mirror determines the far posterior domain in butterfly wings

Martik Chatterjee[1,2,3,*], Xin Y. Yu[1,3], Noah K. Brady[1], Gabriel C. Hatto[1], Robert D. Reed[1,**]

[1]Dept. of Ecology and Evolutionary Biology, Cornell University, Ithaca NY-14853

[2]Dept. of Molecular Biology and Genetics, Cornell University, Ithaca NY-14853

[3]Authors contributed equally

*Corresponding author: mc2548@cornell.edu

**Corresponding author: robertreed@cornell.edu

ABSTRACT

Insect wings, a key innovation that contributed to the explosive diversification of insects, are recognized for their remarkable variation and many splendid adaptations. Classical morphological work subdivides insect wings into several distinct domains along the antero-posterior (AP) axis, each of which can evolve relatively independently. There has been little molecular evidence, however, for AP subdivision beyond a single compartment boundary described from *Drosophila melanogaster*. Here we show that the transcription factor *mirror* acts as a selector gene to differentiate a far posterior domain in the butterfly wing, classically defined as the vannus, and has wide-ranging effects on wing shape, scale morphology, and color pattern. Our results confirm that insect wings can have more than one posterior developmental domain, and support models of how selector genes may facilitate evolutionarily individuation of distinct AP domains in insect wings. Our results also suggest that the alula, a small *mirror*-dependent structure at the base of the *D. melanogaster* wing, may be an evolutionary derivative of the vannus, and therefore that the *D. melanogaster* wing blade is a solitary remigium that represents only a fraction of the archetypal insect wing.

RESULTS AND DISCUSSION

Insect wings vary dramatically in their size and morphology. Butterflies and moths in particular have attracted attention for their rapidly evolving wings that show extreme variation in shape and color pattern. Interestingly, lepidopteran color pattern diversity appears to follow some simple rules where basic spot and stripe color pattern elements change size, color, and presence/absence relatively independently based on which domain along the antero-posterior (AP) axis they occur in.[3] Analyses of wing pattern diversity across butterflies, considering both natural variation and genetic mutants, suggest that wings can be subdivided into at least five AP domains, bounded by the M1, M3, Cu2, and 2A veins respectively, within each of which there are strong correlations in color pattern variation and wing morphology (Figure 1A).[4-6] The anterior-most of these domains, bordered by the M1 vein, appears to correspond to an AP compartment boundary originally described by cell lineage tracing in *Drosophila melanogaster*,[7] and later supported in butterfly wings by expression of the Engrailed transcription factor.[8] Interestingly, however, *D. melanogaster* work has yet to reveal clear evidence for

additional AP domain boundaries in the wing. It has thus been a matter of significant interest and controversy whether additional AP wing domains might occur in other insects, [9-11] as their existence would provide an explanatory model for the modular diversification of insect wing morphology.

The goal of this study was to identify the molecular basis of AP domain specification in the butterfly wing, posterior of the Engrailed boundary. We initially identified the transcription factor *mirror* as a potential candidate gene based on previous mRNA-seq studies,[12] coupled with *D. melanogaster* work showing that *mirror* specifies the alula – a small membranous lobe at the posterior base of the wing in some dipterans (Figure 1C,D).[2,13] We used *in situ* hybridization to visualize *mirror* mRNA localization in the last-instar imaginal discs of the common buckeye butterfly, *Junonia coenia*, and observed *mirror* expression throughout the wing domain posterior to the 2A vein boundary (Figures 2A and S1). This expression precisely marks the vannus, or anal region (Figure 1E), of the butterfly wing – a distinct posterior field of the wing blade populated by the anal (A) veins, and bordered anteriorly by the claval fold between the cubital (Cu) and A veins in the hindwing. This region was also previously proposed to represent the posterior-most color pattern domain in butterflies (Figure 2B).[3]

To functionally assess the role of *mirror* in wing development, we produced CRISPR/Cas9 deletion mosaic knockouts (mKOs) targeting the coding region of *J. coenia mirror*. The resulting mutants revealed vannus-specific phenotypes, consistent with the *mirror* expression patterns. The unique identity of the vannus was lost in mutants, so that the vannus resembled the remigium (Figure 2C-E; Figure S2). In brief, the vannus took on a morphology consistent with anterior wing domains, including loss or reduction of the vannal fold (along with an overall shape along the posterior wing margin), reduction of light silvery iridescent coloration, and continuation of normally truncated wing margin color patterns along the posterior distal margin of the wing (Figure 2E; Figure S2). Some mutants also displayed aberrations in vannal venation, including loss of the 3A vein, discontinuous and bifurcated 2A veins, and ectopic veins (Figures 2C-E; Figures S2 and S3). Together, our expression and knockout data show that *mirror* acts as a selector gene necessary for defining the identity of the far posterior wing domain that corresponds to the vannus.

Our finding is of interest for several reasons. First, it provides a molecular explanation for how different AP domains of the butterfly wing may be individuated to independently evolve their own shape and color pattern variations. Previous authors have proposed the existence of such individuated domains, and speculated that they may be specified by selector genes.[5,10] Our data provide experimental support for this model, and now motivate us to identify factors that specify other domain boundaries between the M1 and A2 veins. Second, the role of *mirror* in specifying the vannus provides an important link between molecular developmental biology and Snodgrass' classical anatomical designations of the insect wing fields – i.e., the remigium, the vannus, and the jugum (Figure 1E).[14] Importantly, *mirror* knockdown in the milkweed bug *Oncopeltus fasciatus* causes developmental defects in the claval and anal furrows in the posterior wing,[15] which leads us to infer that *mirror*'s role in determining the vannus may be deeply conserved in insects.

Finally, the function of *mirror* in determining the alula in *D. melanogaster*[2] suggests that the alula may represent an evolutionarily reduced vannus. The ramifications of this are significant for reconstructing the history of the insect wing, as they suggest that the *D. melanogaster* wing blade is a

lone remigium – only one-third of the archetypal insect wing (Figure 1E). The dipteran jugum appears to have been lost entirely, perhaps except as a calypter in the Calypterata lineage, as speculated by Snodgrass.[14] Thus, while flies have been an important model for characterizing genes and processes that build insect wings, a more comprehensive understanding of the development and evolution of insect wings will require work in species that have wings more representative of the complete ancestral blueprint.

ACKNOWLEDGMENTS

We thank Kate Siegel, Nigel Williams, and Rick Fandino for assistance with rearing butterflies; Johanna Dela Cruz for help with confocal microscopy at Biotechnology Resource Center (BRC) Imaging Facility (RRID:SCR_021741) at the Cornell Institute of Biotechnology. We also thank Dr. Eirene Markenscoff-Papadimitriou for use the Leica Stellaris 5 confocal microscope for imaging. This work was supported by the United States National Science Foundation grants NSF IOS-1753559 and IOS-2128164 awarded to R.D.R., and a Cornell Summer Experience Grant and an Office of Undergraduate Biology fellowship to X.Y.

MATERIALS AND METHODS

Identity of *Iroquois* family gene *mirror* ortholog in *J. coenia*

We performed a reciprocal BLAST of amino acid sequences of *D. melanogaster Iroquois* complex genes *araucan*, *caupolican*, and *mirror* to identify their orthologs in the latest *J. coenia* and *Heliconius erato lativitta* genome sequences on lepbase.org and *Tribolium castaneum* and *Apis mellifera* on NCBI. The amino acid sequences of the top hits – from *J. coenia* (JC_02265-RA, JC_02269-RA), *H. erato lativitta* (HEL_009655-RA, HEL_009656-RA), *A. mellifera* (LOC412840, LOC412839) and *T. castaneum* (LOC652944, LOC660345) were aligned to the *Iroquois* genes of *D. melanogaster* using MUSCLE.[16] We used these alignments to build a maximum likelihood gene phylogeny on IQ-TREE 2[17] using ModelFinder[18] for estimating the most accurate substitution model for our amino acid sequences. We used FigTree to visualize the tree depicted in Figure S4.

HCR fluorescent *in situ* hybridization of *mirror* mRNA

This protocol was adapted and modified from Bruce et al. (dx.doi.org/10.17504/protocols.io.bunznvf6). The coding sequence of JC_02269-RA was used by Molecular Instruments, Inc. (Los Angeles, CA, USA) to design a *mirror*-specific hybridization chain reaction (HCR) probe set (Lot #PRQ901).
The probes are unique to the JC_02269-RA transcript as verified by BLAST search using the latest *J. coenia* genome assembly (v2) on lepbase.org, and do not share significant sequence similarity with any other transcripts, including the other *Iroquois* family member JC_02265-RA (see Figure S4).

Last (5th) instar larval wing imaginal discs were dissected in cold 1x PBS. The dissected discs collected from left and ride sides of the larvae were randomly assigned to control or treatment batches. Both controls and treatment batches were fixed in 1x cold fix buffer (750uL PBS, 50mM EGTA, 250uL 37% formaldehyde) on ice. The discs were then washed three times in 1x PBS with 0.1% Tween 20 (PTw) on ice spending approximately 30sec – 2min for each wash. Tissues were then gradually dehydrated by washing in 33%, 66%, and 100% MeOH (in PTw) for 2-5min / wash on ice, and stored in 100% MeOH in -20°C for days until HCR protocol was started. On the day of the HCR protocol, the wing discs were progressively rehydrated in cold 75% MeOH, 50% MeOH, and 25% MeOH in PTw spending 2-5min per wash. After rehydrating, discs were washed once for 10min and twice for 5min with PTw. The discs were then permeabilized in 300-500µL of detergent solution (1.0% SDS, 0.5% Tween, 50mM Tris-HCL pH 7.5, 1mM EDTA pH 8.0, 150mM NaCl) for 30min at room temperature. While waiting, hybridization buffer (Molecular Instruments) was warmed to 37°C (200µL/tube). After permeabilizing with detergent solution, each tube of wing discs was incubated in 200µL of pre-warmed hybridization buffer for 30min at 37°C. Probe solution was prepared by adding 0.8pmol (0.8 µl of probe from 1uM stock solution) of each probe mixture to 200µL of probe hybridization buffer at 37°C. The pre-hybridization buffer was removed, and the probe solution was added. For negative control tissues, hybridization buffer without probes was added instead. All wing discs were incubated overnight (12-16hr) at 37°C. Before resuming the protocol, probe wash buffer (Molecular Instruments) was warmed to 37°C, amplification buffer (Molecular Instruments) was calibrated to room temperature, and a heat block was set to 95°C. The probe solution was removed and saved at -20°C to be reused. Wing discs were then washed four times in 1mL pre-warmed probe wash buffer at 37°C for 15min per wash. After the last wash step, discs were washed twice (5min/wash) with 1mL of 5x SSCT (5X sodium chloride sodium citrate, 0.1% Tween 20) at room temperature. The wing discs were pre-amplified with 1mL of equilibrated amplification buffer for 30min at room temperature. During this step, 2µL of each of hairpin h1 and 2uL of each hairpin h2 were mixed in 100µL of amplification buffer at 95°C for 90sec and cooled to room temperature in the dark for 30 min. For *mirror*-only *in situ* hybridizations, we used B1 hairpin amplifiers tagged with AlexaFluor 594 fluorophore (Molecular Instruments). For *mirror* and *wingless* double stains, we used *mirror* probes with B1 amplifier tagged with AlexaFluor 647 fluorophore and *wingless* probes (Lot #PRG129) with B3 hairpin amplifiers tagged with AlexaFluor 546 fluorophore. After 30min, the pre-amplification buffer was removed from the discs, hairpin solution added, and the tissues were then incubated in the dark at room temperature for 2-16hr. The hairpins were removed and saved at -20°C to be reused later. Excess hairpins were removed by washing 5 times (twice for 5min, twice for 30min and once for 5min) with 1mL of 5x SSCT at room temperature. The wing discs were then incubated in 50% glycerol solution (in 1X PBS) with DAPI (0.01µg/mL) overnight at 4°C. The wing discs were then mounted and visualized on a Zeiss 710 or Leica Stellaris 5 confocal microscope.

CRISPR-Cas9 mediated mutagenesis of *mirror* in *J. coenia*

Two single-guide RNAs (sgRNAs; sgRNA1- 5'-GAATGGACTTGAACGGGGCA; sgRNA2- 5'-AGAAACAGGGTCGATGATGA) targeting the homeobox domain of *mirror* were mixed with 500ng/µl of Cas9 nuclease and injected to *J. coenia* eggs 0.5-4hr after oviposition (n = 1042 eggs) as previously described. [19] Injected G0 individuals were reared on standard artificial diet, until they emerged and were immediately frozen in -20˚C upon emergence. Wings of surviving G0 adults (n=99) were assayed for anomalies to detect mosaic knockout (mKO) phenotypes (n=29). Injection results and mutation phenotype frequencies are detailed in Table S1.

Validating *mirror* mutants by genotyping

DNA was extracted using E.Z.N.A. Tissue DNA Kit (Omega Bio-Tek) from the thorax of individuals that showed mutations in the wings. Extracted genomic DNA was amplified using a pair of primers flanking the sgRNA cut sites (Forward: 5'-CGCTTGTGCCCACCTTAAAC, Reverse: 5'-GTATGGCTCGGGGGATTCTG). Amplified DNA was run on a 2% agarose gel and were excised and purified using the MicroElute Cycle-Pure Kit (Omega Bio-Tek). Purified DNA was Sanger sequenced by Cornell Institute of Biotechnology. Example mutant alleles of *mirror* are shown in Figure S5.

AUTHOR CONTRIBUTIONS

Conceptualization, M.C and R.D.R; Methodology, M.C and R.D.R; Investigation, X.Y.Y, M.C, N.K.B and G.C.H; Writing- Original Draft, M.C, X.Y.Y and R.D.R; Writing- Review & Editing, M.C and R.D.R; Visualization, M.C and X.Y.Y; Funding Acquisition, R.D.R and X.Y.Y; Supervision, M.C and R.D.R

REFERENCES

1. McAlpine, J., Peterson, B., Shewell, G., Teskey, H., Vockeroth, J., and Wood, D. (1981). Manual of Nearctic Diptera. Volume 1.
2. Kehl, B.T., Cho, K.-O., and Choi, K.-W. (1998). mirror, a Drosophila homeobox gene in the Iroquois complex, is required for sensory organ and alula formation. Development *125*, 1217-1227.
3. Nijhout, H.F. (1991). The development and evolution of butterfly wing patterns (Smithsonian Institution).
4. Monteiro, A., Chen, B., Ramos, D.M., Oliver, J.C., Tong, X., Guo, M., WANG, W.K., Fazzino, L., and Kamal, F. (2013). Distal-L ess Regulates Eyespot Patterns and Melanization in Bicyclus Butterflies. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution *320*, 321-331.
5. McKenna, K.Z., Kudla, A.M., and Nijhout, H.F. (2020). Anterior–posterior patterning in lepidopteran wings. Frontiers in Ecology and Evolution *8*, 146.

6. Banerjee, T.D., and Monteiro, A. (2020). Molecular mechanisms underlying simplification of venation patterns in holometabolous insects. Development *147*, dev196394.
7. Guillén, I., Mullor, J.L., Capdevila, J., Sánchez-Herrero, E., Morata, G., and Guerrero, I. (1995). The function of engrailed and the specification of Drosophila wing pattern. Development *121*, 3447-3456.
8. Carroll, S.B., Gates, J., Keys, D.N., Paddock, S.W., Panganiban, G.E., Selegue, J.E., and Williams, J.A. (1994). Pattern formation and eyespot determination in butterfly wings. Science *265*, 109-114.
9. Beldade, P., and Brakefield, P.M. (2002). The genetics and evo–devo of butterfly wing patterns. Nature Reviews Genetics *3*, 442-452.
10. Abbasi, R., and Marcus, J.M. (2017). A new AP compartment boundary and organizer in holometabolous insect wings. Scientific reports *7*, 1-11.
11. Lawrence, P.A., Casal, J., Celis, J.F.d., and Morata, G. (2019). A refutation to 'A new AP compartment boundary and organizer in holometabolous insect wings'. Scientific Reports *9*, 1-6.
12. Hanly, J.J., Wallbank, R.W., McMillan, W.O., and Jiggins, C.D. (2019). Conservation and flexibility in the gene regulatory landscape of heliconiine butterfly wings. Evodevo *10*, 1-14.
13. Foronda, D., Pérez-Garijo, A., and Martín, F.A. (2009). Dpp of posterior origin patterns the proximal region of the wing. Mechanisms of development *126*, 99-106.
14. Snodgrass, R.E. (1935). Principles of insect morphology (McGraw-Hill Book Company Inc).
15. Fisher, C.R., Kratovil, J.D., Angelini, D.R., and Jockusch, E.L. (2021). Out from under the wing: reconceptualizing the insect wing gene regulatory network as a versatile, general module for body-wall lobes in arthropods. Proceedings of the Royal Society B *288*, 20211808.
16. Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics *5*, 1-19.
17. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Molecular biology and evolution *37*, 1530-1534.
18. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., Von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods *14*, 587-589.
19. Zhang, L., and Reed, R.D. (2017). A practical guide to CRISPR/Cas9 genome editing in Lepidoptera. Diversity and evolution of butterfly wing patterns: An integrative approach, 155-172.

FIGURES

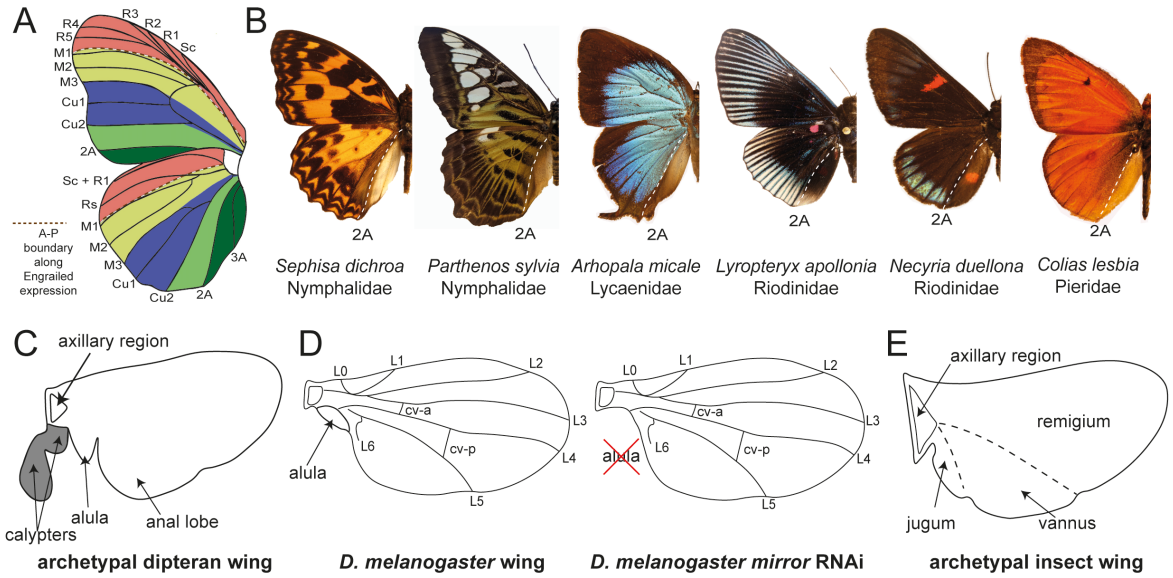

Figure. 1. Proposed AP domains of insect wings. (A) Comparative morphology has identified boundaries associated with color pattern variation demarcated by the M1, M3, Cu2, and 2A veins, as summarized in McKenna et al.[5] (B) Examples of the 2A vein marking a clear posterior color pattern boundary across multiple butterfly families. (C) Proximal posterior features of the archetypal dipteran wing include the calypters, which are only found in Calypterata clade, and the alula, a lobe at the base of the wing blade.[1] (D) *mirror* RNAi knockdowns result in highly specific loss of the alula in *D. melanogaster*.[2] (E) Snodgrass' model of the archetypal insect wing specifies three major domains along the anteroposterior axis: the remigium, the vannus, and the jugum. Homologies of these domains with the features of the dipteran wing (C), or butterfly wing pattern boundaries (A), has remained unclear.

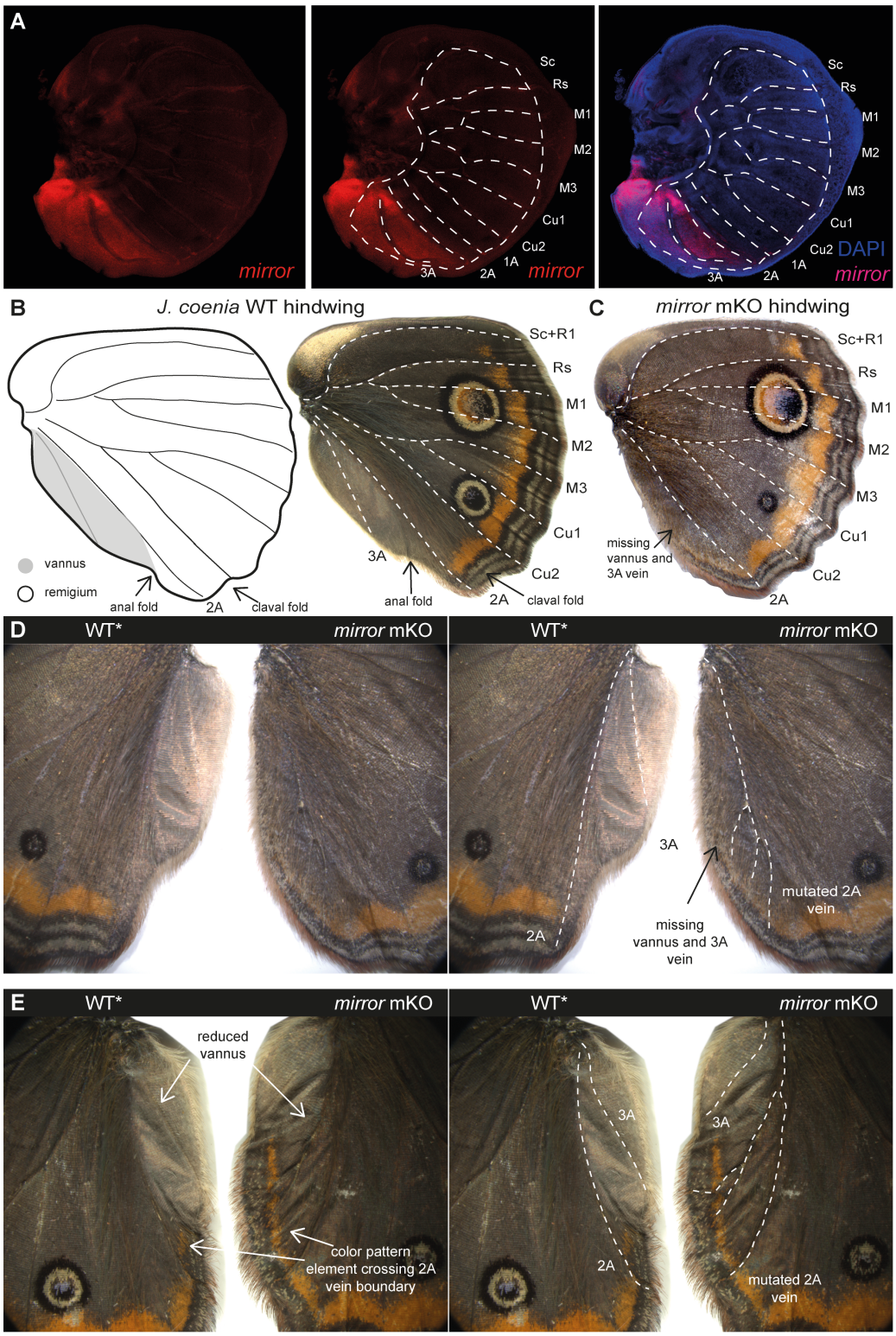

Figure 2. *mirror* determines the identity of the far posterior vannus domain in *J. coenia*. (A) In *J. coenia mirror* (red) is expressed posterior to the 2A vein in the late last-instar hindwing imaginal discs, marking the region that will develop into the vannus. See Figure S1 for forewing expression and controls. (B) In adult butterfly wings the vannus is the wing field posterior to the 2A vein, as shown here in *J. coenia*. (C) Targeted mosaic knockouts (mKOs) of *mirror* results in loss of vannal identity in *J. coenia*, including (D) reduction of the anal fold, and (E) continuation of margin color patterns along the posterior wing margin. Some *mirror* mKOs (C-E) also show mutation in the 2A vein along with loss (D) or reduction (E) of the 3A vein, likely associated with mosaic clone boundaries. mKOs result in asymmetric effects observed in the wings where (D,E) one wing resembles a wild-type wing (WT*) more than the other wing from the respective individual. For more *mirror* mKOs see Figures S2 and S3.

SUPPLEMENTAL INFORMATION

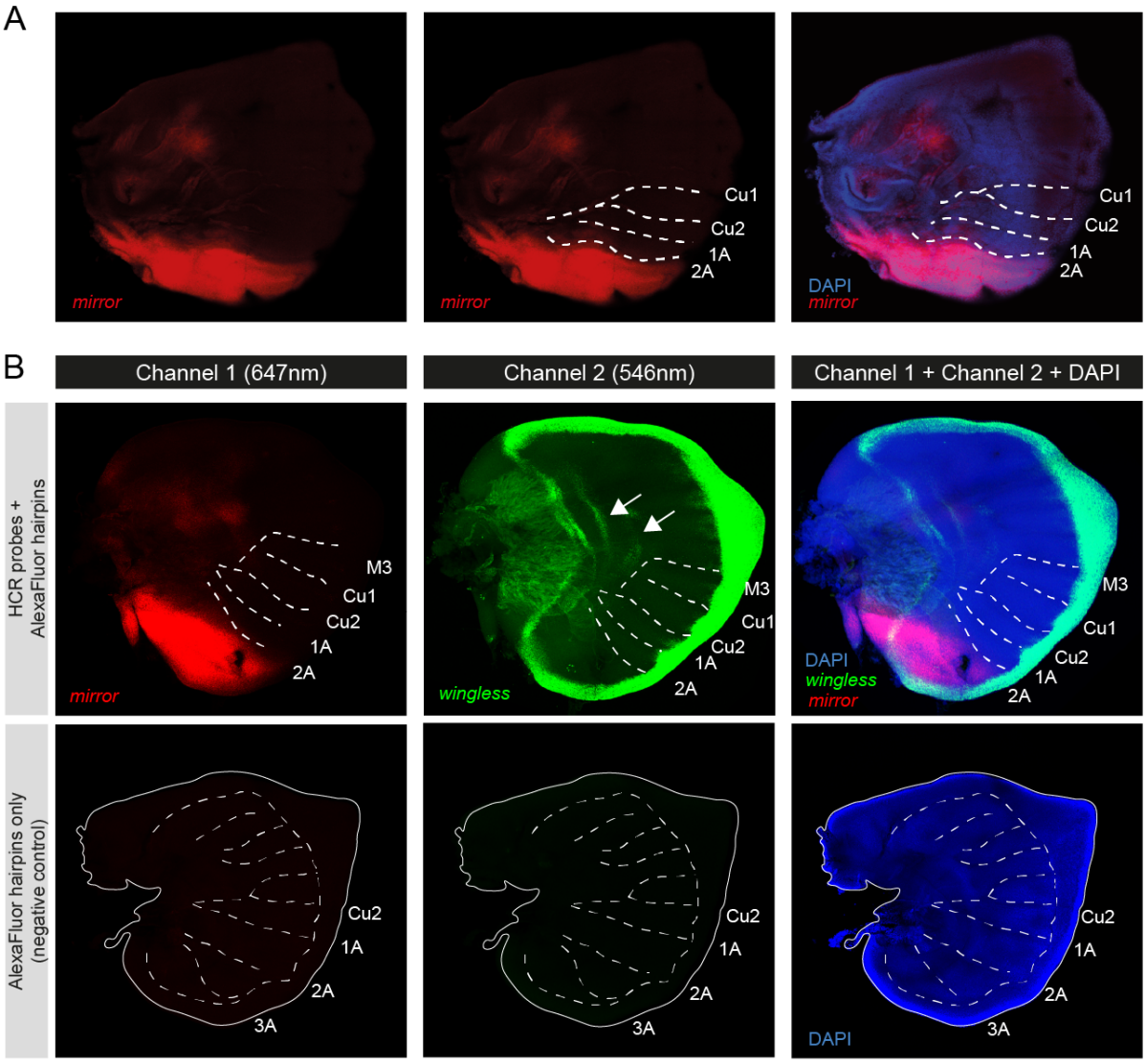

Figure S1. HCR *in situ* hybridization of *mirror* highlight posterior domain expression in *J. coenia* wing imaginal discs. Related to Figure 2A. (A) *mirror* is expressed posterior to the 2A vein in the last-instar forewing imaginal disc of *J. coenia.* (B) Double *in situ* of *mirror* (Channel 1: 647nm) and *wingless* (Channel 2: 546 nm) in last instar *J. coenia* hindwing imaginal discs. Expression of *wingless* observed by HCR precisely matches expression in wing margin peripheral tissue and discal bands (white arrow) as previously described from chromogenic in situs.[5] Bottom panel: Negative controls, lacking HCR probes, showed no signal in the 647nm and 546nm channels. Dashed lines show position of posterior veins.

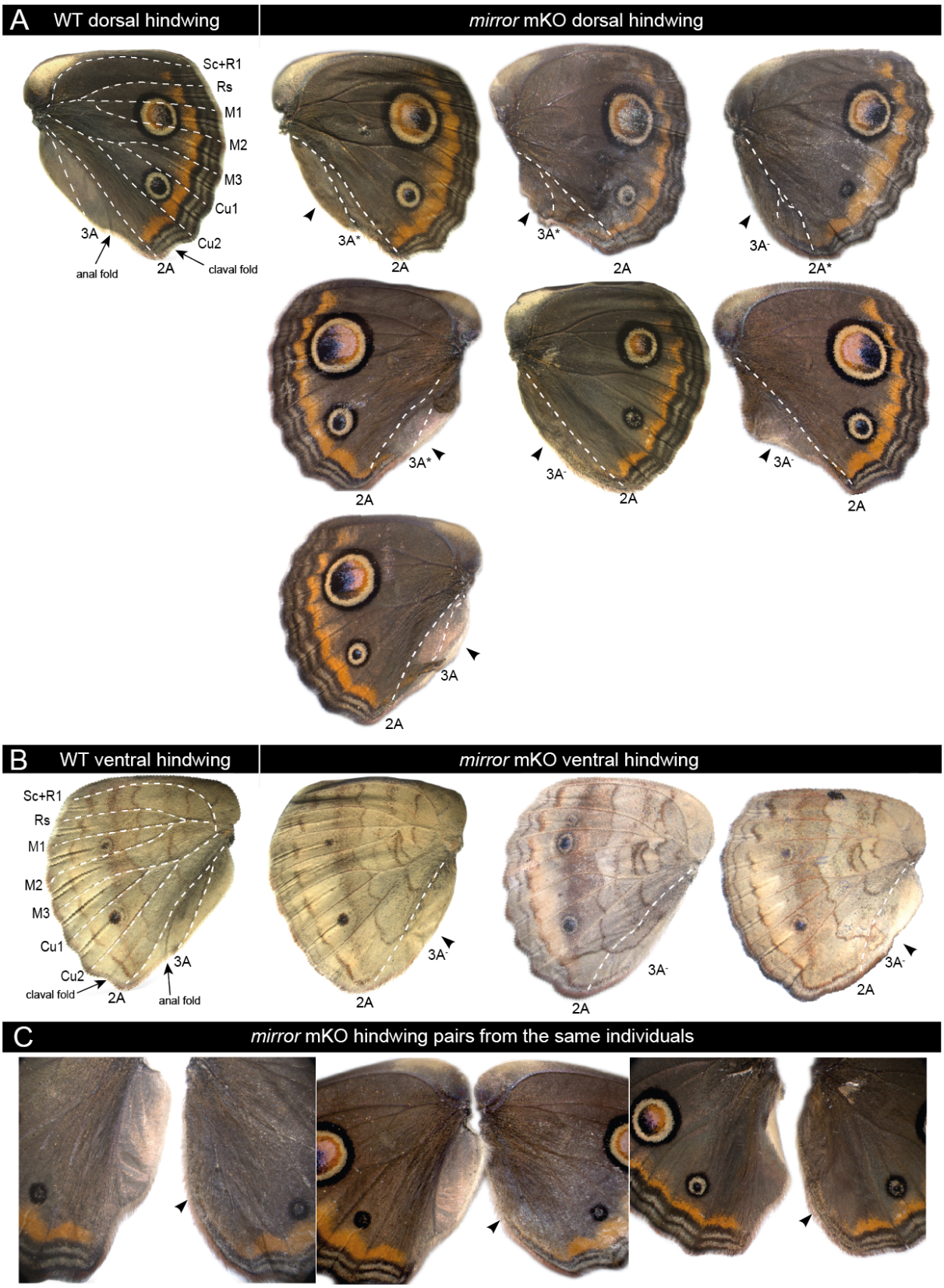

Figure S2. **mKOs of *mirror* result in partial or entire loss of vannus identity in *J. coenia* hindwings**. **Related to Figure 2B-E.** Arrows highlight vannal reduction-related phenotypes in the (A) dorsal side and (B) ventral side of the hindwings. Dashed lines show positions of veins. Next to vein labels, asterisks denote vein anomalies, and "-" denote missing veins. (C) Comparison of left and right dorsal hindwings from the same individual, with vannus reduction phenotype highlighted with arrows.

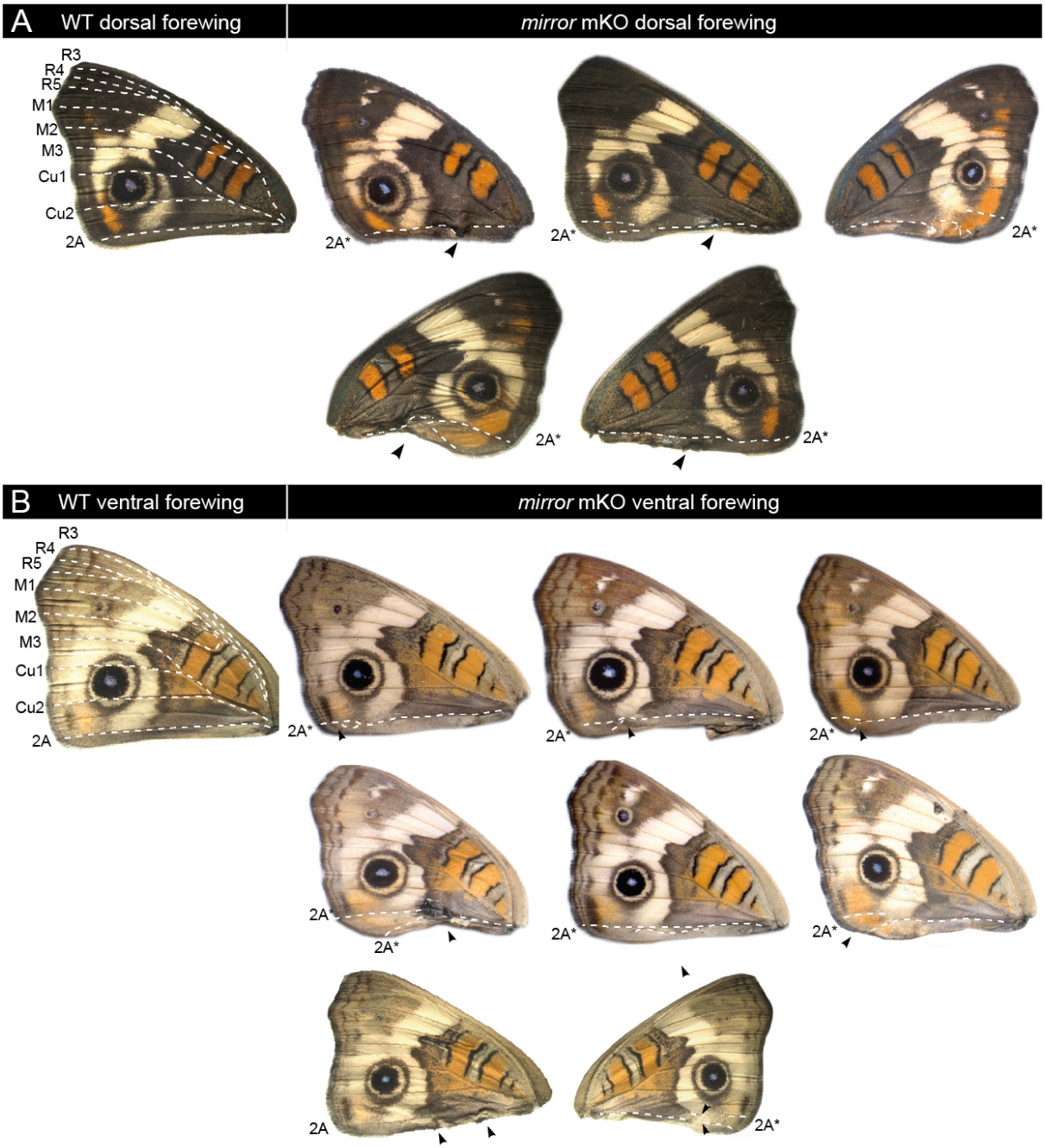

Figure S3. ***mirror* mKOs show post-2A morphological anomalies *J. coenia* forewings. Related to Figure 2B-E.** mKOs of *mirror* result in anomalies along the forewing 2A vein, including vein bifurcation and ectopic vein fragments, which are often associated with color pattern disruption. Dashed lines show the position of 2A vein; asterisks denote 2A vein anomalies.

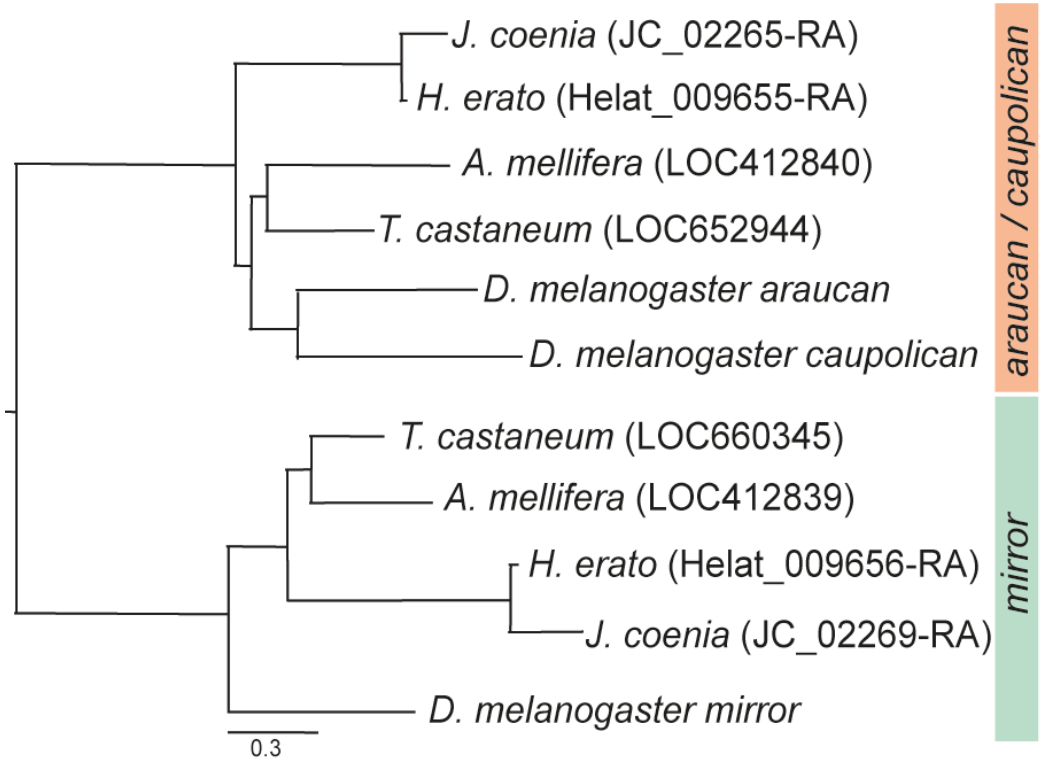

Figure S4. Identification of *mirror* ortholog in *J. coenia*. A maximum likelihood phylogeny of *J. coenia*, *H. erato lativitta*, *Tribolium castaneum*, *Apis mellifera*, and *D. melanogaster Iroquois* complex genes confirms that JC_02269-RA is the ortholog of mirror.

Individuals	Indels	sg RNA1 PAM
mKO 1	WT	GAATGGACTTGAACGGGGCAAGGAGAAAAAATGCAACCAGAGAGACCACA
	-7	GAATGGACTTGAACGG----------AGAAAAAATGCAACCAGAGAGACCACA
	-2	GAATGGACTTGAACGGG----AAGGAGAAAAAATGCAACCAGAGAGACCACA
	+1	GAATGGACTTGAACGG--NGCAAGGAGAAAAAATGCAACCAGAGAGACCAC
	-2	GAATGGACTTGAACGG----CAAGGAGAAAAAATGCAACCAGAGAGACCACA
mKO 2	WT	GAATGGACTTGAACGGGGCAAGGAGAAAAAATGCAACCAGAGAGACCACA
	-7	GAATGGACTTGAACGG----------AGAAAAAATGCAACCAGAGAGACCACA
	-2	GAATGGACTTGAACGG----CAAGGAGAAAAAATGCAACCAGAGAGACCACA
	-2	GAATGGACTTGAACGGG----AAGGAGAAAAAATGCAACCAGAGAGACCACA
	+4	GAATGGACTTGAACGGGNNNNCAAGGAGAAAAAATGCAACCAGAGAGACC
mKO 3	WT	GAATGGACTTGAACGGGGCAAGGAGAAAAAATGCAACCAGAGAGACCACA
	-6	GAATGGACTTGAACGGG----------AGAAAAAATGCAACCAGAGAGACCACA
	-7	GAATGGACTTGAACGG----------AGAAAAAATGCAACCAGAGAGACCACA
	-3	GAATGGACTTGAACGG------AAGGAGAAAAAATGCAACCAGAGAGACCACA
	-1	GAATGGACTTGAACGG--GCAAGGAGAAAAAATGCAACCAGAGAGACCACA
	-6	GAATGGACTTGAACGG----------GAGAAAAAATGCAACCAGAGAGACCACA
	-2	GAATGGACTTGAACGGG----AAGGAGAAAAAATGCAACCAGAGAGACCACA

Figure S5. Sanger sequencing confirms mutations at CRISPR sgRNA site in *mirror* mKOs. Sanger sequencing results from three different *mirror* mutant individuals. Each line denotes a different wildtype (WT) or mutant allele with indels indicated in red, and the PAM site in green.

Table S1. *mirror* CRISPR/Cas9 injection results.

sgRNA concentration (ng/μL)	Eggs Injected	Hatched	Adults Emerged	Mutants
487 ng/μL	142	25	4	2
391 ng/μL	96	33	6	3
335 ng/μL	163	32	15	4
281 ng/μL	98	11	1	0
175 ng/μL	149	52	19	8
125 ng/μL	252	79	24	9
88.5 ng/μL	142	41	30	3
Total	**1042**	**273**	**99**	**29**