


# Novel *Doublesex* Duplication Associated with Sexually Dimorphic Development of Dogface Butterfly Wings

Fernando Rodriguez-Caro,<sup>\*1</sup> Jennifer Fenner,<sup>2</sup> Shivam Bhardwaj,<sup>2</sup> Jared Cole,<sup>3</sup> Caleb Benson,<sup>2</sup> Alexandra M. Colombara,<sup>2</sup> Riccardo Papa,<sup>4</sup> Matthew W. Brown,<sup>5</sup> Arnaud Martin,<sup>6</sup> Ryan C. Range,<sup>2</sup> and Brian A. Counterman <sup>\*2</sup>

<sup>1</sup>Division of Biological Sciences, University of Montana, Missoula, MT, USA

<sup>2</sup>Department of Biological Sciences, Auburn University, Auburn, AL, USA

<sup>3</sup>Department of Integrative Biology, University of Texas, Austin, TX, USA

<sup>4</sup>Department of Biological Sciences, University of Puerto Rico—Rio Piedras, San Juan, PR, USA

<sup>5</sup>Department of Biological Sciences, Mississippi State University, Mississippi State, MS, USA

<sup>6</sup>Department of Biological Sciences, The George Washington University, Washington, DC, USA

\*Corresponding authors: E-mails: luis.rodriguezcaro@umontana.edu; bac0071@auburn.edu.

Associate editor: John True

## Abstract

Sexually dimorphic development is responsible for some of the most remarkable phenotypic variation found in nature. Alternative splicing of the transcription factor gene *doublesex* (*dsx*) is a highly conserved developmental switch controlling the expression of sex-specific pathways. Here, we leverage sex-specific differences in butterfly wing color pattern to characterize the genetic basis of sexually dimorphic development. We use RNA-seq, immunolocalization, and motif binding site analysis to test specific predictions about the role of *dsx* in the development of structurally based ultraviolet (UV) wing patterns in *Zerene cesonia* (Southern Dogface). Unexpectedly, we discover a novel duplication of *dsx* that shows a sex-specific burst of expression associated with the sexually dimorphic UV coloration. The derived copy consists of a single exon that encodes a DNA binding but no protein-binding domain and has experienced rapid amino-acid divergence. We propose the novel *dsx* paralog may suppress UV scale differentiation in females, which is supported by an excess of *Dsx*-binding sites at cytoskeletal and chitin-related genes with sex-biased expression. These findings illustrate the molecular flexibility of the *dsx* gene in mediating the differentiation of secondary sexual characteristics.

**Key words:** wing pattern, Lepidoptera, gene expression, ultraviolet, sexual dimorphism.

## Introduction

Explaining the origin of morphological variation is a major challenge in evolutionary and developmental biology. Some of the most dramatic intraspecific variation in nature emerges from sexual dimorphism such as the striking colorations of birds of paradise, the antlers of male elks, and elaborate horns of rhinoceros beetles that result from the evolution of sex-specific developmental pathways (Sibley 1957; Jarman 1983; Emlen et al. 2007; Coyne et al. 2008). Such traits often evolve under the influence of sexual selection (Emlen 2001; Emlen et al. 2007; Irestedt et al. 2009) and are key elements of the evolutionary ecology of species. The development of these sexually dimorphic traits in males and females that have nearly identical genomes is often achieved through patterns of sex-specific gene expression.

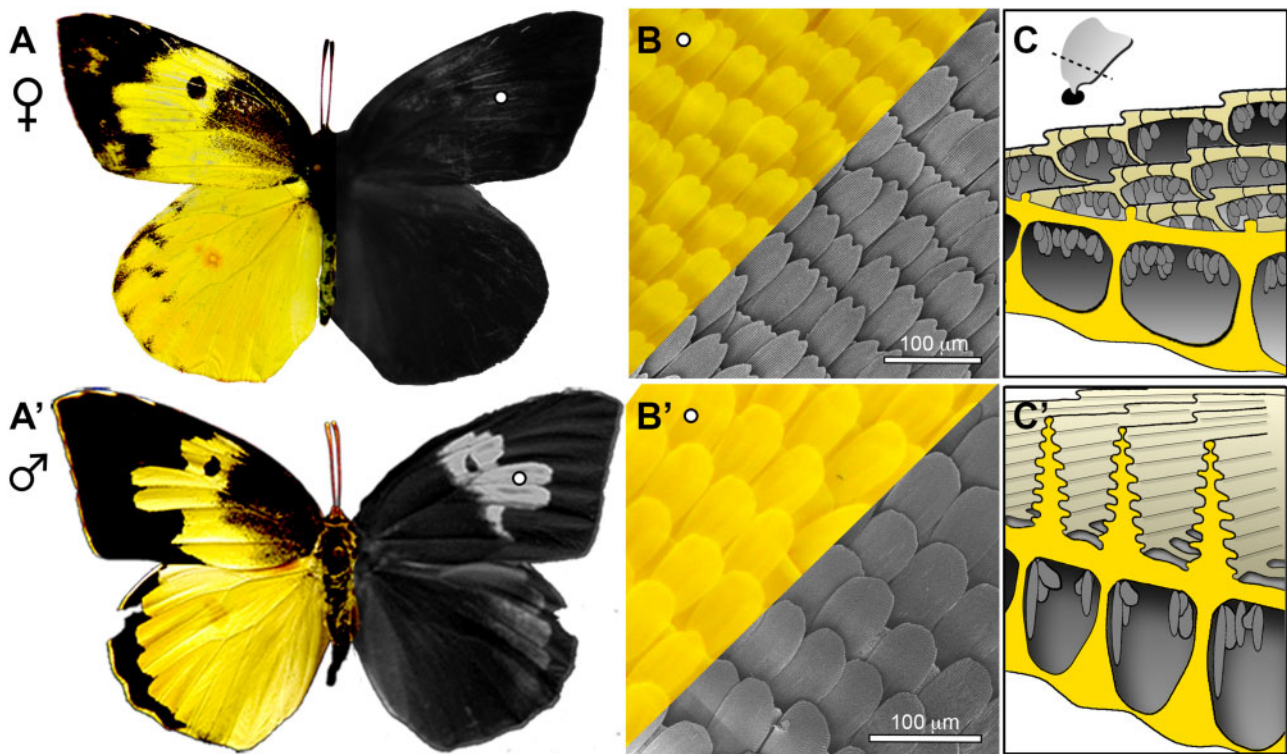
Sexually dimorphic development involves the *doublesex*/*mab-3* related (*Dmrt*) family of transcription factors that share a common DNA-binding domain (Kopp 2012). *Dmrt* genes are responsible for initiating cascades of tissue-specific

gene expression patterns that drive populations of cells to develop with male or female fates. In insects, the *doublesex* gene (*dsx*, also known as *Dmrt1*) functions as a molecular switch at the base of the sex determination cascade through alternative splicing, a highly conserved mechanism across all orders of insects studied (Price et al. 2015; Hopkins and Kopp 2021). Interestingly, *dsx* has been independently co-opted multiple times during the evolution of novel secondary sexual dimorphisms (Kunte et al. 2014; Deshmukh et al. 2020). For example, in *Drosophila*, tissue-specific expression of alternative *dsx* isoforms is essential for proper development of a myriad of sexually dimorphic traits (Verhulst and van de Zande 2015). In butterflies, *dsx* underlies the repeated evolution of female-limited mimetic wing patterns (Nishikawa et al. 2015; Iijima et al. 2019; Palmer and Kronforst 2020) and is required for the development of male-specific scent organs (Prakash and Monteiro 2020). These findings highlight the importance of *dsx* in the evolution of insects, not only as an ancient and highly conserved developmental switch, but

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access



**FIG. 1.** Sexual dimorphic UV patterns on *Zerene cesonia* wings. (A–A') Female and male visual (left) and UV (right) color patterns. (B–B') Light and scanning electron microscope (SEM) images show differences in cell arrangements in females and males (images taken from wing region denoted by white circle in A). (C–C'). Drawing renditions of scale cross-sections highlighting differences in chitin cytoskeletal differences between female non-UV reflecting scales and male UV-reflecting scales (based on focused ion beam SEM images from Fenner et al. [2019]).

also as a ubiquitous element in the evolution of novel sexually dimorphic traits (Burtis and Baker 1989; Kimura et al. 2008; Kijimoto et al. 2012; Ito et al. 2013).

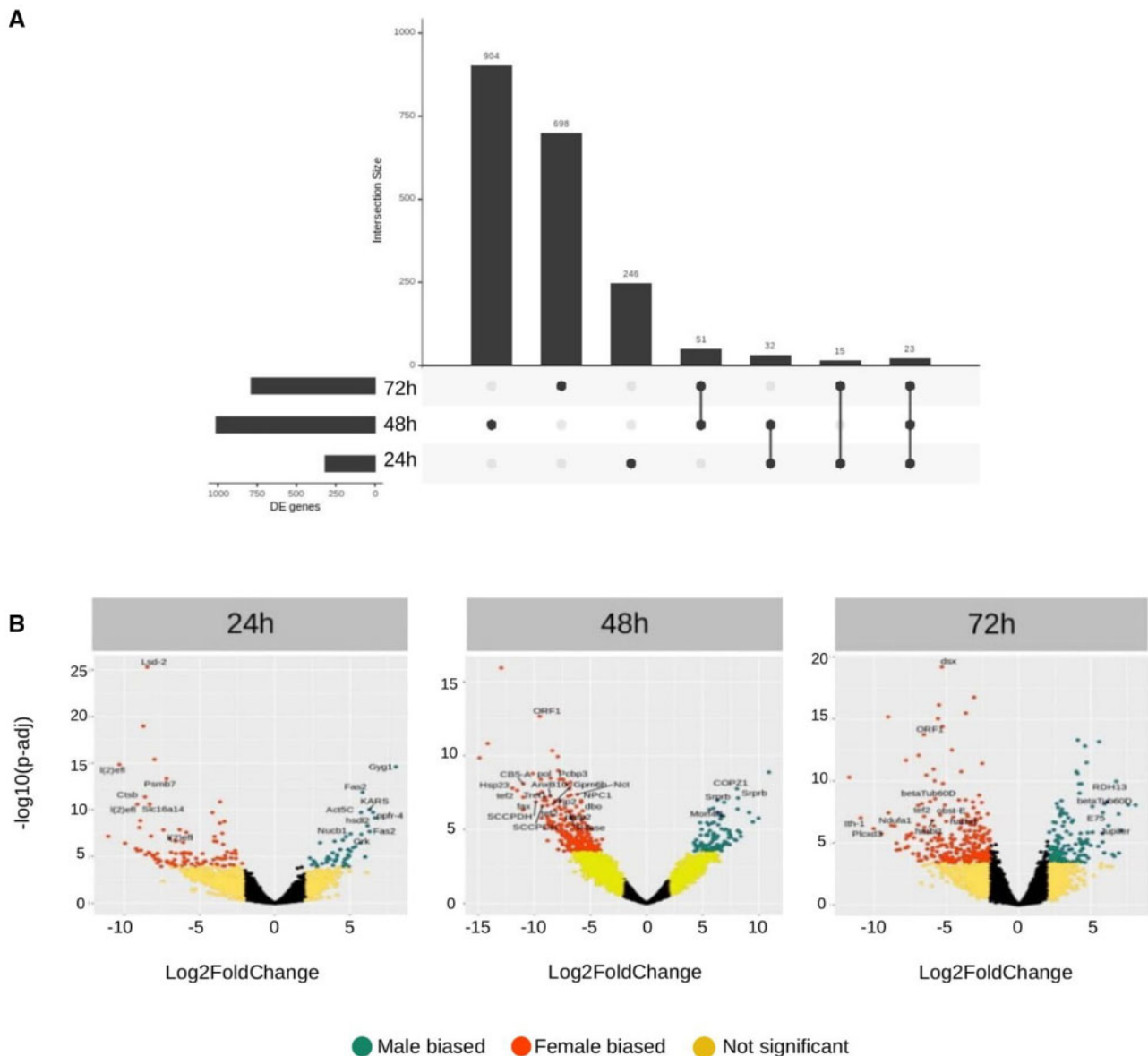
Coliadinae butterflies provide an opportunity to test the ubiquity of *dsx* in the evolution of novel phenotypes. In this group, sexually dimorphic ultraviolet (UV) wing patterns have evolved rapidly, resulting in dramatic diversity within and between species (Silberglied and Taylor 1973; Rutowski 1977; Kemp and Rutowski 2011; Stella et al. 2018; Pecháček et al. 2019). Studies across the Coliadinae clade have found that UV dimorphism is frequently a sexually selected trait involved in speciation (Kemp et al. 2005; Kemp 2008). Therefore, UV reflectance has been a highly influential trait in the diversification of Coliadinae butterflies. In the *Colias* genus, UV brightness has been shown to be the greatest indicator of male mating success and important for interspecific mate choice (Rutowski 1977; Silberglied and Taylor 1978; Papke et al. 2007). In the sister genus, *Zerene*, UV patterns also differ between species and likely influence mate choice and reproductive isolation when they come into contact and hybridize in nature (Gerould 1914). Here, our objective is to examine the potential role of *dsx* and sex-specific gene expression in the development of UV-reflective patterns on wings of the Southern Dogface Butterfly, *Zerene cesonia* (Lepidoptera: Pieridae) (Fenner et al. 2019). This trait is male limited in *Z. cesonia*, where elaborate structures on the scales in narrow regions of fore and hind wings reflect a bright UV pattern that resembles the profile of a “dog’s face”

(fig. 1). Although the UV-reflective patterns vary dramatically between the sexes of *Z. cesonia*, the pigmentation patterns are largely identical, allowing us to leverage differences in sex-specific expression patterns to explore the genetic basis of novel UV reflecting structures that have evolved in a subset of pierid butterflies (Wilts et al. 2011). Using RNA-seq and immunolocalization, we characterize patterns of sex-specific gene expression through male and female wing development, with a particular focus on 24–72 h APF when scales visibly develop and differentiate (Ghiradella 1989; Dinwiddie et al. 2014). This study provides a detailed examination of the genes involved in sexually dimorphic wing development, reveals a novel duplication and mode of sex-specific expression of *dsx*, and highlights the importance of *dsx* in the evolution of novel phenotypes.

## Results

### Sex-Biased Expression during Wing Development

We identified 2,584 transcripts differentially expressed (DE) in the interval from 24 to 72 h after pupal formation (APF) ( $P\text{-adj} < 0.05$ ,  $|\log_2\text{FC}| > 2$ ) with a peak of differential expression at 48 h APF and little overlap of DE transcripts between stages (fig. 2A). Of these, 774 showed male-biased expression and 1,810 female-biased expression (fig. 2B). The distribution of sex-biased genes often varies between autosomes and sex chromosomes (Ellegren and Parsch 2007). In Lepidoptera, the Z chromosome often has an excess of genes with male-



**Fig. 2.** Sexually dimorphic gene expression in *Zerene cesonia* wings. (A) UpSet plot showing overlap between three sets of DE genes (24, 48, and 72 h AFP). The horizontal bars show the number of DE genes at each stage, with 48 h AFP exhibiting the greatest number of DE genes. Vertical bars show the number of unique DE genes for each stage and set of stages. (B) Volcano plots of DE genes at each developmental stage from 24 to 72 h showing the distribution of female and male-biased genes. Note that *dsx* is the most DE gene at 72 h AFP.

biased expression (Huylmans et al. 2017). However, in *Z. cesonia* wing development the numbers of male-biased and female-biased genes on the Z chromosome were not significantly different from expectations based on random distributions (Fisher exact test  $P$  values of 0.72 for females and 0.31 for males). Despite there being no excess of male-biased genes on the Z chromosome, the average expression of Z-linked genes was significantly higher in males compared with females (Wilcoxon test,  $P = 0.03$ ), indicating that dosage compensation is not complete. When compared with autosomal genes, the median expression level of Z-linked genes is 21% lower in males (Z:A ratio = 0.79) and 27% lower in females (Z:A ratio = 0.73). These results suggest that in *Z. cesonia* wing development, the Z chromosome exhibits incomplete dosage

compensation, with overall lower dosage from the Z chromosome compared with autosomes in both sexes but higher expression of Z-linked genes in males relative to females.

### Novel *Dsx* Duplication Shows Sexually Dimorphic Expression

*Dsx* had the most sexually dimorphic expression of all the DE genes during wing disc development ( $\log_2FC = 5.321621$ ,  $P_{adj} = 3.2e-16$ ). Unique *dsx<sup>F</sup>* and *dsx<sup>M</sup>* transcripts were expressed in female and male wings, respectively; similar to what has been previously reported in *Drosophila* and *Bombyx*. De novo transcriptome assembly of individual RNA-seq libraries resulted in *dsx<sup>F</sup>* transcripts for every female and *dsx<sup>M</sup>* transcripts for every male (supplementary fig. S1,



Supplementary Material online). Unexpectedly, the  $dsx^F$  and  $dsx^M$  transcripts mapped to two distinct copies of  $dsx$  in the *Z. cesonia* genome assembly (supplementary fig. S2, Supplementary Material online), suggesting that instead of alternative splicing,  $dsx^F$  and  $dsx^M$  originate from the expression of paralog sequences.

In the *Z. cesonia* genome assembly, there are two tandem annotated copies of  $dsx$  on chromosome 8 (fig. 3A). Duplications of  $dsx$  have never been reported in a Lepidoptera species, thus we first sought to annotate and confirm the duplication was not an artifact of the genome assembly. A full-length  $dsx$  gene (denoted  $dsxA$ ) sequence was annotated that spanned  $\sim 10$  kb and included the six exons and UTR sequences found in  $dsx$  genes of other Lepidoptera as well as *Drosophila* (Baral et al. 2019) (fig. 3B). A partial duplicate of  $dsx$  (denoted  $dsxB$ ) was also found that includes multiple partial duplications of the 5' UTR and exon 1. The first  $\sim 535$  bp of  $dsxB$  contains a near complete duplication of the  $dsxA$  5' UTR and CDS of exon 1 with  $>90\%$  sequence similarity (fig. 3B and Supplementary fig. S3, Supplementary Material online). This region is followed by  $\sim 30$  bp of unknown origin or sequence similarity with  $dsxA$ , and then  $\sim 100$  bp that includes a second duplication of the 5' UTR and exon 1, again with  $>90\%$  sequence similarity to  $dsxA$ . The next  $\sim 550$  bp show no sequence similarity to exons or introns of  $dsxA$  and include a  $\sim 100$ -bp novel transposable element (TE) insertion (DNA Zator Pieris 5178). Following the TE insertion, there is another  $\sim 50$  bp of the duplicated exon 1 that again is  $>90\%$  similar to  $dsxA$  and appears to have been split from the nearest upstream duplicated segment of exon 1, likely due to the TE insertion. Collectively, this annotation suggests  $dsxB$  is the result of multiple duplication events and a recent TE insertion.

The  $dsx$  duplication appears to have occurred recently in the lineage leading to *Zerene* butterflies, followed by a relatively rapid accumulation of amino acid substitutions. We were able to successfully amplify and sequence  $dsxB$  from the sister species *Zerene eurydice*, confirming the duplication occurred before the species split. However, we were unable to confirm the presence of a  $dsx$  duplicate in the sister genus *Colias* using sequence similarity searches in the *Colias eurytheme* or *Colias croceus* genome assemblies. Attempts to amplify  $dsx$  in *C. eurytheme* using primers from conserved sequence regions and  $dsxB$ -specific primers from *Z. cesonia* also failed to amplify more than one  $dsx$  copy. Phylogenetic analyses of the  $dsx$  sequences from *Zerene* sps. ( $dsxA$  and  $dsxB$ ), *Colias eurytheme* ( $dsxA$ ), and *Pieris rapae* ( $dsxA$ ), support that the duplication event likely occurred relatively recently ( $\sim 10.8$  Ma) after the split with *Colias* ( $\sim 25.85$  Ma), but before the split of *Z. cesonia* and *Z. eurydice* ( $\sim 1.76$  Ma), the only two species in the *Zerene* genus (supplementary fig. S6, Supplementary Material online). The  $dsxB$  sequences in *Zerene* show a relatively high number of amino acid substitutions from  $dsxA$  sequences compared with copies of  $dsxA$  in *Z. cesonia*, *Z. eurydice*, or *Colias eurytheme*, which show no amino acid differences between species (supplementary table S1, Supplementary Material online). Correspondingly, the ratio of substitution rates at nonsynonymous and synonymous

sites ( $dN/dS$ ) are 10-fold higher or more between paralogs within *Zerene* species (0.2–0.209), than for  $dsxA$  sequences between species (0.001–0.019), further supporting rapid divergence of the novel  $dsx$  paralog.

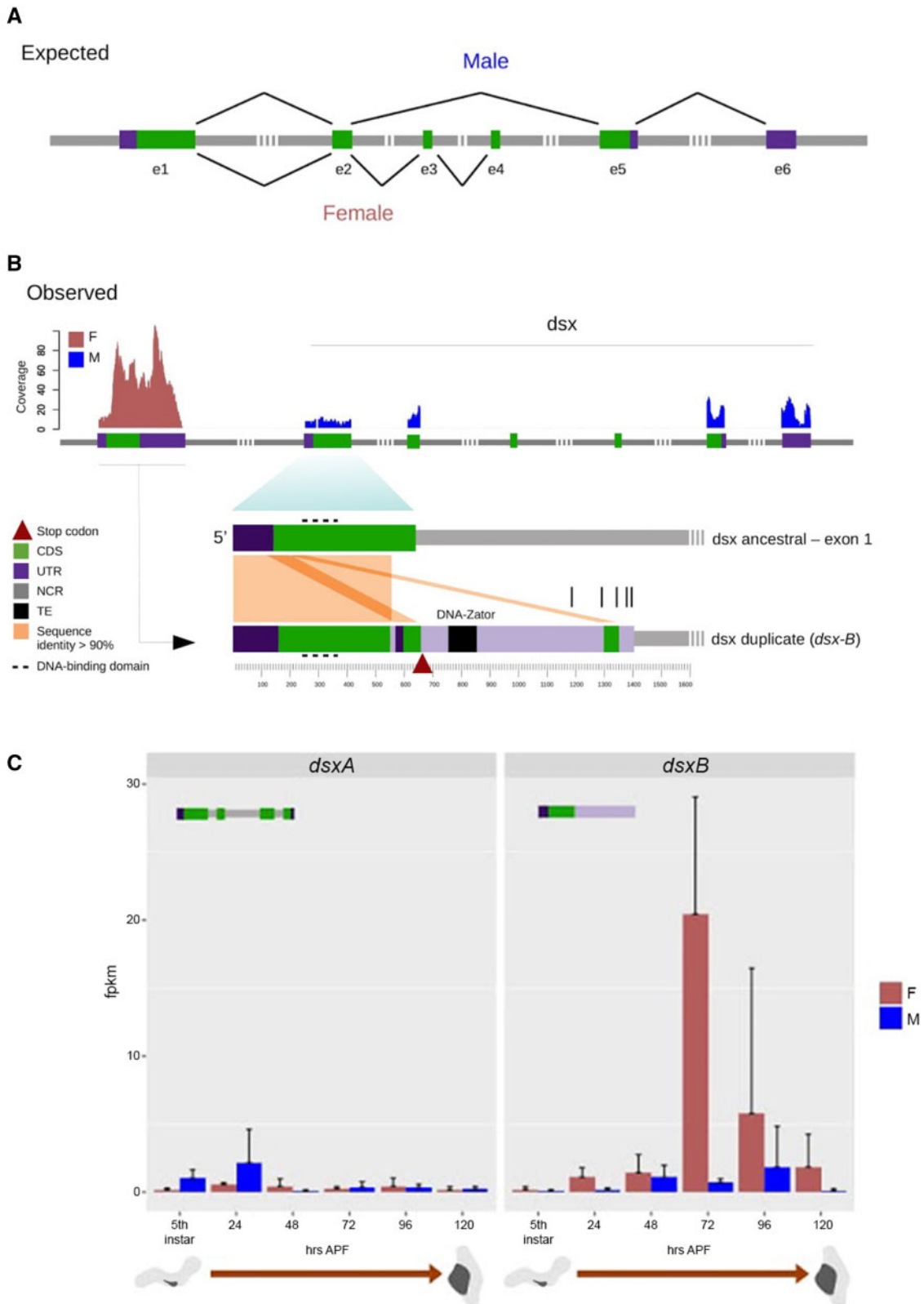
The sequence divergence between the  $dsx$  paralogs allowed RNA sequence reads to be unambiguously mapped to  $dsx$  copies and revealed sexually dimorphic expression of the  $dsx$  paralogs during wing development (supplementary figs. S1–S3, Supplementary Material online). Nearly all RNA-seq reads from females mapped to  $dsxB$ , whereas the male reads primarily mapped to the  $dsxA$  paralog (fig. 3B and C). In addition, among the de novo RNA-seq assemblies every  $dsx$  transcript assembled from a male ( $dsx^M$ ) aligned best to  $dsxA$  and every  $dsx$  transcript from a female ( $dsx^F$ ) aligned best to  $dsxB$  (supplementary fig. S2, Supplementary Material online). Figure 3B shows reads mapped for a single male and female to the  $dsxA$  and  $dsxB$  copies, respectively, illustrating the sex-specific expression of the  $dsx$  paralogs.

Female  $dsxB$  transcripts often extended beyond the stop codon at nucleotide position 660 and included 3' noncoding sequence (fig. 3B). These transcripts include the DM domain characteristic of *dmrt* genes (Raymond et al. 1998) including the DNA-binding domain (fig. 3B), a nuclear localization signal (NLS), and oligomerization domain 1 (OD1) (Verhulst and van de Zande 2015). However, for  $dsxB$  transcripts, we were unable to find evidence of oligomerization domain 2 (OD2) that is involved in cofactor recruitment (Verhulst and van de Zande 2015). This is not surprising as OD2 is typically formed through sex-specific alternative splicing of  $dsx$  exons 3–5 in other insects, but these exons are missing in the *Z. cesonia*  $dsxB$  paralog (fig. 3).

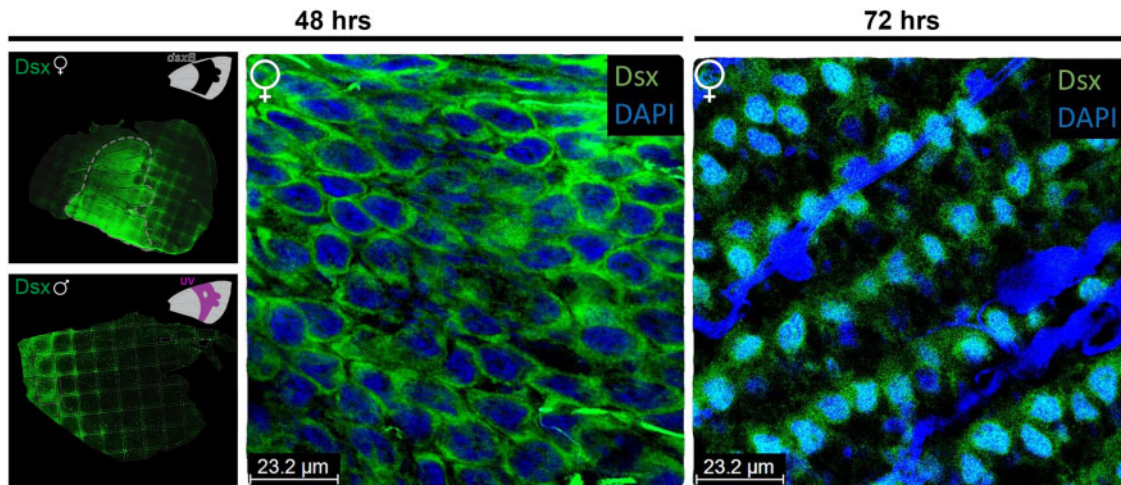
In contrast to findings in the wing, other tissues showed evidence of sex-specific splicing of  $dsxA$ . cDNA sequences from adult and late larval genitalia showed evidence of male- and female-specific isoforms of  $dsxA$ , but no evidence of  $dsxB$  expression (supplementary fig. S4, Supplementary Material online). In all male tissues examined (pupal wings, adult head, adult thorax, adult, and larval testes), a  $dsx$  isoform was found that included  $dsxA$  exons 1, 2, 5, and 6 (supplementary fig. S4, Supplementary Material online). However, in female tissues (adult head, adult thorax, adult, and larval ovaries), the  $dsx$  isoforms included  $dsxA$  exons 1–6 and showed no evidence of exon skipping (supplementary fig. S4, Supplementary Material online). Interestingly, of all tissues examined in males and females, we only found  $dsxB$  sequences in developing female wings.

### Novel Dsx Paralog Involved in Female-Specific Pattern Development

The novel  $dsx$  paralog ( $dsxB$ ) expression in females appears to be associated with wing color pattern elements. Immunostaining for Dsx in developing pupal wings revealed a strong signal across the central region of the wing that corresponds to the region of yellow scales responsible for generating the so-called “dogface” profile (fig. 4). This result is consistent with the RNAseq data that suggest a burst of  $dsxB$  upregulation in 72 h APF female wings (fig. 2B). In contrast, males did not show expression of Dsx in this central



**FIG. 3.** *Dsx* duplication and expression in *Zerene cesonia*. (A) Splicing patterns of *dsxF* and *dsxM*, observed in most arthropods. (B) Observed *dsx* duplication and gene annotations. An incomplete tandem duplication of the *dsx* gene was found in the genome of *Z. cesonia*. The barplots on top show read coverage of *dsx* paralogs from a male (blue) and a female (red) sample, after removing multimapping reads. The bottom diagram shows a sequence alignment of *dsx* paralogs and a detailed annotation of the novel paralog (*dsxB*) including the predicted UTRs, stop codon and coding region based on the longest open-reading frame. In this diagram, vertical bars on top of *dsxB* indicate the different lengths of independently assembled *dsxB* transcripts from single-library assemblies. (C) Quantification of *dsxA* and *dsxB* expression through wing development from RNA-seq data. Error bars show 95% confidence intervals calculated from biological replicates.



**Fig. 4.** Dsx spatial expression during *Zerene cesonia* wing development. Dsx immunofluorescent stains in female (upper) and male (lower) forewings at 48 h APF. Female panel shows Dsx associated with the central pattern region of the wing (dotted white line), where UV-reflecting scales develop in males (shown in wing schematic in upper corner). Higher magnification shows Dsx in cytosol at 48 h APF (middle panel) in female ground and cover scales. At 72 h APF (right panel), Dsx is detected in the nucleus of alternating rows of scales on female wings.

wing region where the UV-reflecting scales are located. Because in-situ hybridization ([supplementary fig. S7, Supplementary Material](#) online) and RNA-seq data showed evidence of *dsxA* in male and not female wings, we can infer that these antibody stains in females reflect the spatial expression of *dsxB*. Therefore, we suggest that the novel *dsx* paralog, *dsxB*, has an expression pattern associated with the development of yellow, non-UV-reflecting scales in the central pattern of female wings.

A closer look at the scale cells of female wings revealed a coordinated change of *dsxB* expression in the cytosol and nucleus between 48 and 72 h APF ([fig. 4](#)). At 48 h APF on female wings, Dsx appeared to be almost exclusively in the cytosol in ground and cover scales. However, at 72 h APF Dsx was diffusely present in both the cytosol and the nuclei. Notable is that at 72 h Dsx was present in alternating rows of scales, suggesting it is expressed in only ground or cover scales, but not both, which contrasts with 48 h APF where rows of both ground and cover scales showed *dsxB* ([fig. 4](#)). These expression patterns are intriguing, but leaves us with new questions, such as what is the possible function of the sharp differences of Dsx in the cytosol and nucleus from 48 to 72 h? Is Dsx in the female scale nuclei at 72 h APF the result of active transport from the cytosol to nucleus? A nuclear localization sequence (NLS) is present in the *dsxB* paralog, suggesting transport is possible. Although the impact of these Dsx spatial and temporal patterns remains an enigma, the timing of the changes and spatial coordination among alternating rows of different scale types (i.e., ground vs. cover) is indicative of *dsxB* playing an important signaling role in scale differentiation.

#### Sex-Biased Expression Enriched for Genes Involved in Cytoskeletal Structure

At 48–72 h APF, morphological differences in scale structure become clearly visible ([Ghiradella 1989](#); [Fenner et al. 2019](#)). Tests for enrichment of Gene Ontology (GO) terms among

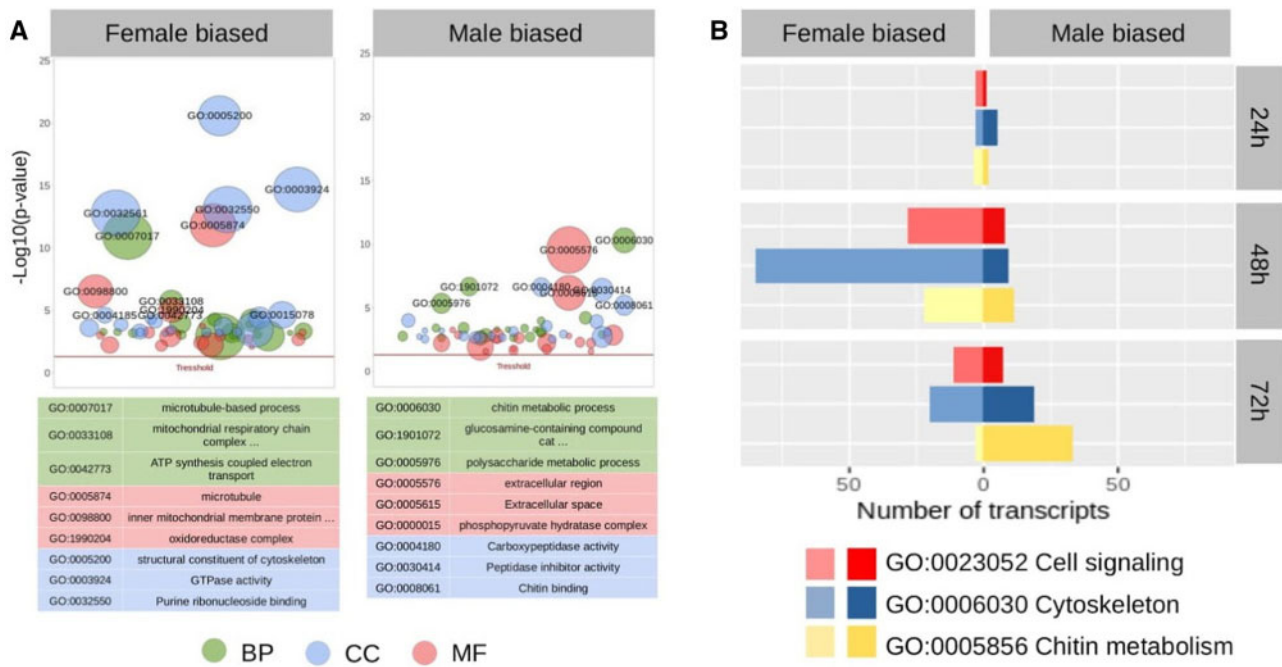
the DE genes between males and females from 48 to 72 h APF identified several biological functions related to cytoskeletal development, chitin metabolism, and cell signaling ([fig. 5](#)). In female-biased genes, an excess of these processes was most pronounced at 48 h ([fig. 5B](#)) and included several microtubule (MT)-related genes ([fig. 5B](#)). This contrasted with the male-biased genes that did not show an excess of MT-related genes, but rather showed an excess of genes involved in chitin metabolism and cytoskeletal-related processes ([fig. 5](#)). The excess of these genes was greatest at 72 h APF ([fig. 5B](#)), which corresponds to when UV-reflecting scale structures become visible during wing development.

To determine if these sex differences in GO results reflect visible differences during scale development, we stained for MTs in developing male and female wings ([fig. 6](#) and [supplementary fig. S7, Supplementary Material](#) online). As predicted, there were clear, large differences in  $\alpha$ -tubulin expression between males ( $n = 5$ ) and females ( $n = 5$ ) at 48 h APF ([fig. 6](#)). Most notable was the lack of  $\alpha$ -tubulin in the UV reflective region of male wings. In accordance with GO enrichment results, females showed high densities of  $\alpha$ -tubulin across the wing ([fig. 6](#) and [supplementary fig. S8, Supplementary Material](#) online). As wing development progressed, there was little difference in MTs among males and females, and by 120 h APF, MT organization clearly reflected differences in scale structure between yellow and black scales in both sexes ([supplementary fig. S8, Supplementary Material](#) online). Together these results suggest the differential expression between males and females was enriched for genes involved in the sexually dimorphic development of the UV-reflecting scale structures.

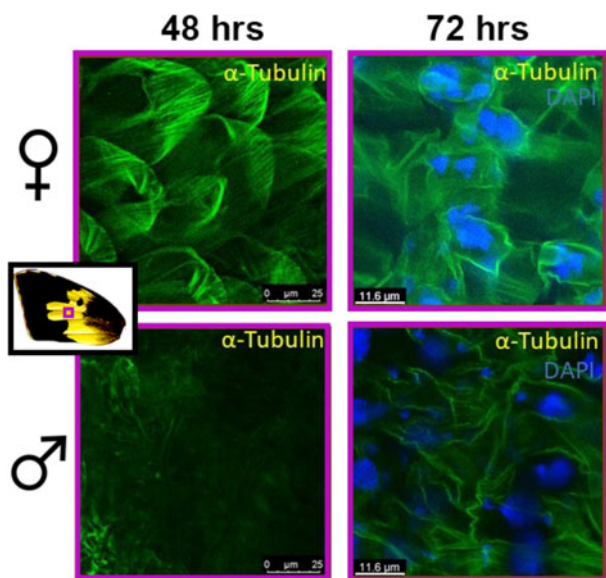
#### Male-Biased Genes Are Enriched for Dsx-Binding Motif

To test if *dsxB* could be acting as an activator and/or repressor of sex-specific gene expression, we used a position weight matrix of *Drosophila* *dsx*-binding motifs ([Clough et al. 2014](#))





**Fig. 5.** Sexually dimorphic gene expression in *Zerene cesonia* wings. (A) Bubble graph results of GO enrichment analyses for female and male-biased genes from 24 to 72 h APF (red line denotes  $P > 0.05$ ); BP = Biological process, CC = Cellular component, MF = Molecular function. (B) Comparison of transcriptome activity related to cell signaling, chitin metabolism, and cytoskeleton development among males and females. Bars reflect the number of annotated transcripts expressed at each developmental stage (FPKM > 1).



**Fig. 6.** MT distributions on male and female wings (48 and 72 h APF). MTs show greater density in female scales and MTs are largely absent in developing UV-reflecting scales of males at 48 h APF.

to test for *dsxB*-binding sites in promoter regions (2 kb 5' and 3' of the coding sequence) of female ( $n = 497$ ) and male-biased genes ( $n = 289$ ) at 72 h APF. At 72 h APF, *dsxA* and *dsxB* showed little or no evidence of expression in males, however, this is when *dsxB* is most highly expressed in females. Interestingly, we found a significant enrichment of *dsxB*-binding motifs among the male-biased genes (Fisher's exact test,  $P = 0.0219$ ), but not the female-biased genes (Fisher's exact

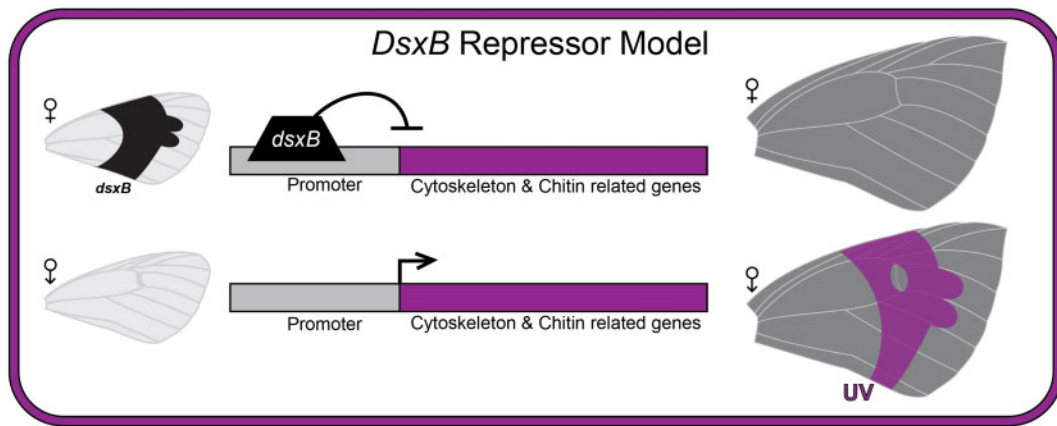
test,  $P = 0.356$ ) (supplementary fig. S9, Supplementary Material online). These results suggest *dsxB* may be repressing genes in females that are highly expressed in males (i.e., male-biased genes).

Unfortunately, attempts to study the functions of *dsxB* in male and female wing development using CRISPR-based editing were not informative, as none of >450 injected individuals survived beyond third larval instar, preventing the examination of wing development. A 1.7-fold increase in hatching was observed by reducing the injected gRNA by half, suggesting *dsxB* edits impact survival, thereby limiting the applicability of CRISPR mosaic knock-out methods to study *dsxB* function. Similarly, a recent study in *Bicyclus* butterflies found a relatively low success rate in using CRISPR to edit and study *dsxB* function in developing wings (Prakash and Monteiro 2020).

## Discussion

### A *Dsx* Repressor Model of UV Scale Development

We propose a genetic model for the regulation of UV pattern that involves the novel *dsxB* paralog, *dsxB*, acting as a repressor of genes required for the development of UV-reflecting scales (fig. 7). At 72 h APF, *dsxB* is highly expressed and present in the nucleus of scales in females, but not males (figs. 2–4). Also at 72 h APF, there is a significant excess of cytoskeletal and chitin-related genes ( $n = 52$ ) that are expressed in males and downregulated in females (fig. 5B). We suggest these are candidate genes involved in the development of the chitinous male-specific UV-reflecting scale structures that occur at 72 h APF (Supplementary table S2, Supplementary Material online).



**Fig. 7.** Hypothetical repressor model of *dsxB*. Model shows *dsxB* binding to promoter regions and repressing expression of cytoskeleton and chitin related genes in female pupal wing development. Wings on the left illustrate *dsxB* spatial expression during pupal wing development. Wings on the right show UV patterns in adult wings. Based on this model, *dsxB* expression in females represses the development of UV-reflecting scales, and the lack of *dsxB* in males results in the development of UV-reflecting scales.

We found that these male-biased genes at 72 h APF also show a significant enrichment of *dsxB* binding sites in their promoter regions. Therefore, it is plausible that *dsxB* binds these promoter regions and represses these genes' expression in females at 72 h APF. If *dsxB* can act as a repressor, we may also predict that the female-biased genes at 72 h APF should not be enriched for *dsxB*-binding sites, or *dsxB* could have repressed their expression. Accordingly, there was no significant enrichment of *dsxB*-binding motifs in the female-biased genes at 72 h APF. The lack of a OD2 protein-binding domain in *dsxB* also supports that it more likely functions as a repressor, than activator of transcription. Without a protein-binding domain to recruit, it is unlikely for *dsxB* to act as an activator that binds other components of the transcription factor complex (i.e., cofactors). In our model, it is unlikely that *dsxB* is repressing transcription by competing with or blocking *dsxA* function, as *dsxA* shows no or little expression in females and males at 72 h APF. Instead, promoter-bound *dsxB* may interfere with regulating the battery of effector genes that mediate the differentiation of male-specific features, including the peculiar UV-reflective scales. Further experiments that assay *dsxB* binding or function, such as ChIP-seq and CRISPR-based gene editing are needed to confirm how *dsxB* and the male-biased genes are involved in UV scale development.

Interestingly, *dsxB* genes in several ant and termite species also lack the protein-binding domain (Jia et al 2018; Miyazaki et al. 2021). In several ant species, sex-specific alternative splicing of *dsxB* occurs as in other insects, despite losing the OD2 protein-binding domain. There appears to be sex- and caste-specific exon usages of ant *dsxB* genes, without the OD2 domain. There is also evidence of rapid, positive selection on *dsxB* genes in ant lineages that lost the OD2 domain. In these ants, the sex-specific *dsxB* isoforms lacking OD2 may have evolved ways to regulate sexually dimorphic development similar to *dsxB* in *Z. cesonia* that also lacks OD2. In termites, there appears to have been an ancestral loss of the *dsxB* OD2 associated with evolution of eusociality (Miyazaki et al 2021).

In these termites, instead of sex-specific alternative splicing, *dsxB* appears to only be expressed in males. These systems offer a comparative approach to studying how alternative isoforms versus gene duplications that lack a protein-binding domain may impact the regulatory evolution of sexual dimorphism.

### An Important Role of MTs in Scale Cytoskeletal Structures and Growth

Like bristles on *Drosophila* wings, butterfly scales can be an excellent model system to study the role of the cytoskeleton during polarized cell growth (Wulfkuhle et al. 1998; Bitan et al. 2012; Dinwiddie et al. 2014). In *Drosophila* bristle development, MTs densities are high on the leading tip, which we also observed in the *Z. cesonia* wing scales. We also see MTs distributed laterally, parallel to where chitinous ridges develop on the upper surface. These ridges are the structures responsible for UV reflection, which are thought to “self-assemble” as chitin is deposited in pulses on the scale surface (Ghiradella 1994). In *Drosophila* bristles, MTs are important for trafficking vesicles of construction materials needed for chitin metabolism (Bitan et al. 2010). In butterfly scales, actin filaments and MTs may provide the tracks for the transport of materials from cellular organelles to sites of chitin deposition. The change in MTs in *Z. cesonia* scales from 48 to 72 h APF in the UV-reflecting region of male wings may be involved in regulating the pulsed deposition of chitin that generates the UV-reflecting structures. Interestingly on female wings, MTs were densely distributed and showed little differences across the wing, which corresponds with females lacking the UV-reflecting structures and in general have less differentiation in scale structure. An important role of MTs in scale biogenesis is also supported in Monarch butterflies, where a MT-related gene was proposed to underlie a depigmented state in a naturally occurring mutant (Kronforst 2015). Further investigations of MTs in butterfly scales may be a promising approach to better understanding the cellular processes that regulate variation in cell growth.



## Shared Genetic Mechanisms of Dosage Compensation and Sex-Specific Expression in Lepidoptera

In the silk moth, *Bombyx mori*, sex-specific splicing of *dsx* and dosage compensation appear to be controlled by a shared suite of genes, *BmMasc*, *BmPSI*, *BmIMP*, and *BmZnf2* (Gopinath et al. 2016; Katsuma et al. 2019; Kiuchi et al. 2019). The masculinization gene, *BmMasc*, is regulated by a piwi-interacting RNA that originates from the W chromosome, and acts as the genetic sex determination switch (Kiuchi et al. 2019). When *BmMasc* is inhibited, it impacts dosage compensation and sex-specific splicing of *dsx* (Kiuchi et al. 2019). *BmZnf2* appears to interact with *BmMasc*, *BmPSI*, and *BmIMP*, and when inhibited impacts both dosage compensation and *dsx* splicing (Gopinath et al. 2016). Interestingly, in *B. mori*, *dsx<sup>F</sup>* isoforms include all exons, and it's only *dsx<sup>M</sup>* that experiences exon skipping when spliced. Therefore, there is no special alternative splicing machinery needed for *dsx<sup>F</sup>* to be transcribed, and when *BmMasc* or *BmZnf2* are inhibited, *dsx<sup>F</sup>* isoforms are still produced (Gopinath et al. 2016; Kiuchi et al. 2019). We were able to identify homologs for *BmMasc*, *BmPSI*, *BmIMP*, and *BmZnf2* in the *Z. cesonia* genome (XP\_037869585, XP\_012545114, XP\_021206051, XP\_004924549). However, we were only able to detect very low levels of expression for *BmPSI*, *BmIMP*, and *BmZnf2*, and we found no RNA-seq reads that mapped to *BmMasc*. This perhaps is not surprising if they are only involved in *dsx<sup>M</sup>* alternative splicing, as *dsx* was very lowly expressed in male wings.

We did find clear evidence of incomplete dosage compensation that suggested expression of Z-linked genes is actively being down-regulated in males of *Z. cesonia*. Down-regulation of Z-linked genes in males is consistent with findings in *B. mori* and *Heliconius* butterflies (Walters and Hardcastle 2011; Walters et al. 2015). In *B. mori*, male down-regulation of the Z was confirmed through RNAi knockdown of *BmMasc*, which led to a global increase in expression of Z-linked genes (Kiuchi et al. 2019). Despite similar observations of Z chromosome down-regulation in *Z. cesonia* males, there was no evidence of *masc* expression, suggesting an alternative molecular mechanism is likely involved. Due to the lack of *masc* expression, the low level of *dsx<sup>M</sup>* expression, and lack of alternative splicing for female *dsx* isoforms in *Z. cesonia* wings, we are unable to address if the same genes may regulate both dosage compensation and *dsx* splicing, as has been found in *B. mori*. However, it would be very interesting to explore this in *Z. cesonia* with the incomplete dosage compensation and potential involvement of the novel *dsx* duplication.

### *Dsx* and the Evolution of Sexual Dimorphism

Ultimately, some of the most remarkable morphological variation of traits in nature lies in the differences between males and females. Our understanding of the genetic basis of sexually dimorphic traits in animals has been revolutionized in the past few decades with advances in our understanding of *Dmrt* genes (Yang et al. 2008; Kopp 2012). For example, studies of the iconic horns on rhinoceros beetles and the sex combs on fruit fly appendages have shown *dsx*, the insect homolog of *Dmrt1*, is essential for the development of these sexually

dimorphic traits (Kijimoto et al. 2012; Ito et al. 2013; Morita et al. 2019; Rice et al. 2019). Our findings suggest the DNA-binding motif of *dsx* may be very similar in Lepidoptera and other insects, yet *dsx* may regulate sexually dimorphic development through novel molecular mechanisms in *Zerene*. The well-characterized alternative splicing of male and female isoforms (*dsx<sup>M</sup>* and *dsx<sup>F</sup>*) that drives sexual dimorphic development in insects appears to have been replaced in wing development by sexually dimorphic expression of recently evolved paralogs. These paralogs would be released from the sexually antagonistic mutations that can be difficult to overcome when exons are shared in alternative splicing, like the *dsx<sup>M</sup>* and *dsx<sup>F</sup>* isoforms found across most insects. This would allow for the rapid amino acid evolution we find between the *dsx* paralogs in *Z. cesonia*. In addition, it opens opportunity for rapid diversification of regulatory modules that could independently control sex-specific spatial expression of the *dsx* paralogs. This is very different from *Drosophila* and other insects, where genes such as *sex lethal* and *transformer* are required for the sex-specific alternative splicing, and *cis*-regulatory modules are responsible for tissue-specific expression patterns (Rice et al. 2019). We hypothesize that the duplication of *dsx* would promote the evolution of not only tissue-specific, but also sex-specific regulatory modules. Duplications of *Dmrt* genes have been reported in diverse organisms ranging from vertebrates (Ogita et al. 2020) to hexapods (Price et al. 2015; Panara et al. 2019; Gruzin et al. 2020) and may be a key step in the evolution of complex sexually dimorphic traits. In this regard, our findings of the *dsx* paralogs highlight *Z. cesonia* as a potentially powerful system to study the evolution of molecular mechanisms that regulate sexually dimorphic development.

## Materials and Methods

### Experimental Design

To characterize sex-specific patterns of gene expression, mRNA sequencing was performed across seven developmental stages (late fifth instar larvae, 24, 48, 72, 96, 120, 148 h APF) for replicate males and females (42 samples total; minimum three replicates per sex, for each stage). To identify sex-specific expression patterns involved in sexually dimorphic development of UV-reflecting scales, analyses were focused on 24–72 h APF when scales differentiate and the UV-reflecting structures develop (Ghiradella 1989; Dinwiddie et al. 2014). Copies of the *dsx* gene were further examined through detailed genomic annotations, in situ hybridizations and antibody immunofluorescence. Differential expression of MTs in developing scales in males and females was also examined using antibody immunofluorescence. Lastly, potential targets of *dsx* were identified by performing DNA-binding motif enrichment analyses among subsets of genes with sex-biased expression. Collectively, these approaches provide a detailed view of sex-specific patterns of gene expression during scale development, with particular focus on the involvement of *dsx*.

### Preparation for RNA Sequencing

Butterflies were reared in controlled conditions (Shelby et al. 2020), and wing discs were dissected and preserved in RNAlater. Dissection timing varied within 4 h among samples within the sampled developmental stages.

### Transcriptome Sequencing and Assemblies

RNA-seq libraries were prepared from dissected forewings from individual butterflies using Nextera XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) and sequenced at 150-bp paired ends reads (Fenner et al. 2020). RNA-seq reads were preprocessed using AfterQC (Chen et al. 2017) and Trimmomatic (Bolger et al. 2014).

Two transcriptome assemblies were generated de novo using the evidential gene (EG) pipeline (<http://arthropods.eugeneg.org/EvidentialGene/>; last accessed April 3, 2020) and the Oyster River protocol (MacManes 2018). The quality of the resulting transcriptomes was estimated using QAST scores (Gurevich et al. 2013), BUSCO completeness (Seppey et al. 2019), BLASTX (Altschul et al. 1990) hits against a reference lepidopteran protein library, and mapping quality support using Bowtie2 (Langmead and Salzberg 2012). The assembly with the best quality (EG) was manually curated, removing transcripts smaller than 500 bp and redundant transcripts using CD-HIT-EST (Li and Godzik 2006) with a similarity threshold of 95%. The final transcriptome was annotated using Trinotate following the procedure in Bryant et al. (2017) using the Trinotate pipeline with the uniprot/uniref 3.1.1 database for protein sequences references, and the PFAM 3.1 database for protein domain references. Scripts and reference files are accessible at the GitHub repository for this project ([https://github.com/LF-Rodriguez/Zerene\\_cesonia\\_WingDiscDevelopment](https://github.com/LF-Rodriguez/Zerene_cesonia_WingDiscDevelopment); last access April 3, 2020).

### Transcriptome Analysis of Wing Development

RNA-seq reads were used to estimate gene expression during wing scale development and identify gene networks associated with specific stages of wing scale development. First, reads were mapped to a reference transcriptome using Bowtie2 and gene expression was estimated using RSEM (Li and Dewey 2011) with default parameters. The R package WGCNA (Langfelder and Horvath 2008) was used to perform gene network analysis and measure the associations with specific developmental stages. *Dsx* transcripts and reads were also mapped to the reference genome of *Z. cesonia* (Rodriguez-Caro et al. 2020) to manually annotate *dsx* paralogs and examine read coverage of *dsx* exons.

### Analyses of Sexually Dimorphic Expression

Differences in scale morphology between sexes are evident by 72 h post pupation. We performed differential expression tests between sexes with the RNA-seq data sampled from at three stages of scale development: 24, 48, and 72 h APF. Differential expression analyses were performed using the R package DESeq2 (Love et al. 2014). Transcripts were considered DE between sexes with a *P*-adjusted value lower than 0.05 and log<sub>2</sub>-fold change higher than 2. Genes with male-biased and female-biased expression were identified as those

that were significantly DE between sexes and higher in males or females, respectively. The R package TopGO (Rahnenfuhrer 2019) was used to test for GO enrichment on the set of DE genes using a customized GO annotation.

To test for dosage compensation on the Z chromosome during wing development, gene expression levels were compared between autosomal genes and Z-linked genes separately in males and females using normalized read counts (FPKM). For this analysis, lowly expressed genes (FPKM < 1) were removed, as well as genes that showed sex-biased expression during wing development following Huylmans et al. (2017). Differences in median expression between autosomes and Z chromosomes were detected using a *t*-test, and differences in gene expression levels in Z-linked genes between males and females were detected using a Wilcoxon test.

### Tissue-Specific Expression of *Dsx*

To examine *dsxA* and *dsxB* expression in tissues besides pupal wings, we used cDNA sequencing and explored available RNA-seq datasets. We dissected the following tissues and extracted RNA for 3–6 males and females: adult head, adult thorax, adult abdomen, adult gonads and fifth instar larval gonadal masses. mRNA was extracted with a Qiagen Total RNA isolation kit, and cDNA was synthesized using random primers and reverse transcriptase. Primers were designed from *dsxA* and *dsxB* transcript and reference genome sequence alignments (*dsxA* exon1: P5-ATAGAGCTCAAGGGCCACAA, P2-GGGACCGAATCCACTGAAGT; *dsxB* exon1: P5-ATAGAGCTCAAGGGCCACAA, P8-AAGCCGAATCCACTGAAGT; *dsxA* exons2-6: F3-TCGAGAACTGTCACAAGCTG, R2-AAACAGGCGTAACGAAGTTG). PCR products were amplified, cloned, and sequenced. Sequences were aligned to *dsxA* and *dsxB* transcript and reference genome sequence to characterize *dsx* isoforms present in each sample. We also used available RNA-seq data from male head and thorax previously used to annotate the reference genome (Rodriguez-Caro et al. 2020), which we mapped to the reference transcriptome to determine isoform composition. To determine which *dsx* paralog was present in each RNA-seq and cDNA sequence, we used the substitution differences between *dsxA* exon 1 and *dsxB* shown in supplementary figure S3, Supplementary Material online. In addition, as a positive control to confirm the PCR primers could successfully amplify both *dsxA* and *dsxB*, we conducted PCRs using cDNA from a subset of male and female pupal wing RNA-seq samples that we confirmed had *dsxA* and *dsxB* transcripts present, respectively. These positive controls allowed us to infer that a lack of *dsxB* sequences in any tissue samples reflected a lack of *dsxB* transcripts, rather than an inability to PCR amplify *dsxB*.

### *Dsx* Annotation and Evolutionary Analyses

To annotate exon boundaries in *dsx* transcripts, we aligned each of our *dsx* transcripts to reference *dsx* exons from *Bombyx mori* (Zheng et al. 2019) using MAFFT (Katoh et al. 2002). We then mapped each *dsx* exon to the genome assembly of *Z. cesonia* (Rodriguez-Caro et al. 2020) using Gmap

(Wu and Watanabe 2005). MUMmer (Marçais et al. 2018) was used to align the scaffold containing *dsx* against itself to identify duplications and estimate sequence conservation in the duplicates. TEs in the *dsx* gene were identified using RepeatMasker (<http://www.repeatmasker.org>). The NLS was identified in *dsxB* using NLSdb (<https://roslab.org/services/nlsdb/>) that compares query sequences to a database of known NLS sequences. *Zerene eurydice dsxB* amplicons and sequences were obtained using the primers (FP-TGCATTGTGCACAGCAATAA and RP-AGTTCGGCGCTAGTTTGTGT). *Dsx* exon sequences for *Pieris rapae* (accession GCA\_001856805.1) were downloaded from NCBI and *Colias eurytheme* sequences were acquired from a draft genome assembly provided by A. Martin and Chris Wheat. *Dsx* exon1 alignments were generated using the Muscle alignment tool and maximum-likelihood phylogenies were generated using a Tamura-Nei model with uniform rates and 1,000 replicate bootstraps to assess support, as implemented in MegaX Version 10.1.8 (Kumar et al. 2018). Trees were created from Newick files using the Ape package in R (Paradis 2012). Divergence time estimates for the *Colias/Zerene* and *Z. cesonia/Z. eurydice* lineage splits were taken from Zhang et al. (2019). Using Zhang et al. (2019) divergence time estimates for the *Colias/Zerene* (28.85 My) and *Z. cesonia/Z. eurydice* (1.76 My) lineage splits, and calibrated branch lengths from the nuclear genome tree of Zhang et al. (2019), we estimated a *Colias-Zerene* nuclear substitution rate of 0.0046 substitutions/my. We used this substitution rate and the number of substitutions between *Z. cesonia dsxA* and *dsxB* to estimate the divergence time for the paralogs (e.g., age of duplication event).

### Dsx-Binding Site Analyses

mRNA transcripts that correspond to genes exhibiting significant male-biased expression on 72 h APF were used to test for enrichment of the *dsx* DNA-binding motif. Aligned genomic intervals were used to extract 2 kb of sequence up- and downstream to use as input. These extracted sequences were tested for enrichment of the *Drosophila melanogaster dsx* DNA-binding motif position weight matrix (Clough et al. 2014) using AME (<https://meme-suite.org/meme/tools/ame>) from MEME Suite 5.3.1, via the available web server. Fisher's exact test was conducted alongside a set of shuffled input sequences with preserved sequence size of 1 bp. The same method was applied to 72 h APF female-biased transcripts.

### Immunostains and In Situ Hybridizations

To corroborate our findings from DE tests, and characterize the spatial expression patterns of MTs, we immunostained developing pupal wings with the Primary Monoclonal Anti- $\alpha$ -Tubulin antibody (catalogue #T6074, Sigma Aldrich) (1:500 dilution) and Alexa 488 Anti-Mouse IgG2a ( $\gamma$ 2a) secondary antibodies (catalogue #SAB4600239, Sigma Aldrich) (1:1000 dilution) for male and female forewings. Doublesex immunostains were performed using primary antibodies DsxDBD (DSHB) at a dilution of 1:50, and Alexa 488 Goat anti-mouse secondary antibody (ThermoFisher #A28175, 1:250 dilution).

The time of pupation was recorded, and wings were dissected at 48, 72, 96 and 120 h post pupation, with a minimum of five replicates for each sex and stage. Dissection timing varied within less than 2 h across samples within the sampled developmental stages. Nuclear counterstaining were performed by mounting wings in a solution of Fluoroshield with DAPI (catalogue #F6057, Sigma Aldrich). A Leica SPE confocal microscope was used to collect fluorescent images of antibody stains.

We focused on characterizing spatiotemporal expression patterns of *dsx* transcripts and isoforms in developing wings using in situ hybridization. A fragment of *dsx* mRNA was amplified from *Z. cesonia* cDNA using primers specific for *dsxA* transcripts (FP-GTGCCTCCAGCTCCGC, RP-GTTCAGGATCACGAGCACGA). The amplicon with T7/SP6 overhangs were used as a template to synthesize DIG-labeled RNA probes. Fore and hind wing discs were collected from male and females at 72 h APF and used for in situ hybridization using a preestablished protocol (Martin and Reed 2014).

### CRISPR/Cas9 Editing

CRISPR/cas9 injections for *dsx* knockouts were attempted on a total of 450 individuals but resulted in lethality of all injected individuals before the fourth larval instar. Two *dsx* single-guide RNAs, *dsx1*: CCCGGTGTATCGTAAATGATAGT and *dsx2*: CCAACATATGATCCACTTTATTG, were designed and purchased from Integrated DNA Technologies as Alt-R crRNA in the Alt-R CRISPR-cas9 system. Embryo injections used borosilicate glass needles, and eggs were affixed to glass slides with double-sided sticky tape. Injections occurred between 3 and 5 h after eggs were laid. Fluorescein dye at 2.5 mg/ml was added to the injection mixture to ensure embryos were equally and evenly injected. After injections, eggs were kept in a petri dish in a rearing chamber with 50% humidity, 16:8 h photoperiod, and 27 °C temperature. After 3 days, host plant cuttings were laid across the slide, allowing larvae to immediately climb onto host plants once hatching occurred on day four of embryo development. Larvae were reared on an unlimited supply of host plant *Dalea purpurea*. Two experiment concentrations of cas9 and guide RNAs were used among the 450 individuals: 187 eggs were injected with 1  $\mu$ g/ $\mu$ l cas9 and 222 ng/ $\mu$ l guide RNAs with a hatching success rate of 14% and 263 eggs were injected with 0.5  $\mu$ g/ $\mu$ l cas9 and 111 ng/ $\mu$ l guide RNAs with a hatching success rate of 17%. Larval growth time was increased by more than double normal developmental times (Shelby et al. 2020) and all larvae died before reaching the fourth larval instar.

### Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

### Acknowledgments

The DsxDBD developed by Baker, B.S. was obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of



Iowa, Department of Biology, Iowa City, Iowa. This work was supported by National Science Foundation awards 1755329 to B.A.C., R.R., and A.M., and 1736026 to B.A.C., R.R., and R.P.

## Author Contributions

F.R.C., J.F., R.R., A.M., and B.A.C. conceived and designed the study. J.C. and J.F. constructed RNAseq libraries. R.P. provided RNA sequencing and assistance with analyses and interpretation. MB provided computational resources and assistance with analyses and interpretation. S.B. conducted in situ hybridizations, immunofluorescent staining, and microscopy. A.M.C., S.B., and F.R.C. conducted gene expression analyses. C.B. performed molecular evolution and binding site enrichment analyses. F.R.C. and B.A.C. wrote the manuscript with assistance from S.B., J.F., A.M., R.R., and C.B. All authors provided input on manuscript drafts.

## Data Availability

The annotated *Zerene cesonia* genome assembly 1.1 and RNA-seq data are accessible through the NCBI Genome Data Viewer (ID 88871) and Short Read Archive (SRA) accession numbers SRX8624474–SRX8624487 (male and female wing discs) and SRX7098269–SRX7098276 (head and thorax). Scripts and reference files are accessible at the GitHub repository for this project ([https://github.com/LF-Rodriguez/Zerene\\_cesonia\\_WingDiscDevelopment](https://github.com/LF-Rodriguez/Zerene_cesonia_WingDiscDevelopment)).

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410.
- Baral S, Arumugam G, Deshmukh R, Kunte K. 2019. Genetic architecture and sex-specific selection govern modular, male-biased evolution of doublesex. *Sci Adv.* 5(5):eaau3753.
- Bitan A, Guild GM, Bar-Dubin D, Abdu U. 2010. Asymmetric microtubule function is an essential requirement for polarized organization of the *Drosophila* bristle. *Mol Cell Biol.* 30(2):496–507.
- Bitan A, Rosenbaum I, Abdu U. 2012. Stable and dynamic microtubules coordinately determine and maintain *Drosophila* bristle shape. *Development* 139(11):1987–1996.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.
- Bryant DM, Johnson K, DiTommaso T, Tickle T, Couger MB, Payzin-Dogru D, Lee TJ, Leigh ND, Kuo T-H, Davis FG, et al. 2017. A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell Rep.* 18(3):762–776.
- Burtis KC, Baker BS. 1989. *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* 56(6):997–1010.
- Chen S, Huang T, Zhou Y, Han Y, Xu M, Gu J. 2017. AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. *BMC Bioinformatics.* 18(Suppl 3):80.
- Clough E, Jimenez E, Kim Y-A, Whitworth C, Neville MC, Hempel LU, Pavlou HJ, Chen Z-X, Sturgill D, Dale RK, et al. 2014. Sex- and tissue-specific functions of *Drosophila* doublesex transcription factor target genes. *Dev Cell.* 31(6):761–773.
- Coyne JA, Kay EH, Pruett-Jones S. 2008. The genetic basis of sexual dimorphism in birds. *Evolution* 62(1):214–219.
- Deshmukh R, Lakhe D, Kunte K. 2020. Tissue-specific developmental regulation and isoform usage underlie the role of doublesex in sex differentiation and mimicry in *Papilio swallowtails*. *R Soc Open Sci.* 7(9):200792.
- Dinwiddie A, Null R, Pizzano M, Chuong L, Leigh Krup A, Ee Tan H, Patel NH. 2014. Dynamics of F-actin prefigure the structure of butterfly wing scales. *Dev Biol.* 392(2):404–418.
- Ellegren H, Parsch J. 2007. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet.* 8(9):689–698.
- Emlen DJ. 2001. Costs and the diversification of exaggerated animal structures. *Science* 291(5508):1534–1536.
- Emlen DJ, Lavine LC, Ewen-Campen B. 2007. On the origin and evolutionary diversification of beetle horns. *Proc Natl Acad Sci U S A.* 104(1 Suppl 1):8661–8668.
- Fenner J, Benson C, Rodriguez-Caro L, Ren A, Papa R, Martin A, Hoffmann F, Range R, Counterman BA. 2020. Wnt genes in wing pattern development of coliadinae butterflies. *Front Ecol Evol.* 8:197. doi:10.3389/fevo.2020.00197.
- Fenner J, Rodriguez-Caro L, Counterman B. 2019. Plasticity and divergence in ultraviolet reflecting structures on Dogface butterfly wings. *Arthropod Struct Dev.* 51:14–22.
- Gerould JH. 1914. Species-building by hybridization and mutation. *Am Nat.* 48(570):321–338.
- Ghiradella H. 1989. Structure and development of iridescent butterfly scales: lattices and laminae. *J Morphol.* 202(1):69–88.
- Ghiradella H. 1994. Structure of butterfly scales: patterning in an insect cuticle. *Microsc Res Tech.* 27(5):429–438.
- Gopinath G, Arunkumar KP, Mita K, Nagaraju J. 2016. Role of Bmzmf-2, a *Bombyx mori* CCH zinc finger gene, in masculinisation and differential splicing of Bmtra-2. *Insect Biochem Mol Biol.* 75:32–44.
- Gruzin M, Mekheal M, Ruhlman K, Winkowski M, Petko J. 2020. Developmental expression of doublesex-related transcripts in the common house spider, *Parasteatoda tepidariorum*. *Gene Expr Patterns.* 35:119101.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29(8):1072–1075.
- Hopkins BR, Kopp A. 2021. Evolution of sexual development and sexual dimorphism in insects. *Curr Opin Genet Dev.* 69:129–139.
- Huylmans AK, Macon A, Vicoso B. 2017. Global dosage compensation is ubiquitous in lepidoptera, but counteracted by the masculinization of the Z chromosome. *Mol Biol Evol.* 34(10):2637–2649.
- Iijima T, Yoda S, Fujiwara H. 2019. The mimetic wing pattern of *Papilio polytes* butterflies is regulated by a doublesex-orchestrated gene network. *Commun Biol.* 2(1):257.
- Irestedt M, Jönsson KA, Fjeldså J, Christidis L, Ericson PG. 2009. An unexpectedly long history of sexual selection in birds-of-paradise. *BMC Evol Biol.* 9:235.
- Ito Y, Harigai A, Nakata M, Hosoya T, Araya K, Oba Y, Ito A, Ohde T, Yaginuma T, Niimi T. 2013. The role of doublesex in the evolution of exaggerated horns in the Japanese rhinoceros beetle. *EMBO Rep.* 14(6):561–567.
- Jia LY, Chen L, Keller L, Wang J, Xiao JH, Huang DW. 2018. Doublesex evolution is correlated with social complexity in ants. *Genome Biol Evol.* 10(12):3230–3242.
- Jarman P. 1983. Mating system and sexual dimorphism in large terrestrial, mammalian herbivores. *Biol Rev.* 58(4):485–520.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–3066.
- Katsuma S, Shoji K, Sugano Y, Suzuki Y, Kiuchi T. 2019. Msc-induced dosage compensation in silkworm cultured cells. *FEBS Open Biol.* 9(9):1573–1579.
- Kemp DJ. 2008. Female mating biases for bright ultraviolet iridescence in the butterfly *Eurema hecabe* (Pieridae). *Behav Ecol.* 19(1):1–8.
- Kemp DJ, Rutowski RL. 2011. Chapter 2 - The role of coloration in mate choice and sexual interactions in butterflies. In: Brockmann HJ, Roper TJ, Naguib M, Mitani JC, Simmons LW, editors. *Advances in the study of behavior*. Vol. 43. Amsterdam, Netherlands: Academic Press. p. 55–92.
- Kemp DJ, Rutowski RL, Mendoza M. 2005. Colour pattern evolution in butterflies: a phylogenetic analysis of structural ultraviolet and

- melanic markings in North American sulphurs. *Evol Ecol Res.* 7(1):133–141.
- Kijimoto T, Moczek AP, Andrews J. 2012. Diversification of doublesex function underlies morph-, sex-, and species-specific development of beetle horns. *Proc Natl Acad Sci U S A.* 109(50):20526–20531.
- Kimura K, Hachiya T, Koganezawa M, Tazawa T, Yamamoto D. 2008. Fruitless and doublesex coordinate to generate male-specific neurons that can initiate courtship. *Neuron* 59(5):759–769.
- Kiuchi T, Sugano Y, Shimada T, Katsuma S. 2019. Two CCCH-type zinc finger domains in the Masc protein are dispensable for masculinization and dosage compensation in *Bombyx mori*. *Insect Biochem Mol Biol.* 104:30–38.
- Kopp A. 2012. Dmrt genes in the development and evolution of sexual dimorphism. *Trends Genet TIG.* 28(4):175–184.
- Kronforst MR. 2015. Exploring the molecular basis of monarch butterfly color pattern variation. *Pigment Cell Melanoma Res.* 28(2):127–130.
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 35(6):1547–1549.
- Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR. 2014. Doublesex is a mimicry supergene. *Nature* 507(7491):229–232.
- Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 9(1):559.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 9(4):357–359.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* 12:323.
- Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658–1659.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15(12):550.
- MacManes MD. 2018. The Oyster River Protocol: a multi-assembler and kmer approach for de novo transcriptome assembly. *PeerJ.* 6:e5428.
- Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput Biol.* 14(1):e1005944.
- Martin A, Reed RD. 2014. Wnt signaling underlies evolution and development of the butterfly wing pattern symmetry systems. *Dev Biol.* 395(2):367–378.
- Miyazaki S, Fujiwara K, Kai K, Masuoka Y, Gotoh H, Niimi T, Hayashi Y, Shigenobu S, Maekawa K. 2021. Evolutionary transition of *doublesex* regulation in termites and cockroaches: from sex-specific splicing to male-specific transcription. *Sci Rep.* 11(1):15992.
- Morita S, Ando T, Maeno A, Mizutani T, Mase M, Shigenobu S, Niimi T. 2019. Precise staging of beetle horn formation in *Trypoxylus dichotomus* reveals the pleiotropic roles of doublesex depending on the spatiotemporal developmental contexts. *PLoS Genet.* 15(4):e1008063.
- Nishikawa H, Iijima T, Kajitani R, Yamaguchi J, Ando T, Suzuki Y, Sugano S, Fujiyama A, Kosugi S, Hirakawa H, et al. 2015. A genetic mechanism for female-limited Batesian mimicry in *Papilio* butterfly. *Nat Genet.* 47(4):405–409.
- Ogita Y, Mawaribuchi S, Nakasako K, Tamura K, Matsuda M, Katsumura T, Oota H, Watanabe G, Yoneda S, Takamatsu N, et al. 2020. Parallel evolution of two *dmrt1*-derived genes, *dmy* and *dm-W*, for vertebrate sex determination. *iScience* 23(1):100757.
- Palmer DH, Kronforst MR. 2020. A shared genetic basis of mimicry across swallowtail butterflies points to ancestral co-option of doublesex. *Nat Commun.* 11(1):6.
- Panara V, Budd GE, Janssen R. 2019. Phylogenetic analysis and embryonic expression of panarthropod *Dmrt* genes. *Front Zool.* 16(1):23.
- Papke RS, Kemp DJ, Rutowski RL. 2007. Multimodal signalling: structural ultraviolet reflectance predicts male mating success better than pheromones in the butterfly *Colias eurytheme* L. (Pieridae). *Anim Behav.* 73(1):47–54.
- Paradis E. 2012. Analysis of phylogenetics and evolution with R. 2nd ed. New York: Springer.
- Pecháček P, Stella D, Kleisner K. 2019. A morphometric analysis of environmental dependences between ultraviolet patches and wing venation patterns in *Gonepteryx* butterflies (Lepidoptera, Pieridae). *Evol Ecol.* 33(1):89–110.
- Prakash A, Monteiro A. 2020. Doublesex mediates the development of sex-specific pheromone organs in bicyclus butterflies via multiple mechanisms. *Mol Biol Evol.* 37(6):1694–1707.
- Price DC, Egizi A, Fonseca DM. 2015. The ubiquity and ancestry of insect doublesex. *Sci Rep.* 5(1):13068.
- Rahnenfuhrer A. 2021. topGO: enrichment analysis for Gene Ontology. R package version 2.44.0.
- Raymond CS, Shamu CE, Shen MM, Seifert KJ, Hirsch B, Hodgkin J, Zarkower D. 1998. Evidence for evolutionary conservation of sex-determining genes. *Nature* 391(6668):691–695.
- Rice GR, Barmina O, Luecke D, Hu K, Arbeitman M, Kopp A. 2019. Modular tissue-specific regulation of doublesex underpins sexually dimorphic development in *Drosophila*. *Development* 146(14):dev178285. doi:10.1242/dev.178285.
- Rodriguez-Caro L, Fenner J, Benson C, Van Belleghem SM, Counterman BA. 2020. Genome assembly of the dogface butterfly *Zerene cesonia*. *Genome Biol Evol.* 12(1):3580–3585.
- Rutowski RL. 1977. The use of visual cues in sexual and species discrimination by males of the small sulphur butterfly *Eurema lisa* (Lepidoptera, Pieridae). *J Comp Physiol.* 115(1):61–74.
- Seppy M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. *Methods Mol Biol Clifton Biol.* 1962:227–245.
- Shelby EA, Fenner J, Counterman BA. 2020. Impact of artificial diet on development and coloration in *Zerene cesonia*. *J Lepidopterists Soc.* 74(2):65–72.
- Sibley CG. 1957. The Evolutionary and taxonomic significance of sexual dimorphism and hybridization in birds. *The Condor* 59(3):166–191.
- Silberglied RE, Taylor OR. 1973. Ultraviolet differences between the sulphur butterflies, *Colias eurytheme* and *C. philodice*, and a possible isolating mechanism. *Nature* 241(5389):406–408.
- Silberglied RE, Taylor OR. 1978. Ultraviolet reflection and its behavioral role in the courtship of the sulfur butterflies *Colias eurytheme* and *C. philodice* (Lepidoptera, Pieridae). *Behav Ecol Sociobiol.* 3(3):203–243.
- Stella D, Fric Z, Rindos M, Kleisner K, Pechacek P. 2018. Distribution of ultraviolet ornaments in *Colias* butterflies (Lepidoptera: Pieridae). *Environ Entomol.* 47(5):1344–1354.
- Verhulst EC, van de Zande L. 2015. Double nexus—Doublesex is the connecting element in sex determination. *Brief Funct Genomics.* 14(6):396–406.
- Walters JR, Hardcastle TJ. 2011. Getting a full dose? Reconsidering sex chromosome dosage compensation in the silkworm, *Bombyx mori*. *Genome Biol Evol.* 3:491–504.
- Walters JR, Hardcastle TJ, Jiggins CD. 2015. Sex chromosome dosage compensation in *Heliconius* butterflies: global yet still incomplete? *Genome Biol Evol.* 7(9):2545–2559.
- Wilts BD, Pirih P, Stavenga DG. 2011. Spectral reflectance properties of iridescent pierid butterfly wings. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol.* 197(6):693–702.
- Wu TD, Watanabe CK. 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21(9):1859–1875.
- Wulfskuhl JD, Petersen NS, Otto JJ. 1998. Changes in the F-actin cytoskeleton during neurosensory bristle development in *Drosophila*: the role of singed and forked proteins. *Cell Motil Cytoskeleton.* 40(2):119–132.
- Yang Y, Zhang W, Bayrer JR, Weiss MA. 2008. Doublesex and the regulation of sexual dimorphism in *Drosophila melanogaster* structure, function, and mutagenesis of a female-specific domain. *J Biol Chem.* 283(11):7280–7292.
- Zhang J, Cong Q, Shen J, Opler PA, Grishin NV. 2019. Genomics of a complete butterfly continent. bioRxiv 829887. doi: 10.1101/829887.
- Zheng Z-Z, Sun X, Zhang B, Pu J, Jiang Z-Y, Li M, Fan Y-J, Xu Y-Z. 2019. Alternative splicing regulation of doublesex gene by RNA-binding proteins in the silkworm *Bombyx mori*. *RNA Biol.* 16(6):809–820.