The gene *cortex* controls scale colour identity in *Heliconius*

Luca Livraghi[1,2], Joseph J. Hanly[1,2,3], Ling Sheng Loh[2], Anna Ren[2], Ian A. Warren[1], Carolina Concha[2], Charlotte Wright[1], Jonah M. Walker[1], Jessica Foley[2], Henry Arenas-Castro[2], Lucas Rene Brenes[2], 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

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

Funding:

This work was funded by a grant from the BBSRC to CJ and supported LL (BB/R007500/1); the National Science Foundation awards IOS-1656553 and IOS-1755329 to AM; a Wellcome Trust PhD studentship awarded to JJH, a Smithsonian Institution grant to WOM and a Balfour-Browne Trust studentship to J.M.W.

Keywords:

Evolution, wing patterning, *cortex*, *Heliconius*, cell fate, CRISPR, Lepidoptera.

Abstract

The wing patterns of butterflies are an excellent system with which to study phenotypic evolution. The incredibly diverse patterns are generated from an array of pigmented scales on a largely two-dimensional surface, resulting in a visibly tractable system for studying the evolution of pigmentation. In *Heliconius* butterflies, much of this diversity is controlled by a few genes of large effect that regulate pattern switches between races and species across a large mimetic radiation. One of these genes – *cortex* - has been repeatedly mapped in association with colour pattern evolution in both *Heliconius* and other Lepidoptera, but we lack functional data supporting its role in modulating wing patterns. Here we carried out CRISPR knock-outs in multiple *Heliconius* species and show that *cortex* is a major determinant of scale cell identity. Mutant wing clones lacking *cortex* showed shifts in colour identity, with melanic and red scales acquiring a yellow or white state. These homeotic transformations include changes in both pigmentation and scale ultrastructure, suggesting that *cortex* acts during early stages of scale cell fate specification rather than during the deployment of effector genes. In addition, mutant clones were observed across the entire wing surface, contrasting with other known *Heliconius* mimicry loci that act in specific patterns. Cortex is known as a cell-cycle regulator that modulates mitotic entry in *Drosophila*, and we found the Cortex protein to accumulate in the nuclei of the polyploid scale building cells of the butterfly wing epithelium, speculatively suggesting a connection between scale cell endocycling and colour identity. In summary, and while its molecular mode of action remains mysterious, we conclude that *cortex* played key roles in the diversification of lepidopteran wing patterns in part due to its switch-like effects in scale identity across the entire wing surface.

Introduction

Evolutionary hotspots have become a recurrent theme in evolutionary biology, whereby variation surrounding homologous loci at both micro- and macro-evolutionary scales have driven parallel cases of phenotypic change. Notably, a remarkable 138 genes have been linked to phenotypic variation in 2 or more species (GepheBase; Courtier-Orgogozo et al., 2020). In some cases, parallel adaptation has occurred through the alteration of downstream effector genes, such as pigmentation enzymes with functions clearly related to the trait under selection (e.g. *tan*, *ebony*). In other cases, upstream patterning factors are important, and these are typically either transcription factors (e.g. *optix*, *pitx1, Sox10*) or components of signalling pathways such as ligands or receptors (e.g. *WntA*, MC1R). These classes of genes influence cell fate decisions during development by modulating downstream gene regulatory networks (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 is conspicuous in its failure to conform to this paradigm is *cortex*, a gene implicated 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* (Concha et al., 2019; Mazo-Vargas et al., 2017) and two transcription factors *optix* (Lewis et al., 2019; Zhang et al., 2017) and *aristaless1* (Westerman et al., 2018). The fourth is *cortex*, 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 (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, that can be readily applied to other major *Heliconius* patterning loci.

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 1a) and the polymorphism in yellow, white, black and orange elements in *H. numata*, 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 *Bicyclus anynana* and *Papilio clytia* (Beldade et al., 2009; 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.

While *cortex* remains the most likely candidate driving yellow and white scale evolution in *Heliconius*, other genes at the locus may also be playing a role in establishing scale colour identity. Most notably, the genes *domeless* (*dome*), a JAK-STAT pathway receptor and *washout* (*wash*), a cytoskeleton regulator, which also show associations with colour pattern phenotypes in *H. melpomene* and *H. numata* (Nadeau et al., 2016; Saenko et al., 2019). These genes neighbour a non-coding region linked to the Bigeye mutation in *B. anynana,* have been implicated in shaping eyespot patterns in CRISPR mutagenesis experiments, and are within an interval of around 100kb which also includes *cortex* (Beldade et al., 2009; Lopes da Silva, 2015). It is thus possible that multiple linked genes are contributing to the evolution of wing patterning across Lepidoptera (Joron et al., 2006; Saenko et al., 2019).

While fantastically diverse, most of the pattern variation in *Heliconius* is created by the differences in the distribution of only 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., 1987). 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 colour is the result of the scale cell morphology (i.e. structural) (Gilbert et al., 1987). In contrast, Type II scale cells are pigmented with melanin and Type III scale cells contain the red ommochrome pigments xanthommatin and dihydroxanthommatin.

Here we focus on the role of *cortex* 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*, *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 races of *H. melpomene* and *H. erato* (Figure 1b).

Despite the fact that *cortex* does not follow the prevailing paradigm of patterning loci, we demonstrate for the first time 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 the phenotypic effects of *cortex* extend across the fore- and hindwing surface. 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.

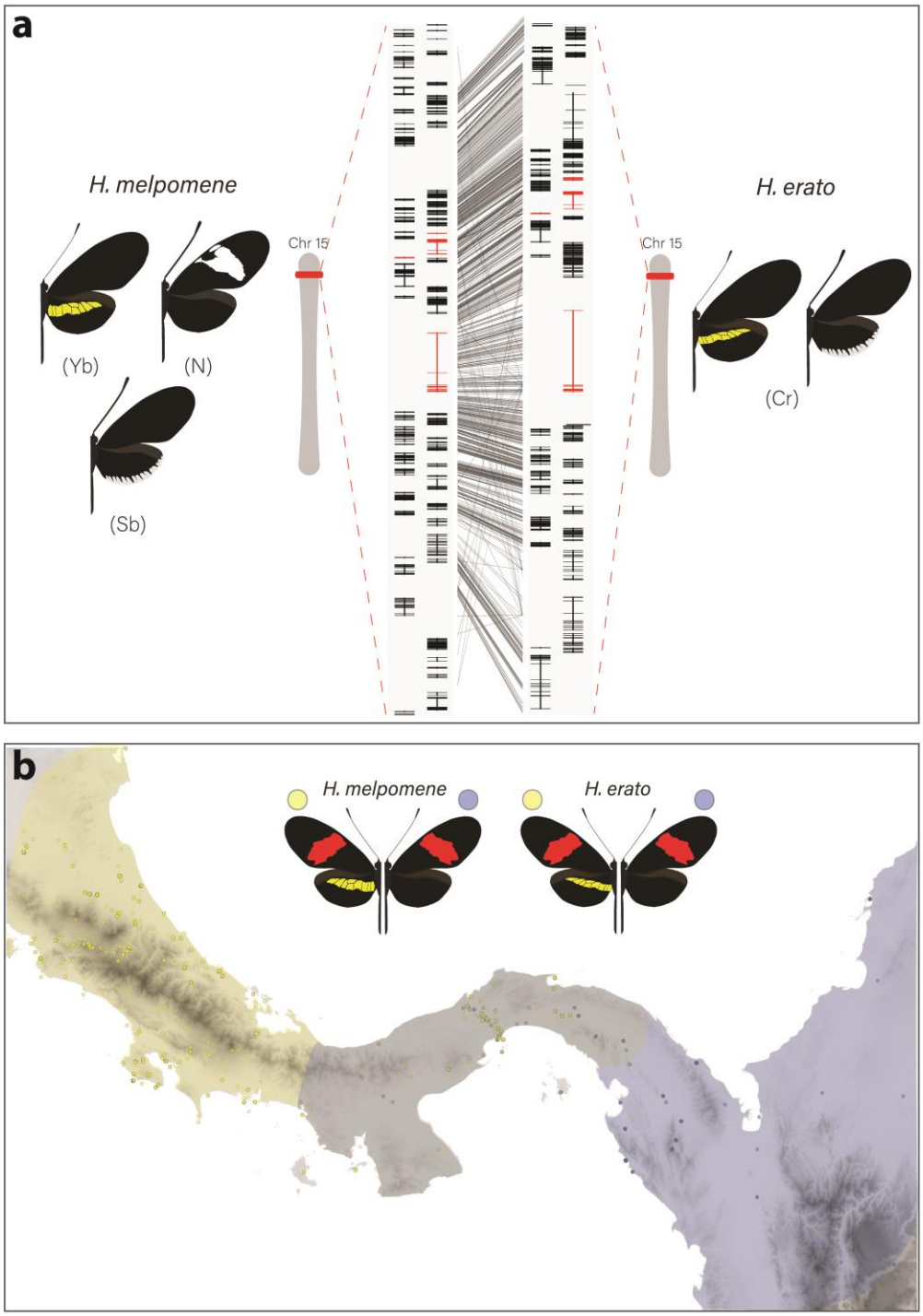


Figure 1 – Ranges of *Heliconius* butterflies differing at Yb phenotypes in Central America and associated loci

(a) Homologous loci in both species are associated with variation in yellow and white patterns between races. In *H. melpomene* three tightly linked genetic elements located at chromosome 15 control variation for hindwing yellow bar, forewing band and white margin elements (Yb, N and Sb respectively) while in *H. erato* variation has been mapped to one element (Cr). Genes previously associated with wing patterning differences in Lepidoptera are highlighted in red within a specific region of chromosome 15 (from bottom up; *cortex*, *domeless-truncated*, *domeless* and *washout*) and alignment between the two co-mimetic species at the locus is shown (grey lines, 95% alignment identity). **(b)** Focal co-mimetic races of *Heliconius erato* and *Heliconius melpomene* used in this study, differing for 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 races, blue: black hindwing races, grey: range overlap.

Results

RNA-seq and reannotation of key intervals reveals the presence of duplications and bi-cistronic transcription of candidate genes

In order to identify genes associated with differences in yellow pattern elements, we performed differential gene expression (DGE) analysis using developing wings sampled from colour pattern races in *H. erato* and *H. melpomene* differing only in the presence or absence of the hindwing yellow bar (Figure 1b and Figure 2a). 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 races in each of the two species, with hindwings divided into two parts for the pupal stages (Figure 2a). 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 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 showed differential expression and are likely targets of selection: *cortex*, *dome* and *wash* (Figure 2c). This led us to further explore the annotation of these genes prior to further analysis.

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). Independent tandem duplications of *dome* have occurred in several other Lepidoptera. 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.

When examining the *RNA-seq* reads mapping to the *dome* and *wash* genes, we observed several individual reads splicing over the 5' UTR of *wash* and into the coding region of *dome*. It was not possible to unambiguously assign reads that map to this overlapping portion of the annotation to either gene, suggesting the possibility that *dome/wash* are transcribed as a single, bi-cistronic transcript. To look for further evidence of co-transcription, we searched the Transcription Shotgun Assembly (TSA) sequence archive on NCBI for assembled transcripts containing the open reading frames (ORFs) of both genes in other Lepidoptera (Supplementary File 2 – Figure S2). We found several instances where ORFs encoding for both *dome* and *wash* can be found in a single transcript, suggesting bi-cistronic transcription is a conserved feature across butterflies. Furthermore, an examination of published *ATAC-seq* peaks (Lewis et al., 2019), shows the presence of a single promoter at the start of *dome* for *H. erato lativitta*, suggesting both genes share a single transcription start site (Supplementary File 2 – Figure S2). Given this result, we repeated the DGE analysis with *dome/wash* as a single annotation.

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

149 *RNA-seq* data show *cortex* transcripts were most abundant in 5th instar larvae, almost depleted in Day 1
150 pupae, but were again detected at relatively high levels in Day 2 pupae in *H. melpomene*, suggesting
151 dynamic expression in this species (Figure 2b). In *H. erato*, *cortex* transcripts are found in high
152 abundance in 5th instar larvae but are almost depleted in Day 1 and Day 2 pupae. Both *dome* paralogs
153 remain relatively constant in terms of expression across all stages in *H. melpomene* whereas *dome-trunc*
154 expression increases in pupal stages in *H. erato*. *Dome/wash* transcripts are detected in relatively low
155 and constant amounts in both species.

156 The two species were analysed separately, with both showing only *cortex* and *dome/wash* as
157 significantly differentially expressed between morphs among the 47 genes in the candidate region, with
158 *cortex* differential expression occurring earlier in development. In fifth instar larvae, *cortex* is
159 differentially expressed in both species between the two colour pattern races, with *cortex* showing the
160 highest adjusted *p*-value for any gene in the genome at this stage in *H. erato* (Figure 2c). Interestingly,
161 *cortex* transcripts were differentially expressed in opposite directions in the two species, with higher
162 expression in the melanic hindwing race in *H. melpomene*, and in the yellow banded race in *H. erato*.
163 This pattern is reversed for *dome/wash* in Day 1 pupae, where a statistically higher proportion of
164 transcripts are detected in *H. melpomene rosina* (yellow), and in *H. erato hydara* (melanic). No
165 differential expression of these genes was found at Day 2 pupae. In order to confirm this inverted pattern
166 was not due to a sampling error, we performed a diagnostic SNP analysis by correlating coding SNPs
167 found within protein coding genes at the *cortex* locus from whole genome sequence data to the
168 corresponding *RNA-seq* datasets (Supplementary File 3 – Tables S3.1 and S3.2).

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

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

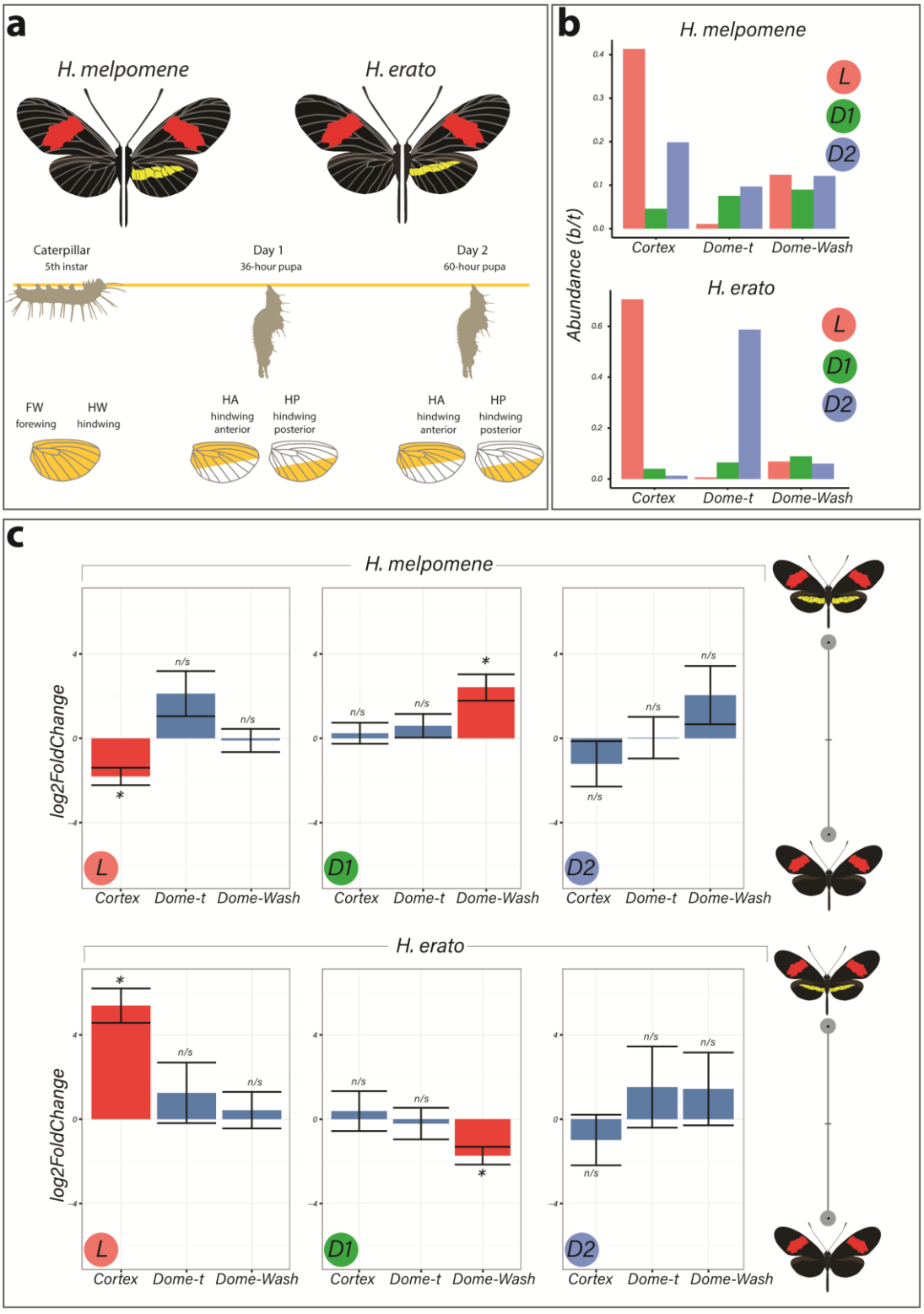


Figure 2 – Differential expression of genes at Chromosome 15 implicate *cortex* as most likely candidate driving yellow bar differences

(a) Hindwing tissue from co-mimetic races 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. *Cortex* expression decreases from larval to pupal stages *domeless-truncated* expression increases, whereas *domeless*/*washout* stay relatively constant at all three stages. **(c)** Log_2FoldChange for the genes *cortex*, *domeless-truncated*, *domeless*/*washout* across developmental stages. Comparisons are for whole wing discs (Larvae, L) and for contrast C for pupae (D1 and D2; see Supplementary File 4: Figure S4.3 for depiction of contrasts analysed).

Cortex transcripts localise distally in 5th instar larvae

Two studies have reported that *cortex* mRNA expression correlates with melanic patch 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). In *H. erato demophoon*, expression was strongest at the intervein midline, but extends across vein compartments covering the distal portion of both forewing and hindwing. By contrast, in *H. erato hydara*, *cortex* transcripts are more strongly localised to the intervein midline forming a distally localised intervein expression domain.

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*. In *H. melpomene melpomene*, hindwing *cortex* expression extends across most of the hindwing, and does not appear to be restricted to the intervein midline.

Given that *cortex* has been implicated in modulating wing patterns in many divergent lepidoptera, we examined localisation in a *Heliconius* species displaying distinct patterns: *H. hecale melicerta* (Figure 3). 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.

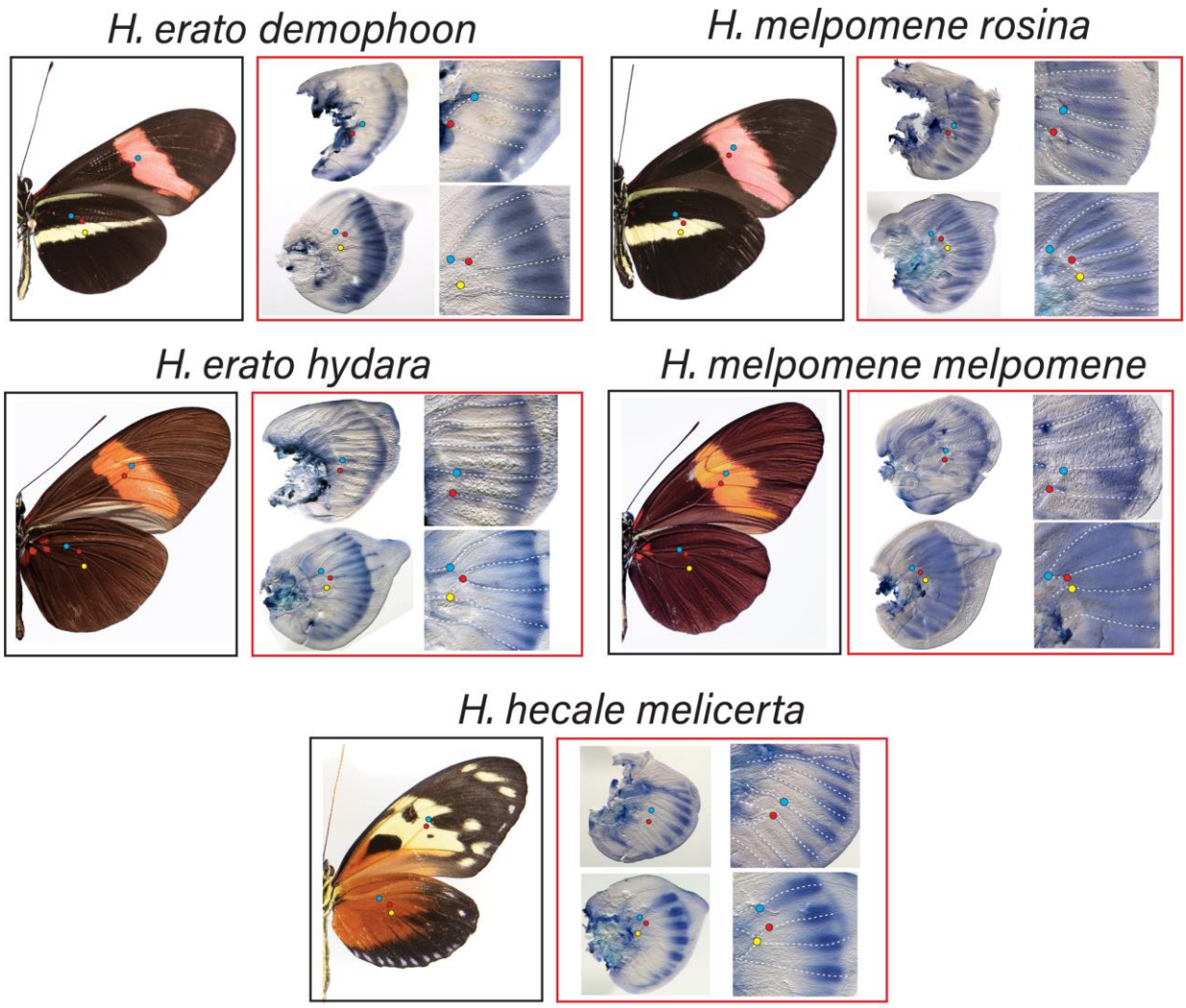


Figure 3 – Expression of *cortex* transcripts in *H. melpomene*, *H. erato* and *H. hecale* 5th instar wing discs

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*, whereas it is more strongly localised to the intervein midline in *H. erato hydara*. In *H. melpomene rosina*, *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* extends more proximally. Coloured dots represent homologous vein landmarks across the wings.

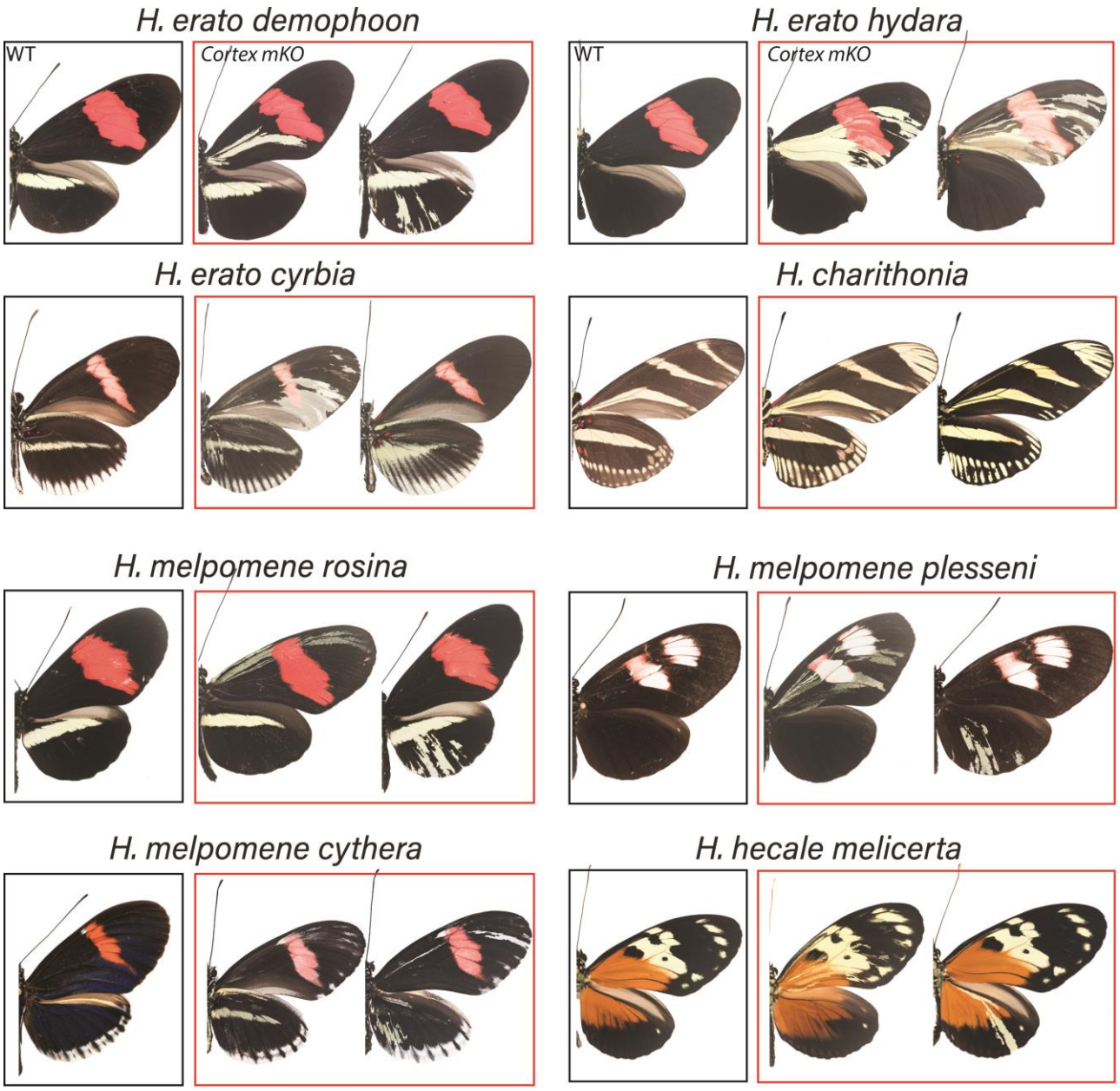
Cortex establishes Type II and III scale identity in *Heliconius* butterflies

To assay the function of *cortex* during wing development, we used CRISPR/Cas9 to generate G_0 somatic mosaic mutants (crispants) (Mazo-Vargas et al., 2017; Zhang et al., 2017). We targeted multiple exons using a combination of different guides and genotyped the resulting mutants through PCR amplification, cloning and Sanger sequencing (Supplementary File 6 – Figure S6). Overall KO efficiency was low when compared to similar studies in *Heliconius* (Concha et al., 2019; Mazo-Vargas et al., 2017), with observed wing phenotype to hatched eggs ratios ranging from 0.3% to 4.8%. Lethality

was also high, with hatched to adult ratios ranging from 8.1% to 29.8% (Supplementary File 7 – Table S7.1).

Targeting of the *cortex* gene in *H. erato* produced patches of ectopic yellow and white scales spanning regions across both forewings and hindwings (Figure 4 and Supplementary File 8 – Figures S8.1-S8.7). Both colour pattern races 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 the co-mimetic morphs in *H. melpomene*, with mutant clones affecting both distal and proximal regions in forewings and hindwings. In *H. erato hydara*, we recovered a mutant individual where clones spanned the dorsal forewing band. 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 (Figure 5 and Supplementary File 9 – Figure S9).

As this locus has been associated with differences in white hindwing margin phenotypes (Jiggins and McMillan, 1997) (Figure 1b), we also targeted *cortex* in mimetic races showing this phenotype, *H. erato cyrbia* and *H. melpomene cythera*. Mutant scales in these colour pattern races were also localised across both wing surfaces, with both white and yellow ectopic scales. In these races, a positional effect was observed, where ectopic scales in the forewing and anterior compartment of the hindwing shifted to yellow, and posterior hindwing scales became white (Figure 4 and Supplementary File 9 – Figure S9). This positional effect likely reflects differential uptake of the yellow pigment 3-OHK across the wing surface (Reed et al., 2008). For one individual of *H. erato cyrbia*, clones also extended across the red band where a shift to white scales was observed, as in *H. erato hydara*.



242

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

Phenotypes of *cortex* mKO across *Heliconius* species and 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. A positional effect is observed in some races, where ectopic Type I scales are either white or yellow depending on their position along the wing (*e.g H. erato cyrbia*). Ectopic Type I scales can be induced from both melanic and red scales, switching to either white or yellow depending on wing position and race. Boundaries between Wild-type (WT) to mutant scales are highlighted (dotted white line).

243 To further test the conservation of *cortex* function across the *Heliconius* radiation, we knocked out
244 *cortex* in *H. charithonia* and *H. hecale melicerta*, outgroups to *H. erato* and *H. melpomene* respectively.
245 Again, ectopic yellow and white scales appeared throughout the wing surface in both species,
246 suggesting conserved function with respect to scale development among *Heliconius* butterflies. In *H.*

hecale melicerta, we also recovered a mutant where we saw transformation from orange ommochrome scales to yellow.

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 race and wing position (Figure 5). 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 (Figure 5 and Supplementary File 9 - Figure S9).

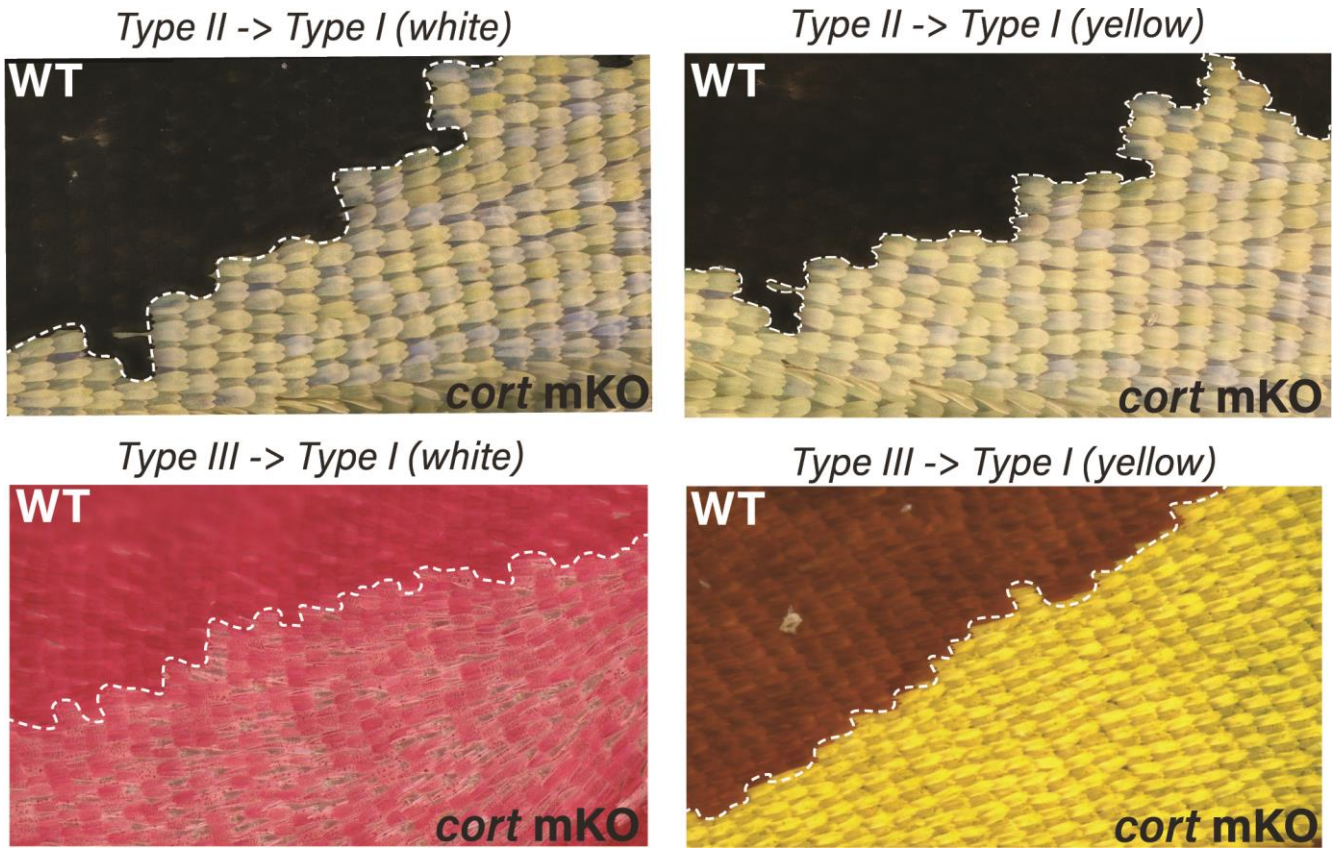


Figure 5 – CRISPR KOs induce Type I scale identity

Ectopic Type I scales can be induced from both melanic and red scales, switching to either white or yellow depending on wing position and race. Boundaries between Wild-type (WT) to mutant scales are highlighted (dotted white line).

Nuclear localization of Cortex 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 long 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 (Figure 6 and supplementary File 9 – Figure S9).

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 element (Figure 6). This nuclear localization overlapped with DNA, also included a strong signal in the large nucleoli of both the polyploid scale building cells, and their adjacent, non-polyploid epithelial cells (Greenstein, 1972). Following previous reports suggesting a correlation between pigmentation state and ploidy level (Cho and Nijhout, 2013; Henke and Pohley, 1952; Iwata and Otaki, 2016), we tested if nuclear volume or nucleoli number would associate with the yellow band, but failed to find a consistent pattern in the distribution of Cortex protein (Figure 6 and Supplementary File 10 - Figure S10). However, currently we cannot rule an association of Cortex protein with colour pattern elements at other developmental stages, and given the apparent dynamic nature of *cortex* expression, a more precise developmental time series will be required to make more conclusive statements.

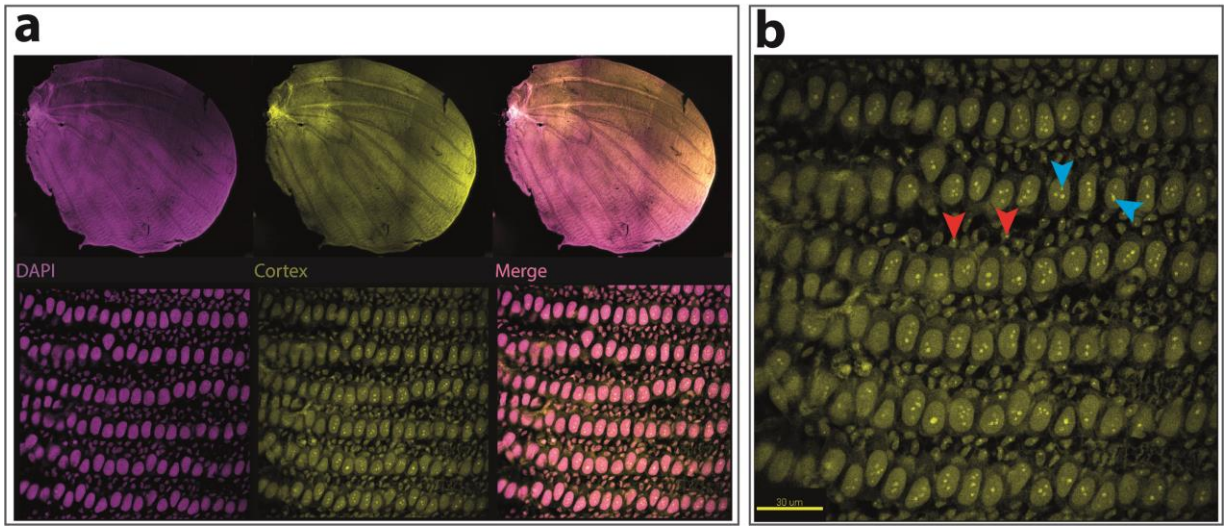


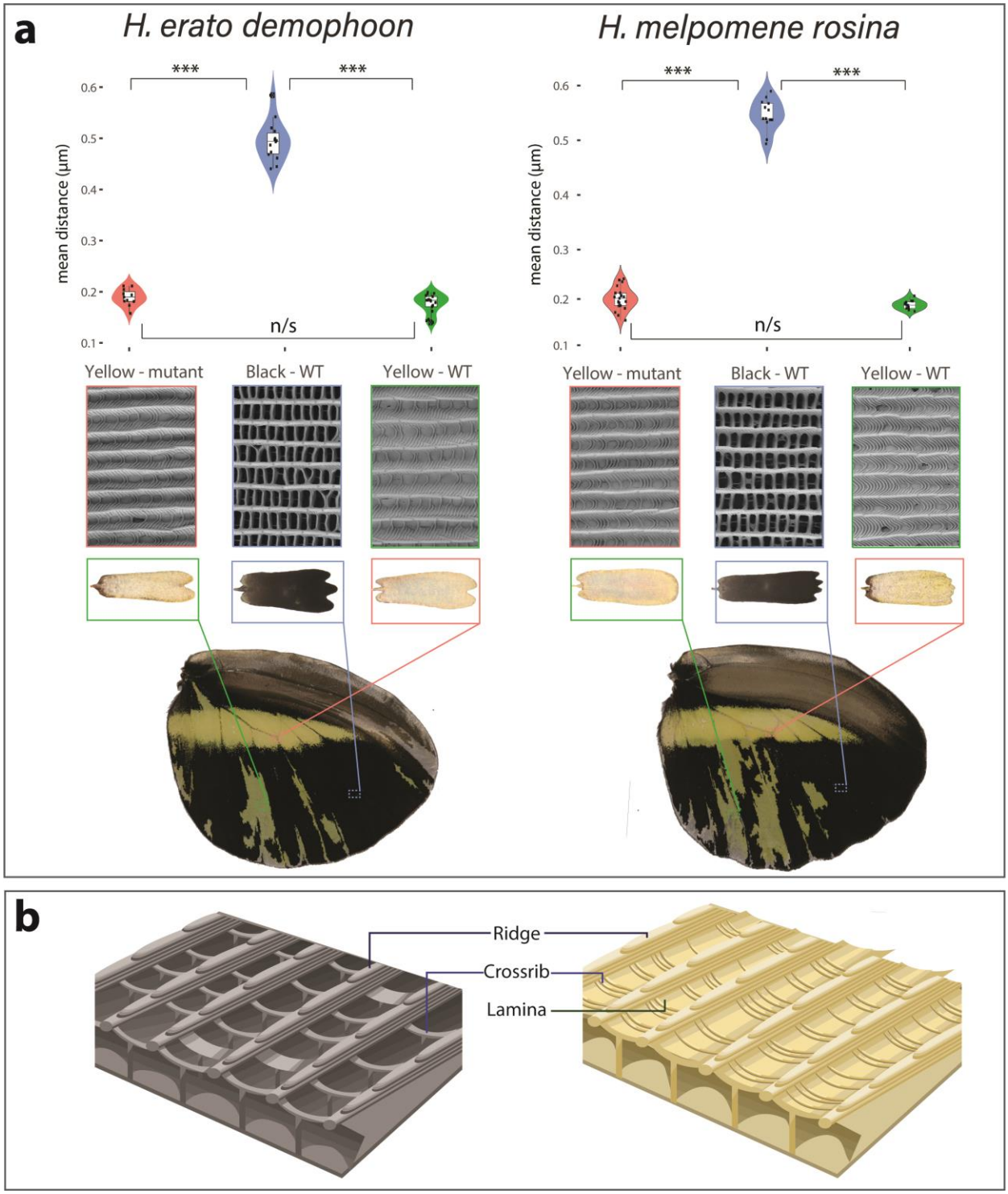
Figure 6 – Cortex protein localises throughout pupal hindwings in *H. erato demophoon*

(a) Cortex immnostaining reveals presence of Cortex across the hindwing of H erato demophoon. DAPI (left) Cortex (middle) and merged channels (right) are shown. **(b)** Cortex localises as puncta in the developing pupal cells. Multiple puncta per cell are visible in the large polyploid nuclei (blue arrows) while single puncta localise to the uninuclear epithelium below (red arrows). Scale bar = 30µm.

Cortex KO causes 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; Zhang et al., 2017). With this in mind, we tested whether ectopic yellow/white scales were accompanied by structural homeosis using Scanning Electron Microscopy. To account for known positional effects on scale structure we compared wild-type and mutant scales from homologous locations across the wing surface.

Ultrastructural differences are consistent with homeosis in *cortex* mutant scales in both *H. melpomene* and *H. erato* (Figure 7). Cross-rib distance is the same between yellow wild-type and *cortex* mutant scales, and significantly different between distally located wild-type black scales. A similar relationship was observed for scale length in both species, but inter-ridge distance and scale width was consistent with homeosis only in *H. melpomene* (Supplementary File 11 – Figure S11). A consistent difference between all Type I scales (mutant and wild-type) is the presence of a lamina covering the inter-ridge space (Figure 7b), suggesting this structure is an important morphological feature of yellow/white scales (Matsuoka and Monteiro, 2018), and that *cortex* is necessary for the differentiation of lamellar tissue in *Heliconius* scales.



297

Figure 7 – SEM reveals structural homeosis is induced in *cortex* KO scales.

Structural homeosis is induced in *cortex* KO scales in both *H. melpomene* and *H erato*. Mutant and wild-type scale comparisons from homologous wing positions are shown, illustrating clear ultrastructural homeosis between wild-type and KO yellow scales. Mean cross-rib distance between wild-type and mutant yellow scales is not significantly different, while significantly different between both wild-type yellow and mutant yellow with wild-type black scales (Wilcoxon test, *** indicates $p<0.001$).

298

Discussion:

Cortex is a key scale cell specification gene

The genetic locus containing the gene *cortex* represents a remarkable case of parallel evolution, where repeated and independent mutations surrounding the gene are associated with shifts in scale pigmentation state in at least 8 divergent species of Lepidoptera (Beldade et al., 2009; 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. Moreover, 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). Cortex 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.

Heliconius wing patterning is controlled by interactions among major patterning genes

Functional knockouts now exist for all the 4 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

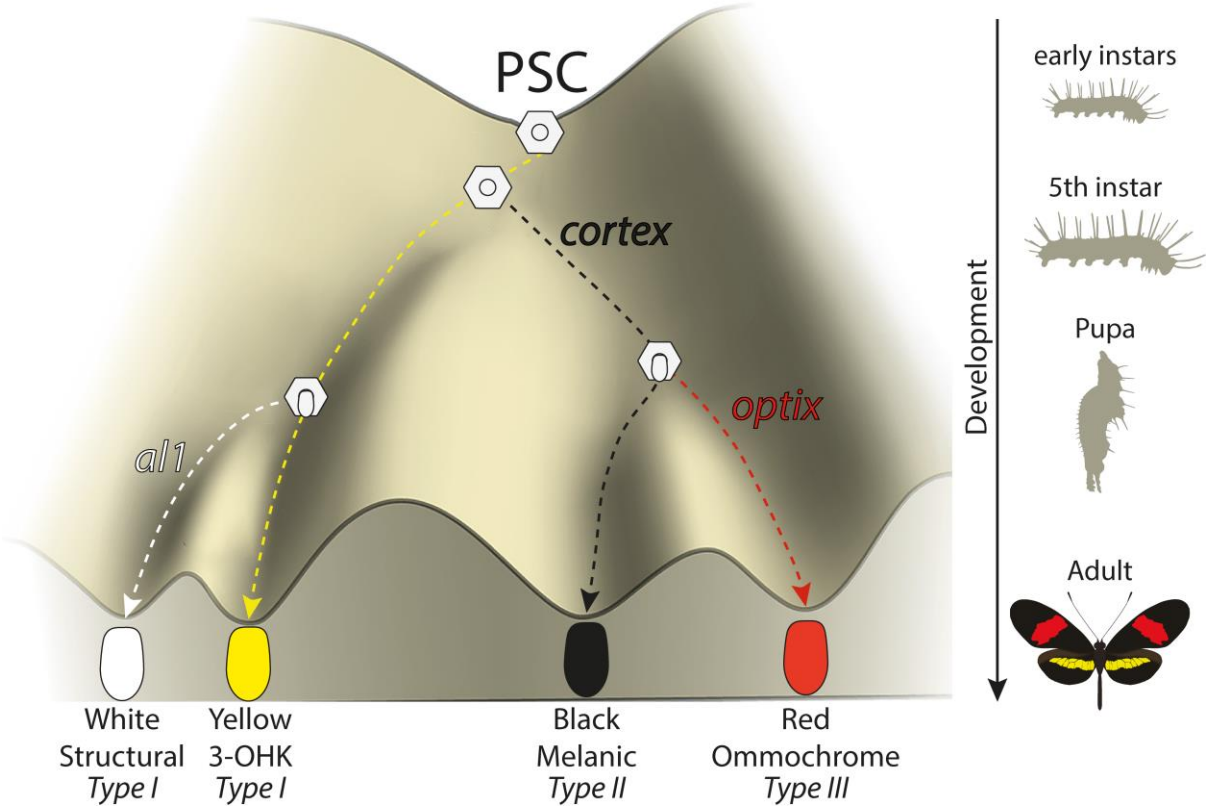
regulators of scale cell identity. Both are upregulated early in wing development and have broad effects on pattern variation (Concha et al., 2019; Nadeau et al., 2016). The signalling molecule *WntA* modulates forewing band shape in *Heliconius* by delineating boundaries around patterns elements, and 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 either that *cortex* is required as an inductive signal to allow *WntA* to signal further melanisation, or that two, independent ways to melanise a scale are available to the developing wing. The latter hypothesis is supported by 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 suggests 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.

Under a simple model (Figure 8), *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

368 (Supplementary File 8 and 9). This pigmentation state is likely controlled by differences in *al1*
369 expression varying between wing sections in different ways across races.



370

Figure 8 – 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).

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

available on 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), using the GLMl;

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

(Compartments: Anterior Hindwing (HA), Posterior Hindwing (HPo)). *H. melpomene* and *H. erato* were analysed separately; homology between genes was determined by reciprocal BLAST. Contrasts were then extracted for comparison of race, compartment, and race given the effect of compartment, alternating the race used as the base level.

***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. Image processing was done using FIJI plugins Trainable Weka Segmentation and BioVoxxel (Arganda-Carreras et al., 2017; Brocher, Jan, 2014; Schindelin et al., 2012). Manual tracing of nuclei was input for machine learning and processing of images to obtain final thresholded images, then an overlay of Cortex puncta with DAPI nuclei staining identified regions of nuclei

475 containing Cortex puncta. Spatial analysis of image data was conducted using R software 4.0.0 package
476 Spatstat (Baddeley and Turner, 2005).

477 **CRISPR/Cas9 genome editing**

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

489 **Genotyping**

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

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

501 **Morphometrics analysis**

502 Morphometric measurements of scale widths and ridge distances were carried out on between 10 and
503 20 scales of each type, using a custom semi-automated R pipeline that derives ultrastructural
504 parameters from large SEM images (Day et al., 2019). Briefly, ridge spacing was assessed by Fourier
505 transforming intensity traces of the ridges acquired from the FIJI software (Schindelin et al.,
506 2012). Scale width was directly measured in FIJI by manually tracing a line, orthogonal to the ridges,
507 at the section of maximal width.

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., I.A.W., C.C., C.W., J.M.W., J.F., H.A.C., L.R.B. performed research. L.L wrote the paper.

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, Steven Van Belleghem 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 forenabling SEM, and in particular Christine Brantner and Anastas Popratiloff for their technical assistance

Competing interests

The authors declare no competing interests.

References

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169. doi:10.1093/bioinformatics/btu638

Arganda-Carreras I, Kaynig V, Rueden C, Eliceiri KW, Schindelin J, Cardona A, Sebastian Seung H. 2017. Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**:2424–2426. doi:10.1093/bioinformatics/btx180

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

Baddeley A, Turner R. 2005. spatstat: An R Package for Analyzing Spatial Point Patterns. *Journal of Statistical Software* **12**:1–42. doi:10.18637/jss.v012.i06

Bassett A, Liu J-L. 2014. CRISPR/Cas9 mediated genome engineering in Drosophila. *Methods* **69**:128–136. doi:10.1016/j.ymeth.2014.02.019

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

538 Brocher, Jan. 2014. Qualitative and Quantitative Evaluation of Two New Histogram Limiting
539 Binarization Algorithms. *Computer Science Journals* **8**:30–48.
540 Brown KS. 1981. The Biology of Heliconius and Related Genera. *Annual Review of Entomology*
541 **26**:427–457. doi:10.1146/annurev.en.26.010181.002235
542 Challi RJ, Kumar S, Dasmahapatra KK, Jiggins CD, Blaxter M. 2016. Lepbase: the Lepidopteran
543 genome database. *bioRxiv* 056994. doi:10.1101/056994
544 Cho EH, Nijhout HF. 2013. Development of polyploidy of scale-building cells in the wings of Manduca
545 sexta. *Arthropod Struct Dev* **42**:37–46. doi:10.1016/j.asd.2012.09.003
546 Chu T, Henrion G, Haegeli V, Strickland S. 2001. Cortex, a Drosophila gene required to complete
547 oocyte meiosis, is a member of the Cdc20/fizzy protein family. *Genesis* **29**:141–152.
548 doi:10.1002/gene.1017
549 Concha C, Wallbank RWR, Hanly JJ, Fenner J, Livraghi L, Rivera ES, Paulo DF, Arias C, Vargas M,
550 Sanjeev M, Morrison C, Tian D, Aguirre P, Ferrara S, Foley J, Pardo-Diaz C, Salazar C, Linares
551 M, Massardo D, Counterman BA, Scott MJ, Jiggins CD, Papa R, Martin A, McMillan WO. 2019.
552 Interplay between Developmental Flexibility and Determinism in the Evolution of Mimetic
553 Heliconius Wing Patterns. *Current Biology* S0960982219313168.
554 doi:10.1016/j.cub.2019.10.010
555 Courtier-Orgogozo V, Arnoult L, Prigent SR, Wiltgen S, Martin A. 2020. Gephebase, a database of
556 genotype–phenotype relationships for natural and domesticated variation in Eukaryotes.
557 *Nucleic Acids Res* **48**:D696–D703. doi:10.1093/nar/gkz796
558 Davey JW, Chouteau M, Barker SL, Maroja L, Baxter SW, Simpson F, Merrill RM, Joron M, Mallet J,
559 Dasmahapatra KK, Jiggins CD. 2016. Major Improvements to the Heliconius melpomene
560 Genome Assembly Used to Confirm 10 Chromosome Fusion Events in 6 Million Years of
561 Butterfly Evolution. *G3: Genes, Genomes, Genetics* **6**:695–708. doi:10.1534/g3.115.023655
562 Day CR, Hanly JJ, Ren A, Martin A. 2019. Sub-micrometer insights into the cytoskeletal dynamics and
563 ultrastructural diversity of butterfly wing scales. *Dev Dyn* **248**:657–670. doi:10.1002/dvdy.63
564 Enciso-Romero J, Pardo-Díaz C, Martin SH, Arias CF, Linares M, McMillan WO, Jiggins CD, Salazar C.
565 2017. Evolution of novel mimicry rings facilitated by adaptive introgression in tropical
566 butterflies. *Molecular Ecology* **26**:5160–5172. doi:10.1111/mec.14277
567 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
568 analysis tools APIs in 2019. *Nucleic Acids Res* **47**:W636–W641. doi:10.1093/nar/gkz268
569 Finkbeiner SD, Fishman DA, Osorio D, Briscoe AD. 2017. Ultraviolet and yellow reflectance but not
570 fluorescence is important for visual discrimination of conspecifics by Heliconius erato.
571 *Journal of Experimental Biology* **220**:1267–1276. doi:10.1242/jeb.153593

572 Gilbert LE, Forrest HS, Schultz TD, Harvey DJ. 1987. Correlations of ultrastructure and pigmentation
573 suggest how genes control development of wing scales of Heliconius butterflies. *The Journal*
574 *of research on the Lepidoptera (USA)*.
575 Greenstein ME. 1972. The ultrastructure of developing wings in the giant silkmoth, Hyalophora
576 cecropia. II. Scale-forming and socket-forming cells. *J Morphol* **136**:23–51.
577 doi:10.1002/jmor.1051360103
578 Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and
579 Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of
580 PhyML 3.0. *Systematic Biology* **59**:307–321. doi:10.1093/sysbio/syq010
581 Hanly JJ, Wallbank RWR, McMillan WO, Jiggins CD. 2019. Conservation and flexibility in the gene
582 regulatory landscape of heliconiine butterfly wings. *EvoDevo* **10**:15. doi:10.1186/s13227-
583 019-0127-4
584 Henke K, Pohley H-J. 1952. Differentielle Zellteilungen und Polyploidie bei der Schuppenbildung der
585 Mehlmotte Ephestia kühniella Z. *Zeitschrift für Naturforschung B* **7**:65–79. doi:10.1515/znb-
586 1952-0201
587 Huber B, Whibley A, Poul YL, Navarro N, Martin A, Baxter S, Shah A, Gilles B, Wirth T, McMillan WO,
588 Joron M. 2015. Conservatism and novelty in the genetic architecture of adaptation in
589 Heliconius butterflies. *Heredity* **114**:515–524. doi:10.1038/hdy.2015.22
590 Ito K, Katsuma S, Kuwazaki S, Jouraku A, Fujimoto T, Sahara K, Yasukochi Y, Yamamoto K, Tabunoki H,
591 Yokoyama T, Kadono-Okuda K, Shimada T. 2016. Mapping and recombination analysis of two
592 moth colour mutations, Black moth and Wild wing spot, in the silkworm Bombyx mori.
593 *Heredity* **116**:52–59. doi:10.1038/hdy.2015.69
594 Iwata M, Otaki JM. 2016. Spatial patterns of correlated scale size and scale color in relation to color
595 pattern elements in butterfly wings. *Journal of Insect Physiology* **85**:32–45.
596 doi:10.1016/j.jinsphys.2015.11.013
597 Jiggins CD. 2017. The Ecology and Evolution of Heliconius Butterflies. Oxford University Press.
598 Jiggins CD, McMillan WO. 1997. The genetic basis of an adaptive radiation: warning colour in two
599 Heliconius species. *Proc Biol Sci* **264**:1167–1175. doi:10.1098/rspb.1997.0161
600 Joron M, Papa R, Beltrán M, Chamberlain N, Mavárez J, Baxter S, Abanto M, Bermingham E,
601 Humphray SJ, Rogers J, Beasley H, Barlow K, H. ffrench-Constant R, Mallet J, McMillan WO,
602 Jiggins CD. 2006. A Conserved Supergene Locus Controls Colour Pattern Diversity in
603 Heliconius Butterflies. *PLoS Biol* **4**:e303. doi:10.1371/journal.pbio.0040303
604 Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S,
605 Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated

606 and extendable desktop software platform for the organization and analysis of sequence
607 data. *Bioinformatics* **28**:1647–1649. doi:10.1093/bioinformatics/bts199
608 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and
609 genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**:907–915.
610 doi:10.1038/s41587-019-0201-4
611 Koch PB. 1993. Production of [14C]-Labeled 3-Hydroxy-L-Kynurenine in a Butterfly, Heliconius
612 charitonia L. (Heliconidae), and Precursor Studies in Butterfly Wing Ommatins. *Pigment Cell*
613 *Research* **6**:85–90. doi:10.1111/j.1600-0749.1993.tb00586.x
614 Kronforst MR, Papa R. 2015. The Functional Basis of Wing Patterning in Heliconius Butterflies: The
615 Molecules Behind Mimicry. *Genetics* **200**:1–19. doi:10.1534/genetics.114.172387
616 Lamas G, editor. 2004. Atlas Of Neotropical Lepidoptera: Checklist Pt. 4a Hesperioidea-papilionoidea.
617 Gainesville: Scientific Pub.
618 Lewis JJ, Geltman RC, Pollak PC, Rondem KE, Belleghem SMV, Hubisz MJ, Munn PR, Zhang L, Benson
619 C, Mazo-Vargas A, Danko CG, Counterman BA, Papa R, Reed RD. 2019. Parallel evolution of
620 ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. *PNAS* **116**:24174–
621 24183. doi:10.1073/pnas.1907068116
622 Lopes da Silva C. 2015. Genetics of diversification: a hotspot locus for pigmentation evolution
623 (Masters Thesis). Lisboa: Universidad de Lisboa.
624 Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq
625 data with DESeq2. *Genome Biol* **15**:550. doi:10.1186/s13059-014-0550-8
626 Luo S, Tang M, Frandsen PB, Stewart RJ, Zhou X. 2018. The genome of an underwater architect, the
627 caddisfly Stenopsyche tienmushanensis Hwang (Insecta: Trichoptera). *Gigascience* **7**.
628 doi:10.1093/gigascience/giy143
629 Martin A, Courtier-Orgogozo V. 2017. Morphological Evolution Repeatedly Caused by Mutations in
630 Signaling Ligand Genes In: Sekimura T, Nijhout HF, editors. Diversity and Evolution of
631 Butterfly Wing Patterns: An Integrative Approach. Singapore: Springer. pp. 59–87.
632 doi:10.1007/978-981-10-4956-9_4
633 Martin A, Papa R, Nadeau NJ, Hill RI, Counterman BA, Halder G, Jiggins CD, Kronforst MR, Long AD,
634 McMillan WO, Reed RD. 2012. Diversification of complex butterfly wing patterns by
635 repeated regulatory evolution of a Wnt ligand. *Proc Natl Acad Sci USA* **109**:12632–12637.
636 doi:10.1073/pnas.1204800109
637 Martin A, Reed RD. 2014. Wnt signaling underlies evolution and development of the butterfly wing
638 pattern symmetry systems. *Developmental Biology* **395**:367–378.
639 doi:10.1016/j.ydbio.2014.08.031

640 Matsuoka Y, Monteiro A. 2018. Melanin Pathway Genes Regulate Color and Morphology of Butterfly
641 Wing Scales. *Cell Reports* **24**:56–65. doi:10.1016/j.celrep.2018.05.092
642 Mazo-Vargas A, Concha C, Livraghi L, Massardo D, Wallbank RWR, Zhang L, Papador JD, Martinez-
643 Najera D, Jiggins CD, Kronforst MR, Breuker CJ, Reed RD, Patel NH, McMillan WO, Martin A.
644 2017. Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity.
645 *PNAS* **114**:10701–10706. doi:10.1073/pnas.1708149114
646 Moest M, Belleghem SMV, James JE, Salazar C, Martin SH, Barker SL, Moreira GRP, Mérot C, Joron M,
647 Nadeau NJ, Steiner FM, Jiggins CD. 2020. Selective sweeps on novel and introgressed
648 variation shape mimicry loci in a butterfly adaptive radiation. *PLOS Biology* **18**:e3000597.
649 doi:10.1371/journal.pbio.3000597
650 Nadeau NJ. 2016. Genes controlling mimetic colour pattern variation in butterflies. *Current Opinion*
651 *in Insect Science*, Global change biology * Molecular physiology **17**:24–31.
652 doi:10.1016/j.cois.2016.05.013
653 Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko SV, Wallbank RWR, Wu GC, Maroja L,
654 Ferguson L, Hanly JJ, Hines H, Salazar C, Merrill RM, Dowling AJ, ffrench-Constant RH,
655 Llaurens V, Joron M, McMillan WO, Jiggins CD. 2016. The gene cortex controls mimicry and
656 crypsis in butterflies and moths. *Nature* **534**:106–110. doi:10.1038/nature17961
657 Pesin JA, Orr-Weaver TL. 2007. Developmental Role and Regulation of cortex, a Meiosis-Specific
658 Anaphase-Promoting Complex/Cyclosome Activator. *PLOS Genetics* **3**:e202.
659 doi:10.1371/journal.pgen.0030202
660 Prud'homme B, Gompel N, Carroll SB. 2007. Emerging principles of regulatory evolution. *PNAS*
661 **104**:8605–8612. doi:10.1073/pnas.0700488104
662 Raff JW, Jeffers K, Huang J. 2002. The roles of Fzy/Cdc20 and Fzr/Cdh1 in regulating the destruction
663 of cyclin B in space and time. *J Cell Biol* **157**:1139–1149. doi:10.1083/jcb.200203035
664 Reed RD, McMillan WO, Nagy LM. 2008. Gene expression underlying adaptive variation in Heliconius
665 wing patterns: non-modular regulation of overlapping cinnabar and vermilion prepatterns.
666 *Proc Biol Sci* **275**:37–46. doi:10.1098/rspb.2007.1115
667 Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, Jiggins CD, Chamberlain NL,
668 Kronforst MR, Chen R, Halder G, Nijhout HF, McMillan WO. 2011. optix drives the repeated
669 convergent evolution of butterfly wing pattern mimicry. *Science* **333**:1137–1141.
670 doi:10.1126/science.1208227
671 Rosser N, Phillimore AB, Huertas B, Willmott KR, Mallet J. 2012. Testing historical explanations for
672 gradients in species richness in heliconiine butterflies of tropical America: DIVERSIFICATION

673 OF BUTTERFLIES. *Biological Journal of the Linnean Society* **105**:479–497. doi:10.1111/j.1095-
674 8312.2011.01814.x
675 Saenko SV, Chouteau M, Piron-Prunier F, Blugeon C, Joron M, Llaurens V. 2019. Unravelling the
676 genes forming the wing pattern supergene in the polymorphic butterfly Heliconius numata.
677 *Evodevo* **10**:16. doi:10.1186/s13227-019-0129-2
678 Schaeffer V, Althauser C, Shcherbata HR, Deng W-M, Ruohola-Baker H. 2004. Notch-Dependent
679 Fizzy-Related/Hec1/Cdh1 Expression Is Required for the Mitotic-to-Endocycle Transition in
680 Drosophila Follicle Cells. *Current Biology* **14**:630–636. doi:10.1016/j.cub.2004.03.040
681 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C,
682 Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A.
683 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682.
684 doi:10.1038/nmeth.2019
685 Shcherbata HR. 2004. The mitotic-to-endocycle switch in Drosophila follicle cells is executed by
686 Notch-dependent regulation of G1/S, G2/M and M/G1 cell-cycle transitions. *Development*
687 **131**:3169–3181. doi:10.1242/dev.01172
688 Stern DL, Orgogozo V. 2009. Is Genetic Evolution Predictable? *Science* **323**:746–751.
689 doi:10.1126/science.1158997
690 Swan A, Schüpbach T. 2007. The Cdc20 (Fzy)/Cdh1-related protein, Cort, cooperates with Fzy in
691 cyclin destruction and anaphase progression in meiosis I and II in Drosophila. *Development*
692 **134**:891–899. doi:10.1242/dev.02784
693 Thayer RC, Allen FI, Patel NH. 2020. Structural color in Junonia butterflies evolves by tuning scale
694 lamina thickness. *eLife* **9**:e52187. doi:10.7554/eLife.52187
695 Turner JRG. 1981. Adaptation and Evolution in Heliconius: A Defense of NeoDarwinism. *Annual*
696 *Review of Ecology and Systematics* **12**:99–121. doi:10.1146/annurev.es.12.110181.000531
697 Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J,
698 Lewis JJ, Hines HM, Ruiz M, Salazar C, Linares M, Moreira GRP, Jiggins CD, Counterman BA,
699 McMillan WO, Papa R. 2017. Complex modular architecture around a simple toolkit of wing
700 pattern genes. *Nature Ecology & Evolution* **1**:1–12. doi:10.1038/s41559-016-0052
701 VanKuren NW, Massardo D, Nallu S, Kronforst MR. 2019. Butterfly mimicry polymorphisms highlight
702 phylogenetic limits of gene re-use in the evolution of diverse adaptations. *Mol Biol Evol*.
703 doi:10.1093/molbev/msz194
704 Van't Hof AE, Campagne P, Rigden DJ, Yung CJ, Lingley J, Quail MA, Hall N, Darby AC, Saccheri IJ.
705 2016. The industrial melanism mutation in British peppered moths is a transposable
706 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

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. 2017. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. *PNAS* **114**:10707–10712. doi:10.1073/pnas.1709058114