Version of Record: https://www.sciencedirect.com/science/article/pii/S0959437X21000198
Manuscript_fff8f38d2882acb938c2a5ec0438249c

Heliconius butterflies: A window into the evolution and development of diversity.

Steven M. Van Belleghem[1*], James J. Lewis[2,3*], Edgardo Santiago Rivera[1,4] and Riccardo Papa[1,5].

Addresses

[1] Department of Biology, University of Puerto Rico–Rio Piedras, San Juan, Puerto Rico.
[2] Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA.
[3] Baker Institute for Animal Health, Cornell University, Ithaca, NY, USA.
[4] Chairs of Biomaterials, University of Bayreuth, Bayreuth, Bayern, Germany.
[5] Molecular Sciences and Research Center, University of Puerto Rico, San Juan, Puerto Rico.

[*]Contributed equally

Orcid IDs

Steven M. Van Belleghem: 0000-0001-9399-1007
James J. Lewis: 0000-0002-4170-9118
Edgardo Santiago Rivera: 0000-0001-6890-109X
Riccardo Papa: 0000-0002-7986-9993

Abstract

Butterflies have become prominent models for studying the evolution and development of phenotypic variation. In *Heliconius,* extraordinary within species divergence and between species convergence in wing color patterns has driven decades of comparative genetic studies. However, connecting genetic patterns of diversification to the molecular mechanisms of adaptation has remained elusive. Recent studies are bridging this gap between genome and function and have driven substantial advances in deciphering the genetic architecture of diversification in *Heliconius*. While only a handful of large-effect genes were initially identified in the diversification of *Heliconius* color patterns, recent experiments have begun to unravel the underlying gene regulatory networks and how these have evolved. These results reveal an evolutionary story of many interacting loci and partly independent genetic architectures that underlie convergent evolution.

Keywords: *Heliconius*, convergence, functional genomics, epigenomics, CRISPR/Cas9, *cis*-regulatory elements

Introduction

Evolutionary radiations and bursts of animal diversification present clear opportunities to study the genetic, and developmental mechanisms that underlie the evolution of novel morphologies, population divergence, and speciation. *Heliconius* butterflies, a group of more than 40 species that rapidly diversified in the Neotropics within the last 12 million years, have become a model system for the study of adaptive diversification and morphological evolution (Figure 1A) [1,2]. *Heliconius* are well known for their aposematic wing color patterns that advertise their unpalatability to predators [2]. These distinct signals are under strong selection [3,4], and multispecies Müllerian mimicry rings, where separate lineages converge on the same phenotype to share the mortality imposed by the learning process of their bird predators [5], are common in *Heliconius*. A few large-effect loci and associated genes that control convergent wing phenotypes have been identified over the last decade (Figure 1B). This has prompted investigation into the role of introgression [6,7], structural variation [8,9], chromosomal rearrangements [10,11], and nucleotide polymorphisms [12–14] in adaptive trait evolution [15,16], as well as their role in reproductive isolation [17–19]. Despite these studies, the precise genetic mechanisms of wing color pattern determination have remained largely uncharacterized. Using advances in developmental genetics and functional genomics, current studies are starting to reveal the connection between color pattern genes and causal regulatory variants that underlie wing phenotypes. Thus, the application of developmental, functional, and comparative genomic approaches in this diverse group of butterflies has opened a window into the mechanisms of phenotypic diversity and begun to test previous hypotheses on the developmental basis of wing pattern evolution [20–22].

In this review, we present some of the most recent molecular and genomic discoveries in *Heliconius* wing color pattern diversification and convergence. We discuss these results in light of three emerging themes: (i.) the population and speciation genomic architecture, (ii.) the molecular mechanisms of wing pattern evolution, and (iii.) the predictability of morphological evolution. These areas of emphasis provide novel insights into the process of adaptive diversification but raise more questions than answers. We therefore conclude with a discussion of some future directions for *Heliconius* biology.

Genomic architecture of phenotypic diversity

High-quality genome assemblies and whole genome resequencing in *Heliconius* have begun to unlock the genomic architecture of phenotypic diversity within and between species. For example, adaptive introgression of wing color pattern alleles has been found to underlie convergence of color pattern phenotypes between many cases of co-mimetic *Heliconius* lineages [7,23–25]. Subsequent sequencing and *de novo* assembly of 20 Heliconiini genomes probed the evolutionary role of introgression and admixture during speciation and found widespread exchange of genetic content between lineages during this process [6]. This result, accompanied by studies of pairwise species comparisons [13,26,27], suggests that a significant period of genetic exchange may persist before the completion of speciation, often resulting in reticulate phylogenies rather than well-defined bifurcating trees (Figure 2A). From these features observed during speciation in *Heliconius* butterflies, we can extrapolate a broader hypothesis, which

proposes that the processes and evolutionary history of adaptive diversification may be disjoint from those that give rise to the average species phylogeny. This hypothesis is supported by similar results in birds [28], sticklebacks [29], and cichlid fishes [30] and highlights that frequent exchange of adaptive genetic content may be a common feature in natural biodiversity.

Within *Heliconius* species, population resequencing and Quantitative Trait Loci (QTL) studies have made remarkable progress in detecting the genetic architecture of geographic diversity in wing patterns. Independent studies of *H. erato* [24,31,32] and *H. melpomene* [23,25,33], an iconic example of phenotypic convergence in *Heliconius*, generated hypotheses on the genes and *cis*-regulatory architecture defining wing color patterns. Segregating wing pattern phenotypes in these species primarily map to the signaling ligand *WntA,* involved in black pattern formation [32,34,35], *cortex,* which associates with a variety of pattern elements [33,36,37], and the transcription factor *optix*, involved in red/orange pattern formation [31,38] (Figure 1B). Analyses of phenotypic variation in *Heliconius* crosses suggested that wing scale structure underlying iridescence is a quantitative trait controlled by multiple genes [39,40], and knockouts in another species (*Junonia coenia*) implicated *optix* [38]. While a few other *Heliconius* phenotypes have been associated with *vvl* (apical portion of the medial forewing band in *H. erato*), *aristaless1* (yellow/white switch in the *H. cydno*), we mainly focus here on the role of *WntA*, *cortex* and *optix* in wing color pattern diversification. A detailed review of the developmental function of color patterns genes in *Heliconius* is given by [41].

Analyses of introgression, hybrid zone divergence and selective sweeps, give a comprehensive picture of the genomic variation associated with wing pattern diversity at *WntA*, *cortex* and *optix* (Figure 2B). These studies identified a few non-coding regions around these genes predicted to control wing pattern diversity across the *H. erato* and *H. melpomene* radiations [4,23–25]. Sequence phylogenies of the non-coding genomic regions suggested that wing color pattern associated genomic intervals are shared between populations and could be recombined to produce new functional combinations. This shuffling and sharing of haplotypes between distinct species and populations indicated that modular *cis*-regulatory element (CRE) architectures around the major color pattern genes could be a flexible mechanism for rapid morphological diversification (Figure 2C) [23,24]. Modular *cis*-regulatory architectures may also be present in genes regulating traits related to perception and behavior in *Heliconius* [17,18]. Confirmation of modular CREs for wing patterning and behavior would thus indicate that combinatorial evolution of modular elements is an important driver of diversity within *Heliconius.* While wing pattern studies suggest a genetic architecture composed of distinct CRE modules associated with discrete pattern elements, it also shows the limits of population genomics to identify the causative mutations. Linkage disequilibrium between the color pattern genes and the neighboring non-coding regions limited the ability to identify single CREs, though advances in linked-read sequencing may provide improved resolution for this challenge [42].

Molecular and regulatory mechanisms of diversification

While comparative population genomics has demonstrated the importance of introgression in *Heliconius* diversification, connecting the genomic signatures of wing pattern variability to the molecular mechanisms (e.g., *cis*-regulatory sequences and *trans*-regulatory factors) of adaptation

has been challenging. Several recent studies have used functional genomic approaches to elucidate the mechanisms of diversification in *Heliconius*. Comparisons of gene expression [43–45] chromatin accessibility (assayed with ATAC-seq) and CRE activity divergence (assayed with ChIP-seq of histone modifications) [45] between wing pattern morphs show trends similar to genome sequence-based inferences. Population genomic comparisons suggest that a few loci strongly differentiate between hybridizing parapatric populations, while reproductively isolated sympatric species and allopatric populations display greater differentiation at many loci with highly elevated population-specific divergence [13,45]. Similarly, gene regulatory and expression divergence becomes increasingly common with geographical and phylogenetic distance between taxa. In contrast to nucleotide differentiation, however, changes in gene expression and CRE activity between populations of *H. erato* show that divergence of chromatin accessibility and CRE activity is more apparent after relatively little separation, and even parapatric populations display a moderate degree of population-specific gene expression and *cis*-regulatory activity. This putative functional divergence is associated with evidence of selection, and thus likely reflects population-specific adaptative evolution [45,46]. Additional functional validation (e.g., CRISPR/Cas9 deletion) of these population-associated *cis*-regulatory variants will be critical for deciphering the complex adaptive landscape of *Heliconius* diversification. Cumulatively, these studies show that intraspecific variation in CRE activity is common, associated with population structure and changes in gene expression, and likely an important force underlying local adaptation and genomic divergence.

Targeted investigation of the major color pattern genes and CREs that underlie variation in wing phenotype between morphs has been limited in *Heliconius*, though several studies of this nature are ongoing. Two studies of red wing color pattern variation generated interesting results that contrast with previous population genetic inferences. A first study used chromatin accessibility data (ATAC-seq) and chromatin interaction profiles (Hi-C) to identify and target five interdependent regulatory elements near *optix* necessary for red/orange color pattern development in *H. erato* (Figure 3A) [47]. In contrast with the predicted modular CRE architecture, these regulatory loci were pleiotropic and resulted in red scale mutants in both fore- and hindwings when excised using CRISPR/Cas9 (Figure 3A). Two CRE mutants also showed evidence of epistasis between wing pattern loci and indicate that repression may be an important mechanism for wing pattern development. Thus, functional genomic assays of *optix* CREs do not support the hypothesis of independent CRE modules underlying the development of distinct color pattern elements, as hypothesized from hybrid zone scans (e.g., a separate CRE module for proximal red on the forewing and red hindwing rays; Figure 2B).

Admixture and homogenization of the genetic background in hybrid zones likely explain the disparity between signals of a modular CRE architecture in population genomic inferences and pleiotropic effects in functional studies. Current evidence indicates that specific haplotypes at multiple color pattern genes, such as *WntA*, *cortex*, and *optix*, are necessary for proper red color pattern formation (Figure 3B). In hybrid zones, however, a few small genomic regions may be sufficient to disrupt CRE integrity at *optix*, and thus act as modules capable of swapping between phenotypes within a single hybrid zone. These "evolutionary modules" [48] are insufficient to explain the evolutionary origin and development of red wing phenotypes, but may well explain the separation of geographically adjacent morphs. What remains unidentified are the specific CRE mutations and *trans*-regulatory interactions that preserve the integrity of the interdependent CREs and function as tissue specific switches.

A second study that investigated the downstream targets of the *optix* protein further indicated the importance of the genetic background in the divergence of color patterns. Tests for selection and Optix protein ChIP-seq (which identified the genomic locations where *optix* binds to CREs of targeted genes) revealed that dozens of Optix-bound loci displayed strong signatures of selection (Figure 3C) [46]. These Optix-bound sequences also displayed significantly elevated signals of genomic differentiation between neighboring morphs and matched genes previously associated with coloration [46]. While these studies are limited to red wing patterns, we speculate that epistasis may be similarly inherent to the developmental genetics of other color pattern loci as well. Similarly, the complex adaptive and differentiation landscapes of red wing associated loci suggests that simple genomic islands of diversification models may fail to accurately capture the evolutionary history of mimicry phenotypes.

Predictability of evolution

Studies mapping the molecular basis of wing pattern variation in *Heliconius* found that homologous genomic loci control the independently evolved mimicry phenotypes in distantly related species [15,16]. Further work showed that the same genes underlie repeated evolution of wing patterns within and between species [49]. This result was seemingly confirmed by multiple studies of gene expression across the *Heliconius* genus [32,43,50]. While initial findings proposed predictable evolution of a conserved developmental ground plan for *Heliconius* wing patterns [15,16,22], recent studies of gene function using CRISPR/Cas9 have indicated that, so-called, gene regulatory networks have diverged before or during the convergent evolution of wing color patterns. Thus, exploration of the *trans*-regulatory landscape of wing pattern development in different species [44] and studies of gene function [34,35,38] are rewriting our earlier hypotheses for the predictability of evolution of wing color pattern diversity and convergence.

To date, the developmental functions of *optix* [38], *cortex* [36], *WntA* [34,35], and *aristaless1* [51] have been tested with CRISPR/Cas9, though only *optix* and *WntA* have data for multiple butterfly species. Mutation of *optix*, which is active later in pupal development, consistently presented the same loss of ommochrome (red or orange) patterning in all species [38]. Importantly, there was no obvious alteration of function between species, and thus *optix* appears to be a master regulator of red and orange scale development that processes inputs from earlier wing pattern genes. In contrast to the stability of *optix* function, the most striking finding from WntA-deficient wing phenotypes was the non-homologous modification of color pattern observed between co-mimics (Figure 4A) [35]. The effect of *WntA* knockouts on convergent wing patterns in co-mimetic pairs was always different, where *H. erato* clade species consistently demonstrated greater alteration of the wing phenotype than observed in *H. melpomene.* These differences suggest a considerable divergence in gene regulatory networks between *H. erato* and *H. melpomene*, though whether this is merely a consequence of evolutionary time and phylogenetic separation or adaptive processes remains unclear. Nevertheless, the cryptic divergence in wing development between co-mimics highlights how adaptive processes can converge on the same phenotypes in different ways. Furthermore, *WntA* knockouts among *H. erato* morphs induced changes in the patterns of every major scale color

type, thus providing further evidence of widespread functional activity and epistatic interactions between the major color pattern genes (Figure 4B).

The degree of wing color pattern variation within *Heliconius* populations, and the magnitude of resemblance between convergent species, has mostly remained descriptive with little quantitative analysis of color pattern similarity. A recent study provided a quantitative assessment of pattern variation within and between species and revealed a greater amount of pattern dissimilarity than expected (Figure 4C) [52]. While the relative size of the forewing color patterns compared to the wing background is nearly identical between co-mimics, the placement of these elements is not perfectly matched between species. Combined quantitative analysis of natural variation and WntA-deficient wing phenotypes identified distinct boundaries of *WntA* expression in the wing that prevent *H. erato* and *H. melpomene* from achieving perfect mimicry. This mismatch is likely due to divergence in the underlying gene regulatory network that defines wing color patterns in both species, which places constraints on morphological evolution even between related species. QTL analysis of quantitative variation in red forewing color patterns further demonstrated that genetic parallels are indeed weak outside of the major color pattern genes, thus strengthening the hypothesis that the genetic basis of convergent traits may be more different than assumed from the major color pattern genes [53].

In sum, despite the appearance of similarity in wing pattern evolution, recent studies of wing pattern development across *Heliconius* indicate that the genetic and developmental mechanisms of evolution may be less predictable than once thought. Indeed, even when the same genomic loci are targets of selection, developmental constraints and strong selection pressures may drive rapid diversification via non-homologous mechanisms (Figure 4D). The extent to which this result in *Heliconius* is representative of other examples of phenotypic convergence is uncertain and further study of additional taxa will be needed to ascertain the generality of this observation.

Conclusion and future directions

The *Heliconius* radiation – with phenotypically segregated populations, frequent hybrids between closely related species, and multi-species mimicry rings with strong selection pressures – presents a unique opportunity to study the mechanisms of adaptive diversification along the speciation continuum. While a small set of key genes have been repeatedly associated with *Heliconius* wing color pattern evolution, recent work has demonstrated that parallels beyond these few genes may be rare and that the genetic architecture of adaptation may be more complex than expected. This is exemplified by a lack of genetic parallels for quantitative trait variation [53], divergence in gene regulatory landscapes [44], different phenotypic effects of gene knockouts in divergent lineages [35], and a demonstration of polygenic signals of adaptation during the evolution of color patterns [46]. These studies suggest that a combination of shared and novel genomic mechanisms have contributed to the diversity of phenotypes showcased in the *Heliconius* radiation. Although these studies were limited in scope to *Heliconius*, we can speculate that the process of diversification may often rely on both repeated change at a few key "hotspots" of evolution (e.g., *MC1R* and *Agouti* in vertebrates [54], and *WntA* [34,55,56] and *cortex* [57–59] in Lepidoptera) and singular modifications of multiple gene network components and phenotype specific *cis*-regulatory loci.

Three modes of developmental divergence may underlie the lack of fine-scale genetic parallels during phenotypic convergence beyond the repeated adaptation at several color pattern genes. First, RNA-seq studies suggest that upstream components of wing pattern gene networks may differ between taxa, and thus constrain the spatial and/or temporal expression of color pattern genes [44]. Second, differences in the *cis*-regulatory landscape between populations and species at known color pattern genes may require evolution of novel *cis*-regulatory architectures for convergent traits. Though some degree of parallel evolution may occur at CREs shared between species [47], this could prove to be infrequent. Finally, color pattern gene networks can be complex [46], and evolution at multiple genes in a developmental pathway may facilitate similar phenotypic change. How each of these possibilities contribute to genetic parallelism, or the lack thereof, in phenotypic convergence will be a major focus of ongoing research in *Heliconius*. We expect that use of single cell transcriptomics and functional genomics will help address these standing questions in the coming years. Though difficult, elucidating these developmental and genetic mechanisms in *Heliconius* will add to a conceptual foundation for a comprehensive model of the mechanisms and processes that guide diversification.

Acknowledgments

R.P. and S.M.V.B were funded by NSF EPSCoR RII Track-2 FEC (grant no. OIA 1736026) and NSF IOS 1656389. R.P. was also supported by the Fondo Institucional para la Investigación (FIPI), Universidad de Puerto Rico - Recinto de Río Piedras, Decanato de Estudios Graduados e Investigación. S.M.V.B. was also supported by a Puerto Rico Science, Technology & Research Trust catalyzer award (#2020-00142) and in part by National Institutes of Health-NIGMS COBRE Phase 2 Award – Center for Neuroplasticity at the University of Puerto Rico (Grant No. 1P20GM103642). E.S.R. was supported by a RISE fellowship (5R25GM061151-18).

References

1. Merrill RM, Dasmahapatra KK, Davey JW, Dell'Aglio DD, Hanly JJ, Huber B, Jiggins CD, Joron M, Kozak KM, Llaurens V, et al.: **The diversification of *Heliconius* butterflies: What have we learned in 150 years?** *J Evol Biol* 2015, **28**:1417–1438.

2. Jiggins CD: *The ecology and evolution of* Heliconius *butterflies*. Oxford University Press; 2017.

3. Mallet J: **Speciation, raciation, and color pattern evolution in *Heliconius* butterflies: evidence from hybrid zones.** In *Hybrid zone and the Evolutionary process*. Edited by Harrison RG. Oxford University Press; 1993:226–260.

4. Moest M, Van Belleghem SM, James J, Salazar C, Martin S, Barker S, Moreira G, Mérot C, Joron M, Nadeau N, et al.: **Selective sweeps on novel and introgressed variation shape mimicry loci in a butterfly adaptive radiation**. *PLoS Biol* 2020, **18**:e3000597.

5. Mallet J, Joron M: **Evolution of diversity in warning color and mimicry:**

Polymorphisms, shifting balance and speciation. *Annu Rev Ecol Syst* 1999, **30**:201–233.

6. Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow R, García-accinelli G, Van Belleghem SM, Patterson N, Daniel E, et al.: **Genomic architecture and introgression shape a butterfly radiation**. *Science* 2019, **366**:24174–24183.

7. Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Adler S, Ahn S-J, Baker DA, et al.: **Butterfly genome reveals promiscuous exchange of mimicry adaptations among species.** *Nature* 2012, **487**:94–98.

8. Davey JW, Barker SL, Rastas PM, Pinharanda A, Martin SH, Durbin R, McMillan WO, Merrill RM, Jiggins CD: **No evidence for maintenance of a sympatric *Heliconius* species barrier by chromosomal inversions**. *Evol Lett* 2017, **1**:138–154.

9. Pinharanda A, Martin SH, Barker SL, Davey JW, Jiggins CD: **The comparative landscape of duplications in *Heliconius melpomene* and *Heliconius cydno***. *Heredity* 2017, **118**:78–87.

10. Davey JW, Chouteau M, Barker SL, Maroja L, Baxter SW, Simpson F, Merrill RM, Joron M, Mallet J, Dasmahapatra KK, et al.: **Major improvements to the *Heliconius melpomene* genome assembly used to confirm 10 chromosome fusion events in 6 million years of butterfly evolution**. *G3 Genes|Genomes|Genetics* 2016, **6**:695–708.

11. Jay P, Whibley A, Frézal L, Rodríguez de Cara MÁ, Nowell RW, Mallet J, Dasmahapatra KK, Joron M: **Supergene evolution triggered by the introgression of a chromosomal inversion**. *Curr Biol* 2018, **28**:1839–1845.

12. Van Belleghem SM, Salazar C, Jiggins CD, Baquero M, Papa R, Counterman BA, Mcmillan WO, Martin SH: **Patterns of Z chromosome divergence among *Heliconius* species highlight the importance of historical demography**. *Mol Ecol* 2018, **27**:3852–3872.

13. Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD: **Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies.** *Genome Res* 2013, **23**:1817–1828.

14. Martin SH, Möst M, Palmer WJ, Salazar C, McMillan WO, Jiggins FM, Jiggins CD: **Natural selection and genetic diversity in the butterfly Heliconius melpomene**. *Genetics* 2016, **203**:525–541.

15. Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, Ffrench-Constant RH, et al.: **Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*.** *PLoS Genet* 2010, **6**:e1000796.

16. Baxter SW, Nadeau NJ, Maroja LS, Wilkinson P, Counterman B a, Dawson A, Beltran M, Perez-Espona S, Chamberlain N, Ferguson L, et al.: **Genomic hotspots for adaptation:**

the population genetics of Müllerian mimicry in the *Heliconius melpomene* clade. *PLoS Genet* 2010, **6**:e1000794.

17. van Schooten B, Meléndez-Rosa J, van Belleghem SM, Jiggins CD, Tan JD, Owen McMillan W, Papa R: **Divergence of chemosensing during the early stages of speciation**. *Proc Natl Acad Sci U S A* 2020, **117**:16438–16447.

18. Rossi M, Hausmann AE, Thurman TJ, Montgomery SH, Papa R, Jiggins CD, McMillan WO, Merrill RM: **Visual mate preference evolution during butterfly speciation is linked to neural processing genes**. *Nat Commun* 2020, **11**.

19. Merrill RM, Rastas P, Martin SH, Melo MC, Barker S, Davey J, McMillan WO, Jiggins CD: **Genetic dissection of assortative mating behavior**. *PLoS Biol* 2019, **17**:e2005902.

20. Gilbert LE: **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*. Edited by Boggs CL, Watt WB, Ehrlich PR. University of Chicago Press; 2003:1–34.

21. Sheppard PM, Turner JRG, Brown KS, Benson WW, Singer MC: **Genetics and the evolution of Muellerian mimicry in *Heliconius* Butterflies**. *Philos Trans R Soc B Biol Sci* 1985, **308**:433–610.

22. Nijhout HF: *The development and evolution of butterfly wing patterns*. Smithsonian Institution Press; 1991.

23. Wallbank RWR, Baxter SW, Pardo-diaz C, Hanly J, Martin SH, Mallet J, Dasmahapatra KK: **Evolutionary novelty in a butterfly wing pattern through enhancer shuffling**. *PLoS Biol* 2016, **14**:e1002353.

24. Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J, Lewis JJ, Hines HM, et al.: **Complex modular architecture around a simple toolkit of wing pattern genes**. *Nat Ecol Evol* 2017, **1**:0052.

25. Morris J, Hanly J, Martin S, Van Belleghem S, Salazar C, Jiggins CD, Dasmahapatra KK: **Deep convergence, shared ancestry, and evolutionary novelty in the genetic architecture of *Heliconius* mimicry**. *Genetics* 2020, **216**:765–780.

26. Martin SH, Davey JW, Salazar C, Jiggins CD: **Recombination rate variation shapes barriers to introgression across butterfly genomes**. *PLOS Biol* 2019, **17**:e2006288.

27. Van Belleghem S, Cole J, Montejo-Kovacevich G, Bacquet C, McMillan WO, Papa R, Counterman B: **Selection and gene flow define polygenic barriers between incipient butterfly species**. *bioRxiv* 2020,

28. Toews DPL, Taylor SA, Vallender R, Brelsford A, Butcher BG, Messer PW, Lovette IJ: **Plumage genes and little else distinguish the genomes of hybridizing warblers**. *Curr Biol* 2016, **26**:2313–2318.

29. Schluter D, Conte GL: **Genetics and ecological speciation.** *Proc Natl Acad Sci U S A* 2009, **106**:9955–9962.

30. Meier JI, Marques DA, Mwaiko S, Wagner CE, Excoffier L, Seehausen O: **Ancient hybridization fuels rapid cichlid fish adaptive radiations**. *Nat Commun* 2017, **8**:14363.

31. Reed RD, Papa R, Martin A, Hinas HM, Counterman BA, Pard-Diaz C, Jiggins CD, Chamberlain NL, Kronforst MR, Chen R, et al.: ***optix* drives the repeated convergent evolution of butterfly wing pattern mimicry**. *Science* 2011, **333**:1137–1141.

32. Martin A, Papa R, Nadeau NJ, Hill RI, Counterman BA, Halder G, Jiggins CD, Kronforst MR, Long AD, McMillan WO, et al.: **Diversification of complex butter flywing patterns by repeated regulatory evolution of a *Wnt* ligand**. *Proc Natl Acad Sci U S A* 2012, **109**:12632–12637.

33. Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko S V., Wallbank RWR, Wu GC, Maroja L, Ferguson L, Hanly JJ, et al.: **The gene cortex controls mimicry and crypsis in butterflies and moths**. *Nature* 2016, **534**:106–110.

34. Mazo-vargas A, Concha C, Livraghi L, Massardo D, Wallbank RWR, Zhang L: **Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity**. *Proc Natl Acad Sci* 2017, **2**.

35. Concha C, Wallbank R, Hanly J, Fenner J, Livraghi L, Santiago E, Paulo D, Arias C, Vargas M, Sanjeev M, et al.: **Interplay between developmental flexibility and determinism in the evolution of mimetic Heliconius wing patterns**. *Curr Biol* 2019, **29**:3996–4009.

36. Livraghi L, Hanly J, Loh LS, Ren A, Warren I, Concha C, Wright C, Walker J, Foley J, Arenas-Castro H, et al.: **The gene *cortex* controls scale colour identity in Heliconius**. *bioRxiv* 2020,

37. Saenko S V., Chouteau M, Piron-Prunier F, Blugeon C, Joron M, Llaurens V: **Unravelling the genes forming the wing pattern supergene in the polymorphic butterfly *Heliconius numata***. *Evodevo* 2019, **10**:1–12.

38. Zhang L, Mazo-vargas A, Reed RD: **Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence**. *Proc Natl Acad Sci* 2017, **114**:10707–10712.

39. Brien MN, Enciso-Romero J, Parnell AJ, Salazar PA, Morochz C, Chalá D, Bainbridge HE, Zinn T, Curran E V., Nadeau NJ: **Phenotypic variation in *Heliconius erato* crosses shows that iridescent structural colour is sex-linked and controlled by multiple genes**. *Interface Focus* 2019, **9**.

40. Parnell AJ, Bradford JE, Curran E V., Washington AL, Adams G, Brien MN, Burg SL, Morochz C, Fairclough JPA, Vukusic P, et al.: **Wing scale ultrastructure underlying convergent and divergent iridescent colours in mimetic *Heliconius* butterflies**. *J R Soc*

Interface 2018, **15**.

41. McMillan WO, Livraghi L, Concha C, Hanly JJ: **From patterning genes to process: Unraveling the gene regulatory networks that pattern *Heliconius* wings**. *Front Ecol Evol* 2020, **8**:1–15.

42. Meier JI, Salazar PA, Ku M, Davies RW, Dréau A, Aldás I, Power OB, Nadeau NJ, Bridle JR, Rolian C, et al.: **Haplotype tagging reveals parallel formation of hybrid races in two butterfly species**. *bioRxiv* 2020,

43. Hines HM, Papa R, Ruiz M, Papanicolaou A, Wang C, Nijhout HF, McMillan WO, Reed RD: **Transcriptome analysis reveals novel patterning and pigmentation genes underlying Heliconius butterfly wing pattern variation**. *BMC Genomics* 2012, **13**:288.

44. Hanly JJ, Wallbank RWR, McMillan WO, Jiggins CD: **Conservation and flexibility in the gene regulatory landscape of heliconiine butterfly wings**. *Evodevo* 2019, **10**:1–14.

45. Lewis JJ, Reed RD: **Genome-wide regulatory adaptation shapes population-level genomic landscapes in *Heliconius***. *Mol Biol Evol* 2019, **36**:159–173.

46. Lewis JJ, Van Belleghem SM, Papa R, Danko CG, Reed RD: **Many functionally connected loci foster adaptive diversification along a neotropical hybrid zone**. *Sci Adv* 2020, **6**:1–11.

47. Lewis JJ, Geltman RC, Pollak PC, Rondem KE, Belleghem SM Van: **Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry**. *Proc Natl Acad Sci* 2019, **116**:24174–24183.

48. Lewis JJ, Van Belleghem SM: **Mechanisms of Change: A population-based perspective on the roles of modularity and pleiotropy in diversification**. *Front Ecol Evol* 2020, **8**:261.

49. Kronforst MR, Papa R: **The functional basis of wing patterning in *Heliconius* butterflies: The molecules behind mimicry**. *Genetics* 2015, **200**:1–19.

50. Martin A, McCulloch KJ, Patel NH, Briscoe AD, Gilbert LE, Reed RD: **Multiple recent co-options of *Optix* associated with novel traits in adaptive butterfly wing radiations**. *Evodevo* 2014, **5**:7.

51. Westerman EL, Vankuren NW, Massardo D, Buerkle N, Palmer SE, Kronforst MR: ***Aristaless* controls butterfly wing color variation used in mimicry and mate choice**. *Curr Biol* 2018, **28**:1–6.

52. Van Belleghem SM, Alicea Roman PA, Gutierrez HC, Counterman BA, Papa R: **Perfect mimicry between *Heliconius* butterflies is constrained by genetics and development**. *Proc R Soc B* 2020, **287**:20201267.

53. Bainbridge HE, Brien MN, Morochz C, Salazar PA, Rastas P, Nadeau NJ: **Limited**

genetic parallels underlie convergent evolution of quantitative pattern variation in mimetic butterflies. *J Evol Biol* 2020, **33**:1516–1529.

54. Hoekstra HE: **Genetics, development and evolution of adaptive pigmentation in vertebrates**. *Heredity* 2006, **97**:222–234.

55. VanKuren N, Massardo D, Nallu S, Kronforst MR: **Butterfly mimicry polymorphisms highlight phylogenetic limits of gene reuse in the evolution of diverse adaptations**. *Mol Biol Evol* 2019, **36**:2842–2853.

56. Martin A, Papa R, Nadeau NJ, Hill RI, Counterman BA, Halder G, Jiggins CD, Kronforst MR, Long AD, McMillan WO, et al.: **Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand**. *Proc Natl Acad Sci* 2012, **109**:12632–12637.

57. Nadeau N, Pardo-Diaz C, Whibley A, Supple M, Saenko SV, Wallbank RWR, Wu GC, Maroja L, Ferguson L, Hanly JJ, et al.: **The gene *cortex* controls mimicry and crypsis in butterflies and moths**. *Nature* 2016, **534**:106–110.

58. Van't Hof AE: **The industrial melanism mutation in British peppered moths is a transposable element**. *Nature* 2016, **534**:102–105.

59. van der Burg KRL, Lewis JJ, Brack BJ, Fandino RA, Mazo-Vargas A, Reed RD: **Genomic architecture of a genetically assimilated seasonal color pattern**. *Science* 2020, **370**:721–725.

Reference highlights (interest*, outstanding interest)**

*[4] **Moest et al. 2020** – This study investigated genomic patterns of recent adaptive evolution in a set of 600 butterflies from 53 populations from South America. Dynamic and recent signals of natural selection at butterfly color pattern genes were identified, despite theoretically expected stabilizing selective forces.

*[6] **Edelman et al. 2019** – Using recombination rate analysis and a new method to separate genomic signals of incomplete lineage sorting from introgression, this study provides a rigorous analysis of the extent of introgression in *Heliconius* butterflies.

[35] **Concha et al. 2019 – This study knocks out the gene *WntA* in a variety of *Heliconius* butterflies, demonstrating that this gene's phenotypic effect differs between lineages.

*[36] **Livraghi et al. 2020 –** This study demonstrates that *cortex* played a key role in the diversification of lepidopteran wing patterns in part due to its switch-like effects in scale identity across the entire wing surface. This is in contrast with other known *Heliconius* mimicry loci that act in specific patterns.

*[42] **Meier et al. 2020** – This study describes a new linked-read sequencing technique that enables whole genome haplotyping in large populations and uses it to study parallel divergence in parallel hybrid zones of co-mimicking *Heliconius* butterflies.

*[44] **Hanly et al. 2019** – By investigating the spatial transcriptomic landscape across *Heliconius* wings, this study provides a first attempt to identity the factors that regulate color pattern switch genes in *Heliconius*.

[47] **Lewis et al. 2019 – First demonstration of functional effect of *cis*-regulatory elements near the gene *optix*. Targeted *cis*-regulatory elements effect multiple red color pattern elements, demonstrating interdependence and pleiotropy of *cis*-regulatory elements for red color pattern development.

[46] **Lewis et al. 2020 – This study combined (i.) chromatin immunoprecipitation, (ii.) ATAC-seq, (iii.) Hi-C, (iv.) RNA sequencing, and (v.) whole-genome sequencing to show that evolution of a red color pattern in *Heliconius* is accompanied by adaptive evolution at many pigmentation genes bound by the transcription factor Optix. By integrating these methods, functional connections between loci and signals of polygenic evolution are unraveled.

*[52] **Van Belleghem et al. 2020** – Using quantitative analysis of color patterns and CRISPR/Cas9 mutants from Concha et al. 2019 [35], this study identifies parts in the color pattern of co-mimicking butterflies that may be constrained to evolve perfect mimicry due to divergence in the phenotypic effect of *WntA*.

*[53] **Bainbridge et al. 2020** – This study uses Quantitative Trait Locus (QTL) analysis together with a quantitative phenotyping approach to map variation in red pattern elements across forewings of *Heliconius erato* and *Heliconius melpomene* and demonstrated a remarkably low level of genetic parallelism between these co-mimicking species, apart from the red patterning locus including *optix*.

Figures

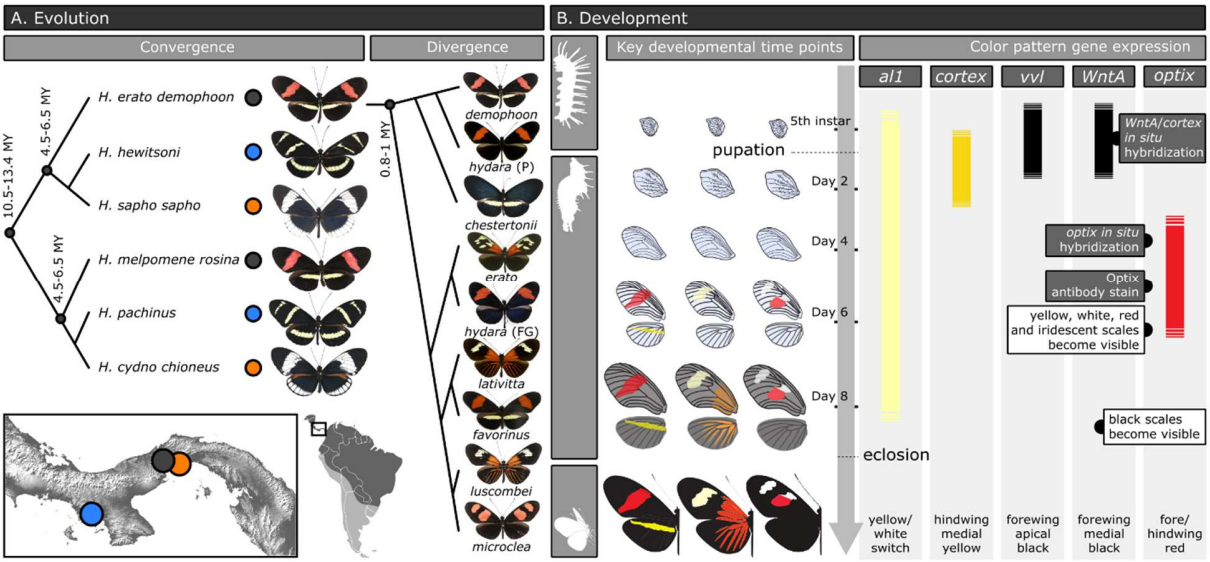

Figure 1. Diversity and development of *Heliconius* butterfly wing coloration. (A.) Example of convergence and divergence in the *Heliconius* genus. Left, mimicry between two clades that have diverged over 10 million years ago and converged on the same color pattern (Modified from [35]). Right, color pattern diversity within the *Heliconius erato* species (tree and time estimates from [27]). Dark gray shading on the map of South America indicates the distribution area of *H. erato*. Colored circles indicate localities in Panama where mimicking species considered in this review occur. **(B.)** Within the *Heliconius* species, four genes, including *aristaless1*, *cortex*, *WntA* and *optix*, have been functionally linked to color pattern development, with *vvl* being a strong candidate for an effect in the apical part of the medial forewing pattern. The genes *cortex*, *vvl* and *WntA* are known to be expressed in the late larval instar stage and early pupal stages [33,43,56]. The gene *optix* is expressed later in development and its expression pattern is dependent on the color patterns genes expressed earlier in development [31,34]. The gene *aristaless1* (*al1*) is expressed during 5th instar larvae and pupal development in white *H. cydno* [51].

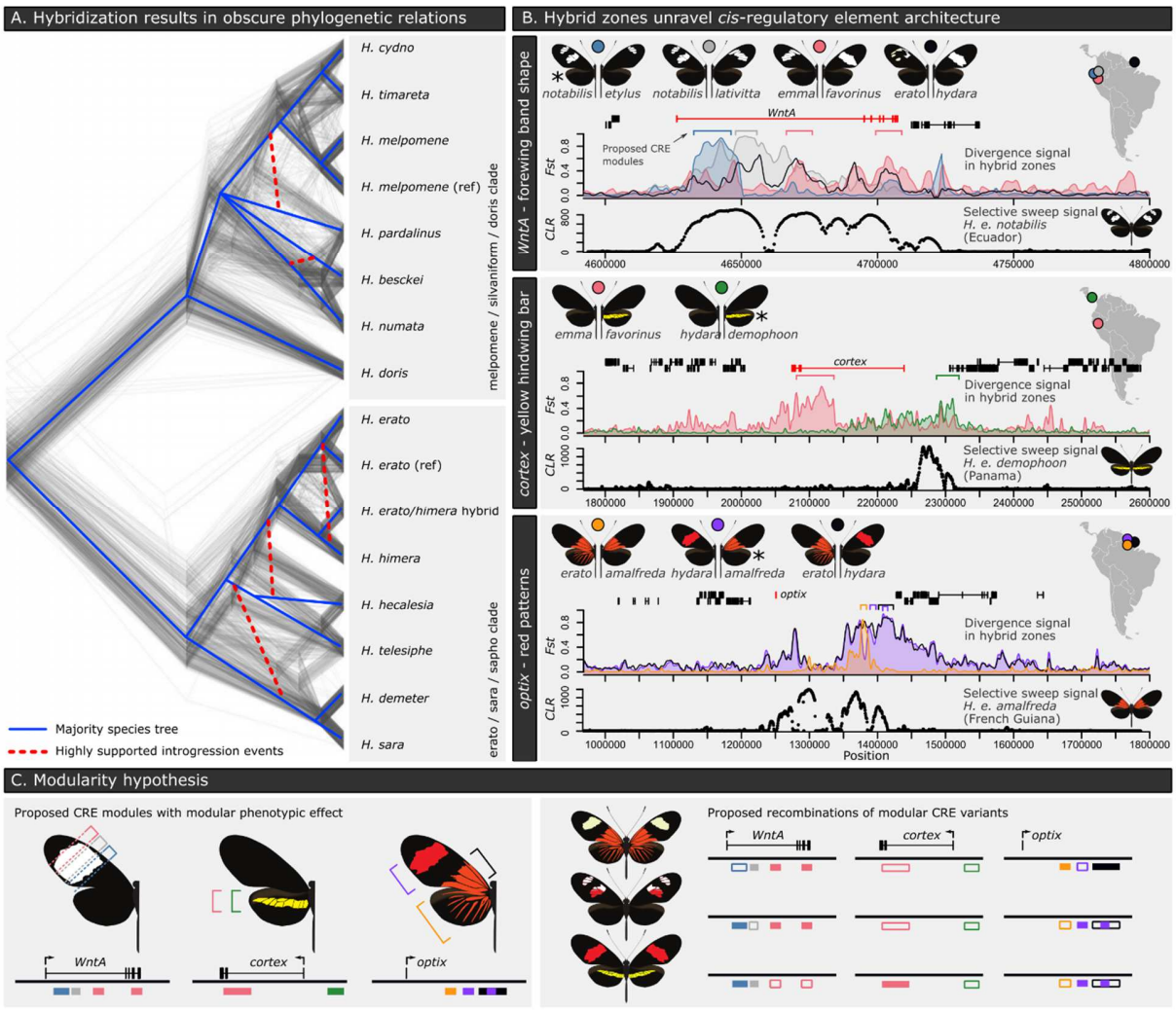

Figure 2. Introgression, divergence and selection in *Heliconius* species and populations. (A.) Densitree calculated from 10 kb windows showing the reticulate (non-bifurcating) phylogenetic relations within two *Heliconius* clades. Blue solid lines indicate the tree supported by the majority of 10 kb windows. Red dashed lines indicate introgression events. Tree obtained from [6], supplemental Figure S21. **(B.)** Hybrid zone divergence (F_{ST}) and selective sweep signals (Sweepfinder2 Composite Likelihood Ratio (CLR) statistic) at the *WntA* (top), *cortex* (middle), and *optix* (bottom) gene. Butterfly cartoons highlight phenotypic differences that have been associated with these genes. Localities of hybrid zones in Central and South America are indicated as colored circles. Asterisks indicate populations for which selective sweep support is shown. Colored brackets above F_{ST} signals indicate proposed CRE modules. F_{ST} data obtained from [24], selective sweep data obtained from [4]. **(C.)** A hypothesis of *cis*-regulatory element modularity was proposed from these data as populations or species that share different combinations of pattern elements (brackets) often also share haplotypes at genomic intervals in different combinations (filled and empty rectangles indicate modular CRE variants).

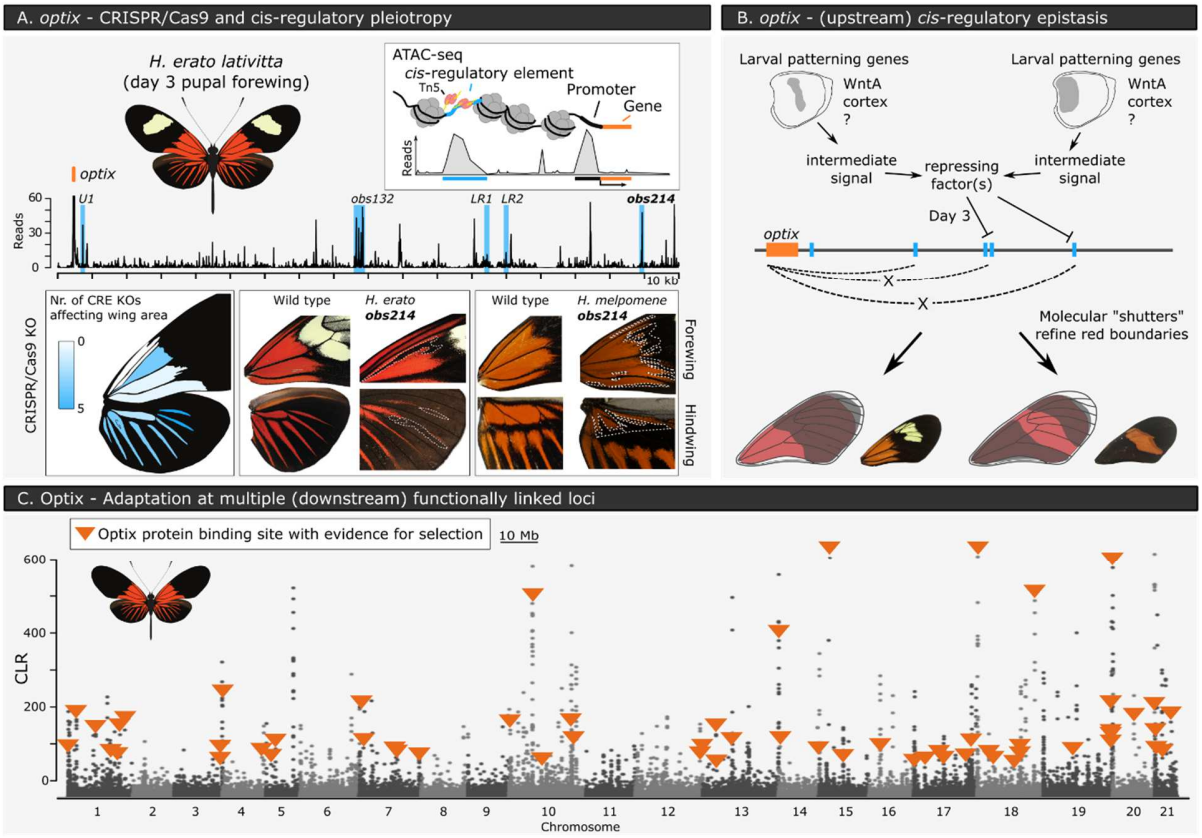

Figure 3. Functional genomics on the color pattern gene *optix* (A.) ATAC-seq profile from a 3-day old pupal forewing in *H. e. lativitta* and CRISPR/Cas9 mutants of the *obs214* CRE in *H. e. lativitta* and *H. m. aglaope*. CREs (shaded in blue) were identified based on differential accessibility between morphs, wing tissue and functional effect after excision. Wing diagram on the left shows number of identified CREs that affected red color pattern elements. Note that four out of five CREs affected both hindwing and forewing pattern. Data and images obtained from [47]. **(B.)** Expression of *optix* is dependent on patterning genes. Both divergence in these patterning genes and the *cis*-regulatory landscape near *optix* may result in morphological diversification. Diagram modified from [47]. **(C.)** ChIP-seq of the Optix protein identified many loci with evidence for selection (Sweepfinder2 Composite Likelihood Ratio (CLR) statistic) that are bound by Optix (orange triangles). Data obtained from [46].

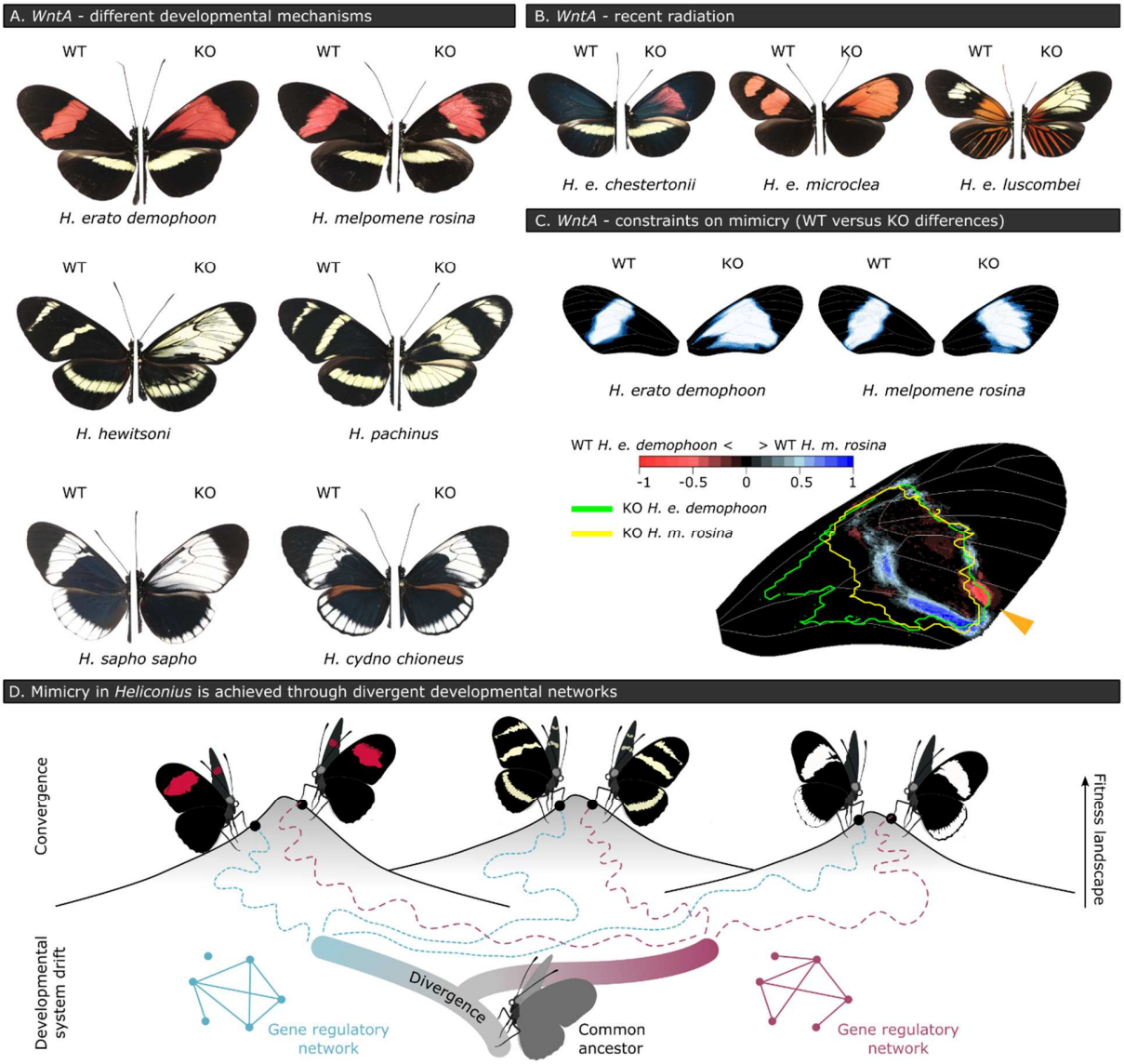

Figure 4. CRISPR/Cas9 on the color pattern gene *WntA*. (A.) *WntA* Wildtype (WT) and CRISPR/Cas9 (KO) phenotypes on co-mimicking butterflies from two clades (left, erato/sara/sapho clade; right, melpomene/silvaniform/doris clade) that diverged over 10 million years ago. For phylogenetic relations and localities see Figure 1A. Images obtained from [35]. **(B.)** WT and KO phenotypes on geographic color pattern morphs of *H. erato*. Images obtained from [35]. **(C.)** Quantitative comparison of WT and KO variation between *H. erato* and *H. melpomene* from Panama with a red forewing pattern. Heatmaps at the top demonstrate the consistency of the forewing pattern in WT and KO samples with white indicating consistent presence of forewing patterns and blue gradient indicating less consistent presence. Bottom shows inter-species differences in forewing pattern with red indicating higher presence of forewing pattern in *H. erato* and blue indicating higher presence in *H. melpomene*. The yellow and green outlines show the *WntA* KO area. Orange triangle indicates overlap in mismatch between WT co-mimicking populations and respective *WntA* KO patterning area. Images

obtained from [52]. **(D.)** Diagram of divergence in the gene regulatory network with which *WntA* interacts, and subsequent selection for co-mimicking phenotypes (convergence). Tips of mountains represent peaks in the fitness landscape which are hypothesized to have been reached by independently rewiring the gene regulatory network underlying the development of color pattern traits. Diagram modified from [35].