Cortex cis-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*

Luca Livraghi[1,2], Joseph J. Hanly[1,2,3], Steven M. Van Belleghem[4], Gabriela Montejo-Kovacevich[1], Eva S. M. van der Heijden[1,2], Ling Sheng Loh[3], Anna Ren[3] Ian A. Warren[1], James J. Lewis[5], Carolina Concha[2], Laura H. López[1,2], Charlotte Wright[1], Jonah M. Walker[1], Jessica Foley[2], Zachary H. Goldberg[6], Henry Arenas-Castro[2], Michael W. Perry[6], Riccardo Papa[4], Arnaud Martin[3], W. Owen McMillan[2] and Chris D. Jiggins[1,2]

Author affiliations:

1. Department of Zoology, University of Cambridge, Downing St., Cambridge, CB2 3EJ, UK
2. Smithsonian Tropical Research Institute, Gamboa, Panama
3. The George Washington University Department of Biological Sciences, Science and Engineering Hall 6000, 800 22nd St NW Washington, DC 20052, USA
4. Department of Biology, Centre for Applied Tropical Ecology and Conservation, University of Puerto Rico, Rio Piedras, Puerto Rico
5. Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853
6. Cell & Developmental Biology, Division of Biological Sciences, UC San Diego, La Jolla, CA

Corresponding author: Luca Livraghi, University of Cambridge, Cambridge, UK. Dept. of Zoology. ll566@cam.ac.uk.

Keywords:

Evolution, wing patterning, *cortex*, *Heliconius*, cell fate, CRISPR, *cis*-regulation, *ATAC-seq*.

Abstract

In *Heliconius* butterflies, wing pattern diversity is controlled by a few genes of large effect that regulate colour pattern switches between morphs and species across a large mimetic radiation. One of these genes, *cortex,* has been repeatedly associated with colour pattern evolution in butterflies. Here we carried out CRISPR knock-outs in multiple *Heliconius* species and show that *cortex* is a major determinant of scale cell identity. Chromatin accessibility profiling and introgression scans identified *cis*-regulatory regions associated with discrete phenotypic switches. CRISPR perturbation of these regions in black hindwing genotypes recreated a yellow bar, revealing their spatially limited activity. In the *H. melpomene/timareta* lineage, the candidate CRE from yellow-barred phenotype morphs is interrupted by a transposable element, suggesting that *cis*-regulatory structural variation underlies these mimetic adaptations. Our work shows that *cortex* functionally controls scale colour fate and that its *cis*-regulatory regions control a phenotypic switch in a modular and pattern-specific fashion.

Introduction

Butterfly wing pattern diversity provides a window into the ways genetic changes underlie phenotypic variation that is spatially limited to specific parts or regions of the organism (McMillan et al., 2020; Orteu and Jiggins, 2020; Rebeiz et al., 2015). Many of the underlying genetic loci controlling differences in colour patterns have been mapped to homologus "hotspots" across disparate taxa. In some cases, this repeated adaptation has occurred through the alteration of downstream effector genes, such as pigment biosynthetic enzymes with functions clearly related to the trait under selection, for example, the genes *tan* and *ebony* that control insect melanin pigmentation (reviewed in Massey and Wittkopp, 2016). In other cases, upstream regulatory genes are important, and these are typically either transcription factors (e.g. *optix, MITF, Sox10*) or components of signalling pathways such as ligands or receptors (e.g. *WntA, MC1R, Agouti*). These 'developmental toolkit genes' influence pigment cell fate decisions by modulating gene regulatory networks (GRNs) (Kronforst and Papa, 2015; Martin and Courtier-Orgogozo, 2017; Prud'homme et al., 2007), and are commonly characterised by highly conserved functions, with rapid evolutionary change occurring through regulatory fine-tuning of expression patterns. One gene that has been repeatedly implicated in morphological evolution but does not conform to this paradigm is *cortex*, a gene implicated by mapping approaches in the regulation of adaptive changes in the wing patterning of butterflies and moths.

Cortex is one of four major effect genes that act as switch loci controlling both scale structure and colour patterns in *Heliconius* butterflies, and has been repeatedly targeted by natural selection to drive differences in pigmentation (Nadeau, 2016; Van Belleghem et al., 2017). Three of the four major effect genes correspond to the prevailing paradigm of highly conserved patterning genes; the signalling ligand *WntA* (Martin et al., 2012; Mazo-Vargas et al., 2017) and two transcription factors *optix* (Lewis et al., 2019; Reed et al., 2011; Zhang et al., 2017) and *aristaless1* (Westerman et al., 2018). *Cortex,* on the other hand, is an insect-specific gene showing closest homology to the *cdc20/fizzy* family of cell cycle regulators (Chu et al., 2001; Nadeau et al., 2016; Pesin and Orr-Weaver, 2007). The lepidopteran orthologue of *cortex* displays rapid sequence evolution, and has acquired novel expression domains that correlate with melanic wing patterns in *Heliconius* (Nadeau et al., 2016; Saenko et al, 2019). It therefore seems likely that the role of *cortex* in regulating wing patterns has involved a major shift in function, which sits in contrast to the classic model of regulatory co-option of deeply conserved patterning genes.

The genetic locus containing *cortex* was originally identified in the genus *Heliconius* as controlling differences in yellow and white wing patterns in *H. melopmene* and *H. erato* (Figure 1, **a**) and the polymorphism in yellow, white, black and orange elements in *H. numata*. This was inferred using a combination of association mapping and gene expression data (Joron et al., 2006; Nadeau et al., 2016). The same locus has also been repeatedly implicated in controlling colour pattern variation among divergent Lepidoptera, including the peppered moth *Biston betularia* and other geometrids, the silkmoth

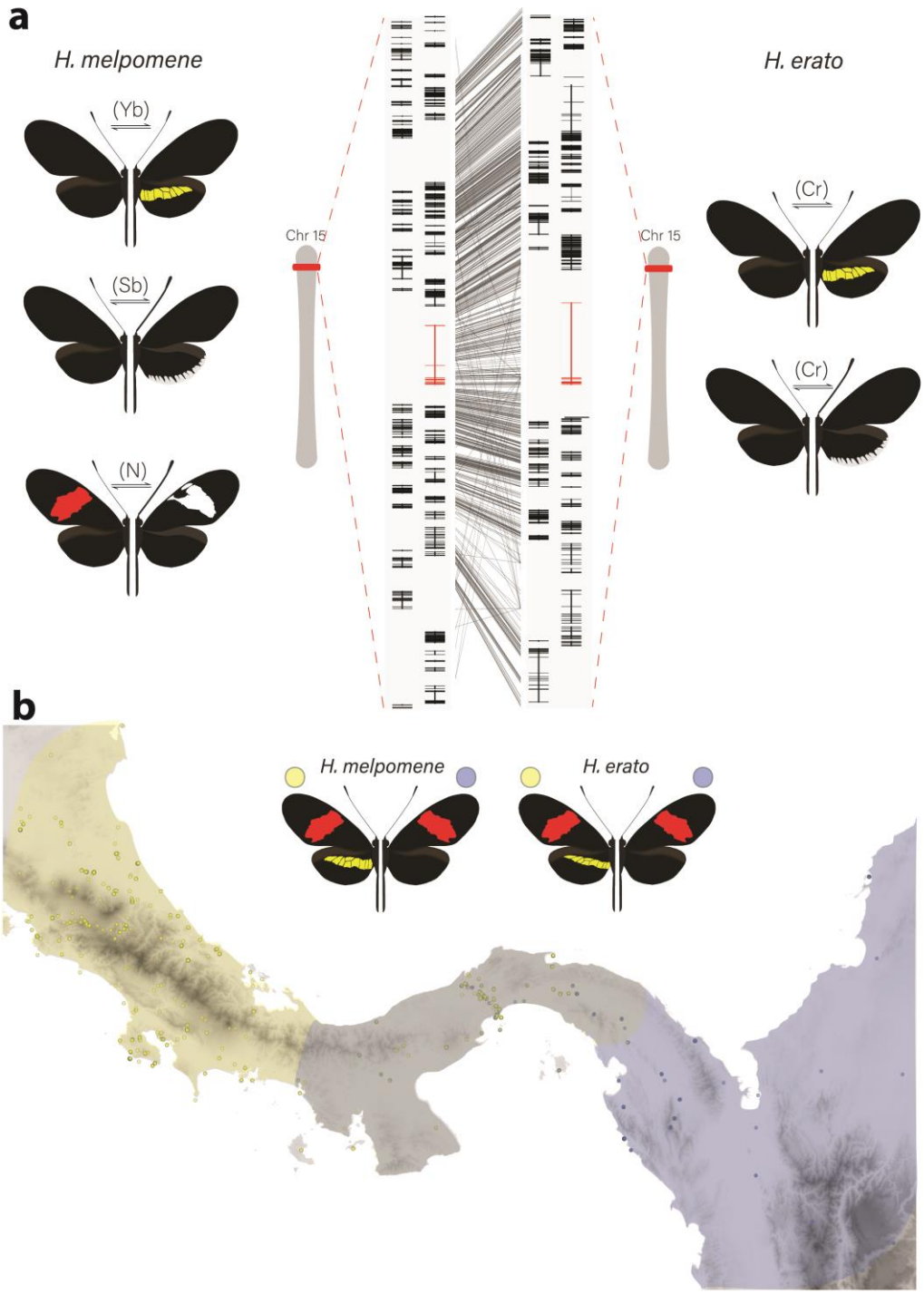
Bombyx mori and other butterflies such as *Junonia coenia*, *Bicyclus anynana* and *Papilio clytia* (Beldade et al., 2009; van der Burg et al., 2020; Ito et al., 2016; VanKuren et al., 2019; van't Hof et al., 2019; Van't Hof et al., 2016). This locus therefore contains one or more genes that have repeatedly been targeted throughout the evolutionary history of the Lepidoptera to generate phenotypic diversity.

In *Heliconius* butterflies, population genomic data suggest that *cis*-regulatory modules surrounding *cortex* underlie adaptive variation of yellow and white colour pattern elements (Enciso-Romero et al., 2017; Van Belleghem et al., 2017). These studies predict the existence of modular elements that compartmentalise expression of colour pattern genes across developing wings. However, developmental genes have complex regulatory domains and recent work has suggested that pleiotropy among different enhancers may be more common than is currently appreciated (Lewis et al., 2019; Murugesan et al., 2021). Further dissection of the regulatory elements controlling wing pattern variation is thus necessary to assess the relative contribution of pleiotropy versus modularity at colour pattern loci (Lewis and van Belleghem, 2020).

While fantastically diverse, most of the pattern variation in *Heliconius* is created by differences in the distribution of just three major scale cell types; Type I (yellow/white), Type II (black), and Type III (red/orange/brown) (Aymone et al., 2013; Gilbert et al., 1988). Each type has a characteristic nanostructure and a fixed complement of pigments. Type I yellow scales contain the ommochrome precursor 3-hydroxykynurenine (3-OHK) (Finkbeiner et al., 2017; Koch, 1993; Reed et al., 2008), whereas Type I white scales lack pigment, and the white colour is the result of the scale cell ultrastructure (i.e. structural white) (Gilbert et al., 1988). Structurally, Type I scales are characterised by the presence of a lamina covering the scale windows and by micro-ribs joining the larger longitudinal ridges. In contrast, Type II scale cells are pigmented with melanin, have larger crossribs and lack a lamina covering the scale windows. Quantitative variation in scale structures between populations (but not within individuals) can cause Type II scales to range from matte black to iridescent blue, (Brien et al., 2019; Parnell et al., 2018). Finally, Type III scale cells contain the red ommochrome pigments xanthommatin and dihydroxanthommatin and are characterised by larger spacing between cross-ribs and ridges.

Here we focus on the role of *cortex* in specifying these scale types in *Heliconius* butterflies, an adaptive radiation with over 400 different wing forms in 48 described species (Jiggins, 2017; Lamas, 2004) and where diversity in wing patterns can be directly linked to the selective forces of predation and sexual selection (Brown, 1981; Turner, 1981). Specifically, we combine expression profiling using *RNA-seq*, *ATAC-seq*, *in situ* hybridization and antibody staining experiments, as well as CRISPR/Cas9 gene knock-outs to determine the role that this locus plays in pattern variation of two co-mimetic morphs of *H. melpomene* and *H. erato* (Figure 1, **b**).

Despite *cortex* not following the prevailing paradigm of patterning loci, we demonstrate that the gene plays a fundamental role in pattern variation by modulating a switch from Type I scale cells to Type II and Type III scale cells. Moreover, we show that while the phenotypic effects of *cortex* extend across the entire fore- and hindwing surface, modular enhancers have evolved in two distantly related *Heliconius* species that control spatially restricted, pattern specific expression of *cortex*. Our findings, coupled with recent functional experiments on other *Heliconius* patterning loci, are beginning to illuminate how major patterning genes interact during development to determine scale cell fate and drive phenotypic variation across a remarkable adaptive radiation.



115

Figure 1 – Phenotypic switches of yellow and white colour pattern elements are controlled by homologous loci in *Heliconius* species.

(a) Homologous loci in both *H. erato* and *H. melpomene* are associated with variation in yellow and white patterns between morphs. In *H. melpomene* three tightly linked genetic elements located at chromosome 15 have been identified that control variation for the hindwing yellow bar, white margin elements and forewing band (Yb, Sb and N respectively) while in *H. erato* variation has been mapped to one element (Cr). The gene *cortex* is highlighted in red and alignment between the two co-mimetic species at the locus is shown (grey lines, 75% alignment identity). **(b)** Focal co-mimetic morphs of *H. erato* and *H. melpomene* used in this study, differing in the presence of a hindwing yellow bar, and their ranges across Central America are shown (ranges based on Rosser et al., 2012). Yellow: yellow banded morphs, blue: black hindwing morphs, grey: range overlap.

Results

The genes *cortex* and *domeless/washout* are differentially expressed between colour pattern morphs, and between wing sections differing in the presence of the hindwing yellow bar

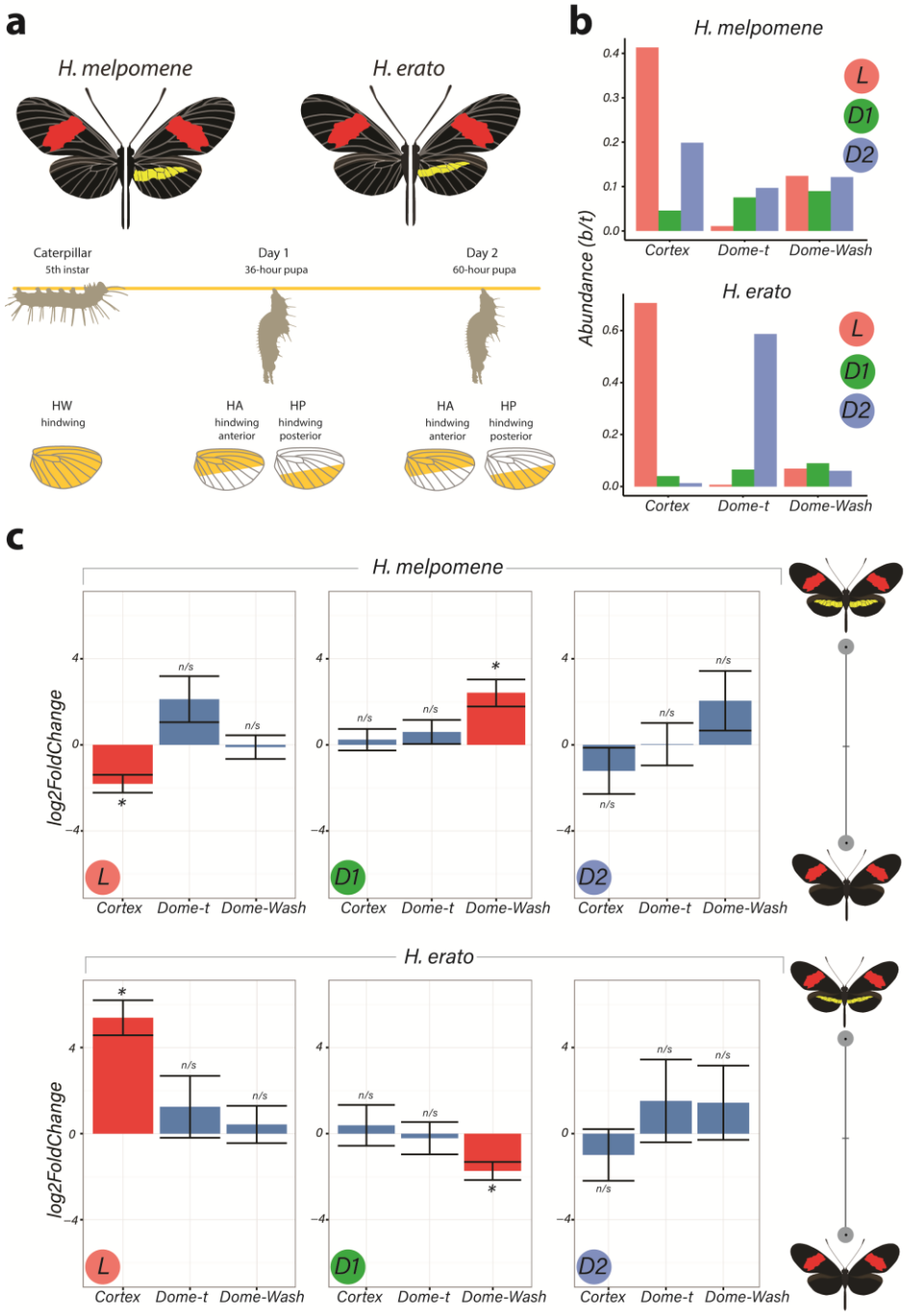
To identify genes associated with the yellow bar phenotype, we performed differential gene expression (DGE) analysis using developing hindwings sampled from colour pattern morphs in *H. erato* and *H. melpomene* differing only in the presence or absence of the hindwing yellow bar (Figure 1, **b** and Figure 2, **a**). In total, we sequenced 18 samples representing three developmental stages (larval, 36h +/-1.5h (Day 1 pupae) and 60h +/- 1.5h (Day 2 pupae)) from two morphs in each of the two species, with hindwings divided into two parts for the pupal stages (Figure 2, **a**). We focused our attention on genes centred on a 47-gene interval on chromosome 15 previously identified as the minimal associated region with yellow band phenotypes by recombination and population genetic association mapping (Nadeau et al., 2016, supp table 1; Joron et al., 2006; Moest et al., 2020; Van Belleghem et al., 2017). Both our initial expression analysis and recent analysis of selective sweeps at this locus (Moest et al., 2020) indicate that three genes show differential expression and are likely targets of selection: *cortex*, *domeless* (*dome*) and *washout* (*wash*) (Figure 2, **c**). In *Heliconius*, *dome* appears to have duplicated in the ancestor of *H. erato* and *H. melpomene*, resulting in a full-length copy (referred to here as *domeless*) and a further copy exhibiting truncations at the C-terminus (*domeless-truncated*) (Supplementary File 1 – Figure S1). Protein alignments indicate that in both *H. erato* and *H. melpomene*, *dome-trunc* maintains only the N-terminal half of the gene, suggesting *dome-trunc* is undergoing pseudogenisation. Transcriptomic evidence also indicates that *dome/wash* are transcribed as a single bi-cistronic gene (See supplementary File 2 for details and Lewis et al., 2020). Differential expression analysis was thus performed with *dome/wash* as a single annotation.

The two species were analysed separately, with both showing only *cortex* and *dome/wash* as significantly differentially expressed between morphs among the 47 genes in the candidate region, with *cortex* differential expression occurring earlier in development. In fifth instar larvae, *cortex* is differentially expressed in both species between the two colour pattern morphs, with *cortex* showing the highest adjusted *p*-value for any gene in the genome at this stage in *H. erato* (Figure 2, **c**). Interestingly, *cortex* transcripts were differentially expressed in opposite directions in the two species, with higher expression in the melanic hindwing morph in *H. melpomene*, and in the yellow banded morph in *H. erato*. We confirmed this pattern of expression through a SNP association analysis and RT-qPCR (Supplementary File 3). This pattern is reversed for *dome/wash* in Day 1 pupae, where a statistically higher proportion of transcripts are detected in *H. melpomene rosina* (yellow), and in *H. erato hydara* (melanic) (Supplementary File 4 – Tables S4.1 and S4.2). No differential expression of these genes was found at Day 2 pupae.

When comparing across hindwing sections differing for the yellow bar phenotype, 22 genes out of the associated 47-gene interval were differentially expressed at Day 1 between relevant wing sections in *H. melpomene*, including *cortex* and *dome/wash* (Supplementary File 4 – Figures S4.1 and S4.2). In contrast in *H. erato* Day 1 pupae, only *dome/wash* was differentially expressed. At Day 2 pupae, there were no differentially expressed genes in either species between relevant wing sections at this locus.

Given the strong support for the involvement of *cortex* in driving wing patterning differences, we re-analysed its phylogenetic relationship to other cdc20 family genes with more extensive sampling than previous analyses (Nadeau et al., 2016). Our analysis finds strong monophyletic support for *cortex* as an insect-specific member of the cdc20 family, with no clear *cortex* homologs found outside of the Neoptera (Supplementary File 5 – Figure S5.1). Branch lengths indicate *cortex* is evolving rapidly within the lineage, despite displaying conserved APC/C binding motifs, including the C-Box and IR tail (Supplementary File 5 – Figure S5.2) (Chu et al., 2001; Pesin and Orr-Weaver, 2007).

In summary, *cortex* is the most consistently differentially expressed gene and showed differential expression earlier in development as compared to the other candidate *dome/wash*. We therefore focus subsequent experiments on *cortex*, although at this stage we cannot rule out an additional role for *dome*/*wash* in yellow pattern specification.



166

Figure 2 – Differential expression of genes at Chromosome 15 implicate *cortex* as most likely candidate driving yellow bar differences

(a) Hindwing tissue from co-mimetic morphs of *H. melpomene* and *H. erato* were collected at three developmental stages (5th instar caterpillar, Day 1 Pupae (36hAPF) and Day 2 Pupae (60hAPF). For pupal tissue, hindwing tissue was dissected using the wing vein landmarks shown, corresponding to the future adult position of the hindwing yellow bar (dissection scheme based on Hanly et al., 2019). **(b)** Relative abundance of transcripts corresponding to the genes *cortex*, *domeless-truncated*, *domeless*/*washout* throughout developmental stages. **(c)** Log_2FoldChange for the genes *cortex*, *domeless-truncated*, *domeless*/*washout* across developmental stages. Comparisons are for whole wing discs (Larvae, L) and across wing sections differing in the presence of a yellow bar in pupal wings (D1 and D2; see Supplementary File 4: Figure S4.3 for depiction of contrasts analysed). * = adjusted p<0.05; n/s = not significant.

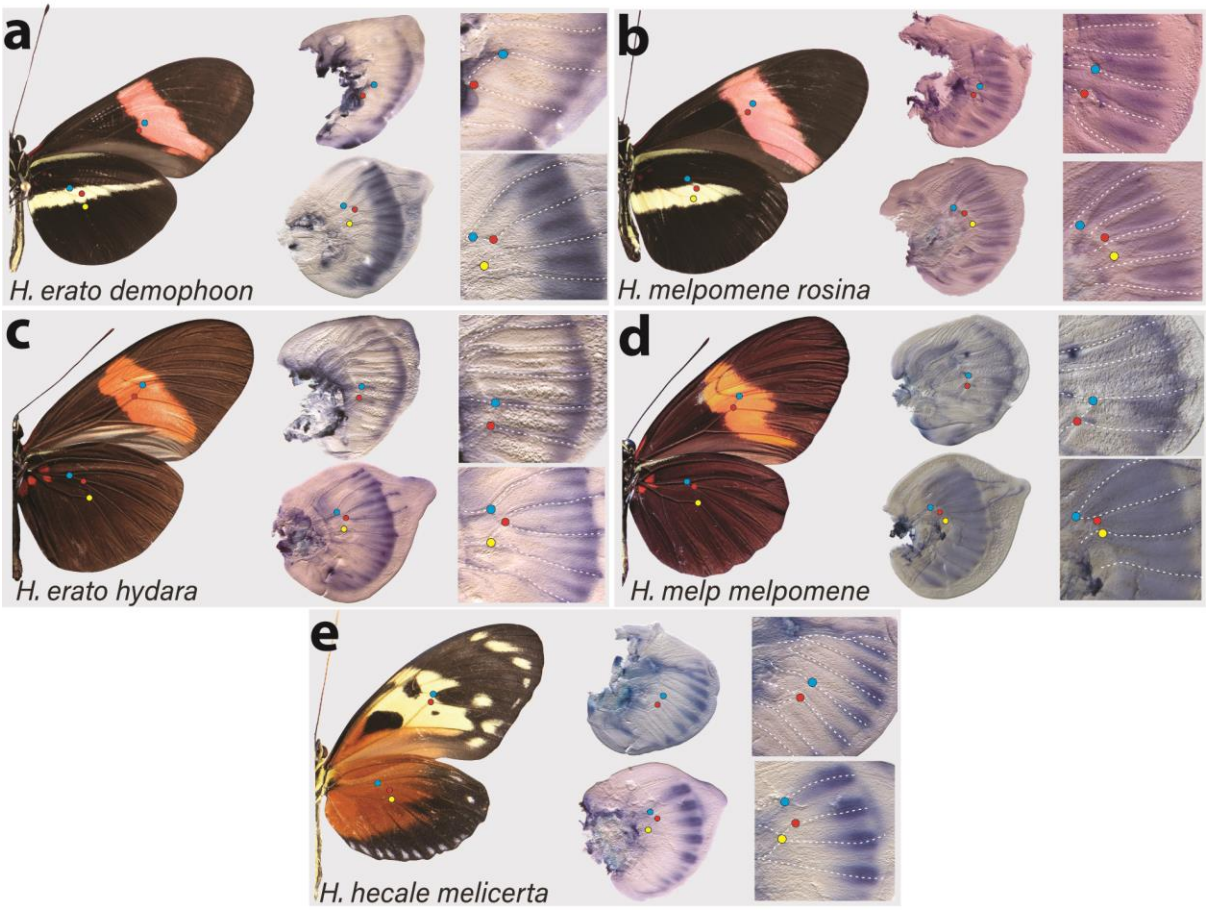
Cortex transcripts localise distally in 5th instar larval wing discs

Two studies have reported that *cortex* mRNA expression correlates with melanic patches in two species of *Heliconius* (Nadeau et al., 2016 and Saenko et al., 2019). To further assess this relationship between *cortex* expression and adult wing patterns, we performed *in situ* hybridisation on developing wing discs of 5th instar larvae, where we observed largest *cortex* transcript abundance, in both the yellow-barred and plain hindwing morphs of *H. erato* and *H. melpomene*. *Cortex* transcripts at this stage localised distally in forewings and hindwings of both species (Figure 3, Supplementary File 6 – Figure S6). In *H. erato demophoon* hindwings, expression was strongest at the intervein midline, but extends across vein compartments covering the distal portion of both forewing and hindwing (Figure 3, **a**). By contrast, in *H. erato hydara* hindwings, *cortex* transcripts are more strongly localised to the intervein midline forming a sharper intervein expression domain (Figure 3, **c**).

Expression in *H. melpomene rosina* is similar to *H. erato demophoon* at comparable developmental stages, again with stronger expression localised to the intervein midline but extending further proximally than in *H. erato demophoon* (Figure 3, **b**). In *H. melpomene melpomene*, hindwing *cortex* expression extends across most of the hindwing, and does not appear to be restricted to the intervein midline (Figure 3, **c**).

Given that *cortex* has been implicated in modulating wing patterns in many divergent lepidoptera, we also examined localisation in a *Heliconius* species displaying distinct patterns: *H. hecale melicerta* (Figure 3, **e**). Interestingly, in this species transcripts appear strongest in regions straddling the wing disc veins, with weak intervein expression observed only in the hindwings. Previous data has shown variation in yellow spots (Hspot) is also controlled by a locus located a chromosome 15 (Huber et al., 2015). Expression in *H. hecale melicerta* forewings corresponds to melanic regions located in between yellow spots at the wing margins, indicating *cortex* may be modulating Hspot variation in *H. hecale*.

Overall, our results suggest a less clear correlation to melanic elements than reported expression patterns (Nadeau et al., 2016; Saenko et al., 2019) where *cortex* expression in 5th instar caterpillars is mostly restricted to the distal regions of developing wings, but appears likely to be dynamic across 5th instar development (Supplementary File 6).



194

Figure 3 – Expression of *cortex* transcripts in *Heliconius melpomene*, *Heliconius erato* and *Heliconius hecale* 5th instar wing discs.

Cortex expression in 5th instar wing discs is restricted to the distal end of both forewings and hindwings in all species and morphs analysed. In *H. erato*, expression is strongest at the intervein midline but extends across vein compartments in *H. erato demophoon* **(a)**, whereas it is more strongly localised to the intervein midline in *H. erato hydara* **(c)**. In *H. melpomene rosina* **(b)**, *cortex* localises in a similar manner to *H. erato demophoon*, with stronger expression again observed at the intervein midline, whereas expression in *H. melpomene melpomene* **(d)** extends more proximally. Expression in *H. hecale melicerta* **(e)** is strongest at the distal wing vein margins. Coloured dots represent homologous vein landmarks across the wings.

195 ## *Cortex* establishes Type II and III scale identity in *Heliconius* butterflies

196 To assay the function of *cortex* during wing development, we generated G_0 somatic mosaic mutants
197 using CRISPR/Cas9 knock outs (crispants). We targeted multiple exons using a combination of
198 different guides and genotyped the resulting mutants through PCR amplification, cloning and Sanger
199 sequencing (Supplementary File 7 – Figure S7). Overall KO efficiency was low when compared to
200 similar studies in *Heliconius* (Concha et al., 2019; Mazo-Vargas et al., 2017), with observed wing
201 phenotype to hatched eggs ratios ranging from 0.3% to 4.8%. Lethality was also high, with hatched to
202 adult ratios ranging from 8.1% to 29.8% (Supplementary File 8 – Table S8.1).

203 Targeting of the *cortex* gene in *H. erato* morphs produced patches of ectopic yellow and white scales
204 spanning regions across both forewings and hindwings (Figure 4 and Supplementary File 9 – Figures

S9.1-S9.3). All colour pattern morphs were affected in a similar manner in *H. erato*. Mutant clones were not restricted to any specific wing region, affecting scales in both proximal and distal portions of wings. The same effect on scale pigmentation was also observed in *H. melpomene* morphs, with mutant clones affecting both distal and proximal regions in forewings and hindwings (Figure 5, **a**-**c**). In *H. erato hydara*, we recovered mutant individuals where clones also spanned the red forewing band (Figure 4, **b** and Supplementary File 9 - Figure S9.3). Clones affecting this region caused what appears to be an asymmetric deposition of pigment across the scales, as well as transformation to white, unpigmented scales (Supplementary File 10 – Figure S10).

As this locus has been associated with differences in white hindwing margin phenotypes (Jiggins and McMillan, 1997) (Figure 1, **b**), we also targeted *cortex* in mimetic morphs showing this phenotype, *H. erato cyrbia* and *H. melpomene cythera*. Mutant scales in these colour pattern morphs were also localised across both wing surfaces, with both white and yellow ectopic scales (Figure 4, **c** and Figure 5, **c**). Both the white and blue colouration in these co-mimics is structurally derived, indicating that *cortex* loss-of-function phenotype also affects the scale ultrastructure. Furthermore, we observed a positional effect, where ectopic scales in the forewing and anterior compartment of the hindwing shifted to yellow, and posterior hindwing scales became white (Figure 4, **c** and Supplementary File 10 – Figure S10, **d**). This positional effect likely reflects differential uptake of the yellow pigment 3-OHK across the wing surface, which may be related to cryptic differential expression of the yellow-white switch *aristaless-1* (Reed et al., 2008; Westerman et al., 2018).

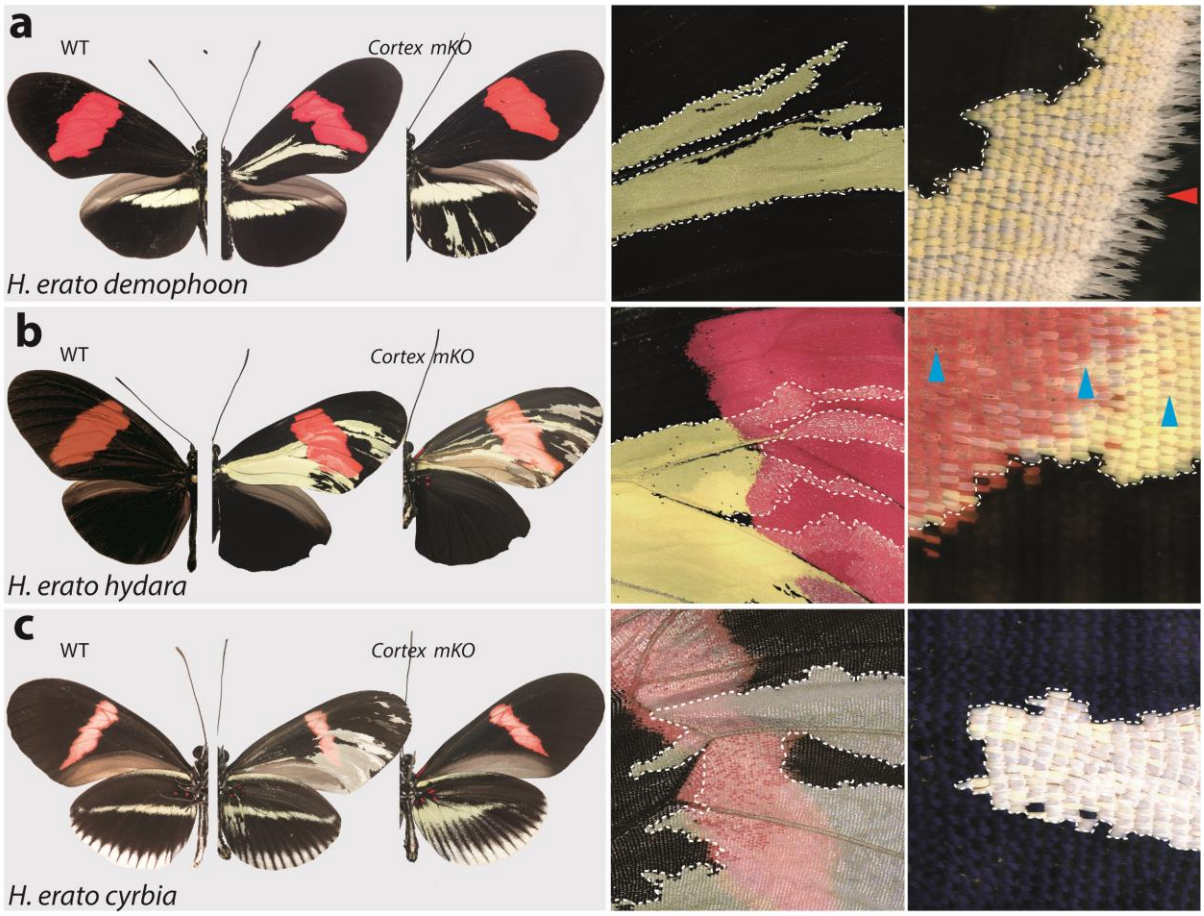


Figure 4 – *Cortex* loss-of-function transforms scale identity across the entire wing surface of *Heliconius erato*

Phenotypes of *cortex* mKO across *H. erato* morphs reveals a loss of melanic (Type II) and red (Type III) scales, and transformation to Type I (yellow or white) scales. Affected regions are not spatially restricted, and span both distal and proximal portions of forewings and hindwings. The scale transformation extends to all scale types, including the wing border scales (red arrow head in (**a**)), and across the red band elements, where mutant scales transform to white, as well as some showing an intermediate phenotype (blue arrow heads in (**b**)). A positional effect is observed in some morphs, where ectopic Type I scales are either white or yellow depending on their position along the wing (*H. erato cyrbia*, (**c**)). Ectopic Type I scales can be induced from both melanic and red scales, switching to either white or yellow depending on wing position and morph. Boundaries between Wild-type (WT) to mutant scales are highlighted (dotted white line).

To further test the conservation of *cortex* function across the *Heliconius* radiation, as well as nymphalids as a whole, we knocked out *cortex* in *H. charithonia* and *H. hecale melicerta*, outgroups to *H. erato* and *H. melpomene* respectively and *Danaus plexippus* as an outgroup to all Heliconiini. Again, ectopic yellow and white scales appeared throughout the wing surface in all species, suggesting a conserved function with respect to scale development among butterflies (Figure 5, **d-f** and Supplementary File 9, Figures S9.6-S9.8).

In summary, *cortex* crispants appear to not be restricted to any specific wing pattern elements, and instead affect regions across the surface of both forewings and hindwings. Mutant scales are always Type I scales, with differing pigmentation (3-OHK, yellow) or structural colouration (white) depending

on morph and wing position. The high rate of mosaicism combined with high mortality rates suggests *cortex* is likely developmentally lethal. Furthermore, the sharp boundaries observed between wild-type and mutant scales suggest *cortex* functions in a cell-autonomous manner, with little or no communication between neighbouring cells (Supplementary File 10 - Figure 10).

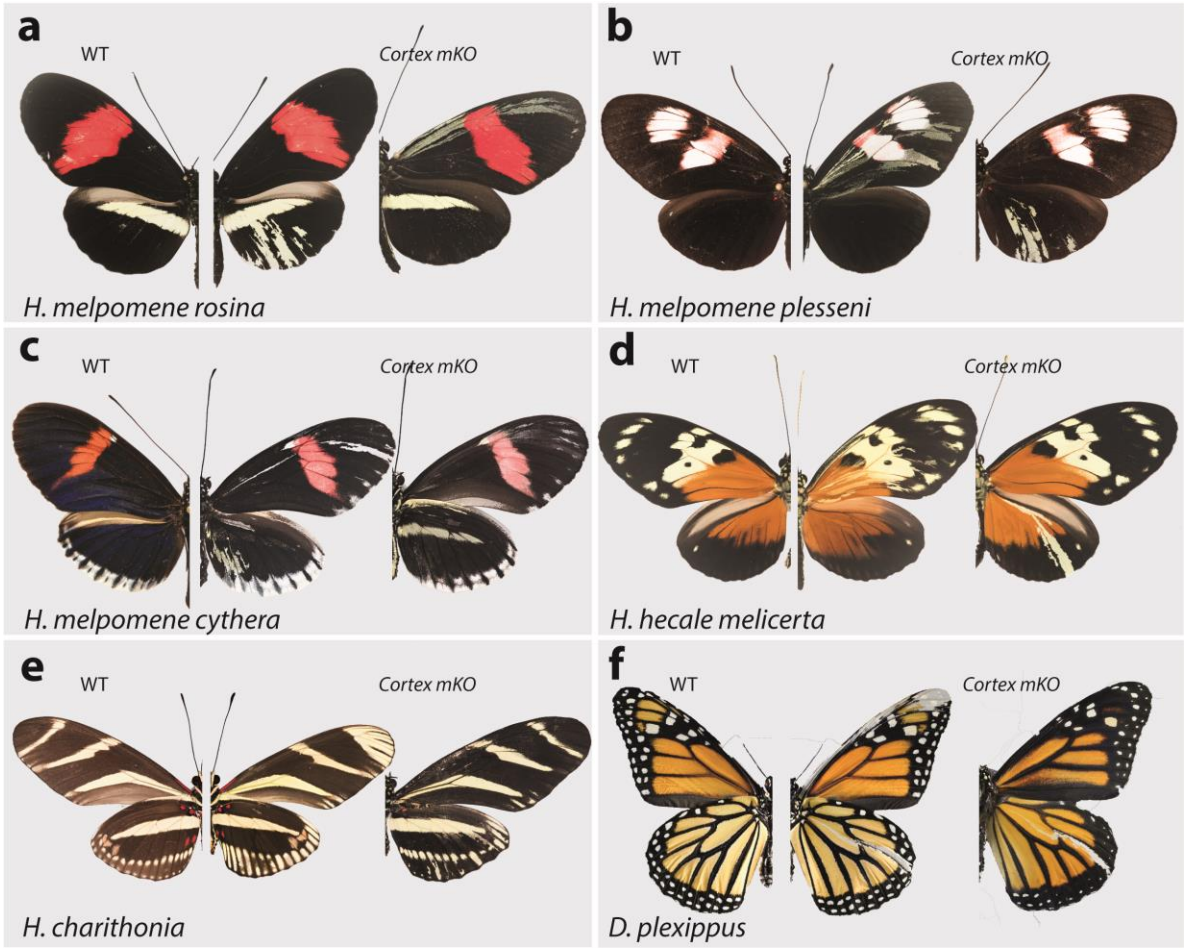


Figure 5 – *Cortex* function is conserved across *Heliconius* and Nymphalids

Phenotypes of *cortex* mKO across *H. melpomene* colour pattern morphs (**a**-**c**) reveal *cortex* has a conserved function in switching scale cell fates, as in *H. erato*. This function is also conserved in outgroups to *H. melpomene* and *H. erato* (*H. hecale melicerta* and *H. charithonia* respectively (**d**-**e**)) as well as in distantly diverged nymphalids (*D. plexippus* (**f**)). Left; wild-type, middle and right; *cortex* mKO.

Nuclear localization of Cortex protein extends across the wing surface in pupal wings

The *cortex* mRNA expression patterns in larval imaginal disks suggest a dynamic progression in the distal regions, and in a few cases (Figure 3; Nadeau et al., 2016; Saenko et al., 2019) a correlation with melanic patterns whose polymorphisms associate with genetic variation at the *cortex* locus itself. We thus initially hypothesized that like for the *WntA* mimicry gene (Martin et al., 2012, Mazo-Vargas 2017 et al., Concha et al., 2020), the larval expression domains of *cortex* would delimit the wing territories where it is playing an active role in colour patterning. However, our CRISPR based loss-of-function experiments challenge that hypothesis because in all the morphs that we assayed, we found mutant scales across the wing surface.

This led us to re-examine our model and consider that post-larval stages of *cortex* expression could reconcile the observation of scale phenotypes across the entire wing, rather than in limited areas of the wing patterns. To test this hypothesis, we developed a Cortex polyclonal antibody, and found nuclear expression across the epithelium of *H. erato demophoon* pupal hindwings without restriction to specific pattern elements (Figure 6). In 5th instar larvae, Cortex protein was detected in a similar pattern to mRNA, with expression visible at the intervein midline of developing wings (Figure 6, **a**). Cortex was then detected across the entire wing surface from 24h after pupal formation (a.p.f), until 80h a.p.f in our time series (Figure 6, **b-d** and Supplementary File 11 - Figure S11). Localisation remained nuclear throughout development and appears equal in intensity across hindwing colour pattern elements (Figure 6, **e**).

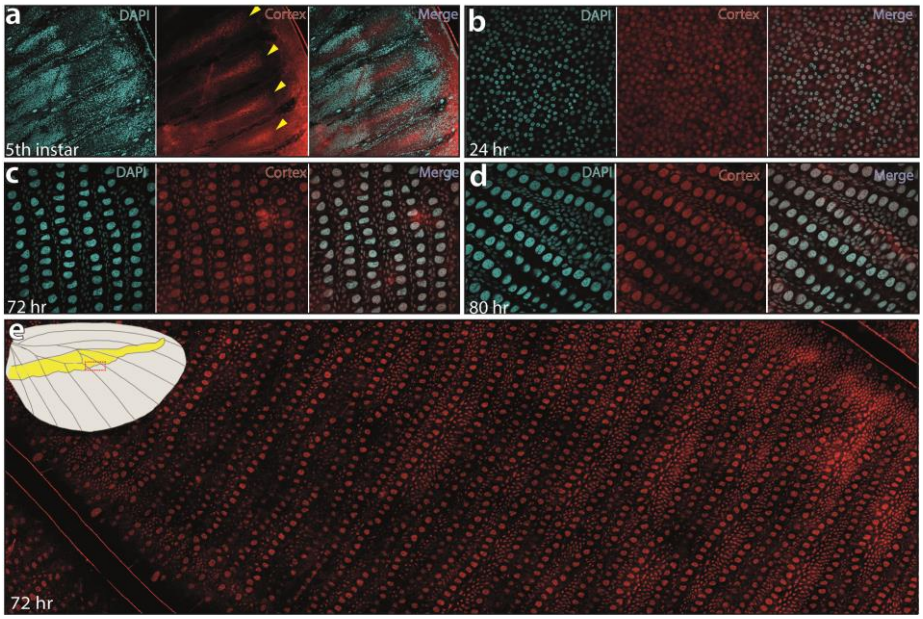


Figure 6 – Cortex protein localises across the wings in *H. erato demophoon*.

Cortex protein is localised at the distal intervein midline in 5th instar wing discs (**a**). At 24h a.p.f, the protein is detected across the wing, and localised strongly to the cell nuclei (**b**). This localisation continues at 72hr a.p.f (**c**) and 80 a.p.f (**d**). No appreciable difference in localisation is detected across presumptive pattern elements (**e**).

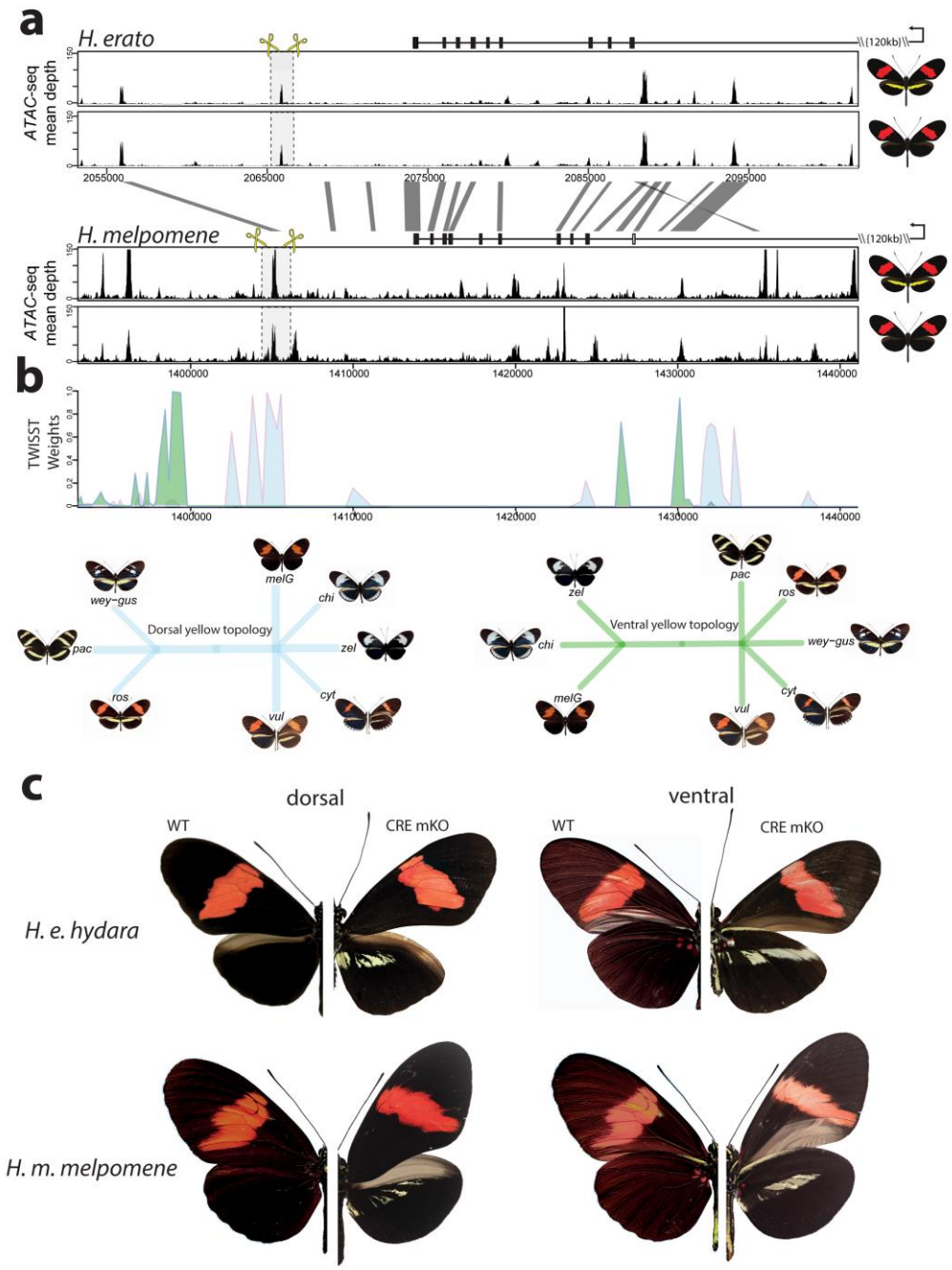
Modular *cis*-regulatory elements drive the evolution of the mimetic yellow bar

Given the broad effect observed for *cortex* across both wing surfaces, we next tested whether specific expression might be under the control of pattern specific *cis*-regulatory elements (CREs). In order to look for potential CREs, we performed an Assay for Transposase-Accessible Chromatin using sequencing (*ATAC-seq*) in 5th instar hindwings of both co-mimetic morphs differing in the presence of the yellow bar in *H. erato* and *H. melpomene*. We observed many accessible chromatin "peaks" surrounding *cortex* (Supplementary File 12 – Figure S12), each of which could represent a potentially active CRE. To narrow down candidate peaks that could be regulating *cortex* in a pattern specific manner, we overlayed association intervals with the *ATAC-seq* signals, which indicate evolved regions between populations of *H. melpomene* differing in the hindwing yellow bar phenotype. Specifically, we applied the phylogenetic weighting strategy Twisst (topology weighting by iterative sampling of subtrees; Martin and Belleghem, 2017) to identify shared or conserved genomic intervals between sets of *H. melpomene* populations (obtained from Moest et al., 2020) with similar phenotypes around *cortex*. This method identified a strong signal of association ~8kb downstream of the annotated *cortex* stop codon, that overlapped with a clear *ATAC-seq* peak (Figure 7, **a-b** and Supplementary File 12 – Figure S12).

We next sought to knock out this CRE, by designing a pair of sgRNA guides flanking the *ATAC-seq* signal. We reasoned that since *cortex* controlled the switch to melanic scales across the entire wing, knocking-out an enhancer in the melanic morph (*H. melpomene melpomene*), or in F1 hybrids between *H. melpomene rosina* and *H. melpomene melpomene*, should result in the appearance of yellow scales in a yellow bar-like pattern. Indeed, upon KO of this CRE we recovered crispants consistent with a modular enhancer driving *cortex* expression in a yellow bar-specific pattern, with no clones exhibiting yellow scales extending out of the region that forms the yellow bar (Figure 7, **c** and Supplementary File 9 Figure S9.8).

To test whether the same element was driving the evolution of the yellow bar phenotype in the co-mimetic morph of *H. erato*, we first targeted the homologous peak, which shares both 95% sequence identity with *H. melpomene*, as well as the presence of an accessible chromatin mark (Figure 7, **a**). While none of our CRISPR trials resulted in a visible phenotype at this locus (number of injected adults eclosed = 36), we did observe the presence of a further accessible region ~10kb 3' of the *H. melpomene* conserved CRE. We reasoned that a different but positionally close peak could be driving the yellow bar phenotype in *H. erato*. Remarkably, targeting of this CRE resulted in a yellow bar phenotype in the melanic *H. erato hydara*, with no clones containing yellow scales extending beyond the region that forms the yellow bar (Figure 7, **c** and Supplementary File 9 Figure S9.9). Deletions at each of the loci were confirmed by PCR amplification, cloning and Sanger sequencing (Supplementary File 7 – Figure S7).

299 Finally, to confirm that the CREs were interacting with the *cortex* promoter, we took advantage of a
300 previously published set of Hi-C samples in *H. erato* populations (Lewis et al., 2019), to check for
301 enhancer/promoter interactions through the implementation of virtual 4C plots. For both CREs, we
302 found a statistically significant interaction between CRE and promoter, indicating the observed effect
303 is likely due to the CRE interacting with the *cortex* promoter, and not a different gene at the locus
304 (Supplementary File 13 – Figure S13).

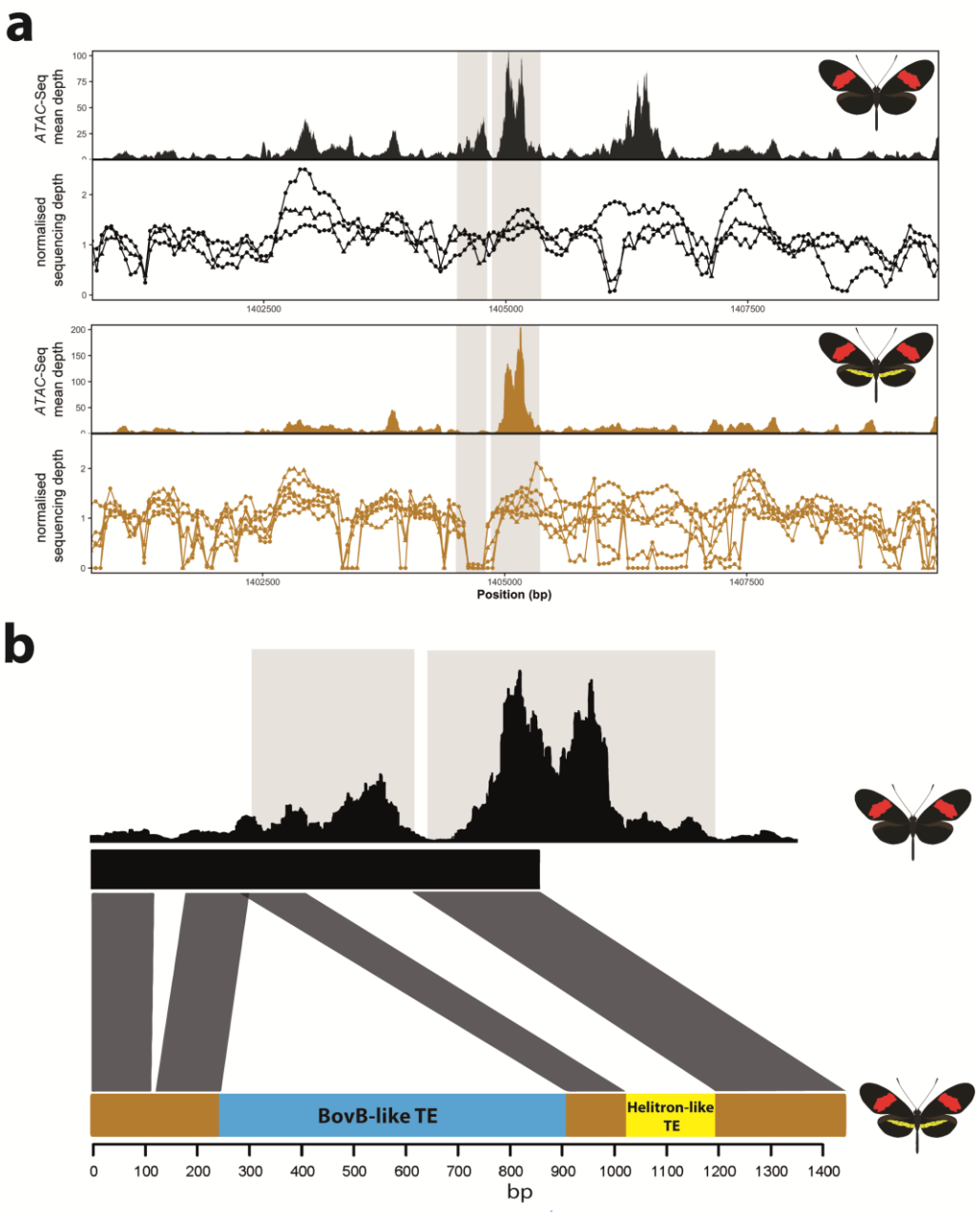


305 **Figure 7 – Modular CREs control the presence of the yellow band in *Heliconius melpomene* and *Heliconius erato.***

(a) Chromatin accessibility as measured by mean sequence depth for *ATAC-seq* traces in *H. erato* (top) and *H. melpomene* (bottom) in 5th instar caterpillar hindwings in yellow banded and black morphs. The gene model for *cortex* is shown above the traces (black rectangles are coding exons, white rectangle non-coding exon, lines are introns, direction of transcription is indicated by an arrow. The transcription start site is around 120kb downstream). Positions of sgRNAs used for peak excision are shown (yellow scissors). Regions with >75% sequence identity between *H. melpomene* and *H. erato* are indicated by grey lines. **(b)** Twisst analysis results showing high genetic association for the presence of a yellow bar in *H. melpomene* populations with a ventral (ventral topology) or dorsal (dorsal topology) yellow bar. Abbreviations for twisst morphs: wey-gus = *H. cydno weymeri-gustavi*, chi = *H. cydno chioneus*, zel = *H. cydno zelinde*, pac = *H. pachinus*, ros = *H. melpomene rosina*, melG = *H. melpomene melpomene* French Guiana), vul = *H. melpomene vulcanus*, cyt = *H. melpomene cythera*. **(c)** *Cortex* loss-of-function at the yellow bar CREs affect scales only in the presumptive yellow bar region. CRE KO affects both dorsal (left) and ventral (right) hindwings.

Transposable Element insertions are associated with the yellow bar phenotype in geographically distinct *H. melpomene* populations.

Given that we were able to induce a yellow bar phenotype by the deletion of a modular CRE, we next asked whether natural populations with this phenotype might also show a similar deletion. To test for the presence of deletions at the candidate CRE, we used extensive publicly available whole genome re-sequence data for geographically isolated populations differing in the presence of the hindwing yellow bar (Darragh et al., 2019; Enciso-Romero et al., 2017; Kozak et al., 2018; Martin et al., 2019; Van Belleghem et al., 2017). In total, we assayed 16 geographically isolated subspecies across central and south America and looked for signature coverage drops at the targeted *ATAC-seq* peak, which could be indicative of deletions (Chan et al., 2010; Kemppainen et al., 2021) (Supplementary File 14 – Table S14). We observed a characteristic drop in coverage at the targeted CRE in all *H. melpomene* and *H. timareta* morphs exhibiting a yellow bar phenotype, while no drop was detected in morphs with a melanic hindwing (Figure 8, **a**). Given this characteristic signature associated with the presence of a yellow bar in the sequencing data, we next genotyped across the putative deletion using Sanger sequencing in *H. melpomene rosina* and *H. melpomene melpomene* individuals. Surprisingly, we found two Transposable Element (TE) insertions in *H. melpomene rosina* with a Helitron-like TE found spanning the CRE peak, suggesting that the coverage drop is instead due to an insertion of repetitive sequence, rather than a deletion. Enhancer disruption is therefore likely caused by TE sequence in yellow bar morphs (Liu et al., 2019). We next assayed three other yellow barred morphs (*H. melpomene bellula*, *H. melpomene amaryllis* and *H. timareta tristero*), and found the same TE signatures in all three populations (Supplementary File 14 – Figure S14.2), suggesting the TE insertions are likely shared through introgression. No similar signature of reduced coverage was observed in co-mimetic morphs of *H. erato*, suggesting that sequence divergence is responsible for the evolution of the yellow bar CRE in this species (Supplementary File 14 – Figure S14).



330

Figure 8 – Coverage drop indicative of deletions in yellow barred populations of *H. melpomene*

(a) Mean sequence depth for *ATAC-seq* traces for the excised CRE are shown above normalised depth in sequencing coverage for populations of *H.melpomene* (circles) and H. *timareta* (triangles) differing for the presence of a yellow bar. Yellow barred populations display a drop in coverage for both ATAC signal and sequencing depth at a position corresponding to a portion of the targeted CRE (highlighted by the grey lines). Morphs analysed for melanic hindwing *H. melpomene* and *H. timareta*: *H. m. maletti*, *H. m. melpomene*, *H. t. florencia*. Morphs analysed for yellow barred *H. melpomene* and *H. timareta*: *H. m. bellula, H. m. amaryllis, H. m. rosina, H. m. burchelli, H t. thelxione, H. t. tristero*. (b) Sanger sequencing of target regions in *H. m. melpomene* and *H. m. rosina* reveals an insertion of two TE elements surrounding the yellow bar CRE. A larger BovB-like element of 690bp (blue) and a smaller 163bp Helitron-like element (yellow) are present in the *H. m. rosina* sequences, but not the *H. m. melpomene* sequences. Base pair positions of the consensus Sanger sequencing traces are shown below.

331

332

Cortex coding KO causes partial homeotic shifts in scale structure.

Previous studies have shown an association between scale ultrastructure and pigmentation in *Heliconius* butterflies (Concha et al., 2019; Gilbert et al., 1987; Janssen et al., 2001; Zhang et al., 2017). In particular, it has been reported that perturbation by wounding transforms both the pigment content and structure of scales in a tightly coupled way (Janssen et al., 2001). We thus asked whether ectopic yellow/white scales generated through *cortex* knockout were accompanied by structural transformations using Scanning Electron Microscopy (SEM) in the same way as ectopic colour scales generated through wounding or *WntA* knockouts (Janssen et al, 2001; Concha et al., 2019). To account for known positional effects on scale structure we compared wild-type and mutant scales from homologous locations across the wing surface.

We observed ultrastructural shifts that are consistent with partial homeosis in *cortex* mutant scales in both *H. melpomene* and *H. erato* (Figure 9 and Supplementary File 15). In all cases where a yellow or white (Type I) clone was present in a region that would otherwise be black or blue (Type II) in the wild type, the ultrastructure of the scale was notably different. Wild type blue and black scales have crossribs at a spacing of ~0.6 µm, lack lamina between ridges and crossribs, and have no prominent microribs, while both wild-type and mutant Type I scales have no prominent crossribs, lamina that fills the spaces between the microribs and ridges, and prominent microribs at a spacing of ~0.2 µm (Figure 9, **a-d**, and Supplementary File 15). A consistent difference between all Type I scales (mutant and wild-type) is the presence of a lamina covering the inter-ridge space (Figure 9, **f**), suggesting this structure is an important morphological feature of yellow/white scales (Matsuoka and Monteiro, 2018), and that *cortex* is necessary for the development of lamellar structures in *Heliconius* scales.

Red scales (Type III) that are within a coding KO clone take on an aberrant structure and pigmentation. Scales were frequently found to be curled up laterally, and while ommochrome pigment is sometimes visibly deposited in the scale, it is granular in appearance rather than diffuse throughout the scale (Figure 9, **e**). These 'granular' pigment accumulations could not be observed as a distinct structure by SEM, suggesting that they are under the surface of the scale. As with wild type and mutant Type I scales, prominent microribs can also be observed on these rolled scales, but due to the topological deformity of these scales it was not possible to take accurate measurements.

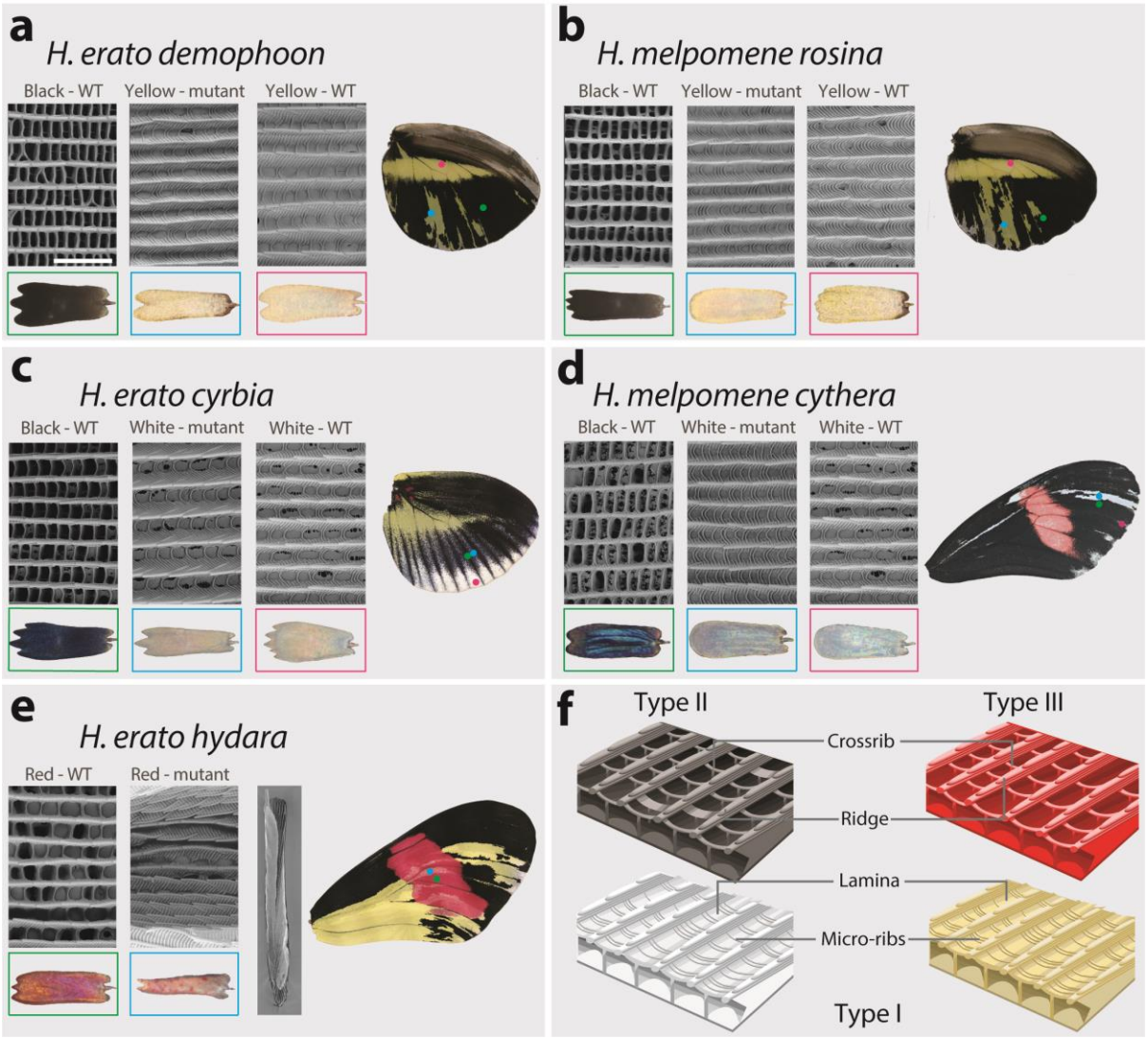


Figure 9 – SEM reveals structural changes induced by *cortex* KO.

Scanning electron microscopy images showing major differences between wild-type scales and mutant Type I yellow scales (**a,b**), Type I white scales (**c**,**d**) and Type III red scales (**e**). Cartoon depiction of scale ultrastructure illustrating differences between scale Types (**f**).

Discussion:

Cortex is a key scale cell specification gene

The genetic locus containing the gene *cortex* represents a remarkable case of convergent evolution, where repeated and independent mutations surrounding the gene are associated with shifts in scale pigmentation state in at least nine divergent species of Lepidoptera (Beldade et al., 2009; van der Burg et al., 2020; Nadeau et al., 2016; Van Belleghem et al., 2017; VanKuren et al., 2019; van't Hof et al., 2019; Van't Hof et al., 2016). While these studies have linked putative regulatory variation around *cortex* to the evolution of wing patterns, its precise effect on scale cell identity and pigmentation has remained speculative until now. Here, we demonstrate that *cortex* is a causative gene that specifies

melanic and red (Type II and Type III) scale cell identity in *Heliconius* and acts by influencing both downstream pigmentation pathways and scale cell ultrastructure. We also show that *cortex* is under the control of modular enhancers that appear to control the switch between mimetic yellow bar phenotypes in both *H. melpomene* and *H. erato*. Our combination of expression studies and functional knock-outs demonstrate that this gene acts as a key early scale cell specification switch across the wing surface of *Heliconius* butterflies, and thus has the potential to generate much broader pattern variation than previously described patterning genes.

While we have shown that *cortex* is a key scale cell specification gene, it remains unclear how a gene with homology to the *fizzy/cdc20* family of cell cycle regulators acts to modulate scale identity. In *Drosophila*, Fizzy proteins are known to regulate APC/C activity through the degradation of cyclins, leading to the arrest of mitosis (Raff et al., 2002). In particular, *fizzy-related* (*fzr*), induces a switch from the mitotic cycle to the endocycle, allowing the development of polyploid follicle cells in *Drosophila* ovaries (Schaeffer et al., 2004; Shcherbata, 2004). Similarly, *cortex* has been shown to downregulate cyclins during *Drosophila* female meiosis, through its interaction with the APC/C (Pesin and Orr-Weaver, 2007; Swan and Schüpbach, 2007). Immunostainings show that Cortex protein localises to the nucleus in *Heliconius* pupal wings, suggesting a possible interaction with the APC/C in butterfly scale building cells. Ploidy levels in Lepidoptera scale cells have been shown to correlate with pigmentation state, where increased ploidy and scale size lead to darker scales (Cho and Nijhout, 2013; Iwata and Otaki, 2016). *cortex* may thus be modulating ploidy levels by inducing endoreplication cycles in developing scale cells. However, we currently have no direct evidence for a causal relationship between ploidy state and pigmentation output, and a mechanistic understanding of this relationship and any role *cortex* may be playing in modulating ploidy levels will require future investigation.

A curious result reported from our *RNA-seq* dataset is that differential expression appears to occur in opposite directions between the two co-mimetic morphs. While this could represent some difference in the precise role of *cortex* between *H. melpomene* and *H. erato*, it may be more likely that this result reflects a dynamic expression of *cortex* in developing larval wings, in which case some relatively subtle developmental heterochrony between the two species would capture the state of differentially expressed genes in a different dynamic step.

The mimetic yellow bar phenotype switch is controlled by the evolution of modular enhancers

In *H.melpomene*, we were able to narrow down a clear peak of association with the presence of accessible chromatin marks, and showed that KO of this region results in the appearance of a yellow bar phenotype in black hindwing morphs (Figure 7). Interestingly, when targeting the homologous peak in *H. erato*, we failed to recover any type of phenotype, but were able to induce the appearance of a

yellow bar through the targeting of an adjacent peak, not present in the *H.melpomene* datasets indicating that an independently evolved CRE is driving this phenotype in *H. erato*.

These results, coupled with the coding KOs, suggests that the CREs are enhancers that are able to drive *cortex* expression in a yellow bar specific manner. It is therefore puzzling that both the *in situ* hybridisation and antibody experiments failed to recover an association between Cortex localisation and the yellow bar phenotype. One possibility is that, because *cortex* is expressed throughout the wing, the differences in *cortex* expression that drive the pattern difference are either highly discrete in time and therefore hard to observe, or are the consequence of subtle changes in concentration that we could not detect with immunofluorescence. Moreover, *cortex* is known to have complex patterns of alternative splicing (Nadeau et al, 2016), suggesting that perhaps both our polyclonal antibody and *in situ* probes lack the specificity to detect localisation of specific alternatively spliced variants. This lack of a conspicuous link between expression and function is a puzzling result that will require further investigation in future. The ideal experiments would utilise the identified enhancers as enhancer traps, to show they are able to drive expression in a pattern specific manner, as well as perform knock-in experiments in the reciprocal co-mimetic morph, to show that these regions are sufficient to drive the phenotypic switches.

In *H. melpomene*, we found a clear association between the absence of an accessible chromatin peak in yellow barred populations with a characteristic drop in coverage over the same region, that overlaps with both the targeted CRISPR and association intervals. The mapped profiles show that this drop in coverage is explained by phenotype, rather than geography, in contrast to other adjacent regions. Upon further investigation, we found a large 690bp TE insertion 5' of the peak of interest as well as a Helitron-like sequence overlapping the peak in *H. melpomene rosina*. This raises the interesting possibility that this portion of the enhancer might contain the binding sites necessary to drive *cortex* in a yellow bar specific manner, and that recurrent TE insertions across this region are driving the evolution of this phenotype in *H. melpomene* populations. We also note that this insertion is observed in mimetic morphs of a different species, *H. timareta*, with which *H. melpomene* has previously been described to share regulatory regions at other patterning loci via adaptive introgression (Morris et al., 2019; Wallbank et al., 2016). Thus, adaptive introgression of this region and its structural variants is likely facilitating mimicry in this system (Dasmahapatra et al., 2012).

***Heliconius* wing patterning is controlled by interactions among major patterning genes**

Functional knockouts now exist for all the four major loci known to drive pigmentation differences in *Heliconius* (Mazo-Vargas et al., 2017; Westerman et al., 2018; Zhang et al., 2017). These loci represent the major switching points in the GRNs that are ultimately responsible for determining scales cell identity. This work underscores the importance of two patterning loci, *cortex* and *WntA*, as master regulators of scale cell identity. Both are upregulated early in wing development and have broad effects

on pattern variation (Concha et al., 2019; Nadeau et al., 2016). The signalling molecule *WntA* modulates forewing band shape in *Heliconius* by delineating boundaries around patterns elements, and is expressed in close association with future pattern elements (Concha et al., 2019; Martin et al., 2012). Unlike *cortex* mutants, *WntA* KOs shift scale cell identity to all three cell Types (I, II and III), depending on genetic background. Thus, *WntA* acts as a spatial patterning signal inducing or inhibiting colour in specific wing elements, in contrast to *cortex*, which acts as an "on-off" switch across all scales on the butterfly wing.

Interestingly, *cortex* knockouts lead to shifts in scale fate irrespective of *WntA* expression. This suggests *cortex* interacts with *WntA* to melanise a scale in the developing wing. In certain *H. erato* colour pattern *WntA* mutants, where even in putatively *cortex* positive regions, scales are able to shift to Type I in the absence of *WntA* alone (Concha et al., 2019). This indicates that while under certain conditions *cortex* is sufficient to induce the development of black scales, *WntA* is also required as a further signal for melanisation in some genetic backgrounds. Under this scenario, colour pattern morphs may be responding epistatically to different *WntA*/*cortex* alleles present in their respective genetic backgrounds. This is also consistent with genetic evidence for epistasis between these two loci seen in crossing experiments, whereby the yellow bar in *H. erato favorinus* results from an interaction between the *Cortex* and *WntA* loci (Mallet, 1989).

Under a simple model (Figure 10), *cortex* is one of the earliest regulators and sets scale differentiation to a specific pathway switches between Type I (yellow/white) and Type II/III (black/red) scales. Thus, we can envision a differentiating presumptive scale cell (PSC) receiving a Cortex input as becoming Type II/III competent, with complete Type III differentiation occurring in the presence of *optix* expression (Zhang et al., 2017). This is consistent with our data, which shows *cortex* is also required as a signal for Type III (red) scales to properly develop. Several *cortex* mutant individuals had clones across red pattern elements and failed to properly develop red pigment. The development of red scales in *Heliconius* butterflies is also dependent on expression of the transcription factor *optix* during mid-pupal development (Lewis et al., 2019; Reed et al., 2011; Zhang et al., 2017). Therefore, *cortex* expression is required for either downstream signalling to *optix*, or to induce a permissive scale morphology for the synthesis and deposition of red pigment in future scales. *Cortex* is thus necessary for the induction of Type III scale cells but insufficient for their proper development.

Conversely, a PSC lacking a Cortex input differentiates into a Type I scale, whose pigmentation state depends on the presence of the transcription factor *aristaless1* (*al1*), where *al1* is responsible for inducing a switch from yellow to white scales in *Heliconius* by affecting the deposition of the yellow pigment 3-OHK (Westerman et al., 2018). The uptake of 3-OHK from the haemolymph occurs very late in wing development, right before the adult ecloses (Reed et al., 2008). Our *cortex* crispants revealed a shift to both yellow and white scales, with their appearance being positionally dependent; more distally located scales generally switch to white, while more proximal scales become yellow

(Supplementary File 9 and 10). This pigmentation state is likely controlled by differences in *al1* expression varying between wing sections in different ways across morphs.

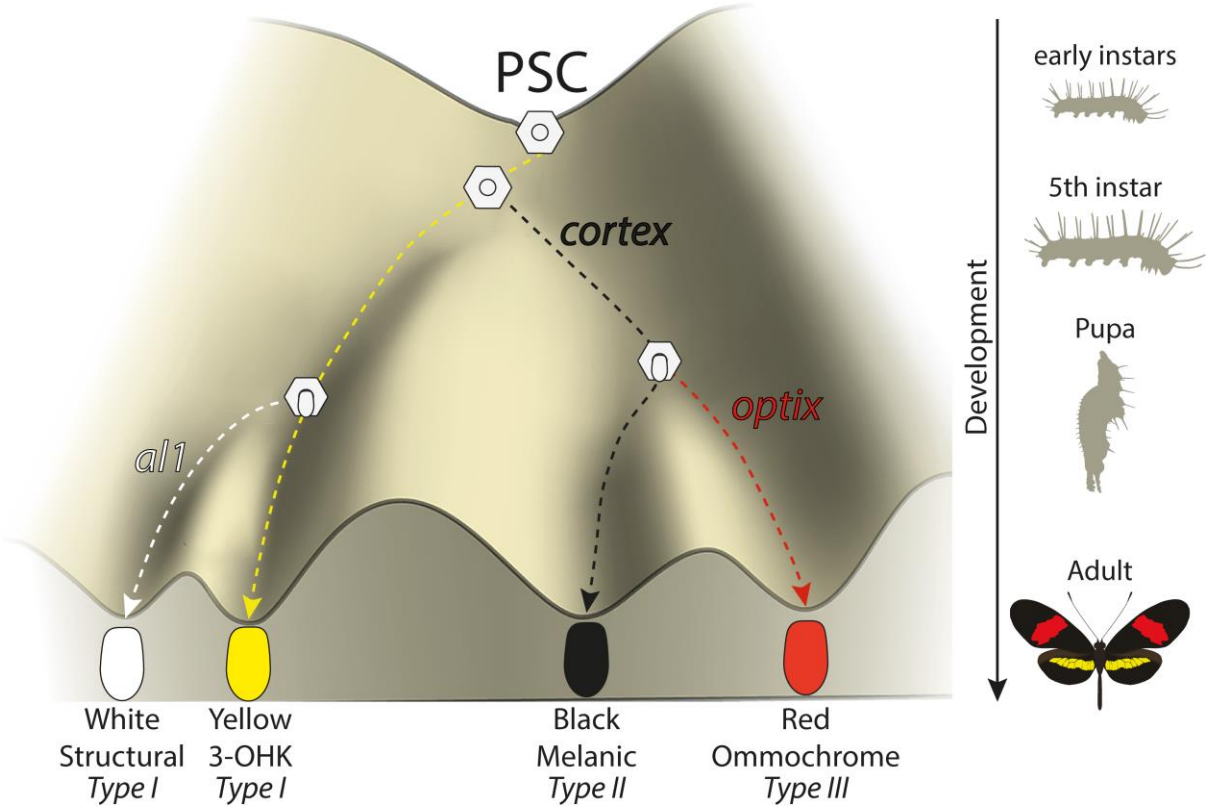


Figure 10 – Expression of key genes affect scale fate decisions and influence downstream pigmentation state

During early instar development, wing disc cells differentiate into presumptive scale cells (PSCs). Throughout 5th instar growth, PSCs express key scale cell specification genes such as *cortex*, which induce differentiation into Type II (*optix* -) scales or Type III (*optix* +) scales. In the absence of *cortex*, scale cells differentiate into Type I scales which differ in pigmentation state based on 3-OHK synthesis controlled by *aristaless1* expression. Model based on the epigenetic landscape (Waddington) and by observations made by Gilbert (1987).

However, the switch induced by Cortex under this model is likely not a simple binary toggle, and is perhaps dependent on a given protein threshold or heterochrony in expression rather than presence/absence, as we find that Cortex protein also localises to the presumptive yellow bar in developing pupal wings. Moreover, the *RNA-seq* data presented suggests other linked genes may also be playing a role in controlling pattern switches between *Heliconius* morphs. In particular, we report the presence of a bi-cistronic transcript containing the ORFs of the genes *dome* and *wash*, which are differentially expressed during early pupal wing development. While a precise role for *dome/wash* in wing patterning remains to be demonstrated, it raises the possibility that multiple linked genes co-operate during *Heliconius* wing development to drive pattern diversity. It is noteworthy that in the locally polymorphic *H. numata*, all wing pattern variation is controlled by inversions surrounding *cortex*

and *dome/wash*, both of which are also differentially expressed in *H. numata* (Saenko et al., 2019). This raises the interesting possibility that evolution has favoured the interaction of multiple genes at the locus that have since become locked into a supergene in *H. numata*.

Conclusions:

The utilization of 'hotspots' in evolution has become a recurring theme of evolutionary biology, with several examples in which independent mutations surrounding the same gene have driven adaptive evolution (e.g *Pitx1, Scute*) (Stern and Orgogozo, 2009). One proposed facilitator of such hotspots is through the action of genes acting as "input-output" modules, whereby complex spatio-temporal information is translated into a co-ordinated cell differentiation program, in a simple switch like manner. One prediction of the nature of such genes would be a switch-like behaviour such as that observed for *cortex* in this study, as well as the presence of a complex modular *cis*-regulatory architecture surrounding the gene that is able to integrate the complex upstream positional information into the switch-like output. A conserved feature of the *cortex* locus in Lepidoptera is the presence of large intergenic regions surrounding the gene, with evidence these may be acting as modular *cis*-regulatory switches in *Heliconius* (Enciso-Romero et al., 2017; Van Belleghem et al., 2017), fitting the predicted structure of input-output genes. Unlike canonical input-output loci however, *cortex* expression appears not to be restricted to any particular colour pattern element in any given species/morph, and yet is capable of producing a switch-like output (Type I vs Type II/III scales).

The genetic locus containing the gene *cortex* has now been implicated in driving wing patterning differences in many divergent Lepidoptera, and represents one of the more striking cases of convergent evolution to date. We have shown that it is spatially regulated during larval development, and yet shows wing-wide cell fate phenotypes leading to a switch in scale cell fate. Furthermore, our work shows that two independent CREs in H. melpomene and H. erato evolved to control the presence/absence of a yellow hindwing bar. The amenability of *cortex* to evolutionary change suggests it may be occupying an unusual position in the GRN underlying scale cell identity, and may be acting as an input/output gene (Stern and Orgogozo, 2009) that integrates upstream positional information into a simple on-off switch for scale differentiation. However, it is still unclear how *cortex* mechanistically affects pigmentation differences, and given its widespread usage throughout Lepidoptera, it is of general interest to understand its role in driving scale pigmentation.

Materials and Methods

Butterfly husbandry

Heliconius butterflies were collected in the tropical forests of Panama and Ecuador. Adults were provided with an artificial diet of pollen/glucose solution supplemented with flowers of *Psiguria, Lantana* and/or *Psychotria alata* according to availability. Females were provided with Passiflora plants for egg laying (*P. menispermifolia* for *H. melpomene*, *P. biflora* for *H. erato* and *H. charithonia,* and *P. vitifolia* for *H. hecale*). Eggs were collected daily, and caterpillars reared on fresh shoots of *P. williamsi* (*melpomene*), *P. biflora* (*erato* and *charithonia*) and *P. vitifolia* for *H. hecale.* Late 5th (final) instar, caterpillars were separated into individual pots in a temperature-monitored room for *RNA-seq* experiments, where they were closely observed for the purpose of accurate developmental staging.

Phylogenetic analysis of *domeless* and *cortex*

To identify orthologs of *dome* across the Lepidoptera we performed tBLASTn searches using the previously annotated *H. melpomene* Hmel2 (Hm) and *H.erato demophoon* V1 (Hed) *dome* sequences against the genomes of *Operophtera brumata* V1 (Ob), *Trichoplusia ni* Hi5.VO2 (Tn), *Bombyx mori* ASM15162v1 (Bm), *Manduca sexta* 1.0 (Ms), *Plodia interpunctella* V1 (Pi), *Amyeolis transitella* V1 (At), *Phoebis sennae* V1.1 (Ps), *Bicyclus anynana* V1.2 (Ba), *Danaus plexippus* V3 (Dp), *Dryas iulia* helico3 (Di), *Agraulis vanillae* helico3 (Av), *Heliconius erato lativitta* V1 (Hel) genomes found on Lepbase (Challis et al., 2016). As a trichopteran outgroup we used a recently published Pacbio assembly of *Stenopsyche tienmushanensis* (St) (Luo et al., 2018). Recovered amino acid translations were aligned using clustal omega (F. et al., 2019). The resulting alignments were used to produce a phylogenetic tree using PhyML (Guindon et al., 2010), based on a best fit model using AIC criterion (selected model was JTT + G + I + F). The tree was visualised and re-rooted to the Trichopteran outgroup using FigTree.

To confirm *cortex* as a cdc20 gene, we retrieved full-length protein homologs from TBLASTN searches and used them to generate a curated alignment with MAFFT/Guidance2 with a column threshold of 0.5. We then constructed a maximum-likelihood tree using W-IQ-TREE with the “Auto” function to find a best-fit model of substitution.

Tissue sampling and *RNA-seq*

H. melpomene rosina and *H. erato demophoon* butterflies were collected around Gamboa, Panama; *H. melpomene melpomene* and *H. erato hydara* butterflies were collected around Puerto Lara, Darien, Panama. Methodology for sample preparation and sequencing was performed as previously described (Hanly et al., 2019). The datasets generated and/or analysed during the current study are available in the SRA repository (PRJNA552081). Reads from each species were aligned to the respective genome assemblies Hmel2 (Davey et al., 2016) and Herato_demophoon_v1 (Van Belleghem et al., 2017), using Hisat2 with default parameters (Kim et al., 2019). The genomes and annotations

used are publicly available at www.lepbase.org. Reads were counted with HTSeq-count in union mode (Anders et al., 2015) and statistical analysis performed with the R package DESeq2 (Love et al., 2014). Comparisons for larvae were for whole hindwings, grouping samples by pattern form. Samples for pupal stages included wings that were dissected into anterior and posterior compartment as in Hanly et al (2019), and were analysed in DESeq2 using the GLM;

$$\sim \text{individual} + \text{compartment}*\text{morph}$$

(Compartments: Anterior Hindwing (HA), Posterior Hindwing (HPo)). *H. melpomene* and *H. erato* were analysed separately; homology between genes was determined by reciprocal BLAST. The fold-changes and adjusted P-values given in figure 2 reflect the primary contrast, showing the effect of pattern form given the effect of compartment. Read counts were determined for whole hindwings at all stages.

RT-qPCR

The expression level of *cortex* in larval hindwings was further analysed by qPCR in *H. e. demophoon* and *H. e. hydara*. Three individuals were used for each morph. Each individual was an independent replicate (i.e. no pooling of samples). RNA was extracted from the hindwing tissue of larva using Trizol followed by DNase-treatment. An mRNA enrichment was performed using the Dynabeads mRNA purification kit (Thermo Fisher). The mRNA was then converted to cDNA by reverse transcription using the iScript cDNA synthesis kit (Bio-Rad). All reactions had a final cDNA concentration of 2ng μl^{-1} and a primer concentration of 400nM. The RT-qPCR was carried out using Brilliant III Ultra-fast SYBR green qPCR master mix (Agilent Technologies), on a AriaMx Real-time PCR system (Agilent Technologies) according to manufacturer's instructions. The PCR programme consisted of 95°C for 2 min followed by 40 cycles of 95°C for 5 seconds, 58°C for 30 seconds and 70°C for 30 seconds. qPCR experiments were performed using three biological replicates, three technical replicates and a no template control. Expression levels were normalised using the geometric mean of three housekeeping genes, *eF1α*, *rpL3* and *polyABP* that have previously been validated for *Heliconius numata* (Piron Prunier et al., 2016). The relative expression levels were analysed using the $R = 2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). Primer specificity was confirmed using melting curve analysis and the PCR products were checked on a 2% (w/v) agarose gel. The primer efficiency of each gene was calculated using the standard curve given by a 10-fold serial dilution of cDNA (1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}) and regression coefficient (R^2) values.

***In situ* hybridizations**

Fifth instar larval wing disks and whole mount *in situ* hybridizations were performed following a published procedure (Martin and Reed, 2014) and imaged using a Leica S4E microscope (Leica Microsystems). Riboprobe synthesis was performed using the following primers from a 5th instar wing

disc cDNA library extracted from *H. melpomene*:

Forward primer 5' – CCCGAGATTCTTTCAGCGAAAC -3' and Reverse primer 5' – ACCGCTCCAACACCAAGAAG – 3'. Templates for riboprobes were designed by attaching a T7 promoter through PCR and performing a DIG labelled transcription reaction (Roche). The same *H. melpomene* probe was used in all in situ hybridisation experiments. The resulting probe spanned from Exon 2 to Exon 7 and was 841bp long.

Immunohistochemistry and image analysis

Pupal wings were dissected around 60 to 70 h post pupation in PBS and fixed at room temperature with fix buffer (400 µl 4% paraformaldehyde, 600 µl PBS 2mM EGTA) for 30 min. Subsequent washes were done in wash buffer (0.1% Triton-X 100 in PBS) before blocking the wings at 4°C in block buffer (0.05 g Bovine Serum Albumin, 10 ml PBS 0.1% Triton-X 100). Wings were then incubated in primary antibodies against Cortex (1:200, monoclonal rabbit anti-Cortex) at 4°C overnight, washed and added in secondary antibody (1:500, donkey anti-rabbit lgG, AlexaFlour 555, ThermoFisher Scientific A-31572). Before mounting, wings were incubated in DAPI with 50% glycerol overnight and finally transferred to mounting medium (60% glycerol/ 40% PBS 2mM EGTA) for imaging. Z-stacked 2-channelled confocal images were acquired using a Zeiss Cell Observer Spinning Disk Confocal microscope.

CRISPR/Cas9 genome editing

Guide RNAs were designed corresponding to $GGN_{20}NGG$ sites located within the *cortex* coding region and across putative CREs using the program Geneious (Kearse et al., 2012). To increase target specificity, guides were checked against an alignment of both *H. melpomene* and *H. erato* re-sequence data at the scaffolds containing the *cortex* gene (Moest et al., 2020; Van Belleghem et al., 2017), and selected based on sequence conservation across populations. Based on these criteria, each individual guide was checked against the corresponding genome for off-target effects, using the default Geneious algorithm. Guide RNAs with high conservation and low off-target scores were then synthesised following the protocol by Bassett and Liu, (2014). Injections were performed following procedures described in Mazo-Vargas et al., (2017), within 1-4 hours of egg laying. Several combinations of guide RNAs for separate exons at different concentrations were used for different injection experiments (Supplementary File 7). For *H. charithonia* we used the *H. erato* specific guides and for *H. hecale* we used the *H. melpomene* guides.

Genotyping

DNA was extracted from mutant leg tissue and amplified using oligonucleotides flanking the sgRNAs target region (Supplementary File 6). PCR amplicons were column purified, subcloned into the pGEM-T Easy Vector System (Promega) and sequenced on an ABI 3730 sequencer.

620

621 ***ATAC-seq***

622 *H. melpomene rosina* and *H. erato demophoon* butterflies were collected around Gamboa, Panama; *H.*
623 *melpomene melpomene* and *H. erato hydara* butterflies were collected around Puerto Lara, Darien,
624 Panama. Caterpillars of each species were reared on their respective host plants and allowed to grow
625 until the wandering stage at 5th instar. Live larvae were placed on ice for 1-2 minutes and then pinned
626 and dissected in 1X ice cold PBS. The colour of the imaginal discs, as well as the length of the lacunae,
627 gradually change throughout the larva's final day of development, so that they can be used to confirm
628 the staging inferred from pre-dissection cues (Reed et al., 2007). All larvae used for this project were
629 stage 3.5 or older. *ATAC-seq* protocol was based on previously described methodology (Lewis and
630 Reed, 2019) and edited as follows. The imaginal discs were removed and suspended whole in 350µl of
631 sucrose solution (250mM D-sucrose, 10mM Tris-HCl, 1mM. MgCl2, 1x protease inhibitors (Roche))
632 inside labelled 2ml dounce homogenisers (Sigma-Aldrich) for nuclear extraction. Imaginal discs
633 corresponding to the left and right hindwings were pooled. After homogenising the tissue on ice, the
634 resulting cloudy solution was centrifuged at 1000 rcf for 7 minutes at 4˚C. The pellet was then
635 resuspended in 150µl of cold lysis buffer (10mM Tris-HCl, 10mM NaCl, 3mM MgCl2, 0.1% IGEPAL
636 CA-630 (SigmaAldrich), 1x protease inhibitors) to burst the cell membranes and release nuclei into the
637 solution. Samples were then checked under a microscope with a counting chamber following each
638 nuclear extraction, to confirm nuclei dissociation and state and to assess the concentration of nuclei in
639 the sample. Finally based on these observations a calculation to assess number of nuclei, and therefore
640 DNA, to be exposed to the transposase was performed. This number was fixed on 400,000 nuclei, which
641 is the number of nuclei with ~0.4Gb genomes (*H. erato* genome size) required to obtain the amount of
642 DNA for which *ATAC-seq* is optimised (Buenrostro et al., 2013). For *H. melpomene* this number was
643 500,333, since the genome size of *H. melpomene* is 0.275Gb. (Buenrostro et al., 2013). For *H.*
644 *melpomene* this number was 500,333, where the genome size of *H.melpomene* is 0.275Gb. For quality
645 control, a 15µl aliquot of nuclear suspension was stained with trypan blue, placed on a hemocytometer
646 and imaged at 64x. After confirmation of adequate nuclear quality and assessment of nuclear
647 concentration, a subsample of the volume corresponding to 400,000 nuclei (*H. erato*) and 500,333 (*H.*
648 *melpomene*) was aliquoted, pelleted 1000 rcf for 7 minutes at 4˚C and immediately resuspended in a
649 transposition mix, containing Tn5 in a transposition buffer. The transposition reaction was incubated at
650 37°C for exactly 30 minutes. A PCR Minelute Purification Kit (Qiagen) was used to interrupt the
651 tagmentation and purify the resulting tagged fragments, which were amplified using custom-made
652 Nextera primers and a NEBNext High-fidelity 2x PCR Master Mix (New England Labs). Library
653 amplification was completed between the STRI laboratory facilities in Naos (Panama) and Cambridge
654 (UK). The amplified libraries were sequenced as 37bp paired-end fragments with NextSeq 500 Illumina
655 technology at the Sequencing and Genomics Facility of the University of Puerto Rico.

656

657 **Topology Weighting by Iterative Sampling of Subtrees (Twisst) Analysis**

658 We applied the phylogenetic weighting strategy Twisst (topology weighting by iterative sampling of
659 subtrees; Martin and Belleghem 2017) to identify shared or conserved genomic intervals between sets
660 of *Heliconius melpomene* and *Heliconius cydno* populations with similar phenotypes around the *cortex*
661 gene locus on chromosome 15. Given a tree and a set of pre-defined groups Twisst determines a
662 weighting for each possible topology describing the relationship of groups or phylogenetic hypothesis.
663 Similar to (Enciso-Romero et al. 2017) we evaluated the support for two alternative phylogenetic
664 hypotheses using genomic data obtained from (Moest et al. 2020). Hypothesis one tested for monophyly
665 of samples that have a dorsal yellow hindwing bar. This comparison included the geographic colour
666 patterns morphs with a dorsal hindwing bar *H. m. rosina*, *H. c. weymeri weymeri*, *H. c. weymeri gustavi*
667 and *H. pachinus* versus the all-black dorsal hindwing morphs *H. m. vulcanus*, *H. m. melpomene* (French
668 Guiana), *H. m. cythera*, *H. c. chioneus* and *H. c. zelinde*. Hypothesis two tested for monophyly of
669 samples that have a ventral yellow hindwing bar versus an all-black ventral hindwing. This comparison
670 included the geographic colour patterns morphs with a ventral hindwing bar *H. m. rosina*, *H. m.*
671 *vulcanus*, *H. m. cythera*, *H. c. weymeri weymeri*, *H. c. weymeri gustavi* and H. *pachinus* versus the all-
672 black ventral hindwing morphs *H. m. melpomene* (French Guiana), *H. c. chioneus* and *H. c. zelinde*.
673 Maximum likelihood trees were built from sliding windows of 50 SNPs with a step size of 20 SNPs
674 using PhyML v3.0 (Guindon et al. 2010) and tools available at https://github.com/simonhmartin/twisst.
675 Only windows were considered that had at least 10 sites for which each population had at least 50% of
676 its samples genotyped. Twisst was run with a fixed number of 1000 subsampling iterations.

677 **Hi-C and virtual 4C plots**

678 Analysis of chromatin contacts between distal cis-regulatory loci and the *cortex* promoter region was
679 performed as previously described (Lewis et al. 2019 PNAS, Lewis et al. 2020). In brief, Hi-C data
680 produced from day 3 pupal *H. e. lativitta* wings was used to generate an empirical expected distribution
681 and read counts for Hi-C contacts between 5kb windows centred on the distal CRE and *cortex* promoter
682 were used to determine the observed contacts between loci. Fisher's exact test was then performed to
683 determine significance of the observed contacts relative to those expected from the background model.
684 Virtual Hi-C signal plots were generated using a custom python script (Ray et al. 2019).

685 **Coverage depth analysis and TE genotyping**

686 High-depth whole-genome sequences of 16 *H. melpomene*, *H. timareta*, and *H. erato* subspecies were
687 obtained from the European Nucleotide Archive, accession numbers can be found in Supplementary
688 File 14 – Table S14. To assess structural variation putatively affecting the yellow phenotype, reads were
689 mapped to the reference genomes of subspecies that lacked the yellow bar stored in the genome browser

Lepbase, "hmel2.5" for *H. melpomene* and *H. timareta*, and "Heliconius_erato_lativitta_v1" for *H. erato* (Challis et al., 2016) with BWA mem (Li et al., 2009). Median sequencing depths across the scaffold containing cortex were computed for all individuals (n=79) in 50bp sliding windows and a mapping quality threshold of 30 with the package Mosdepth (Perdersen and Quinlan, 2018). Window median depths were normalised by dividing them by the mean depth for the full scaffold per individual. We then averaged the normalised median depths of all individuals per subspecies, to visualise deviations from the mean sequencing depth across the region in geographically widespread subspecies with and without the yellow bar. We then genotyped across the putative *H. melpomene* deletion using the primers employed for CRISPR genotyping (see Supplementary File 8). The products were then cloned into the pGEM-T Easy Vector System (Promega) and sequenced them on an ABI 3730 sequencer from both directions using T7 forward and M13 reverse primers. Sequencing was performed from three separate colonies, and a consensus sequence was created based on an alignment of the three replicates from populations of *H. m. melpomene* and *H. m. rosina*.

Scanning Electron Microscopy (SEM) Imaging

Individual scales from wild type and mutant regions of interest were collected by brushing the surface of the wing with an eyelash tool, then dusted onto an SEM stub with double-sided carbon tape. Stubs were then colour imaged under the Keyence VHX-5000 microscope for registration of scale type. Samples were sputter-coated with one 12.5 nm layer of gold for improving sample conductivity. SEM images were acquired on a FEI Teneo LV SEM, using secondary electrons (SE) and an Everhart-Thornley detector (ETD) using a beam energy of 2.00 kV, beam current of 25 pA, and a 10 µs dwell time. Individual images were stitched using the Maps 3.10 software (ThermoFisher Scientific).

Morphometrics analysis

Morphometric measurements of scale widths and ridge distances were carried out on between 10 and 20 scales of each type, using a custom semi-automated R pipeline that derives ultrastructural parameters from large SEM images (Day et al., 2019). Briefly, ridge spacing was assessed by Fourier transforming intensity traces of the ridges acquired from the FIJI software (Schindelin et al., 2012). Scale width was directly measured in FIJI by manually tracing a line, orthogonal to the ridges, at the section of maximal width.

Acknowledgements

We thank Oscar Paneso, Elizabeth Evans, Rachel Crisp and Cruz Batista, for technical support with rearing of butterflies and CRISPR larvae, and to Markus Möest, and Tim Thurman for assistance with butterfly collection. We are also grateful to Krzysztof ''Chris'' Kozak and Chi Yun for thoughtful discussions and feedback on the manuscript. We thank the GW Nanofabrication and Imaging Center

for enabling SEM, and in particular Christine Brantner and Anastas Popratiloff for their technical assistance.

Competing interests

The authors declare no competing interests.

Data availability

The ATAC-Seq sequencing reads reported in this paper have been deposited under ENA BioProject (accession number PRJEB43672).

Author Contributions

C.D.J., L.L., J.J.H., A.M., and W.O.M. designed the research; L.L., J.J.H., L.S.L., A.R., E.S.M.H., S.M.B, J.L., Z.G., I.A.W., C.C., C.W., J.M.W., J.F., L.H.L., G.M.K., H.A.C., L.R.B. R.P., M.P. performed research. L.L wrote the paper. C.D.J and W.O.M contributed equally.

Funding

This work was funded by a grant from the BBSRC to C.J and supported L.L, I.A.W., C.W., J.M.W., (BB/R007500/1); the National Science Foundation awards IOS-1656553 and IOS-1755329 to A.M and supported J.J.H., L.S.L., A.R.,; a Wellcome Trust PhD studentship awarded to JJH, a Smithsonian Institution grant to WOM and supported C.C., J.F., H.A.C., L.R.B.

References

Aymone ACB, Valente VLS, de Araújo AM. 2013. Ultrastructure and morphogenesis of the wing scales in Heliconius erato phyllis (Lepidoptera: Nymphalidae): What silvery/brownish surfaces can tell us about the development of color patterning? *Arthropod Structure & Development* **42**:349–359. doi:10.1016/j.asd.2013.06.001

Beldade P, Saenko SV, Pul N, Long AD. 2009. A Gene-Based Linkage Map for Bicyclus anynana Butterflies Allows for a Comprehensive Analysis of Synteny with the Lepidopteran Reference Genome. *PLOS Genetics* **5**:e1000366. doi:10.1371/journal.pgen.1000366

Brien MN, Enciso-Romero J, Parnell AJ, Salazar PA, Morochz C, Chalá D, Bainbridge HE, Zinn T, Curran EV, Nadeau NJ. 2019. Phenotypic variation in Heliconius erato crosses shows that iridescent structural colour is sex-linked and controlled by multiple genes. *Interface Focus* **9**:20180047. doi:10.1098/rsfs.2018.0047

Brown KS. 1981. The Biology of Heliconius and Related Genera. *Annual Review of Entomology* **26**:427–457. doi:10.1146/annurev.en.26.010181.002235

Burg KRL van der, Lewis JJ, Brack BJ, Fandino RA, Mazo-Vargas A, Reed RD. 2020. Genomic architecture of a genetically assimilated seasonal color pattern. *Science* **370**:721–725. doi:10.1126/science.aaz3017

Challi RJ, Kumar S, Dasmahapatra KK, Jiggins CD, Blaxter M. 2016. Lepbase: the Lepidopteran genome database. *bioRxiv* 056994. doi:10.1101/056994

758 Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, Southwick AM, Absher DM,
759 Grimwood J, Schmutz J, Myers RM, Petrov D, Jónsson B, Schluter D, Bell MA, Kingsley DM.
760 2010. Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1
761 Enhancer. *Science* **327**:302–305. doi:10.1126/science.1182213
762 Cho EH, Nijhout HF. 2013. Development of polyploidy of scale-building cells in the wings of Manduca
763 sexta. *Arthropod Struct Dev* **42**:37–46. doi:10.1016/j.asd.2012.09.003
764 Chu T, Henrion G, Haegeli V, Strickland S. 2001. Cortex, a Drosophila gene required to complete
765 oocyte meiosis, is a member of the Cdc20/fizzy protein family. *Genesis* **29**:141–152.
766 doi:10.1002/gene.1017
767 Concha C, Wallbank RWR, Hanly JJ, Fenner J, Livraghi L, Rivera ES, Paulo DF, Arias C, Vargas M,
768 Sanjeev M, Morrison C, Tian D, Aguirre P, Ferrara S, Foley J, Pardo-Diaz C, Salazar C, Linares
769 M, Massardo D, Counterman BA, Scott MJ, Jiggins CD, Papa R, Martin A, McMillan WO. 2019.
770 Interplay between Developmental Flexibility and Determinism in the Evolution of Mimetic
771 Heliconius Wing Patterns. *Current Biology* S0960982219313168.
772 doi:10.1016/j.cub.2019.10.010
773 Courtier-Orgogozo V, Arnoult L, Prigent SR, Wiltgen S, Martin A. 2020. Gephebase, a database of
774 genotype–phenotype relationships for natural and domesticated variation in Eukaryotes.
775 *Nucleic Acids Res* **48**:D696–D703. doi:10.1093/nar/gkz796
776 Darragh K, Byers KJRP, Merrill RM, McMillan WO, Schulz S, Jiggins CD. 2019. Male pheromone
777 composition depends on larval but not adult diet in Heliconius melpomene. *Ecological*
778 *Entomology* **44**:397–405. doi:https://doi.org/10.1111/een.12716
779 Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Hughes DST,
780 Ferguson LC, Martin SH, Salazar C, Lewis JJ, Adler S, Ahn S-J, Baker DA, Baxter SW,
781 Chamberlain NL, Chauhan R, Counterman BA, Dalmay T, Gilbert LE, Gordon K, Heckel DG,
782 Hines HM, Hoff KJ, Holland PWH, Jacquin-Joly E, Jiggins FM, Jones RT, Kapan DD, Kersey P,
783 Lamas G, Lawson D, Mapleson D, Maroja LS, Martin A, Moxon S, Palmer WJ, Papa R,
784 Papanicolaou A, Pauchet Y, Ray DA, Rosser N, Salzberg SL, Supple MA, Surridge A, Tenger-
785 Trolander A, Vogel H, Wilkinson PA, Wilson D, Yorke JA, Yuan F, Balmuth AL, Eland C, Gharbi
786 K, Thomson M, Gibbs RA, Han Y, Jayaseelan JC, Kovar C, Mathew T, Muzny DM, Ongeri F, Pu
787 L-L, Qu J, Thornton RL, Worley KC, Wu Y-Q, Linares M, Blaxter ML, ffrench-Constant RH,
788 Joron M, Kronforst MR, Mullen SP, Reed RD, Scherer SE, Richards S, Mallet J, Owen McMillan
789 W, Jiggins CD, The Heliconius Genome Consortium. 2012. Butterfly genome reveals
790 promiscuous exchange of mimicry adaptations among species. *Nature* **487**:94–98.
791 doi:10.1038/nature11041
792 Day CR, Hanly JJ, Ren A, Martin A. 2019. Sub-micrometer insights into the cytoskeletal dynamics and
793 ultrastructural diversity of butterfly wing scales. *Dev Dyn* **248**:657–670. doi:10.1002/dvdy.63
794 Enciso-Romero J, Pardo-Díaz C, Martin SH, Arias CF, Linares M, McMillan WO, Jiggins CD, Salazar C.
795 2017. Evolution of novel mimicry rings facilitated by adaptive introgression in tropical
796 butterflies. *Molecular Ecology* **26**:5160–5172. doi:10.1111/mec.14277
797 F M, Ym P, J L, N B, T G, N M, P B, Arn T, Sc P, Rd F, R L. 2019. The EMBL-EBI search and sequence
798 analysis tools APIs in 2019. *Nucleic Acids Res* **47**:W636–W641. doi:10.1093/nar/gkz268
799 Finkbeiner SD, Fishman DA, Osorio D, Briscoe AD. 2017. Ultraviolet and yellow reflectance but not
800 fluorescence is important for visual discrimination of conspecifics by Heliconius erato.
801 *Journal of Experimental Biology* **220**:1267–1276. doi:10.1242/jeb.153593
802 Gilbert LE, Forrest HS, Schultz TD, Harvey DJ. 1987. Correlations of ultrastructure and pigmentation
803 suggest how genes control development of wing scales of Heliconius butterflies. *The Journal*
804 *of research on the Lepidoptera (USA)*.
805 Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and
806 Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of
807 PhyML 3.0. *Systematic Biology* **59**:307–321. doi:10.1093/sysbio/syq010

Huber B, Whibley A, Poul YL, Navarro N, Martin A, Baxter S, Shah A, Gilles B, Wirth T, McMillan WO, Joron M. 2015. Conservatism and novelty in the genetic architecture of adaptation in Heliconius butterflies. *Heredity* **114**:515–524. doi:10.1038/hdy.2015.22

Ito K, Katsuma S, Kuwazaki S, Jouraku A, Fujimoto T, Sahara K, Yasukochi Y, Yamamoto K, Tabunoki H, Yokoyama T, Kadono-Okuda K, Shimada T. 2016. Mapping and recombination analysis of two moth colour mutations, Black moth and Wild wing spot, in the silkworm Bombyx mori. *Heredity* **116**:52–59. doi:10.1038/hdy.2015.69

Iwata M, Otaki JM. 2016. Spatial patterns of correlated scale size and scale color in relation to color pattern elements in butterfly wings. *Journal of Insect Physiology* **85**:32–45. doi:10.1016/j.jinsphys.2015.11.013

Janssen JM, Monteiro A, Brakefield PM. 2001. Correlations between scale structure and pigmentation in butterfly wings. *Evolution & Development* **3**:415–423. doi:https://doi.org/10.1046/j.1525-142X.2001.01046.x

Jiggins CD. 2017. The Ecology and Evolution of Heliconius Butterflies. Oxford University Press.

Jiggins CD, McMillan WO. 1997. The genetic basis of an adaptive radiation: warning colour in two Heliconius species. *Proc Biol Sci* **264**:1167–1175. doi:10.1098/rspb.1997.0161

Joron M, Papa R, Beltrán M, Chamberlain N, Mavárez J, Baxter S, Abanto M, Bermingham E, Humphray SJ, Rogers J, Beasley H, Barlow K, H. ffrench-Constant R, Mallet J, McMillan WO, Jiggins CD. 2006. A Conserved Supergene Locus Controls Colour Pattern Diversity in Heliconius Butterflies. *PLoS Biol* **4**:e303. doi:10.1371/journal.pbio.0040303

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649. doi:10.1093/bioinformatics/bts199

Kemppainen P, Li Z, Rastas P, Löytynoja A, Fang B, Yang J, Guo B, Shikano T, Merilä J. n.d. Genetic population structure constrains local adaptation in sticklebacks. *Molecular Ecology* **n/a**. doi:https://doi.org/10.1111/mec.15808

Koch PB. 1993. Production of [14C]-Labeled 3-Hydroxy-L-Kynurenine in a Butterfly, Heliconius charitonia L. (Heliconidae), and Precursor Studies in Butterfly Wing Ommatins. *Pigment Cell Research* **6**:85–90. doi:10.1111/j.1600-0749.1993.tb00586.x

Kozak KM, McMillan WO, Joron M, Jiggins CD. 2018. Genome-wide admixture is common across the Heliconius radiation. *bioRxiv* 414201. doi:10.1101/414201

Kronforst MR, Papa R. 2015. The Functional Basis of Wing Patterning in Heliconius Butterflies: The Molecules Behind Mimicry. *Genetics* **200**:1–19. doi:10.1534/genetics.114.172387

Lamas G, editor. 2004. Atlas Of Neotropical Lepidoptera: Checklist Pt. 4a Hesperioidea-papilionoidea. Gainesville: Scientific Pub.

Lewis JJ, Belleghem SMV, Papa R, Danko CG, Reed RD. 2020. Many functionally connected loci foster adaptive diversification along a neotropical hybrid zone. *Science Advances* **6**:eabb8617. doi:10.1126/sciadv.abb8617

Lewis JJ, Geltman RC, Pollak PC, Rondem KE, Belleghem SMV, Hubisz MJ, Munn PR, Zhang L, Benson C, Mazo-Vargas A, Danko CG, Counterman BA, Papa R, Reed RD. 2019. Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *PNAS* **116**:24174–24183. doi:10.1073/pnas.1907068116

Lewis JJ, Reed RD. 2019. Genome-Wide Regulatory Adaptation Shapes Population-Level Genomic Landscapes in Heliconius. *Molecular Biology and Evolution* **36**:159–173. doi:10.1093/molbev/msy209

Liu Y, Ramos-Womack M, Han C, Reilly P, Brackett KL, Rogers W, Williams TM, Andolfatto P, Stern DL, Rebeiz M. 2019. Changes throughout a Genetic Network Mask the Contribution of Hox Gene Evolution. *Curr Biol* **29**:2157-2166.e6. doi:10.1016/j.cub.2019.05.074

857 Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative
858 PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**:402–408.
859 doi:10.1006/meth.2001.1262
860 Luo S, Tang M, Frandsen PB, Stewart RJ, Zhou X. 2018. The genome of an underwater architect, the
861 caddisfly Stenopsyche tienmushanensis Hwang (Insecta: Trichoptera). *Gigascience* **7**.
862 doi:10.1093/gigascience/giy143
863 Martin A, Courtier-Orgogozo V. 2017. Morphological Evolution Repeatedly Caused by Mutations in
864 Signaling Ligand Genes In: Sekimura T, Nijhout HF, editors. Diversity and Evolution of
865 Butterfly Wing Patterns: An Integrative Approach. Singapore: Springer. pp. 59–87.
866 doi:10.1007/978-981-10-4956-9_4
867 Martin A, Papa R, Nadeau NJ, Hill RI, Counterman BA, Halder G, Jiggins CD, Kronforst MR, Long AD,
868 McMillan WO, Reed RD. 2012. Diversification of complex butterfly wing patterns by
869 repeated regulatory evolution of a Wnt ligand. *Proc Natl Acad Sci USA* **109**:12632–12637.
870 doi:10.1073/pnas.1204800109
871 Martin A, Reed RD. 2014. Wnt signaling underlies evolution and development of the butterfly wing
872 pattern symmetry systems. *Developmental Biology* **395**:367–378.
873 doi:10.1016/j.ydbio.2014.08.031
874 Martin SH, Davey JW, Salazar C, Jiggins CD. 2019. Recombination rate variation shapes barriers to
875 introgression across butterfly genomes. *PLOS Biology* **17**:e2006288.
876 doi:10.1371/journal.pbio.2006288
877 Massey J, Wittkopp PJ. 2016. The genetic basis of pigmentation differences within and between
878 Drosophila species. *Curr Top Dev Biol* **119**:27–61. doi:10.1016/bs.ctdb.2016.03.004
879 Matsuoka Y, Monteiro A. 2018. Melanin Pathway Genes Regulate Color and Morphology of Butterfly
880 Wing Scales. *Cell Reports* **24**:56–65. doi:10.1016/j.celrep.2018.05.092
881 Mazo-Vargas A, Concha C, Livraghi L, Massardo D, Wallbank RWR, Zhang L, Papador JD, Martinez-
882 Najera D, Jiggins CD, Kronforst MR, Breuker CJ, Reed RD, Patel NH, McMillan WO, Martin A.
883 2017. Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity.
884 *PNAS* **114**:10701–10706. doi:10.1073/pnas.1708149114
885 McMillan WO, Livraghi L, Concha C, Hanly JJ. 2020. From Patterning Genes to Process: Unraveling
886 the Gene Regulatory Networks That Pattern Heliconius Wings. doi:10.17863/CAM.55961
887 Moest M, Belleghem SMV, James JE, Salazar C, Martin SH, Barker SL, Moreira GRP, Mérot C, Joron M,
888 Nadeau NJ, Steiner FM, Jiggins CD. 2020. Selective sweeps on novel and introgressed
889 variation shape mimicry loci in a butterfly adaptive radiation. *PLOS Biology* **18**:e3000597.
890 doi:10.1371/journal.pbio.3000597
891 Moest M, Belleghem SMV, James JE, Salazar C, Martin SH, Barker SL, Moreira GRP, Mérot C, Joron M,
892 Nadeau NJ, Steiner FM, Jiggins CD. 2019. Classic and introgressed selective sweeps shape
893 mimicry loci across a butterfly adaptive radiation. *bioRxiv* 685685. doi:10.1101/685685
894 Morris J, Navarro N, Rastas P, Rawlins LD, Sammy J, Mallet J, Dasmahapatra KK. 2019. The genetic
895 architecture of adaptation: convergence and pleiotropy in Heliconius wing pattern evolution.
896 *Heredity* **123**:138–152. doi:10.1038/s41437-018-0180-0
897 Murugesan SN, Connahs H, Matsuoka Y, Gupta M das, Huq M, Gowri V, Monroe S, Deem KD, Werner
898 T, Tomoyasu Y, Monteiro A. 2021. Butterfly eyespots evolved via co-option of the antennal
899 gene-regulatory network. *bioRxiv* 2021.03.01.429915. doi:10.1101/2021.03.01.429915
900 Nadeau NJ. 2016. Genes controlling mimetic colour pattern variation in butterflies. *Current Opinion*
901 *in Insect Science*, Global change biology * Molecular physiology **17**:24–31.
902 doi:10.1016/j.cois.2016.05.013
903 Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko SV, Wallbank RWR, Wu GC, Maroja L,
904 Ferguson L, Hanly JJ, Hines H, Salazar C, Merrill RM, Dowling AJ, ffrench-Constant RH,
905 Llaurens V, Joron M, McMillan WO, Jiggins CD. 2016. The gene cortex controls mimicry and
906 crypsis in butterflies and moths. *Nature* **534**:106–110. doi:10.1038/nature17961

907 Orteu A, Jiggins CD. 2020. The genomics of coloration provides insights into adaptive evolution.
908 *Nature Reviews Genetics* **21**:461–475. doi:10.1038/s41576-020-0234-z
909 Parnell AJ, Bradford JE, Curran EV, Washington AL, Adams G, Brien MN, Burg SL, Morochz C,
910 Fairclough JPA, Vukusic P, Martin SJ, Doak S, Nadeau NJ. 2018. Wing scale ultrastructure
911 underlying convergent and divergent iridescent colours in mimetic Heliconius butterflies.
912 *Journal of The Royal Society Interface* **15**:20170948. doi:10.1098/rsif.2017.0948
913 Pesin JA, Orr-Weaver TL. 2007. Developmental Role and Regulation of cortex, a Meiosis-Specific
914 Anaphase-Promoting Complex/Cyclosome Activator. *PLOS Genetics* **3**:e202.
915 doi:10.1371/journal.pgen.0030202
916 Piron Prunier F, Chouteau M, Whibley A, Joron M, Llaurens V. 2016. Selection of Valid Reference
917 Genes for Reverse Transcription Quantitative PCR Analysis in Heliconius numata
918 (Lepidoptera: Nymphalidae). *Journal of Insect Science* **16**. doi:10.1093/jisesa/iew034
919 Prud'homme B, Gompel N, Carroll SB. 2007. Emerging principles of regulatory evolution. *PNAS*
920 **104**:8605–8612. doi:10.1073/pnas.0700488104
921 Raff JW, Jeffers K, Huang J. 2002. The roles of Fzy/Cdc20 and Fzr/Cdh1 in regulating the destruction
922 of cyclin B in space and time. *J Cell Biol* **157**:1139–1149. doi:10.1083/jcb.200203035
923 Rebeiz M, Patel NH, Hinman VF. 2015. Unraveling the Tangled Skein: The Evolution of Transcriptional
924 Regulatory Networks in Development. *Annu Rev Genom Hum Genet* **16**:103–131.
925 doi:10.1146/annurev-genom-091212-153423
926 Reed RD, McMillan WO, Nagy LM. 2008. Gene expression underlying adaptive variation in Heliconius
927 wing patterns: non-modular regulation of overlapping cinnabar and vermilion prepatterns.
928 *Proc Biol Sci* **275**:37–46. doi:10.1098/rspb.2007.1115
929 Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, Jiggins CD, Chamberlain NL,
930 Kronforst MR, Chen R, Halder G, Nijhout HF, McMillan WO. 2011. optix drives the repeated
931 convergent evolution of butterfly wing pattern mimicry. *Science* **333**:1137–1141.
932 doi:10.1126/science.1208227
933 Saenko SV, Chouteau M, Piron-Prunier F, Blugeon C, Joron M, Llaurens V. 2019. Unravelling the
934 genes forming the wing pattern supergene in the polymorphic butterfly Heliconius numata.
935 *Evodevo* **10**:16. doi:10.1186/s13227-019-0129-2
936 Schaeffer V, Althauser C, Shcherbata HR, Deng W-M, Ruohola-Baker H. 2004. Notch-Dependent
937 Fizzy-Related/Hec1/Cdh1 Expression Is Required for the Mitotic-to-Endocycle Transition in
938 Drosophila Follicle Cells. *Current Biology* **14**:630–636. doi:10.1016/j.cub.2004.03.040
939 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C,
940 Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A.
941 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682.
942 doi:10.1038/nmeth.2019
943 Shcherbata HR. 2004. The mitotic-to-endocycle switch in Drosophila follicle cells is executed by
944 Notch-dependent regulation of G1/S, G2/M and M/G1 cell-cycle transitions. *Development*
945 **131**:3169–3181. doi:10.1242/dev.01172
946 Stern DL, Orgogozo V. 2009. Is Genetic Evolution Predictable? *Science* **323**:746–751.
947 doi:10.1126/science.1158997
948 Swan A, Schüpbach T. 2007. The Cdc20 (Fzy)/Cdh1-related protein, Cort, cooperates with Fzy in
949 cyclin destruction and anaphase progression in meiosis I and II in Drosophila. *Development*
950 **134**:891–899. doi:10.1242/dev.02784
951 Turner JRG. 1981. Adaptation and Evolution in Heliconius: A Defense of NeoDarwinism. *Annual*
952 *Review of Ecology and Systematics* **12**:99–121. doi:10.1146/annurev.es.12.110181.000531
953 Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J,
954 Lewis JJ, Hines HM, Ruiz M, Salazar C, Linares M, Moreira GRP, Jiggins CD, Counterman BA,
955 McMillan WO, Papa R. 2017. Complex modular architecture around a simple toolkit of wing
956 pattern genes. *Nature Ecology & Evolution* **1**:1–12. doi:10.1038/s41559-016-0052

VanKuren NW, Massardo D, Nallu S, Kronforst MR. 2019. Butterfly mimicry polymorphisms highlight phylogenetic limits of gene re-use in the evolution of diverse adaptations. *Mol Biol Evol*. doi:10.1093/molbev/msz194

Van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ. 2016. The industrial melanism mutation in British peppered moths is a transposable element. *Nature* **534**:102–105. doi:10.1038/nature17951

van't Hof AE, Reynolds LA, Yung CJ, Cook LM, Saccheri IJ. 2019. Genetic convergence of industrial melanism in three geometrid moths. *Biology Letters* **15**:20190582. doi:10.1098/rsbl.2019.0582

Wallbank RWR, Baxter SW, Pardo-Diaz C, Hanly JJ, Martin SH, Mallet J, Dasmahapatra KK, Salazar C, Joron M, Nadeau N, McMillan WO, Jiggins CD. 2016. Evolutionary Novelty in a Butterfly Wing Pattern through Enhancer Shuffling. *PLOS Biology* **14**:e1002353. doi:10.1371/journal.pbio.1002353

Westerman EL, VanKuren NW, Massardo D, Tenger-Trolander A, Zhang W, Hill RI, Perry M, Bayala E, Barr K, Chamberlain N, Douglas TE, Buerkle N, Palmer SE, Kronforst MR. 2018. Aristaless Controls Butterfly Wing Color Variation Used in Mimicry and Mate Choice. *Curr Biol* **28**:3469-3474.e4. doi:10.1016/j.cub.2018.08.051

Zhang L, Mazo-Vargas A, Reed RD. 2017a. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *PNAS* **114**:10707–10712. doi:10.1073/pnas.1709058114

Zhang L, Mazo-Vargas A, Reed RD. 2017b. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *PNAS* **114**:10707–10712. doi:10.1073/pnas.1709058114