Supergene evolution via gain of auto-regulation

Nicholas W. VanKuren*, Sofia I. Sheikh, Darli Massardo, Wei Lu, Marcus R. Kronforst*

Department of Ecology & Evolution, The University of Chicago, Chicago IL USA

***Corresponding authors:**
The University of Chicago
Department of Ecology & Evolution
Zoology 202
1101 E 57th Street
Chicago IL 60637

E-mail: nvankuren@uchicago.edu, mkronforst@uchicago.edu

Keywords: polymorphism, sexual dimorphism, butterfly, co-option, mimicry

The development of complex phenotypes requires the coordinated action of many genes across space and time, yet many species have evolved the ability to develop multiple discrete, alternate phenotypes[1–3]. Such polymorphisms are often controlled by supergenes, sets of tightly-linked mutations in one or more loci that function together to produce a complex phenotype[4]. Although theories of supergene evolution are well-established, the mutations that cause functional differences between supergene alleles remain essentially unknown. *doublesex* is the master regulator of insect sexual differentiation but functions as a supergene in multiple *Papilio* swallowtail butterflies, where divergent *dsx* alleles control development of discrete non-mimetic or mimetic female wing color patterns[5–7]. Here we demonstrate that the functional elements of the mimetic allele in *Papilio alphenor* are six new *cis*-regulatory elements (CREs) spread across 150 kb that are bound by DSX itself. Our findings provide experimental support to classic supergene theory and suggest that the evolution of auto-regulation may provide a simple route to supergene origination and to the co-option of pleiotropic genes into new developmental roles.

While sex chromosomes are the most obvious supergenes, numerous complex polymorphisms spanning alternative social structures in ants[2,8], flower structures[9], and male plumage[10] have now been traced back to allelic variation in supergenes. Theory predicts that supergenes evolve when selection favors linkage disequilibrium between alleles at multiple loci that function together to control development of an adaptive phenotype[11]. Reduced recombination between supergene alleles, caused by a variety of mechanisms including inversions, allows supergene alleles to accumulate subsequent mutations that refine the supergene's function and therefore the adaptive phenotype. Although the genetic architectures of several supergenes are well-characterized, the genetic variants that cause functional differences between supergene alleles, and therefore the evolutionary origins of the alleles, their architectures, and polymorphisms they control, remain elusive.

doublesex functions as a supergene in at least five species of *Papilio* swallowtail butterflies, where it controls the switch between discrete female wing color patterns[5–7,12]. The switch between male-like non-mimetic color patterns and a novel mimetic color pattern in the closely-related species *P. polytes* and *P. alphenor* is caused by the dominant *H* allele[5,6]. *H* is inverted relative to the ancestral *h* allele, has an extremely divergent sequence from *h*, and contains *dsx* and a novel non-coding gene, *untranslated three exons* (*U3X*; **Fig 1A-B**)[5–7]. A

unique spike of *dsx*^{H} expression in early pupal wings initiates the mimetic color development program, but the genetic basis of this expression pattern remained unknown (**Fig 1C**)[13].

We therefore investigated the *cis*-regulatory architecture and evolution of the *P. alphenor* supergene alleles (**Fig 1**). We first narrowed our search for CREs controlling *dsx* wing expression by identifying topologically associating domains (TADs) containing each allele, which we expected to define the local regulatory region[14]. Each allele was contained within a single TAD harboring *dsx* and the adjacent genes *sir-2, rad51*, and *nach*; the mimetic *H* TAD additionally contained *U3X* (**Fig 1D**). While the right TAD boundaries coincided with the right inversion breakpoint in both alleles, the left TAD boundaries were 21.9 kb and 17.9 kb outside of the inversion, suggesting that the inversion itself caused minor changes to the local topology despite high sequence and structural divergence between the alleles (**Fig 1B**). TAD boundaries were similar in male and mid-pupal female and male wings (Supplementary Figure 1). We therefore expected CREs controlling *dsx* wing expression to be within or just to the left of the inversion.

We then identified potential *dsx* CREs using the assay for transposase-accessible chromatin (ATAC)[15]. We expected that the CRE(s) controlling *dsx*^{H} expression, and therefore the mimicry polymorphism, would be 1) unique to the *H* allele, 2) differentially accessible (DA) between female and male wings, and 3) not DA or potentially inactive in other tissues. Alternatively, CREs may be shared between alleles but have accumulated *H*-specific divergence that altered their function in the developing wing. To identify CREs involved in the mimicry switch, we performed ATAC in early and mid-pupal wings and heads from *H/H* and *h/h* females and males (**Fig 1E**; Supplementary Figures 2-5)[13,15].

We first focused on peaks in early pupal wings, where *dsx*^{H} expression spikes in mimetic females (**Fig 1C**)[13]. The *H* allele contained 28 ATAC peaks in early pupal wings, 12 of which were wing-specific (**Fig 1D**). Most peaks fell within or near *dsx*^{H}: 22 within introns plus peaks in exon 1, exon 6, and 0.47 kb and 3.9 kb upstream of the *dsx*^{H} promoter. Peaks were also found at the *UXT* and *U3X* promoters. The *h* allele contained 33 ATAC peaks in early pupal wings, 16 of which were wing-specific. These *h* peaks were located in similar relative locations to *H* peaks, but were absent from ~4 kb upstream of *dsx*^{h} (**Fig 1E**).

We assigned orthology between these putative *h* and *H* CREs and polarized CRE gain and loss using BLAST to search for *P. alphenor* CRE sequences in *P. alphenor, P. polytes* and three additional *Papilio* species (**Fig 1F;** Supplementary Tables 4-5). Despite the enormous sequence divergence between the *h* and *H* alleles (**Fig 1B**), we found surprising conservation of

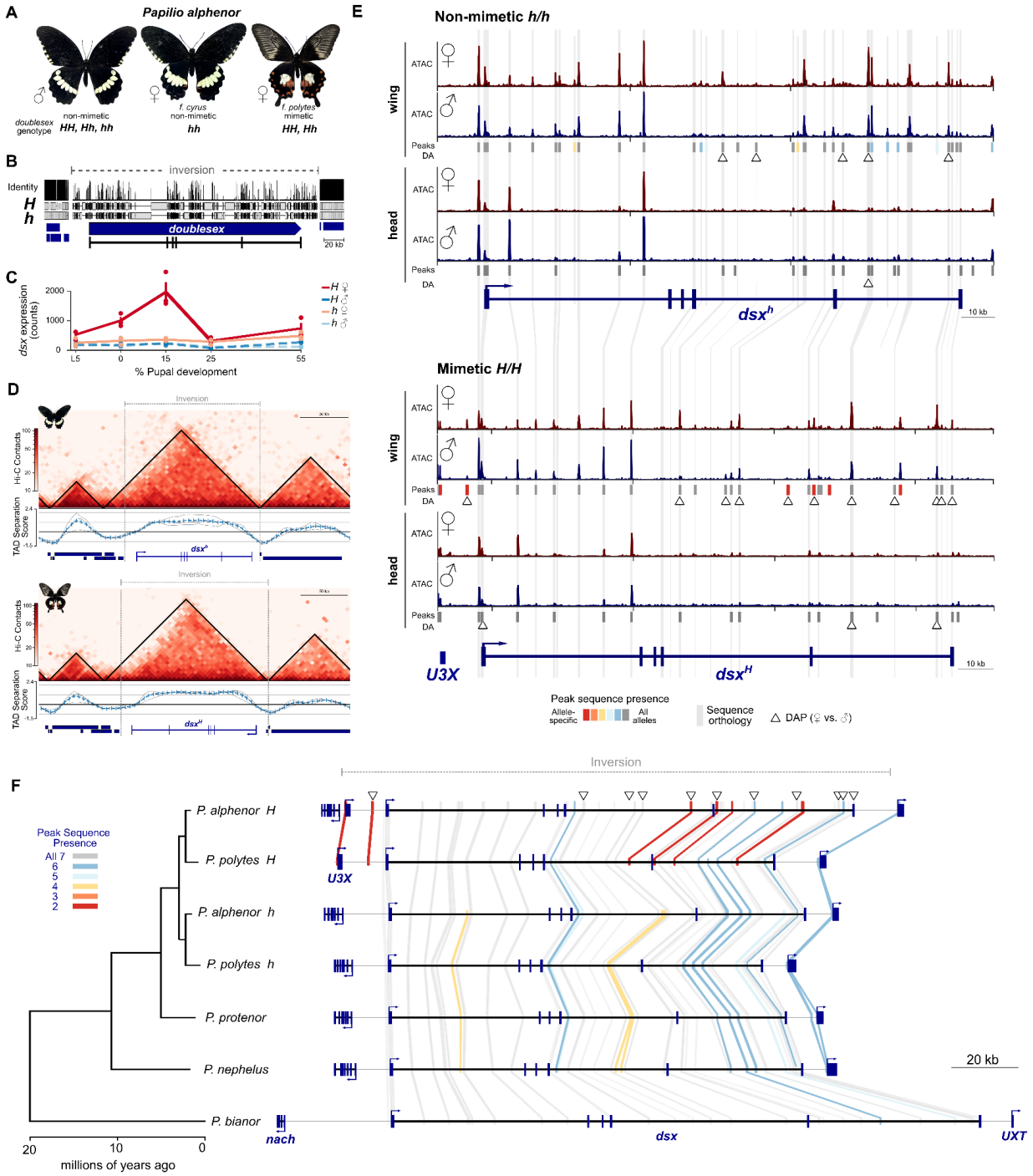

Fig 1. Regulatory architecture of the *doublesex* supergene. (**A**) Female-limited mimicry polymorphism in *Papilio alphenor*. (**B**) Pairwise alignment of *dsx* alleles and flanking 20 kb. The mimetic *H* allele is un-inverted for display. (**C**) *dsx* expression in each group across development. (**D**) Hi-C contacts, topologically associating domains (TADs), and TAD separation scores near the inversion in early pupal female wings. (**E**) Transposase-accessible chromatin within the inversion in early pupal (15%) wings and heads. All tracks are normalized and shown on the same scale [0-450]. Peak calls and differentially accessible (DA) peaks between females and males are shown below. Orthologous peak sequences between the alleles are connected

by shaded lines. Additional information and replicates are shown in Supplementary Figures 2-5 and Supplementary Table 1. (**F**) Orthology and synteny between *P. alphenor* ATAC-seq peaks and outgroup alleles. Orthology was determined by BLASTing *P. alphenor* peak sequences to the genome regions bounded by *nach* and *UXT,* which flank the inverted region, in each species. The *P. alphenor* and *P. polytes H* alleles are un-inverted for display. Peak conservation (i.e. the number of sequences in which the peak sequence is present) is denoted by color and lines connect orthologous regions in each allele. Lines are thickened to make them easier to view. DA peaks between *H/H* female and male early pupal wings are shown above. Additional details and results can be found in the Methods and Supplementary Tables 2 - 7.

CRE sequences and synteny over 20 million years of evolution. Importantly, all *h* CRE sequences were found in the *P. polytes h* allele and at least two other species, but six *H* CRE sequences were unique to the *P. alphenor* and *P. polytes H* allele, strongly suggesting that the *H* allele has gained multiple novel *dsx* CREs (**Fig 1F**). Both *dsx* alleles perform their roles in sexual differentiation equally well, so *H*-specific CREs cannot be essential for dsx^H expression.

We expected that CREs responsible for dsx^H expression would show a correspondingly unique pattern of accessibility in mimetic females. The extreme divergence between the alleles prevented direct comparisons of peak accessibility between mimetic and non-mimetic females (**Fig 1B**). Instead, we identified differentially accessible (DA) peaks between males and females of the same genotype and compared them based on their orthology assignments (**Fig 1F;** Supplementary Tables 6-7). We found that 39.2% (11/28) of *H* CREs were differentially accessible (DA) between the sexes, including half (3/6) of *H* allele-specific CREs. In contrast, 15.1% (5/33) of *h* CREs were DA, perhaps reflecting the similar *dsx* expression patterns and small amount color pattern dimorphism between non-mimetic females and males (**Fig 1E**)[16]. This sexually dimorphic accessibility is likely involved in color pattern development, as no peaks were DA specifically in heads and only one CRE was DA in both heads and wings.
These results strongly support the role of a small number of dsx^H-specific CREs in activating the *P. alphenor* mimicry switch. We experimentally tested if these *H*-specific CREs and six conserved CREs were required for dsx^H expression, and therefore mimetic color pattern development, using CRISPR/Cas9 knockouts (**Fig 2**; Supplementary Figure 6; Supplementary Tables 8 - 9). We expected that knocking out CREs required for dsx^H expression should cause mimetic females to develop non-mimetic color patterns.

Consistent with their requirement for *dsx* wing expression and mimetic color pattern development, knockouts of three conserved CREs caused mimetic females to develop mosaic non-mimetic color patterns (**Fig 2**; Supplementary Figure 6). We observed no color pattern differences in non-mimetic females when we targeted these six orthologous CREs. Importantly, mKOs of the *H*-specific CRE 22670 contained numerous patches of mimetic to the non-mimetic

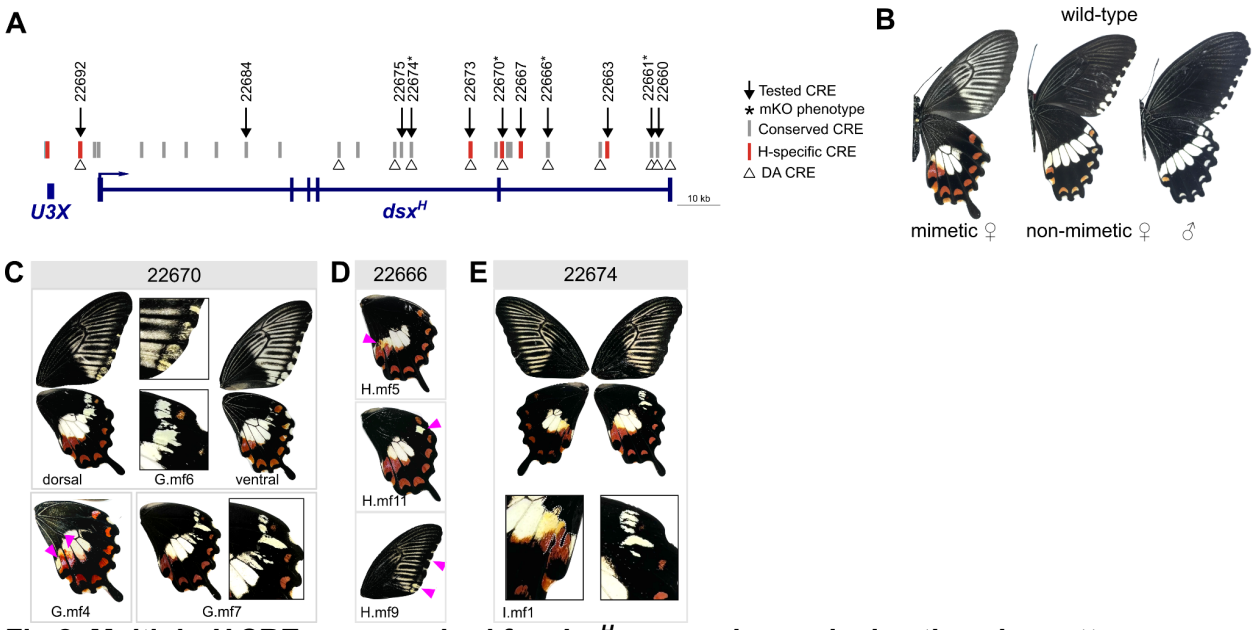

Fig 2. Multiple *H* CREs are required for *dsx*^{H} expression and mimetic color pattern formation. Mosaic CRISPR/Cas9 mutants were generated by injecting 2 - 4 sgRNAs targeting a CRE of interest (**A**). Tested CREs are labeled, with CRE orthology and DA in early pupal wings carried over from Fig 1. Mutants are identified by the appearance of non-mimetic color patterns in mimetic *H/h* or *H/H* females. (**B**) Wild-type color patterns. (**C-E**) Examples of females with mosaic patches from knockouts of CREs 22670 (**C**), 22666 (**D**), and 22674 (**E**). Additional mKOs of additional CREs can be found in Supplementary Figure 6.

color pattern, strongly suggesting that this CRE controls the unique *dsx*^{H} expression pattern in mimetic females (**Fig 2D**). We therefore hypothesize that the *dsx* supergene originated through recruitment of one or more novel CREs that significantly increased expression in early pupal wings. The inversion that linked this novel CRE and the *dsx* promoter, two functional elements located 100 kb apart, without deleterious pleiotropic effects on *dsx* and nearby gene expression would have been favored by selection for mimicry.

We next sought to identify the transcription factor(s) (TFs) that bind to these CREs and control *dsx* wing expression. The TFs that directly control *dsx* expression are unknown in any organism. Genetic analyses in *Drosophila* showed that Hox and other TFs, including Sex-combs reduced, Abdominal-B, and Caudal[17], are required for DSX expression in certain contexts, but it is unknown if these TFs directly regulate *dsx* transcription. On the other hand, ChIP-seq experiments identified binding sites for over 230 different DNA binding proteins within *dsx* in whole adult *Drosophila*, chiefly DSX itself (25 sites)[18].

To identify TFs that regulate *dsx* wing expression, we first searched for known TF binding site motifs enriched in *dsx*^{H} CREs using HOMER (Supplementary Figure 7). CREs were CUT&RUN, the active promoter/enhancer histone mark H3K4 tri-methylation, and negative control

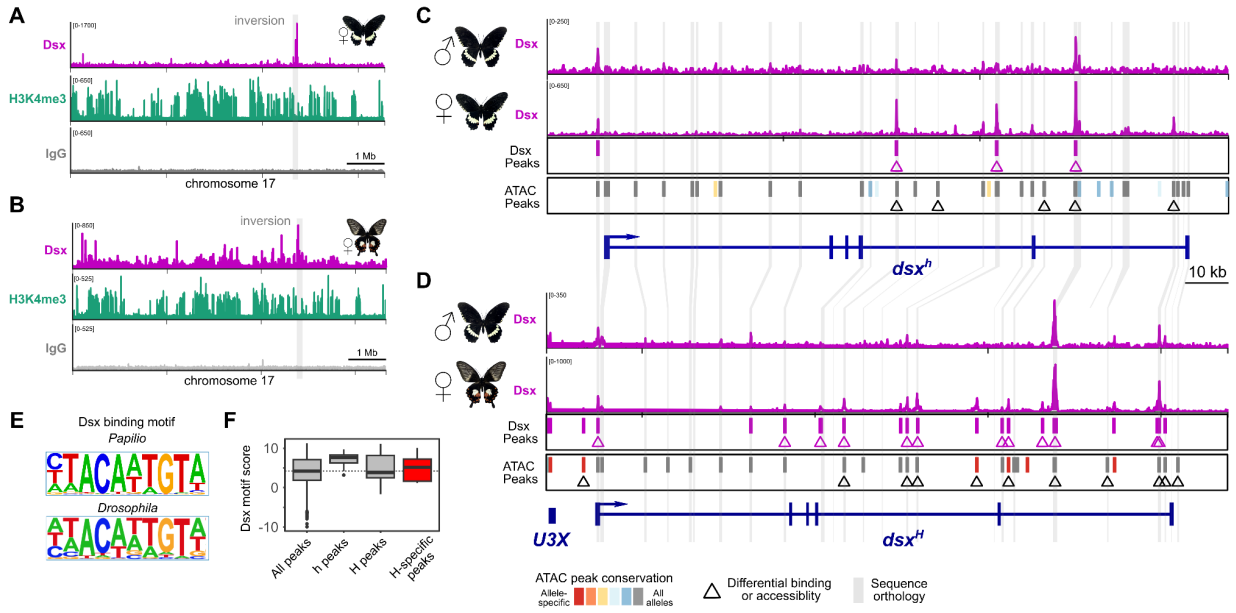

Fig 3. *dsx^H* gained multiple novel CREs bound by DSX itself. (**A-B**) Normalized coverage tracks for DSX immunoglobulin in non-mimetic *h/h* (**A**) and mimetic *H/H* (**B**) pupal wings. The whole of chromosome 17 is shown. (**C-D**) DSX binding in the *h* allele (**C**) and *H* allele (**D**) in early pupal wings. DSX peaks, differentially bound (DB) peaks between males and females, and ATAC data from Fig 1 are shown below. (**E**) The *Papilio* DSX binding motif relative to the *Drosophila* DSX binding motif. Letter heights are proportional to the base frequencies at each position. (**F**) The distribution of log-odds probability scores for DSX motifs in DSX CUT&RUN peaks. Higher values indicate better matches. Scores are shown for motifs in all peaks genome-wide (all peaks); peaks within orthologous *h* and *H* CREs; or in *H*-specific CREs. No pairwise comparisons were significantly different (all Welch's *t*-test *p*-values >0.1).

most significantly enriched with motifs for Paired (p = 1e-13), Caudal (p =1e-11), Extradenticle (p = 1e-10), and DSX (p = 1e-9), immediately suggesting that the mimetic supergene allele could be auto-regulated. To test this hypothesis and identify the direct targets of DSX in wing development, we assayed genome-wide patterns of DSX binding in early and mid-pupal *h/h* and *H/H* wings using CUT&RUN (**Fig 3;** Supplementary Figures 8-10; Supplementary Table 10)[19].

Overall, we found 9160 DSX peaks genome-wide among all samples, with the majority of peaks (84.0%) initially called only in mimetic *H/H* female samples where DSX expression is highest. DSX peaks mostly overlapped ATAC peaks (89.5%), and DSX peaks were enriched with a motif similar to the known *Drosophila* DSX binding site (p = 1e-596; Supplementary Figure 9), supporting the quality of the data. For comparison, 82.3% (9264/11339) H3K4me3 CUT&RUN peaks overlapped ATAC peaks. In addition, we found strong DSX peaks in genes known to be bound by DSX in *Drosophila*, including *bric-a-brac 1* (Supplementary Figure 11).

Importantly, we found multiple DSX CUT&RUN peaks within and just upstream of *dsx* in all samples (**Fig 3**). In fact, the strongest peaks on chromosome 17 in non-mimetic females

were found within the inversion (**Fig 3C**). The *h* allele contained four DSX peaks in early pupal wings, all within conserved CREs, while the mimetic *H* allele contained 22 DSX peaks: 19 intronic, one 0.8 kb and one 4.1 kb upstream of dsx^H, and one at the *U3X* promoter (**Fig 3D**). Surprisingly, five of the six *H*-specific CREs contained DSX peaks, strongly suggesting that the mimetic allele gained novel auto-regulatory interactions (**Fig 3D**). Importantly, the *H*-specific CRE 22670 that yielded strong mKO phenotypes (**Fig 2D**), also contained a strong DSX peak that was differentially bound between females and males.

Overall, 55% of DSX peaks (12/22) in the *H* allele and three of four *h* peaks were significantly differentially bound (DB) between males and females in early pupal wings. All three DB *h* peaks were conserved and also DB in the *H* allele, suggesting that DSX binding in these CREs is involved in sexual dimorphism. Thus, differential binding in the *H* allele may primarily contribute to the unique spike of dsx^H expression. The different patterns of DSX binding in conserved CREs appears to be caused by differential use of those CREs rather than mutations that disrupt binding sites, as the binding sites are well-conserved between the supergene alleles and across species (**Fig 3E-F**). Log-odds probabilities, measurements of the strength of matches between predicted DSX binding sites and the consensus motif, were not significantly different between genome-wide DSX peaks, conserved peaks, or *H*-specific peaks (**Fig 3F;** all Welch's *t*-test *p*-values > 0.10). In fact, two of the top three matches are found in *H*-specific CREs (Supplementary Table 11). These results suggest that the mimetic *H* allele has gained multiple new CREs with strong DSX binding sites that regulate *dsx* expression in the developing wing.

Finally, in addition to the spike of widespread dsx^H expression in early pupal wings, DSX^H becomes uniquely expressed in regions of the wing that will become white in mid-pupal wings[13]. Interestingly, no *H*-specific CREs were even accessible in mid-pupal wings (Supplementary Figure 4). Instead, three conserved peaks were DB by DSX between the sexes in both *h/h* and *H/H* comparisons, and a single conserved peak was DB specifically between *H/H* males and females (Supplementary Figure 10). These results suggest that *H*-specific CREs and DSX binding may be required early in development to trigger the developmental switch, but that conserved CREs are differentially used to sustain the mimetic color pattern program throughout development. In other words, changes in conserved wing CREs may have helped refine the mimetic program after its origination via novel, auto-regulated *dsx* expression in the early wing.

Antibody stains showed that DSX^H expression never fully pre-figured the adult mimetic pattern, suggesting that DSX^H initiates mimetic pattern development in early pupal wings but

quickly becomes decoupled from it[13]. We expected that DSX^{H} initiated the mimetic program by directly regulating one or more downstream genes in early pupae that execute the mimetic program[5,6,13]. To identify these direct targets and characterize the consequences of DSX binding on chromatin accessibility and gene expression, we directly compared genome-wide patterns of DSX CUT&RUN, ATAC-seq, and differential gene expression (DE) between mimetic and non-mimetic females (**Fig 4**).

We first identified DE genes by re-analyzing pupal wing RNA-seq data from five developmental stages[13], then intersected those genes with our DSX and ATAC peak data. Consistent with our previous work, 1523 genes were DE between mimetic and non-mimetic

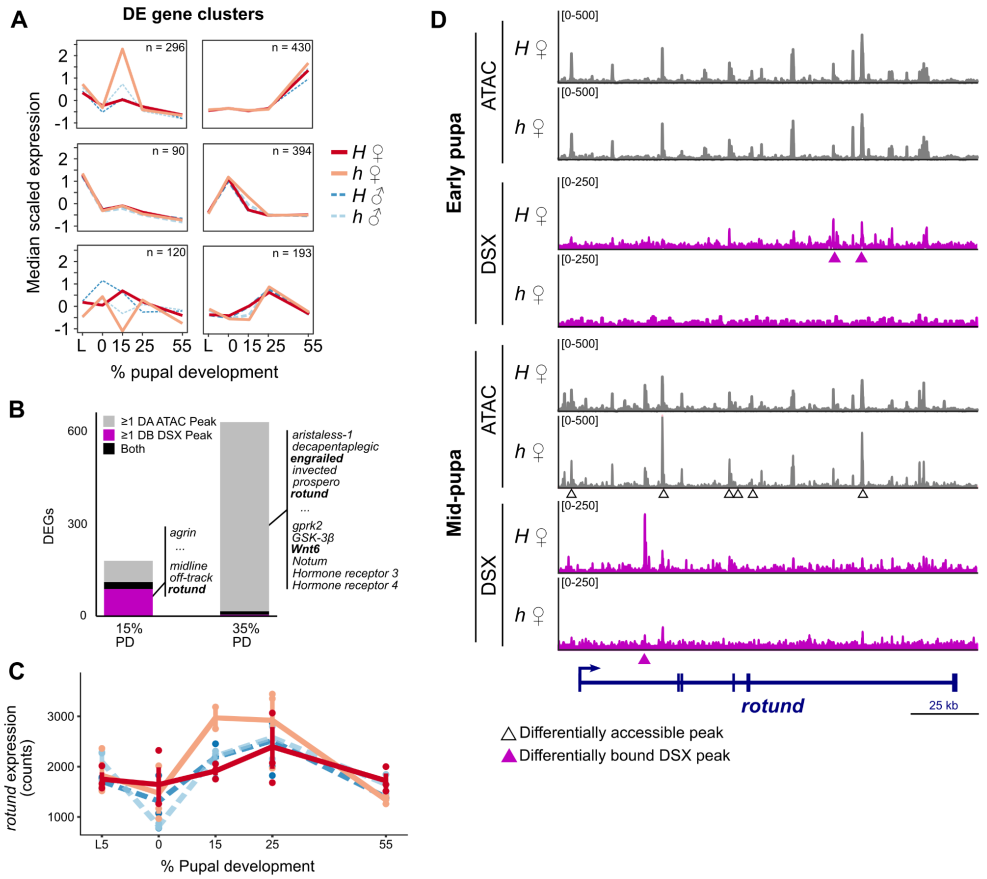

Fig 4. Acute and long-term consequences of DSX binding in early pupal wings. (**A**) Median expression profiles of clustered differentially expressed (DE) genes between mimetic and non-mimetic females. Hierarchical clustering was performed using Mclust, using BIC to choose *k*[20]. L: 5th instar larva. (**B**) Differential DSX binding and differential ATAC peak accessibility in DE genes. Select genes known to be involved in wing color pattern development, transcription, or signaling are shown. Genes known from previous studies to be involved in the *Papilio* mimicry switch are bolded[13,21]. Full dataset can be found in Supplementary Table 13. (**C**) *rotund* expression during *P. alphenor* wing development. (**D**) ATAC and CUT&RUN data in *rotund*.

females across development, with the majority (54.8%) of those genes being DE specifically in early pupal (15%) wings (**Fig 4A**; Supplementary Figure 12; Supplementary Table 13). First, 1103 (12.0%) DSX peaks in 776 genes were DB between early pupal mimetic and non-mimetic females, including 62 peaks in 49 DE genes. These potentially direct targets include *rotund* (a TF previously implicated in the mimicry switch in *P. polytes*[21]), the T-box TF *midline*, and at least three other DNA-binding proteins (**Fig 4B**; Supplementary Table 13). The small number of genes that DSX appears to directly regulate in the *P. alphenor* wing is consistent with work in *Drosophila* that showed DSX binds many targets genome-wide, but affects expression of a small number of those genes, presumably due to the presence of the appropriate co-factors[22]. By mid-pupal development, only 177 (1.9%) DSX peaks were DB, with just 15 peaks in 15 DE genes, including *rotund*.

Patterns of chromatin accessibility were opposite those of DSX binding. While 1.7% of ATAC peaks were DA between early pupal female wings, 27.7% were DA in mid-pupal female wings, suggesting that DSX binding early in development alters the wing regulatory landscape throughout development (Supplementary Figure 13). More specifically, 5.6% of DE genes contained DA peaks in early pupal wings, yet 40.8% contained DA ATAC peaks in mid-pupal wings, including genes known to be involved in the *dsx* mimicry switch and to specify color patterns in other butterflies, including *invected, engrailed,* and *aristaless-1* (**Fig 4**)[13]. Thus, while DSX directly regulates few genes early in pupal development, those effects are propagated to later stages while being decoupled from DSX binding itself.

Our results provide several new insights into the structures and evolution of supergenes that have remained hidden by the extensive LD and complexity that characterize many of the best-characterized supergenes. First, we experimentally showed that multiple functional elements are required for supergene function, supporting predictions that inversions and other mechanisms of recombination suppression link distant functional loci together[11]. Second, our results suggest that the *dsx* supergene originated via the gain of a novel auto-regulatory element(s) in the 3' half of the gene that enabled positive reinforcement of dsx^H expression in the early pupal wing, causing a spike in *dsx* expression that initiates the mimetic pattern program[6,13]. An inversion then linked this novel DSX binding site to the *dsx* promoter and facilitated recruitment of additional *H*-specific CREs that reinforced this positive feedback loop and refined DSX expression across wing development.

Beyond supergenes, our findings provide novel insight into the genetic mechanisms by which pleiotropic genes are co-opted into new developmental roles. Auto-regulation may avoid

many of the deleterious pleiotropic effects from ectopic expression because it can only occur in tissues or stages where the gene is already expressed. Second, auto-regulation, particularly up-regulation, increases the probability that novel alleles are dominant and therefore immediately exposed to natural selection. Auto-regulation could play an identical role in the evolution of multi-gene supergenes, where key TFs regulate their own and/or nearby gene expression. Recombination suppression and subsequent divergence between alleles, in CREs or protein-coding sequences, would then refine the supergene alleles' functions.

Acknowledgements

We thank the University of Chicago greenhouse staff and the University of Chicago Functional Genomics Facility for research support; the University of Chicago's Research Computing Center for computational support.

Author contributions

NWV - Conceptualization, Investigation, Visualization, Writing - original draft, Writing - review & editing; SIS - Conceptualization, Investigation; DM - Investigation; WL - Investigation; MRK - Conceptualization, Funding acquisition, Writing - review & editing

Funding

This work was supported by NIH R35 GM131828 to MRK.

Data Availability

Illumina sequencing data is publicly available in the National Center for Biotechnology Information (NCBI) under BioProject PRJNA1062051. Genome assemblies, annotations, R projects, and full analysis results are publicly available in the Dryad repository under accession doi://AAAAAAAAAA.

Materials and Methods

Butterfly care

Papilio alphenor pupae were purchased from Philippines breeders and allowed to emerge in the University of Chicago greenhouses. New adults were sexed, labeled with a unique number on the forewing, and the sexes separated until use. We determined each individual's *dsx* genotype using DNA from a single leg and a custom TaqMan (Life Technologies, USA) assay. We set up crosses between multiple virgin adults carrying the desired alleles in $2m^3$ mesh cages, allowed them to mate, and provided *Citrus* (Meyer lemon) shrubs for oviposition. Adults were fed Bird's Choice artificial nectar and supplied with *Lantana*. Pre-pupae were collected and placed into labeled boxes in an incubator set to 25°C, 16h:8h light/dark cycle, and constant 65% humidity. Pupal development takes approximately 15 days under these conditions; experiments focused on two days and five days after pupation (15% and 35% pupal development, respectively).

Genome sequencing and assembly

We extracted high molecular weight (HWM) genomic DNA from thorax of freshly killed *P. alphenor* females that were homozygous for the *h* or *H* alleles using the QIAgen GenomicTip G-100 kit (QIAgen, USA). Extractions followed the manufacturer's instructions, except we incubated chopped fresh tissue in lysis buffer and proteinase K overnight at 50°C and shaking at 200 rpm before purification. We then constructed Oxford Nanopore sequencing libraries using the ONT Ligation Sequencing Kit (LSK-110) and eliminated reads <10 kb using the PacBio SRE XS kit before sequencing on a MinION Mk1b and R9.4.3 flow cells to 30X - 40X coverage.

We called bases using Guppy and super high quality base calling (dna_r9.4.1_450bps_sup.cfg), then assembled the mimetic *H* and non-mimetic *h* genomes separately using these raw reads and Flye v2.9.1 with default settings with expected genome size set to 250 Mb. The initial Flye assemblies were each polished using the Guppy basecalls and Medaka v1.7.2 (medaka_consensus) with the appropriate error model (r941_min_sup_g507). We then purged duplicates using purge_dups v1.2.5.

We then scaffolded contigs together into 31 chromosomes using Hi-C data (see below) and the 3d-dna pipeline[23]. We identified the Z and W chromosomes by analyzing coverage of re-sequencing data from five males and five females[5]. The remaining 29 chromosomes were ordered and labeled by decreasing size. We assembled the *alphenor* mitochondrial genome using NOVOplasty v4.2[24] using sequencing data from SRR1108726 and the RefSeq mtDNA assembly for *polytes* (NC_024742.1) as the seed sequence. This resulted in a single circularized sequence of 15,247 bp. We added this sequence as chrM to each assembly.

We identified repeat sequences in the mimetic assembly using RepeatModeler, then used our custom library, the RepBase 20181026 "arthropoda" database, and Dfam 20181026 database to identify and mask repeats genome-wide in both assemblies using RepeatMasker. Finally, we hard masked regions in the nuclear genomes with homology to chrM (blastn e-value < 1e-50).

We sequenced and assembled a new *Papilio nephelus* genome following the same protocol, using DNA from a single random female. *Papilio nephelus* is a sexually monomorphic, non-mimetic species that diverged ~15 mya from the *P. alphenor* lineage. Final assembly statistics can be found in Supplementary Table 14.

TAD Plot

Mimetic: chr17:5850000-6183000
Non-mimetic: chr17:5704000-6023000

Inversion:
Mimetic chr17:5937900-6095400
Non-mimetic chr17:5791937-5930539

Annotation

We annotated the mimetic assembly using EvidenceModeler 1.1.1[25]. We first assembled a high-quality transcript database using PASA [25], SE50 data generated in VanKuren et al. [13], and PE100 and SE50 data from Nallu et al.[26]. After adapter trimming, we performed *de novo* and genome-guided assembly using Trinity v2.10.0[27] and genome-guided assembly using StringTie v1.3.3[28]. RNA-seq data was also mapped to the mimetic *alphenor* assembly using STAR 2.6.1d[29], and the resulting alignments used to generate genome-guided assemblies with Trinity and StringTie 1.3.1[28]. We combined *de novo* and genome-guided assemblies using PASA 2.4.1[30]. Evidence for protein-coding regions came from mapping the UniProt/Swiss-Prot (2020_06) database and all Papilionoidea proteins available in NCBI's GenBank nr protein database (downloaded 6/2020) using exonerate[31]. We identified high-quality multi-exon protein-coding PASA transcripts using TransDecoder (transdecoder.github.io), then used these models to train and run Genemark-ET 4[32] and GlimmerHMM 3.0.4[33]. We also predicted gene models using Augustus 3.3.2[34], the supplied *heliconius_melpomene1* parameter set, and hints derived from RNA-seq and protein mapping above. Augustus predictions with >90% of their length covered by hints were considered high-quality models. Transcript, protein, and *ab initio* data were integrated using EVM with the weights in Supplementary Table 15.

Raw EVM models were then updated twice using PASA to add UTRs and identify alternative transcripts. Gene models derived from transposable element proteins were identified using BLASTp and removed from the annotation set. We manually curated the *dsx* region and key color patterning genes. The full annotation comprises 20,674 genes encoding 26,074 protein-coding transcripts, containing 97.2% complete and missing 1.9% of endopterygota single-copy orthologs according to BUSCO v5 and OrthoDB v10. We functionally annotated protein models using eggNOG's emapper-2.0.1b utility and the v2.0 eggNOG database[35]. We used liftoff [36] to transfer these annotations to the non-mimetic reference genome assembly.

Downstream analyses used sequence, transcripts, and proteins from only the 31 main chromosomes and chrM.

Hi-C sequencing and analysis

We performed Hi-C on developing hindwings using Dovetail Genomics' (USA) Omni-C kit. Hindwings were dissected from staged pupae then snap frozen in liquid nitrogen and stored at -80°C until use. Wings from three individuals were pooled and pulverized in liquid nitrogen before proceeding with fixation and proximity ligation according to the manufacturer's protocol. We performed Hi-C separately on hindwings from males and females homozygous for each *dsx* allele at two days after pupation. All eight libraries were pooled and sequenced 2 x 150 bp on an Illumina NovaSeq 6000 S1 flow cell at the University of Chicago Functional Genomics Facility (RRID:SCR_019196). We trimmed adapters and low-quality regions using Trimmomatic 0.39 [37], then used Juicer v1.6 [23] and bwa 0.7.17 [38] to map, sort, and de-duplicate reads.

We used these merged_nodups.txt files as input to the 3d-dna pipeline for genome assembly (above). Separately, we used hicExplorer [39] to identify TADs in each sample and to plot results used in Fig 1. Juicer output files were converted to .h5 format using hicConvertFormat and 5 kb resolution (except mfp6 and mmp6, which used 15 kb resolution due to their lower quality), normalized using hicCorrectMatrix and and the Knight-Ruiz method [40], then used to call TADs with hicFindTads and default settings except "--correctForMultipleTesting fdr --minDepth 20000 --maxDepth 50000 --step 10000". Samples mfp6 and mmp6 used settings --minDepth 60000 --maxDepth 150000 --step 30000 to account for their lower resolution.

Genome blacklists

We generated genome blacklists by identifying low-mappability regions using genmap v1.3.0 [41]. We calculated mappability for 50-mers, allowing for 1 mismatch (-k50 -e1), then identified low mappability regions as those with mappability < 1. We merged low map regions within 300 bp of each other using bedtools merge, then kept regions > 100 bp. We combined these regions with all regions with mtDNA homology (identified using BLASTn with E < 1e-50) into a single blacklist. This was performed separately for the mimetic and non-mimetic genome assemblies.

ATAC-seq

We performed ATAC experiments following Lewis and Reed[42] and Buenrostro et al.[15] with minor modifications. Wings were dissected in room temperature PBS then immediately transferred into ice cold sucrose buffer in 2 mL dounce homogenizers. Tissues were dissociated using 30 - 50 strokes with the tight pestle, then transferred to cold 1.5 mL tubes. Cells and nuclei were pelleted by centrifugation for 5 min at 1000 x *g* at 4ºC, then resuspended in 200 uL lysis buffer (wash buffer: 10 mM Tris-HCl pH7.5, 10 mM NaCl, 3 mM MgCl2 plus 0.2% NP-40) and lysed on ice for 5 min. Nuclei were harvested by centrifugation for 5 min at 1000 x *g* for 5 min and 4ºC, then resuspended in 750 uL ice cold wash buffer and counted. Aliquots of 500,000 cells were pelleted for 5 min at 1000 x *g* and 4ºC, then resuspended in transposition mix (25 uL TD buffer, 2.5 uL TDE1 enzyme, 22.5 uL water). Transposition was performed 30 min at 37ºC shaking at 1000 rpm, then cleaned up using the Zymo DNA Clean and Concentrator 5 kit. Libraries were amplified 10 cycles before double-sided cleanup (0.5X - 1.8X) with SPRI Select Beads (Beckman-Coulter, USA). We sequenced libraries PE50 on a single lane of a NovaSeq X 10B flowcell, for an average of 30M read pairs per sample.

ATAC-seq analysis

Raw ATAC-seq sequencing reads were trimmed using Trimmomatic 0.39[37], then mapped to each reference genome using bwa 0.7.17[38]. Duplicate reads were removed using picard[43]. We assessed sample quality using ATACseqQC 1.22.0 [44], and used the cleaned BAM files for downstream analysis. Bigwig tracks for visualization were created and normalized using the RPKM method and deepTools2[45]. We called peaks using F-Seq2 following the authors' recommendations for ATAC-seq data (-pe -l 600 -f 0 -t 4.0 -nfr_upper_limit 150 -pe_fragment_size_range auto) in each sample, then merged peaks from biological replicates using bedtools intersect[46], requiring reciprocal 20% overlap between at least two replicates. Finally, we combined sample peaksets using bedtools merge to generate a comprehensive peakset that we used to identify differentially accessible peaks (DAPs) using DiffBind 3.8.4 [47,48]. We performed all relevant pairwise comparisons between sexes, genotypes, and stages, then corrected *p*-values using the Benjamini-Hochberg method[49]. Only DAPs with global FDR < 0.05 were used in downstream analyses.

We annotated the full ATAC peaksets using annotatePeaks.pl from HOMER 4.11.0 [50].

Peak and peak sequence orthology

We assigned orthology between *h* and *H* ATAC peak sequences and outgroup genomes using BLAST 2.2.24. First, we identified the genome region bounded by *nach* and *UXT* in the *P. polytes* (GCF_000836215.1)[6], *P. protenor* (GCA_029286645.1), *P. nephelus,* and *P. bianor* (available at http://gigadb.org/dataset/100653)[51] genomes by mapping the non-mimetic *P. alphenor nach, dsx,* and *UXT* transcripts to each genome using minimap2 with the "-xsplice" option[52]. We then converted the minimap2 results to GTF format using paftools.js and UCSCtools bedToGenePred and genePredToGtf tools. We extracted and oriented each region, then used blastn to identify *h* and *H* sequences in each target region. We used the following command, for example:

```
blastn -task blastn-short -evalue 1e-3 \
-outfmt "6 qseqid qlen qstart qend sseqid sstart send length \
pident evalue" -max_hsps 1 -query papAlpH.peaks.fa -db papBia.fa
```

We collated BLAST results and *P. alphenor* peak sequences in R, then plotted the results using the gggenomes package[53]. An R project containing the analysis and plotting pipelines can be found in Dryad (AAAAAAAAA).

CUT&RUN

We assessed genome-wide Dsx binding in homozygous males and females at 15% and 35% pupal development using CUT&RUN[54], following the protocol described by Meers et al.[19] with minor modifications. We dissected hindwings in room temperature (RT) PBS, then enzymatically dissociated cells following Prakash and Monteiro[55]. Wings were immediately transferred to 750 uL TrypLE Select Enzyme (Gibco, USA) diluted to 5X in PBS, then broken up by gently pipetting 20X with a p1000 pipet set to 500 uL. Tissues were allowed to dissociate at 32ºC / 1200 rpm in a thermomixer for 20 minutes, pipetting 5X every 5 minutes. Cells were harvested by centrifugation at 600 x *g* for 3 min, washed twice with 1 mL RT wash buffer (WB), then resuspended in 1 mL RT WB. Cells were bound to concanavalin A-bound magnetic beads (Bangs Labs) for 10 min rotating at RT. Beads were separated using a magnetic stand, then immediately resuspended in ice cold antibody buffer (WB + 0.025% digitonin). Permeabilized cells (500 000 per reaction) were aliquoted to 0.2 mL tubes before adding 0.5 ug primary antibody and nutating overnight at 4ºC. Washing, digestion, and purification followed Meers et al. (2019).

We used rabbit anti-Dsx[13], mouse anti-H3K4me3 (Abcam ab8580), or goat anti-rabbit IgG (Cell Signaling Technologies 45262).

We constructed sequencing libraries using the NEBNext Ultra II DNA Library Prep kit following the protocol outlined in Liu (2021) "Library Prep for CUT&RUN with NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (E7645) V.2 (https://www.protocols.io/view/library-prep-for-cut-amp-run-with-nebnext-ultra-ii-kxygxm7pkl8j/v2). The key differences between this protocol and the manufacturer's protocol are lower annealing temperatures during end repair and PCR. Libraries were pooled and sequenced PE50 on a NovaSeq 6000 in a single SP flow cell to yield ~10M read pairs per sample (Supplementary Table 10).

CUT&RUN analysis

Raw CUT&RUN sequencing reads were trimmed using Trimmomatic 0.39[37], then mapped to each reference genome using bwa 0.7.17[38]. Duplicate reads were removed using picard[43]. Bigwig files for visualization were generated from these filtered BAM files using RPKM normalization in deepTools2[45]. We called peaks using MACS3[56] for each sample and merged peak calls from biological replicates using bedtools[46], requiring reciprocal 20% overlap between at least two of three replicates. Finally, we combined peaksets from different samples using bedtools merge to generate a comprehensive peakset that we used to identify differentially bound peaks (DBPs) using DiffBind 3.8.4[47,48]. We performed all pairwise comparisons between sexes, genotypes, and stages, then corrected *p*-values using the Benjamini-Hochberg method[49]. Only DBPs with global FDR < 0.05 were used in downstream analyses.

We identified enriched motifs and annotated the full Dsx peakset relative to gene models using HOMER 4.11.0 [50]. We also identified transcription factor binding site motifs enriched in *dsx* peaks (summits +- 100 bp) using HOMER's findMotifsGenome.pl script with default parameters.

To identify potential TFBSs in *Drosophila melanogaster doublesex*, we downloaded ChIP-seq peak calls for all 546 DNA binding proteins assayed by the modENCODE project in whole flies[18], then intersected those data with the r6.55 *doublesex* gene model (plus 1 kb) using bedtools intersect[46].

CRISPR/Cas9

We designed pairs of guide RNAs (gRNAs) to flank target ATAC peaks using Integrated DNA Technology's (IDT's) gRNA Design Tool, then purchased as 2 nmol single gRNAs. sgRNAs were resuspended to 1 ug/uL (32 nM) in water. Injection mixes with two sgRNAs consisted of 125 ng/uL each sgRNA and 500 ng/uL IDT SpCas9 V3 in PBS. Reagents were mixed in a PCR tube, incubated at 37°C for 10 min to complex, then stored at -80°C in small aliquots until use.

We allowed females to lay on a fresh *Citrus* shrub for 1 - 3 hours, then collected eggs for injections. Eggs were aligned on a small strip of double-sided tape on a glass slide. We injected a small amount of injection mix into the bottom of each egg using pulled borosilicate needles. The double sided tape was then placed directly onto fresh *Citrus* leaves, where the eggs were allowed to hatch and develop. Results are shown in Supplementary Table 9. We clipped wings from adults and extracted gDNA from thorax using the QIAgen QIAwave Blood and Tissue Kit. DNA extractions followed the manufacturer's instructions except we incubated tissue in buffer ATL/proteinase K overnight at 56°C shaking at 600 rpm in a thermomixer.

RNA-seq analysis

We re-analyzed SE50 RNA-seq data from VanKuren et al. (2023) following their pipeline. Sequencing data are available in NCBI BioProject PRJNA882073. We quantified transcript expression levels using the raw reads and Salmon 1.9.0[58]. We used k-mer size of 23 and quantified against the transcripts from the mimetic *P. alphenor* assembly, using the whole genome sequence as a decoy. We allowed Salmon to correct for GC, positional, and sequence bias. We then loaded gene-level quantifications using tximport[59] and identified differentially expressed genes (DEGs) between mimetic and non-mimetic females at each developmental stage. We also identified genes with significantly different developmental expression profiles using maSigpro[60,61] following our previously described protocol[13], keeping significant genes with fit correlations >= 0.9.

References

1. Joron, M. *et al.* A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. *PLoS Biol.* **4**, 1831–1840 (2006).
2. Wang, J. *et al.* A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668 (2013).
3. Li, J. *et al.* Genetic architecture and evolution of the S locus supergene in Primula vulgaris. *Nat Plants* **2**, 16188 (2016).
4. Fisher, R. A. *The Genetical Theory of Natural Selection*. 308 (Clarendon Press, 1930).
5. Kunte, K. *et al.* Doublesex is a mimicry supergene. *Nature* **507**, 229–232 (2014).
6. Nishikawa, H. *et al.* A genetic mechanism for female-limited Batesian mimicry in Papilio butterfly. *Nat. Genet.* **47**, 405–409 (2015).
7. Zhang, W., Westerman, E., Nitzany, E., Palmer, S. & Kronforst, M. R. Tracing the origin and evolution of supergene mimicry in butterflies. *Nat. Commun.* **8**, 1269 (2017).
8. Martinez-Ruiz, C. *et al.* Correction: Genomic architecture and evolutionary antagonism drive allelic expression bias in the social supergene of red fire ants. *Elife* **9**, (2020).
9. Mather, K. The Genetical Architecture of Heterostyly in Primula sinensis. *Evolution* **4**, 340–352 (1950).
10. Küpper, C. *et al.* A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83 (2016).
11. Charlesworth, D. & Charlesworth, B. Theoretical genetics of Batesian mimicry II. Evolution of supergenes. *J. Theor. Biol.* **55**, 305–324 (1975).
12. Iijima, T. *et al.* Parallel evolution of Batesian mimicry supergene in two Papilio butterflies, P. polytes and P. memnon. *Science Advances* **4**, eaao5416 (2018).
13. VanKuren, N. W. *et al.* Acute and Long-Term Consequences of Co-opted doublesex on the Development of Mimetic Butterfly Color Patterns. *Mol. Biol. Evol.* **40**, (2023).
14. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
15. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213 (2013).
16. Yoda, S. *et al.* Genetic switch in UV response of mimicry-related pale-yellow colors in Batesian mimic butterfly, Papilio polytes. *Sci Adv* **7**, (2021).
17. Romero-Pozuelo, J. *et al.* Cooperation of axial and sex specific information controls Drosophila female genitalia growth by regulating the Decapentaplegic pathway. *Dev. Biol.* **454**, 145–155 (2019).
18. Nègre, N. *et al.* A cis-regulatory map of the Drosophila genome. *Nature* **471**, 527–531 (2011).
19. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN chromatin profiling tools. *Elife* **8**, (2019).
20. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J.* **8**, 289–317 (2016).
21. Komata, S. *et al.* Functional unit of supergene in female-limited Batesian mimicry of Papilio polytes. *Genetics* (2022) doi:10.1093/genetics/iyac177.
22. Clough, E. *et al.* Sex- and Tissue-Specific Functions of Drosophila Doublesex Transcription Factor Target Genes. *Dev. Cell* **31**, 761–773 (2014).
23. Dudchenko, O. *et al.* De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
24. Dierckxsens, N., Mardulyn, P. & Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**, (2017).

25. Haas, B. J. *et al.* Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7 (2008).
26. Nallu, S. *et al.* The molecular genetic basis of herbivory between butterflies and their host plants. *Nat Ecol Evol* **2**, 1418–1427 (2018).
27. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
28. Pertea, M. *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
29. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
30. Haas, B. J. *et al.* Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).
31. Slater, G. S. C. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**, 31 (2005).
32. Lomsadze, A., Burns, P. D. & Borodovsky, M. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res.* **42**, e119 (2014).
33. Majoros, W. H., Pertea, M. & Salzberg, S. L. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878–2879 (2004).
34. Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).
35. Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P. & Huerta-Cepas, J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021).
36. Shumate, A. & Salzberg, S. L. Liftoff: accurate mapping of gene annotations. *Bioinformatics* **37**, 1639–1643 (2021).
37. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
38. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
39. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).
40. Knight, P. A. & Ruiz, D. A fast algorithm for matrix balancing. *IMA J. Numer. Anal.* **33**, 1029–1047 (2013).
41. Pockrandt, C., Alzamel, M., Iliopoulos, C. S. & Reinert, K. GenMap: ultra-fast computation of genome mappability. *Bioinformatics* **36**, 3687–3692 (2020).
42. Lewis, J. J. & Reed, R. D. Genome-Wide Regulatory Adaptation Shapes Population-Level Genomic Landscapes in Heliconius. *Mol. Biol. Evol.* **36**, 159–173 (2019).
43. Picard. *Picard Tools*. (Broad Institute).
44. Ou, J. *et al.* ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics* **19**, 169 (2018).
45. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–5 (2016).
46. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
47. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
48. Stark, R. & Brown, G. DiffBind: differential binding analysis of ChIP-Seq peak data. *R package version* (2011).
49. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* **57**, 289–300 (1995).
50. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-

regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
51. Lu, S. *et al.* Chromosomal-level reference genome of Chinese peacock butterfly (Papilio bianor) based on third-generation DNA sequencing and Hi-C analysis. *Gigascience* **8**, (2019).
52. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
53. Hackl, T., Duponchel, S., Barenhoff, K., Weinmann, A. & Fischer, M. G. Virophages and retrotransposons colonize the genomes of a heterotrophic flagellate. *Elife* **10**, (2021).
54. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, (2017).
55. Prakash, A. & Monteiro, A. Cell Dissociation from Butterfly Pupal Wing Tissues for Single-Cell RNA Sequencing. *Methods Protoc* **3**, (2020).
56. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137–R137 (2008).
57. Mazo-Vargas, A. *et al.* Deep cis-regulatory homology of the butterfly wing pattern ground plan. *Science* **378**, 304–308 (2022).
58. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
59. Soneson, C., Love, M. I. & Robinson, M. D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res.* **4**, 1521 (2015).
60. Conesa, A., Nueda, M. J., Ferrer, A. & Talón, M. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* **22**, 1096–1102 (2006).
61. Nueda, M. J., Tarazona, S. & Conesa, A. Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. *Bioinformatics* **30**, 2598–2602 (2014).