancestral isoform of *dsx* expressed in both sexes before the evolution of *dsx*<sup>M</sup> and *dsx*<sup>F</sup>. The conserved presence of a weak splicing acceptor upstream of the female-specific exon provides the selection basis for later recruitment of Tra regulation to generate the female-specific *dsx*<sup>F</sup> from *dsx*<sup>M2</sup> in females, as well as the evolution of a stronger

splicing acceptor to capture male-specific exons and generate *dsx*<sup>M</sup> in males. Such an evolutionary pathway of generating sex-specific Dsx isoforms could fill the gap between the splicing-based mechanism in insects and transcription-based mechanism in vertebrates and nematodes.

The Dsx<sup>M2</sup> isoform contains all common Dsx amino acids and six specific C-terminal residues, thus could act like other Dmrt transcription factors. Our knockdown experiments reveal that Dsx<sup>M2</sup> still functions in the nervous system to promote courtship robustness in *D. melanogaster*, but not in the development of male-specific traits such as genitals and sex combs. In addition, the overexpression experiments indicate that Dsx<sup>M2</sup> has a potential masculinizing role just like Dsx<sup>M</sup>. However, a direct comparison of Dsx<sup>M</sup> and Dsx<sup>M2</sup> function could not be faithfully achieved as knockdown and overexpression levels of *dsx*<sup>M</sup> and *dsx*<sup>M2</sup> may be different. Another caveat of these results is the lack of direct evidence of Dsx<sup>M2</sup> expression in males. It is also possible that Dsx<sup>M2</sup> only serves as an ancient Dsx isoform whose function has largely been replaced during evolution (e.g., by Dsx<sup>M</sup>), and now has limited expression and function. Future studies could generate better reagents to testify Dsx<sup>M2</sup> expression in a temporal and spatial manner in *D. melanogaster* and other insect species to better understand its function as a possible origin of sex-specific Dsx isoforms underlying sexual development and sexual dimorphism.

## Methods

**Fly stocks.** Flies were raised at 22 °C or 25 °C at 12 hr light/12 hr dark cycle at 60% humidity. *Canton-S* (*w<sup>ts</sup>*) and *w<sup>1118</sup>* were used as wild-type strains. *dsx* mutant lines used in Fig. 1b and c include *dsx*<sup>683-7058</sup> and *dsx*<sup>1649-9625</sup>, which were used as previously<sup>13</sup>. *Drosophila mojavensis*, *Drosophila simulans* and *Drosophila virilis*<sup>13</sup>, *R57C10-GAL4* (attP2, BDSC\_39171), *actin-GAL4* (BDSC\_25374), *UAS-dsx<sup>M</sup>RNAi* (attP2)<sup>14</sup>, *UAS-flag-dsx<sup>F</sup>* (attP40) and *UAS-myc-dsx<sup>M</sup>* (attP40)<sup>35</sup>, *dsx<sup>GAL412</sup>* and *dsx<sup>GAL4(Δ2)36</sup>* were used as described previously. *UAS-dsx<sup>M2</sup>RNAi* (attP2) and *UAS-flag-dsx<sup>M2</sup>* (attP40) were generated in this study and described below in details. Detailed information about fly stocks and other materials used in this study is listed as Table 1.

***dsx* sequences and multiple sequence alignment.** We downloaded *dsx* gene sequence of *D. melanogaster* from FlyBase (<http://flybase.org/>). We used NCBI (<https://www.ncbi.nlm.nih.gov/>) to find the location and gene sequences of *dsx* in following species: *D. simulans* (Gene ID: 6727147), *D. mojavensis* (Gene ID: 6574377), *D. virilis* (Gene ID: 6633147) and *B. germanica*<sup>17,32</sup>. Gene sequence or amino acid sequences comparisons were performed using Clustal Omega and LaserGene software.

We also used public RNA-seq data<sup>27</sup> (GSM694258 and GSM694259 for *D. melanogaster* females, GSM694260 and GSM694261 for *D. melanogaster* males, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28078>) to analyze splicing events of *dsx* (Supplementary Fig. 1b) using the Integrative Genomics Viewer (IGV)<sup>37</sup>.

**Generation of *UAS-dsx<sup>M2</sup>RNAi* line.** To generate the *UAS-dsx<sup>M2</sup>RNAi* construct, the *UAS-dsx<sup>F</sup>RNAi* plasmid (based on the pVALIUM20)<sup>14</sup> was digested to be linearized with *NheI* and *EcoRI*. We selected a 21 nucleotides sequence (tggctgtgaagtgaattgta) targeting the 114 bp intron, which was synthesized and generated into hairpin by annealing. The resulting DNA fragment and linearized vector were ligated. The resulting constructs were injected into embryos of attP2 site with PhiC31-mediated transgenesis. The correct insertion was screened by *vermillion* (*vermillion* positive eye color) and verified by PCR and followed by DNA sequencing. Oligo primers are as following:

Oligo forward:  
5'-ctagcagTGGCTGTGAAGTGAATGTGAtagttatattcaagcataTACAATTTC  
ACTTCACAGCCAgcg-3'  
Oligo reverse:  
5'-aattcgcTCGCACCTGTAGCCAGATCTAtatgcttgaataactaTACAATTTC  
CTTCACAGCCAactg-3'.

**Generation of *UAS-flag-dsx<sup>F</sup>*, *UAS-myc-dsx<sup>M</sup>* and *UAS-flag-dsx<sup>M2</sup>* lines.** *UAS-flag-dsx<sup>F</sup>* and *UAS-myc-dsx<sup>M</sup>* lines were generated previously in our lab<sup>35</sup>. The *UAS-flag-dsx<sup>M2</sup>* line was generated in this study. In brief, the 3×Flag tags (MDYKDHGD-DYKDHDI-DYKDDDDKL) were added to the N terminus of Dsx<sup>F</sup> and Dsx<sup>M2</sup> protein separately. We simultaneously amplified the *flag* sequence and the full-length CDS fragment of *dsx<sup>F</sup>* and *dsx<sup>M2</sup>* from wild-type cDNAs using

**Table 1 Detailed information for key resources and fly stocks.**

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	Mouse monoclonal anti-Flag	Sigma-Aldrich	Cat# F1804	IHC (1:500)
Antibody	Mouse monoclonal anti-Myc	MBL	MO47-3	IHC (1:200)
Antibody	Donkey polyclonal anti-Mouse, Alexa Fluor 488	Thermo Fisher Scientific	Cat# A-21202, RRID: AB_141607	IHC (1:500)
Antibody	Donkey polyclonal anti-Mouse, Alexa Fluor 555	Thermo Fisher Scientific	Cat# A-31570, RRID: AB_2536180	IHC (1:500)
Plasmid	pVALIUM20	Tsinghua University	#26214	
Plasmid	pJFRC2-10×UAS-IVS-mCD8::GFP	Addgene	Code# 005-000-121	3% NGS in 1×PBS
Chemical compound drug	Normal Goat Serum (NGS)	Jackson ImmunoResearch Laboratories	RRID: AB_2336990	
Chemical compound drug	Paraformaldehyde (PFA)	Sigma-Aldrich	CAS# 30525-89-4	4% PFA in 1×PBS
Chemical compound drug	TRIZOL™ reagent	Invitrogen	Cat# 15596026	
Chemical compound drug	SuperScript™ IV DNA Polymerase High Fidelity	Invitrogen	Cat# 18091050	
Chemical compound drug	EvaGreen Dye	Transgen	Cat# AS131-21	
Genetic reagent (D. melanogaster)	UAS-dsx <sup>M</sup> /RNAi	Biotium <sup>14</sup>	Cat# 31000	
Genetic reagent (D. melanogaster)	dsx <sup>GAL4</sup>	<sup>12</sup>		
Genetic reagent (D. melanogaster)	dsx <sup>GAL4(Δ2)</sup>	<sup>36</sup>		
Genetic reagent (D. melanogaster)	R57C10-GAL4	Bloomington <i>Drosophila</i> Stock Center	BDSC_39171	
Genetic reagent (D. melanogaster)	actin-GAL4	Bloomington <i>Drosophila</i> Stock Center	BDSC_25374	
Genetic reagent (D. melanogaster)	UAS-traRNAi	Bloomington <i>Drosophila</i> Stock Center	BDSC_28512	
Genetic reagent (D. simulans)	<i>Drosophila simulans</i>	<sup>13</sup>		
Genetic reagent (D. simulans)	<i>Drosophila mojavensis</i>	<sup>13</sup>		
Genetic reagent (D. virilis)	<i>Drosophila virilis</i>	<sup>13</sup>		
Genetic reagent (D. melanogaster)	UAS-dsx <sup>M</sup> RNAi	This study		Described below
Genetic reagent (D. melanogaster)	UAS-flag-dsx <sup>M2</sup>	This study		Described below
Genetic reagent (D. melanogaster)	UAS-flag-dsx <sup>F</sup>	<sup>35</sup>		Described below
Genetic reagent (D. melanogaster)	UAS-myc-dsx <sup>M</sup>	<sup>35</sup>		Described below
Software,algorithm	LaserGene	DNASar	<a href="http://www.dnastar.com/">http://www.dnastar.com/</a>	
Software,algorithm	Clustal Omega	EMBL-EBI	<a href="https://www.ebi.ac.uk/Tools/msa/clustalo">https://www.ebi.ac.uk/Tools/msa/clustalo</a>	
Software,algorithm	ImageJ	ImageJ National Institutes of Health	<a href="https://imagej.nih.gov/ij/">https://imagej.nih.gov/ij/</a>	
Software,algorithm	Prism 8	GraphPad	<a href="https://www.graphpad.com/">https://www.graphpad.com/</a>	
Software,algorithm	Integrative Genomics Viewer	<sup>37</sup>	<a href="https://www.igv.org/">https://www.igv.org/</a>	



**Table 2 Primers used for RT-PCR experiments.**

Usage	Primer names	Sequences (5'-3')
Fig. 1d, e	Primer 1	Forward: CAATCGCTGGAGGGTCTCTG Reverse: TCATCCACATTGCCGCGTTG
Fig. 1d, e	Primer 2	Forward: GCAAAGCACACCTCGGGAG Reverse: TCATCCACATTGCCGCGTTG
Fig. 1d, e	Primer 3	Forward: TTCCGCTATCCTTGGGAGCT Reverse: TTGGCTTGATGCCTATTCC
Fig. 3a, b	Primer 1	Forward: GAATCATGGTTTCGGAGGAGAACTG Reverse: CTACGTGGCAGCCGTGGAGCTCACC
Fig. 3a, b	Primer 2	Forward: GAATCATGGTTTCGGAGGAGAACTG Reverse: TCATCCACATTGCCGCGTTGTGTTTG
Fig. 3d, e	Primer 1	Forward: ATGGTTTCGGAGGAGAACTGG Reverse: TCATCCACATTGCCGCGTTGT
Fig. 3d, e	Primer 2	Forward: ACGCAAGAATGTGCCACTCG Reverse: TCATCCACATTGCCGCGTTGT
Fig. 3g, h	Primer 1	Forward: ATGGTTTCAGAGGAGAATTGGAACA Reverse: TCATCCACATTGCCGCGTTGTGTTG
Fig. 3g, h	Primer 2	Forward: AGCACGCAAGAATGTGCCACTGG Reverse: TCATCCACATTGCCGCGTTGTGTTG
Fig. 3j, k	Primer	Forward: AGAACGGCAGCGAGACAGGC Reverse: TGTGGAGACGGCGATGAGG
Supplementary Fig. 8b, c	Primer 1	Forward: ATGGTTTCGGAGGAGAACTG Reverse: TCATCCACATTGCCGCGTTG
Supplementary Fig. 8b, c	Primer 2	Forward: CAATCGCTGGAGGGTCTCTG Reverse: TCATCCACATTGCCGCGTTG

the following overhang primers and then cloned into pJFRC2-10XUAS-IVS-mCD8::GFP plasmid (Addgene #26214) to replace the mCD8::GFP sequence.

*dsx<sup>F</sup>*-forward: CCGAGATCTATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGCTTGTTCGGAGGAGAACTGG

*dsx<sup>F</sup>*-reverse: CCATCTAGATCATCCACATTGCCGCGTTG.

*dsx<sup>M2</sup>*-forward: CCGAGATCTATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACGATGACAAGCTTGTTCGGAGGAGAACTGG

*dsx<sup>M2</sup>*-reverse: CCATCTAGATCACTGATCAACAGACTTACC.

For the *UAS-myc-dsx<sup>M</sup>* transgenic line, the 6×Myc tag (MEQKLISEEDL) was added to the N terminus of *Dsx<sup>M</sup>*. We first cloned the 6×myc sequence into pJFRC2-10XUAS-IVS-mCD8::GFP between the *Bgl*III and the *Xho*I sites, and then amplified the full-length CDS of *dsx<sup>M</sup>* and cloned into C terminus of 6×Myc to replace the mCD8::GFP sequences.

*myc-F*: CCGAGATCTTCCCATCGACTTAAAGCTATG

*myc-R*: CCACTCGAGACTAGTCTCAAGAGGCCTTGAGTTC

*dsx<sup>M</sup>-F*: CCGACTAGTGTTCGGAGGAGAACTGG

*dsx<sup>M</sup>-R*: CCATCTAGACTACGTGGCAGCCGTGGAG

The constructed plasmids were injected and integrated into the attP40 sites on the second chromosome through phiC31 integrase mediated transgenesis. The correct strains were screened by *mini-white* (orange eyes) and verified by PCR and followed by DNA sequencing.

**RT-PCR analysis.** To acquire cDNA template, *Canton-S*, *w<sup>1118</sup>* and other species including *D. simulans*, *D. mojavensis*, *D. virilis* and *Blattella germanica* were used as wild-type strains for RT-PCR. We obtained total RNA from approximate thirty adult females and/or males using a commercial TRizol™ reagent (15596026, Invitrogen, USA) and purified RNA with DNA-free™ kit (AM1906, Ambion) according to manufacturer's protocol. Purified RNA was finally resuspended in 60μL of DEPC-treated water. First strand cDNA was synthesized for each RNA sample using SuperScript™ IV reverse transcriptase (18091050, Invitrogen). *dsx*-specific primers are listed in Table 2 to detect different *dsx* transcripts.

**qPCR.** Total RNA extraction and first strand cDNA template were acquired as described above, except for the data in Supplementary Fig. 3, which used approximately 100 flies to obtain different tissues (head, thorax, abdomen, forelegs, and wings). qPCR was performed using a LightCycler® 96 SW 1.1 system (Roche). We used EvaGreen Dye (31000, Biotium, USA) and High Fidelity PCR SuperMix (AS131-21, TransGen, Beijing) to conduct qPCR. *actin* was amplified as an internal control for normalization. The primers for qPCR used in this study are listed in Table 3.

**Male courtship assay.** Newly enclosed males were collected and group housed for 4–7 days at 25°C and 60% humidity with a 12 hr: 12 hr light/dark cycle. Virgin *Canton-S* females were group-housed and aged under similar conditions. To

**Table 3 Primers used for qPCR experiments.**

	Sequences (5'-3')
<i>actin</i>	Forward: CAGGCGGTGCTTCTCTCTA Reverse: AGTGTAACCGCGCTCAGTA
<i>dsx<sup>M</sup></i>	Forward: GAAGAGGCTTCCC GGCGAAT Reverse: GGACAAATCTGTGTGGAGCGG
<i>dsx<sup>F</sup></i>	Forward: TTCCGCTATCCTTGGGAGCT Reverse: CATCCACATTGCCGCGTTGT
<i>dsx<sup>M2</sup></i>	Forward: GCCGATCTCAGTTTCCGTC Reverse: TCACTGATCAACAGACTTAC
<i>yp2</i>	Forward: TGGGTCAATCCAGTGAAGT Reverse: ACAATGTAGCCCCTGATCTG
<i>yp3</i>	Forward: GAAGCCGACCAAGTGGCTGA Reverse: TCCAGACGGGCACATTGCTC

measure courtship, a male of each genotype and a *Canton-S* virgin female were loaded individually into two-layer chambers (diameter: 1 cm; height: 3 mm per layer) which were separated by a plastic transparent barrier until courtship test for 30 min. Courtship index (CI), which is the percentage of observation time a fly displayed any courtship step in 10 min, and percentages mated were counted based on successful copulation in 30 min.

**Tissue dissection, staining, and imaging.** Adult flies were reared at 25°C and aged for 4–7 days old, and were dissected in Schneider's insect medium (Thermo Fisher Scientific, Waltham, MA) and fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) for 20–30 min at room temperature (RT). Tissues were washed at least 4 times for 15–20 min with PAT3 (0.5% Triton X-100, 0.5% bovine serum albumin in PBS), and then blocked in 3% normal goat serum (NGS) for 60 min at RT. In Supplementary Fig. 5e and f, samples were incubated with anti-Flag mouse (Sigma, F1804, 1:500) antibody or anti-Myc mouse (MBL, M047-3, 1:200) antibody diluted in 3% NGS for overnight at 4°C, then washed four times in PAT3, and incubated in secondary antibodies anti-mouse IgG conjugated to Alexa 488 (Invitrogen, A21202, 1:500) diluted in 3% NGS for 1–2 days at 4°C. Tissues were then washed thoroughly in PAT3 and mounted for confocal imaging.

For visualizing morphological appearances of flies (Fig. 2f, g and Supplementary Fig. 8a), 4–7 days old flies were frozen at -80°C for 30 min. Fly forelegs and abdomens were dissected and imaged by a Nikon Shuttle pix P400RV stereoscopic microscope.

**Statistics and reproducibility.** All statistical analyses were performed using the Prism 8 (GraphPad software). Experimental flies and genetic controls were tested at the same condition, and data are collected from at least two independent experiments and are reproducible. The Mann-Whitney U test was used for pairwise comparisons. For comparing mating success, Chi-square tests were performed to compare two different groups at 30 min time point.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### Data availability

All data generated or analyzed during this study are included in the manuscript and its supplementary information files. Source data underlying figures are presented in Supplementary Data 1. Uncropped gel images are provided in Supplementary Fig. 9. All other relevant data supporting the findings of this study are available from the corresponding author upon reasonable request.

Received: 3 January 2022; Accepted: 29 June 2022;

Published online: 22 July 2022

### References

- Kopp, A. Dmrt genes in the development and evolution of sexual dimorphism. *Trends Genet* **28**, 175–184 (2012).
- Matson, C. K. & Zarkower, D. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nat. Rev. Genet* **13**, 163–174 (2012).
- Williams, T. M. & Carroll, S. B. Genetic and molecular insights into the development and evolution of sexual dimorphism. *Nat. Rev. Genet* **10**, 797–804 (2009).
- Raymond, C. S. et al. Evidence for evolutionary conservation of sex-determining genes. *Nature* **391**, 691–695 (1998).
- Veitia, R. et al. Deletions of distal 9p associated with 46,XY male to female sex reversal: definition of the breakpoints at 9p23.3-p24.1. *Genomics* **41**, 271–274 (1997).
- Smith, C. A., McClive, P. J., Western, P. S., Reed, K. J. & Sinclair, A. H. Conservation of a sex-determining gene. *Nature* **402**, 601–602 (1999).
- Kim, S., Bardwell, V. J. & Zarkower, D. Cell type-autonomous and non-autonomous requirements for Dmrt1 in postnatal testis differentiation. *Dev. Biol.* **307**, 314–327 (2007).
- Matson, C. K. et al. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* **476**, 101–104 (2011).
- Lints, R. & Emmons, S. W. 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 (2002).
- Burtis, K. C. & Baker, B. S. Drosophila doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* **56**, 997–1010 (1989).
- Robinett, C. C., Vaughan, A. G., Knapp, J. M. & Baker, B. S. Sex and the single cell. II. There is a time and place for sex. *PLoS Biol.* **8**, e1000365 (2010).
- Rideout, E. J., Dornan, A. J., Neville, M. C., Eadie, S. & Goodwin, S. F. Control of sexual differentiation and behavior by the doublesex gene in *Drosophila melanogaster*. *Nat. Neurosci.* **13**, 458–466 (2010).
- Pan, Y. & Baker, B. S. Genetic identification and separation of innate and experience-dependent courtship behaviors in *Drosophila*. *Cell* **156**, 236–248 (2014).
- Peng, Q. et al. The sex determination gene doublesex is required during adulthood to maintain sexual orientation. *J. Genet. Genomics* **49**, 165–168 (2022).
- Zarkower, D. Establishing sexual dimorphism: conservation amidst diversity? *Nat. Rev. Genet* **2**, 175–185 (2001).
- Hopkins, B. R. & Kopp, A. Evolution of sexual development and sexual dimorphism in insects. *Curr. Opin. Genet. Dev.* **69**, 129–139 (2021).
- Wexler, J. et al. Hemimetabolous insects elucidate the origin of sexual development via alternative splicing. *eLife* **8**, e47490 (2019).
- Zhuo, J. C. et al. A feminizing switch in a hemimetabolous insect. *Sci. Adv.* **7**, eabf9237 (2021).
- Kiuchi, T. et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* **509**, 633–636 (2014).
- Nagoshi, R. N., McKeown, M., Burtis, K. C., Belote, J. M. & Baker, B. S. The control of alternative splicing at genes regulating sexual differentiation in *D. melanogaster*. *Cell* **53**, 229–236 (1988).
- Goralski, T. J., Edstrom, J. E. & Baker, B. S. The sex determination locus transformer-2 of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell* **56**, 1011–1018 (1989).
- McKeown, M., Belote, J. M. & Baker, B. S. A molecular analysis of transformer, a gene in *Drosophila melanogaster* that controls female sexual differentiation. *Cell* **48**, 489–499 (1987).
- Williams, T. M. et al. The regulation and evolution of a genetic switch controlling sexually dimorphic traits in *Drosophila*. *Cell* **134**, 610–623 (2008).
- Burtis, K. C., Coschigano, K. T., Baker, B. S. & Wensink, P. C. The doublesex proteins of *Drosophila melanogaster* bind directly to a sex-specific yolk protein gene enhancer. *EMBO J.* **10**, 2577–2582 (1991).
- Coschigano, K. T. & Wensink, P. C. Sex-specific transcriptional regulation by the male and female doublesex proteins of *Drosophila*. *Genes Dev.* **7**, 42–54 (1993).
- Sanders, L. E. & Arbeitman, M. N. Doublesex establishes sexual dimorphism in the *Drosophila* central nervous system in an isoform-dependent manner by directing cell number. *Dev. Biol.* **320**, 378–390 (2008).
- Graveley, B. R. et al. The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471**, 473–479 (2011).
- Brown, J. B. et al. Diversity and dynamics of the *Drosophila* transcriptome. *Nature* **512**, 393–399 (2014).
- Jenett, A. et al. A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**, 991–1001 (2012).
- Belote, J. M., Handler, A. M., Wolfner, M. F., Livak, K. J. & Baker, B. S. Sex-specific regulation of yolk protein gene expression in *Drosophila*. *Cell* **40**, 339–348 (1985).
- Baral, S., Arumugam, G., Deshmukh, R. & Kunte, K. Genetic architecture and sex-specific selection govern modular, male-biased evolution of doublesex. *Sci. Adv.* **5**, eaau3753 (2019).
- Li, S. et al. The genomic and functional landscapes of developmental plasticity in the American cockroach. *Nat. Commun.* **9**, 1008 (2018).
- Hughes, A. L. Runaway evolution of the male-specific exon of the doublesex gene in Diptera. *Gene* **472**, 1–6 (2011).
- Kato, Y., Kobayashi, K., Watanabe, H. & Iguchi, T. Environmental sex determination in the branchiopod crustacean *Daphnia magna*: deep conservation of a Doublesex gene in the sex-determining pathway. *PLoS Genet* **7**, e1001345 (2011).
- Peng, Q. et al. The sex determination gene doublesex regulates expression and secretion of the basement membrane protein Collagen IV. *J. Genet. Genomics* **49**, 636–644 (2022).
- Pan, Y., Robinett, C. C. & Baker, B. S. Turning males on: activation of male courtship behavior in *Drosophila melanogaster*. *PLoS One* **6**, e21144 (2011).
- Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform* **14**, 178–192 (2013).

### Acknowledgements

We thank the Bloomington *Drosophila* Stock Center and the Tsinghua Fly Center for fly stocks, and Dr. Sheng Li and Nan Chen (South China Normal University) for generously providing cDNA of *Blattella germanica*. This work was supported by grants from National Key R&D Program of China (2019YFA0802400), the National Natural Science Foundation of China (31970943 and 31700905), and the Jiangsu Innovation and Entrepreneurship Team Program.

### Author contributions

C.H., Q.P. and Y.P. designed research; C.H., Q.P., X.S. and L.X. generated transgenic flies; C.H., and X.S. performed behavioral experiments; C.H. and X. Ji. performed immunostaining and qPCR experiments; C.H., Q.P. and Y.P. analyzed data; and Y.P. wrote the manuscript with input from all authors.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-022-03664-7>.

**Correspondence** and requests for materials should be addressed to Yufeng Pan.

**Peer review information** *Communications Biology* thanks Maroun Bou Sleiman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Luke R. Grinham. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022