Tissue-specific developmental regulation and isoform usage underlie the role of doublesex in sex differentiation and mimicry in Papilio swallowtails

Riddhi Deshmukh, Dhanashree Lakhe and Krushnamegh Kunte

National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKVK Campus, Bellary Road, Bengaluru 560065, India

Adaptive phenotypes often arise by rewiring existing developmental networks. Co-option of transcription factors in novel contexts has facilitated the evolution of ecologically important adaptations. doublesex (dsx) governs fundamental sex differentiation during embryonic stages and has been co-opted to regulate diverse secondary sexual dimorphisms during pupal development of holometabolous insects. In Papilio polytes, dsx regulates female-limited mimetic polymorphism, resulting in mimetic and non-mimetic forms. To understand how a critical gene such as dsx regulates novel wing patterns while maintaining its basic function in sex differentiation, we traced its expression through metamorphosis in P. polytes using developmental transcriptome data. We found three key dsx expression peaks: (i) eggs in pre- and post-ovisposition stages; (ii) developing wing discs and body in final larval instar; and (iii) 3-day pupae. We identified potential dsx targets using co-expression and differential expression analysis, and found distinct, non-overlapping sets of genes—containing putative dsx-binding sites—in developing wings versus abdominal tissue and in mimetic versus non-mimetic individuals. This suggests that dsx regulates distinct downstream targets in different tissues and wing colour morphs and has perhaps acquired new, previously unknown targets, for regulating mimetic polymorphism. Additionally, we observed that the three female isoforms of dsx were differentially expressed across stages (from eggs to adults) and tissues and differed in their protein structure. This may promote differential protein–protein

© 2020 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
interactions for each isoform and facilitate sub-functionalization of \textit{dsx} activity across its isoforms. Our findings suggest that \textit{dsx} employs tissue-specific downstream effectors and partitions its functions across multiple isoforms to regulate primary and secondary sexual dimorphism through insect development.

1. Introduction

Evolutionary novelties, which often lead to diversification of life forms, have varied developmental origins. While complex adaptations could result from the emergence of novel genes [1], they could also arise by rewiring existing pathways. The pleiotropic nature of several genes, such as transcription factors, allows co-option of developmental circuits in novel contexts at relatively short timescales [2]. For instance, horns in beetles [3], wings in insects [4,5] and mimicry in butterflies [6–11] have resulted from tissue- and developmental stage-specific co-option of existing genes and pathways. Mutations in regulatory regions, synthesis of different gene products by alternative splicing, and mutations in binding sites that enable interactions with new partners, are a few mechanisms by which transcription factors may acquire functions that lead to novel adaptive phenotypes [12,13].

\textit{Doublesex (dsx)} is a pleiotropic transcription factor that governs sexual dimorphism in insects at various developmental stages. Apart from governing embryonic sex differentiation, it is often co-opted in developing pupa to govern secondary sexually dimorphic phenotypes in adults, such as sex combs in male \textit{Drosophila} [14], exaggerated horns and mandibles in male beetles [15–17] and caste differentiation in social ants [18]. In \textit{Papilio} butterflies, \textit{dsx} regulates female-limited Batesian mimicry and mimetic polymorphism [8,19]. \textit{Papilio polytes}, through most of its geographical range, exhibits two female forms, the mimetic form \textit{polytes}—which mimics the aposomatic \textit{Pachliopta aristolochiae}—and non-mimetic, male-like form \textit{cyrus} [20]. These forms are governed by differential expression of alternative alleles of \textit{dsx}, one of which is contained in an inversion and results in \textit{f. polytes} [8,19]. While the inversion maintains two separate alleles and prevents maladaptive intermediates in \textit{P. polytes}, in \textit{Papilio memnon}, mimetic polymorphism is maintained in the absence of an inversion with two alleles of \textit{dsx} [10]. It appears as if this critical gene has evolved multiple mechanisms to maintain and govern different morphs even in closely related species. Exploring the molecular role of \textit{dsx} in the regulation of sexually dimorphic and polymorphic traits in developing insects in all embryonic and other stages during the metamorphosis will help us understand the extent of its molecular pliability to accommodate novel functions and co-option in different developmental contexts while maintaining its basic function in sex differentiation. However, our understanding in this area is somewhat fragmented: the function of \textit{dsx} in embryonic sex differentiation has been studied very well in a few model insect species [21–32], whereas the co-opted functions have been studied in a larger number of non-model organisms [8,10,15–18]. It is necessary that we bridge this disconnect by comparing the action of pleiotropic genes in embryonic sex differentiation as well as that in late-development regulation of secondary sexual traits in a wide range of model and non-model species.

In this study, we investigate the mechanisms by which \textit{dsx} regulates sexual dimorphism in \textit{P. polytes} by using a developmental transcriptome and quantitative polymerase chain reaction (qPCR) datasets across its metamorphosis from ova and eggs to adults. Previously, only the expression profiles and isoform usage in pre-pupal and pupal stages had been studied in this species [8,19]. Thus, we provide new data on gene expression patterns as well as isoform usage in ova and eggs when sex differentiation occurs, and put this in context of the well-known pattern of expression and isoform usage in late pupal development when different wing colour morphs are laid down. Additionally, we identify potential interacting partners of this gene from fifth instar larvae to pupal stages to help elucidate how gene networks have been tweaked to accommodate the pleiotropic functions of \textit{dsx}. This will help generate a more coherent understanding of the action of \textit{dsx} from early sex differentiation to late-stage secondary sexual dimorphism in this emerging model species.

2. Materials and methods

2.1. Sample collection and sequencing

We sampled mimetic \textit{P. polytes} females and non-mimetic males from a greenhouse population of \textit{P. polytes} at different stages of their life cycle (electronic supplementary material, table S1) and separately preserved the tissues in TRIzol™. While sampling pre-oviposition eggs from females, we cleaned out as much somatic
tissue from around the eggs as feasible. We extracted RNA from these tissues using the chloroform-isopropanol-based extraction method and prepared libraries using TruSeq® RNA Sample Preparation Kit v2. We quantified the libraries using Qubit fluorometric quantification and checked their profile using Bioanalyzer. At the time of library preparation and then sequencing, we took equal RNA concentrations across samples to normalize differences between tissue mass (e.g. wing discs have less tissue mass compared to abdomen). Thus, the differential expression revealed in further analyses reflected true differences in gene expression across tissue and lifestages, and not differences in starting tissue mass. We sequenced the transcriptomes using 2 × 100 paired-end sequencing runs on Illumina HiSeq 2500 to obtain nearly 20 million reads per sample (accession numbers for raw data are in the electronic supplementary material, table S2).

2.2. Transcriptome assembly, differential expression and weighted gene correlation network analysis

We quality-checked the raw sequencing data, trimmed reads and aligned them to the P. polytes reference genome [19] using STAR 2.6 [33]. We obtained raw counts from the alignment using HTSeq [34], which we normalized to plot gene expression using the edgeR package [35] in R [36]. To determine differential expression between male and female tissues of 3-day pupal stage, which is critical for wing patterning [19], we used genewise negative binomial generalized linear models. We chose an false discovery rate cut-off of less than 0.001 and annotated differentially expressed genes using the P. polytes annotation on the National Centre for Biotechnology Information (NCBI; https://www.ncbi.nlm.nih.gov/genome/annotation_euk/Papilio_polytes/100/). To represent these in the form of a heatmap, we used the heatmap3 package in R [37]. We used the weighted gene correlation network analysis (WGCNA) package [38] to obtain modules of genes co-expressed with dsx in male wings, female wings, male abdomen and female abdomen separately, using read counts from fifth instar larvae, pre-pupae, 3-day-, 6-day- and 9-day-old pupae. We chose the smallest soft threshold power for which our data showed scale-free topology. These powers were different for each run (ranging from 6 to 13) based on the gene composition and connectivity of each tissue. We used a minimum module size of 30 genes, and in each run, dsx did not get sorted with unassigned genes. We used the dynamic tree cut function to merge modules with similar expression using the default parameters. We annotated the genes in the dsx-containing module using the P. polytes annotation from NCBI. For assembling isoforms of dsx, we performed de-novo assembly of raw reads using TRINITY [39] on the Galaxy server (https://usegalaxy.org) and used BLAST+ [40,41] to identify individual isoforms in each sample.

2.3. Expression and structure prediction of doublesex and its isoforms

We used reverse transcriptase-PCR (RT-PCR) and qPCR to test for the presence and expression levels of dsx and its isoforms F1, F2, F3 and the male isoform (M), across tissues and developmental stages of P. polytes (primer details in the electronic supplementary material, table S3). We used ProtoScript® II First Strand cDNA Synthesis Kit to synthesize cDNA and performed RT-PCR and qPCR using ribosomal protein RPL3 as the internal control. We used PowerUp SYBR Green Master Mix (ThermoFisher), for performing qPCR for samples that tested positive for isoform-specific RT-PCR, with the same primers. The three-dimensional structure of dsx largely comprises of loops, making tertiary structure prediction unreliable. Therefore, we predicted secondary structures of dsx isoforms F1, F2, F3 and M, using PSIPRED [42,43].

2.4. doublesex-binding motif discovery

We used Hypergeometric Optimization of Motif EnRichment (HOMER) v4.11 [44] to scan the P. polytes reference genome for dsx-binding motifs, using the canonical dsx-binding sequence from Drosophila melanogaster [45]. With an 11 bases long binding motif, we found nearly 4200 loci that had a motif similar to that of Drosophila dsx. We screened these using candidates from WGCNA and differential expression and found dsx-binding motifs in 246 of 478 genes (electronic supplementary material, table S5).

3. Results and discussion

3.1. doublesex shows multiple activity peaks through metamorphosis in Papilio polytes

We tracked dsx expression across various tissues during P. polytes metamorphosis using whole-transcriptome data and dsx-specific RT-PCR. The whole-transcriptome sequencing showed little dsx expression in eggs and
early larval instars compared to wings in pupal stages, suggesting the inherent expression differences at these two stages and tissues (figure 1a). However, sampling of eggs at a finer timescale using RT-PCR elucidated the pattern of dsx expression in developing eggs before and after oviposition (figure 1b). These two methods combined showed two stages with elevated and differential dsx activity: (i) egg stage, where ova and eggs before oviposition showed very high expression, and eggs from 6–36 h after oviposition showed lesser but prominent expression levels (figure 1b). This is similar to embryonic dsx activity in Bombyx mori [21], where ovarian eggs exclusively express female isoform of dsx (dsxF), whereas the male isoform (dsxM) starts expressing several hours post-oviposition, both contributing to early sex differentiation [22,25]. The transition period between the expression of female isoform to male isoform might account for variation in dsx expression after oviposition. Overall, the embryonic patterns of expression in P. polytes are in line with the patterns known in B. mori, showing that the basic function of dsx in sex differentiation may be conserved across these widely separated species; and (ii) peak dsx expression was observed in the developing wings of 3-day-old female pupae (figure 1a), as previously reported [19], followed by a gradual decrease through pupal development (figure 1a). In addition, there was discernable sexually dimorphic expression in wing discs of the final instar caterpillars, where female wing discs and body showed marginally higher expression levels compared to males (figure 1a). Taken together, it is possible that relatively low expression of dsx is sufficient for sex differentiation at the early egg developmental stage, but significant upregulation is perhaps required to lay down wing pattern differences in pupal stages. Moreover, it appears that the regulation of wing pattern polymorphism by dsx initiates in wing discs of the final larval instar and peaks at 3 days after pupation before slowly trailing off, but maintaining the female-biased upregulation throughout (figure 1a).

### 3.2. doublesex targets different genes in wings and abdomen of mimetic and non-mimetic wing forms

We identified potential dsx targets in tissues with high dsx activity, such as developing wings and abdomen in all stages where sexes could be distinguished (fifth instar larvae onwards). We used WGCNA to identify
genes that co-expressed with dsx in mimic and non-mimetic butterflies. This analysis revealed stark differences between the functionally relevant and non-overlapping pools of downstream targets in mimic and non-mimetic wings. Genes involved in various physiological and metabolic processes co-expressed with dsx in each comparison (figure 2 and electronic supplementary material, table S4), including genes that participate in pathways that dsx is involved in and genes that are important in wing development and patterning, such as: (i) dishevelled protein DVL-3 homologue, which is involved in the Wnt signalling pathway (figure 2). Wnt ligands are involved in wing patterning in several butterfly species: WntA in nymphal butterflies [46,47] and Wnt1 and Wnt6 in f. polytes of P. polytes [48]. The activity of Wnt1 and Wnt6 is tied to that of dsx in f. polytes and both genes are downregulated in dsx knockdowns; (ii) trithorax protein osa-like, a gene that shows female-biased expression in 3-day pupae and its counterpart in Drosophila is required for activation of genes such as Antp, Ubx and Eve, as well as for repression of wingless-regulated genes such as dll, and dpp, in the absence of Wg signal [49,50]; and (iii) lin, a regulator of Abd-B in Drosophila. However, lin was not highly expressed in the 3-day pupal stage, a characteristic of dsx expression in mimic females, so it is perhaps not involved in regulating mimic wing morph in P. polytes. There was no overlap in co-expressed genes between wings and abdominal tissue, implying the existence of discrete dsx-associated molecular networks during wing versus genitalia development, with dsx potentially regulating different downstream genes in both
contexts. Several tissue-specific genes co-expressed with dsx: wing development and patterning genes in mimetic wings, and genes involved in lipid metabolism in mimetic abdomens (figure 2). On the other hand, non-mimetic wings showed few wing patterning genes co-expressing with dsx. This may have resulted from low expression of dsx in non-mimetic wings compared to mimetic wings leading to spurious co-expression hits. Some genes that co-expressed with dsx in mimetic and non-mimetic wings did not show sex-specific expression in the wings. We retained these genes in figure 2 as they may have a more general role to play in wing development. Overall, nearly all the context-specific genes for each comparison also contained at least one dsx-binding motif (denoted by an asterisk in figure 2, complete list in the electronic supplementary material, table S5). Thus, dsx may have acquired non-canonical targets such as Wnt pathway genes [48] in addition to its known targets in mimetic individuals to govern wing pattern polymorphism. However, this needs to be verified by chromatin-binding and protein interaction assays such as ChIP-seq and co-IP in P. polytes. We also assessed differentially expressed genes in 3-day pupae (dsx expression peak) of non-mimetic and mimetic butterflies to find relevant potential dsx targets (electronic supplementary material, figure S1). Among genes that showed high expression in mimetic wings compared to non-mimetic wings, and that also contained dsx-binding motifs, hedgehog homologue (desert hedgehog B) and hairy emerged as potential candidates owing to their involvement in morphogenesis and bristle formation, respectively (bristles are equivalent to scales in Lepidoptera; electronic supplementary material, figure S1).

3.3. Female isoforms of doublesex co-express in most tissues, but differ in expression levels and protein structures

Three female isoforms of dsx occur in P. polytes, which differ in their extent of usage of exons 3 and 4 (figure 3a) [19,51]. We traced stage- and tissue-specific isoform activity using a combination of isoform-specific search in de-novo assembled transcriptomes and PCRs using isoform-specific primers (figure 3b and electronic supplementary material, table S6). We observed that all three isoforms were expressed across most dsx-expressing tissues during development. Although poorly detected using RT-PCR (electronic supplementary material, table S6), qPCR identified measurable F1 isoform activity in several developmental stages and tissues of P. polytes females. In most tissues, however, this expression was low compared to that of isoforms F2 and F3 which seemed to contribute most to dsx expression (figure 3b and electronic supplementary material, table S6). This was also observed in non-mimetic females (f. cyrus) in 3-day pupae (peak dsx expression). Isoform-specific expression was similar between mimetic and non-mimetic females, while male-specific isoform showed low expression compared to female isoforms in most tissues except abdomen (figure 3c). The difference in relative expression in 3-day pupae in panels (b) and (c) can be attributed to developmental differences at the individual level among pupae sampled at the 3-day stage. The dichotomy in the expression of isoforms may be owing to varying strength of splice sites between exons 3 and 4. Selective mutation or blocking of splice sites might illustrate their role in isoform expression. Alternatively, isoform F1 might perform a subset of dsx functions, which may be specific to early embryonic development and pupal abdomens, where it shows higher expression. The secondary structure prediction for dsx isoforms showed that they differed at their C-terminal ends (figure 4b and electronic supplementary material, figure S2). Isoforms F1 and F3 showed the presence of helices, while isoform F2 lacked a secondary structure in this region. The C-terminal domain of dsx contains the dimerization domain and sex-specific region that is important for the recruitment of cofactors for regulatory functions. Moreover, dimerization or protein–protein interactions of dsx can influence DNA binding [54]. It is possible that female isoforms regulate different sets of genes and partition dsx activity. However, the relevance of these structural differences needs further investigation.

3.4. Proposed mechanism of developmental regulation of sex differentiation and secondary sexual traits by doublesex

dsx is a crucial gene that regulates various aspects of sexual dimorphism, from embryonic sex differentiation to pupal and adult expression of secondary sexual traits. Its modular architecture with individual genic regions evolving under different selection pressures [51] might facilitate maintenance of its core function in sex differentiation while accommodating the evolution of novel ones in later development. How does a highly conserved gene such as dsx regulate different phenotypic outcomes with a limited set of developmental tricks? To address this question, we propose a mode of action for
Figure 3. Expression of female isoforms of dsx in developing Papilio polytes. (a) dsx occurs as three isoforms in P. polytes females depending on the usage of female-specific exons 3 and 4. (b) Expression of female isoforms of dsx using qPCR. The number at the base of the bar indicates sample size for each tissue and stage. The female isoforms of dsx show different expression levels across tissues and stages, with F1 usually expressing at lower levels in most tissues compared to F2 and F3, which contribute to most of the dsx expression. (c) Expression of dsx isoforms in mimetic and non-mimetic female forms and males at 3-day pupal stage.

dsx in different tissues based on our findings and other studies (figure 4). Figure 4a is largely based on what we know from other studies [21,22,25,52], and our hypotheses in P. polytes are drawn upon this existing information. Our co-expression analysis in addition to previous work [48] indicates that dsx might be governing different targets in mimetic and non-mimetic wings, from which figure 4b is drawn. Panel (c), and specifically our hypothesis that the three isoforms regulate different, potentially overlapping pools of genes, is based largely on our data on dsx isoforms (figure 3 and electronic supplementary material, figure S2). We also take into account the allelic status and tissue specificity of
dsx in regulating wing pattern polymorphism and reproductive traits in male and female *P. polytes*, which has not been considered fully before. *doublesex* may have fundamentally different modes of action for early embryonic sex differentiation and late developmental (final instar and pupal) regulation of secondary sexual traits.

The embryonic sex differentiation may be based purely on male-specific (*dsxM*) and female-specific (*dsxF*) isoforms and their downstream targets that set in motion basic sex differences, as seen in other insects [51,55]. In *P. polytes*, mimetic alleles do not seem to affect larval phenotype and development [52], and allele-specific phenotypic changes only start appearing late during pupal stages [53]. Therefore, embryonic sex differentiation may occur irrespective of the allelic status of an individual (figure 4a). While the role of *dsx* isoforms in sex-determination has been explored in *B. mori*, we lack such information in *P. polytes*, which has the additional complexity of two *dsx* alleles. It is conceivable that secondary sexual traits in simple sexual dimorphisms (i.e. those without polymorphisms), such as genitalia structures and production of eggs and associated fat bodies, may be regulated similarly by sex-specific isoforms. By regulating similar [56,57] or different downstream genes, *dsxM* and *dsxF* can govern sex-specific developmental cascades establishing basic sexual differentiation, irrespective of multiple *dsxF* isoforms in *P. polytes* (figure 4a). However, polymorphic secondary sexual traits might be regulated by *dsx* through more complex developmental processes owing to their sex-limited nature. For example, caste-related polymorphism in social insects, and polymorphic Batesian mimicry in *P. polytes* are female-limited and both governed by *dsx*. The switch between mimetic and non-mimetic

![Figure 4. Proposed mechanism outlining the role of dsx in early-development sex differentiation and novel, late-development wing pattern polymorphism in *Papilio polytes*. (a) Early embryonic sex differentiation and late developmental sex-specific reproductive traits (e.g. genitalia and eggs) are highly canalized and strictly sex-specific, irrespective of allelic polymorphism and alternatively spliced isoforms [52,53]. (b) Late developmental novel mimetic wing patterns are produced by a combination allelic polymorphism, differential expression of alternatively spliced isoforms, and sex-limitation, expressing mimetic wing patterns only in females that contain at least one copy of the mimetic H allele of dsx (figure 2, [8,19,48]). (c) A graphic representation of relative Dsx isoform activity in regulation of polymorphic wings and other tissues (based on data from [8,19] and figure 3 from this study). Dsx F1 is downregulated in relation to F2 and F3 during the production of secondary sexual traits, but expressed comparably during sex differentiation and basic reproductive traits. The C terminal ends of the three DsxF isoforms are different (based on the electronic supplementary material, figure S2), and molecular identities of downstream targets of *dsx* that regulate mimetic wing patterns ([48] and figure 2; electronic supplementary material, figure S1 in this study) have recently been identified. This model has several unexplored aspects that need to be studied in the future, including upstream regulators of *dsx* in each context, precise downstream targets, and both allele- and isoform-specific and protein-interacting partners of *dsx*. Such genes, regulators and networks with as yet unknown molecular genetic identities and mechanisms are marked with ?.
wing phenotypes can be attributed to \textit{dsx} alleles, of which the mimetic allele shows high expression in developing wings early in pupal development [19]. The amino acid substitutions (or possibly \textit{cis}-regulatory mutations) between the two alleles might enable the mimetic allele to regulate different downstream targets in mimetic females compared to the non-mimetic allele in female form \textit{cyrus} and \textit{dsxM} in males [48] irrespective of allelic composition (figure 4b). Targeted genetic manipulation in coding and non-coding regions of the \textit{dsx} gene, in addition to chromatin-binding assays can help elucidate the repertoire of targets \textit{dsx} interacts with in mimetic and non-mimetic female forms and identify the corresponding sites in the \textit{dsx} sequence crucial for these interactions.

\textit{dsx} can obtain finer control over the regulation of tissue and developmental stage-specific phenotypes by means of alternative isoforms. Our results showed that isoforms F2 and F3 contributed to most of the \textit{dsx} expression in mimetic wings, whereas eggs and 6-day pupal tissues showed somewhat comparable expression of all three isoforms (figures 3b, 4c). This suggests that developing eggs and abdominal tissue may employ the entire repertoire of \textit{dsx} regulation by all three isoforms, but in wings, elevated expression of \textit{dsx} isoforms F2 and F3, without F1, might be involved in regulating mimetic phenotype. The occurrence of \textit{dsx} isoforms is not universal across insects: (i) isoform F2 from Lepidoptera shares sequence similarity with \textit{dsxF} isoform in Diptera and Coleoptera, (ii) isoform F3 from Lepidoptera shares sequence similarity only with \textit{dsxF} from Coleoptera, and (iii) isoform F1 appears to be unique to Lepidoptera [51]. It is possible that isoform F2, which is most conserved, plays a larger role in basic sex differentiation, which is universal, while F3—which is shared between insect orders Lepidoptera and Coleoptera with super-diverse wing patterns—might play a larger role in regulating wing patterning, among other sexually dimorphic traits. Similarly, the Lepidoptera-specific F1 might be involved in regulating aspects of early and late sex differentiation that are unique to Lepidoptera. It is yet unknown what might drive this differential regulation of downstream targets of \textit{dsx}. One intriguing possibility is that this differential, possibly non-overlapping regulation of downstream targets of \textit{dsx} is facilitated by distinct C-terminal ends of \textit{dsxF} isoforms, which we are showing here, to our knowledge for the first time (figure 4c). This can be experimentally validated with genetic manipulation of the exonic regions unique to each isoform along with isoform-specific ChIP-seq.

The work presented here makes two advances. First, we show that, similar to \textit{B. mori}, \textit{P. polytes} embryos produce sex-specific \textit{dsxF} and \textit{dsxM} isoforms. These sex-specific isoforms presumably have similar functions in early sex differentiation cascades, irrespective of multiple \textit{dsxF} isoforms in \textit{P. polytes}. Our expression analysis also appears to suggest that the low \textit{dsx} expression during early embryonic development may be sufficient for sex differentiation, whereas the very high 3-day pupal expression may be necessary for regulating mimetic polymorphism. Both these hypotheses about the expression levels and isoform functions should be experimentally tested in the future. Second, we demonstrate with WGCNA that \textit{dsx} might employ different downstream effectors in developing wings and abdomens of mimetic and non-mimetic individuals. The acquisition of new \textit{dsx}-binding sites may also help it acquire new targets, such as \textit{Wnt} signalling pathway genes [48]. The added complexity of having multiple female isoforms might assist \textit{dsx} in partitioning some of its functions and enabling tissue-specific regulation. Besides Lepidoptera, Coleoptera and Blattodea have two female-specific exons of \textit{dsx} [51,58]. Therefore, these insect groups may have multiple female isoforms, with the inclusion of one or both female exons. Comparison of developmental transcriptomes across species or even orders might help illustrate how regulation by \textit{dsx} has evolved across different sexually dimorphic traits. Further, morph-specific validation of chromatin binding of \textit{dsx} and identification of its protein co-factors in sexually dimorphic adaptations would help understand the mechanisms by which \textit{dsx} has become the jack of all sexually dimorphic trades.

Data accessibility. The raw RNA seq data are deposited in the NCBI SRA database (BioProject PRJNA634605, Accession nos. SAMN15001929 to SAMN15002046; https://dataview.ncbi.nlm.nih.gov/object/PRJNA634605?reviewer=mf4q72cfvq67spj56gl1sduop). Accession details for individual raw transcriptome sequences in the NCBI SRA database are given in the electronic supplementary material, table S2. Primer sequences, isoform sequences, sample details, etc. are provided in the electronic supplementary material, Information.

Authors’ contributions. R.D. and K.K. designed research; R.D. generated and analysed transcriptome data and wrote the paper; R.D. generated molecular data and D.L. performed motif discovery; K.K. conceived and supervised the project, and wrote the paper with R.D.

Competing interests. The authors declare no competing interests.

Funding. This work was partially funded by a Ramanujan Fellowship from the Department of Science and Technology, Government of India, and an NCBS Research Grant to K.K., support of the Department of Atomic Energy, Government of India, under project nos. 12-R&D-TFR-5.04-0800 and 12-R&D-TFR-5.04-0900 to TIFR/NCBS, a CSIR Shyama Prasad Mukherjee Fellowship to R.D. and an NCBS student fellowship to D.L.
Acknowledgements. We thank Bhavya Dharmaraj, Athulya Girish and Tulsamma for maintenance of greenhouse populations of *Papilio polytes*; Vaishali Bhaumik and Dipendra Nath Basu for help with R codes; Sai Guha and Sarvesh Menon for help with WGCNA analysis and NCBS Greenhouse Facility for breeding of butterflies.

References

1. Santos ME, Le Bouquin A, Crumière AJH, Khila A. 2017 Taxon-restricted genes at the origin of a novel trait allowing access to a new environment. *Science* 358, 386–390. (doi:10.1126/science.aan2748)

2. Oliver JC, Tong X-L, Gall LF, Piel WH, Monteiro A. 2012 A single origin for Nymphalid butterfly eye-spots followed by widespread loss of associated gene expression. *PLoS Genet.* 8, e1002893. (doi:10.1371/journal.pgen.1002893)

3. Hu Y, Lin D, Moczek AP. 2019 Beetle horns evolved from wing serial homologs. *Science* 366, 1004–1007. (doi:10.1126/science.aaw2980)

4. Clark-Hachtel CM, Linz DM, Tomoyasu Y. 2013 Insights into insect wing origin provided by functional analysis of vestigial in the red flour beetle, Tribolium castaneum. *Proc. Natl Acad. Sci.* USA 110, 16 951–16 956. (doi:10.1073/pnas.1303432110)

5. Tomoyasu Y, Arakane Y, Kramer KJ, Denell RE. 2009 Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. *Curr. Biol.* 19, 2057–2065. (doi:10.1016/j.cub.2009.11.014)

6. Reed RD et al. 2011 optic drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* 333, 1137–1141. (doi:10.1126/science.1208227)

7. Nadeau NJ et al. 2016 The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* 534, 106–110. (doi:10.1038/nature17961)

8. Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kenforst MR. 2014 doublesex is a mimicry superngene. *Nature* 507, 229–232. (doi:10.1038/nature13112)

9. Timmermans MUTI et al. 2014 Comparative genomics of the mimicry switch in Papilio dardanus. *Proc. R. Soc. B* 281, 20140465–. (doi:10.1098/rspb.2014.0465)

10. Ijima T, Kajitani R, Komata S, Lin C-P, Sota T, Itoh T, Fujiwara H. 2018 Parallel evolution of genomics of the mimicry switch in *Bombyx mori*. *Nature* 554, 232–236. (doi:10.1038/s41586-018-0021-3)

11. Deshmukh R, Baral S, Gandhimathi A, Gotoh H, Zinna RA, Warren I, DeNieu M, Niimi T, Vaishali Bhaumik and Dipendra Nath Basu for help with R codes; Sai Guha and Sarvesh Menon for help with WGCNA analysis and NCBS Greenhouse Facility for breeding of butterflies.

12. Monteiro A. 2012 Gene regulatory networks reused to build novel traits. *Bioessays* 34, 181–186. (doi:10.1002/bies.201100160)

13. Tanaka K, Barmina O, Sanders LE, Arbeitman MN, Kopp A. 2011 Evolution of sex-specific traits through changes in HOX-dependent doublesex expression. *PLoS Biol.* 9, e1001131. (doi:10.1371/journal.pbio.1001131)

14. Gottho H, Zinna RA, Warren I, Delieu M, Niimi T, Dworkin I, Enmien DJ, Miura T, Lavine LC. 2016 Identification and functional analyses of sex determination in the genetic sex of dimorphic stag beetle *Clytomaus metallifer*. *BMC Genomics* 17, 250. (doi:10.1186/s12864-016-2522-4)

15. Ito Y et al. 2013 The role of doublesex in the evolution of exaggerated horns in the Japanese rhinoceros beetle, *EMBO Rep.* 14, 561–567. (doi:10.1038/embor.2013.50)

16. Gottho H, Ishiguro M, Nishikawa H, Morita S, Okada K, Miyatake T, Yaginuma T, Niiimi T. 2016 Molecular cloning and functional characterization of the sex-determination gene doublesex in the sexually dimorphic broad-horned beetle *Gonocephala cornutus* (Coleoptera, Tenebrionidae). *Sci. Rep.* 6, 29337. (doi:10.1038/srep29337)

17. Klein A, Schulten F, Lowak H, Schrader L, Heinze J, Holman L, Oettler J. 2016 Evolution of social insect polyphenism facilitated by the sex differentiation cascade. *PLoS Genet.* 12, e1005952. (doi:10.1371/journal.pgen.1005952)

18. Nishikawa H et al. 2015 A genetic mechanism for female-limited Batesian mimicry in *Papilio butterfly*. *Nature Genet.* 47, 1–7. (doi:10.1038/ng.3241)

19. Clarke CA, Sheppard PM. 1972 The genetics of courtship song. *Proc. R. Soc. B* 263, 431–458.

20. Sakai H, Aski F, Suzuki MG. 2014 Identification of the key stages for sex determination in the silkworm, *Bombyx mori*. *Dev. Genes Evol.* 224, 119–123. (doi:10.1007/s00427-013-0461-9)

21. Suzuki MG, Funaguma S, Kanda T, Tamura T, Shimada T. 2005 Role of the male BdDSX protein in the sexual differentiation of *Bombyx mori*. *Dev. Biol.* 37, 58–68. (doi:10.1111/j.1525-1648.2005.00507.x)

22. Mellert DJ, Robbinett CC, Baker BS. 2012 doublesex functions early and late in gustatory sense organ development. *PLoS ONE* 7, e51489. (doi:10.1371/journal.pone.0051489)

23. Baker BS. 1989 Sex in flies: the splice of life. *Nature* 340, 521–524. (doi:10.1038/340521a0)

24. 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. (doi:10.1007/s00427-003-0334-8)

25. Siegal ML, Baker BS. 2005 Functional conservation and divergence of intersex, a gene required for female differentiation in *Drosophila melanogaster*. *Dev. Genes Evol.* 215, 1–12. (doi:10.1007/s00427-004-0445-x)

26. Zhao S, Guo Y, Sheng O, Shyr Y. 2014 Heatmap: an improved heatmap package with more powerful and convenient features. *BMC Bioinf.* 15, P16. (doi:10.1186/1471-2105-15-510-P16)

27. Langfelder P, Horvath S. 2008 WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. (doi:10.1186/1471-2105-9-559)
39. Grabherr MG et al. 2011 Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. (doi:10.1038/nbt.1883)

40. Altschul S. 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402. (doi:10.1093/nar/25.17.3389)

41. Gertz EM, Yu Y-K, Agarwala R, Schäffer AA, Altschul SF. 2006 Composition-based statistics and translated nucleotide searches: improving the tBLASTn module of BLAST. BMC Biol. 4, 41. (doi:10.1186/1471-2180-4-41)

42. Buchan DW A, Jones DT. 2019 The PSIPRED programs. A new generation of protein database search with a reference genome. Nucleic Acids Res. 25, 3389–3402. (doi:10.1093/nar/25.17.3389)

43. Buchan DW A, Jones DT. 2019 The PSIPRED programs. A new generation of protein database search with a reference genome. Nucleic Acids Res. 25, 3389–3402. (doi:10.1093/nar/25.17.3389)

44. Heinz S et al. 2010 Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589. (doi:10.1016/j.molcel.2010.05.004)

45. Shazman S, Lee H, Socoll Y, Mann RS, Honig B. 2014 OnTheFly, a database of Drosophila melanogaster transcription factors and their binding sites. Nucleic Acids Res. 42, D167–D171. (doi:10.1093/nar/gkt1165)

46. Mazo-Vargas A et al. 2017 Macromolecular networks of Wnt4 function potentiates butterfly wing-pattern diversity. Proc. Natl Acad. Sci. USA 114, 10 701–10 706. (doi:10.1073/pnas.1708149114)

47. Martin A et al. 2012 Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. Proc. Natl Acad. Sci. USA 109, 12 632–12 637. (doi:10.1073/pnas.1204800109)

48. Iijima T, Yoda S, Fujiiwara H. 2019 The mimetic wing pattern of Papilio polyxenes butterflies is regulated by a doublesex-orchestrated gene network. Commun. Biol. 2, 257. (doi:10.1038/s42003-019-0510-7)

49. Vázquez M, Moore L, Kennison JA. 1999 The doublesex proteins. J. Mol. Biol. 292, 195–202. (doi:10.1006/jmbi.1999.3091)

50. Collins RT, Treisman JE. 2000 Osa-containing brahma chromatin remodeling complexes are required for the repression of Wingless target genes. Genes Dev. 14, 3140–3152. (doi:10.1101/gad.85400)

51. Baral S, Gandhimathi A, Deshmukh R, Kunte K. 2013 Macroevolutionary shifts in the brahma chromatin-remodeling factor to regulate transcription. Development 140, 733–742.

52. Kopp A, Duncan I, Carroll SB. 2000 Genetic control and evolution of sexually dimorphic characters in Hemimetabolous insects elucidate the origin of sexual development via alternative splicing. Elife 8, e47490. (doi:10.7554/eLife.47490)

53. Nishikawa H, Iga M, Yamaguchi J, Saito K, Kakeka H, Suzuki Y, Sugano S, Fujiiwara H. 2013 Molecular basis of wing coloration in a Batesian mimic butterfly, Papilio polyxenes. Sci. Rep. 3, 1–9. (doi:10.1038/srep03184)

54. Cho S, Wensink PC. 1998 Linkage between oligonization and DNA binding in Drosophila Doublesex proteins. Biochemistry 37, 589–606. (doi:10.1021/bi872916a)

55. Verhulst EC, van de Zande L. 2015 Double nexus—Doublesex is the connecting element in sex determination. Brief. Funct. Genomics 14, 396–406. (doi:10.1093/bfgp/elv005)

56. Kopp A, Duncan I, Carroll SB. 2000 Genetic control and evolution of sexually dimorphic characters in Drosophila. Nature 408, 553–559. (doi:10.1038/35046017)

57. Burtis KC, Coschigano KT, Baker BS, Wensink PC. 1998 Linkage between oligonization and DNA binding in Drosophila Doublesex proteins. Biochemistry 37, 589–606. (doi:10.1021/bi872916a)

58. Wexler J, Delaney EK, Belles X, Schal C, Wada-Katsuruma A, Amicucci MJ, Kopp A. 2019 Genetic architecture and sex-specific selection govern modular, male-biased evolution of doublesex. Sci. Adv. 5, eaaw3753. (doi:10.1126/sciadv.aaw3753)

59. Komata S, Lin C-P, Sota T. 2018 Do juvenile developmental and adult body characteristics differ among genotypes at the doublesex locus that controls female-limited Batesian mimicry polymorphism in Papilio memnon?: a test for the ‘cost of mimicry’ hypothesis. J. Insect Physiol. 107, 1–6. (doi:10.1016/j.jinsphys.2018.02.001)