THE DEVELOPMENTAL BASIS AND EVOLUTION OF WING COLOR PATTERN IN BUTTERFLIES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Anyimilehidi Mazo-Vargas

August, 2020

THE DEVELOPMENTAL BASIS AND EVOLUTION OF WING COLOR PATTERN IN BUTTERFLIES

Anyimilehidi Mazo-Vargas, Ph. D.

Cornell University 2020

Evolutionary biologists look to understand the processes underlying biodiversity. One aim of the field is to link the phenotype to its instructive genetic code, and to take advantage of homologies between species to ask what genetic changes generate the diversity of forms and shapes we see in nature. In this work, I study the evolution of wing color patterns in butterflies—a system famous for its striking variation in form and complexity. The specific goals of my work were to test *Wnt* gene functions in patterning wing colors, and to characterize the regulatory architecture modulating color pattern-related function of the *WntA* gene. Our results revealed that *WntA* is essential to establish multiple pattern elements, even in highly divergent species of nymphalids and papilionids–two distant and very diverse families of butterflies. In nymphalids, I observed lineage-specific effects in various wing regions, including spatial shifts in gene expression and novel expression domains, that are associated with color pattern evolution. In papilionids, a basal butterfly family, *WntA* function is restricted to an expanded wing margin system that largely dominates *Papilio* wing patterning. Interestingly, I confirmed that the main pattern systems driving nymphalid and *Papilio* wing diversity, i.e., the *central symmetry system* and the *glauca*,

respectively, are not homologous. In *Papilio*, I identified a new set of *Wnt6*-related wing pattern elements, the *submarginal spots*, that show substantial variation between papilionid species and deserve further characterization in future studies. *WntA*'s versatility and flexibility in patterning a diverse repertoire of elements prompted me to explore this gene's cis-regulatory landscape. I implemented a novel 'shotgun' CRISPR/Cas9 strategy to produce mosaic knock-outs of *WntA* cis-regulatory elements across multiple species. The results contrasted with traditional postulates about gene regulatory modularity, and instead pointed to a highly interdependent network of both ancestral and recently evolved cis-regulatory elements that interact to generate *WntA* patterns. Monarch butterflies stood out from other nymphalids as having a highly divergent cis-regulatory architecture, likely associated with its distinctive *WntA* wing pattern. Finally, the implementation of a 'shotgun' deletion strategy revealed that most CREs have the potential to act as both enhancers or silencers, which is new evidence for our emerging understanding of regulatory element functionality.

BIOGRAPHICAL SKETCH

Anyi was born in Cali, Colombia, to Alba M. Vargas Oviedo and Oscar A. Mazo. Her grandmother and mother were very influential in her view of the natural world. She was a good student, but due to her socioeconomic position there were low expectations for a college degree. It was a dream. Despite the odds, Anyi managed to get into the public higher education system at the Universidad del Valle. In 2005, after finding her enthusiasm for insects, she obtained a bachelor's degree in Biology with a minor in Entomology. The was the first university degree in her family. In Anyi's first year of college, she learned about graduate school and the seed was planted for her to eventually seek out a higher level of education. Following her undergraduate training, Anyi worked at the International Center for Tropical Agriculture (CIAT) in Cali. There, she experienced a research environment and saw it as a life path. The accumulation of experiences and her conviction led her to pursue an MSc in Biology at the University of Puerto Rico, Mayaguez. This program allowed her to understand the significance of graduate school education. She benefitted from valuable mentors who fomented the search for knowledge, while she acquired practical skills in molecular biology, phylogenetics, biogeography, and undergraduate mentoring. Subsequently, Anyi obtained a position as a laboratory manager at Duke University, which was a great experience at a university with very high research activity. In working with yeast genetics and cell cycle biology, she obtained a valuable quantitative and experimental perspective to study biological processes. Her interests in insects, coupled with her new skillset in cell biology and genetics, then combined to motivate her to pursue a Ph.D. degree studying the developmental genetic basis of butterfly color patterns at Cornell University.

Dedicate to mita Rosa y mami Alba, who fought for our family, to give us the chance of a better life.

To my daughter to reinvent and illuminate the future, and to Pedro for following me on this roller coaster.

ACKNOWLEDGMENTS

It has been an incredible honor to work with so many talented people during these last six years. Many thanks to my advisor, Robert Reed, for his support and understanding throughout my Ph.D. Bob was always ready to provide advice and inspire me with his vast wealth of knowledge and positivism. My other committee members, Mariana Wolfner and Charles Danko, for their caring criticism and input, helping me to convey ideas more clearly. I was also privileged to work with collaborators outside of Cornell that were a vital part in the development of my research project. Thus, I would like to thank Arnaud Martin for the opportunity to work in the *WntA* project with a large group of butterfly researchers and for his support and encouragement. To Riccardo Papa, Steven Van Belleghem, and Joe Hanly, for their crucial comments throughout all the stages of this research.

I am very grateful to the Reed lab past and present members for their support, and for providing spaces for discussion and brainstorming ideas, and the shared ice cream. Thanks to Linlin Zhang for being an exceptional role model. Bethany Wasik, for being my editor-in-chief and sharing family moments. Matthew Verosloff for all his help, friendship, and fun moments. To Karin Van der Burg, Noah Brady, Martik Chatterjee, Jeanne McDonald, and Richard A. Fandino for the many conversations and teaching moments. I will also highlight the vital role that undergrads had in me to forge a complete Ph.D. experience and their vast contribution to me as a researcher and mentor. In particular, Ceili Peng, Benjamin Brack, Laura Echávez, Etinosa Osagie, Joshua Kwan, Rachel Geltman, and Caitlyn Robertson. My research work was

possible with the support of the greenhouse staff, notably Trey Ramsey, Joshua A. Manser, and Melissa Lynn Brechner. Also, I am in debt with Agrawal lab members for providing monarchs and milkweed plants, Anurag, Amy Hastings, Tyler Coverdale, and Patty Jordan.

I have been very fortunate to be in the Departments of Entomology and also Ecology and Evolutionary Biology. These departments congregate an excellent academic environment, with many laboratories open to sharing equipment, knowledge, and food when life happens. The graduate students had been inspirational and illuminating; their work in and out of academic settings is an abundant source of inspiration. I would also like to thank faculty that facilitated different opportunities and conversations, Monica Geber, Kelly Zamudio, Jeremy Searle, Katja Poveda, and Jennifer Thaler. I am also grateful to Sara Xayarath Hernandez, Colleen McLinn, Susi Varvayanis, Azucena Ortega, Anitra Douglas-McCarthy, for all the resources and unveiling obvious not obvious.

Finally, the support of family and close friends had been essential for the conclusion of this life chapter. Thanks Mariely, Christian, Lina, Jon, Angela and Ernesto. My mom, mita, papi, Pedro and Luna. Nothing I do would be conceivable without the love and support of my family. They are my inspiration and energy source to keep up with life's challenges.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

CREs	Cis-regulatory elements
sgRNA	single-guide RNA
mKO	Mosaic knock-out
NGP	Nymphalid groundplan
BSS	Basalis symmetry system
D1	Discalis element 1
D2	Discalis element 2
CSS	Central symmetry system
BoSS	Border oceli symmetry system
ESS	Externa symmetry system
dPF	Distal parafocal elements
CDS	Coding sequence

PREFACE

Evolutionary biology is a field that asks about the origin of life, the mechanisms of adaptation, and the process of diversification [1]. One of its aims has been to identify genetic variants – i.e., genotypes that affect phenotypes – upon which natural selection operates. The field of population genetics has exploded in data in recent decades, and there are many successful case studies that link adaptive phenotypic variation to specific genomic loci[2]. But the question still largely remains: how is genotypic information (mutations) decoded to manifest the diversity of phenotypes we observe? This question is a driving motivation of the discipline of evolutionary developmental biology (evo-devo), which seeks to characterize how developmental repatterning occurs during the evolution of organisms[1]. Evo-devo is also growing, especially with the advance of techniques that allow us to use a comparative approach to test long-standing hypothesis about trait homology, gene function, evolutionary novelty, modularity, pleiotropy, and plasticity[3]. In the work presented here, I used a combination of genetic and genomic tools to dissect, in a comparative framework, the function of the gene *WntA* in establishing the diversity of wing patterns in two butterfly families: Nymphalidae (Chapter 1) and Papilionidae (Chapter 3). Then, I went further to explore how *WntA* can pattern the variety of wing designs by examining its evolving cis-regulatory architecture (Chapter 2).

Butterflies and moths (Lepidoptera) are an extraordinary example of diversification and innovation. This megadiverse group of insects has more than 160,000 species, and

wing coloration differentiates many of them[4]. From evolutionary and ecological perspective, wing patterns play essential roles in mimicry, plasticity, mate choice, and thermoregulation[5–7]. Lepidopteran wing color patterns have been a magnet for naturalists, and there are from very early in the taxonomic literature, descriptions of wing design variation, and 'aberrations'. In the early 1900s, two entomologists, Schwanwitsch[8] and Süffert[9], independently proposed similar schemas, or prototypes, of wing pattern homologies for several families. These conceptual depictions show a series of parallel and symmetric patterns set over a wing background, and are a projection of all the possible elements observed in a lineage, usually a family or genus. H. Frederik Nijhout[10] revived this early literature and framed it in a developmental context, underlining wing color patterns in Lepidoptera as a system of homology hypotheses analogous to the tetrapod limb development. Since then, the Nymphalid Groundplan (NGP), the prototype for the brush-footed butterfly family, has been studied in detail using pharmaceutical manipulation, cold and heat-shock perturbations, and surveys of natural variation[11–15].

Lepidopteran wing patterns are a practical and integral model to study development. Even outside of the established comparative framework of the NGP, there are aspects of this group that make it an excellent model system to study morphological evolution. The wing imaginal discs develop from the caterpillar through the pupae stage, with a very dynamic trajectory. Starting with a simple two-cell layer 'naked' tissue, to an appendage infiltrated by trachea, with diverse types of scales, and different kinds of pigments[10,16], all arranged according to the spatial and temporal information of a small

number of critical genes, for which patterning role seems to be conserved between distant species. Lepidopteran wing patterns can be highly variable within and between species, with known selective forces shaping them[5,17]. The accessibility to whole-genome sequences has increased drastically in the last decade[18]. Also, CRISPR/Cas9 is now routinely implemented for gene knockouts[19]. Although not all species are suitable for continuous breeding in the laboratory, many species can be maintained for at least a few generations, which permits us to produce knockouts in species with very divergent phenotypes across the butterfly phylogeny. Plus, there is a network of enthusiasts and butterfly farms that collect, rear, and sell livestock in the US and other countries. In short, the dramatic diversity in Lepidopterans, the ecological relevance of wing patterns, and the feasibility of genetic work, make wing color patterns a unique resource to study developmental and evolutionary processes underlying adaptive morphological traits.

Thanks to these advantages, I examined the role of *Wnt* genes in patterning wing color elements in different species of butterflies. Wnts are a large and highly conserved group of secreted proteins that mediate animal development by creating concentration gradients, causing different cell environments and distinctive phenotypes. Different *Wnt* genes are expressed in the developing imaginal discs of butterflies and months, among them are: *wingless* (*wg*), *Wnt6*, *Wnt7*, *Wnt10*, and *WntA*[15,20,21]. Remarkably, the *WntA* gene has emerged as one of the most significant loci underlying the adaptive radiation of the mimetic butterflies *Heliconius* [22,23] and the *Limenitis arthemis* [24,25] species complex.

In Chapter 1, I tested the function of the *WntA* gene in seven species of nymphalid butterflies, to measure its role in the formation of the different elements of the NGP between species with known ancestral and divergent *WntA* expression patterns. I used a combination of CRISPR/Cas9 mosaic deletions, mRNA wing expression (*in situ* hybridization), and signaling manipulation experiments using heparin and dextran sulfate. Overall, the results show the versatility of *WntA* in the imaginal discs, driving lineage-specific effects in different wing regions, including the evolution of novel expression domains that diverge from the NGP.

In Chapter 2, so far, we learn the vital function of the *WntA* gene setting different color elements in brush-footed butterflies. The expression patterns of this gene have diverged drastically in some species, but in other stayed similar even though they still drive lineage-specific effects. Hence, I wanted to understand how and what switches, i.e., cis-regulatory elements (CREs), are activating and/or repressing *WntA*. With that goal, I implemented ATAC-seq, HiC-seq, and CRISPR/Cas9 experiments to dissect *WntA* CREs systematically and test for their function in five species of nymphalid butterflies. The results provided a portrait for the regulation of this signaling gene. We detected multiple CREs modulating *WntA* expression in a non-redundant manner, in many cases with broad wing effects and multifunctionality.

In Chapter 3, using similar methods as Chapter 1, I tested the function of *WntA* and *Wnt6* in the genus *Papilio*, the major genus of the swallowtail family Papilionidae–the

most basal lineage of butterflies. In this case, I wanted to update the *Papilio* groundplan and make inferences about the possible deep homologies between distantly related families, building from the knowledge obtained in Chapter 1. Indeed, we found *WntA* is necessary to pre-pattern major elements during imaginal disc development; however, those elements do not seem homologous to elements in the NGP. Additionally, I found that *Wnt6* is a positive regulator or a new color pattern system in the margin of the wing I dub the '*submarginal spots'*.

Taken together, my dissertation research demonstrates the evolutionary flexibility of a developmental patterning gene and the strengths of the comparative framework associated with wing color patterns to study morphological evolution. *Wnt* genes underline the morphological complexity of wings during the evolution of a diverse group of insects. *WntA* is associated with the establishment of multiple color elements; some are the product of spatial expression shifts, but also evolved *de novo* expression domains. Its regulatory landscape is complex, where multiple distant CREs are necessary to produce discreet wing pattern elements. Regions with open chromatin around *WntA* are relatively well conserved between nymphalids. However, in butterflies with a prototype arrangement (NGP) and species that diverged from it, *WntA* expression is modulated by both conserved and recently evolved CREs. Nonetheless, monarch butterflies have a divergent phenotype and regulatory architecture. Finally, the results presented here revealed enhancer- and silencer- like phenotypes, suggesting bifunctionality of CREs as an attribute of wing color pattern regulation.

The study of cis-regulation using traditional methods, mainly reporter constructs, gave us the basic principles of gene regulation, include patterning sufficiency of “modular enhancers”, but with new tools and type of data we are starting to build on, and in some cases, challenge some of the established assumptions. This research opened many questions about homology between elements, its diversification, the interaction between *Wnt* genes setting color patterns, and the traditional "modular enhancer" notions of gene regulation. Future avenues can be taken to address the new hypotheses; some examples are listed below. (1) Analysis of single-cell expression and chromatin profiling at different type points in caterpillars will help to build gene networks underlying color patterns and might detect pattern specific CREs. With this type of data, we can model gene interactions and refine the regulatory landscape of separate wing compartments or pattern elements. (2) Generation of CRE mutant lines with enhancer and silencer reporter constructs in Lepidoptera will be necessary to study the mechanisms producing the effects we observed from CRISPR/Cas9 experiments. Specifically, we can use our newly validated CREs to test for redundancy and the mechanisms behind bifunctionality. (3) Tagging proteins will open the door to assess the epistatic effect or the interactions between Wnt proteins during wing development. (4) Study of the 'trans' landscape – conserved CREs underlie very divergent expression patterns of *WntA*, begging the question, how much is really conserved, are the transcription factors binding these elements the same between species? What about novel patterns? Studying the regulators of *WntA* will provide insights into the trans- vs. cis-regulatory architecture of evolution. (5) Finally, are there different properties for the gene regulation of robust and stable traits vs.

rapidly evolving ones, like plastic and adaptive phenotypes? *WntA* could be a model to study this question by comparing wing vs. appendage/segment patterning in caterpillars.

REFERENCES

1. Arthur, Wallace. *Evolution: a developmental approach.* (John Wiley & Sons, 2010).
2. Courtier-Orgogozo, V., Arnoult, L., Prigent, S. R., Wiltgen, S. & Martin, A. Gephebase, a database of genotype–phenotype relationships for natural and domesticated variation in Eukaryotes. *Nucleic Acids Res.* **48**, D696–D703 (2020).
3. Moczek, A. P. The shape of things to come: EvoDevo perspectives on causes and consequences in evolution. in *Evolutionary causation: biological and philosophical reflections* (eds. Uller, T. & Laland, K. N.) (2019).
4. Kawahara, A. Y. *et al.* Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc Natl Acad Sci U A* **116**, 22657–22663 (2019).
5. Beldade, P. & Brakefield, P. M. The genetics and evo-devo of butterfly wing patterns. *Nat Rev Genet* **3**, 442–52 (2002).
6. Krebs, R. A. & West, D. A. Female mate preference and the evolution of female-limited batesian mimicry. *Evolution* **42**, 1101–1104 (1988).
7. Järvi, V. V., van der Burg, K. R. L. & Reed, R. D. Seasonal plasticity in *Junonia coenia* (nymphalidae): linking wing color, temperature dynamics, and behavior. *J. Lepidopterists Soc.* **73**, 34 (2019).

8. Schwanwitsch, B. N. On the ground-plan of wing-pattern in nymphalids and certain other families of the Rhopaloeerous Lepidoptera. *Proc. Zool. Soc. Lond.* **94**, 509–528 (1924).
9. Süffert, F. Zur vergleichende analyse der schmetterlingszeichnung. *Biol. Zentralblatt* **47**, 385–413 (1927).
10. Nijhout, H. F. *The development and evolution of butterfly wing patterns*. (Smithsonian Institution Press, 1991).
11. Otaki, J. M. Physiologically induced color-pattern changes in butterfly wings: mechanistic and evolutionary implications. *J Insect Physiol* **54**, 1099–112 (2008).
12. Otaki, J. M. Color pattern analysis of nymphalid butterfly wings: revision of the nymphalid groundplan. *Zoolog. Sci.* **29**, 568–576 (2012).
13. Nijhout, H. F. Symmetry systems and compartments in Lepidopteran wings: the evolution of a patterning mechanism. *Development* 9 (1994).
14. Serfas, M. S. & Carroll, S. B. Pharmacologic approaches to butterfly wing patterning: sulfated polysaccharides mimic or antagonize cold shock and alter the interpretation of gradients of positional information. *Dev Biol* **287**, 416–24 (2005).
15. Martin, A. & Reed, R. D. Wnt signaling underlies evolution and development of the butterfly wing pattern symmetry systems. *Dev Biol* **395**, 367–78 (2014).
16. Nijhout, H. F. & McKenna, K. Z. Wing morphogenesis in Lepidoptera. *Biol. Chall. Morphog.* **137**, 88–94 (2018).
17. Brakefield, P. M. *et al.* Development, plasticity and evolution of butterfly eyespot patterns. *Nature* **384**, 236–242 (1996).

18. Challi, R. J., Kumar, S., Dasmahapatra, K. K., Jiggins, C. D. & Blaxter, M. *Lepbase: the Lepidopteran genome database*. http://biorxiv.org/lookup/doi/10.1101/056994 (2016) doi:10.1101/056994.
19. Zhang, L. & Reed, R. D. A practical guide to CRISPR/Cas9 genome editing in Lepidoptera. in *Diversity and Evolution of Butterfly Wing Patterns* 155–172 (Springer, Singapore, 2017).
20. Ding, X. *et al.* Genome-wide identification and expression profiling of Wnt family genes in the silkworm, *Bombyx mori*. *Int J Mol Sci* **20**, (2019).
21. Fenner, J. *et al.* Wnt genes in wing pattern development of Coliadinae butterflies. *Front. Ecol. Evol.* **8**, 197 (2020).
22. Martin, A. *et al.* Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proc Natl Acad Sci U A* **109**, 12632–7 (2012).
23. Concha, C. *et al.* Interplay between developmental flexibility and determinism in the evolution of mimetic *Heliconius* wing patterns. *Curr Biol* **29**, 3996-4009 e4 (2019).
24. Mullen, S. P. *et al.* Disentangling population history and character evolution among hybridizing lineages. *Mol. Biol. Evol.* **37**, 1295–1305 (2020).
25. Gallant, J. R. *et al.* Ancient homology underlies adaptive mimetic diversity across butterflies. *Nat Commun* **5**, 4817 (2014).

CHAPTER 1

MACRO-EVOLUTIONARY SHIFTS OF *WNTA* FUNCTION POTENTIATE BUTTERFLY WING PATTERN DIVERSITY

Anyi Mazo-Vargas, Carolina Concha, Luca Livraghi, Darli Massardo, Richard W. R. Wallbank, Linlin Zhang, Joseph D. Papador, Daniel Martinez-Najera, Chris D. Jiggins, Marcus R. Kronforst, Casper J. Breuker, Robert D. Reed, Nipam H. Patel, W. Owen McMillan, Arnaud Martin[1].

Abstract

Butterfly wing patterns provide a rich comparative framework to study how morphological complexity develops and evolves. Here we used CRISPR/Cas9 somatic mutagenesis to test a patterning role for *WntA*, a signaling ligand gene previously identified as a hotspot of shape-tuning alleles involved in wing mimicry. We show that *WntA* loss-of-function causes multiple modifications of pattern elements in seven nymphalid butterfly species. In three butterflies with a conserved wing pattern arrangement, *WntA* is necessary for the induction of stripe-like patterns known as symmetry systems, and acquired a novel eyespot activator role specific to *Vanessa* forewings. In two *Heliconius* species, *WntA* specifies the boundaries between melanic fields and the light-color patterns they contour. In the passionvine butterfly *Agraulis*, *WntA* removal shows opposite effects on adjacent pattern elements, revealing a dual

[1] This chapter is already published: Mazo-Vargas et al., "Macroevolutionary Shifts of *WntA* Function Potentiate Butterfly Wing-Pattern Diversity." PNAS 114, no. 40 (October 3, 2017): 10701–6. https://doi.org/10.1073/pnas.1708149114. It is reprinted with permission of PNAS.

role across the wing field. Finally, *WntA* acquired a divergent role in the patterning of interveinous patterns in the monarch, a basal nymphalid butterfly that lacks stripe-like symmetry systems. These results identify *WntA* as an instructive signal for the pre-patterning of a biological system of exuberant diversity, and illustrate how shifts in the deployment and effects of a single developmental gene underlie morphological change.

Introduction

The multitude of patterns found in developing organisms is achieved by a small number of conserved signaling pathways, which raises an important question. How does biodiversity arise from the sharing of constituents across a single tree of life? One explanation for this apparent paradox is that conserved regulatory genes evolve new "tricks" or roles during development [1]. Assessing this phenomenon requires comparing the function of candidate genes across a dense phylogenetic sampling of divergent phenotypes. Here, the patterns on butterfly wings provide an ideal test case. The development of scale-covered wings, their structural and pigment complexity and an elaborate patterning system are key features of the Lepidoptera (moths and butterflies), which form about 10% of all species known to mankind [2]. Wing patterns across the group are fantastically diverse and are often shaped by natural and sexual selection [3]. Studies in fruit flies, butterflies and moths have implicated secreted Wnt signaling ligands as color pattern inducers [4–8]. In butterfly wings, two lines of evidence suggest a prominent patterning role for the Wnt ligand gene *WntA* in particular. First, *WntA* was repeatedly mapped as a locus driving pattern shape adaptations involved in

mimicry, and a total of 18 *WntA* causative alleles have been identified across a wide phylogenetic spectrum [9–13]. Second, *WntA* expression marks developing wing domains that prefigure the position and shape of pattern elements of various color compositions [10,14].

The nymphalid groundplan provides a conceptual framework to understand pattern variation in butterflies [3]. Under this framework, patterns are organized into parallel sub-divisions of autonomous color pattern complexes known as "symmetry systems", which are arranged across the dorsal and ventral surfaces of both the fore- and hindwing [14–19] (Figure 1-1A-C). This arrangement is thought to represent a putative archetype of a butterfly wing pattern, and diversity is created by modifying elements within and among these symmetry systems [3]. *WntA* is typically expressed in three of the four symmetry systems [14]: the small proximal pattern called Basalis (B), the large median pattern called the Central Symmetry System (CSS), and the Marginal Band System (MBS), which features laminar stripes bordering the wing. Here we used CRISPR/Cas9 mutagenesis to impair *WntA* function and assess its patterning roles in Nymphalidae, the largest butterfly family that radiated around 90 MYA [20]. We characterize the developmental function of *WntA* in species representative of the nymphalid groundplan, and then show that *WntA* has acquired divergent patterning roles in several lineages.

Methods

Butterflies.

Insect stock origins, rearing conditions, and oviposition host plants are described in Table S3.

In situ hybridizations.

WntA cDNA sequences cloned, or amplified with T7 overhang primers, were used as a template to synthesize DIG-labeled RNA probe as described previously [14,21]. Primers for amplification of template DNA are shown in Table S4. *In situ* hybridization of imaginal discs from fifth instar larvae were performed as described in [14].

Gene annotation

A cDNA sequence containing the full CDS of the *V. cardui WntA* gene was retrieved from a published wing transcriptome (18). The predicted peptide was used as a TBLASTN query across genome sequence assemblies available on the Release 4 of Lepbase (37) for: *P. aegeria* (38), *H. erato demophoon* (9), *D. plexippus* (39), *H. sara* (Lepbase: *H. sara helico3*), and *A. vanillae* (Lepbase: *A. vanillae helico3* assembly). *V. cardui* and *J. coenia* genomic sequences were retrieved from preliminary Single Molecule Real-Time sequencing data and the *WntA* region scaffolds are available upon request. The Phobius web server (40) was used to verify the presence of an N-terminal 20 amino-acid signal peptide in the first coding exons of all the retrieved sequences. *WntA* predicted peptide sequences showed 89% to 98% identity in pairwise alignments (Figure S1-8*A*).

sgRNA design

We implemented two approaches for invalidating the *WntA* coding sequence in mKO experiments: simultaneous injections of two or four single guide-RNAs (sgRNAs) was performed to generate large deletions by double DNA cleavage, thus facilitating genotyping (17, 18, 41); and single sgRNA injections, which induce frameshifts in the coding region by imperfect Non-Homologous End Joining repair (42). We used online tools such as CasBLASTR (http://www.casblastr.org) to select for GGN16-18 or N20 sgRNA targets in *WntA* exons (Table S1, Figure S1-8*B*). Template for *in vitro* transcription was generated by PCR assembly (43). Alternatively, synthesized target sites were annealed and ligated into the plasmid pDR274 linearized with *BsaI* (44). PCR product or *DraI*- linearized plasmids were transcribed using the MEGAscript or MEGAshortscript T7 Transcription Kits (Thermo Fisher Scientific). sgRNAs were purified with phenol–chloroform extraction and isopropanol precipitation, or using the MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific).

Egg injections

Butterfly eggs laid on host plant leaves were collected after 1–6 h (Tables S2-3). *J. coenia* and *V. cardui* eggs were then washed for 20-100 s in 5% benzalkonium chloride (Sigma-Aldrich), rinsed in water, and dried in a desiccation chamber or by air ventilation for softening the chorion. To soften and separate egg mass in *H. sara*, clumps were treated with a 1:20 dilution of Milton sterilizing fluid (Procter and Gamble) for 4 minutes, rinsed with water, and dried. Eggs were arranged on a double-sided adhesive tape or glued to a glass slide, usually with the micropyle facing up. CRISPR mixtures containing pre-assembled sgRNAs and recombinant Cas9 protein

(PNA Bio Inc.) were injected, using pulled quartz or borosilicate needles. The concentration of sgRNAs and Cas9 varied between butterfly species and experiments (Table S2).

Genotyping

DNA was extracted from wing muscles or single legs using the Phire animal tissue direct PCR kit (Thermo Fisher Scientific), and amplified using oligonucleotides flanking the sgRNAs target region (Table S2). PCR amplicons were gel-purified, sub-cloned into the pGEM®-T Easy Vector System (Promega) and sequenced on an ABI 3730 sequencer.

Drug treatments

Dextran sulfate (Sigma, Cat # D8906) was injected in the basal wing region of 4- 12 h monarch pupae using a hand pulled glass capillary needles mounted on a micro-injector system. Ten out of forty-nine pupae survived the injection and reached the adult stage.

Imaging and morphometrics

In situ hybridizations were photographed on a Nikon SMZ800N stereomicroscope with a trinocular adapter mounted with a Nikon D5300 digital camera. Adult butterflies were spread with their ventral side up. Digital photos of full specimens were taken with a Nikon D5300 digital camera mounted with an AF-S VR Micro-Nikkor 105mm f/2.8G lens and an 80-LED ring light. High-resolution images of wing patterns were taken using a Keyence VHX-5000 digital microscope fitted with a VH-Z100T or VH-Z00T lens. Pattern sizes were measured using the "Area Measurement"

modules of the Keyence VHX-5000 software. *V. cardui* dorsal forewing eyespot surfaces were analysed using a two-tailed Mann-Whitney U- test. *A. vanillae* discal spot sizes were normally distributed and analyzed using a two-tailed Student t-test. Normalization of the *V. cardui* and *A. vanillae* pattern measurements by a wing size index did not affect the results presented in this study. *P. aegeria* patterns were measured using the ImageJ software, normalized by total wing area, and analyzed using a two-tailed Student t-test.

Results and Discussion

We injected Cas9/sgRNA duplexes into 1-6 h butterfly embryos at a syncytial stage (N=5794 eggs). As only a fraction of the dividing nuclei are edited, the resulting mosaicism can bypass the deleterious effects of developmental mutations and yields G_0 escapers that survive until the adult stage for phenotypic analysis [22–24]. We performed CRISPR injections in seven nymphalid species in order to induce frameshift mutations in *WntA* coding exons. About 10% of hatchlings (240 out of 2293 survivors) yielded adult butterflies with mosaic knock-out (mKO) pattern defects on their wings (Figures S1-9 and Tables S1-2).

WntA induces Central Symmetry Systems.

First, we used CRISPR to test the effects of *WntA* loss-of-function on the wing patterns of the Common Buckeye *Junonia coenia* (tribe: Junoniini). *WntA* mKOs resulted in a complete loss of the CSS, consistent with *WntA* expression that prefigures its shape and position in the wing imaginal disks (Figure 1-1*D-E*, and Figure S1-1). The *WntA*-positive forewing B element was lost while the *wg*-positive D^1-D^2 elements

[8,25] were unaffected (Figure 1-1*F*). The B-D[1]-D[2] patterns have a similar color composition, indicating that *WntA* and *wg* play interchangeable roles in their induction. In contrast, the double loss of the distinct B and CSS patterns also illustrates the regional specificity of *WntA* signaling color outputs across the wing surface. In the marginal section of the wing (Figure 1-1*G*), *WntA* mKOs resulted in a contraction of the MBS and in a shift of chevron patterns known as the distal parafocal elements (dPF) [17,19]. *WntA* may impact these distal elements by participating in complex patterning dynamics in the marginal section of the wing [26].

Variations on the WntA groundplan theme.

Next, we asked if the instructive roles of *WntA* were phylogenetically conserved, using two other nymphalid butterflies with a groundplan organization, the Specked Wood *Pararge aegeria* (tribe: Satyrini) [15], and the Painted Lady *Vanessa cardui* (tribe: Nymphalini) [18]. *WntA* mKOs yielded consistent effects by eliminating the CSS and distalizing the parafocal elements in these two species (Figure 1-1*H-K*, Figure 1-*2A-H*, and Figure S1-10). Of note, in the *V. cardui* hindwing, the complex wave-like patterns of the CSS were lost upon severe *WntA* mKO and reduced in more intermediate forms (Figure 1-2*I-J*). These two species also highlighted other aspects of *WntA* phenotypic effects. In *P. aegeria* hindwings, the mKO-mediated disruption of the marginal system resulted in an apparent expansion of the eyespot outer rings (Figure S1-10D). *V. cardui WntA* mKOs resulted in the reduction of each dorsal forewing eyespots (p-values$<10^{-4}$; Figure 1-2*K-M*), and generated color composition defects in the ventral forewing eyespots (Figure 1-2*N*, and Figure S1-3). Only *V. cardui* forewings are

known to express *WntA* in their eyespots [14]. We thus infer that *WntA* was co-opted in the eyespot gene regulatory network of the *V. cardui* lineage to elaborate upon the patterning of this complex feature [27]. Overall, comparisons in three species show that multifaceted modulations of *WntA* function have shaped variations on the basic nymphalid groundplan theme.

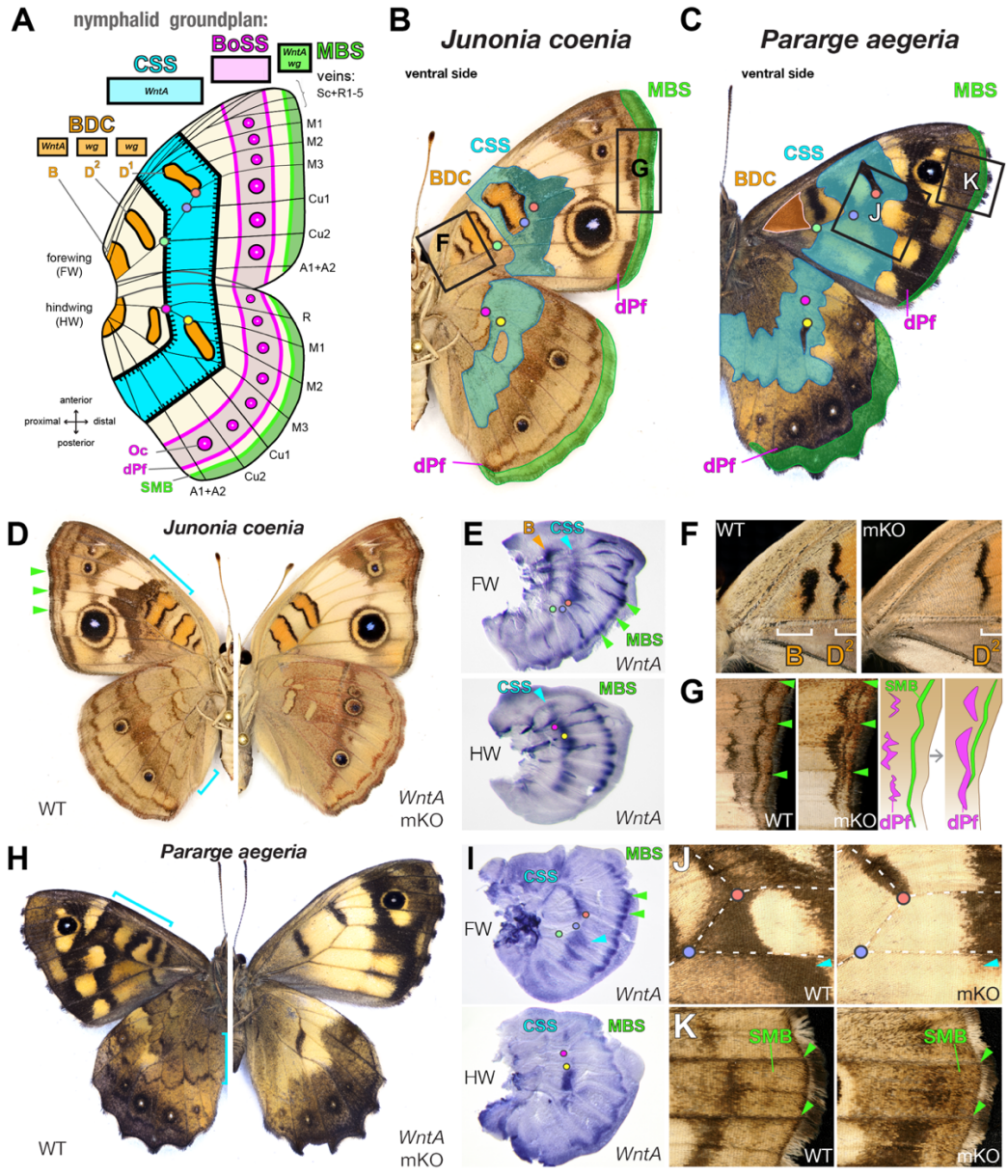


Figure 1-1. ***WntA* loss-of-function effects in groundplan-like nymphalids.** (***A-C***) The nymphalid groundplan consists of consecutive symmetry systems organized along the antero-

posterior axis. Color code indicates groundplan elements in subsequent panels (orange: BDS patterns; blue: CSS; fuchsia: Bordel Ocelli Symmetry System, including dPf; green: MBS), and dots show wing topological landmarks corresponding to vein crossings. (***D-G***) *WntA* mKO in *J. coenia* results in the loss of $WntA^{+}$ patterns: (***D***) whole wing phenotypes; (***E***) *in situ* hybridization of *WntA* in WT fifth instar imaginal disks; (***F***) Blow-up of proximal forewing area showing the loss of B upon *WntA* mKO. (***G***) Blow-up of proximal forewing area showing the distalization of dPf and Sub-Marginal Band (SMB) elements (***H-K***) Replication of the *J. coenia* results in *P. aegeria.* ***H:*** whole wing phenotypes; ***I:*** *in situ* hybridization of *WntA* in WT fifth instar imaginal disks; ***J*:** loss of the forewing CSS; **K:** distalization of dark brown dPF, and SMB; arrowheads point at corresponding *WntA* expression domains in ***I***

WntA induces pattern boundaries in Heliconius.

We next focused on species that departed more markedly from the nymphalid groundplan configuration, starting with the hyper-diverse *Heliconius* clade (tribe: Heliconiini). We performed CRISPR mKOs in Central American morphs of two species, *Heliconius erato demophoon* and *Heliconius sara sara*. *WntA* removal resulted in an expansion of light-color patterns in both cases (Figures 3*A,F*). In *H. erato demophoon, WntA* expression marked melanic patches that contour forewing red and hindwing yellow stripes (Figures 3*B,D*). Predictably, its loss-of-function resulted in the loss of the corresponding boundaries, with black contours being replaced by expansions of red or yellow (Figures 3*C, E*). *H. sara* forewing disks showed a proximal and a central *WntA* expression domains which each correspond to melanic fields that frame the signature yellow stripes of this butterfly (Figure 1-3*G*). Both melanic intervals were lost following *WntA* mKOs (Figure 1-3*H*), yielding an almost uniformly yellow forewing surface. Hindwings showed a similar effect, with *WntA* deficiency resulting in melanic-to-yellow switches in the antero-proximal half of the wing (Figure 1-3*J*). Interestingly, this treatment also revealed a cryptic stripe of red patches. A similar phenotype is observed in subspecies of *H. sara,* as well as in its

sister species *Heliconius leucadia* (Figure S1-11), suggesting that modulations of Wnt signaling could underlie these cases of natural variation. Overall, these data support previous predictions that groundplan elements such as the CSS can be homologized to what form the apparent contours of *Heliconius* patterns [28–30].

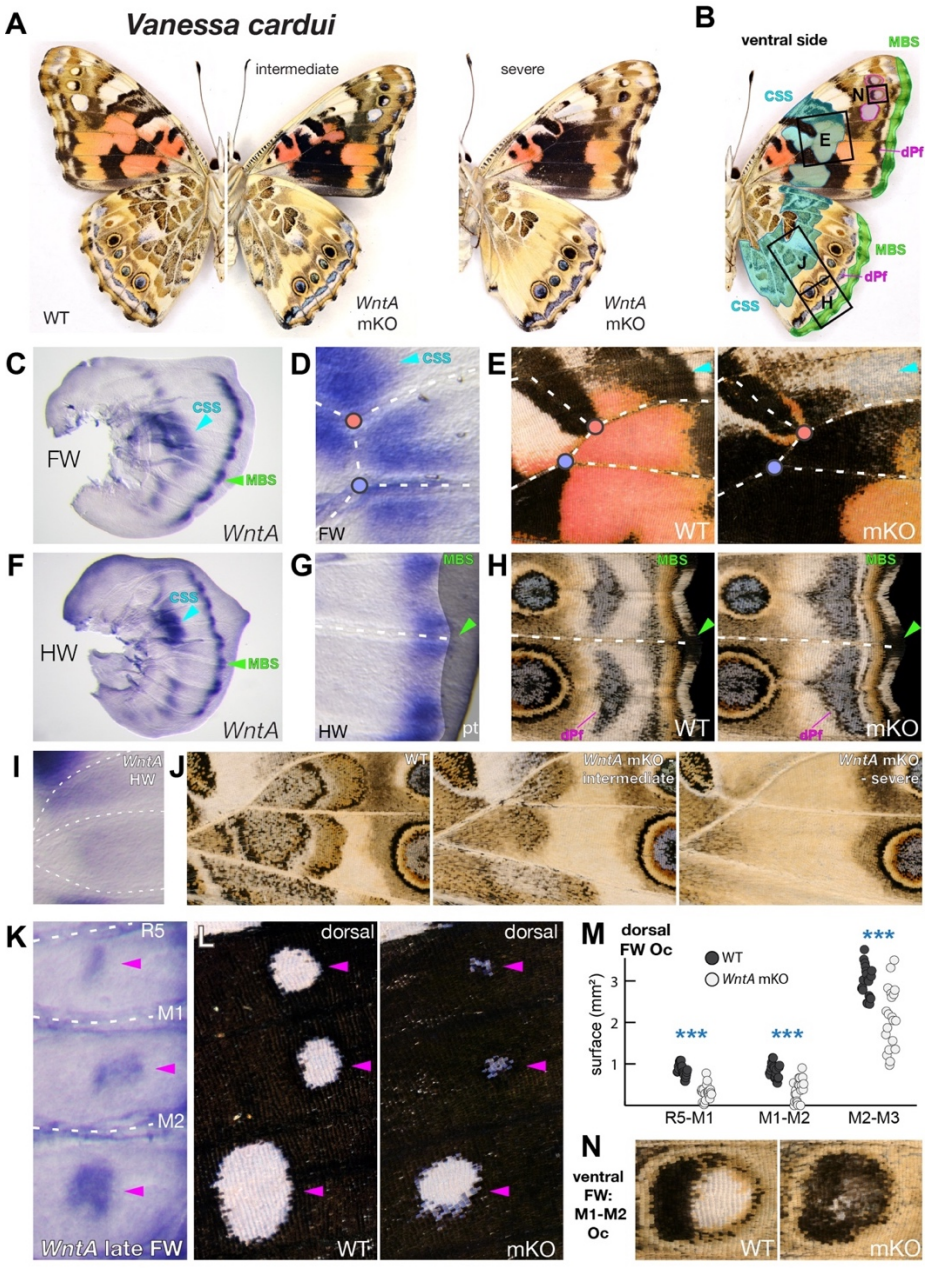


Figure 1-2. Conserved and novel aspects of *WntA* function in Painted Lady butterflies. (***A-B***) *WntA* mKOs in *V. cardui* result in defects or loss of the CSS, highlighted in cyan in ***B***.

(***C-D***) *WntA* forewing expression is associated with the CSS (magnified in ***D***) and the MBS in wild-type fifth instar wing disks. (***E***) Blow-up of a CSS section showing pattern disruption upon *WntA* mKO. (***F-G***) *WntA* hindwing expression in the CSS and the MBS (magnified in ***G***; pt: peripheral tissue) in wild-type fifth instar wing disks. (***H***) *WntA* mKO results in distal shifts of dPf elements. (***I***) Blow-up of *WntA* expression in the hindwing CSS. (***J***) Magnification of intermediate and severe levels of CSS reduction observed upon *WntA* mKO. (***K***) *WntA* expression as observed in the presumptive forewing eyespots in late fifth instar wing disks of *V. cardui*. (***L-M***) Reduction of dorsal forewing eyespots following *WntA* mKO. (***N***) Color change in ventral mKO forewing eyespots

WntA is best thought as a pre-patterning factor that determines boundaries between color fields, a view that is compatible with the replacement effects of mKOs, where *WntA*-deficient cells acquire the color fate of the adjacent territory. This property may explain why *cis*-regulatory tinkering of *WntA* expression seems to underlie the repeated modification of color pattern shapes across this explosive radiation [9–12], as it allows the coordinated modulation of color fate on either side of a moving boundary.

Antagonistic roles of WntA in adjacent patterns.

Compared to *Heliconius*, the closely related Gulf Fritillary butterfly (*Agraulis vanillae*) has modified the nymphalid groundplan differently to produce its distinctive wing pattern [28]. Rather than continuous stripes, *A. vanillae* shows dispersed silver spots of identical color composition, each consisting of a core of highly reflective "mirror" scales [31] and an outline of black scales. A subset of silver spots express *WntA* or *wg* [14], and accordingly, all the $WntA^{+}$ patterns contracted or disappeared in *WntA* mKOs (Figure 1-3*K-N*, and Figure S1-6). Among the wg^{+} elements (forewing D^{1} and D^{2}), only D^{1} co-expressed *WntA* and was specifically reduced in *WntA* mKOs (Figure

S1-12), suggesting that silver spots respond to overall Wnt dosage. *WntA* mKOs also resulted in a drastic expansion of *WntA*-free (*WntA*$^{-}$) patterns (Figure 1-3*O*).

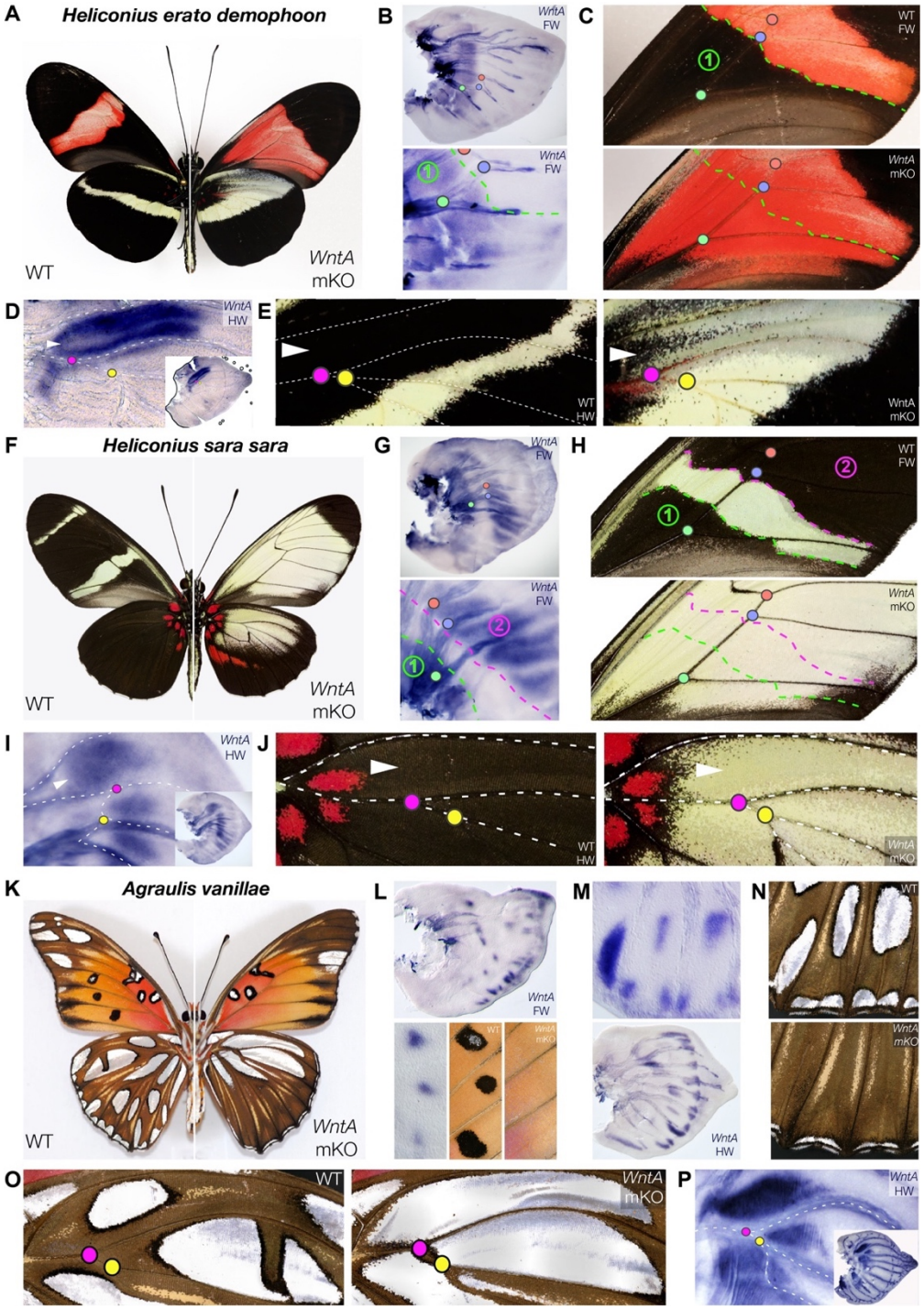


Figure 1-3. Variegated *WntA* loss-of-function phenotypes in passionvine butterflies. (***A-E***) Effects of *WntA* mKO in *H. e. demophoon*. (***A***) Whole wings. (***B***) Detection of WT

proximal *WntA* expression by larval forewing *in situ* hybridization (zone 1). (***C***) Loss of proximal pattern boundary in *WntA*-positive zone 1. (***D***) Antero-proximal expression of *WntA* in WT late larval hindwings. (***E***) Loss of antero-proximal pattern boundary in mKO hindwings. (***F-J***) Effects of *WntA* mKO in *H. sara*. (***F***) whole wings. (***G***) Detection of proximal (zone 1) and median (zone 2) *WntA* in larval forewings (***H***) Loss of proximal (green line) and median (fuchsia line) pattern boundaries resulting in loss of melanic identity in zones 1-2. (***I***) Antero-proximal expression of *WntA* in larval hindwings. (***J***) Widespread antero-proximal color identity shift in mKO hindwings. (***K-P***) Effects of *WntA* mKO in *A. vanillae*: (***K***) whole wings). (***L***) Silver spot-related expression of *WntA* in larval forewings, and loss in mKO forewings; M3-A1 spot triad shown in bottom panels. (***M-N***) Silver spot-related expression of *WntA* in larval hindwings (***M***) and loss/reduction in mKO hindwings (***N***); M3-A1 spot complex shown in top panels. (***O***) Silver spot pattern expansion in proximal mKO hindwings. (***P***) Secondary expression of *WntA* in the proximal region of late larval hindwings. Color dots: wing topological landmarks (vein crossings).

Importantly, butterflies treated with exogenous heparin, a ligand-binding molecule with Wnt gain-of-function effects [9,14,32,33], showed the opposite outcome: expanded $WntA^{+}$ and reduced $WntA^{-}$ patterns [14]. These reverse effects of CRISPR loss-of-function *vs.* heparin gain-of-function suggest that *WntA* activates and represses two distinct sets of patterns, and the repressed domain in fact shows a secondary wave of *WntA* expression in late larval instar wing disks (Figure 1-3*P*). This observation leads us to propose that the dual effect of *WntA* may be due to a biphasic deployment, with a first wave of *WntA* pattern-activating expression, followed by an inhibitory event in the Wnt-repressed territory. Testing this working model will require the identification and expression profiling of *WntA* signaling targets in *A. vanillae*.

Repurposing of WntA in a reduced groundplan.

Last, we used the lack of visible CSS in monarchs (*Danaus plexippus*; tribe: Danaini) as an example of extreme divergence from the nymphalid groundplan. *WntA* lacked a CSS median stripe expression as expected, and was instead detected around the

presumptive veins, indicative of a potential role in the induction of vein-dependent patterns [34]. *WntA* mKO adults showed drastic expansions of the white interveinous patterns (Figure 1-4), which are usually visible as thin outlines of the veins in WT ventral wings.

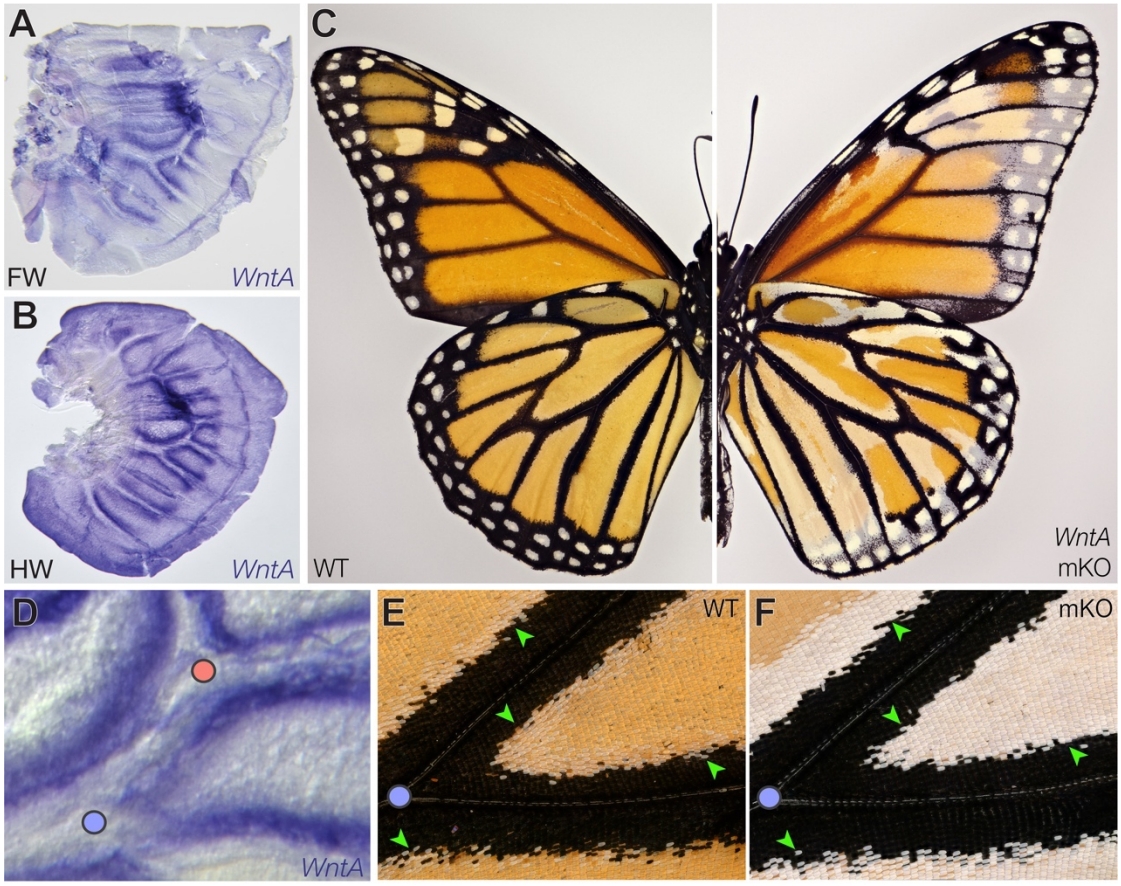


Figure 1-4. *WntA* loss-of-function in monarch butterflies induces interveinous white scales. (***A-B***) *In situ* detection of *WntA* in *D. plexippus* last larval instar forewing (***A***) and hindwing disks (***B***). (***C***) Whole wing phenotypes *WntA* mKO. All effects consist of expansion of white scale patterns. (***D***) Close-up view of larval hindwing *WntA* interveinous expression in the periphery of tracheal vein precursors. (***E-F***) Expansion of interveinous white scale fate in the hindwing region corresponding to panel ***D***. Color dots: wing topological landmarks (vein crossings).

In addition, white dot elements that decorate the marginal region expanded and fused following *WntA* mKO. Other *WntA* mKO monarchs showed a small dorsal patch of ectopic interveinous scales in the crossvein region showing maximal *WntA* expression

in hindwings (Figure S1-13). Consistent with a Wnt loss-of-function, this mild phenotype was reproduced by injection of Dextran Sulfate, a drug treatment that emulates Wnt signal inhibition in other butterflies [14,33] (Figure S1-14). Overall, expression and functional data suggest that *WntA* was again repurposed, in this case as a repressor of interveinous white scales in the monarch lineage.

Lessons from somatic CRISPR phenotypes.

Somatic mutagenesis yielded loss-of-function data in the G_0 adults of seven butterfly species, an achievement that would have been unrealistic in the pre-CRISPR era. Experimental replication using various sgRNA targets ruled out a contribution of off-target lesions, and genotyping experiments revealed a predominance of frameshift, presumably null *WntA* alleles (Figures S8-9). Variations in clone size, allelic dosage, and the possible occurrence of hypomorphic mutations could underlie complex cases of mosaicism, explaining the range of observed effects (Figure 1-2*J* and Figures S1-7). Inferring the allelic composition of wing mutant clones from their genotyping is complicated by the movement of insect wing epithelial cells following adult emergence [35], as well as by the presence of cell contaminants that are unlikely to underlie the pattern phenotype (*eg.* tracheal cells, neurons, haemocytes). We attempted the generation of germline mutations in *V. cardui* to bypass the experimental limitations of somatic heterogeneity. Following the injection of a single sgRNA targeting the *WntA* stop codon, we obtained an adult female bearing a modification of the forewing CSS (Figure S1-15). Six G_1 offsprings displayed the same phenotype and were all heterozygous for a 16bp indel mutation, resulting in a C-terminal Cys-Asn-

Stop→Gly-Ser-Arg-Stop editing of the predicted *WntA* protein. This allele was passed to a second generation but was subsequently lost due to an episode of high-mortality in our stock. Nonetheless, this preliminary result illustrates the potential of CRISPR to induce a variety of loss-of-function alleles, which could be propagated via the germline for tackling future developmental questions where mosaicism is a concern.

Conclusions

The Nymphalidae family comprise about 6,000 butterfly species, most of which can be identified by their wing patterns. We used this system as a proxy of morphological evolution and found that a single signal articulates its underlying complexity, as shown by the variety of *WntA* mKO phenotypes obtained across different wing regions and species. Our data highlight three major results. First, *WntA* is associated with multiple pattern elements within the same individual, including within the same wing surface, *eg.* the adjacent Basalis and CSS patterns both require *WntA* in *J. coenia* forewings, in spite of distinct color compositions; whereas CSS stripes often differ between wing surfaces (dorsal *vs.* ventral, forewing *vs.* hindwing). Wnt signaling may combine with selector genes that mark distinct wing domains to mediate these regional-specific outputs within a single individual [25,36]. Second, spatial shifts in *WntA* expression cause pattern shape evolution, exemplified by the multitude of species-specific manifestations of the CSS. *Cis*-regulatory variants of *WntA* [9–12], or alternatively, modulations of the *trans*-regulatory landscape that controls *WntA* expression may have fashioned these macro-evolutionary shifts. And last, *WntA* evolves new patterning functions. It was co-opted into forewing eyespot formation in

the *V. cardui* lineage, evolved a localized pattern-inhibiting role in *A. vanillae*, and was re-purposed for the patterning of vein-contouring markings in monarchs. In summary, *WntA* instructs the formation of multiple wing pattern elements in the nymphalid radiation, demonstrating the importance of pre-patterning processes in the unfolding of complex anatomy. The versatility of this signaling factor illustrates how the repeated tinkering of a developmental gene can foster boisterous evolutionary change.

Acknowledgments

We thank Saad Arif, Nora Braak, Chris Day, Melanie Gibbs, Jonah Heller, José Hermina-Perez, Colin Morrison, Oscar Paneso, Manu Sanjeev, David Tian, Camille Tulure, Matthew Verosloff, and Hans Van Dyck, for assistance with rearing, injecting, dissecting, and imaging butterflies; Lawrence Gilbert for sharing photographs and insights on *Heliconius* wing patterning; and Karin van der Burg, Bernardo Clavijo, David Jaffe, James Lewis, and James Mallet for providing access to genomic data. This work was supported by grants from the NSF (DGE-1650441; IOS-1354318; IOS-1557443; IOS-1452648), the NIH (GM108626), the Leverhulme Trust (RPG-2014-167), the Pew Charitable Trust, a Nigel Groome PhD studentship (Oxford Brookes University), and the Smithsonian Institution.

REFERENCES

1. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
2. Kristensen, N. P., Scoble, M. J. & Karsholt, O. Lepidoptera phylogeny and systematics: the state of inventorying moth and butterfly diversity. *Zootaxa* **1668**, e747 (2007).
3. Nijhout, H. F. *The Development and Evolution of Butterfly Wing Patterns*. (Smithsonian Institution Press, 1991).
4. Werner, T., Koshikawa, S., Williams, T. M. & Carroll, S. B. Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* **464**, 1143–1148 (2010).
5. Yamaguchi, J. *et al.* Periodic Wnt1 expression in response to ecdysteroid generates twin-spot markings on caterpillars. *Nat. Commun.* **4**, 1857 (2013).
6. Koshikawa, S. *et al.* Gain of cis-regulatory activities underlies novel domains of wingless gene expression in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 7524–7529 (2015).
7. Özsu, N., Chan, Q. Y., Chen, B., Gupta, M. D. & Monteiro, A. wingless is a positive regulator of eyespot color patterns in Bicyclus anynana butterflies. *Dev. Biol.* (2017).
8. Martin, A. & Reed, R. D. Wingless and aristaless2 define a developmental groundplan for moth and butterfly wing pattern evolution. *Mol. Biol. Evol.* **27**, 2864–2878 (2010).

9. Martin, A. *et al.* Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 12632–12637 (2012).
10. Gallant, J. R. *et al.* Ancient homology underlies adaptive mimetic diversity across butterflies. *Nat Commun* **5**, (2014).
11. Huber, B. *et al.* Conservatism and novelty in the genetic architecture of adaptation in Heliconius butterflies. *Heredity* **114**, 515–524 (2015).
12. Van Belleghem, S. M. *et al.* Complex modular architecture around a simple toolkit of wing pattern genes. *Nat. Ecol. Evol.* **1**, (2017).
13. Martin, A. & Courtier-Orgogozo, V. Morphological Evolution Repeatedly Caused by Mutations in Signaling Ligand Genes. in *Diversity and Evolution of Butterfly Wing Patterns - an Integrative Approach* (Springer Singapore, 2017).
14. Martin, A. & Reed, R. D. Wnt signaling underlies evolution and development of the butterfly wing pattern symmetry systems. *Dev. Biol.* **395**, 367–378 (2014).
15. Schwanwitsch, B. Evolution of the wing-pattern in Palaearctic Satyridae: Pararge and five other genera. *Acta Zool.* **16**, 143–281 (1935).
16. Schwanwitsch, B. Color-pattern in Lepidoptera. *Entomol. Obozr.* **35**, 530–546 (1956).
17. Otaki, J. M. Color pattern analysis of nymphalid butterfly wings: revision of the nymphalid groundplan. *Zoolog. Sci.* **29**, 568–576 (2012).
18. Abbasi, R. & Marcus, J. M. Color pattern evolution in Vanessa butterflies (Nymphalidae: Nymphalini): non-eyespot characters. *Evol. Dev.* **17**, 63–81 (2015).

19. Taira, W., Kinjo, S. & Otaki, J. M. The marginal band system in nymphalid butterfly wings. *Zoolog. Sci.* **32**, 38–46 (2015).

20. Wahlberg, N. *et al.* Nymphalid butterflies diversify following near demise at the Cretaceous/Tertiary boundary. *Proc. R. Soc. Lond. B Biol. Sci.* **276**, 4295–4302 (2009).

21. Carter, J.-M., Gibbs, M. & Breuker, C. J. Divergent RNA localisation patterns of maternal genes regulating embryonic patterning in the butterfly Pararge aegeria. *PLoS ONE* **10**, e0144471 (2015).

22. Perry, M. *et al.* Molecular logic behind the three-way stochastic choices that expand butterfly colour vision. *Nature* **535**, 280–284 (2016).

23. Zhang, L. & Reed, R. D. Genome editing in butterflies reveals that spalt promotes and Distal-less represses eyespot colour patterns. *Nat. Commun.* **7**, (2016).

24. Zhang, L. *et al.* Genetic basis of melanin pigmentation in butterfly wings. *Genetics* (2017) doi:10.1534/genetics.116.196451.

25. Carroll, S. B. *et al.* Pattern formation and eyespot determination in butterfly wings. *Science* **265**, 109–115 (1994).

26. Nijhout, H. F. A comprehensive model for colour pattern formation in butterflies. *Proc. R. Soc. Lond. B Biol. Sci.* **239**, 81–113 (1990).

27. Monteiro, A. & Gupta, M. Identifying Coopted Networks and Causative Mutations in the Origin of Novel Complex Traits. *Curr. Top. Dev. Biol.* **119**, 205–226 (2016).

28. Nijhout, H. F. & Wray, G. A. Homologies in the Color Patterns of the Genus Heliconius (Lepidoptera, Nymphalidae). *Biol. J. Linn. Soc.* **33**, 345–365 (1988).

29. Nijhout, H. F., Wray, G. A. & Gilbert, L. E. An analysis of the phenotypic effects of certain color pattern genes in Heliconius (Lepidoptera, Nymphalidae). *Biol. J. Linn. Soc.* **40**, 357–372 (1990).

30. Gilbert, L. E. Adaptive novelty through introgression in Heliconius wing patterns: Evidence for shared genetic 'tool box' from synthetic hybrid zones and a theory of diversification. in *Ecology and Evolution Taking Flight: Butterflies as Model Systems* (eds. Boggs, C. L., Watt, W. B. & Ehrlich, P. R.) (Univ. of Chicago Press, 2003).

31. Dinwiddie, A. *et al.* Dynamics of F-actin prefigure the structure of butterfly wing scales. *Dev. Biol.* **392**, 404–418 (2014).

32. Binari, R. C. *et al.* Genetic evidence that heparin-like glycosaminoglycans are involved in wingless signaling. *Development* **124**, 2623–2632 (1997).

33. Serfas, M. S. & Carroll, S. B. Pharmacologic approaches to butterfly wing patterning: sulfated polysaccharides mimic or antagonize cold shock and alter the interpretation of gradients of positional information. *Dev. Biol.* **287**, 416–424 (2005).

34. Nijhout, H. F. Molecular and physiological basis of colour pattern formation. *Adv. Insect Physiol.* **38**, 219–265 (2010).

35. Tögel, M., Pass, G. & Paululat, A. The Drosophila wing hearts originate from pericardial cells and are essential for wing maturation. *Dev. Biol.* **318**, 29–37 (2008).

36. Rebeiz, M., Patel, N. H. & Hinman, V. F. Unraveling the tangled skein: the evolution of transcriptional regulatory networks in development. *Annu. Rev. Genomics Hum. Genet.* **16**, 103–131 (2015).

CHAPTER 2

CIS-REGULATORY ARCHITECTURE OF BUTTERFLY WING PATTERN EVOLUTION

Anyi Mazo-Vargas and Robert D. Reed

Abstract

Cis-regulatory variation is a known mechanism underlying the evolution of morphological complexity and diversity. However, a systematic characterization of the regulatory architecture of adaptive traits awaits. We take advantage of butterflies' diverse color wing patterns as a comparative framework to characterize the cis-regulation of the signaling protein *WntA*, a locus known for its role in the evolution of mimetic color patterns. We first defined the chromatin accessibility landscape in five brush-footed butterflies to identify candidate cis-regulatory elements (CREs) for *WntA*. We then used our CRISPR-Cas9 'shotgun' deletion approach to test their function in wing patterning. We revealed that CREs modulating *WntA* work in concert to produce the diversity of wing designs. Hence, independent deletions of individual CREs are enough to disrupt the phenotype, i.e., multiple CREs are *'necessary'* to establish wing color patterns. We included nymphalid species with an ancestral wing pattern groundplan arrangement (*Junonia* and *Vanessa*) and species that have diverged from this groundplan (*Agraulis*, *Heliconius,* and monarchs). In both groups, we found ancient and recently evolved CREs interacting to set the wing elements. However, monarch butterflies present a highly divergent cis-regulatory architecture

underlying color pattern formation. Finally, our shotgun deletion approach revealed that most CREs have the potential to act as both enhancers and silencers. Our results provide one of the first portraits of the cis-regulatory architecture of a developmental gene that underlies a complex and fast-evolving adaptive character system.

Introduction

A key goal of evolutionary developmental biology is to identify the genomic instructions guiding body pattern formation and to understand how changes to those instructions generate diversity[1]. The study of variation in natural populations has been very insightful for this and continues to highlight an increasing number of loci underlying adaptive phenotypes in various study systems[2–5]. In many cases, the identified genomic regions correspond to non-coding sequences, likely containing cis-regulatory elements (CREs) that regulate the expression of nearby genes. Still, methodological constraints have prevented systematic functional characterization of candidate regulatory regions underlying the evolution of adaptive traits[6]. From model organism research, which has relied predominantly on reporter assays, it is widely believed that CREs provide to gene expression a modular structure (in time, location, and strength), which reduces off-target effects of pleiotropic mutations and results in accurate and robust developmental pathways[1,6]. By extension, modular CRE architecture is a core tenet of evolutionary developmental biology, even though very few studies have directly tested this idea by looking at actual phenotypic effects of specific CREs.

CREs are broadly classified into enhancers or silencers, although most previous research focused on enhancer function, particularly on the loss of individual CREs underlying loss or simplification of traits[7–9]. However, recent studies are highlighting greater functional complexity in the regulatory genome, where CREs can function as both enhancers and silencers[10,11], where CRE redundancy plays a role in maintaining gene expression robustness[12], and where chromatin structure itself plays a role in *in vivo* gene regulation[9,13]. To gain insights into the regulatory evolution of cis-regulation, here we assess the function and evolution of multiple candidate CREs implicated in regulating *WntA*, an essential signaling ligand gene that underlies wing pattern development in many groups of butterflies[14–19].

Butterfly wing patterns are a popular model for studying the developmental genetic basis of morphological diversification. Lepidoptera (butterflies and moths) consist of 160,000 extant species[20] – about 10% of the earth's biodiversity. Butterflies alone include over 18,000 described species, with most of the extant lineages having diverged after the Cretaceous-Paleogene (K-Pg) mass-extinction 65 mya[21]. Colorful wing designs are enough to distinguish many butterfly species, and many patterns are known or suspected to be involved in various adaptive roles, including seasonal plasticity, camouflage and mimicry, and mate choice[22,23]. Also, only a handful of genes underlie patterning at large areas of the wings, and variation at some of these loci drive local adaptation in several model species[5,14,16,24–27]. This oligogenic architecture suggests a certain evolutionary flexibility to reprogram the wing-patterning developmental program.

One wing-patterning gene that has attracted significant interest is *WntA*, which encodes a signaling ligand that induces the major central pattern systems in the brush-footed butterfly family Nymphalidae[15,17], including mimicry patterns in admiral (*Limenitis*) and longwing (*Heliconius*) butterflies[14,16,18,19]. Extensive work shows that *WntA* function in patterning these mimicry traits also plays a role in the deeper diversification of wing patterns across nymphalids [5,10,18], which highlights *WntA* as an important bridge between micro- and macro-evolutionary processes. *WntA* signaling defines multiple elements of the nymphalid groundplan, including the so-called "central symmetry system," which sets up the major stripe patterns in the center of the wing. Also, sequence association data have led to the deduction that the cis-regulatory evolution of *WntA* is responsible for the evolution of the symmetry system and its presumptive derivatives (e.g., melanic mimicry patterns in *Heliconius*) within and between species[14,15,17,18]. Here, we specifically sought to characterize the cis-regulatory basis of *WntA* color patterning in nymphalids in order to gain a basic understanding of the genetic basis of color pattern evolution. We used ATAC-seq on developing wings from five representative brush-footed butterfly species to identify candidate CREs regulating *WntA* expression, followed by a novel mosaic "shotgun" deletion screen using CRISPR/Cas9 to test CRE functionality in wing patterning. We were surprised to find that the cis-regulatory landscape of *WntA* is characterized by relatively deep conservation of many CREs, coupled with functional interdependence of CREs, and fragility of the regulatory architecture. Another major surprise was that most CRE deletions showed both enhancer- and silencer- like phenotypes, suggesting that bifunctionality may be an important feature of color pattern regulation. Our

findings challenge many of the assumptions of traditional "modular enhancer" models of regulatory evolution and suggest a new model where arrays of interdependent and functionally complex CREs

Methods

Butterflies

We sampled five distantly related species of nymphalids (i.e., brush-footed butterflies – the largest family of butterflies). The monarch butterfly, *Danaus plexippus*; the common buckeye, *Junonia coenia*; the painted lady, *Vanessa cardui*; the gulf fritillary, *Agraulis vanillae*; and a neotropical longwing, *Heliconius himera*. All butterfly colonies were reared at the Reed Lab at Cornell University. We used a 16/8 hr light/dark cycle at a temperature of 27-30 C and relative humidity of 60%. Monarch butterflies were a gift of the Anurag Agrawal lab at Cornell University, and larvae fed on the *Asclepia curassavica*. The *J. coenia* butterflies derived from a lab colony originated by Fred Nijhout lab at Duke University, and the larvae were fed an artificial diet as described in Nijhout (1980). *V. cardui* caterpillars and artificial food were purchased from US Commercial Provider, Carolina Biological Supplies, catalog # 144070, and 144040. *A. vanillae* pupae were purchase from Shady Oak Butterfly Farm, and *H. himera* was imported from Ecuador through LPS LLC. Larvae of the last two species were fed on *Passiflora biflora*.

ATAC-seq

For each species, we collected three replicates of the head, forewing, and hindwing tissue from the middle of their last developmental instar. Nuclei were extracted and processed as previously described[28], with the following modifications: For each library, 3-4 larvae were dissected in ice-cold PBS, and isolated wing or heads tissues were homogenized with a dounce homogenizer. The homogenate was then stained with trypan blue and nuclei were counted using a hemocytometer. The formula used to calculate the number of nuclei was (size human genome (3.2Gb) / butterfly genome size) x 50,000. Lysis and transposition were done according to the original ATAC-seq protocol[29]. Qiagen PCR cleanup kits were used to clean transposed DNA, followed by PCR amplification for 10-11 cycles, and a final PCR cleanup kit purification step. Libraries were pooled and sent to sequence at the BRC Genomics Facility at Cornell University, for 2 x 37 bp paired-end (PE) read sequencing on an Illumina NextSeq 500. Read alignment and filtering for samples were performed as previously described[30]. Briefly, ATAC-seq reads were aligned to each species' reference genome (Table S1-1) using Bowtie2[31]. We used ATACseqQC package[32] to assess data quality (Table S1-2). Alignments were filtered, and duplicates removed. We measured sample correlations using the multiBigwigSummary / plotCorrelation from Deeptools[33]. Peak calls were made using Fseq[34], and DESeq2[35] was used to assess the changes in ATAC-seq peak signals between tissues. Differential accessibility of peaks was called with Benjamini-Hochberg adjusted p-value < 0.05.

Hi-C-seq

We pooled 26 *J. coenia* forewing imaginal discs collected from 13 larvae in the middle of their final developmental instar. We followed the published in-situ Hi-C protocol[36] with modifications as described in Lewis *et al*. [11]. Hi-C sequences were analyzed with HiCExplorer tools[37,38]. Reads were mapped against *J. coenia* reference genome (jc_genv2) to produce a contact matrix (HicBuildMatrix) at different resolution bins (1kb, 5kb, 10kb, 20kb, 50k, 100kb). Matrices were corrected using hicCorrectMatrix (--correctionMethod 'KR'). Finally, we estimated the TAD-separation score to identify the degree of separation between right and left directions of the analyzed bin with tool hicFindTADs, and visualize them with hicPlotTADs.

Whole-genome alignment and ortholog prediction

We identified evolutionarily conserved, orthologous CREs using a known phylogenetic tree[21] and a multiple species alignment. Eleven representative genomes from Nymphalidae plus a *Papilio* genome (Table S1-1) were aligned using progressiveCactus[39]. HAL tools[40] were then used to validate and to export alignments to MAF format. PhastCons[41] was used to calculate conservation scores and predictions of the "most conserved" elements. Each of the five species in this study was used as a reference to generate conservation tracks for the *WntA* scaffold. Also, we used HALPER[42] to identify open chromatin orthologs between the five species of interest; this tool allowed us to construct continuous orthologs segments using ATAC-seq peaks coordinates and summits for each species.

CRISPR/Cas9 mosaic shotgun CRE deletions

ATAC-seq signal from heads was subtracted from signal from wings using bamCompare from Deeptools[33] and inspected in IGV browser[43]. We restricted the focus area to the region around the *WntA* coding sequence, limited by the upstream and downstream genes. The signal in the focal area was visually inspected in each species for sharp differentially accessible peaks with a positive signal in wings. From these, we picked peaks that did not overlap with exons to be incised from the genome using CRISPR/Cas9[44]. We designed between 2-4 sgRNAs per peak (Table S1-3), following the motif N20NGG, and purchased sgRNAs from Synthego. This "shotgun" approach of injecting multiple sgRNAs is expected to produce a range of different deletion lengths in locations within individual CREs in order to screen for different possible sequence-specific functions of the element. To avoid or reduce off-target effects, we checked sgRNA sequences for either: (1) being unique BLAST hits in the reference assembly (optimal case) or, (2) with few hits but with more than two mismatches in the first eight nucleotides adjacent to the PAM, and no PAM present. For each species, oviposition was stimulated by placing a host plant in their cage[17]; eggs were collected after 1 – 2 hr. *D. plexippus*, *A. vanillae*, and *H. himera* eggs were injected without further treatment. *J. coenia* and *V. cardui* eggs were washed with 5% benzalkonium chloride (Sigma-Aldrich, St Louis, MO, USA) for 60 and 30 seconds, respectively, and then dried in a desiccation chamber for 10 min. Finally, eggs were glued to a slide and injected using either borosilicate or aluminosilicate needles. Larvae were reared as described above.

Interpretation of G0 mosaics

G0 adult wings were screened for mosaic color pattern phenotypes. Successfully mutated individuals can present clones with different deletion alleles of dissimilar sizes[17,45,46]. To positively identify mutants, we look for asymmetric disruptions in the color patterns between left and right wings, either dorsally or ventrally, since such asymmetry is unlikely due to natural variation. Importantly, for the species in this study, we already know the *WntA* loss-of-function phenotypes[17,18] and presumptive enhancement-of-function effects from heparin injections[15,47,48]. We therefore scored CRE deletions that phenocopied *WntA* knockout phenotypes as having "enhancer-like" functionality. Conversely, mutants that showed expansion of *WntA* color patterns, thus phenocopying heparin injections, were scored as having "silencer-like" functionality.

Long read DNA genotyping

Our CRISPR genome editing strategy using 2-4 pooled sgRNAs creates indels of different sizes at targeted loci. To confirm correct targeting of sgRNAs *in vivo*, and to assess relative frequencies of edit lengths, we use Nanopore long-read sequencing[49]. DNA was extracted from bodies of butterflies that showed mutant phenotypes using the OMEGA E-Z 96 Tissue DNA Kit. We then PCR amplified DNA using primer pairs at least 400 bps away from the most external sgRNAs for each CRE (Table S1-4). Amplified samples were purified with Roche KAPA pure beads. Library preparation was done using the Ligation Sequencing Kit (SQK-LSK109) following the manufacturer's specifications. The flow cell was placed on MiniON (ONT) and allowed to run for 48 hours. MinKNOW produced FAST5 files, and Guppy[50]

performed base-calling with a high-accuracy model (dna_r9.4.1_450bps_hac). Porechop[51] was used to remove adapters and multiplex samples using PCR primer sequences. Minimap2[52] was then used to map long read sequences to the amplicon wild type sequence. Alignments were then inspected in Geneious Prime 2020.1.2. (https://www.geneious.com).

Results

Deep conservation of WntA wing pattern regulatory elements

WntA has been functionally characterized in several butterfly species[17]; however, we still know very little about the CREs that regulate this gene. Our first step to identify *WntA* wing patterning CREs was to estimate the topological association domain (TAD)[53] of the *WntA* locus using Hi-C-seq analysis. Inside individual TADs, CREs and gene promoters preferentially interact with each other[54]. Thus, by identifying the TAD boundaries around *WntA,* we can highlight the non-coding interval that is most likely to contain salient CREs. Furthermore, fine-scale analysis of the Hi-C-seq data can allow us to identify which non-coding regions physically interact with the *WntA* promoter itself, and to prioritize these as candidate regulatory sequences. We performed Hi-C-seq using *J. coenia* forewing tissue from the mid-last larval stage when *WntA* expression presages adult wing patterns. We selected *J. coenia* because it is a major lab model species in color pattern studies, and displays prominent and easily resolvable *WntA*-induced nymphalid groundplan elements. In *J. coenia,* we identified a distinct TAD containing *WntA* and two other genes. *WntA* is the central feature of this TAD, with its exons and introns collectively occupying 62% of its length (Figure

2-1). The Hi-C chromatin interaction plot pinpointed the strongest promoter interactions as occurring just upstream of the *WntA* promoter, and in the long first intron. Based on these data from *J. coenia*, coupled with previous sequence association studies also highlighting the first intron, we focused our subsequent functional screens for *WntA* CREs on these regions (Figure 2-1, Table S1-3).

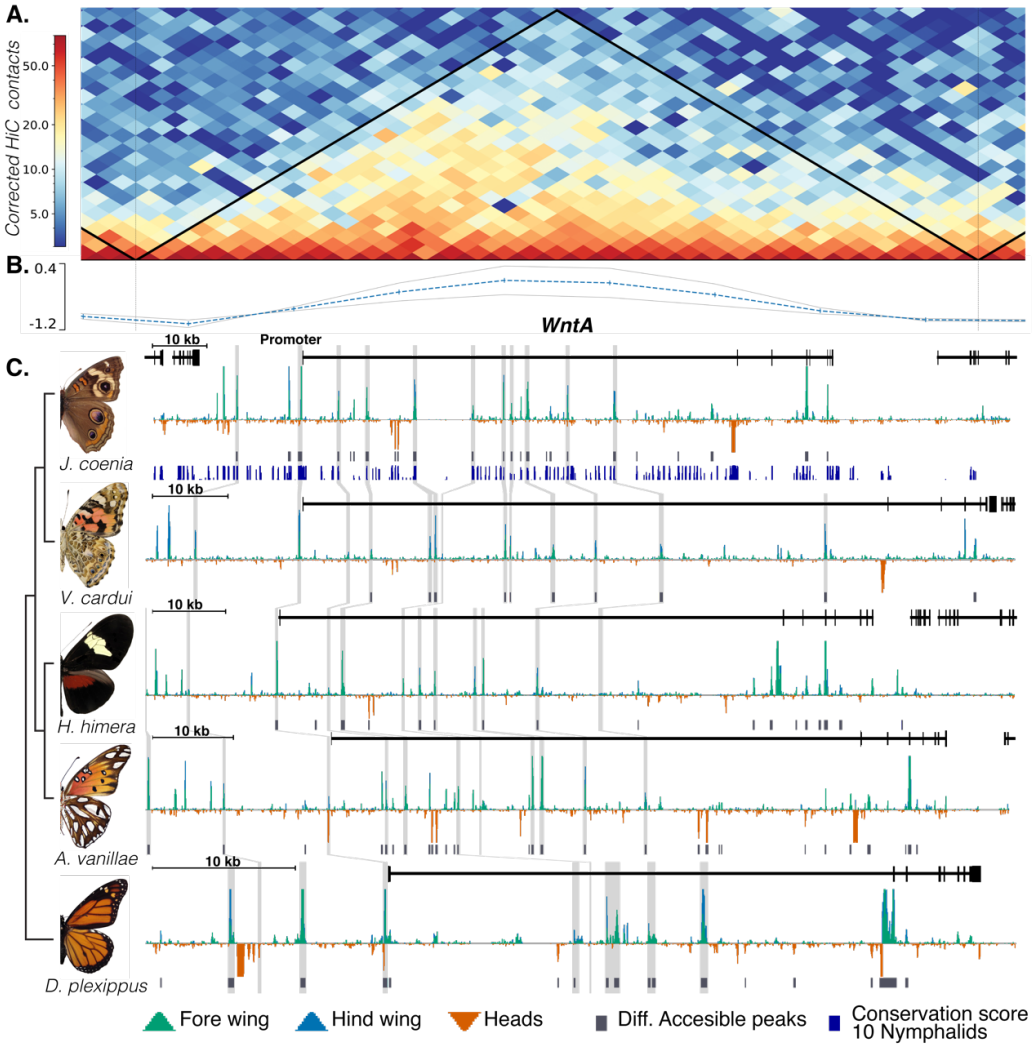


Figure 2-1. Chromatin landscape of the *WntA* locus. **(A)** Hi-C reveals abundant chromatin interactions across the upstream and first intron regions of *WntA*. Color intensity corresponds to the contact frequency per bin. Black lines depict the TAD boundaries as predicted by the **(B)** TAD-separation score. **(C)** Chromatin accessibility tracks (ATAC-seq) for each species sampled in this study. A phylogeny portrays the sampled species, from top to bottom: *J. coenia, V. cardui, H. himera, A. vanillae,* and *D. plexippus*. Grey bars connecting different species represent orthologies of CREs tested in this study.

To identify specific *WntA* CREs active during wing pattern determination, we profiled the chromatin accessibility state of developing heads, forewings, and hindwings of five butterfly species using ATAC-seq (Figure 2-1). We sampled imaginal wing discs and developing head tissue during the mid-last larval stage when *WntA* expression first occurs in wing color patterns for all the species in this study[15,17]. By comparing developing head and wing ATAC-seq profiles, we were able to identify a subset of CREs with wing-specific activity, thus raising their priority as candidate wing patterning elements. Depending on the species, we found that11-33% of the total peak calls are differentially accessible between developing heads and wings, with the majority showing accessibility specifically in wing discs (Figure S2- 1). In sum, across all five species, the ATAC-seq data provided high-resolution annotation of dozens of CREs active during specification of *WntA* color patterns, including in the upstream and first intron regions highlighted by the Hi-C chromatin interaction data.

We next wanted to look at the identified CREs in an evolutionary context in order to understand to what extent individual wing disc CREs are conserved vs. lineage-specific. To discern the histories of the identified CREs, we generated genome alignments for eleven nymphalid species and produced sequence conservation scores. By overlapping the most conserved sequences with the differentially accessible chromatin regions in the *WntA* TAD, we observed 69-88% of wing-specific CREs are located in areas with strong sequence conservation between nymphaline, satyrine, and heliconiine butterfly subfamilies. The interesting exception to this trend of

conservation was the monarch, *D. plexippus*, for which 70.6% of the ATAC-seq peaks were in non-conserved sequences specific to the subfamily Danainae (Figure 2-1). While each CRE has its own history, and every species shows both conserved and derived CREs that would benefit from more detailed future analyses, our conservation analysis shows a clear pattern. Namely, there is a largely conserved cis-regulatory architecture shared amongst the core nymphalid subfamilies, while the basal Danainae clade has a distinctly different set of CREs (Figure 2-2).

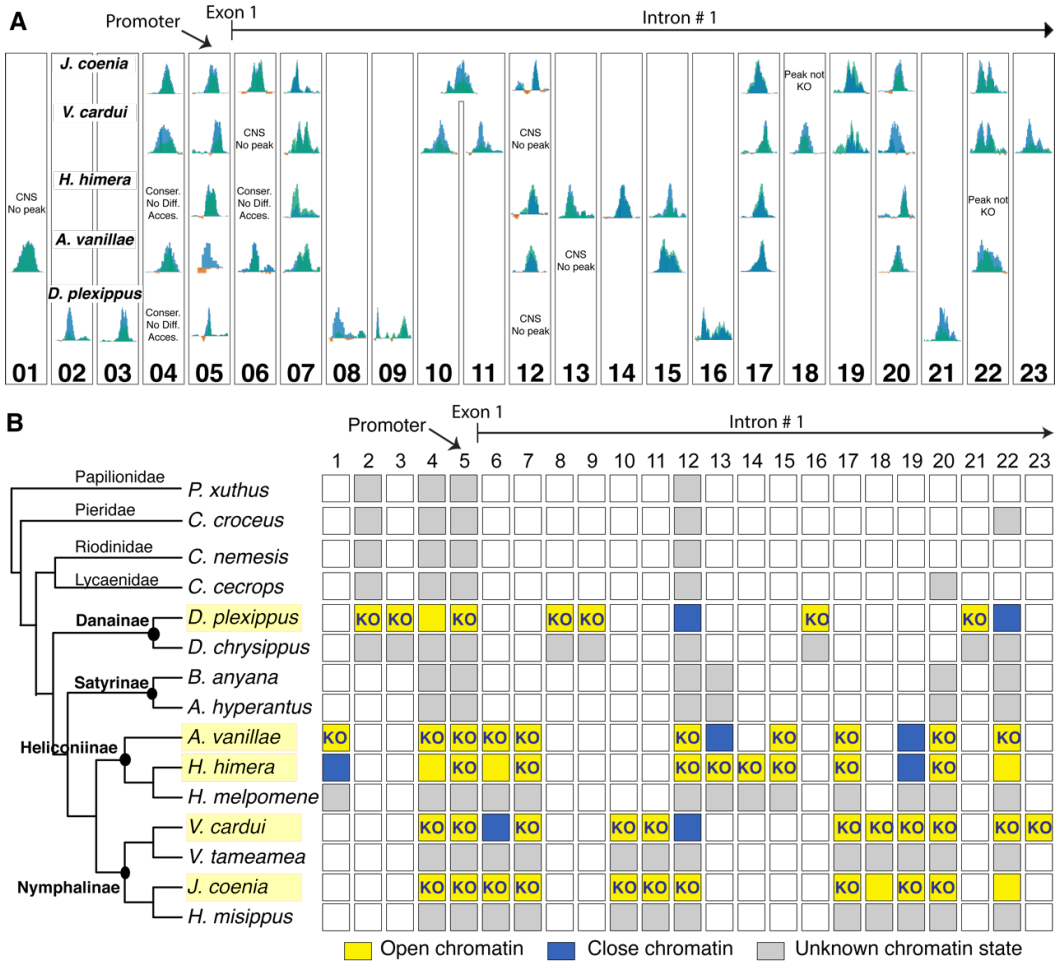


Figure 2-2. *WntA* CREs validated by 'shotgun' mosaic deletion. (A) ATAC-seq profiles of for each of the 46 candidate *WntA* CREs that were individually knocked out. Each column shows the open chromatin state (developing forewing in green, hindwing in blue, heads in orange) and their homology between species, adding up to a total of 23 unique elements. Element 05 is the presumptive promoter. **(B)** Phylogenetic histories of *WntA* CREs in this

study, sequence conservation of an expanded selection of butterflies reveals deep conservation of many CREs, as well as a handful of recently derived elements in several lineages. Solid color boxes indicate orthologs regions.

Multiple interdependent CREs regulate WntA wing patterns

Hi-C and ATAC-seq allowed us to identify several candidate *WntA* CREs across nymphalid butterflies, including both conserved and lineage-specific elements (Figure 2-1, 2-2). We next sought to functionally validate a subset of these using a CRISPR/Cas9 mosaic deletion approach. Our first goal was to test CREs that potentially play a role in establishing ancestral nymphalid groundplan elements, such as the central symmetry system. For this, we targeted candidate CREs active in developing wings, that were also conserved in *J. coenia* and *V. cardui* – two species with *WntA*-positive wing patterns thought to be representative of the ancestral groundplan[15,17]. If the orthologous elements were present in additional species with divergent wing patterns, we targeted those as well when possible. We also targeted another set of elements that are specific to species with wing patterns that have diverged from the nymphalid groundplan: the heliconiines *A. vanillae* and *H. himera*, and the danaine *D. plexippus*. We avoided selecting candidate CREs close to, or overlapping, exons, in order to avoid generate coding mutations. We used a pool of 2-4 guide RNAs targeting each CRE in order to produce a range of deletion sizes across mutants, to ask if different regulatory effects could be partitioned within sequences of individual CREs. Upon emergence, wings of gRNA-injected animals were screened for deviations from the wild type wing patterns, and consistent with previously reported *WntA* knockout or enhancement phenotypes. In total, we successfully

generated deletions in 46 CREs, representing a total of 23 distinct elements across five butterfly species (Figure 2-2, Table S2-5).

We first focused on CRE deletion results from *J. coenia* and *V. cardui* (Figure 2-3), the two species that display *WntA* expression consistent with the presumptive ancestral groundplan. Our initial expectation was that we would identify individual CREs that would have effects limited mainly to specific individual groundplan pattern elements, consistent with traditional models of cis-regulatory modularity (Figure 2-3A-B, F-G). We were surprised to observe a quite different trend, however, where many distinct and spatially distant CREs had similar effects on the same color patterns elements. Furthermore, many of the CREs affected multiple elements of the groundplan (i.e., basal, central, and external pattern elements) (Figure 2-3C-E, H-K, Table S2-5). These observations are important for understanding the general regulatory architecture of *WntA*, because they demonstrate that (1) no single CRE is alone sufficient to activate a distinct pattern element, and (2) multiple CREs are necessary for establishing the major groundplan pattern elements. This functional interdependence of CREs for generating color patterns is consistent with the Hi-C-seq data as well, which reveals extensive physical interactions across the entire first intron indicative of intricate looping between the CREs themselves, as well as the *WntA* promoter. Thus, our data support a model in which color patterns are built by a collective array of interdependent, physically interacting non-coding elements, as opposed to a series of functionally independent elements underlying specification of discrete pattern sub-elements.

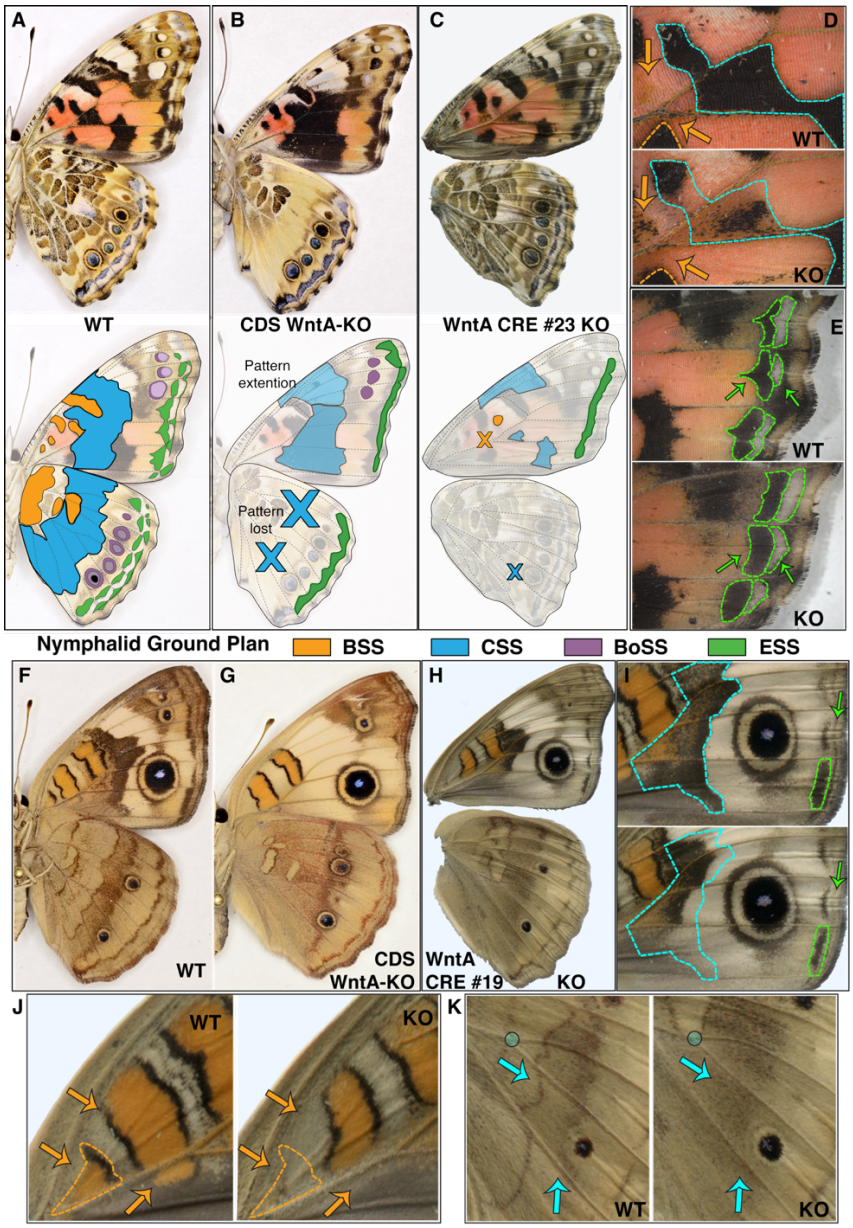


Figure 2-3. CREs modulate ancestral wing patterns in *V. cardui* and *J. coenia*. **(A)** *V. cardui* wild type wing patterns, with color coded illustration of Basal Symmetry System (BSS), Central Symmetry System (CSS), Border Ocelli Symmetry System (BOSS), and Externae Symmetry System (ESS). **(B)** *V. cardui WntA* coding knockout shows effects on all annotated symmetry systems. **(C-E)** An example of a mosaic deletion of the species-specific CRE #23 showing effects across multiple patterns elements (see illustration). **(D)** Blow-up of proximal forewing area showing lost melanic reduction in the CSS (blue dot-line) and modification of the BSS (orange arrows and dot-line). **(E)** Close-up of the forewing margin showing the ESS (green dot line) shape distortion. Arrows and color circles are landmarks. **(F)** J. coenia wild type and **(G)** coding knockout compared to (H) deletion of Nymphalini CRE #19 producing effects in the **(I)** Forewing CSS (blue dot-line), ESS (green dot-line), **(J)** and BSS (Orange dot-line and arrows); **(K)** and in the hindwing CSS.

This functional interdependence of color pattern CREs then prompted us to ask if there is an association between the age of a CRE and the nature of its function. Can recently evolved CREs show the same degree of interdependency as deeply conserved elements? To test this, we deleted a CRE (#23) that appears to be unique to *V. cardui* – this CRE is not found in the sister species *Vanessa tameamea* or any other butterfly for which we have genomic sequences. Remarkably, we found that the deletion of this element causes the reduction and removal of elements in the basal, central, and marginal symmetry system (Fig 2-3C-E). This result suggests that even recently-originated elements can be functionally integrated into the interdependent CRE network.

Ancient groundplan CREs regulate derived Heliconius color patterns

Our work on *J. coenia* and *V. cardui* reveals an overall deeply conserved *WntA* cis-regulatory landscape in nymphaline butterflies displaying the prototypical color pattern groundplan. There are many nymphalids, however, that have highly derived and/or simplified color patterns that depart significantly from the ancestral groundplan. In particular, color pattern homologies of *Heliconius* butterflies have been a matter of speculation and debate[55]. Although gene knockouts have shown that the broad fields of melanin typical of this genus are specified by *WntA* [17,18] – the same ligand as the central symmetry system – it is an open question to what degree patterning of the black fields is achieved through orthologous groundplan CREs vs. derived *Heliconius* CREs. ATAC-seq and comparative sequence analysis suggest that a large number of *WntA* CREs are deeply conserved between heliconiines and

nymphalines (Figure 2-1, 2-2), including CREs that pattern the nymphaline groundplan symmetry system. We used CRISPR/Cas9 to delete five of these conserved CREs in *H. himera* to assess their wing pattern functions. Remarkably, we observed that deletions in all five CREs similarly affected melanic patterns in the forewing (Figure 2-4A-B, Table S2-5). The deletion of two heliconiine-specific CREs revealed similar effects as well (Figure 2-4C-D). Thus, two main conclusions derive from these experiments. First, *Heliconius WntA* has a similar cis-regulatory architecture as nymphaline butterflies, where multiple interdependent CREs are all necessary for the specification of a common set of color pattern elements. Second, and most notably, the highly derived mimicry-related color patterns of *Heliconius* appear to share deep regulatory homology with the nymphalid groundplan.

The homology between *Heliconius* color patterns and the nymphalid groundplan has been a subject of debate for decades[55,56], and it is compelling to find that *Heliconius'* characteristic melanic fields are derived from the ancient regulatory logic governing the ancestral nymphalid groundplan itself. In more basal heliconiines, that typically show fragmented black, brown, and silver spots in the center of the wing, it has been proposed that those derived color elements are simplified central symmetry system patterns, and represent an intermediate state of pattern reduction on the way to the *Heliconius* pattern archetype. We tested this in the basal heliconiine *A. vanillae,* by producing deletions in the same set of conserved CREs that we tested in *H. himera*, along with several additional heliconiine-specific CREs. Again, we found similar results (Figure 2-4E-H, Table S2-5), deletions in eight CREs have overlapping effects distributed across all *WntA*-induced color patterns (black and silver spots, in this case).

This supports previous models of gradual evolution of heliconiine color pattern through fragmentation and/or simplification of the nymphalid central symmetry system, and extends it to further suggest that the process took place through the tinkering of a broadly conserved, ancestral cis-regulatory apparatus.

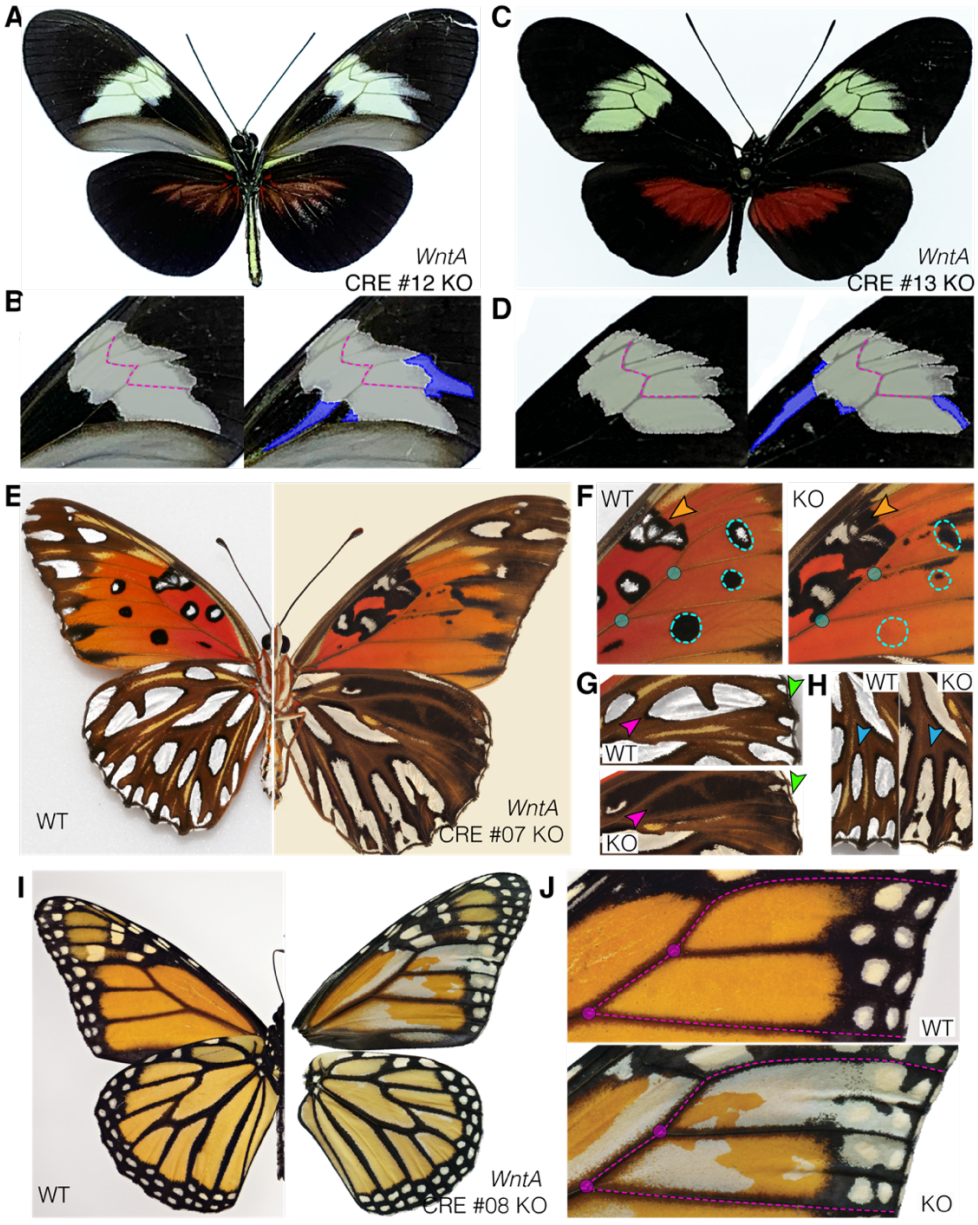


Figure 2-4. Diverged and conserved CREs effects in passion vine and monarch butterflies. **(A-D)** *H. himera* mosaic butterflies showing effects of **(A)** deeply conserved CRE #12 KO whole individual and **(B)** blow-up showing the proximal and distal extension of the

yellow band. Effects of CRE #13 KO is showing similar effects **(C)** in whole wings **(D)** and wing blow-up. **(E-H)** Effects of CRE #07 KO in *A. vanillae*, (A) mosaic butterfly with mutant clones affecting **(F)** Fore wing melanic expansion in BSS (arrowhead), removal of CSS black dots (blue line). **(G)** Silver spot loss in the anterior hind wing (arrowhead) and **(H)** fusion of CSS and ESS marginal silver spots. **(I-J)** Monarch butterflies' effects of *WntA* CRE #08 KO, extending intravenous white patterns **(I)** over regions with normally occurring orange and black colors. Color dots are wing landmarks for reader reference and are based on vein organization.

A novel array of CREs determines unique WntA patterns in monarchs

Our work on heliconiines shows that highly divergent color patterns can evolve from a set of largely conserved CREs. We next wanted to ask to what extent this groundplan cis-regulatory architecture may extend to other more phylogenetically diverged butterfly lineages. We thus turned our attention to *D. plexippus*, the monarch butterfly. The color patterns of monarchs are notable because, like *Heliconius*, they are highly derived and challenging to homologize with the nymphalid groundplan based on morphology alone. In monarchs, *WntA* shows a very distinctive vein-associated expression pattern, and *WntA* deletions cause loss of these patterns (Figure 2-4I-J, Table S2-5). These vein-associated color patterns have little obvious association with the central symmetry system; thus, it is a major question of how these patterns may be evolutionarily related to the groundplan[17].

From sequence comparisons and ATAC-seq, we found that most of the open chromatin regions in monarchs, in both developing wing and head tissue, are located in danaine-specific genomic sequences (Figure 2-2 and Figure S2-1). Generally, the non-coding regions of the monarch *WntA* locus share relatively little sequence similarity with those of other nymphalids and show an overall reduced number of

ATAC-seq peaks. Although there is a handful of conserved CREs, including the *WntA* promoter, it appears that the cis-regulatory architecture of *WntA* has highly diverged between danaines and other butterflies. To test the wing patterning function of monarch CREs, we performed CRISPR mosaic deletions on six danaine-specific CREs and one ancestrally conserved CREs. We found that all of the CREs affected wing color patterns consistent with previous functional work on *WntA* (Figure 2-4I-J, Figure S2-5, Table S2-5). The mutant clones consisted of white patches extending from the veins to replace the orange and black portions of the wing (Figure 2-4I-J). These results show that within butterflies, there can be highly diverged cis-regulatory architectures for color patterning – i.e., despite the heliconiine case above, not all color patterning architectures are necessarily derived directly from ancestral regulatory sequences.

Most WntA CREs can act as both enhancers and silencers in wing pattern development

From previous work[17,18], we know the effects caused by the deletion of the *WntA* coding sequence (i.e., loss of color patterns), and our CREs knockouts overwhelmingly produced perfect phenocopies of the *WntA* knockout (Figure 2-3, 2-4, Table S2-5), thus validating the activity of those CREs as enhancers. We were extremely surprised to find, however, that our shotgun deletion screen resulted in many examples of spatial expansion of *WntA* color patterns. Indeed, the vast majority of CREs showed the ability to both promote and repress *WntA* color patterns. Importantly, the pattern expansion effects we observed phenocopied injections of

heparin sulfate proteoglycans, which are known to enhance *WntA* signaling during color pattern formation[15,17,47], likely by facilitating the spread of Wnt ligands[57]. Thus, the spatial expansion of *WntA* patterns we observe is consistent with increase in *WntA* expression and shows that many CREs can have silencer-like functions in addition to the enhancer-like functions described above. We therefore scored mutants from our shotgun deletion screen for presenting phenotypes similar to the heparin-induced patterns as 'silencer-like,' and 'enhancer-like' if mutations resembled *WntA* knockout phenotype.

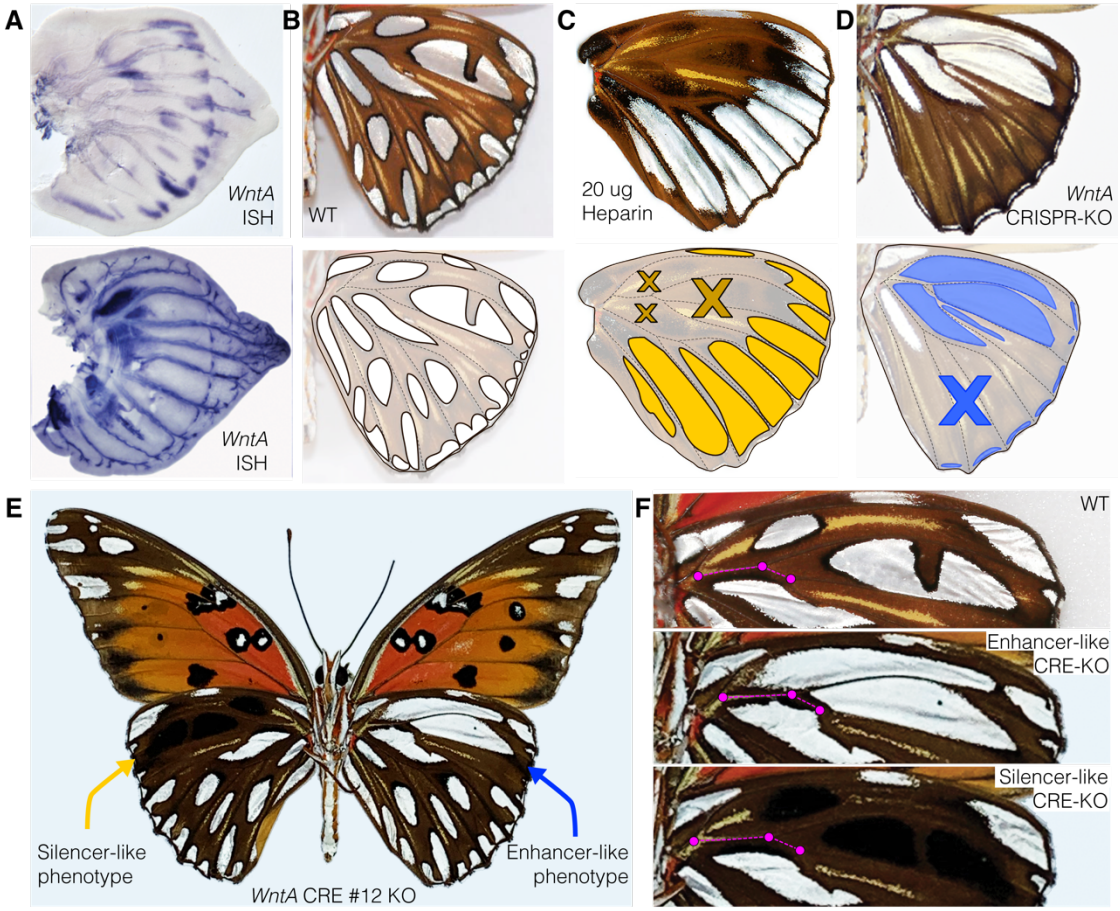


Figure 2-5. Bifunctional activity of *WntA* CREs knockouts in *A. vanillae*. (A) *In-situ* detection of *WntA* in hind wings at two developmental time points during the middle of caterpillar development. **(B)** Wild type hindwing phenotype compared to effects of **(C)** Heparin and **(D)** *WntA* CDS CRISPR-Cas9 injections. Yellow and blue colors indicate the

contrasting effects of these perturbations, with silver spots extending or loss (X). **(E-F)** *WntA* CRE #12 knockout enhancer and silencer-like effects in **(E)** mosaic butterfly and **(F)** blow-up of the silver spot in the anterior region of the hindwing, which is extended or lost in symmetric wing regions. Magenta lines and dots are wing landmarks, based on vein organization.

In Figure 2-5, we illustrate these two opposing effects in the species *A. vanillae*. This species presents discrete silver and black spots surrounded by a gradient of brown-orange coloration. We know that *WntA* is expressed in a subset of posterior-distal silver and black spots early in the imaginal disc development (Figure 2-5A). When heparin is injected after pupation, *WntA*-positive spots expand (Figure 2-5 C, yellow regions in the cartoon), while *WntA*-negative spots shrink (Figure 2-5C, see the anterior area of hindwing with the yellow 'X'). Conversely, deletion of the *WntA* coding region presents the opposite results – the silver spots disappear in the posterior region of the hindwing and expand in the anterior compartment of the wing (Figure 2-5D). Using this logic across species, we identified 27 of the 46 CREs behaving in both enhancer- and silencer-like capacities (Figure 2-6, Table S2-5). Importantly, we observed cases of same CRE producing both enhancer-like and silencer-like effects at the same wing location in different mosaic individuals (e.g., Figure 2-5E-F), suggesting that these effects are caused by different mutations, and are not simply an effect of location on the wing.

An advantage of our shotgun deletion approach, where we inject multiple sgRNAs tiled across each given CRE, is that mutant clones are expected to have a range of mutations of slightly different lengths and positions within the CRE. Thus, we should observe different phenotypic effects if certain mutations affect sequences with

different regulatory functions (e.g., different transcription factor binding sites) within a CRE.

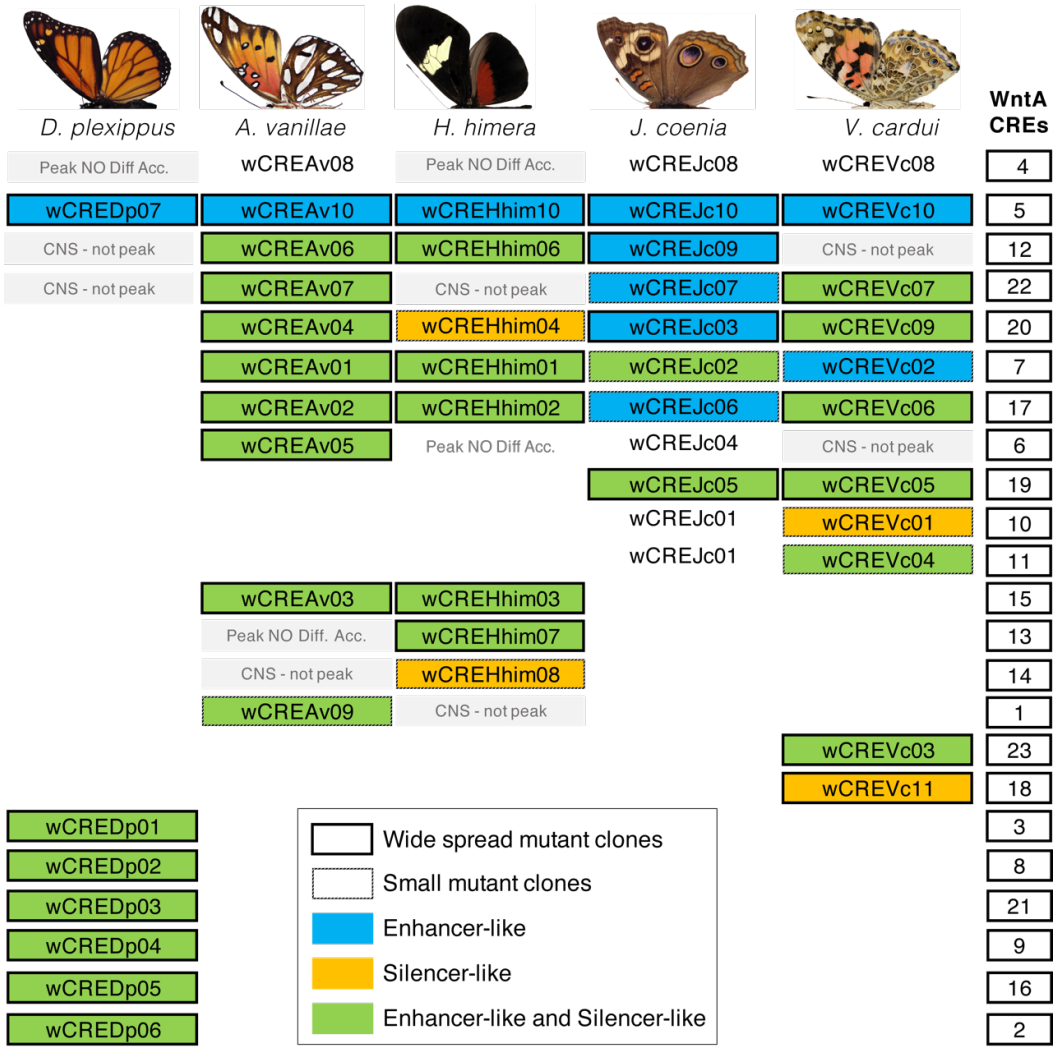


Figure 2-6. *WntA* CREs dual effects in wing phenotypes after CRISPR-Cas9 'shotgun' strategy. A large number of *WntA* CREs knockouts revealed Enhancer and silencer-like functions in patterning wing colors. *Enhancer-like:* Removal of target DNA sequence results in effects similar to loss of *WntA* by CDS KOs. *Silencer-like:* Removal of target DNA sequence results in effects similar to the ones produced by heparin injections.

To confirm the efficiency and nature of deletions produced by shotgun deletions *in vivo*, we genotyped our mosaic mutants using Nanopore long-read sequencing of edited CRE loci. After aligning them against the amplicon wild type sequence, we observed different deletion lengths and locations, matching sgRNA positions, as

predicted (Figure 2-7 and Figure 2-S7). Our results showing both enhancer-like and silencer-like effects for are consistent with this variety of deletion alleles, and lead us to surmise that bifunctional CREs contain a combination of both activator and repressor binding sites that are being disrupted in different mosaic clones. This result highlights the potential of many, or most, CREs to act as both enhancers and silencers, depending on the sequence context.

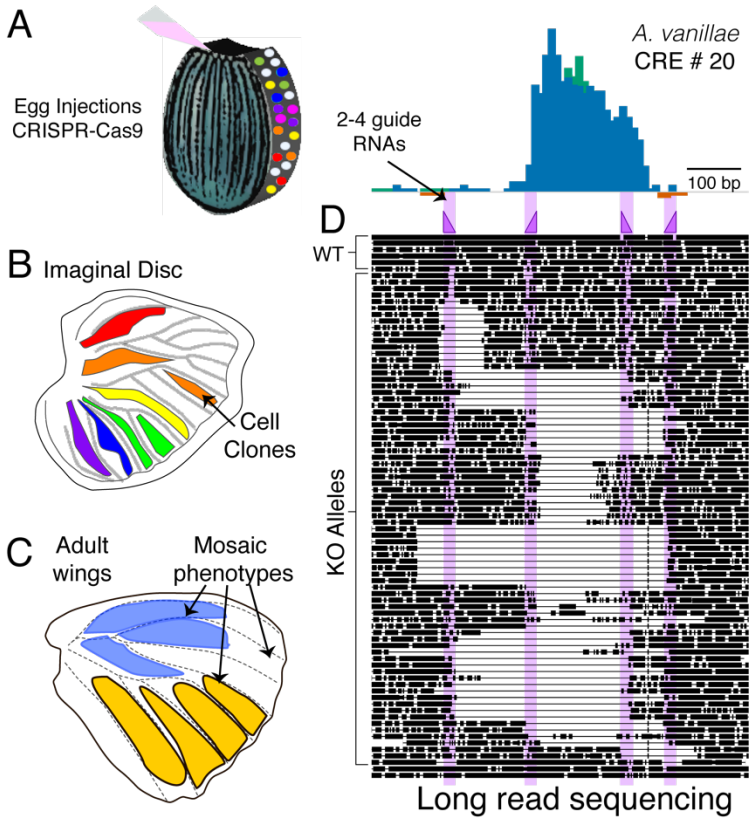


Figure 2-7. CRISPR-Cas9 'shotgun' screen reveals different CREs effects in wing color patterns. **(A)** Knockouts using two or more guide RNAs produced an array of cuts around the candidate open chromatin region (dots in egg represent different mutations), producing **(B)** cell clones with different KO alleles during development and **(C)** adult phenotypes. **(D)** The KO alleles can be trace back by using long read sequencing.

Discussion

CRE interdependence and bifunctionality

Here we tested the function of a total of 46 ATAC-seq predicted CREs on wing patterning in five strikingly different butterfly species. The diversification of butterfly wing designs is thought to be facilitated by the evolutionary and developmental flexibility that arises from the compartmentalized organization of wing pattern elements, where individual patterns can evolve independently[58,59,22]. A gene known to be essential in patterning one set of independent components in butterfly wings is the signaling ligand gene *WntA*[17,18]. Because of its ability to establish a series of dramatically divergent color patterns across species, it is hypothesized that its evolutionary flexibility can be traced to its cis-regulatory architecture[5,17]. From this, we expected that individual *WntA* CRE knockouts would have specific effects on certain wing compartments or pattern elements, as has been predicted using report constructs in other systems, including famous examples like *even skipped* stripe enhancers[1,60,61]. We also predicted that many of the CRE knockouts would present minor, or even undetectable phenotypes due to redundancy – another prediction deriving mainly from reporter-construct work[12,62].

In contrast to these predictions, we found that most individual CRE deletions acted globally in the forewing and hindwing tissue in the species we studied, without restriction to a wing compartment or a specific *WntA* color pattern element (e.g., basal vs. central symmetry system). Thus, each CRE is necessary for proper *WntA* patterning across the wing. Because lesions in any single CRE can cause global

disruption of *WntA* color patterns, we can conclude that no single CRE is sufficient to activate color pattern formation in a specific element. This interdependency and apparent crosstalk between CREs impart a regulatory fragility to wing patterning. We see little evidence of developmental robustness in *WntA's* cis-regulatory architecture. We know that multiple CREs regulate most developmental genes[12,62,63]. In the case of enhancers, the simple case of interaction is additive, where the sum of expression driven by each enhancer will amount for the total phenotype[1,63]. Other types of synergistic interactions are known, such as cooperative and competitive binding of trans effectors, and their interplay could prevent or produce ectopic gene expression[64–66]. It is now an open question to understand the mechanisms maintaining the interdependency and fragility of the regulatory architecture in butterfly color pattern loci[11], but also determine if this is a unique feature of fast-evolving traits, as are the wing color patterns in butterflies, contrary to evolutionary developmental stable programs with recognized robustness [12,67].

Another surprising result of our CREs knockouts was the detection of opposite phenotypic effects from mutations of the same elements. In all the species, we found CREs mutants that phenocopied *WntA* knockouts, i.e., enhancer-like function. However, we also observed phenotypes that phenocopied expansion of *WntA* color patterns, i.e., silencer-like effects. Previous data have suggested dual role of CREs from ChIP-seq data in the same tissue[68], bi-functionality at different tissues[10], or in different regions of the same tissue[11,64]. Our findings are an important addition to the emerging bifunctionality discussion since we show this dual functionality plays a vital

role in *in vivo* pattern formation and, is a widespread feature of a patterning gene's regulatory architecture. This all poses the question of how different deletions in the same CRE could drive opposite effects? Does the CRE contain a shared sequence that will allow the activation or repression of the gene? Or are there different adjacent sequences producing the two effects?[69]. We observe that our shotgun deletion method creates a range of deletion alleles in the mosaic mutant butterflies; thus, we speculate that different alleles have differentially balanced silencer or enhancer functionality. Our results imply the integration of these functions in individual CREs, instead of across adjacent silencers and enhancers[69]; however, a more fine-scale examination of these bifunctional elements is needed to fully understand the mechanisms at play.

Wing pattern evolution

A major strength of the butterfly study system is that we can compare species that display the presumptive ancestrally conserved nymphalid groundplan (e.g., *J. coenia*), with butterflies that have highly divergent patterns (e.g., *Heliconius* and monarchs). We expected to detect different CREs driving the evolution of the ancestral and diverged expression patterns. Instead, in the case of heliconiines vs. nymphalines, we found mostly a common set of orthologous CREs driving expression of both ancestral and derived color patterns. One possibility is that the combinatorial interaction of CREs can modulate the expression of *WntA*, the most conserved can act at the scale of whole wing tissue, while lineage-specific CREs restrict the expression to certain compartments/elements in the wing. However, this remains to be tested.

In addition to assessing highly conserved CREs, we also targeted lineage-specific CREs with no detectable sequence conservation. The most striking case is in monarch butterflies, where five of the seven elements deleted corresponded to *Danaus*-specific CREs. We found that all of these played a role in the characteristic vein-associated *WntA* color patterning in this clade. These results indicate that divergent regulatory sequences underlie divergent phenotypes in the case of danaine butterflies. Thus, the cis-regulatory architecture of *WntA* is not necessarily functionally constrained and can undergo major shifts over evolutionary time. Deeper sampling of more diverged butterfly, and even moth, species will be important for determining to what extent the distinctive danaine CRE landscape evolved de novo, or is perhaps more reflective of an original, simpler architecture that predates the groundplan as seen in nymphalines.

Acknowledgments

We thank Jon Sanders for building a laboratory in the basement of AMV's house–his equipment and input that allowed us to perform Nanopore genotyping in the middle of the COVID-19 pandemic. We also are thankful to Amy Picard Hastings and Anurag Agrawal for sharing the monarch eggs and milkweed plants. To Arnaud Martin, Linlin Zhang, Joe Hanly, Steven Van Belleghem, and Riccardo Papa for comments and experimental input. James Lewis for sharing Hi-C protocol. Ceili Peng, Joshua Kwan, Benjamin J. Brack, Rachael Geltman, and Caitlyn E. Robertson at the Reed lab for assistance injecting and rearing caterpillars. This work was supported by National Foundation Grants DGE-1650441 and IOS-1656514.

REFERENCES

1. Carroll, S. B. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**, 25–36 (2008).
2. Linnen, C. R. *et al.* Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312–6 (2013).
3. Campagna, L. *et al.* Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci Adv* **3**, e1602404 (2017).
4. Verta, J. P. & Jones, F. C. Predominance of cis-regulatory changes in parallel expression divergence of sticklebacks. *Elife* **8**, (2019).
5. Van Belleghem, S. M. *et al.* Complex modular architecture around a simple toolkit of wing pattern genes. *Nat Ecol Evol* **1**, 52 (2017).
6. Tomoyasu, Y. & Halfon, M. S. How to study enhancers in non-traditional insect models. *J. Exp. Biol.* **223**, jeb212241 (2020).
7. Chan, Y. F. *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**, 302–5 (2010).
8. Kvon, E. Z. *et al.* Progressive loss of function in a limb enhancer during snake evolution. *Cell* **167**, 633-642 e11 (2016).
9. Kragesteen, B. K. *et al.* Dynamic 3D chromatin architecture contributes to enhancer specificity and limb morphogenesis. *Nat. Genet.* **50**, 1463–1473 (2018).

10. Gisselbrecht, S. S. *et al.* Transcriptional silencers in *Drosophila* serve a dual role as transcriptional enhancers in alternate cellular contexts. *Mol. Cell* **77**, 324-337.e8 (2020).
11. Lewis, J. J. *et al.* Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *Proc Natl Acad Sci U A* **116**, 24174–24183 (2019).
12. Osterwalder, M. *et al.* Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature* **554**, 239–243 (2018).
13. Paliou, C. *et al.* Preformed chromatin topology assists transcriptional robustness of *Shh* during limb development. *Proc. Natl. Acad. Sci.* **116**, 12390–12399 (2019).
14. Martin, A. *et al.* Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proc Natl Acad Sci U A* **109**, 12632–7 (2012).
15. Martin, A. & Reed, R. D. Wnt signaling underlies evolution and development of the butterfly wing pattern symmetry systems. *Dev Biol* **395**, 367–78 (2014).
16. Gallant, J. R. *et al.* Ancient homology underlies adaptive mimetic diversity across butterflies. *Nat Commun* **5**, 4817 (2014).
17. Mazo-Vargas, A. *et al.* Macroevolutionary shifts of *WntA* function potentiate butterfly wing-pattern diversity. *Proc. Natl. Acad. Sci.* **114**, 10701–10706 (2017).
18. Concha, C. *et al.* Interplay between developmental flexibility and determinism in the evolution of mimetic *Heliconius* wing patterns. *Curr Biol* **29**, 3996-4009 e4 (2019).

19. Mullen, S. P. *et al.* Disentangling population history and character evolution among hybridizing lineages. *Mol. Biol. Evol.* **37**, 1295–1305 (2020).
20. Kawahara, A. Y. *et al.* Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proc Natl Acad Sci U A* **116**, 22657–22663 (2019).
21. Espeland, M. *et al.* A comprehensive and dated phylogenomic analysis of butterflies. *Curr Biol* **28**, 770-778 e5 (2018).
22. Beldade, P. & Brakefield, P. M. The genetics and evo-devo of butterfly wing patterns. *Nat Rev Genet* **3**, 442–52 (2002).
23. Daniels, E. V., Murad, R., Mortazavi, A. & Reed, R. D. Extensive transcriptional response associated with seasonal plasticity of butterfly wing patterns. *Mol Ecol* **23**, 6123–34 (2014).
24. Reed, R. D. *et al. optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science* **333**, 1137–1141 (2011).
25. Nadeau, N. J. *et al.* The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–10 (2016).
26. Westerman, E. L. *et al. Aristaless* controls butterfly wing color variation used in mimicry and mate choice. *Curr. Biol.* **28**, 3469-3474.e4 (2018).
27. Komata, S., Lin, C.-P., Iijima, T., Fujiwara, H. & Sota, T. Identification of *doublesex* alleles associated with the female-limited Batesian mimicry polymorphism in *Papilio memnon*. *Sci. Rep.* **6**, 34782 (2016).
28. Lewis, J. J. & Reed, R. D. Genome-wide regulatory adaptation shapes population-level genomic landscapes in *Heliconius*. *Mol Biol Evol* **36**, 159–173 (2019).

29. 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**, 1213–8 (2013).

30. Lewis, J. J., van der Burg, K. R. L., Mazo-Vargas, A. & Reed, R. D. ChIP-Seq-annotated *Heliconius erato* genome highlights patterns of cis-regulatory evolution in lepidoptera. *Cell Rep* **16**, 2855–2863 (2016).

31. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).

32. Ou, J. *et al.* ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics* **19**, 169 (2018).

33. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* **44**, W160–W165 (2016).

34. Boyle, A. P., Guinney, J., Crawford, G. E. & Furey, T. S. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* **24**, 2537–2538 (2008).

35. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550–550 (2014).

36. Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–80 (2014).

37. Ramírez, F. *et al.* High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).

38. Wolff, J. *et al.* Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **46**, W11–W16 (2018).

39. Paten, B. *et al.* Cactus: Algorithms for genome multiple sequence alignment. *Genome Res.* **21**, 1512–1528 (2011).

40. Hickey, G., Paten, B., Earl, D., Zerbino, D. & Haussler, D. HAL: a hierarchical format for storing and analyzing multiple genome alignments. *Bioinformatics* **29**, 1341–1342 (2013).

41. Siepel, A. *et al.* Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).

42. Zhang, X., Kaplow, I. M., Wirthlin, M., Park, T. Y. & Pfenning, A. R. HALPER facilitates the identification of regulatory element orthologs across species. *Bioinformatics* btaa493 (2020) doi:10.1093/bioinformatics/btaa493.

43. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **14**, 178–192 (2013).

44. Zhang, L. & Reed, R. D. A practical guide to CRISPR/Cas9 genome editing in Lepidoptera. in *Diversity and Evolution of Butterfly Wing Patterns* 155–172 (Springer, Singapore, 2017).

45. Zhang, L. & Reed, R. D. Genome editing in butterflies reveals that *spalt* promotes and Distal-less represses eyespot colour patterns. *Nat Commun* **7**, 11769 (2016).

46. Zhang, L., Mazo-Vargas, A. & Reed, R. D. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *Proc Natl Acad Sci U A* **114**, 10707–10712 (2017).

47. Serfas, M. S. & Carroll, S. B. Pharmacologic approaches to butterfly wing patterning: sulfated polysaccharides mimic or antagonize cold shock and alter the interpretation of gradients of positional information. *Dev Biol* **287**, 416–24 (2005).

48. Sourakov, A. Leopards and giants, tigers and woolly bears: casting a broader net in exploring heparin effects on Lepidoptera wing patterns. *F1000Research* **7**, (2018).

49. Deamer, D., Akeson, M. & Branton, D. Three decades of nanopore sequencing. *Nat. Biotechnol.* **34**, 518–524 (2016).

50. Wick, R. R., Judd, L. M. & Holt, K. E. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol.* **20**, 129 (2019).

51. Wick, R. R. *Porechop. v0.2.4*. https://github.com/rrwick/Porechop. (2018).

52. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

53. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).

54. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

55. Nijhout, H. F. *The development and evolution of butterfly wing patterns*. (Smithsonian Institution Press, 1991).

56. Nijhout, H. F. & Wray, G. A. Homologies in the colour patterns of the genus Heliconius (Lepidoptera: Nymphalidae). *Biol. J. Linn. Soc.* **33**, 345–365 (1988).
57. Fuerer, C., Habib, S. J. & Nusse, R. A study on the interactions between heparan sulfate proteoglycans and Wnt proteins. *Dev. Dyn.* **239**, 184–190 (2010).
58. Nijhout, H. F. Symmetry systems and compartments in Lepidopteran wings: the evolution of a patterning mechanism. *Development* 9 (1994).
59. Nijhout, H. F. Elements of butterfly wing patterns. *J. Exp. Zool.* **291**, 213–225 (2001).
60. Prud'homme, B., Gompel, N. & Carroll, S. B. Emerging principles of regulatory evolution. *Proc Natl Acad Sci U A* 104 Suppl 1, 8605–12 (2007).
61. Arnoult, L. *et al.* Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* **339**, 1423–6 (2013).
62. Cannavo, E. *et al.* Shadow enhancers are pervasive features of developmental regulatory networks. *Curr Biol* **26**, 38–51 (2016).
63. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-changing landscapes: transcriptional enhancers in development and evolution. *Cell* **167**, 1170–1187 (2016).
64. Perry, M. W., Boettiger, A. N. & Levine, M. Multiple enhancers ensure precision of gap gene-expression patterns in the *Drosophila* embryo. *Proc Natl Acad Sci U A* **108**, 13570–5 (2011).
65. Dunipace, L., Ozdemir, A. & Stathopoulos, A. Complex interactions between cis-regulatory modules in native conformation are critical for *Drosophila snail* expression. *Development* **138**, 4075–84 (2011).

66. Iampietro, C., Gummalla, M., Mutero, A., Karch, F. & Maeda, R. K. Initiator elements function to determine the activity state of BX-C enhancers. *PLoS Genet* **6**, e1001260 (2010).

67. Frankel, N. *et al.* Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature* **466**, 490–3 (2010).

68. Yasuoka, Y. *et al.* Occupancy of tissue-specific cis-regulatory modules by *Otx2* and *TLE/Groucho* for embryonic head specification. *Nat. Commun.* **5**, 4322 (2014).

69. Halfon, M. S. Silencers, enhancers, and the multifunctional regulatory genome. Trends Genet. **36**, 149–151 (2020).

CHAPTER 3

SWALLOWTAIL BUTTERFLY WING PATTERNS INDUCED BY WING MARGIN WNT SIGNALING

Anyi Mazo-Vargas, Alan Liang, Brian Lian, Jeanne McDonald, Robert Reed[2]

Abstract

Butterfly wing color patterns are conceptualized as a group of homologous elements, described as a 'prototype.' Modifications to the prototype can explain the diversity of designs that we see in nature. Brush-footed butterflies (Nymphalidae) have an established and well-studied wing groundplan, and this prototype has been extended to other families of Lepidoptera (butterflies and moths). However, the establishment of homologies is limited and remains hypothetical because of the great morphological divergence between families. Here, we study swallowtails, an iconic group of butterflies with a wealth of shapes and color patterns on its wings, for which we are far from understanding the molecular basis of their development. We present the first molecular assessment of the *Papilio* wing groundplan, which we determined using a combination of pharmaceutical assays, CRISPR/Cas9 knockouts, and *in situ* expression experiments. Our results revealed that *WntA* and *Wnt6* are expressed in the margin of the developing wing and are necessary to pattern two major wing elements, the *glauca,* and a novel patterning system we dub the *submarginal spots.* Notably, the distinctive central symmetry system of nymphalids is significantly reduced or

[2] This chapter was done with the contributions of Entomology undergraduates Alan and Brian Liang, who reared and injected *Papilio* eggs and pupae, and EEB graduate student Jeanne McDonald, who synthesized the *in situ* hybridization probes.

completely absent in the swallowtails tested here, with the marginal systems have expanded proximally to become the dominant color patterning system. Overall our data revealed the action of two inductive signals (*WntA* and *Wnt6*) in patterning *Papilio* wings, with interactions between these signals establishing the most prominent elements in the papilionid groundplan. The evidence points to a very versatile and flexible patterning system in the basal lineage of butterflies.

Introduction

Butterfly wing patterns provide an extraordinary example of rapid diversification of a complex patterning system. The colors and shapes we see on butterfly wings are composed of a discrete set of homologous pattern elements[1,2], analogous in many ways to the vertebrate limb groundplan[3]. In the early twentieth century, Schwanwitsch[1,4–6] and Süffert[7] formulated conceptual frameworks for homologies of the color pattern elements in butterfly and moth wings. So far, the most recent articulation of these schemes is commonly referred to as the Nymphalid Groundplan (NGP)[1,2,8–11], which constitutes a theoretical construct of pattern homologies for the family Nymphalidae (i.e., the brush-footed butterflies). The NGP, like the vertebrate limb groundplan, leads to very specific hypothesis of pattern homology that we can now assess with some rigor, thanks to advances in phylogenetics and developmental genetics. Hence, this system provides the opportunity to study underlying developmental and evolutionary questions related to topics like trait homology[8,10,12], the evolution of gene function[11], evolutionary novelty[13], and modularity[14]. The simplicity of the groundplan model, and the great diversity of wing patterns in

Lepidoptera, has prompted many scholars to extrapolate the NGP to other butterflies and moth families[14–18]. However, the scarcity of developmental data has limited this broader comparative work to speculation. Here, we use a combination of genetic techniques to study the wing groundplan of *Papilio* butterflies, which represent Papilionidae – the swallowtail butterflies – the earliest diverging branch of the butterfly clade, and the sister group to all other butterflies[19,20] (Figure 3-1A).

Swallowtail butterflies present a broad spectrum of color wing patterns, sizes, and shapes. The speciose genus *Papilio* is of particular interest because it includes classic examples of mimicry and sex-linked melanism[21–23]. Multiple studies have been done to map the genetics behind mimicry[24–27]; however, few studies have examined the deeper wing pattern homologies of *Papilio* butterflies[2]. Schwanwitsch[4] and Süffert[7] suggested the first groundplan models for swallowtail butterflies, proposing similarities with the NGP, and Nijhout[2] later reconciled their observations and provided interesting hypotheses about papilionid pattern formation (Figure 3-1B-D). Overall, they found color patterns in Papilionidae to differ considerably from the nymphalids, and used positional information between elements, wing veins and natural variants, to define a groundplan[2,4,6]. Depending on author, subfamily, or genus, distinct elements are named differently. For instance, while Süffert and Nijhout[2] identify a *Central Symmetry System* (*CSS*) in papilionids, Schwanwitsch calls for the presence of the *Rubrae*, a different pattern (Figure 3-1B,C). In *Papilio dardanus*, a polymorphic mimetic species from Africa, and other species with very derived patterns as *Papilio polytes*, the element homology determination has been more

challenging, with assumptions about the expansions or shifts of *CSS*, *Border ocelli* (*BoSS*) and *Externa* or Marginal symmetry systems (*ESS*)[2]. We still lack comparative developmental data that allow us to assess color pattern homologies between papilionids and other butterflies confidently.

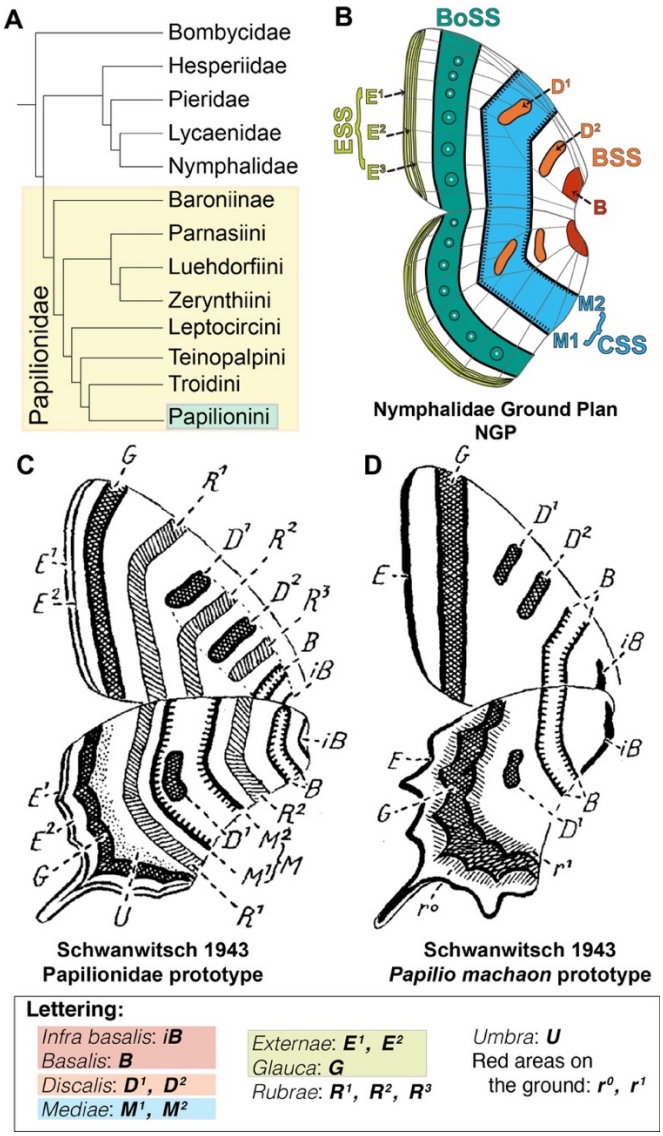


Figure 3-1. Proposed wing pattern groundplan homologies in butterflies. (A) Phylogenetic tree of Papilionidae[20] and their relationship with Nymphalidae. **(B)** Nymphalid GroundPlan derived from Schawanwitsch[4], Sutter[7], and Nijhout[2] **(C)** The Papilionidae and **(D)** *Papilio machaon* species group prototypes described by Schawanwitsch[4]. Colors in 'Lettering' represent elements in the NGP. BSS: *Basalis Symmetry System*, CSS: *Central Symmetry System*, BoSS *Border ocelli Symmetry System*, (ESS) *Externae Symmetry System*.

Comparative gene expression and knockout work has been invaluable for assessing color pattern homologies in nymphalid butterflies. In particular, transcription factors and signaling molecules, including well-known players like distal-less, spalt, Notch, engrailed, and members of the Wnt secreted signaling proteins, have been used to draw homologies between major color pattern systems, even in challenging lineages where color pattern have diverged from the prototypical groundplan[2,28,11]. The Wnt family of secreted proteins is of special interest; the *wg*, *Wnt6*, *Wnt10*, *WntA* genes have all been associated with different color pattern elements in nymphalids[8,10,11,29]. Based on this information, we wanted to explore the *Wnt* family of ligands as a basis for improving our color pattern groundplan model for *Papilio* butterflies. We used *in situ* hybridization, drug perturbation essays, and CRISPR/Cas9 knockouts of *WntA* and *Wnt6*, to obtain insights about the wing diversification of this insect group and the developmental process behind it.

Methods

Papilio sampling

We sampled species distributed throughout the *Papilio* phylogeny[20], specifically: *Papilio zelicaon*, *Papilio polyxenes*, *Papilio glaucus*, *Papilio troilus,* and *Papilio cresphontes* (Figure 3-2A-E). Information about the population origin, caterpillar, and adult diet is reported in Table S1. All colonies were kept in a growth chamber at 26ºC, 60% relative humidity, 16/8 h day/night cycle during egg collection.

Drug injections

Pupal injections of heparin and dextran sulfate were done as previously described[10,30]. Each individual was injected between 4-20 hours after pupation (Table S2), in the basal wing region, taking care to avoid tissue damage. We used hand-pulled glass capillary needles mounted on a micropipette to deliver the drug. Pupae were reared under the conditions described above.

In situ hybridization

We dissected imaginal discs from the last caterpillar instar from *P. zelicaon* in cold PBS. Imaginal discs were stored in TRIzol™ Reagent (Invitrogen™) to extract total RNA using the PureLink™ RNA Mini Kit. cDNA synthesis was then performed using M-MuLV reverse transcriptase (New England Labs). PCR oligos were designed based on conservation between the published *Papilio* genomes (lepbase.org, Table S3). After amplification, pGEM TA-cloning was made, with subsequent plasmid verification using sanger sequencing. The plasmids were linearized and used to create DIG-labeled riboprobe synthesis following the manufacturer's instructions (Roche). To perform the *in situ* hybridization of the DIG-riboprobes, we dissected imaginal discs, then fixed and stained, following a previous protocol [10].

Genetic manipulation

We amplified and sequenced *WntA* and *Wnt6* from *P. zelicaon* and *P. polyxenes* genomic DNA and cDNA to aid the design of CRISPR/Cas9 sgRNAs. We verified sequence conservation and used the same sgRNAs for both species (Table S3). Freshly

laid eggs were treated with 5% benzalkonium chloride (Sigma-Aldrich) for 90 seconds, then thoroughly washed with water. Eggs were glued to a glass slide. We injected 2 or 4 sgRNAs per gene to generate large deletions that could knock out the gene function, as previously described[11]. For each injection, a mix of Synthego synthetized sgRNAs (final 250 ng/ul each) plus recombinant Cas9 protein (0.5ug/ul, PNA Bio) were injected using aluminosilicate glass needles (Sutter). Hatched caterpillars were reared as described above. G0 butterflies were inspected for mutant phenotypes; in this case, the presence of mutant clones in the wings. For adults showing distorted phenotypes were pinned and photographed using a Canon EOS 60D, with lens Canon EF 100mm f/2.8L USM Macro Lens.

Results

Wing margin morphogen sources set Papilio symmetry systems

Several major NGP color pattern elements described as "symmetry systems" are induced by short- or medium-range Wnt signaling[8,10,31]. Importantly, it has been shown that this Wnt signaling can be manipulated by pharmaceutical treatment, leading to strong alterations of Wnt-associated color patterns[10,30,32,33]. Because of this highly pattern-specific effect, we can use drug injections as a tool to identify which color patterns are likely to be induced by Wnt signaling, and by extension, may be homologous with Wnt-derived NGP elements. Heparin and dextran sulfate are known to affect the activity of signaling pathways[10,30,34]. Specifically, heparin acts like heparan sulfate proteoglycans in the cells, which stabilize Wnt signaling proteins and promotes gradient formation [35], hence increasing concentrations of this drug will

expand Wnt's range of effect, and simulate gain-of-function in neighboring cells. Conversely, dextran sulfate works through an unknown mechanism to produce the opposite effect – reduction of Wnt-induced color patterns [10,30].

Figure 3-2. The effect of heparin perturbations in *Papilio* color patterns. Wild type and heparin induced alteration in ventral wing patterns of males of *P. cresphontes* (A, F, K), *P. troilus* (B, G, L), *P. glaucus* (C, H, M), *P. polyxenes* (D, I, N) and *P. zelicaon* (E, J, O). Dot-lines are veins landmarks fore reference, and arrows, indicate the pattern shift.

We sought to test the consequences of heparin injections in five *Papilio* species (Figure 3-2A-E). These treatments presented a range of changes in the mid-wing and margin patterns but not in the basal region of the wings (Figure 3-2 F-J). Here we

summarize the effects we observed on the major pattern elements in *Papilio* (Figure 3-1C-D), following the nomenclature of Schwanwitsch [4,6]:

Basalis: This element at the base of the wing appeared to be unaffected. This element seems to cover a sizable portion of the proximal wing in *Papilio*, but is much smaller in nymphalids (Figure 3-1B).

Discalis: Elements D1 and D2, are present in *Papilio*, and are the most straightforward elements to assign homology due to their conserved position associated with venation in the discal cell. In the hindwing, only D1 is present for some of the species and its color seems to vary with the *Glauca* coloration (see below), either plain black (P. *glaucus*), decorated with some blue scales (*P. cresphontes*), or red/yellow (*P. polyxenes*). The *Discalis* elements were not affected by heparin treatment.

Glauca (a.k.a. *Externa III*): Usually this is a wide black band with blue scales that runs parallel to the wing margin. Heparin injections caused the most striking effects in this pattern, reducing and inducing different color scales (Figure 3-2K-O). Minor effects start with a scattering of blue scales and a proximal movement of the black band, which suppresses the red/yellow region described by Schwanwitsch as the *Umbrae* (Figure 3-2K, M). In some stronger cases, there is a drastic reduction of the melanic pattern (Figure 3-2N), allowing the extension of the color pattern from the distal neighbor spots (Figure 3-2L, N, O). These results are species-specific but were also associated with relative injection time and concentration. Hence, injections close to the

moment of pupation and higher amounts displayed stronger color pattern perturbations (Figure 3-3) – in this case spreading melanin across most of the wing. Surprisingly, in *P. polyxenes,* the heparin treatment revealed sex-dependent effects. This species is sexually dimorphic dorsally (Figure 3-3A), but not ventrally. However, in the ventral side of females, the blue and black patterns seemed to be dominant over the yellow and red elements, contrasting with male effects (Figure 3-3B), suggesting somewhat different pathways specifying similar wing designs between females and males.

Figure 3-3. Variation in phenotypic effect and sex-specific response to heparin injections in *P. polyxenes*. (A) Sexual dimorphism in dorsal wings and (B) ventral wing pattern in wild type individuals. (C) Ventral wings of butterflies injected with heparin.

Submarginal spots: Distal to the *Glauca*, there is a row of spots running along the wing margin. These spots have long considered to be background color showing in the space between the *Glauca* and the *Externa* elements. However, in butterflies in which heparin had strong effects, we were surprised to observe these spots expand proximally and express dominantly to displace the *Glauca*. The exception to this is in

P. cresphontes, where the spots are enlarged and do not extend upon heparin treatment. Movement and replacement of the color patterns occurred from the margin to the proximal region of the wing, implying the extension of wing margin morphogen gradients.

WntA, a key groundplan marker, patterns the Glauca in Papilio butterflies

WntA is expressed in up to three parallel regions in the imaginal discs of nymphalids that show the prototypical NGP organization[10,11]. Across species, *WntA* signaling activity is enhanced by heparin, expanding the symmetry systems of the NGP. Thus, our results from heparin injections in *Papilio* led us to speculate that *WntA* may also play a role in *Papilio* wing patterning. To test this further we looked at *WntA* spatial expression patterns in wing imaginal discs, and knocked out *WntA* function using CRISPR/Cas9, in two species: *P. zelicaon* (Figure 3-4A-B) and *P. polyxenes* (Figure 3-4C-E). Our results unambiguously confirm that *WntA* is necessary to pattern the *Glauca* pattern in these both species. *Glauca* is one of the most prominent elements in *Papilio* wings, and frequently displays iridescent blue scales (Figure 3-2A-E). *WntA* mRNA is observed as a relatively thick band close to the margin of the imaginal disc in forewings and hindwings in *P. zelicaon* (Figure 3-4B). In *WntA* knockouts, *Glauca* is lost, leaving yellow and/or red background color. The yellow background is evident in forewings of *P. zelicaon* (Figure 3-4A), and on the dorsal side of *P. polyxenes* males, along with the ventral forewings (Figure 3-4C). The red background pattern is apparent in the ventral hindwings of both species (Figure 3-3 C- E). The loss-of-function CRISPR results follow a similar trend to the observations after gain-of-

function treatment with higher concentrations of heparin, which is contradictory, because heparin extends morphogens range of action.

Figure 3-4. *WntA* specifies the *Glauca* color pattern system in *Papilio*. (A) *P. zelicaon* phenotype after *WntA* CRISPR/Cas9 mosaic knockout (mKO). **(B)** *WntA* in-situ hybridization mid-last instar imaginal discs. **(C)** Comparison of *P. polyxenes* wild type and mKO phenotypes. Close-up of **(D)** male and **(E)** female phenotypes, cyan shade highlights the melanic regions in the WT that are affected in the knockouts.

However, this could be explained by *WntA* repressing an unidentified neighboring morphogen source, which spreads after either removing *WntA* gene expression or when heparin concentrations are high enough – these results may provide evidence of at least two signaling sources in *Papilio*. Interestingly, *WntA* knockouts showed no sex differences in *P. polyxenes* ventral hindwings, contrary to the observed disparities in the heparin treatments (Figure 3-3), pointing to a different signaling molecule responding in the female-specific patterning. Finally, we did not detect effects in the basal portion of the wing in *WntA* knockouts, but we saw a slight reduction in the size of the *submarginal spots*, discussed below.

Papilio submarginal spots

The heparin injections and *WntA* knockouts described above suggest the presence of a previously unidentified wing pattern system between the *Glauca* and the *Externa* bands. In butterflies, the wing groundplan is typically described in terms of pattern elements over a 'plain' background[36]. The color patterns are usually pigmented designs determined by concentration levels of various inductive molecules[2,10,30,37]. The color seen when these molecules are removed is considered the background. After knocking out *WntA*, we noticed a row of round or semi-rectangular submarginal spots running along the margins of both forewings and hindwings (Figure 3-4A), but showing yellow coloration similar to the background. In addition, these spots were slightly compressed distally in the knockouts, while in the heparin treatment they extended proximally. We used dextran sulfate injections, which work to antagonize Wnt

signaling, to see if these spots would show a pattern of reduction consistent with a role for Wnt signaling[10,30]. This drug treatment was very toxic to the butterflies, however we obtained enough individuals to assess phenotypes in *P. zelicaon* (Figure 3-5A). We observed a distal movement of the *Glauca* with a reduction, in some cases severely, of the *submarginal spots* (Figure 3-5A). Thus, our results collectively suggest a non-*WntA* Wnt ligand could underlie specification of the *submarginal spots* – i.e., drugs that modulate *Wnt* signaling affect the spots, but *WntA* knockouts do not cause loss of the spots.

Figure 3-5. *Wnt6* contributes to patterning submarginal spots in *Papilio*. (A) *P. zelicaon* wild type and dextran sulfate (D.S.) phenotypes in ventral view of male, cyan shades highlight *Glauca* and submarginal spots in wild type, while magenta shows the *Glauca* expansion and the spot reduction from the drug treatment. **(B)** *P. zelicaon Wnt6* CRISPR/Cas9 mosaic knockout phenotype. **(C)** *Wnt6* in situ hybridization in the margin of the forewing, where the submarginal spots are observed. **(D)** *Wnt6* in situ hybridization in the margin of the hindwing and the reduction of the submarginal spots after CRISPR/Cas9 knockout. **(E)** Closeup illustrating *KO* effects on submarginal spots. **(F)** Wnt6 mosaic knockout in *P. polyxenes* shows reduction of sumarginal spots, as highlighted in hindwing closeup **(G)**. Arrows and magenta shading marks mutant phenotypes, while blue indicates wild type patterns. Dashed lines represent homologous vein landmarks between imaginal discs and adult wings.

Several *Wnt* gene are known to be expressed along with wing margins of nymphalid butterflies, thus providing a handful of candidate factors for *submarginal spot* development. We examined expression of one of these genes, *Wnt6*, and found that in *P. zelicaon*, *Wnt6* is transcribed along the border lacuna (future wing margin) of the forewings and hindwings, although with different shapes (Figure 3-4 B-D). In the forewings *Wnt6* is expressed in discrete spots in the center of each wing cell (i.e., the space between veins). In contrast, *Wnt6* expression in the hindwings is elongated and covers most of the width of the wing cells. These differences in *Wnt6* expression correlates with the oval and semi-rectangular *submarginal spot* shapes found in forewings and hindwing, respectively. We next used CRISPR/Cas9 deletion to test the expression of *Wnt6* during color pattern development. *Wnt6* knockouts caused a drastic reduction of the *submarginal spots* in both *P. zelicaon* and *P. polyxenes* (Figure 3-4B-G). The mutant phenotypes consisted of changes in shape and size of the spots, but not color, similar to the *WntA* effect. Our data highlight these *submarginal spots* in *Papilio* as true, dominant pattern elements instead of 'plain' background, where *WntA* and *Wnt6* are positive regulators of element size and shape. We annotated the *submarginal spots* in different *Papilio* subgenera (Figure S3-6). Although further comparative work will tell us more about this pattern system, this comparison shows a wide diversity of colors and sizes, pointing to a very versatile and flexible patterning system that deserves more attentions in an evo-devo context.

Finally, one puzzling element is the single 'eyespot'-like pattern widely seen along the margin, at the base of the hindwing tail in many *Papilio* species. This eyespot does not

seem to be affected by our *WntA* or *Wnt6* knockouts, and it does not respond to heparin except at very high concentrations. However, it does show a slight reduction in response to dextran sulfate, suggesting that one or more signaling ligands maybe influence the pattern, even if they aren't directly inducing it.

Discussion

The groundplan depictions of butterfly and moth wing color patterns are theoretical models that allow us to generate and test hypotheses related to morphological development and evolution. These models were originally derived from observations of intra- and interspecies variation, however including genetic and developmental data allow us to probe a deeper set of evolutionary questions including the molecular basis convergence, evolution of gene function in pattern diversification, and the genetic origins of evolutionary novelties in this megadiverse group of insects. Here we present a revised model for papilionid wing patterns based on pharmacologic manipulation of morphogen gradients, in situ expression in imaginal discs, and knockouts of two *Wnt* genes. This model is a first step to gain insights into the pattering of a butterfly family with extensive examples of pattern variation, aposematism, and sexual dimorphic wing variation. Furthermore, since papilionids are the oldest diverged clade of butterflies[19], this work provides an important phylogenetic point of reference for future studies looking at the origin and evolution of the wing pattern groundplan itself.

Evidence for a novel proximodistal boundary in Papilio

The major groundplan elements described for both nymphalids and papilionids are: *Basalis*, *Discalis* (*D1*, *D2*), *Central Symmetry System* (*CSS*), *Border Ocelli Symmtery System* (*BoSS*), and the *Externa* or marginal symmetry system (*ESS*)[1,4,6,8]. *Basalis* and *Discalis* were not affected by any of our heparin or CRISPR treatments, suggesting that they may not be induced by Wnt signaling. In the case of *Discalis* patterns, this is somewhat surprising since both *D1* and *D2* are marked by strong *wingless* (i.e., *Wnt1*) expression in nymphalids[8]. Interestingly, in all *Papilio* we surveyed, heparin injection phenotypes showed a hard spatial boundary near the base of the wing, across which presumptive morphogen-induced pattern induction never extended (Figure 3-2) – i.e., heparin effects like pattern expansion never crossed this boundary. This boundary has also been observed in some cold-shock experiments in *P. glaucus*[38], drug perturbations in *P. xuthus*[39] and *P. machaon*[40], the naturally occurring "*fletcheri*" phenotype in *P. canadensis* and *P. glaucus*[41], and other spontaneous aberrations of *P. machaon*[42]. Because this boundary has not been detected in nymphalids using heparin treatments, we speculate that this may represent a novel proximodistal pattern boundary in papilionid butterflies (Figure 3-6).

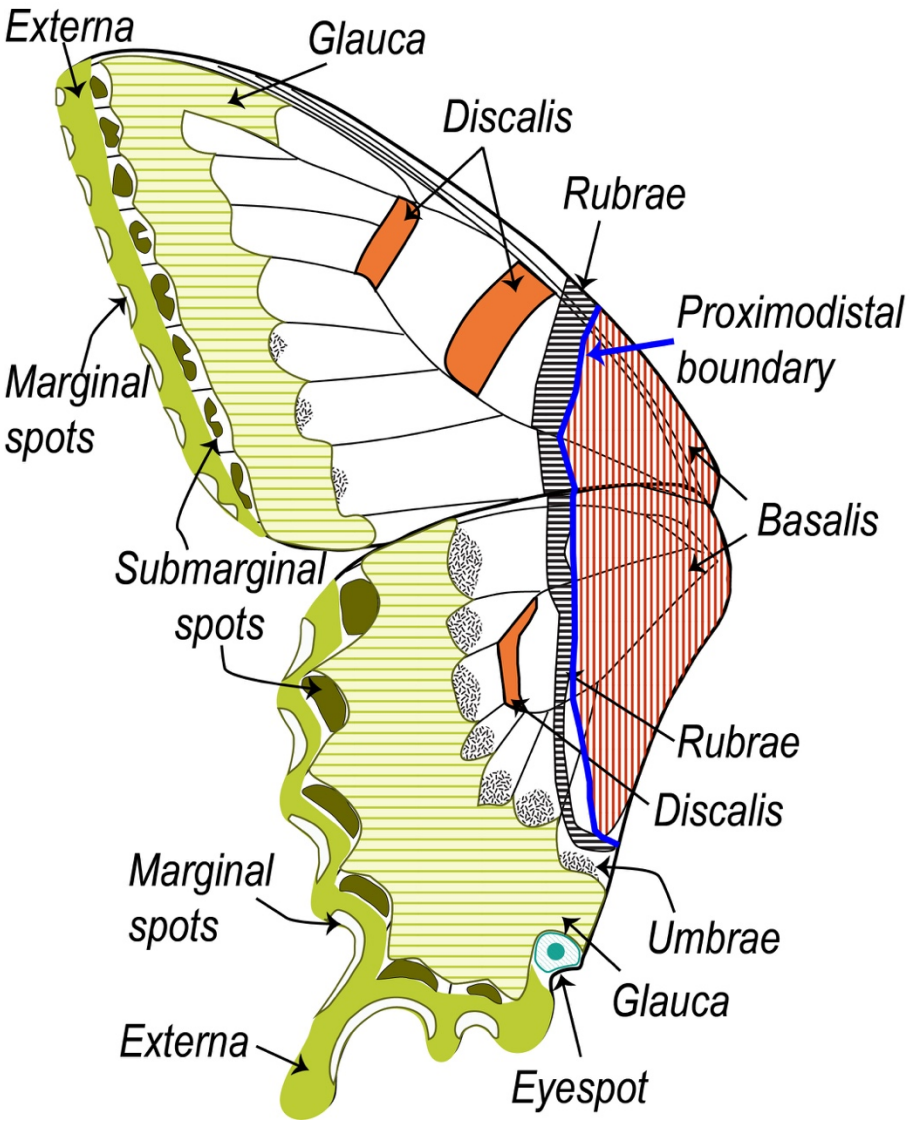


Figure 3-6. Revised model of the *Papilio* wing pattern groundplan.

Loss of the central symmetry system in Papilio

The central symmetry system is perhaps the largest, most diverse, and most important pattern systems in nymphalid butterflies. It is proposed to have a very deep origin in Lepidoptera since similar color patterns are commonly seen throughout the moths[2,16–18]. Remarkably, we found little evidence of the central symmetry system in *Papilio*.

WntA is the key marker for the central symmetry system in nymphalids, however none of our *WntA* expression or perturbation data had any bearing on color patterns in the middle of the *Papilio* wings where the central symmetry system would be expected. In contrast, *WntA* expression only occurs in the wing margins in the *Papilio* species we looked at. Thus, we speculate that the central symmetry system has been lost in *Papilio*, and other patterns have taken over the role of coloring the central wing, including the *Discalis* elements. In this respect it is interesting to note that early *Papilio* color pattern models recognized a different set of central stripes called the *Rubrae*[4,6] that were proposed to be distinct from the central symmetry system, even though they are superficially similar. One final note is that there may be small remnants of the central symmetry system remaining in some *Papilio*, for example the extra central stripe patterns in the hindwing of *Papilio glaucus*. Further work, however, will be required to characterize and fully understand the evolution and development of the *Rubrae* and central symmetry systems in papilionids. Nevertheless, all evidence suggests that the central symmetry system is not a major component of the papilionid groundplan.

***Papilio* pattern diversity is largely derived from the Glauca and submarginal spots**

The most striking results from experimental manipulations were in the *Glauca*. This element was renamed by Schwanwitsch (1956) as *Externa III* and synonymized by Nijhout (1991) as "parafocal elements" in nymphalids. Our knockouts show *WntA* to be necessary to induce the *Glauca*, and when *Glauca* is lost a yellow/red background is observed. In contrast, *WntA* knockouts in nymphalids have much subtler effects on

the *Externa* elements[11] – the *Externa III* shift distally and are misshapen, which suggest a different developmental underpinning that is influenced by, but not wholly reliant, on *WntA*. At this point it is unclear if there is direct homology between the *Papilio Glauca* and the nymphalid *Externa III* elements. That said, immunostainings of the proteins *spalt* and *Distal-less* in the imaginal disc of *P. machaon*[43], a close relative of *P. zelicaon* and *P. polyxenes*, show these proteins are expressed along the wing margin. Both *spalt* and *Distal-less* are involved in the formation of *Externa I, II,* and *III* in nymphalids, with *spalt* being necessary for the establishment of the *Externa III* band[28,44]. Therefore, there may be some ancient homology, although, if this is true, evidence suggests that there has been further developmental elaboration of Externa in *Papilio*.

Distal to the *Glauca*, there is a row of spots that constitutes a dominant pattern system that we here term the *submarginal spots* (Figure 3-5). These spots play an essential role in mimetic species – knockdown *Wnt6* function in *Papilio polytes* with small interfering RNA[45] obtained similar results to our *Wnt6* CRISPR/Cas9 experiments (Figure 3-5), i.e., the reduction in size and shape change of the spots. By examining a small sample of *Papilio* species (Suppl. Figure 3-1), we noticed the considerable diversity of colors and sizes of these spots, which also are involved in mimicry in other species besides *P. polytes*. We conclude that the interplay between *submarginal spots* and the *Glauca* drives most of the color pattern diversity we observe in *Papilio* butterflies. Both *Wnt6* and *WntA* show some overlapping expression in the wing margin, where their gradients appear to extend proximally, and their interactions likely

regulate the shape and size of the *submarginal spots*. Interestingly, mathematical models for *P. dardanus* and *P. polytes* predict that a single gradient originating from the wing margin is sufficient to explain color patterns in these species[46]. Our data largely agree with the central importance of wing margin pattern induction; however, we show clear evidence that gradients of at least two factors – *WntA* and *Wnt6* – interact to induce the *Papilio* color patterns.

Conclusion

Wnt genes expressed along the wing margin are essential for formation of the major color patterns in *Papilio* butterflies. Specifically, *WntA* is necessary to pre-pattern major elements during imaginal disc development in both papilionids and nymphalids, however the major *WntA* patterning systems are not likely homologous between these families, and represent different developmental processes. Here, we also characterize a novel *submarginal spot* pattern system in *Papilio*, in which *Wnt6* signaling plays a promoting role, and could be homologous to the Externa patterns in nymphalids[43,44]. Overall, our findings suggest that much of the color pattern complexity and diversity in *Papilio* can be traced to the interaction between the *WntA Glauca* patterns and the *Wnt6* submarginal patterns, with some occasional ornamentation thanks to *Discalis*, *Rubrae*, and/or highly reduced *CSS* patterns that are not yet well understand in Papilionidae.

Acknowledgments

We thank to Yiwen Zhu for collecting, caring and shipping *Papilio* eggs and larvae from California. To Arnaud Martin and members of the R.D.R. laboratory for discussion and input. This work was supported by National Science Foundation.

REFERENCES

1. Schwanwitsch, B. N. On the ground-plan of wing-pattern in nymphalids and certain other families of the Rhopaloeerous Lepidoptera. *Proc. Zool. Soc. Lond.* **94**, 509–528 (1924).
2. Nijhout, H. F. *The development and evolution of butterfly wing patterns*. (Smithsonian Institution Press, 1991).
3. Hinchliffe, J. R. Developmental basis of limb evolution. *Int. J. Dev. Biol.* **46**, 835–845 (2002).
4. Schwanwitsch, B. N. Wing-pattern in papilionid Lepidoptera. *Entomologist* **76**, 201–203 (1943).
5. Schwanwitsch, B. N. Evolution of the wing-pattern in the Lycaenid Lepidoptera. *Proc. Zool. Soc. Lond.* **119**, 189–263 (1949).
6. Schwanwitsch, B. N. Color-pattern in Lepidoptera. *Entomologicheskoe obozrenie* **35**, 530–546 (1956).
7. Süffert, F. Zur vergleichende analyse der schmetterlingszeichnung. *Biologisches Zentralblatt* **47**, 385–413 (1927).
8. Martin, A. & Reed, R. D. *Wingless* and *aristaless2* define a developmental groundplan for moth and butterfly wing pattern evolution. *Mol Biol Evol* **27**, 2864–78 (2010).
9. Otaki, J. M. Color pattern analysis of nymphalid butterfly wings: revision of the nymphalid groundplan. *Zoological Science* **29**, 568–576 (2012).

10. Martin, A. & Reed, R. D. Wnt signaling underlies evolution and development of the butterfly wing pattern symmetry systems. *Dev Biol* **395**, 367–78 (2014).
11. Mazo-Vargas, A. *et al.* Macroevolutionary shifts of *WntA* function potentiate butterfly wing-pattern diversity. *Proc Natl Acad Sci USA* **114**, 10701–10706 (2017).
12. Monteiro, A., Glaser, G., Stockslager, S., Glansdorp, N. & Ramos, D. Comparative insights into questions of lepidopteran wing pattern homology. *BMC Dev Biol* **6**, 52 (2006).
13. Monteiro, A. Alternative models for the evolution of eyespots and of serial homology on lepidopteran wings. *BioEssays* **30**, 358–366 (2008).
14. Suzuki, T. K. Modularity of a leaf moth-wing pattern and a versatile characteristic of the wing-pattern groundplan. *BMC Evol Biol* **13**, 158 (2013).
15. Gardiner, A. J. & Terblanche, R. F. A suggested terminology for the groundplan of the lycaenid wing pattern using the genus *Erikssonia* Trimen (Lepidoptera: Lycaenidae) as an example. *African Entomology* **18**, 166–170 (2010).
16. Schachat, S. R. & Brown, R. L. Forewing color pattern in Micropterigidae (Insecta: Lepidoptera): homologies between contrast boundaries, and a revised hypothesis for the origin of symmetry systems. *BMC Evol Biol* **16**, 116 (2016).
17. Schachat, S. R. Symmetry systems on the wings of *Dichromodes* Guenée (Lepidoptera: Geometridae) are unconstrained by venation. *PeerJ* **8**, e8263 (2020).
18. Gawne, R. & Nijhout, H. F. The Arctiid archetype: a new lepidopteran groundplan. *Front. Ecol. Evol.* **8**, (2020).

19. Espeland, M. *et al.* A comprehensive and dated phylogenomic analysis of butterflies. *Curr Biol* **28**, 770-778 e5 (2018).
20. Allio, R. *et al.* Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail butterfly evolution. *Systematic Biology* **69**, 38–60 (2020).
21. Brower, J. V. Z. Experimental studies of mimicry in some North American butterflies: Part II. *Battus philenor* and *Papilio troilus*, *P. polyxenes* and *P. glaucus*. *Evolution* **12**, 123–136 (1958).
22. Clarke, C. A. & Sheppard, P. M. The evolution of mimicry in the butterfly *Papilio dardanus*. *Heredity* **14**, 163–173 (1960).
23. Krebs, R. A. & West, D. A. Female mate preference and the evolution of female-limited batesian mimicry. *Evolution* **42**, 1101–1104 (1988).
24. Kunte, K. *et al. doublesex* is a mimicry supergene. *Nature* **507**, 229–32 (2014).
25. Nishikawa, H. *et al.* A genetic mechanism for female-limited batesian mimicry in *Papilio* butterfly. *Nat Genet* **47**, 405–9 (2015).
26. Palmer, D. H. & Kronforst, M. R. A shared genetic basis of mimicry across swallowtail butterflies points to ancestral co-option of *doublesex*. *Nat Commun* **11**, 6 (2020).
27. Thompson, M. J., Timmermans, M. J., Jiggins, C. D. & Vogler, A. P. The evolutionary genetics of highly divergent alleles of the mimicry locus in *Papilio dardanus*. *BMC Evol Biol* **14**, 140 (2014).
28. Zhang, L. & Reed, R. D. Genome editing in butterflies reveals that *spalt* promotes and *Distal-less* represses eyespot colour patterns. *Nat Commun* **7**, 11769 (2016).

29. Carroll, S. B. *et al.* Pattern formation and eyespot determination in butterfly wings. *Science* **265**, 109–14 (1994).
30. Serfas, M. S. & Carroll, S. B. Pharmacologic approaches to butterfly wing patterning: sulfated polysaccharides mimic or antagonize cold shock and alter the interpretation of gradients of positional information. *Dev Biol* **287**, 416–24 (2005).
31. Martin, A. & Courtier-Orgogozo, V. Morphological evolution repeatedly caused by mutations in signaling ligand genes. in *Diversity and Evolution of Butterfly Wing Patterns - an Integrative Approach* (eds. Sekimura, T. & Nijhout, H. F.) 59–87 (Springer, Singapore, 2017).
32. Sourakov, A. Leopards and giants, tigers and woolly bears: casting a broader net in exploring heparin effects on Lepidoptera wing patterns. *F1000Research* **7**, (2018).
33. Sourakov, A. & Shirai, L. T. Pharmacological and surgical experiments on wing pattern development of Lepidoptera, with a focus on the eyespots of saturniid moths. *Trop. Lepid. Res.* **30**, 4–19 (2020).
34. Perrimon, N. & Bernfield, M. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* **404**, 725–728 (2000).
35. Fuerer, C., Habib, S. J. & Nusse, R. A study on the interactions between heparan sulfate proteoglycans and Wnt proteins. *Developmental Dynamics* **239**, 184–190 (2010).
36. Otaki, J. M. Physiologically induced color-pattern changes in butterfly wings: mechanistic and evolutionary implications. *J Insect Physiol* **54**, 1099–112 (2008).

37. Otaki, J. M. Color-pattern analysis of eyespots in butterfly wings: a critical examination of morphogen gradient models. *Zoological Science* **28**, 403–413 (2011).

38. Perlman, D. L. & Perlman, M. P. An investigation into the effects of coldshock on the eastern tiger swallowtail butterfly *Pterourus* (*Papilio*) *glaucus* (C. Linnaeus 1758) (Lepidoptera: Papilionidae). *Holarctic Lepidoptera* **13**, 1–43 (2019).

39. Umebachi, Y. & Osanai, M. Perturbation of the wing color pattern of a swallowtail butterfly, *Papilio xuthus*, induced by acid carboxypeptidase. *Zoolog Sci* **20**, 325–31 (2003).

40. Dabrowski, J. S. & Dobranski, W. Wing pattern variation in diurnal butterflies received by experimental research, with special reference to intrapupae injections. *Atalanta* **27**, 657–664 (1996).

41. Scriber, J. M., Romack, H. & Deering, M. D. Aberrant color patterns in the *Papilio* and an update on the semi-melanic '*Fletcheri*' variants, including females (Lepidoptera: Papilionidae). *Journal of the Lepidopterists' Society* **63**, 118–126 (2009).

42. Spengel, J. W. Über einige Aberrationen von *Papilio machaon*. *Zoologische Jahrbücher. Abteilung für Systematik, Geographie und Biologie der Tiere* **12**, 337–384 (1899).

43. Shirai, L. T. *et al.* Evolutionary history of the recruitment of conserved developmental genes in association to the formation and diversification of a novel trait. *BMC Evol Biol* **12**, 21 (2012).

44. Reed, R. D., Selegue, J. E., Zhang, L. & Brunetti, C. R. Transcription factors underlying wing margin color patterns and pupal cuticle markings in butterflies. *EvoDevo* **11**, 10 (2020).

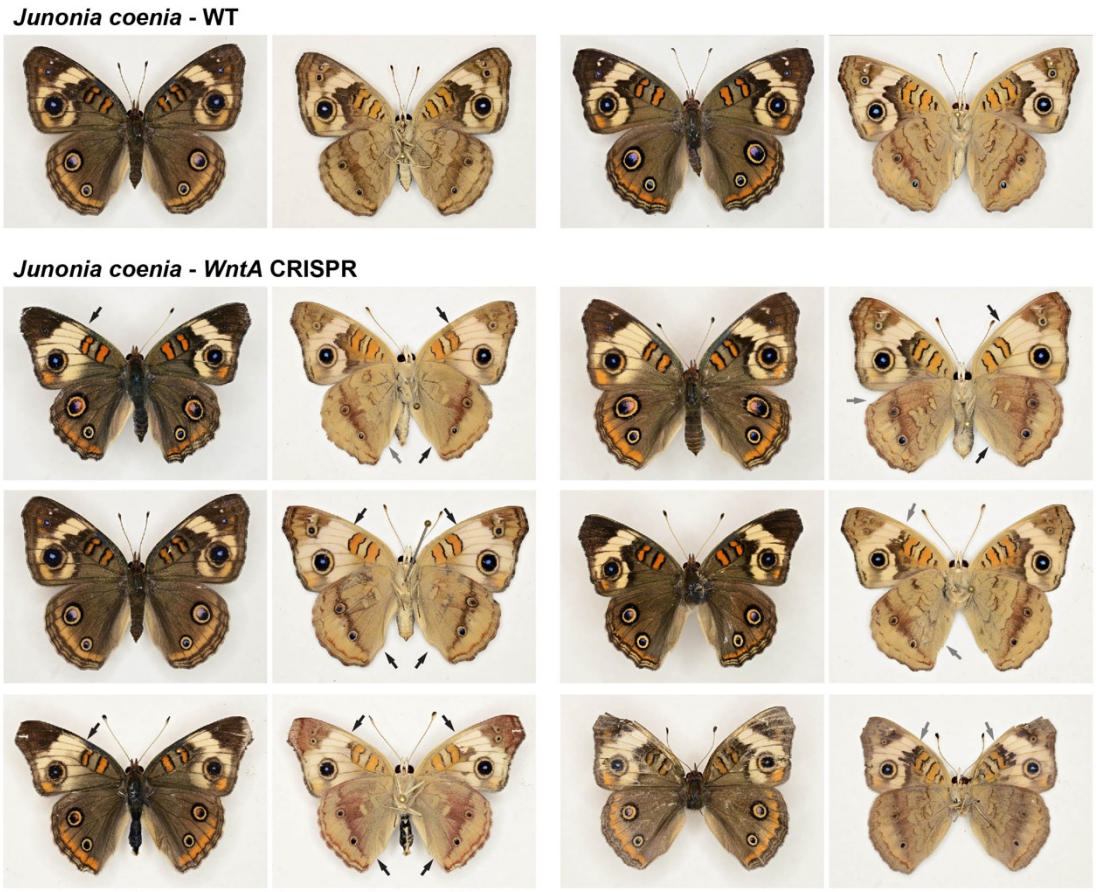
45. Iijima, T., Yoda, S. & Fujiwara, H. The mimetic wing pattern of *Papilio polytes* butterflies is regulated by a *doublesex*-orchestrated gene network. *Commun Biol* **2**, 257 (2019).

46. Sekimura, T., Madzvamuse, A. & Maini, P. K. Pigmentation pattern formation in butterfly wings: global patterns on fore- and hindwing. in *Mathematical Modeling of Biological Systems* (eds. Deutsch, A., Brusch, L., Byrne, H., Vries, G. de & Herzel, H.) 142–148 (Birkhäuser Boston, 2007). doi:10.1007/978-0-8176-4558-8_12.

APPENDIX

Appendix for Chapter 1

Figure S1-1. *WntA* mKO yields losses of several nymphalid groundplan pattern elements in *Junonia coenia.* Dorsal and ventral views of G0 butterflies are shown, with arrows pointing at wing surfaces showing a phenotype. Darker arrows indicate stronger effects.



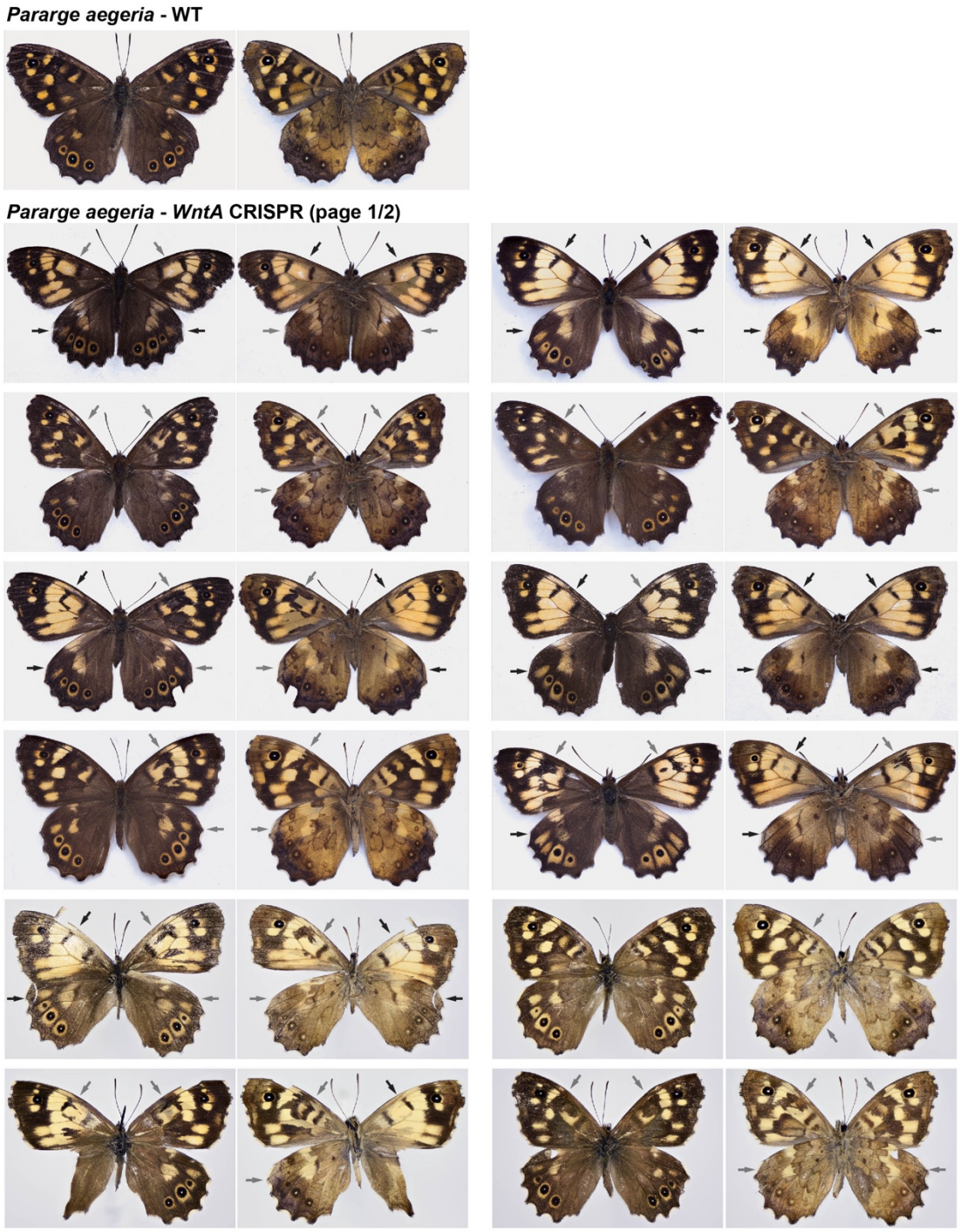


Figure S1-2. ***WntA*** **mKO yields losses of several nymphalid grounplan pattern elements in Pararge aegeria.** Dorsal and ventral views of G0 butterflies are shown, with arrows pointing at wing surfaces showing a phenotype. Darker arrows indicate stronger effects.

Figure S1-2 continuation

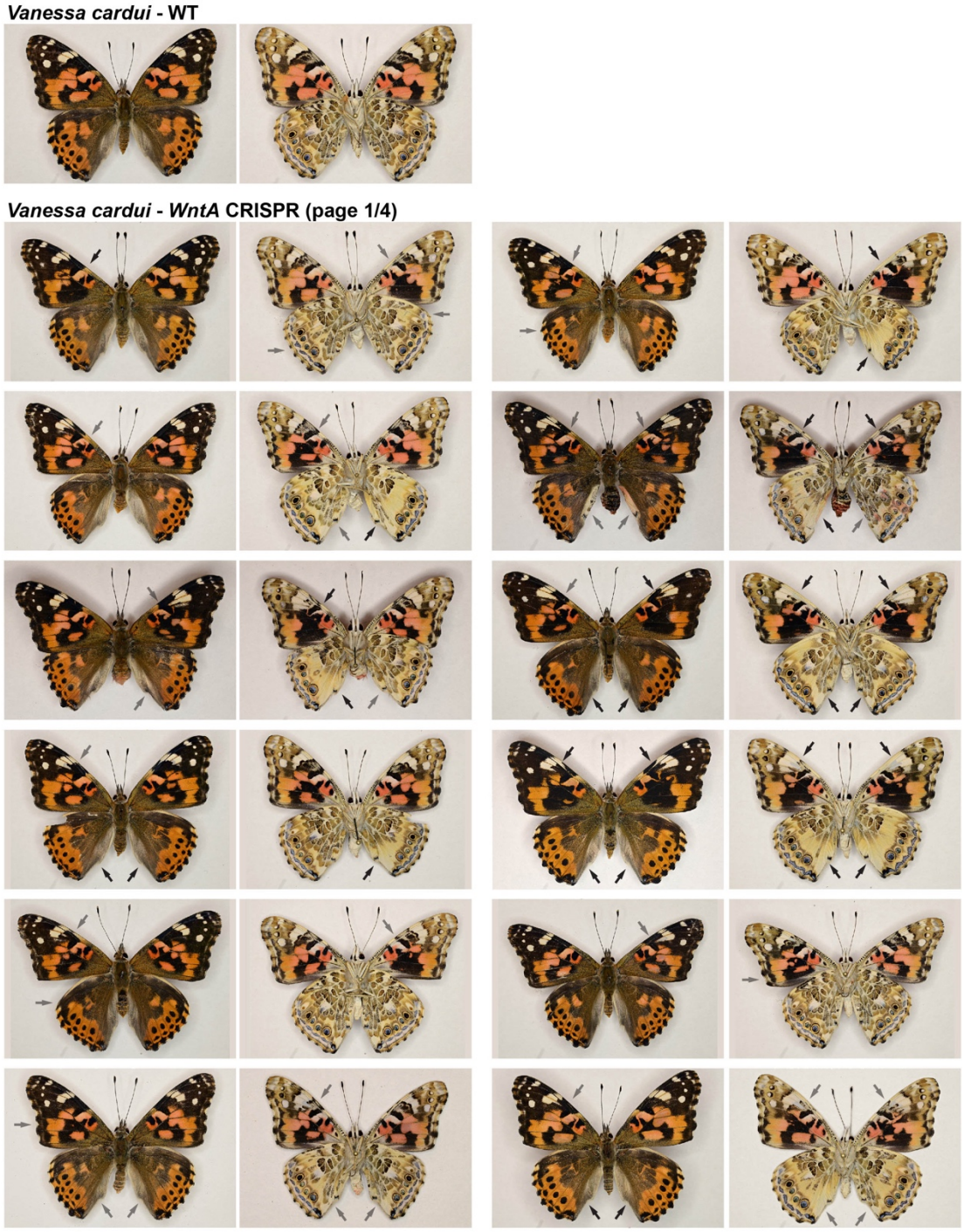


Figure S1-3. ***WntA*** **mKO yields various, discrete wing pattern phenotypes in** ***Vanessa cardui***. Dorsal and ventral views of G0 butterflies are shown, with arrows pointing at wing surfaces showing a phenotype. Darker arrows indicate stronger effects.

Figure S1-3. Continuation.

Figure S1 3. Continuation.

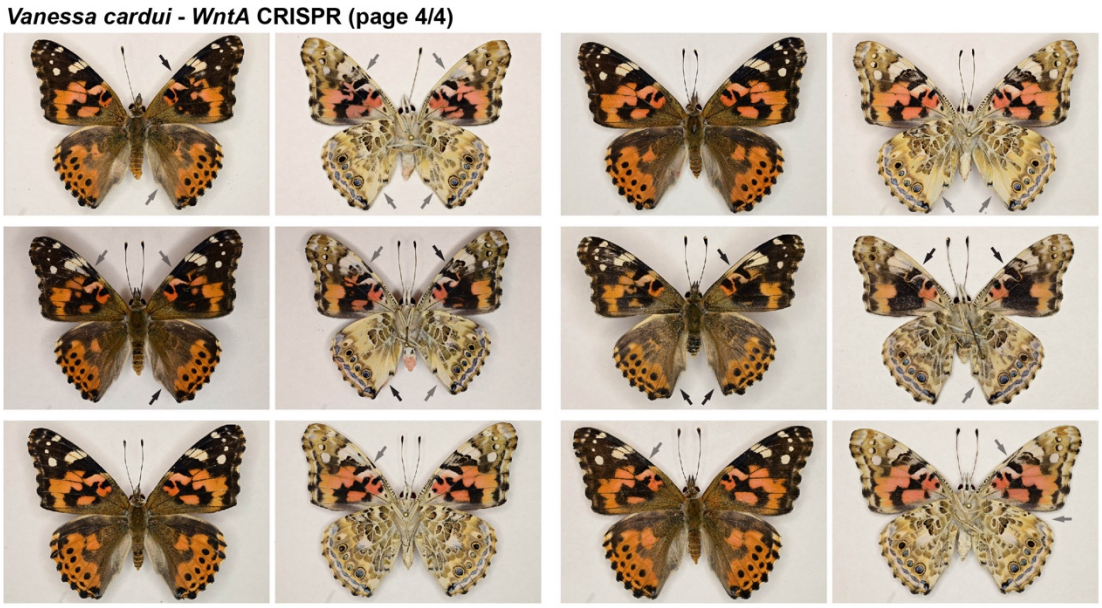


Figure S1 3. Continuation.

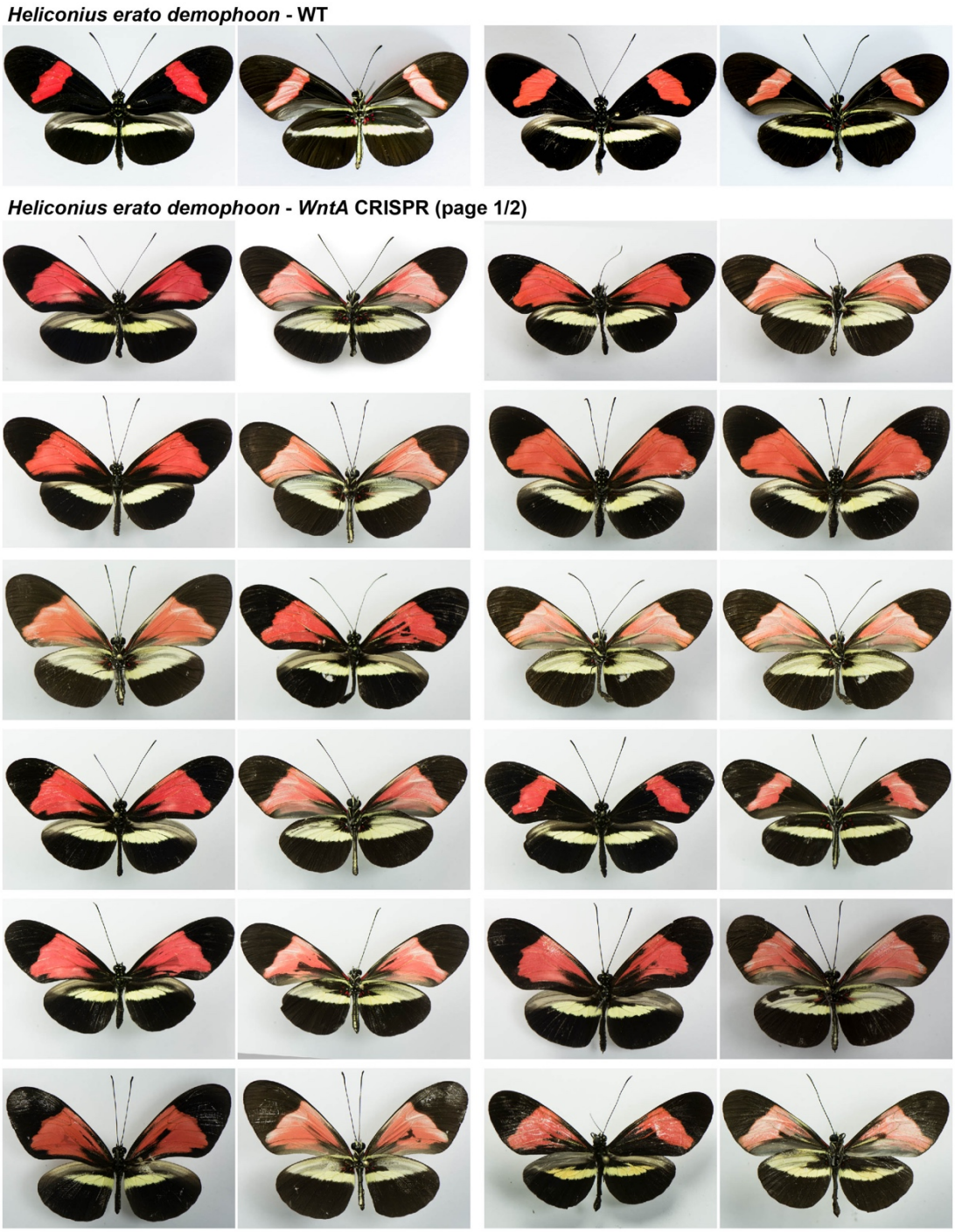


Figure S1-4. ***WntA*** **mKO yields pattern boundary aberrations in Heliconius erato demophoon**. Dorsal and ventral views of G0 butterflies are shown.

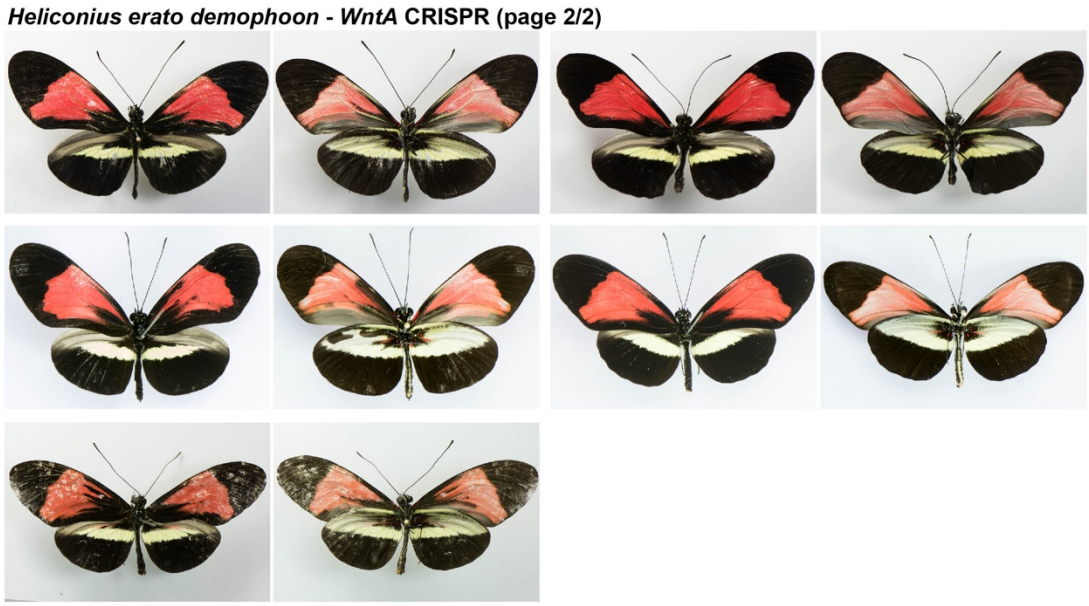


Figure S1-4. Continuation

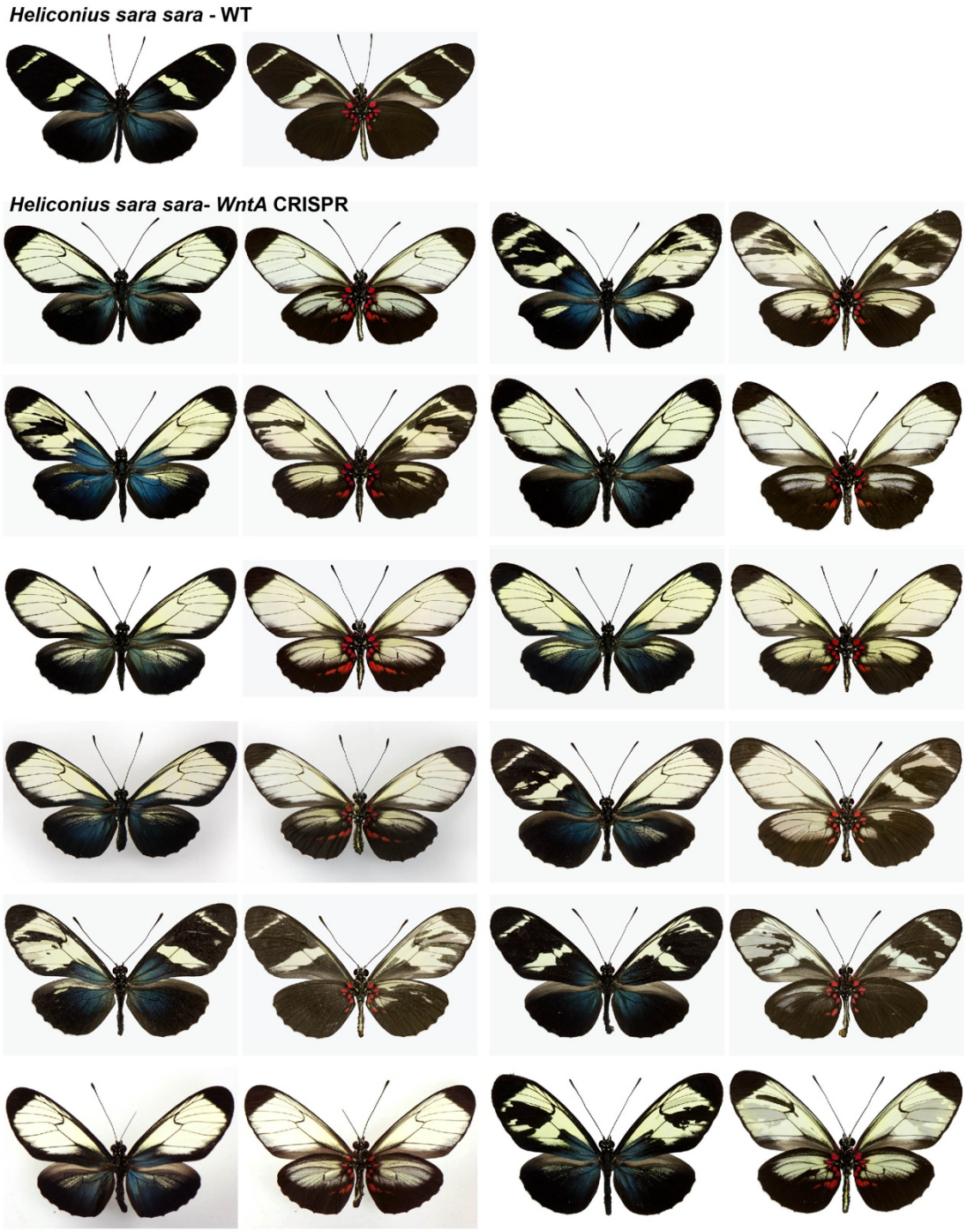


Figure S1-5. *WntA* mKO yields pattern boundary aberrations in *Heliconius sara*. Dorsal and ventral views of G0 butterflies are shown.

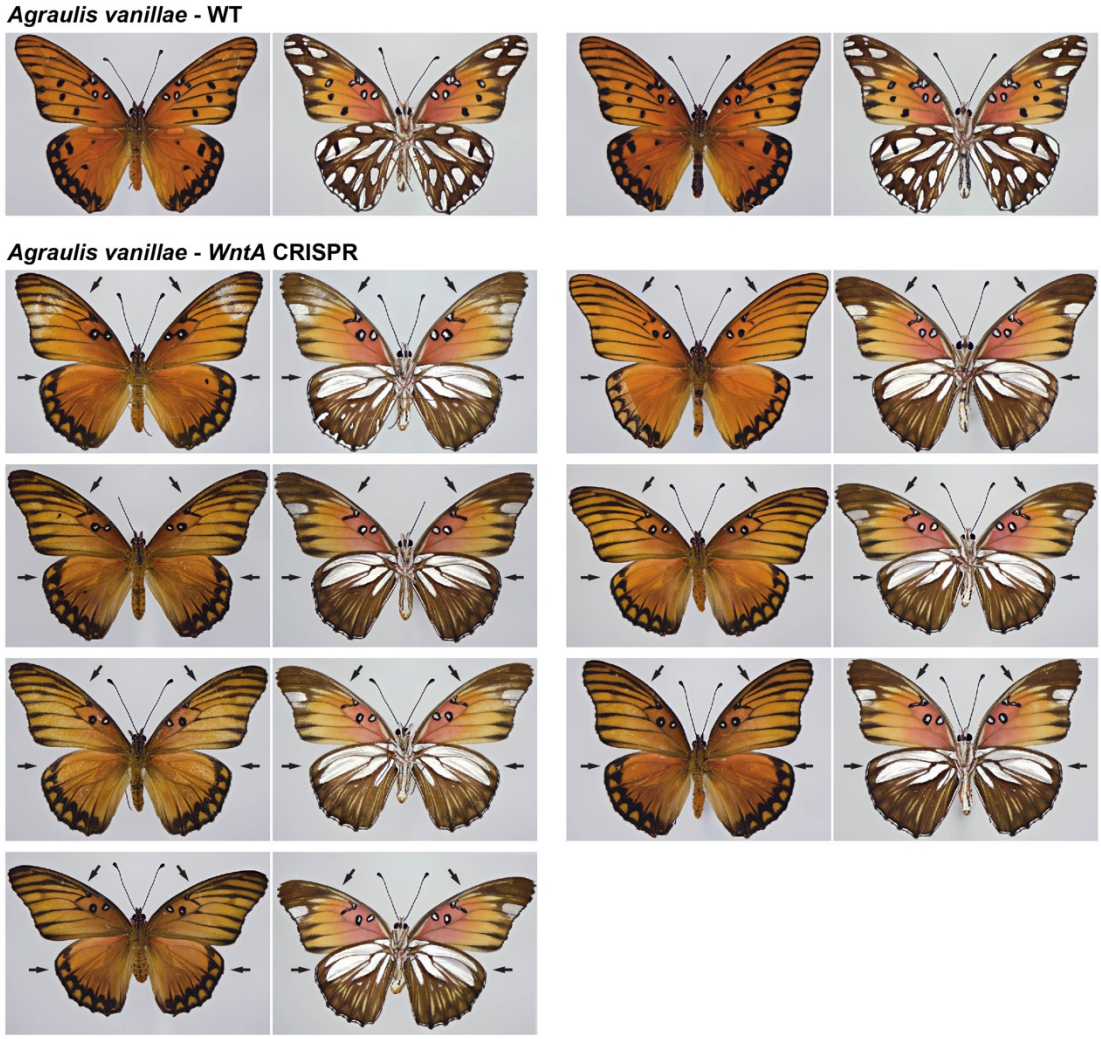


Figure S1-6. *WntA* mKO yields dual loss and expansion of silver spots in *Agraulis vanillae*. Dorsal and ventral views of G0 butterflies are shown, with arrows pointing at wing surfaces showing a phenotype.

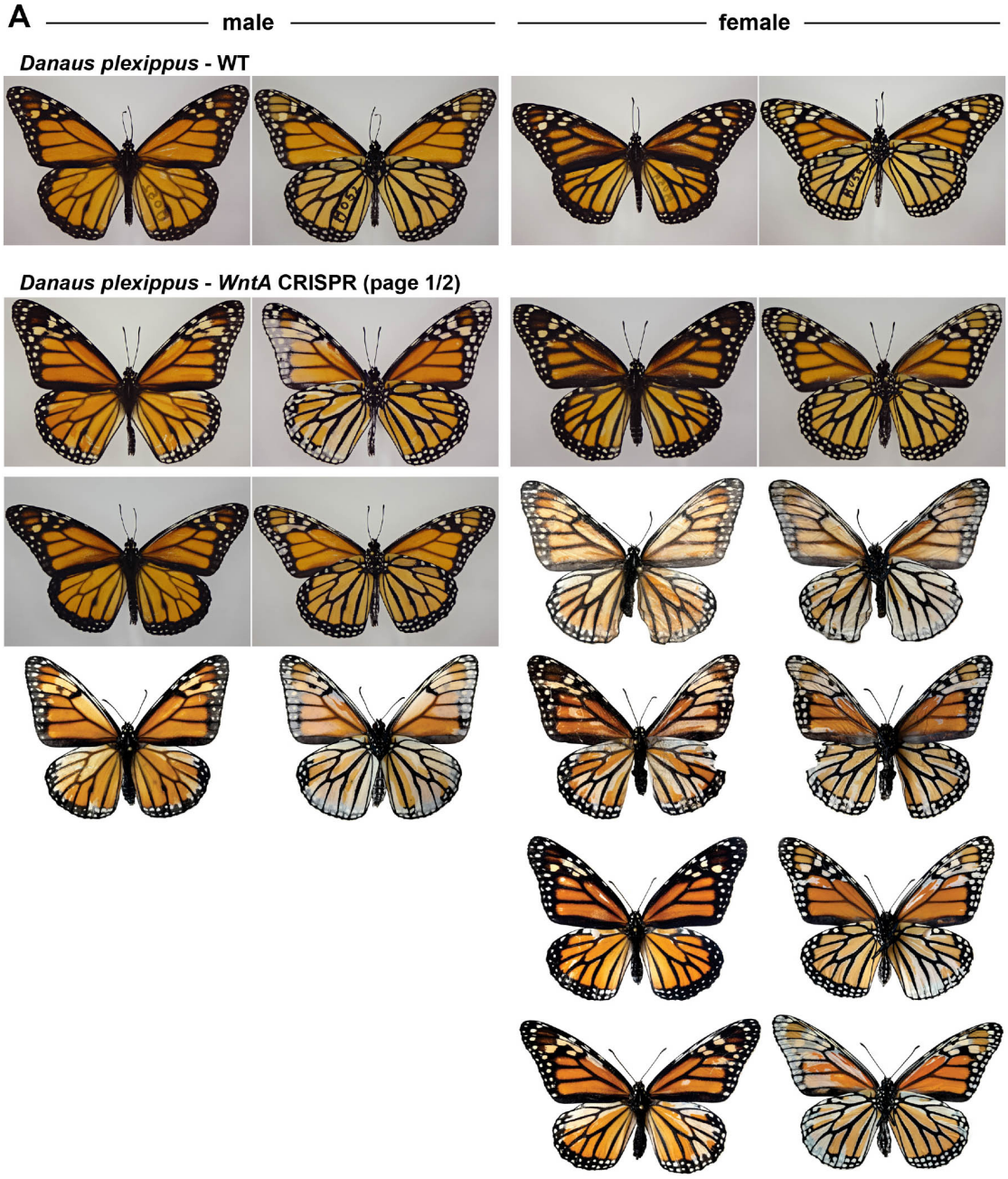


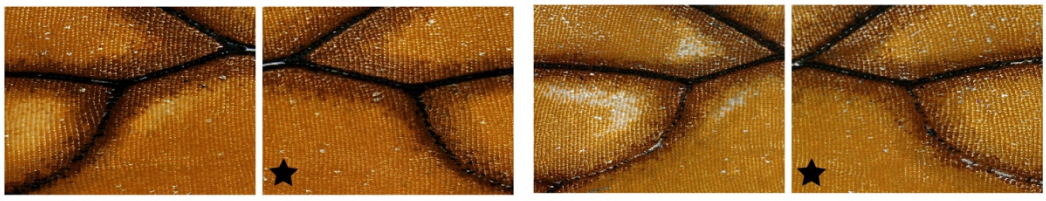
Figure S1-7. *WntA* mKO yields expansions of interveinous white in the monarch butterfly *Danaus plexippus*. (***A***) Dorsal and ventral views of 3 G_0 butterflies (3 males on the left column, 5 female on the right column) are compared to wild-type specimens from the injected stock (top row). (***B*, next page**) Close-up views of ectopic white scales on the left and right dorsal hindwings of 12 G_0 butterflies with a mild phenotype. Two of the mutant females as well as the two mutant males showed left/right asymmetries in their phenotypes, likely due to mosaicism within the individual.

B

***Danaus plexippus* - *WntA* CRISPR (page 2/2)**

Dorsal female hindwings (mild phenotypes)

Dorsal male hindwings (mild phenotypes)



★ : WT-like (no phenotype)

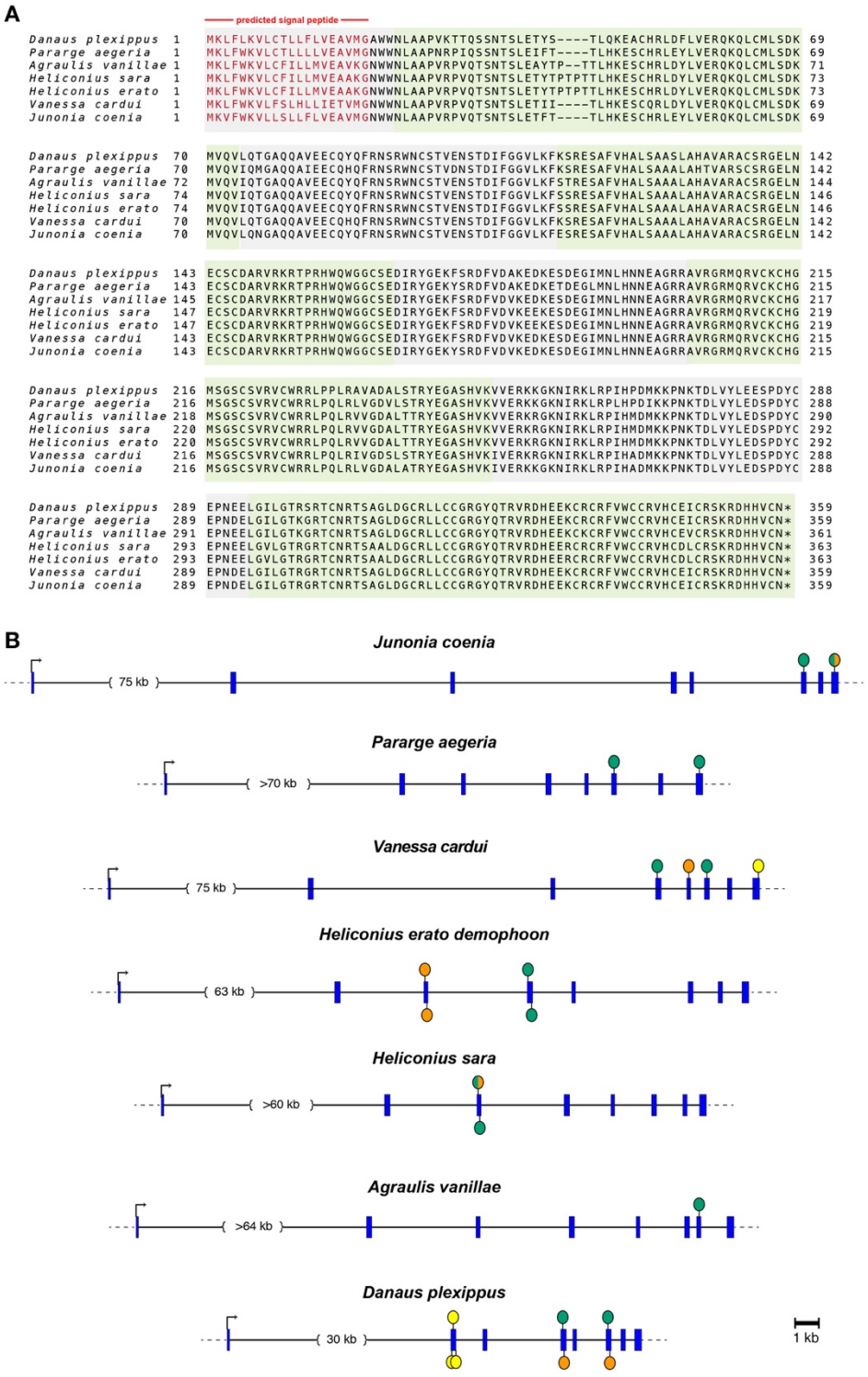


Figure S1-8. ***WntA protein alignment and sgRNA target regions.*** (**A**) Alignment of *WntA* amino-acid sequences (see Material and Methods, Gene annotation), grey and green boxes indicate corresponding coding exon boundaries. The first 20 amino-acids form a predicted signal peptide. (**B**) *WntA* gene structure for each species, exons are shown in blue boxes,

sgRNAs are notated by circles, sgRNAs with different colors indicate that they were used in different experiments (see **Tables S1 and S2**).

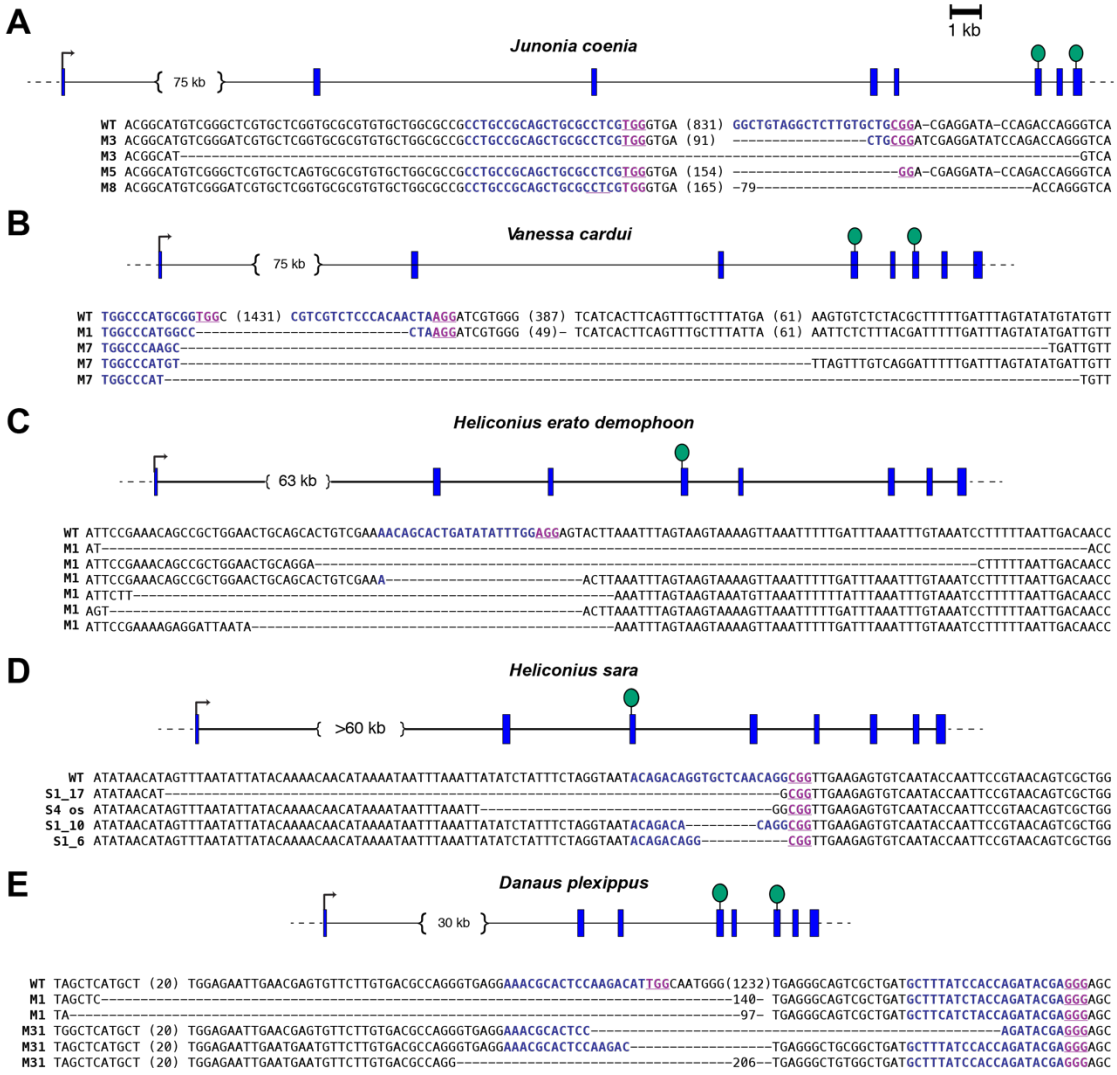


Figure S1-9. Genotyping of CRISPR/Cas9 *WntA* mutants. Position of dual sgRNA injections, and examples of resulting *WntA* coding sequence deletions in *J. coenia*, *V. cardui*, *H. e. demophoon, H. sara* and *D. plexippus*. *Blue letters*: sgRNA targets; *Fuchsia letters*: PAM sequence

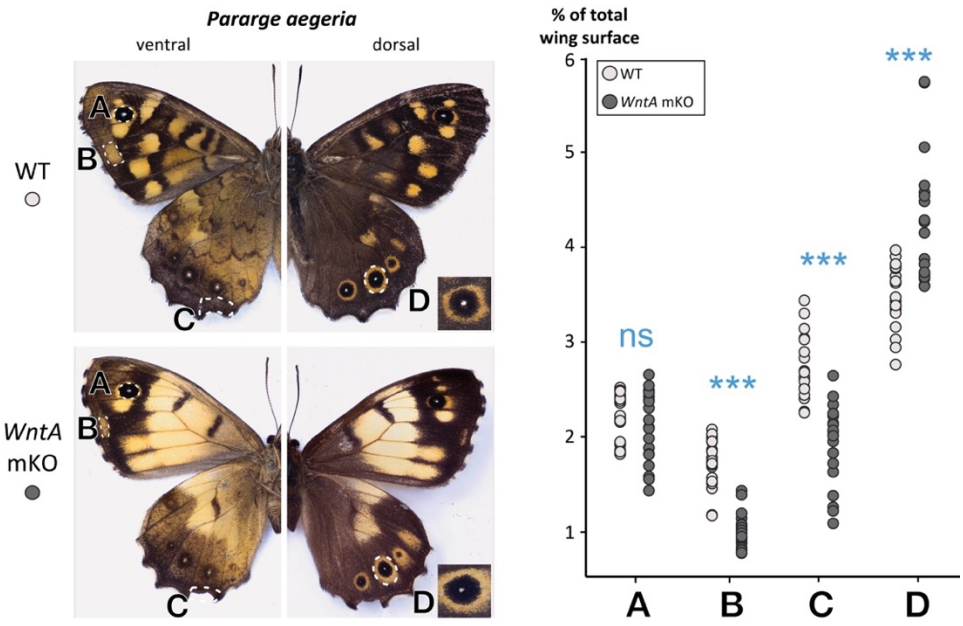


Figure S1-10. Effects of *WntA* mKO on marginal elements of *P. aegeria*. (***A***) No effect of *WntA* loss-of-function in forewing eyespot rings. (***B***) Distalization of the ventral forewing M2-M3 dPF element, resulting in a reduction of the light color- area ($p<0.001$). (***C***) Distalization of the ventral hindwing M1-M2 dPf element ($p<0.001$). (***D***) *WntA* loss-of-function results in expanded dorsal hindwing eyespot outer rings ($p<0.001$).

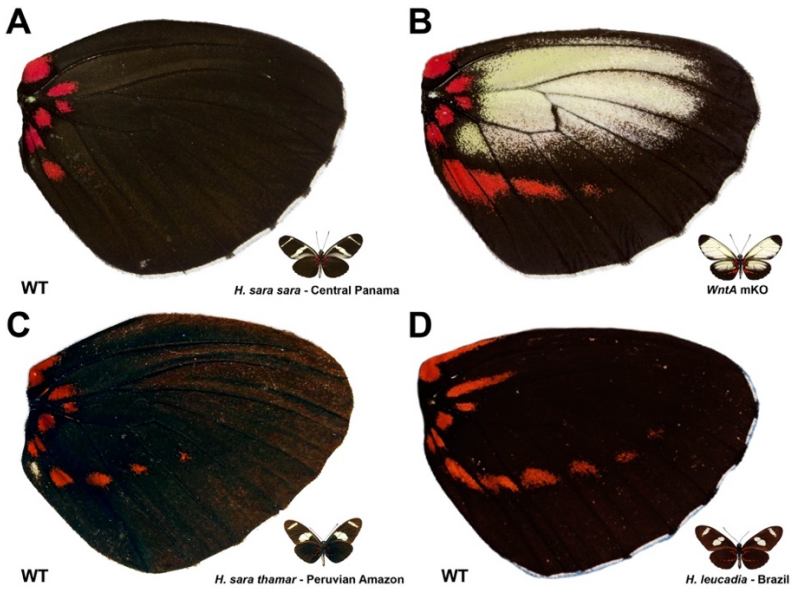


Figure S1-11. Local Wnt modulation may underlie hindwing red pattern variations in the *H. sara* group. (***A-B***) *WntA* mKO results in an expansion of red patterns in the ventral hindwings of *H. sara sara* (Central Panama population). (***C***) *H. sara* shows natural geographic variation in the ventral hindwing red pattern, with extended red stripes common in the

Amazon and in the Brazilian Atlantic forest. (***D***) Ventral hindwing red patterns in *Heliconius leucadia*, a sister species and co-mimic of *H. sara*.

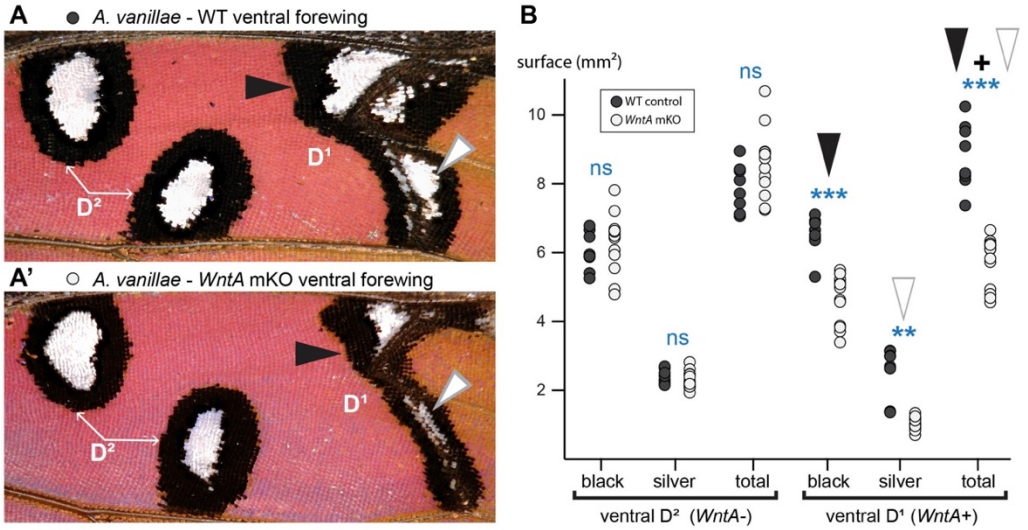


Figure S1-12. *WntA* functions in Discalis patterns of *A. vanillae* ventral forewings. (***A***) Wild type and mutants D1 and D2 patterns. These patterns express *wg*, and only D1 express *WntA* (*10*). (***B***) Both the central and outer color components of D1 patterns are reduced in *WntA* mKO wings.

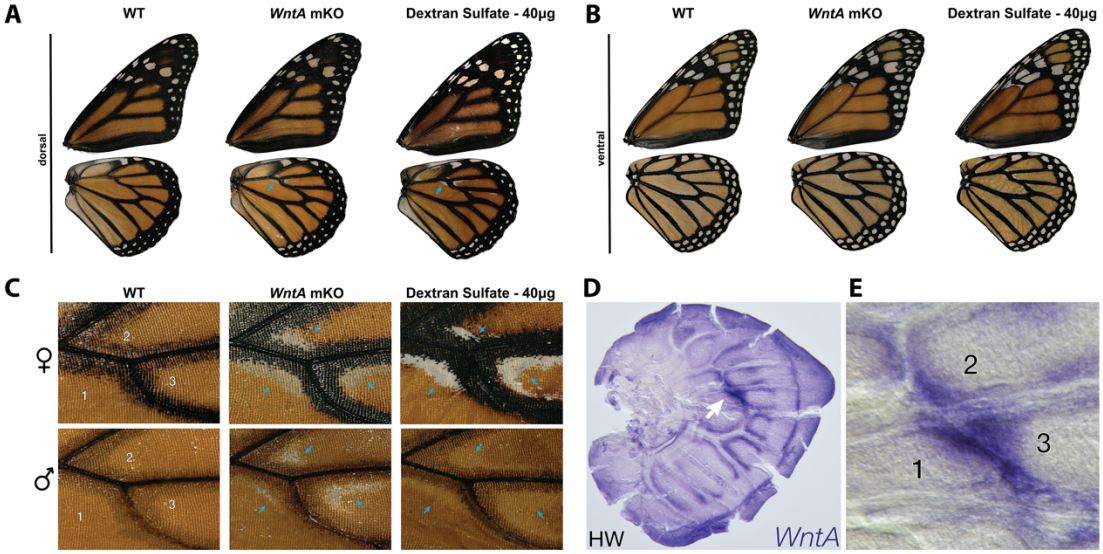


Figure S1-13. Genetic and pharmaceutical induction of ectopic white scales in monarch butterflies. (*A*) Mild phenotypes of *WntA* mKO in *D. plexippus* consist of a dorsal specific patch of ectopic white scales on the hindwings (blue arrows), phenocopied by injections of Dextran Sulfate shortly after pupation. (***B***) Ventral sides of the same butterflies. Dextran Sulfate injections have a visible effect on a forewing white patch, consistent with more pronounced *WntA* loss-of-function phenotypes (**Fig. 4**). (***C***) Close-up views of the dorsal

hindwing ectopic white scales in 3 periveinous domains (numbers). Both *WntA* mKO and Dextran Sulfate injections show a more pronounced effect in females (see also **Fig. S7**). (***D-E***) These phenotypic effects correlate with a domain of strong *WntA* expression in larval hindwing disks (arrow, magnified in ***E***).

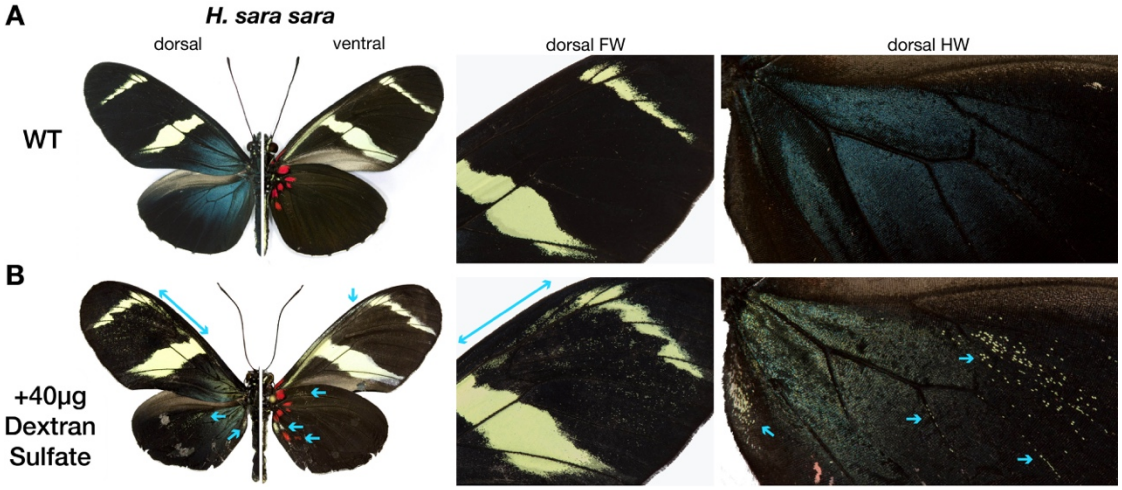


Figure S1-14. Dextran Sulfate injections are consistent with *WntA* loss-of- function effect in *H. sara.* Compared to wild-type (***A***), *H. sara* injected with Dextran Sulfate (***B***) show ectopic yellow and red scales in areas affected by *WntA* mKO (**Fig. 3**).

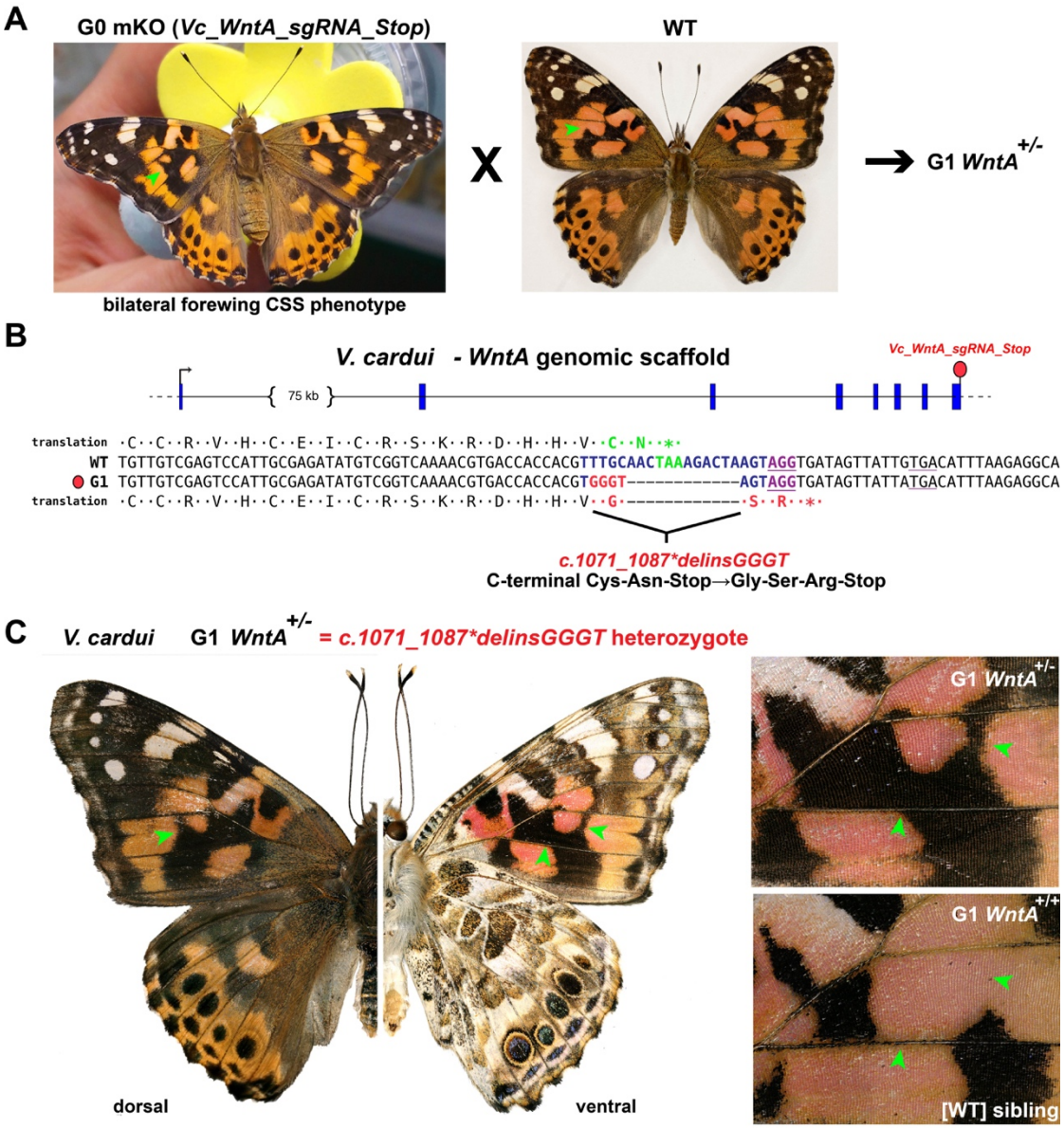


Figure S1-15. Germline transmission of a putative CRISPR-induced dominant negative allele in *V. cardui.* (***A***) CRISPR somatic mutagenesis targeting the Stop codon of the *WntA* ORF generated a G0 female showing a bilateral modification of the forewing CSS (green arrowhead: black stripe extension, visible on both dorsal and ventral sides). This individual was crossed to a wild- type male to generate G1 *WntA+/-* offspring. (***B***) Genotyping of G1 mutant offspring legs revealed a heterozygous state for a *c.1071_1087*delinsGGGT* allele, which results in a short C-terminal amino-acid change in the *WntA* protein. (***C***) *WntA+/-* G1 offspring show a local modification of the forewing CSS characterized by an extension of black elements (arrowheads; right panel: comparison of mutant and wild-type G1 siblings).

Table S1-1. sgRNA target sequences and genotyping primers

Species	sgRNA Name	Target sequence (5' > 3')	Genotyping primers (5' > 3')
Junonia coenia	Jc_WntA_sgRNA1	CCTGCCGCAGCTGCGCCTCG	TGTGCAAGTGCCACGGCATGT
	Jc_WntA_sgRNA2	GGCTGTAGGCTCTTGTGCTG	AAACGTGGTGGTCTCGCTTTGAC
Pararge aegeria	Pa_WntA_sgRNA1	GGCGTCGTTTACCTCAGCTG	–
	Pa_WntA_sgRNA2	GGGTGTCGATTGCTTTGCTG	–
Vanessa cardui	Vc_WntA_sgRNA1	GGCAGCATTGGCCCATGCGG	CCGTGAATCTGCATTCGTCCAT
	Vc_WntA_sgRNA2	GGCGTCGTCTCCCACAACTA	GCGTGTATTGGTCTCAATTTCCTG
	Vc_WntA_sgRNA	GGATTTCGTAGACGTTAAAG	–
	Vc_WntA_sgRNA_Stop	TTTGTAACTAAAGACTAAGT	TCGTGGAAGAACATGTAATAGA
			ACATTACCATCGCAGATCA
Heliconius erato demophoon	He_WntA-sgRNA1	ACAGACAGGTGCTCAACAGG	CCTCAGTCAGTCAGCACAAAGCC
	He_WntA-sgRNA2	AACAGCACTGATATATTTGG	CTCAAGAAGCCAAGTAGTAAGTG
	He_WntA-sgRNA3	CGCGGCACTAGCTCACGCGG	–
	He_WntA-sgRNA4	CCAAGACATTGGCAATGGGG	–
Heliconius sara	Hs_CRI_Wnta_2A	ACAGACAGGTGCTCAACAGG	–
	Hs_CRI_Wnta_2B	TCCAGCGACTGTTACGGAAT	–
Agraulis vanillae	Av_WntA_sgRNA	GGTTGTAGAAAGAAAACG	–
Danaus plexippus	Dp_WntA_E4_gRNA01	TTCTCCACGACTGCAAGCTC	CACTCGTTGACGGAGCTTCTC
	Dp_WntA_E4_gRNA02	CCAAGACATTGGCAATGGGG	TTTCACCTTGACATGACTCGC
	Dp_sgRNA_WntA_03	GGAAACGCACTCCAAGACAT	–
	Dp_WntA_E6_gRNA01	CAGCGTGTATGCAAATGTCA	–
	Dp_sgRNA_WntA_04	GCTTTATCCACCAGATACGA	–
	Dp_WntA_E6_gRNA02	GGCGTCGTCTACCGCCTCTG	–
	Dp_sgRNA_WntA_02	GCGCTTCCCTGGTGCACGTG	GAATGTAGCCCCATCTGCCA
	Dp_sgRNA_WntA_05	ACAGTCATCAAACACGTCCT	GCATTTAGTACCGTAATAGGCCA
	Dp_sgRNA_WntA_01	TTCCAGGAACCTTGCAGCTC	GGTTCAAACCTCGTACATTGCT
	Dp_sgRNA_WntA_08	TGCAGAAGGAAGCCTGCCAC	GCATTTAGTACCGTAATAGGCCA
	Dp_sgRNA_WntA_14	GCTATCCGACAAGATGGTTC	

Table S1-2. Summary of CRISPR/Cas9 injection experiments.

Species	sgRNAs	Final concentration ng/µl [sgRNA1(:sgRNAx):Cas9]	Time of egg collection (Hours)	Number of eggs injected	Number of L1 larvae	Hatching ratio	Number of pupae	Number of adults with phenotype	Wing mutant phenotype ratio from hatched eggs
Junonia coenia	Jc_WntA_sgRNA1, Jc_WntA_sgRNA2	150(:150):100	4	510	143	28.00%	68	44	30.80%
Junonia coenia	Jc_WntA_sgRNA	150:333	2-4.5	400	128	32.00%	23	6	4.70%
Pararge aegeria	Pa_WntA_sgRNA1, Pa_WntA_sgRNA2	200(:200):333	1-2	830	207	24.90%	NA	32	15.50%
Vanessa cardui	Vc_WntA_sgRNA1, Vc_WntA_sgRNA2	75(:75):100	2-4	425	106	24.90%	19	12	11.30%
Vanessa cardui	Vc_WntA_sgRNA	200:400	3.5-6	295	52	17.60%	34	10	19.20%
Vanessa cardui	Vc_WntA_sgRNA	150:300	3-5	241	78	32.40%	30	13	16.70%
Vanessa cardui	Vc_WntA_sgRNA	150:300	1-4	1262	538	42.60%	134	25	4.60%
Vanessa cardui	Vc_WntA_sgRNA	150:300	1-5.5	775	497	64.10%	112	7	1.40%
Heliconius erato	He_WntA-sgRNA1, He_WntA-sgRNA2	200(:200):300	1-5	263	119	45.20%	36	5	4.20%
Heliconius erato	He_WntA-sgRNA1, He_WntA-sgRNA2	200(:200):250	1-2	188	61	32.40%	30	5	8.20%
Heliconius erato	He_WntA-sgRNA1	200:500	4-5	55	32	58.20%	15	3	9.40%
Heliconius erato	He_WntA-sgRNA2	200:500	3-4	42	16	38.10%	4	2	12.50%
Heliconius erato	He_WntA-sgRNA3, He_WntA-sgRNA4	200(:200):500	2-4	107	34	31.80%	3	1	2.90%
Heliconius erato	He_WntA-sgRNA4	200:500	2-4	73	42	57.50%	2	1	2.40%
Heliconius sara	Hs_CRI_Wnta_2A, Hs_CRI_Wnta_2B	250(:250):500	1-4	100-150	28	NA	26	13	46.40%
Heiconius sara	Hs_CRI_Wnta_2A	500:500	1-4	150-200	45	NA	42	33	73.30%
Agraulis vanillae	Av_WntA_sgRNA	150:333	2-5	37	19	51.40%	12	8	42.10%
Danaus plexippus	Dp_WntA_E4_gRNA01, Dp_WntA_E6_gRNA01	60(:60):200	1-2	48	18	37.50%	3	2	5.60%
Danaus plexippus	Dp_WntA_E4_gRNA02, Dp_WntA_E6_gRNA02	60(:60):200	1-2	51	22	43.10%	6	4	18.20%
Danaus plexippus	Dp_WntA_E4_gRNA01, Dp_WntA_E6_gRNA01, Dp_WntA_E4_gRNA02, Dp_WntA_E6_gRNA02	40(:40:40:40):200	1-2	20	6	30.00%	4	2	16.70%
Danaus plexippus	Dp_WntA_E4_gRNA01, Dp_WntA_E6_gRNA01, Dp_sgRNA_WntA_03, Dp_sgRNA_WntA_04	150(:150:150:150):300	3-4	18	7	38.90%	4	2	14.30%
Danaus plexippus	Dp_sgRNA_WntA_02, Dp_sgRNA_WntA_05	200(:200):250	3-3.5	59	40	67.80%	9	2	5.00%
Danaus plexippus	Dp_sgRNA_WntA_01, Dp_sgRNA_WntA_08, Dp_sgRNA_WntA_14	83.3(:83.3:83.3):250	2.5-3.5	97	55	56.70%	31	8	14.55%
			TOTAL =	5796	2293	39.60%	647	240	10.5%

Table S1-3. Butterfly rearing conditions

Species	Stock origin	Lab colony	Light/Dark cycle (hr)	Temperature (°C)	Relative Humidity (%)	Oviposition plants and larval diet
Junonia coenia	North Carolina (gift of Laura Grunert and Fred Nijhout)	Patel Lab at UC Berkeley Reed Lab at Cornell University	14/10 16/8	26 27	60	*Plantago lanceolata* (ribwort plantain) Larval diet : artificial diet prepared with water boiled with 2g/L dry *P. lanceolata*
Pararge aegeria	Belgium (gift of Camille Tulure and Hans Van Dyck)	Breuker Lab at Oxford Brookes University	16/8	23	60	*Poa trivialis (rough bluegrass)* *Brachypodium sylvaticum (false-brome)* *Dactylis glomerata* (orchard grass)
Vanessa cardui	US Commercial Provider (Carolina Biological Supplies)	Patel Lab at UC Berkeley Martin Lab at GW University Reed Lab at Cornell University	14/10 14/10 16/8	26 27	60	*Malva spp.* (mallow plant) *Alcea rosea (*common hollyhocks) *Plantago major* (broadleaf plantain) Artificial diet (Southland Products Inc)
Agraulis vanillae	California	Patel Lab at UC Berkeley	14/10	26	60	*Passiflora incarnata x P. edulis* ("Byron Beauty" passionflower)
Heliconius erato demophoon	Panama (canal zone)	*Heliconius* stock center at the Smithsonian Tropical Research Institute (STRI) in Gamboa, Panama - outdoors	12/12	22-28	60-80	*Passiflora biflora* (two-flowered passion flower)
Heliconius sara	Panama (canal zone)	Cambridge University - indoors Heliconius stock center at STRI- outdoors	12/12	27 22-28	80 60-80	*Passiflora biflora* (two-flowered passion flower)
Dananus plexippus	Florida Costa Rica	Reed Lab at Cornell University Kronforst Lab at the University of Chicago	16/8	26	60	*Asclepia curassavica* (milkweed)

Table S1-4. Primers sequences used to generate templates for riboprobe synthesis.

Species	Primer sequence (5'>3')
Junonia coenia, Vanessa cardui, Agraulis vanillae	CCV MGA CAY TGG CAR TGG GG
	RCA VCG RCA CTT CWC YTC GT
Pararge aegeria	AGA CTT TGT AGA CGC TAA AG
	CAA TTA CAA ACG TGG TGA TC
with T7 promoter	taatacgactcactatagggAGACTTTGTAGACGCTAAAG
	taatacgactcactatagggCAATTACAAACGTGGTGATC
*Heliconius erato**	GGC GGT TGA AGA GTG TCA AT
	GTT GCA GAC GTG GTG GTC
Danaus plexippus	CCA AGA CAT TGG CAA TGG GG
	GCA ACG GCA CTT CTC CTC GT

*For *Heliconius sara*, the *H. erato* probe was used.

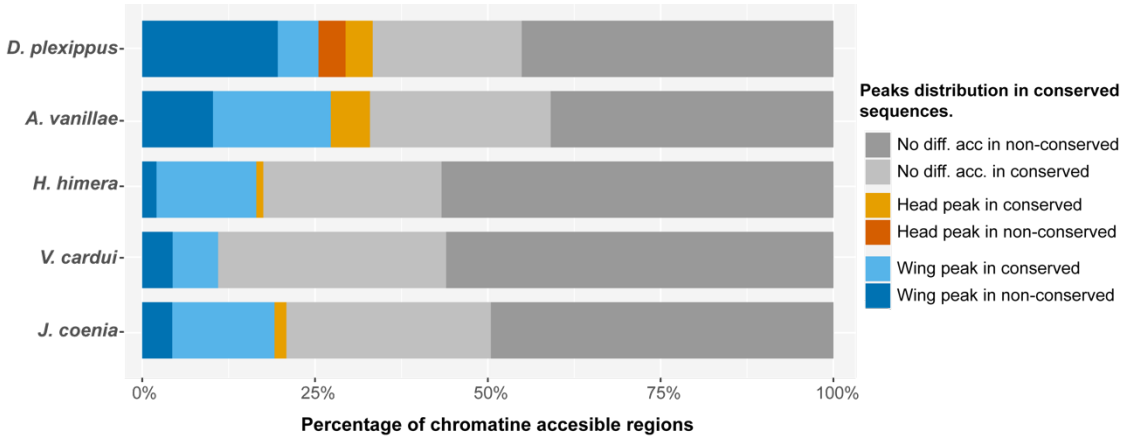


Figure S2-1. Distribution of accessible chromatin regions over conserved sequences. Proportion of peak calls from ATAC-seq profiles overlapping with the 'most conserved' DNA regions as calculated by PhastCons. Most of the peaks are not differentially accessible, the ones that are, arise from wing tissue.

Table S2-1. List of species and genome sequences source used in this study

Species	Family - Subfamily	Genome source	Genome file
Papilio xuthus	Papilionidae - Papilioninae	lepbase.org	Papilio_xuthus_Pxut_1.0_Refseq_-_scaffolds.fa
Danaeus plexippus	Nymphalidae - Danainae	NCBI	GCF_009731565.1_Dplex_v4_genomic.fna
Danaeus chrysippus	Nymphalidae - Danainae	NCBI	GCA_004959915.1_KIT_Dchrysippus_v1.3_genomic.fna
Bicyclus anyana	Nymphalidae - Satyrinae	lepbase.org	Bicyclus_anynana_v1.2_-_scaffolds.fa
Aphantopus hyperantus	Nymphalidae - Satyrinae	NCBI	GCA_902806685.1_iAphHyp1.1_genomic.fna
Agraulis vanillae	Nymphalidae - Heliconiinae	Reed Lab - Cornell	Av.v0.fa
Heliconius erato demophoon	Nymphalidae - Heliconiinae	lepbase.org	Heliconius_erato_demophoon_v1_-_scaffolds.fa
Helicinius melpomene	Nymphalidae - Heliconiinae	lepbase.org	Heliconius_melpomene_melpomene_Hmel2.5.scaffolds.fa
Hypolimnas misippus	Nymphalidae - Nymphalinae	NCBI	GCA_008963455.1_UofC_Hmis_v1.0_genomic.fna
Junonia coenia	Nymphalidae - Nymphalinae	Reed Lab - Cornell	jcgen_v2.fa
Vanessa cardui	Nymphalidae - Nymphalinae	Reed Lab - Cornell	Vcar.v1.scaf.fa
Vanessa tameamea	Nymphalidae - Nymphalinae	NCBI	GCF_002938995.1_ASM293899v1_genomic.fna

Table S2-2. ATAC-seq samples with the alignment and quality control statistics.

Species	Sample name	Sequencing depth	% mapped	NRF	PBC1	PBC2	FRip
D. plexippus	ATAC001_DpM5thFw1	22,576,990	88%	0.8	0.9	11	0.70
	ATAC023_DpM5thFw2	16,738,827	81%	0.8	0.9	17	0.70
	ATAC025_DpM5thFw3	11,962,082	83%	0.7	0.8	7	0.89
	ATAC002_DpM5thHw1	19,418,685	87%	0.8	0.9	9	0.72
	ATAC024_DpM5thHw2	14,673,669	86%	0.7	0.9	8	0.84
	ATAC026_DpM5thHw3	12,508,804	87%	0.7	0.9	8	0.88
	ATAC027_DpM5thHD1	13,479,773	71%	0.8	0.9	18	0.72
	ATAC028_DpM5thHD2	27,142,322	77%	0.7	0.9	9	0.75
	ATAC029_DpM5thHD3	19,927,620	76%	0.8	0.9	11	0.71
J. coenia	ATAC006_JcM5thFw1	13,878,937	81%	0.9	0.9	18	0.78
	ATAC016_JcM5thFw2	16,126,659	88%	0.9	0.9	16	0.79
	ATAC019_JcM5thFw3	52,718,382	86%	0.9	0.9	18	0.61
	ATAC017_JcM5thHw2	18,205,610	75%	0.9	1.0	27	0.60
	ATAC020_JcM5thHw3	30,218,065	77%	0.9	0.9	22	0.60
	ATAC046_JcM5thHw4	51,718,806	87%	0.7	0.9	8	0.79
	ATAC018_JcM5thHD1	23,298,731	79%	0.9	1.0	29	0.48
	ATAC021_JcM5thHD2	20,697,385	81%	0.8	0.9	10	0.63
	ATAC047_JcM5thHD4	32,603,886	82%	0.9	0.9	18	0.55
V. cardui	ATAC004_VcM5thFw1	9,348,925	92%	0.8	0.9	24	0.66
	ATAC010_VcM5thFw2	56,685,433	94%	0.7	0.9	22	0.44
	ATAC030_VcM5thFw3	18,345,041	88%	0.8	0.9	15	0.70
	ATAC005_VcM5thHw1	22,214,567	93%	0.7	0.9	17	0.67
	ATAC011_VcM5thHw2	42,159,245	93%	0.7	0.9	17	0.50
	ATAC031_VcM5thHw3	31,986,703	89%	0.6	0.8	6	0.74
	ATAC012_VcM5thHD1	46,265,224	88%	0.7	0.9	22	0.44
	ATAC032_VcM5thHD2	15,041,762	80%	0.7	0.9	8	0.75
	ATAC033_VcM5thHD3	16,400,602	82%	0.8	0.9	11	0.72
H. himera	ATAC013_HhimM5thFw1	28,390,780	75%	0.8	0.9	17	0.61
	ATAC043_HhimM5thFw2	58,720,883	81%	0.7	0.9	10	0.69
	ATAC048_HhimM5thFw3	32,965,627	78%	0.7	0.9	12	0.67
	ATAC014_HhimM5thHw1	45,342,569	72%	0.8	0.9	19	0.49
	ATAC044_HhimM5thHw2	39,683,955	76%	0.7	0.9	12	0.68
	ATAC049_HhimM5thHw3	32,898,390	79%	0.7	0.9	11	0.75
	ATAC015_HhimM5thHD1	25,895,251	78%	0.8	0.9	23	0.60
	ATAC045_HhimM5thHD2	35,444,279	80%	0.8	0.9	13	0.63
	ATAC050_HhimM5thHD3	31,950,874	79%	0.8	0.9	13	0.68
A. vanillae	ATAC034_AvM5thFw1	17,295,707	92%	0.7	0.9	10	0.82
	ATAC037_AvM5thFw2	20,685,695	92%	0.6	0.8	5	0.83
	ATAC040_AvM5thFw3	17,341,680	89%	0.8	0.9	10	0.79
	ATAC035_AvM5thHw1	24,273,551	92%	0.6	0.8	6	0.81
	ATAC038_AvM5thHw2	15,974,424	91%	0.6	0.8	6	0.86
	ATAC041_AvM5thHw3	15,760,591	89%	0.7	0.9	8	0.80
	ATAC036_AvM5thHD1	26,929,395	90%	0.6	0.8	6	0.75
	ATAC039_AvM5thHD2	16,887,692	91%	0.7	0.9	8	0.84
	ATAC042_AvM5thHD3	18,830,151	87%	0.7	0.8	7	0.77

Table S2-3. List of guide-RNAs used in CRISPR-Cas9 knock-out experiments. wCRE= *WntA* Cis-Regulatory Element. Vc: *Vanessa cardui*, Jc: *Junonia coenia*, Av: Agraulis vanillae, Dp: *Danaus plexippus*, Hhim: *Heliconius himera*.

Element	guide RNA ID	sgRNA sequence
wCREVc01	wCREVc01_01	CGAGGTAAGTGAAACGCACA
wCREVc01	wCREVc01_02	CGTCGACACGTAAAAAGGAA
wCREVc01	wCREVc01_03	TTACATACCGACGATTTCTG
wCREVc01	wCREVc01_04	GATTAATATACTCTAGTGGA
wCREVc02	wCREVc02_01	GCTTTCTCGATTCTGTATAA
wCREVc02	wCREVc02_02	ATTCGTAACGCTATCGAAAC
wCREVc02	wCREVc02_03	GTACGAAGAGGCGGTACGGT
wCREVc02	wCREVc02_04	GCCGCTCTGGTCAAGCGGAT
wCREVc03	wCREVc03_01	CGCCCGCGGATTCATGTTTC
wCREVc03	wCREVc03_02	ACGCCAATAATACGTCGTGG
wCREVc03	wCREVc03_03	AATAACATCGCAAACAGCCT
wCREVc04	wCREVc04_01	GGACGCGTGAACGGGGGTTC
wCREVc04	wCREVc04_02	GGCGGTGGGCATTGTACTTA
wCREVc04	wCREVc04_03	AACACATCTCAATAATCCGA
wCREVc04	wCREVc04_04	ACATGTAACCTCATTCACTT
wCREVc05	wCREVc05_01	ACATTTGCCATCCTTATACG
wCREVc05	wCREVc05_02	AATAAACTCTGTACCTCTGT
wCREVc05	wCREVc05_03	TTACGAATGTGTGAGTGTCG
wCREVc06	wCREVc06_01	ACTGAGATAAAAGACGTCTC
wCREVc06	wCREVc06_02	CTAACTAATGAAATCTACAT
wCREVc06	wCREVc06_03	CATTGACTTGACAGCCTCTA
wCREVc06	wCREVc06_04	AGAGGCCTTGTGTTTAAAGT
wCREVc07	wCREVc07_01	AATCGCCGCTTTGACCACTA
wCREVc07	wCREVc07_02	ATTGGAAGGGGAGTCAGTTT
wCREVc07	wCREVc07_03	TGTGTACCAAGAAAGTGTTA
wCREVc08	wCREVc08_01	GtGACCGTTTCATATGTTTCA
wCREVc08	wCREVc08_02	tgaaaaaggaaCGCGGTCAG
wCREVc08	wCREVc08_03	TTGGGATAAGGCTTCGTAAC
wCREVc08	wCREVc08_04	CGGGCGATGACGTTCGCGCC
wCREVc09	wCREVc09_01	tgaTTCAATTTGAAACAACG
wCREVc09	wCREVc09_02	ATTGTTTTGGCGGGACGCGA
wCREVc09	wCREVc09_03	TGTTTACAAAAGATCTATTG
wCREVc09	wCREVc09_04	TATGTGACTGTTCATTCATT
wCREVc10	wCREVc10_01	GAATGCTGCGAGTATAGATC
wCREVc10	wCREVc10_02	GGCTAGCTGATGCTTATGGT
wCREVc10	wCREVc10_03	CGACGCGCCTGCGATCCGCG
wCREVc10	wCREVc10_04	acaTaaacacaGAACTGAAC
wCREVc11	wCREVc11_01	AAAGTTCGAAATTGATGGTC
wCREVc11	wCREVc11_02	AGCAACCCTGTATCTGGCTA
wCREVc11	wCREVc11_03	ATATTGAAAGCGAATTATTA
wCREJc01	wCREJc01_01	TATGCAAACGTATGTAGTCG
wCREJc01	wCREJc01_02	GGAGTAATAGGTCCCATTAG
wCREJc01	wCREJc01_03	CAGATTACTGACATGACTTT
wCREJc02	wCREJc02_01	TAGGGACAGCTAAAAACGAG
wCREJc02	wCREJc02_02	ACTCGTTTTATTGTTTTGAG
wCREJc02	wCREJc02_03	TACATAAGTCCTTACCAGAC
wCREJc03	wCREJc03_01	ACTATCAACGGTTTTAAGTT
wCREJc03	wCREJc03_02	TAACAATGTTTCGCTTTACG
wCREJc03	wCREJc03_03	CTCTGTATTTACAGCTACAG
wCREJc04	wCREJc04_01	AAATCACGATGTCAGAATAT
wCREJc04	wCREJc04_02	TGCGCGAAACGTCAGCCGCG
wCREJc04	wCREJc04_03	ATGTTGTGTGAAGCGGCTAG
wCREJc05	wCREJc05_01	TTTTATTACAATGCATAGTA
wCREJc05	wCREJc05_02	ACACTTAATAAGGATCTCAT
wCREJc05	wCREJc05_03	GTAATTACGGTGCACTTATT

Element	guide RNA ID	sgRNA sequence
wCREAv01	wCREAv01_03	AATATACAGTCTTGCCTTTT
wCREAv02	wCREAv02_01	TATGGCCAACTCTAAGTGAC
wCREAv02	wCREAv02_02	CCGGCATGATCGGTACGCTA
wCREAv02	wCREAv02_03	CCTTTAATCTTGGTGACATA
wCREAv03	wCREAv03_01	GCACACTTTGGAAATGGCCG
wCREAv03	wCREAv03_02	CACACATCGCCTGCATATTG
wCREAv04	wCREAv04_01	TGTTGTCAACACTAATTCAA
wCREAv04	wCREAv04_02	ATTGCGACGTCACTTGTCAA
wCREAv04	wCREAv04_03	TTGAAAACATTGCTAAGCTT
wCREAv05	wCREAv05_01	TCTGCATAAAGTGCGTGCGT
wCREAv05	wCREAv05_02	TCCTCTACTTTTCGCTTAAA
wCREAv05	wCREAv05_03	TGCGCGAAACGTCAGTCACA
wCREAv05	wCREAv05_04	GTGTTGTATGGCTTGAACGG
wCREAv06	wCREAv06_01	ATATGTGATACCTTCGAAAT
wCREAv06	wCREAv06_02	GGGTGGACCACCTCTGACTT
wCREAv06	wCREAv06_03	CATTAATTTTAGATCACGGC
wCREAv06	wCREAv06_04	CTGGATAGTGAGATGTCACA
wCREAv07	wCREAv07_01	ACAAGCATAAAGGTACGTAT
wCREAv07	wCREAv07_02	AATATTGTATAGAGTGGGAT
wCREAv07	wCREAv07_03	CCAAATATGTCCCAGAAGAT
wCREAv07	wCREAv07_04	CGAACTATATTAAAAATTCG
wCREAv08	wCREAv08_01	TAATGTAATCGAAATCCAGA
wCREAv08	wCREAv08_02	ACAGTGATGACGTTCGCGCC
wCREAv08	wCREAv08_03	tatataATCGGATCGGATTT
wCREAv09	wCREAv09_01	ATTTTGCCTCAGCATTTCAC
wCREAv09	wCREAv09_02	AGTTTAAGCCAAAGCACTTC
wCREAv10	wCREAv10_01	ATTACTTTAGTAACAGCTAT
wCREAv10	wCREAv10_02	AGGCTTCTTAGTAAGGGTTA
wCREAv10	wCREAv10_03	TTTCCGACCGCGCGGATCGC
wCREDp01	wCREDp01_01	GAGCTAACGATATACGATTG
wCREDp01	wCREDp01_02	TCTGACATTTCCGTTGTTGA
wCREDp02	wCREDp02_01	CCGGGTGAGAAGAATGCTAA
wCREDp02	wCREDp02_02	GCTTTAAAATTCTCTTATAG
wCREDp03	wCREDp03_01	ATAAACGTAAATCTACCAGA
wCREDp03	wCREDp03_02	CAACAATAGCATTTGTGGAA
wCREDp04	wCREDp04_01	AAGCGGTAACACTCTTATAG
wCREDp04	wCREDp04_02	CCGCCATCACGGCGGCAGCA
wCREDp04	wCREDp04_03	GAGACGGATCCCTACTCAAA
wCREDp04	wCREDp04_04	TGTTCAAACAAGTCTCCCAT
wCREDp05	wCREDp05_01	AATACTTCACCTTGGTCACG
wCREDp05	wCREDp05_02	ACGGAACAGGATGTGCTGTT
wCREDp05	wCREDp05_03	CGCATATTAAATACGTGTGT
wCREDp05	wCREDp05_04	TAAGTATCTCGAGCTGAACA
wCREDp06	wCREDp06_01	GTCATCGGTTTAAGAATCAA
wCREDp06	wCREDp06_02	CCCCGGGACATGTATGGAGC
wCREDp06	wCREDp06_03	CTTTTACAAGGTGCTTATAC
wCREDp07	wCREDp07_01	TCATTGGACCACTTAAGTGT
wCREDp07	wCREDp07_02	CCATAGGATTCGCTAAGGGA
wCREDp07	wCREDp07_03	CGCAGGCGCGTTACCACTCT
wCREDp07	wCREDp07_04	GAAAACTTTCGTGAAGTTGC
wCREHhim1	wCREHhim1_01	GACCAGGCGCTCTGATCTCC
wCREHhim1	wCREHhim1_02	TTTATGACGGCCTAATATCA
wCREHhim1	wCREHhim1_03	GCATACGAGTCTTCTTAATA
wCREHhim1	wCREHhim1_04	TAGACGTTACGTTACAGTGG
wCREHhim2	wCREHhim2_01	CAATGCCTATGCTTATACAT

Table 2-3 continuation.

Element	guide RNA ID	sgRNA sequence
wCREJc05	wCREJc05_04	TAGATCAGACTTACCTGCAG
wCREJc06	wCREJc06_01	CTACTCAGTTCGTTTTAGAT
wCREJc06	wCREJc06_02	GATATCTTATGAGAGTTATT
wCREJc06	wCREJc06_03	TGACTAGCAGAAAAGCTGAA
wCREJc06	wCREJc06_04	GAGCATTACTTTGCATGTCG
wCREJc07	wCREJc07_01	ACTAATTGAAATTAGAACAT
wCREJc07	wCREJc07_02	CAATCGCTCTATAACCATAG
wCREJc07	wCREJc07_03	ATTTGTCAGTTGAAAGCGAT
wCREJc07	wCREJc07_04	TTAAACCTACGAGAGCTAAA
wCREJc08	wCREJc08_01	CGAATAAAAACATTGCGTTC
wCREJc08	wCREJc08_02	AAATTGATCACACGGTGGCC
wCREJc08	wCREJc08_03	AACTTTTGAGAAAAGGAACG
wCREJc09	wCREJc09_01	TAAAACGATCAACATTTTAG
wCREJc09	wCREJc09_02	AGACGGAATTTATCTACGGC
wCREJc09	wCREJc09_03	TCTGTTGACAACACCTGCCG
wCREJc09	wCREJc09_04	AGATGTCTGGTGAACTTGTT
wCREJc10	wCREJc10_01	TTTGCCCGTCGTACATATTG
wCREJc10	wCREJc10_02	GTTGATCGAAAGAAGATTAA
wCREJc10	wCREJc10_03	GCTGTTGCTTATTCTTGGTG
wCREJc10	wCREJc10_04	GTAATAAACGCGTATCGAAT
wCREAv01	wCREAv01_01	TTAATTTACGTAGGGCTTAA
wCREAv01	wCREAv01_02	GTACGAAGAGGCGGGACGGT

Element	guide RNA ID	sgRNA sequence
wCREHhim2	wCREHhim2_02	GGTCCGCCACGGACCGTAGC
wCREHhim2	wCREHhim2_03	TAATATTCAGCTTAGGTTTT
wCREHhim3	wCREHhim3_01	CGTGTTAAGTGCTCAACGCC
wCREHhim3	wCREHhim3_02	GCAAGTACCAAATACGATGA
wCREHhim3	wCREHhim3_03	GAAGGAATGTGTTACGCATG
wCREHhim4	wCREHhim4_01	TTAAAAAAGATCTATTGTGG
wCREHhim4	wCREHhim4_02	ATTGTTTTGGCGGGACGTGG
wCREHhim4	wCREHhim4_03	TCATGTGAAAGGGATCGCC
wCREHhim6	wCREHhim6_01	AATGTGGCAGCCCATTTCAA
wCREHhim6	wCREHhim6_02	CGAATCTTTCTACCTAAGGG
wCREHhim6	wCREHhim6_03	TATGTGTTTTGATAATATGC
wCREHhim7	wCREHhim7_01	TTACTCGAAGGCATTGAGTA
wCREHhim7	wCREHhim7_02	TAGCTTATTGAAACCGGTTA
wCREHhim7	wCREHhim7_03	AATAGTGCGTTAAGATCTTG
wCREHhim8	wCREHhim8_01	TAGCAATAATTATCTATATT
wCREHhim8	wCREHhim8_02	AAAGCTCTGTTTATGCTTCC
wCREHhim8	wCREHhim8_03	CGTGACACGCGCCGCACCT
wCREHhim8	wCREHhim8_04	CCAAATACTGAGAGCATATC
wCREHhim10	wCREHhim10_01	CTGTTGCTAATGTAATTATA
wCREHhim10	wCREHhim10_02	TAATGTAGGCTTCTTAGTAA
wCREHhim10	wCREHhim10_03	TTTCCTACCGCGCGGATCGC
wCREHhim10	wCREHhim10_04	TGCCGTGATGTGTTGAAGAT

Table S2-4. List of oligos used to amplify regions containing target CREs.

Element	primer name	sequence
wcreAv01	NP_Av01_FWD	ATACTGTCCCACTGGGCGCT
wcreAv01	NP_Av01_REV	TCTGCCCTCAAACGCAACGG
wcreAv02	wcreAv02FWD	GAGATTGAGATCGAGAATCGAGA
wcreAv02	NP_Av02_REV	AACCATACGCACGAAGCTGCA
wcreAv03	NP_Av03_FWD	GGTACGTGAGCGCGTCTTCA
wcreAv03	NP_Av03_REV	TCAGATGCCCTCTCGGTGGA
wcreAv04	NP_Av04_FWD	GGCAATCCGAAATGGCCGTA
wcreAv04	NP_Av04_REV	CCACCTTTCCGTGGCATTGC
wcreAv06	NP_Av06_FWD	GCATTACTCCCGTCGCCGTT
wcreAv06	NP_Av06_REV	ACAGTGACGTACCGCAACCA
wcreAv07	NP_Av07_FWD	CTCAGCGCCGTGGTTGTCTA
wcreAv07	NP_Av07_REV	TTGCCCCCTTAGCCTTCTCA
wcreAv10	NP_Av10_FWD	CGGAATGCTGCCAATGATATGC
wcreAv10	NP_Av10_REV	CCCGACAGTCTGAGGAAGTCAT
wcreDp01	NP_Dp01_FWD	AGATTTTGTGGAGGCCCGGC
wcreDp01	NP_Dp01_REV	AAACCCTCCAGAGACGCCGT
wcreDp02	NP_Dp02_FWD	GCGGCTTCTTGTGGCTTCTC
wcreDp02	wcreDp02FWD	GGCCGTAACAAGCTGAGGCA
wcreDp03	NP_Dp03_FWD	CGCATCGGACGTCCAAGCAA
wcreDp03	NP_Dp03_REV	AGACCGCACACACGTCAAGA
wcreDp04	wcreDp04FWD	TCGCAGAGGATACAGCCGTC
wcreDp04	wcreDp04REV	ACGCCGCTATCATCAAACGC
wcreDp05	NP_Dp05_REV	GCATAGCACAAGTTTGCCCGA
wcreDp05	wcreDp05FWD	TCCAGGCTTGTAGAAACAATGTCA
wcreDp06	NP_Dp06_FWD	GACGCAAGAACCTTGCAGGC
wcreDp06	NP_Dp06_REV	GCACTCAATCAACGCGGCTG
wcreDp07	NP_Dp07_FWD	ATTCCTGGACGCTCCACATG
wcreDp07	NP_Dp07_REV	AGGTTTCGATACATGGAGTGCA
wcreHhim02	NP_Hhim02_FWD	AAACCGACGATGAGGAGGCG
wcreHhim02	NP_Hhim02_REV	ACCAGACAACCTTTGGGTTAGGA
wcreHhim03	NP_Hhim03_FWD	CCGATAGCGCAGAGTGCACC
wcreHhim03	NP_Hhim03_REV	CGTCTAACGACCCACGCCTC
wcreHhim04	NP_Hhim04_FWD	GAGTGCCGCCACAATGCTTC
wcreHhim04	NP_Hhim04_REV	GCCTCCAGCAAAGTCGGGAA
wcreHhim07	NP_Hhim07_FWD	GTCCGGTGTGGTCGAAGTGA
wcreHhim07	NP_Hhim07_REV	AACTGCCCGAATACAGGGCGT

Element	primer name	sequence
wcreHhim10	NP_Hhim10_FWD	GCTGCATCCCGTACACCACAA
wcreHhim10	NP_Hhim10_REV	AAACAGTCTCCCTACCCCGCA
wcreJc02	NP_Jc02_FWD	AGGGCTGATACGCACTAGCGG
wcreJc02	NP_Jc02_REV	GGTCGAGCGGTGTGTCGAA
wcreJc03	NP_Jc03_FWD	ACCCTACAACTGACTCGCCAA
wcreJc03	NP_Jc03_REV	AGACGTCCCGACGAACAACCA
wcreJc04	NP_Jc04_FWD	CGCGCGCACGAAATAGCTT
wcreJc04	NP_Jc04_REV	AGGTTTAACAAGCGCGTTGGG
wcreJc05	NP_Jc05_FWD	GGTCGGCGACACATACGTGA
wcreJc05	NP_Jc05_REV	TCGCGCTGAAGTGGATGTGT
wcreJc07	NP_Jc07_FWD	ACAAACCGCCGCGCTGAAATA
wcreJc07	NP_Jc07_REV	AAACGGATCGTGCCATCTCCA
wcreJc08	wcreJc08FWD	CGTCGCGTCACTAGTAATCAGTACA
wcreJc08	NP_Jc08_REV	CTGGTGCCGAGCCCAAGATA
wcreJc10	NP_Jc10_FWD	AAAATGAGCATCTGCCTGCGGT
wcreJc10	NP_Jc10_REV	GTTTCAACCGTGCTTACGTGCG
wcreVc01	NP_Vc01_FWD	AGAGATCACCAAGTTGACCGT
wcreVc01	NP_Vc01_REV	GCACTATGGACCTGTCGCCA
wcreVc02	NP_Vc02_FWD	GCGCATGGTTCACTAACGCC
wcreVc02	NP_Vc02_REV	AGTGTCGTGTGGCAAGTCGAG
wcreVc03	wcreVc03FWD2	CCGTCTGGGTAGCACCAACT
wcreVc03	wcreVc03REV	GACAGCATCCCTTCCAACAC
wcreVc04	wcreVc04FWD	TCTGCCCACGTTGTACGCTT
wcreVc04	NP_Vc04_REV	TGACAACACCTGCCGTGGTT
wcreVc05	NP_Vc05_FWD	ACGCTGTCTCACTCACGCTA
wcreVc05	NP_Vc05_REV	TCAGTGCAGGCGTGACTGATC
wcreVc06	NP_Vc06_FWD	CCCTACGGCGACTTCGTACG
wcreVc06	NP_Vc06_REV	CGCAAACGGGAGAGTTGACTG
wcreVc08	NP_Vc08_FWD	ATATGGAGCGCGGTTGTGCC
wcreVc08	NP_Vc08_REV	GTCGACAGACGTCATCTGCCTAGT
wcreVc09	NP_Vc09_FWD	GATGGCGCGCCTTTCCAAGT
wcreVc09	wcreVc09REV	AAGCGGTAACGCCTCGTTAAA
wcreVc10	NP_Vc10_FWD	AAAGGACTTGCTACCCGCCAC
wcreVc10	NP_Vc10_REV	TCGCAAGCTGGCATCGAAGA
wcreVc11	wcreVc06FWD	CGCATCGATTCACTGGCAGA
wcreVc11	NP_Vc11_REV	TCAACTCACGGTGACACGCT

Table S2-5. List of all mKOs obtained from CRE shotgun CRISPR/Cas9 experiments. Knock-out effect in adult wings is code as Enhancer-like (E), Silencer-like (S), presence of Enhancer- and Silencer-like effects on the same wing surface (E/S), ambiguous effects (E/S?). DFw=Dorsal Forewing, DHw=Dorsal Hindwing, VFw= Ventral Forewing, VHw=Ventral Hindwing.

Species	Ortholog CRE #	Species CRE ID	Mutant - ID	DFw	DHw	VFw	VHw	Size mutant clones
D. plexippus	CRE #03	wcreDp01	wcreDp01-01	S	E	E	E	Medium to large
D. plexippus	CRE #03	wcreDp01	wcreDp01-02	E/S	S	E	E	Medium to large
D. plexippus	CRE #03	wcreDp01	wcreDp01-03	S	-	E	E	Medium to large
D. plexippus	CRE #03	wcreDp01	wcreDp01-06	-	E	E	E	Medium to large
D. plexippus	CRE #03	wcreDp01	wcreDp01-07	-	E	E	-	Medium to large
D. plexippus	CRE #03	wcreDp01	wcreDp01-10	-	-	E	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-01	S	-	E	-	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-03	S	-	E/S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-04	S	-	E/S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-05	E	-	E/S	E/S	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-06	S	-	S	E/S	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-07	S	S	-	S	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-09	S	S	-	-	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-11	S	S	S	S	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-12	E	-	E	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-13	-	-	S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-14	S	-	-	-	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-18	S	S	S	-	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-20	S	E/S	E/S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-22	S	E/S?	E/S?	E	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-23	S	S	S	E/S	Small

D. plexippus	CRE #08	wcreDp02	wcreDp02-24	S	E/S?	S	E/S	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-25	S	-	S	E	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-26	E/S	E	E	-	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-31	E/S	E/S	S	E	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-34	E/S	-	E/S	E	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-40	S	E	E/S	-	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-41	-	S	E/S?	E	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-42	S	-	S	-	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-47	S	E/S?	-	S	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-48	S	-	S	-	Small
D. plexippus	CRE #08	wcreDp02	wcreDp02-49	S	-	S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-50	S	S	S	E	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-51	S	E/S?	E	S	Medium to large
D. plexippus	CRE #08	wcreDp02	wcreDp02-52	S	-	-	E	Medium to large
D. plexippus	CRE #21	wcreDp03	wcreDp03-02	S	S	E/S	E	Medium to large
D. plexippus	CRE #21	wcreDp03	wcreDp03-14	-	-	E	E	Medium to large
D. plexippus	CRE #21	wcreDp03	wcreDp03-25	S	-	E	-	Medium to large
D. plexippus	CRE #21	wcreDp03	wcreDp03-27	E	-	E	E	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-01	-	E	-	E	Small
D. plexippus	CRE #09	wcreDp04	wcreDp04-07	S	S	E/S	E	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-09	S	-	S	E	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-11	S	-	E/S	E	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-15	-	S	S	E	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-17	S	-	S	-	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-18	-	S	E	E	Medium to large

D. plexippus	CRE #09	wcreDp04	wcreDp04-21	-	S	S	-	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-24	S	S	-	-	Medium to large
D. plexippus	CRE #09	wcreDp04	wcreDp04-28	S	-	E/S	E/S	Medium to large
D. plexippus	CRE #16	wcreDp05	wcreDp05-01	E/S	E	E	E	Medium to large
D. plexippus	CRE #16	wcreDp05	wcreDp05-09	S	E/S?	E	E	Small
D. plexippus	CRE#02	wcreDp06	wcreDp06-01	E/S?	E	E	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-02	S	-	-	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-03	-	S	-	E	Small
D. plexippus	CRE#02	wcreDp06	wcreDp06-04	S	-	E	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-06	E/S	S	-	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-09	S	S	S	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-10	E	-	E	E	Small
D. plexippus	CRE#02	wcreDp06	wcreDp06-11	-	E	-	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-12	E	E	-	E	Small
D. plexippus	CRE#02	wcreDp06	wcreDp06-13	E/S	E	-	E	Medium to large
D. plexippus	CRE#02	wcreDp06	wcreDp06-22	-	S	-	-	Small
D. plexippus	CRE#02	wcreDp06	wcreDp06-23	S	-	-	E	Small
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-01	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-02	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-03	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-04	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-05	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-06	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-07	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-08	E	E	E	E	Medium to large

D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-10	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-11	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-12	E	-	E	E	Small
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-18	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-20	-	-	-	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-22	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-23	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-24	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-25	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-26	E	E	E	E	Medium to large
D. plexippus	CRE #05, Promoter	wcreDp07	wcreDp07-27	E	E	E	E	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-01	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-02	S	S	E/S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-03	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-04	S	E/S	E/S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-05	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-06	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-07	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-08	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-09	S	-	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-10	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-11	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-12	S	S	E/S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-13	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-14	S	S	S	S	Medium to large

A. vanillae	CRE #07	wcreAv01	wcreAv01-15	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-16	S	S	S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-17	S	S	S	E/S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-m1	S	S	E/S	S	Medium to large
A. vanillae	CRE #07	wcreAv01	wcreAv01-m2	E/S	S	E/S	S	Medium to large
A. vanillae	CRE #17	wcreAv02	wcreAv02-01	S	S	E/S	S	Medium to large
A. vanillae	CRE #17	wcreAv02	wcreAv02-02	S	S	S	S	Small
A. vanillae	CRE #17	wcreAv02	wcreAv02-03	-	-	E/S	S	Small
A. vanillae	CRE #17	wcreAv02	wcreAv02-04	E	-	S	-	Small
A. vanillae	CRE #17	wcreAv02	wcreAv02-05	E	S	E/S	E/S	Medium to large
A. vanillae	CRE #17	wcreAv02	wcreAv02-06	S	S	-	E	Medium to large
A. vanillae	CRE #17	wcreAv02	wcreAv02-07	S	S	E	S	Medium to large
A. vanillae	CRE #15	wcreAv03	wcreAv03-01	-	-	-	S	Small
A. vanillae	CRE #15	wcreAv03	wcreAv03-02	E	E	E	E	Medium to large
A. vanillae	CRE #15	wcreAv03	wcreAv03-03	S	S	S	S	Medium to large
A. vanillae	CRE #15	wcreAv03	wcreAv03-04	E	S	E	S	Medium to large
A. vanillae	CRE #20	wcreAv04	wcreAv04-01	S	-	S	S	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-02	S	-	S	-	damage mutant
A. vanillae	CRE #20	wcreAv04	wcreAv04-03	E	E	E	E	Medium to large
A. vanillae	CRE #20	wcreAv04	wcreAv04-04	E	E	E/S	E	Medium to large
A. vanillae	CRE #20	wcreAv04	wcreAv04-05	E	E	E	E	Medium to large
A. vanillae	CRE #20	wcreAv04	wcreAv04-06	-	E	-	-	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-07	-	-	S	-	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-08	-	-	E/S	E	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-09	S	E	-	-	Small

A. vanillae	CRE #20	wcreAv04	wcreAv04-10	E	E	S	E	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-11	S	S	S	S	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-12	-	-	S	S	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-13	-	-	E	-	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-15	-	E	E	E	Small
A. vanillae	CRE #20	wcreAv04	wcreAv04-16	-	-	E	-	Small
A. vanillae	CRE #06	wcreAv05	wcreAv05-01	-	-	E	E	Medium to large
A. vanillae	CRE #06	wcreAv05	wcreAv05-02	-	-	-	S	Small
A. vanillae	CRE #06	wcreAv05	wcreAv05-03	-	S	E/S	S	Medium to large
A. vanillae	CRE #06	wcreAv05	wcreAv05-04	-	-	S	S	Medium to large
A. vanillae	CRE #06	wcreAv05	wcreAv05-07	E	E	E/S	E	Small
A. vanillae	CRE #06	wcreAv05	wcreAv05-08	-	S	-	E	Small
A. vanillae	CRE #12	wcreAv06	wcreAv06-01	S	E/S	E/S	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-02	E/S?	S	S	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-03	S	-	-	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-04	S	S	E/S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-05	S	-	E/S	E	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-06	-	S	E/S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-07	-	-	-	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-08	S	E/S?	E/S	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-09	E/S	E	E/S	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-10	S	S	E	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-11	S	S	S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-12	E/S	E/S?	E/S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-13	-	S	S	E/S	Medium to large

A. vanillae	CRE #12	wcreAv06	wcreAv06-14	S	S	S	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-15	E	E	E	E/S?	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-16	S	S	E/S?	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-17	E	E/S?	E	E	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-18	S	S	S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-20	S	S	E/S?	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-21	S	S	E	E/S?	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-22	S	S	E	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-23	-	-	S	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-24	S	E/S?	-	E/S?	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-25	S	S	E/S	E/S?	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-26	E/S?	S	E/S	E	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-27	-	E/S?	-	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-28	S	-	S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-29	E/S?	E/S?	E/S	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-30	E	-	E/S	E/S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-31	S	S	S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-32	S	S	E/S	S	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-33	E	S	E/S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-34	E/S?	S	E/S	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-35	-	S	-	-	Medium to large
A. vanillae	CRE #12	wcreAv06	wcreAv06-37	-	-	-	E/S	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-01	-	-	-	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-03	-	E	E	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-04	E	E	E/S	S	Medium to large

A. vanillae	CRE #22	wcreAv07	wcreAv07-05	-	-	-	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-06	E/S?	E	E	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-08	E	-	E	S	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-09	E	E	E	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-10	E	-	E	E/S?	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-11	-	-	E	-	Small
A. vanillae	CRE #22	wcreAv07	wcreAv07-12	-	-	E	S	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-13	E	E	E	E/S?	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-14	E	E	E	-	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-15	-	E	-	E/S	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-16	-	E	E	-	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-17	E	-	E	-	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-18	-	E	S	E	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-20	E	E	E	S	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-23	-	-	E	-	Small
A. vanillae	CRE #22	wcreAv07	wcreAv07-25	-	-	E	-	Small
A. vanillae	CRE #22	wcreAv07	wcreAv07-26	E/S?	E/S?	E	-	Small
A. vanillae	CRE #22	wcreAv07	wcreAv07-27	-	-	E	-	Small
A. vanillae	CRE #22	wcreAv07	wcreAv07-28	-	-	E	-	Medium to large
A. vanillae	CRE #22	wcreAv07	wcreAv07-29	E/S?	S	-	-	Medium to large
A. vanillae	CRE #01	wcreAv09	wcreAv09-02	-	-	E S	E/S?	Small
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-01	E	-	E	E	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-02	E	E	E	E	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-03	E	-	E	E	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-04	E	E	E	E	Medium to large

A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-05	E/S?	E	E	-	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-06	E	E	E	E	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-07	E	E	E	E	Medium to large
A. vanillae	CRE #05, Promoter	wcreAv10	wcreAv10-08	E	E	E	E	Medium to large
H. himera	CRE #07	wcreHhim01	wcreHhim01-01	S	-	S	-	Medium to large
H. himera	CRE #07	wcreHhim01	wcreHhim01-02	E	-	E	-	Medium to large
H. himera	CRE #07	wcreHhim01	wcreHhim01-04	S	-	E/S?	-	Small
H. himera	CRE #07	wcreHhim01	wcreHhim01-05	E/S?	-	E/S?	-	doubtful mutant
H. himera	CRE #07	wcreHhim01	wcreHhim01-06	E	-	E	-	Small
H. himera	CRE #07	wcreHhim01	wcreHhim01-07					Medium to large
H. himera	CRE #17	wcreHhim02	wcreHhim02-01	E S	-	E S	-	Medium to large
H. himera	CRE #17	wcreHhim02	wcreHhim02-02	S	-	S	-	Medium to large
H. himera	CRE #17	wcreHhim02	wcreHhim02-03	S		E S		Medium to large
H. himera	CRE #17	wcreHhim02	wcreHhim02-04	E S	-	E S	-	Medium to large
H. himera	CRE #17	wcreHhim02	wcreHhim02-05	S	-	S	-	Small
H. himera	CRE #17	wcreHhim02	wcreHhim02-06	E/S?	-	E/S?	-	doubtful mutant
H. himera	CRE #17	wcreHhim02	wcreHhim02-07	-	-	-	-	Medium to large
H. himera	CRE #15	wcreHhim03	wcreHhim03-02	S	-	S	-	Medium to large
H. himera	CRE #15	wcreHhim03	wcreHhim03-03	E S	-	E S	-	Medium to large
H. himera	CRE #15	wcreHhim03	wcreHhim03-04	E S	-	S	-	Medium to large
H. himera	CRE #15	wcreHhim03	wcreHhim03-05	E/S?	-	E S	-	Medium to large
H. himera	CRE #20	wcreHhim04	wcreHhim04-01	E/S?	-	E/S?	-	
H. himera	CRE #20	wcreHhim04	wcreHhim04-03	E/S?	-	E/S?	-	
H. himera	CRE #20	wcreHhim04	wcreHhim04-05	E/S?	-	E/S?	-	
H. himera	CRE #12	wcreHhim06	wcreHhim06-01	E S	-	E S	E	Small

H. himera	CRE #12	wcreHhim06	wcreHhim06-02	S	-	S	-	Small
H. himera	CRE #12	wcreHhim06	wcreHhim06-05	E/S?	-	E	-	Medium to large
H. himera	CRE #13	wcreHhim07	wcreHhim07-01	E/S?	-	S	-	Small
H. himera	CRE #13	wcreHhim07	wcreHhim07-02	E/S?	-			Small
H. himera	CRE #13	wcreHhim07	wcreHhim07-03	E	E	E	E	Medium to large
H. himera	CRE #13	wcreHhim07	wcreHhim07-04	E S	-	E	E/S?	Medium to large
H. himera	CRE #14	wcreHhim08	wcreHhim08-01	S	-	S	-	Medium to large
H. himera	CRE #14	wcreHhim08	wcreHhim08-02	E/S?	-	S	-	Small
H. himera	CRE #14	wcreHhim08	wcreHhim08-03	S	-	S	-	Small
H. himera	CRE #05, Promoter	wcreHhim10	wcreHhim10-m1	E	-	E	-	Medium to large
J. coenia	CRE #10	wcreJc01	wcreJc01-27	S	S	Y	E	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-30	-	UBX	-	E/S?	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-31	E/S?	-	-	-	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-32	E/S?	-	-	-	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-33	E/S?	-	-	-	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-37	E/S?	-	E	-	Small
J. coenia	CRE #10	wcreJc01	wcreJc01-43	-	-	E, S?	-	Small
J. coenia	CRE #07	wcreJc02	wcreJc02-59	-	UBX	-	E	Small
J. coenia	CRE #07	wcreJc02	wcreJc02-64	-	Y	-	E/S?	Small
J. coenia	CRE #07	wcreJc02	wcreJc02-65	-	-	-	E	Small
J. coenia	CRE #07	wcreJc02	wcreJc02-66	S	-	-	-	Small
J. coenia	CRE #20	wcreJc03	wcreJc03-01	-	-	E	-	Medium to large
J. coenia	CRE #20	wcreJc03	wcreJc03-34	E/S?	-	-	-	Small
J. coenia	CRE #20	wcreJc03	wcreJc03-49	-	-	S	-	Small
J. coenia	CRE #20	wcreJc03	wcreJc03-50	E	Y	-	-	Small

J. coenia	CRE #20	wcreJc03	wcreJc03-52	Y	Y	-	E	small
J. coenia	CRE #20	wcreJc03	wcreJc03-54	E	-	-	Y	small
J. coenia	CRE #20	wcreJc03	wcreJc03-56	E	-	-	-	Small
J. coenia	CRE #06	wcreJc04	wcreJc04-14	ES?	-	-	E	Small
J. coenia	CRE #06	wcreJc04	wcreJc04-15	-	-	E	Y	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-04	-	-	S	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-06	-	-	S	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-09	-	-	E/S?	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-10	-	E/S?	-	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-17a	-	-	E/S?	E?, S	Medium to large
J. coenia	CRE #19	wcreJc05	wcreJc05-17b	E	-	-	E/S?	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-18	-	-	-	E	Medium to large
J. coenia	CRE #19	wcreJc05	wcreJc05-20	S	Y	E	E	Medium to large
J. coenia	CRE #19	wcreJc05	wcreJc05-21	-	-	-	E	Medium to large
J. coenia	CRE #19	wcreJc05	wcreJc05-24	-	-	E	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-25	-	Y	-	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-26	-	-	S	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-27	-	-	Y	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-29	-	-	E/S?	E	Medium to large
J. coenia	CRE #19	wcreJc05	wcreJc05-30	S	-	-	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-32	S	-	-	-	Small
J. coenia	CRE #19	wcreJc05	wcreJc05-33	-	-	E	-	Small
J. coenia	CRE #17	wcreJc06	wcreJc06-24	-	-	E/S?	E/S?	Small
J. coenia	CRE #30	wcreJc06	wcreJc06-30	-	S	-	-	Medium to large
J. coenia	CRE #17	wcreJc06	wcreJc06-38	E/S?	-	-	-	Small

J. coenia	CRE #17	wcreJc06	wcreJc06-41	-	-	E	-	Small
J. coenia	CRE #22	wcreJc07	wcreJc07-13	-	-	E	E	Medium to large
J. coenia	CRE #22	wcreJc07	wcreJc07-15	-	-	E	-	Small
J. coenia	CRE #22	wcreJc07	wcreJc07-16	-	-	E	-	Small
J. coenia	CRE #12	wcreJc09	wcreJc09-01	-	-	E	E	Medium to large
J. coenia	CRE #12	wcreJc09	wcreJc09-02	-	-	-	E	Medium to large
J. coenia	CRE #12	wcreJc09	wcreJc09-03	-	-	E	E	Medium to large
J. coenia	CRE #12	wcreJc09	wcreJc09-04	-	-	E	-	Medium to large
J. coenia	CRE #12	wcreJc09	wcreJc09-05	-	-	-	E	Medium to large
J. coenia	CRE #12	wcreJc09	wcreJc09-06	-	-	E/S?	-	Small
J. coenia	CRE #12	wcreJc09	wcreJc09-07	-	-	E/S?	-	Small
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-01	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-02	E	-	E	-	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-03	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-04	E	-	E	-	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-05	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-06	-	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-07	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-08	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-09	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-10	-	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-11	E	-	E	-	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-12	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-13	-	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-14	E	-	E	E	Medium to large

J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-15	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-16	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-17	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-18	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-19	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-20	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-21	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-22	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-23	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-24	-	-	-	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-25	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-26	E	-	E	E	Medium to large
J. coenia	CRE #05, Promoter	wcreJc10	wcreJc10-27	E	-	E	E	Medium to large
V. cardui	CRE #10	wcreVc01	wcreV01-12	-	-	S	-	medium to large
V. cardui	CRE #10	wcreVc01	wcreV01-13	E/S?	-	E/S?	E/S?	Small
V. cardui	CRE #10	wcreVc01	wcreV01-15	E/S?	-	E/S?	E/S?	Small
V. cardui	CRE #10	wcreVc01	wcreV01-16	-	-	E	-	Small
V. cardui	CRE #10	wcreVc01	wcreV01-17	-	-	-	S	Small
V. cardui	CRE #10	wcreVc01	wcreV01-23	-	-	E	-	Small
V. cardui	CRE #07	wcreVc02	wcreV02-03	E	E	-	-	Small
V. cardui	CRE #07	wcreVc02	wcreV02-04	E/S?	-	-	-	Small
V. cardui	CRE #07	wcreVc02	wcreV02-05	-	-	E	-	Small
V. cardui	CRE #07	wcreVc02	wcreV02-06	E/S?	-	E/S?	-	Small
V. cardui	CRE #07	wcreVc02	wcreV02-07	E/S?	-	-	E	Small
V. cardui	CRE #07	wcreVc02	wcreV02-10	-	-	-	E	Small

V. cardui	CRE #07	wcreVc02	wcreV02-20	E/S?	-	E/S?	E	Medium to large
V. cardui	CRE #07	wcreVc02	wcreV02-21	-	-	E/S?	-	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-08	E	E/S?	E/S?	-	Medium to large
V. cardui	CRE # 23	wcreVc03	wcreV03-12	-	-	E/S?	-	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-17	-	-	S	-	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-18	-	-	E/S?	E/S?	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-20	-	-	-	E	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-21				S	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-23	-	-	E/S?	S	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-29	-	-	E	-	Medium to large
V. cardui	CRE # 23	wcreVc03	wcreV03-34	E/S?	-	E	E	Medium to large
V. cardui	CRE # 23	wcreVc03	wcreV03-35	-	-	E/S?	E/S?	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-39	-	-	-	E	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-41	S	-	-	E	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-46	-	-	E	-	Small
V. cardui	CRE # 23	wcreVc03	wcreV03-53	S	S	S	S	Medium to large
V. cardui	CRE # 23	wcreVc03	wcreV03-57	E	S	E/S?	S	Medium to large
V. cardui	CRE # 23	wcreVc03	wcreV03-62	E	-	E	E	Small
V. cardui	CRE #11	wcreVc04	wcreV04-06	E	-	E	-	Small
V. cardui	CRE #11	wcreVc04	wcreV04-15	-	-	S	-	Small
V. cardui	CRE #11	wcreVc04	wcreV04-21	E	-	-	-	Small
V. cardui	CRE #11	wcreVc04	wcreV04-32	-	-	S	-	Small
V. cardui	CRE #11	wcreVc04	wcreV04-34	-	-	-	S	Medium to large
V. cardui	CRE #19	wcreVc05	wcreV05-01	-	-	-	E	Small
V. cardui	CRE #19	wcreVc05	wcreV05-03	E	-	S	S	Small

V. cardui	CRE #19	wcreVc05	wcreV05-09	-	-	-	E	Small
V. cardui	CRE #19	wcreVc05	wcreV05-12	E	-	E/S	-	Small
V. cardui	CRE #19	wcreVc05	wcreV05-13	E	-	E/S?	E/S?	Small
V. cardui	CRE #19	wcreVc05	wcreV05-17	S	-	E	E	Medium to large
V. cardui	CRE #19	wcreVc05	wcreV05-18	E	-	-	-	Medium to large
V. cardui	CRE #19	wcreVc05	wcreV05-22	E	-	E/S?	-	Small
V. cardui	CRE #19	wcreVc05	wcreV05-23	-	-	E/S?	E	Small
V. cardui	CRE #19	wcreVc05	wcreV05-24	S	-	E/S?	E	Medium to large
V. cardui	CRE #19	wcreVc05	wcreV05-26	E	-	E	-	Small
V. cardui	CRE #19	wcreVc05	wcreV05-28	-	-	-	E/S?	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-01	S	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-02	S	-	-	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-03	S	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-06	S	-	S	E	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-10		-	-	E	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-12	-	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-13	-	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-14	S	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-15	E/S	-	S	S	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-16	S	-	S	-	Medium to large
V. cardui	CRE #17	wcreVc06	wcreV06-19	S	-	S	S	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-01	E/S	-	E/S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-02	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-06	E	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-07	S	-	S	E	Medium to large

V. cardui	CRE #22	wcreVc07	wcreV07-08	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-09	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-10	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-11	-	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-12	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-13	S	-	-	-	Small
V. cardui	CRE #22	wcreVc07	wcreV07-14	-	-	S, E?	E/S?	Small
V. cardui	CRE #22	wcreVc07	wcreV07-17	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-18	S	E	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-19	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-21	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-24	-	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-25	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-31	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-32	S	-	E/S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-33	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-35	-	-	E/S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-36	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-40	S	E/S?	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-42	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-48	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-51	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-52	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-54	E/S	E	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-55	S	-	S	-	Medium to large

V. cardui	CRE #22	wcreVc07	wcreV07-57	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-58	-	-	S	-	Small
V. cardui	CRE #22	wcreVc07	wcreV07-60	-	-	S	-	Small
V. cardui	CRE #22	wcreVc07	wcreV07-62	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-66	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-68	-	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-69	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-70	-	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-73	S	-	S	E	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-78	S	-	S	E/S?	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-80a	S	-	-	S	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-80b	S	-	S	-	Medium to large
V. cardui	CRE #22	wcreVc07	wcreV07-82	S	-	S	-	Medium to large
V. cardui	CRE #20	wcreVc09	wcreVc09-02	S	-	S	-	Medium to large
V. cardui	CRE #20	wcreVc09	wcreVc09-13	-	-	-	E/S?	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-23	-	-	-	E/S?	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-27	-	-	E/S?	-	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-29	-	-	S	-	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-36	-	-	E	-	Medium to large
V. cardui	CRE #20	wcreVc09	wcreVc09-37	S	-	-	-	Medium to large
V. cardui	CRE #20	wcreVc09	wcreVc09-40	-	-	S	S	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-41	S	-	S	E/S?	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-42	-	-	-	E	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-43	-	-	S	-	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-44	S	-	S	-	Medium to large

V. cardui	CRE #20	wcreVc09	wcreVc09-45	S	-	S	E/S?	Medium to large
V. cardui	CRE #20	wcreVc09	wcreVc09-47	E/S?	-	S	-	Small
V. cardui	CRE #20	wcreVc09	wcreVc09-48	S	-	S	-	Small
V. cardui	CRE #05, Promoter	wcreVc10	wcreVc10-01	E	E	E	-	Small
V. cardui	CRE #18	wcreVc11	wcreVc11-01	S	-	S	-	Medium to large
V. cardui	CRE #18	wcreVc11	wcreVc11-03	-	-	S	-	Medium to large
V. cardui	CRE #18	wcreVc11	wcreVc11-05	S	-	-	-	Small
V. cardui	CRE #18	wcreVc11	wcreVc11-06	-	-	-	S	Medium to large
V. cardui	CRE #18	wcreVc11	wcreVc11-07	S	-	S	-	Medium to large
V. cardui	CRE #18	wcreVc11	wcreVc11-08	S	E/S?	S	-	Medium to large

Appendix for Chapter 3

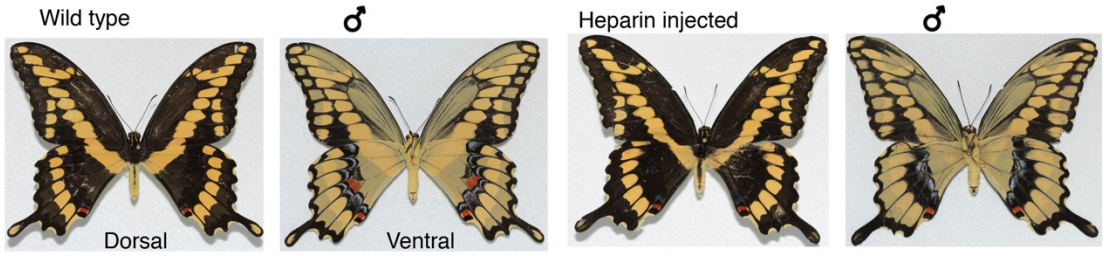


Figure S3-1. Effects of heparin injections in *Papilio cresphontes*.

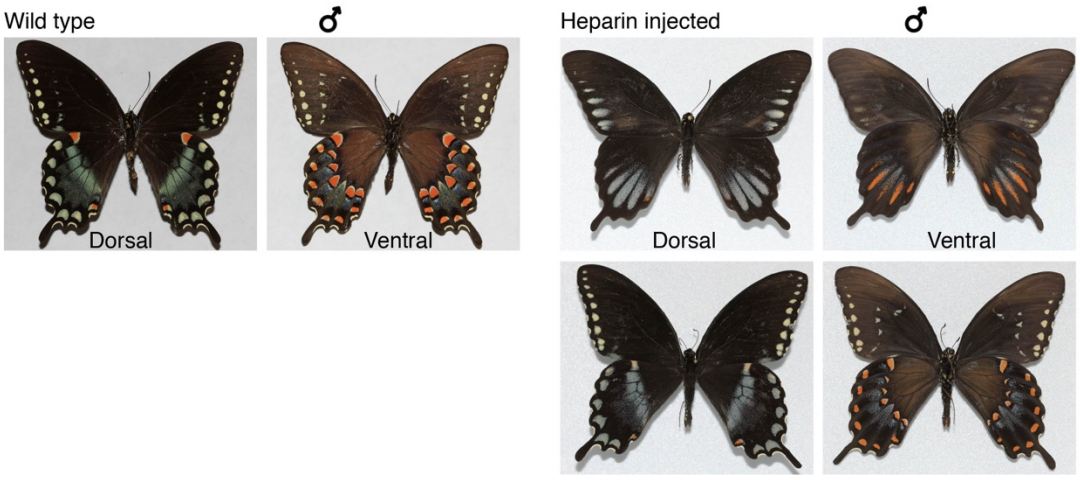


Figure S3-2. Effects of heparin injections in *Papilio trolius*.

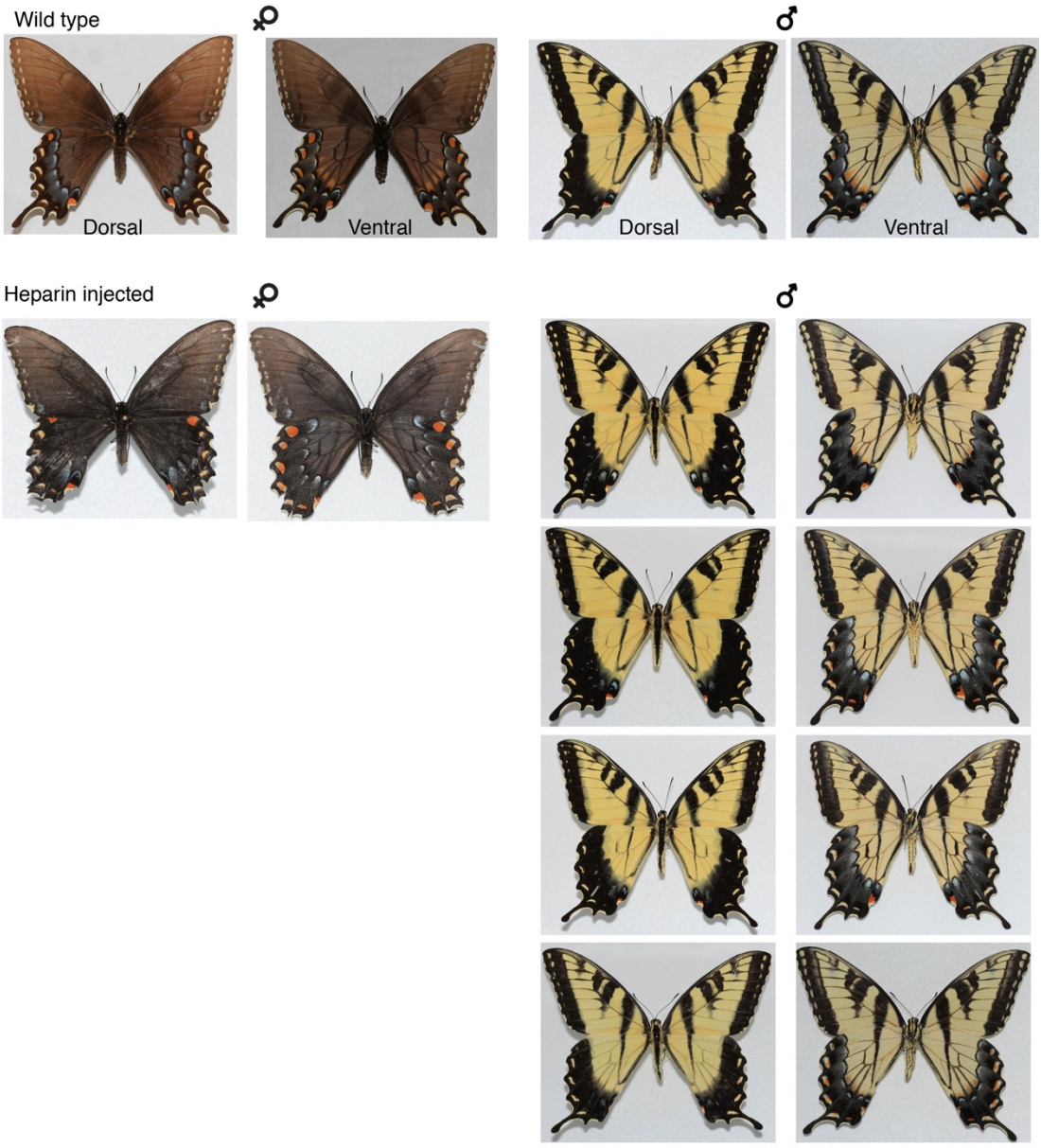


Figure S3-3. Effects of heparin injections in *Papilio glaucus*.

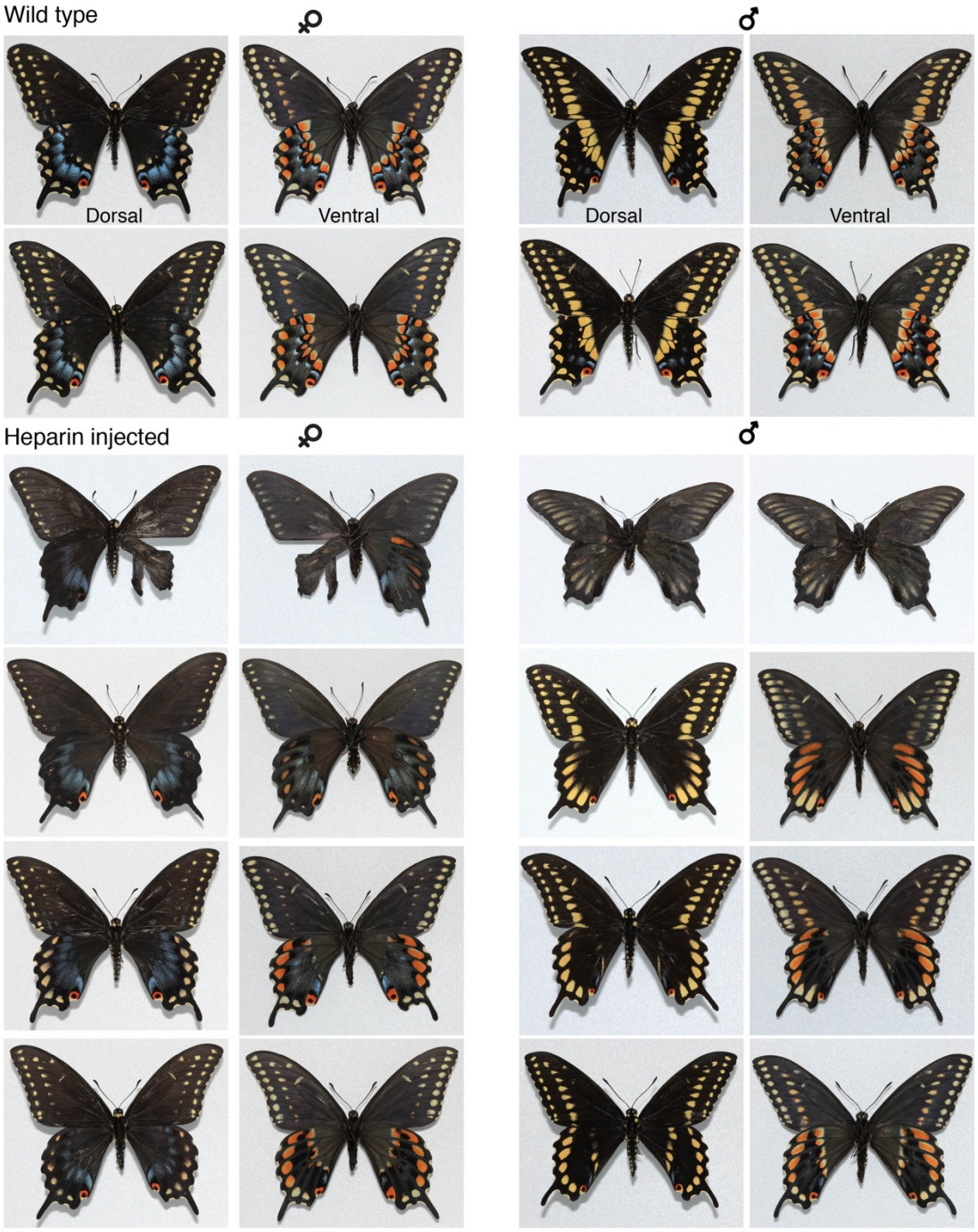


Figure S3-4. Effects of **(A)** heparin injections, **(B, Next Page)** CRISPR/Cas9 *WntA* mkO, **(C, Next Page)** *Wnt6* in *Papilio polyxenes*.

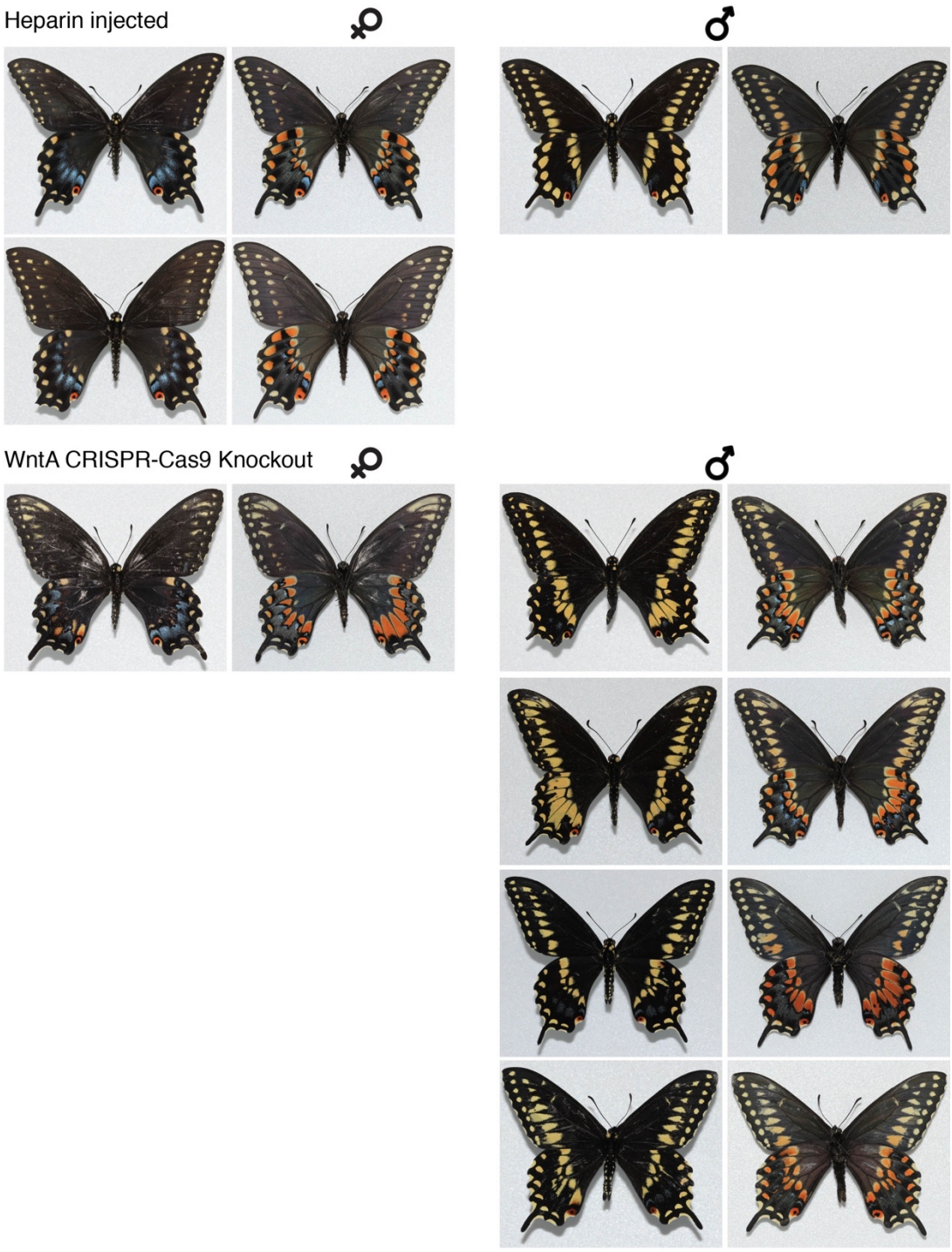


Figure S3-4. Continuation effects in *Papilio polyxenes.*

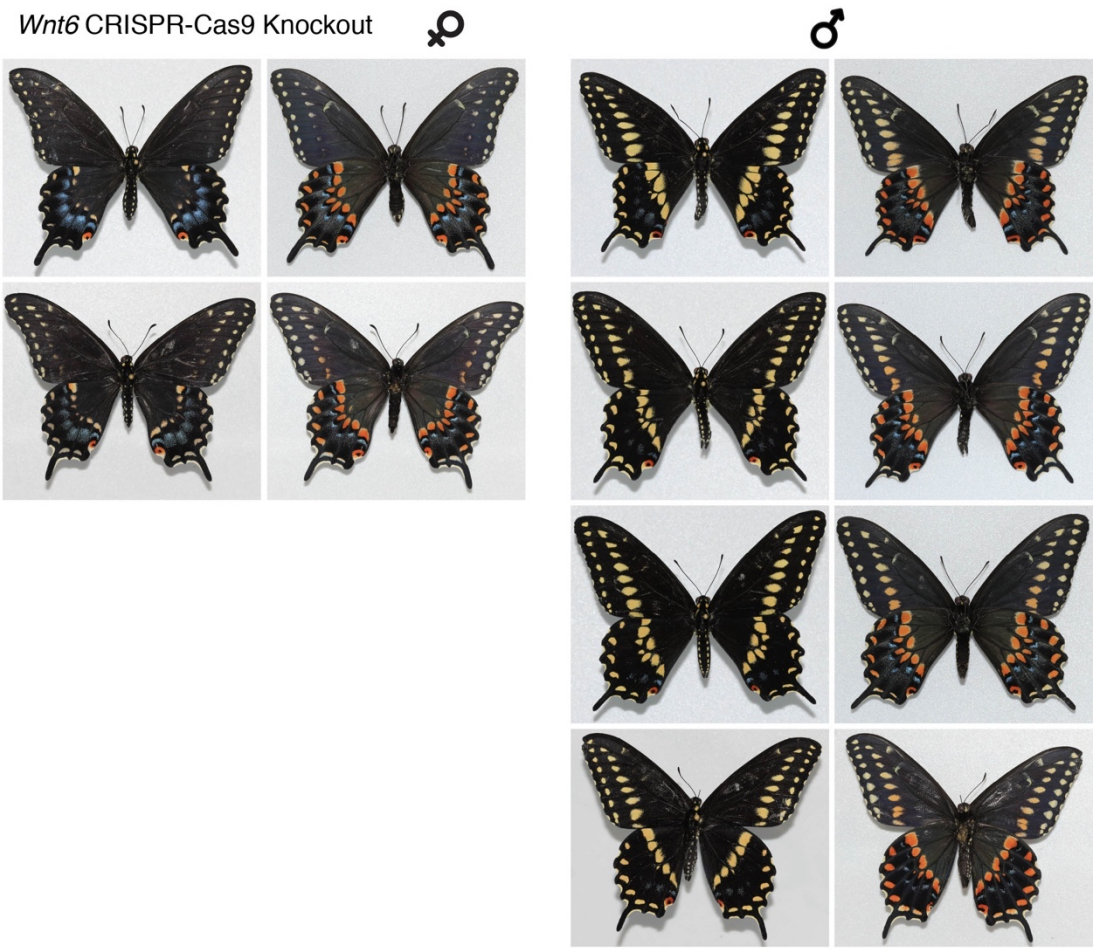


Figure S3-4. Continuation effects in *Papilio polyxenes*.

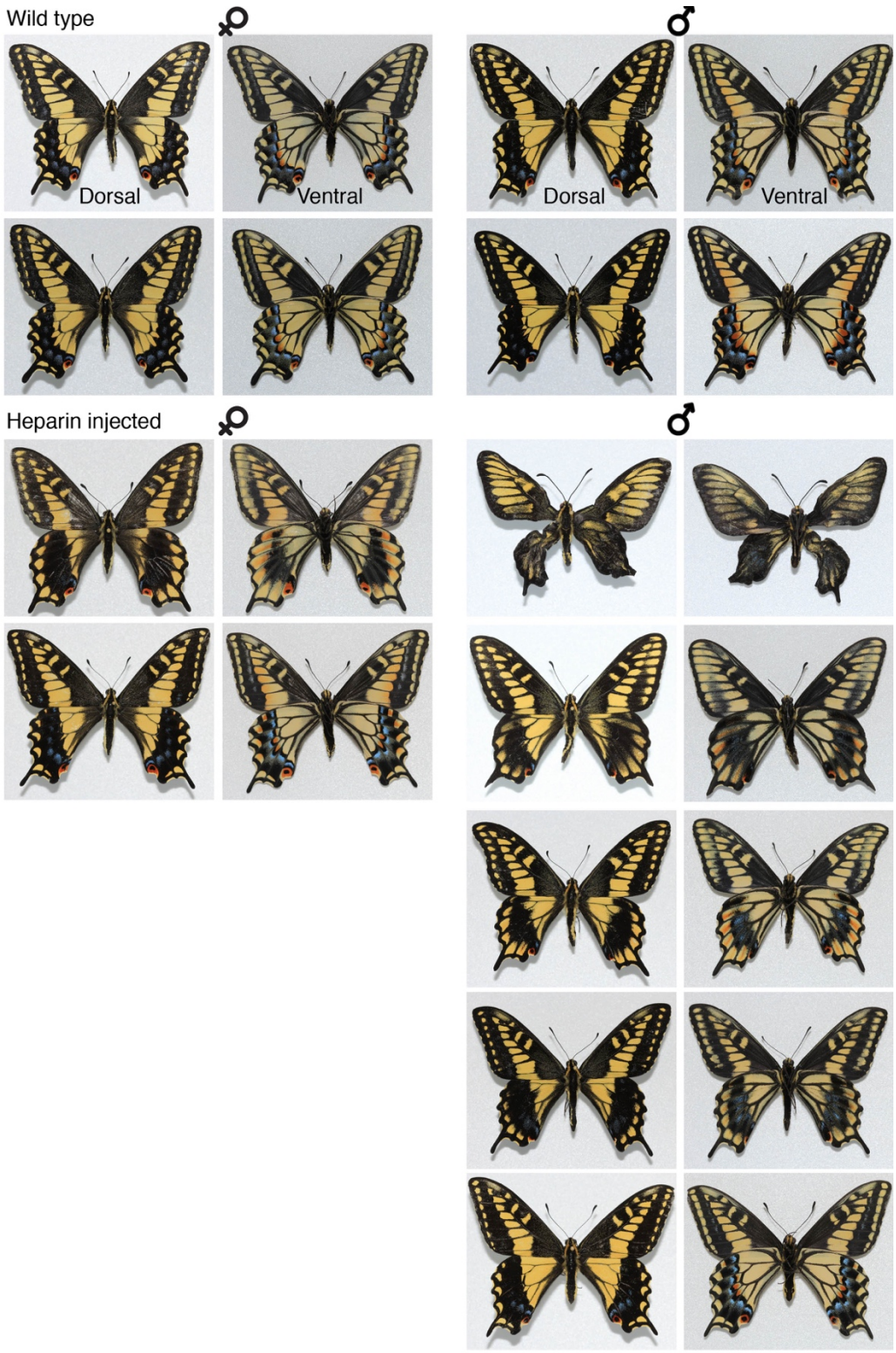


Figure S3-5. Effects of **(A)** heparin injections, **(B, Next Page)** Dextran sulfate, **(C, Next Page)** CRISPR/Cas9 *WntA* mkO, **(D, Next Page)** *Wnt6* in *Papilio zelicaon.*

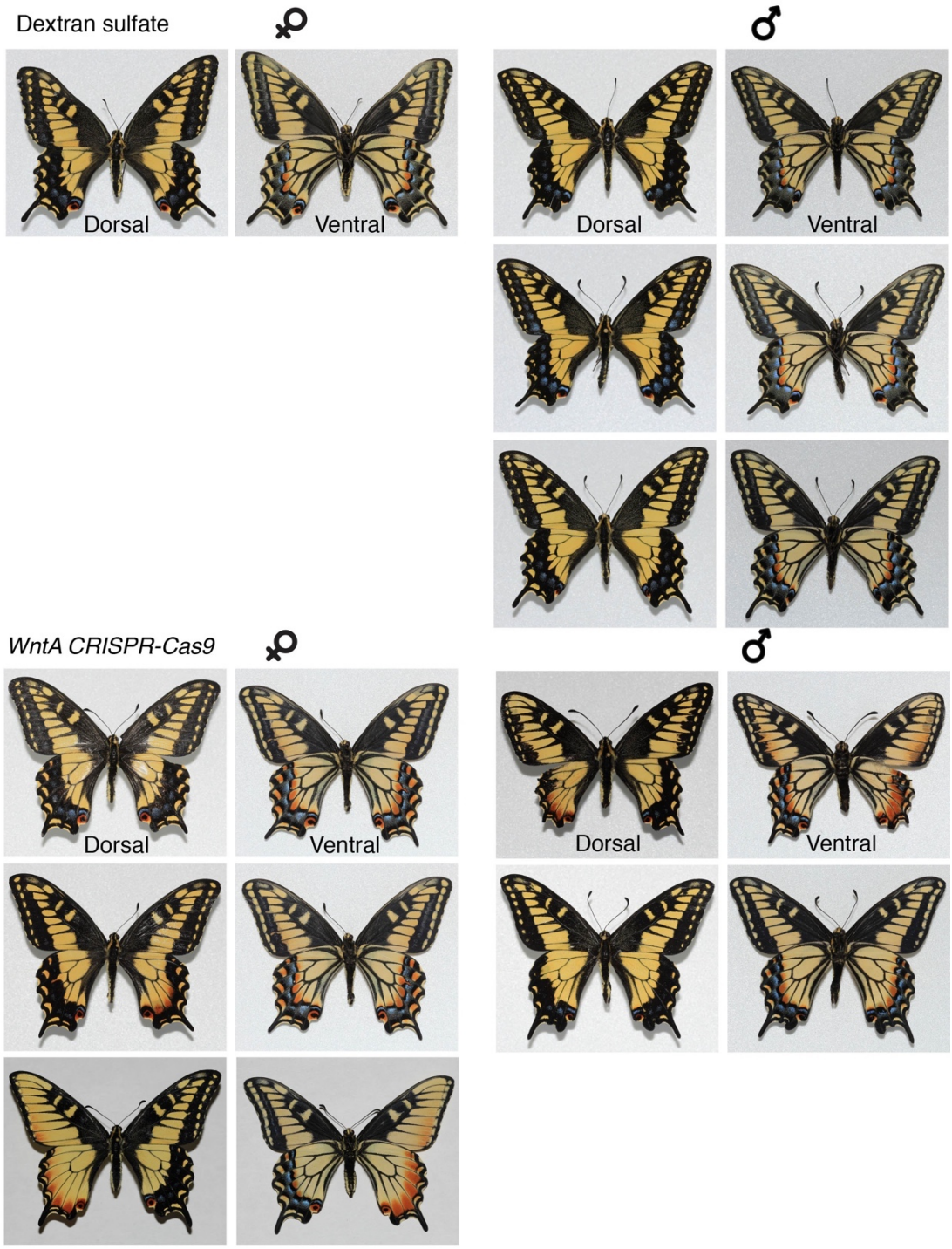


Figure S3-4. Continuation effects in *Papilio zelicaon.*

Figure S3-4. Continuation effects in *Papilio zelicaon.*

Table S3-1. *Papilio* species population source and rearing information. All the adults were hand feed with 10% sugar water solution.

Species	Population origin	Collected	Maiting	Caterpillar host plant
Papilio zelicaon	Albany, CA	Yiwen Zhu. Collected eggs and larvae on fennel	Hand paired ~10-20 females per generation	Fennel (*Foeniculum vulgare*) Parsley (*Petroselinum crispum*)
Papilio polyxenes	Ithaca, NY	Alan Liang, Brian Liang. Collected adult female and larvae on parsley	Hand paired ~10-20 females per generation	Fennel (*Foeniculum vulgare*) Parsley (*Petroselinum crispum*)
	South Carolina	Purchased twelve pupae from Kevin Koffel		
Papilio glaucus	Iowa	purchased ~180 eggs from David Thompson	Hand paired 3 females for second generation	Tuliptree (*Liriodendron tulipifera)* Chokecherry (*Prunus virginiana*) Ash (*Fraxinus sp.*)
Papilio troilus	Ithaca, NY	Alan Liang, Brian Liang (collected larva on spicebush)	Hand paired 2 females	Spicebush (*Lindera benzoin*)
	Allegan, MI	Purchased ten pupae from Kevin Collison		
Papilio cresphontes	Ithaca, NY	Alan Liang, Brian Liang. Collected adult females and eggs on rue and amur cork tree	No mating	Rue (Ruta graveolens)

Table S3-2. Heparin and Dextran sulfate injections experiments.

Species	Age of pupae: Hours after pupation	Treatment	Amount injected (ug)	Optimal* amount (ug)	Pupae injected	Adults with phenotype
P. zelicaon	4-20	Heparin	60-200	80-120	~20	10
P. polyxenes	4-20	Heparin	60-200	80-120	30-40	16
P. glaucus	4-24	Heparin	60-320	>120	~30	6
P. troilus	4-10	Heparin	60-160	100-120	~15	2
P. cresphontes	4-10	Heparin	60-160	120	~20	1
P. zelicaon	2-10	Dextran sulfate	20-60	20-60	~20	13
P. polyxenes	4-20	Dextran sulfate	20-100	?	20-30	0
P. glaucus	4-20	Dextran sulfate	20-200	?	~30	0
P. troilus	4-10	Dextran sulfate	20-100	?	~10	0
P. cresphontes	4-10	Dextran sulfate	20-100	?	~20	0

Table S3-3. List of oligos used to amplify and sgRNA target sequences.

Oligo/sgRNA name	sequence	Function	gene
Papilio_WntApr1_FWD	GATGGTGGCCCATGGTAGG	PCR oligo	WntA
Papilio_WntApr1_REV	CCACCAATTACCGGTGACAGC	PCR oligo	WntA
Papilio_WntA_Fwd_ISH	GCACTGGCAATGGGGAGG	PCR oligo	WntA
Papilio_WntA_REV_ISH	GCAATGTACTCGACAGCACC	PCR oligo	WntA
Papilio_wnt6_Ex3_FWD	GCTGGGGTGACATACGCCAT	PCR oligo	Wnt6
Papilio_wnt6_Ex6_REV	GGTCACAACTATCCGTTCCGGC	PCR oligo	Wnt6
NP_PmWtnA_FWD	GTGCAAATCGCGGTGCAAATC	PCR oligo	WntA
NP_PmWtnA_REV	GCACGTCCAGGAACTGTAAACC	PCR oligo	WntA
NP_PmWtn6_FWD	CGAGGCGTACCAATGACCAC	PCR oligo	Wnt6
NP_PmWtn6_REV	ATCGAACCCAGGAAGAAGCC	PCR oligo	Wnt6
PzPp_WntApr_sgRNA1	AAGGGGAAGGGTCGTACTGG	sgRNA	WntA
PzPp_WntApr_sgRNA2	ATTCGAAAGAGACTTTGGAA	sgRNA	WntA
PzPp_WntApr_sgRNA3	AAGTGATTTAAATACACAAT	sgRNA	WntA
PzPp_WntAE1_sgRNA4	AACTTTCCAAAATAGCTTCA	sgRNA	WntA
Pz_Wnt6_sgRNA1	CATCACACGAGCCTGCACTG	sgRNA	Wnt6
Pz_Wnt6_sgRNA2	CATCACACGAGCCTGCACTG	sgRNA	Wnt6

Table S3-4. Summary of CRISPR/Cas9 injection experiments.

Species	Gene	Time of egg collection (hours)	eggs injected	Number of L1 larvae	Hatching Ratio	Number of pupae	Number of adults with phenotype	Phenotype ratio from hatched eggs
P. zelicaon	*WntA*	1-4	~400	60	15%	12	5	8%
P. polyxenes	*WntA*	1-4	~520	81	16%	19	5	14%
P. zelicaon	*Wnt6*	1-4	~190	35	18%	13	7	20%
P. polyxenes	*Wnt6*	1-4	~350	25	7%	9	6	24%