G3, 2022, 12(4), jkac021
https://doi.org/10.1093/g3journal/jkac021
Advance Access Publication Date: 31 January 2022
Investigation

A large deletion at the cortex locus eliminates butterfly wing patterning

Joseph J. Hanly [1,2,*] Luca Livraghi [2,3] Christa Heryanto [1] W. Owen McMillan [2] Chris D. Jiggins [3] Lawrence E. Gilbert [4] and Arnaud Martin [1]

[1]Department of Biological Sciences, The George Washington University, Washington, DC 20052, USA,
[2]Smithsonian Tropical Research Institute, Panama 0843-03092, Republic of Panama,
[3]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK, and
[4]Department of Integrative Biology, University of Texas, Austin, TX 78712, USA

*Corresponding author. Department of Biological Sciences, The George Washington University, Washington, DC 20052, USA. Email: joe.hanly@gmail.com

Abstract

As the genetic basis of natural and domesticated variation has been described in recent years, a number of hotspot genes have been repeatedly identified as the targets of selection, *Heliconius* butterflies display a spectacular diversity of pattern variants in the wild and the genetic basis of these patterns has been well-described. Here, we sought to identify the mechanism behind an unusual pattern variant that is instead found in captivity, the *ivory* mutant, in which all scales on both the wings and body become white or yellow. Using a combination of autozygosity mapping and coverage analysis from 37 captive individuals, we identify a 78-kb deletion at the *cortex* wing patterning locus, a gene which has been associated with wing pattern evolution in *H. melpomene* and 10 divergent lepidopteran species. This deletion is undetected among 458 wild *Heliconius* genomes samples, and its dosage explains both homozygous and heterozygous *ivory* phenotypes found in captivity. The deletion spans a large 5′ region of the *cortex* gene that includes a facultative 5′UTR exon detected in larval wing disk transcriptomes. CRISPR mutagenesis of this exon replicates the wing phenotypes from coding knock-outs of *cortex*, consistent with a functional role of *ivory*-deleted elements in establishing scale color fate. Population demographics reveal that the stock giving rise to the ivory mutant has a mixed origin from across the wild range of *H. melpomene*, and supports a scenario where the *ivory* mutation occurred after the introduction of *cortex* haplotypes from Ecuador. Homozygotes for the *ivory* deletion are inviable while heterozygotes are the targets of artificial selection, joining 40 other examples of allelic variants that provide heterozygous advantage in animal populations under artificial selection by fanciers and breeders. Finally, our results highlight the promise of autozygosity and association mapping for identifying the genetic basis of aberrant mutations in captive insect populations.

Keywords: structural variant; deletion; Lepidoptera; domestication; evolution; development; cortex

Introduction

Captive populations can harbor variation that is not observed in the wild. Aside from domestication for agricultural or commercial yield, animals and plants have often been selectively bred for "beautiful," "interesting," or "unnatural" esthetic traits, especially color and pattern variants (Driscoll *et al.* 2009; Cieslak *et al.* 2011). In recent decades, the genetic basis of artificially selected variation has begun to be mapped in many organisms used in agriculture, floriculture and the pet trade, including the genetic mapping of a large number of genes responsible for flower color variation (Park *et al.* 2007; Giovannini *et al.* 2021), melanin-based coat color in domestic mammals (Cieslak *et al.* 2011), plumage patterns, and colors in birds (Domyan and Shapiro 2017; Price-Waldman and Stoddard 2021), and chromatophore distribution in squamates and fishes (Guo *et al.* 2021). These variants can affect genes involved in the enzymatic production, transport, and deposition of pigments such as the melanin and carotenoid pathways (Mundy *et al.* 2016; Courtier-Orgogozo and Martin 2020), or can be caused by genes that affect signaling or cell type differentiation [as with chromatophore distribution in squamates (Kuriyama *et al.* 2020)]. The extensive set of artificial and domesticated genetic variants has been used repeatedly both for studies of genotype-to-phenotype relationship, and for understanding the genetics of disease states in humans (Andersson 2016).

Notably, several cases have been identified where a gene underlying an artificially selected variant is also responsible for natural variation in other populations or species. This is the case with *BCO2*, where an artificially selected protein coding variant in cattle affects milk color (Berry *et al.* 2009), and a wild-type regulatory variant in wall lizards causes changes to ventral scale color (Andrade *et al.* 2019). The gene *Agouti* has repeatedly been mapped in cases of pigment differences, both in cases of domestication [e.g. in the chicken (Yu *et al.* 2019), horse (Rieder *et al.* 2001), and rabbit (Fontanesi *et al.* 2010)], and in natural variation [e.g. in warblers (Toews *et al.* 2016), humans (Bonilla *et al.* 2005), and snowshoe hare (Jones *et al.* 2018)]. This reuse of hotspot

Received: December 24, 2021. **Accepted:** January 21, 2022

genes in wild and captive populations can even be observed at the within-species level; multiple separate variants in *Agouti* have been linked to pigment variation in sheep—both in wild and captive populations, and caused by protein coding mutations, *cis*-regulatory mutations, and copy number variation (Norris and Whan 2008; Gratten *et al.* 2010). As such, studies of domesticated variation can inform our understanding of natural variation too, both on macro- and microevolutionary scales.

As the primary selective force in captive populations is the eye of the selector, one may often observe variation that is not seen in nature because it is not optimally fit (Andersson and Georges 2004; Marsden *et al.* 2016; Makino *et al.* 2018; Moyers *et al.* 2018). Fanciers and breeders have purposefully selected for variation that causes deleterious effects in combination with color and pattern differences, applying a regime of selection that favors the phenotype of interest over fitness. This applies especially to distinctive coloration, which can be straightforward to observe and maintain in a captive stock. Examples include coat color variants for Merle dogs, where homozygotes for a retrotransposon insertion in the *SILV* gene have an increased risk of deafness and blindness (Strain *et al.* 2009; Langevin *et al.* 2018), and in overo horses, where foals homozygous for a mutated Endothelin Receptor B (*EDNRB*) develop Lethal White Syndrome (Metallinos *et al.* 1998). "Lemon frost" geckos have been selected for their unique color, but exhibit an increased risk of iridophoroma, the formation of tumors from iridophores (analogous to melanoma) (Guo *et al.* 2021). Such reduced fitness has also been identified in floriculture, where white petunias with mutations to the *AN1* gene have deficient vacuole acidification and a weakened seed coat (Spelt *et al.* 2002), though an artificially selected mutation to the same gene causing a similar effect in the morning glory does not appear to carry the same deleterious fitness effects (Park *et al.* 2007). Whether they are the direct targets of selective breeding, or collateral effects of inbreeding and bottlenecks, deleterious mutations that would be purged by natural selection in the wild are common in captive populations (Bosse *et al.* 2019).

The majority of work mapping artificially selected or domesticated variation in animals has taken place in vertebrates, with the notable exceptions of the economically important silkworm (Futahashi and Osanai-Futahashi 2021). *Heliconius* butterflies, a model system for the genetic study of color pattern adaptations in the wild (Jiggins *et al.* 2017; Van Belleghem *et al.* 2021), have also been maintained in captivity by butterfly breeders continuously since at least the mid-20th century, including at notable tourist attractions like Butterfly World (BW) in Florida. There, they have been selectively bred for hundreds of generations with intentional, artificial selection upon pattern variations that deviate from phenotypes normally observed in the wild. Over an extended period of inbreeding, selection, and occasional outcrossing, novel color variations have appeared in this captive population, including the "Piano Keys" (PK) pattern, and a deleterious mutation that occurred within the PK genetic stock, here dubbed *ivory* (Fig. 1).

The genetics of wing patterning have been extensively studied in *Heliconius* (McMillan *et al.* 2020). Linkage and association mapping have pinpointed regulatory regions around a small toolkit of genes as being responsible for much of the pattern variation seen in wild populations, including the transcription factor *optix* (Wallbank *et al.* 2016; Zhang *et al.* 2017; Lewis *et al.* 2019; Morris *et al.* 2020;), the signaling ligand *WntA* (Mazo-Vargas *et al.* 2017; Van Belleghem *et al.* 2017), and *cortex*, a *cdc20* homolog with a currently unidentified function (Nadeau *et al.* 2016; Livraghi *et al.* 2021). We used whole-genome resequencing, association mapping and autozygosity mapping to determine that the *ivory* mutation is caused by a large deletion at the patterning gene and evolutionary hotspot *cortex*, which likely occurred de novo in the captive population.

Methods

Stock history

The captive population of *Heliconius melpomene* was initially collected by JRG Turner and others from around Central and South America, and was later transferred from his genetic research stocks at the University of Leeds to the former London Butterfly House (LBH) maintained by Clive Farrell from 1981 as a multirace hybrid population. They were then acquired from Tom Fox of LBH by R. Boender at the MetaScience Butterfly Farm in Florida in about 1985. These hybrid stocks formed a core of Boender's inhouse hybrid *Heliconius* display stock when BW opened in 1988. Only one known introduction of wild-caught *H. melpomene* occurred to this stock, of *H. melpomene cythera* collected in the early 1980s by R. Boender and T. Emmel around Tinalandia, Ecuador. After this introduction, Boender began selection for a novel pattern that is termed "Piano Key" in this paper. The exact mix of races that were transferred from LBH to BW are not known, but contain pattern alleles that are characteristic of *H. melpomene* races of Suriname and the Guianas (LEG, personal observation).

Since the Piano Key phenotype was first noticed by R. Boender over 30 years ago, the Piano Key stock population has been maintained in separate compartments for around 400 generations at BW and has long been fixed for the PK phenotype (RB, personal communication). Thus, we referred to it as *H. melpomene* BWPK throughout. Within the last decade, novel mutant butterflies appeared with fewer melanic scales, initially termed "pale PK." These were subsequently separated to form a selected subpopulation at BW. Soon after, virtually pure white butterflies incapable of flight began appearing in this selected stock, herein referred to as *ivory*. These were subjected to intentional artificial selection. In 2015, "Pale PK" were sent to UT Austin and maintained by LEG in climate-controlled greenhouses (see Supplementary Table 1 for phenotypes and collection dates) to investigate the genetics of this mutation.

Imaging

Pinned specimens were imaged with a Nikon D5300 camera mounted with a Micro-NIKKOR 105 mm f/2.8 lens. Scale phenotypes were imaged with a Keyence VHX-5000 microscope mounted with a VH-Z100R lens.

Short-read DNA sequencing

DNA was extracted from thoracic tissues of 37 *H. m.* BWPK individuals using the Qiagen Dneasy Blood & Tissue Kit, RNAse-treated, and used to prepare a multiplexed sequencing library with the TruSeq PCR-free DNA protocol. Samples were sequenced on an Illumina NovaSeq S1 150 bp PE run, yielding 13× mean coverage per individual. Sequencing reads are accessible via NCBI SRA under the project accession number PRJNA610063.

Genomic analyses

Samples were aligned to Hmel2.5 (Pinharanda *et al.* 2019) retrieved from LepBase (Challis *et al.* 2016) with BWA-MEM using default parameters (Li 2013), and variants called using GATK v4.1 with tools HaplotypeCaller and GenotypeGVCFs using default parameters (McKenna *et al.* 2010). Variant sites were accepted if they were biallelic and the quality (QUAL) value was ≥30. SNPs were phased with Beagle 4.1 (Browning and Browning 2007). Phased SNP variants were

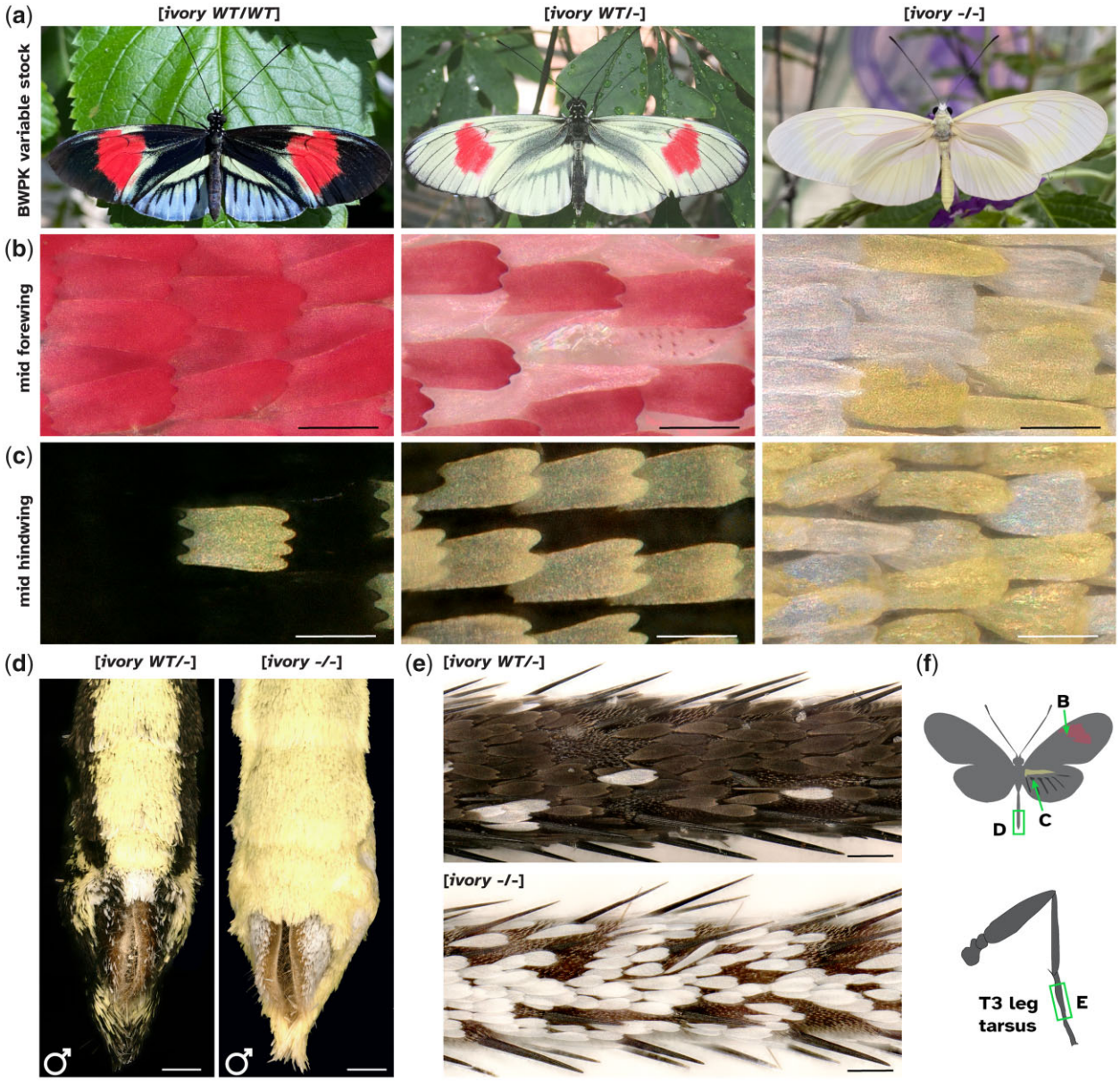

Fig. 1. Phenotypes of *H. melpomene* BWPK butterflies. a) Ivory phenotypes in the *H. melpomene* "Piano Keys" in the UT Austin stock (BWPK). The 3 color states (from left to right "Dark BWPK," "Pale BWPK," and "ivory") depend on the allelic dosage of a codominant mutation. Ivory homozygotes are inviable and only found among offspring from 2 pale heterozygotes. b) Magnified view of the forewing red band region. In the [*ivory WT/–*] state, abnormal scales have formed in the red region, with red pigment in granules and scales curled, while in [*ivory –/–*], all scales are yellow or white. c) Magnified view of a central hindwing region that is black in [*ivory WT/WT*]. The [*ivory WT/–*] wing has all cover scales as yellow or white while all ground scales remain black. All scales are yellow or white in [*ivory –/–*]. Complete replacement of melanic scales by yellow-white scales in the abdomen (d) and legs (e) from [*ivory –/–*] homozygotes. f) Cartoon of image locations for (b–e). Scale bars: b–c = 50 μm; d = 500 μm; e = 100 μm.

used to perform principal component analysis (PCA) using the Eigensoft module SmartPCA (Price *et al.* 2006). SNP association was carried out in PLINK v.1.9, with 1,000 permutations (Purcell *et al.* 2007). For phylogenetic analyses, we followed the TWISST pipeline(Martin and Van Belleghem 2017); briefly, data were phased in Beagle 4.1 with a window size of 10,000 and overlap of 1,000. Trees were built from windows of 50 SNPs, and then analyzed with TWISST using the 5 groups "East," "West," "Ecuador," "Atlantic," and "BWPK" (Supplementary Table 3).

De novo assembly for indel breakpoints

To find precise indel breakpoints, a subset of samples were de novo assembled with Velvet (Zerbino and Birney 2008) using default parameters. The resulting contigs were searched with BLAST (Camacho *et al.* 2009) for the genomic region including the deletion. Scaffolds that bridged the indel breakpoints were selected, and aligned to the Hmel2.5 reference genome with MAFFT (Katoh *et al.* 2002).

CRISPR/Cas9 mutagenesis

We designed sgRNAs against the putative *H. erato cortex* promoter corresponding to N20NGG sites within the distal promoter/5′UTR. In order to mutagenize the locus, we designed 3 sgRNAs (Table 1) using the "find CRISPR sites" algorithm within the Geneious software. Guide specificity and off-target effects were assessed by scoring against the *H. erato* reference genome. sgRNAs displaying

Table 1. CRISPR mutagenesis of the *cortex* distal promoter/5′UTR in *H. erato*.

Batch	sgRNA target	Injected	Pupae	Adults	Crispants	Survival	Penetrance
1	*sgProm1*: 5′AGCGGTAACTTTATCGCGAT	512	42	37	3	7.20%	0.6%
2	*sgProm2*: 5′ATGGGATATGTTTTAATAGT *sgProm3*: 5′AACCTGTCACATCAGTACAG	695	78	73	1	11.2%	0.1%

low off-target scores were then synthesized commercially by Synthego, and mixed with Cas9 at a concentration of 500:500 ng/µl, respectively. Embryonic injections were performed as previously described (Livraghi *et al.* 2021) within 1–3 h after egg laying, after which larvae were allowed to develop on a diet of *P. biflora* until adult emergence.

Results

Heliconius melpomene BWPK—artificially selected wing patterns in an insectary stock

The stock maintained at BW in Florida and at UT Austin is here named *H. melpomene* BWPK (see *Methods—Stock history* section). The insectary-bred, artificially selected line is polymorphic for several wing pattern elements found in wild populations, including the red pattern elements (Fig. 1). Additionally, *H. m.* BWPK includes aberrant pattern features not observed in natural populations. First, in the pattern dubbed "Piano Keys" (Fig. 1a), white or yellow hindwing marginal elements extend along the veins toward the discal cell. While distal yellow hindwing pattern elements are observed in some wild pattern forms of *Heliconius*, including *H. m. cythera* from Ecuador and *H. cydno*, the extent of these marginal elements in *H. m. BWPK* is much greater than any of these, occurring over most of the area of the hindwing except for the intervein regions that would otherwise be taken by the red hindwing rays.

Occasional butterflies with more extensive regions of yellow or white scales were identified (Fig. 1a, center). Other than the wing pattern differences, both forms of butterflies exhibit typical feeding, mating, egg-laying, and flight behaviors to other insectary-reared *Heliconius* (e.g. Supplementary Movie 1). However, the offspring of a mating of 2 pale *H. m.* BWPK will include completely white-winged and -bodied butterflies (Fig. 1c). We dub these butterflies "ivory" in reference to their emergence in the "Piano Keys" stock, and inferred that the pale morph of *H. m.* BWPK represents a heterozygous state for a mutation that causes the ivory phenotype in the homozygous state.

The ivory butterflies have melanin pigments in their body cuticle and eyes, indicating that their capacity to synthesize and deposit melanin has not been perturbed. However, all scales on the wings and body are white or yellow, with no black or red scales, with the exception of a 1-mm patch of red at the base of the hindwing in some individuals. This is in stark contrast to wild-type or BWPK butterflies which have black scales covering most of the body. Unlike the other BWPK butterflies, ivory homozygote butterflies do not fly and have not been observed to successfully mate or lay eggs. The wings appear to be of a normal strength and sturdiness, but flight is weak and will not occur when prompted by dropping (Supplementary Movie 1).

In order to investigate the genetic basis of these wing patterns, we sequenced 37 butterflies from the UT Austin *H. m.* BWPK stock to an average coverage of 13×, including 11 Dark BWPKs, 9 Pale BWPKs and 10 ivory butterflies (Fig. 1a).

Autozygosity mapping identifies associated SNPs

We used a combination of GWAS and patterns of SNP autozygosity to determine the locations of regions associated with Mendelian pattern variants (Fig. 2). As proof of principle, we mapped the Dennis pattern element which was previously described as a *cis*-regulatory element of the gene *optix*. GWAS for presence vs absence of Dennis gave a narrow association peak centered near *optix*, in a region previously identified as associated with the Dennis pattern element (Morris *et al.* 2020) (Supplementary Fig. 1). We then examined the inheritance patterns of SNPs, filtering for SNPs where butterflies with no Dennis element were homozygous for the reference allele, and individuals with the Dennis element were either heterozygous or homozygous for an alternate allele (6/29 butterflies had the Dennis pattern element, Supplementary Table 1). In total, 3,430 SNPs matched this pattern of inheritance, with 3,425 in a cluster centered on the gene *optix*, indicating that autozygosity mapping is appropriate for mapping pattern elements in this data set.

GWAS for ivory gave a broad association peak on chromosome 15 (Fig. 2d). We expected dark morph BWPK butterflies to carry the reference allele, pale morphs to be heterozygous, and ivory morphs to be homozygous for a nonreference allele. We expected that at the causative locus, this allelic segregation pattern should be fixed in every individual. This was the case at just 131 SNPs in the whole genome, all within a 9-kb interval on chromosome 15, within the broad association peak (Fig. 2d, blue marks). These SNPs sit within the first intron of the gene *cortex*. Of particular note, *cortex* is a switch gene necessary for differentiation into black type II and red type III scales (Nadeau *et al.* 2016; Livraghi *et al.* 2021), and is a hotspot gene of wing pattern evolution in many species, including in industrial melanism in the peppered moth (Van't Hof *et al.* 2016, 2019).

We noted that in [*ivory WT/–*] butterflies, scales in red regions exhibited aberrant morphologies (Fig. 2, e and f). Granular clumps of red pigments could be observed in the body of some scales, whereas wild-type scales are solidly pigmented with no granularity. Additionally, we observed scales that were curled at the edges rather than flat. Similar scale phenotypes occur in the wings of *cortex* crispant butterflies (Fig. 2, e and f), thus supporting a role for *cortex* loss-of-function in the ivory color phenotypes (Livraghi *et al.* 2021).

Ivory butterflies carry a large deletion including the *cortex* promoter

Many phenotypes involve structural genomic variation, both under artificial selection (Bruders *et al.* 2020), and in natural adaptation (Junker *et al.* 2020; Mérot *et al.* 2020). As such, we checked read depth across this region in all our sequenced samples. Immediately adjacent to the block of fixed SNPs identified by autozygosity mapping, we found a large region where read depth in [*ivory –/–*] butterflies dropped to zero, while in [*ivory WT/–*] it dropped by half. This indicated that a large deletion spans the first exon and promoter of *cortex* (Fig. 3a).

To determine the precise breakpoints of this deletion, we de novo assembled short-read data for all [*ivory WT/WT*] and [*ivory*

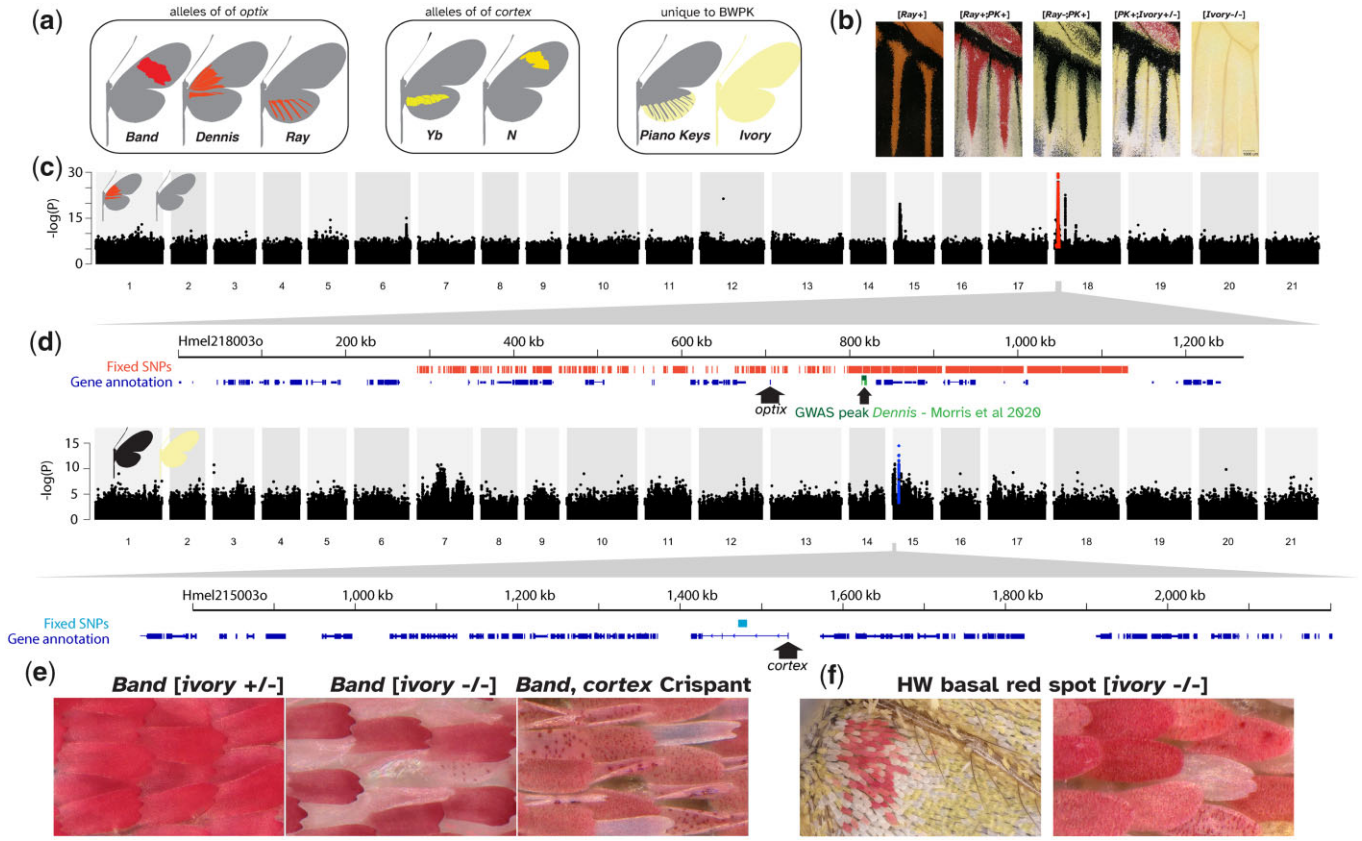

Fig. 2. Genome-wide association and autozygosity mapping in *H. melpomene* BWPK. a) Cartoons of the allelic variation present in the *H. m.* BWPK stock. b) The hindwing veins, showing the composite effects of different alleles depicted in (a). c) GWAS for the Dennis element; Manhattan plot of Wald test *P*-values, with fixed SNPs in red, and a magnified annotation of the region around the gene *optix*. d) GWAS for ivory, with fixed SNPs in blue and a magnified annotation of the *cortex* region, with the point of the arrow at the annotated TSS. e) Comparison of red scales in [*ivory WT/WT*], [*ivory* −/−], and *cortex* crispant wings. f) Some ivory butterflies have a small red dot at the base of the ventral hindwing, the only scales on the entire butterfly that are not white or yellow. They share the atypical phenotype of red scales from *cortex* crispants (e).

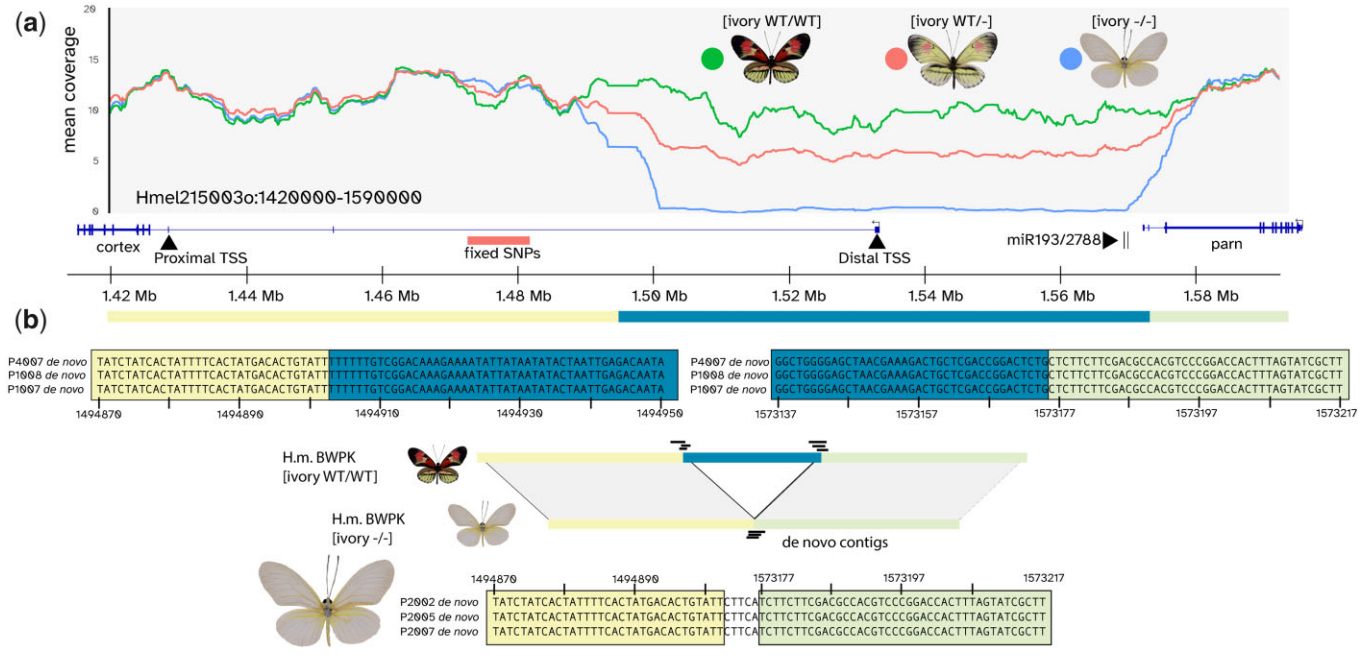

Fig. 3. Ivory is associated with a large deletion in the gene *cortex* (a) mean coverage plot for the region around *cortex*, with [*ivory WT/WT*] in green, [*ivory WT/*−] in red, and [*ivory* −/−] in blue. b) Alignments of de novo scaffolds, indicating the precise breakpoints of the *ivory* deletion. Colors indicate the orientation of sequences relative to the reference scaffold depicted in (a).

−/−] individuals, and BLASTed the resulting contigs against the *cortex* region. This allowed us to recover individual contigs from some individuals that bridged the 2 ends of the deletion, giving a precise breakpoint at positions Hmel215003o:1494902-1573177, with a length of 78,275 bp relative to the reference genome (Fig. 3b).

The *ivory* deletion includes 1 of the 2 *cortex* transcription start sites

The *ivory* deletion contains the only annotated transcription start site (TSS) and promoter for *cortex*. While butterflies carrying this mutation can complete development, even with effects in scale development across the whole organism and the inability to fly, coding knockouts of *cortex* (including clonal G_0 crispants) have been observed to have high levels of embryonic lethality, with very few crispant clones observed (Livraghi *et al.* 2021).

The ability of *ivory* mutants to survive in spite of the loss of the promoter, as well as the presence of a long first intron in *cortex*, led us to hypothesize that there may be a second, alternate TSS. In order to determine if this is the case, we mapped RNAseq from *H. melpomene* for a variety of samples including larval and pupal wings (Hanly *et al.* 2019), adult ovaries (Pinharanda *et al.* 2019), and embryos, adult head and adult abdomen (Dasmahapatra *et al.* 2012), and looked at splicing and alignment of the 5′ end of the transcript. No expression was detected in adult head or abdomen, but in embryos, ovaries, and pupal wings, transcription initiates at position Hmel215003o:1424680 in the third annotated exon, which is adjacent to the coding sequence (the "proximal promoter") (Fig. 4a; Supplementary Figs. 2 and 3). Expression in larval wings initiates at the annotated TSS (the "distal promoter"), and the exon with the proximal promoter is skipped. In pupal wings, transcription again initiated at the proximal promoter. This indicates that *cortex* has 2 alternative TSSs, a proximal promoter used in multiple tissues and a distal promoter specific to the larval wing. Only the distal promoter is included in the *ivory* deletion.

Mutagenesis of the distal promoter phenocopies *cortex* null effects

We reasoned that if the loss of the distal promoter causes the ivory phenotype, knocking out the promoter with CRISPR/Cas9 should phenocopy the effects of the *ivory* deletion. We generated G_0 mosaic knockouts (crispants) of the promoter in the related species *Heliconius erato* (Fig. 4, b–d). Resulting butterflies had extensive yellow/white clones on the wings, however, several failed to emerge from the pupa, and a low survival rate and penetrance were observed (Table 1), suggesting that the few wing crispants obtained are rare mosaic escapers of a loss-of-function experiment with lethal effects (Livraghi *et al.* 2018). Thus, while the pleiotropic effects of deleting the TSS were reminiscent of the previously reported *cortex* protein coding knockouts, which also exhibited low survival and penetrance (Livraghi *et al.* 2021), this experiment did not fully recapitulate the ability of the *ivory* deletion allele to generate large wing clones. This difference in pleiotropy may be due to the use of *H. erato* in the CRISPR assay, and further experiments are needed to test the functionality of deleted elements within a *H. melpomene* stock. However, we conclude from these experiments that the disruption of 5′ distal elements of *cortex* are necessary for normal color scale patterning, consistent with a causal role of the 78-kb deletion in underlying the ivory phenotypes.

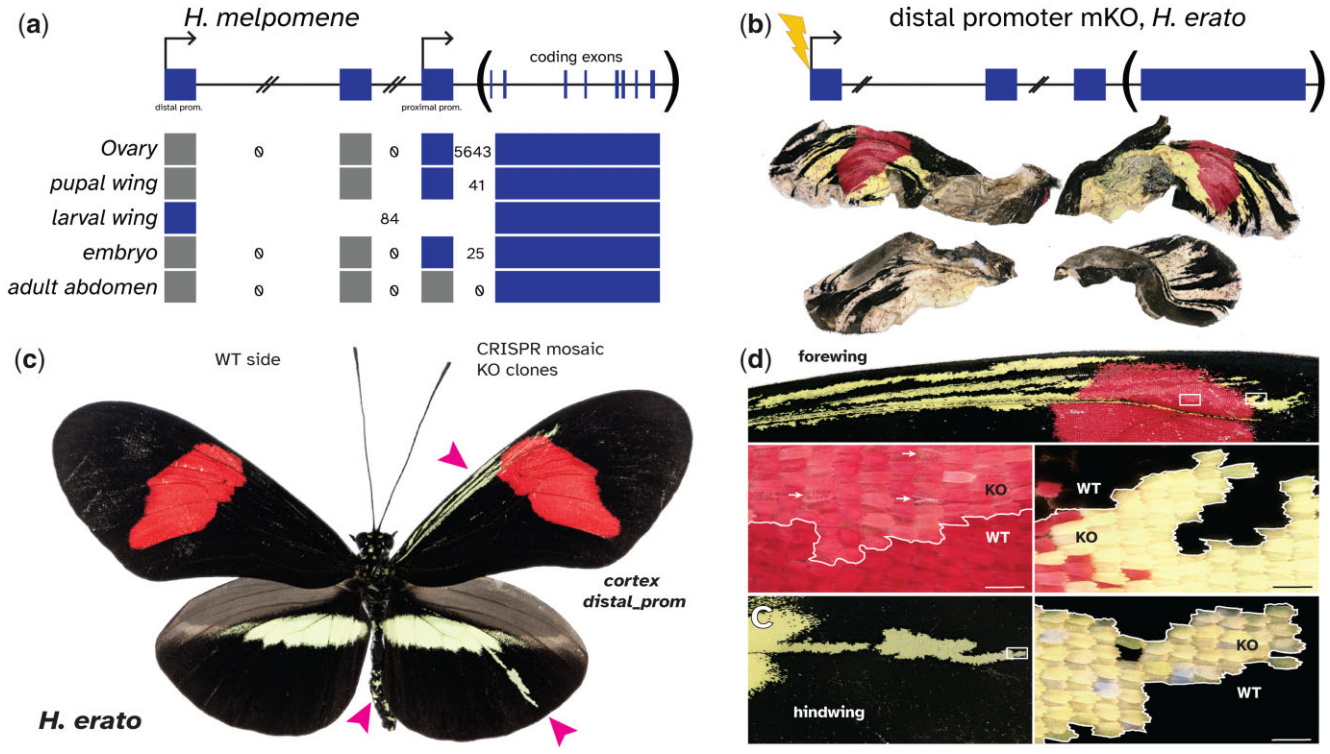

Fig. 4. Knockout of the distal promoter phenocopies *cortex* protein-coding knockouts. a) Mapping of RNAseq reads at *cortex* suggests that there are 2 promoters and TSS. Numbers of intron-spanning reads are indicated (see Sashimi plots in Supplementary Figs. 2 and 3). Knockout of the distal promoter in *H. erato* causes a transformation of black scales to white/yellow in a phenocopy of *cortex* protein coding knockouts. b) A butterfly with extensive clones but that emerged poorly, indicating pleiotropic effects not observed in the *ivory* mutants, while c) much smaller clones similar to those reported by Livraghi *et al.* (2021), with clone positions indicated by pink arrows. d) High magnification images of mutant clones. Scale bars: 100 μm.

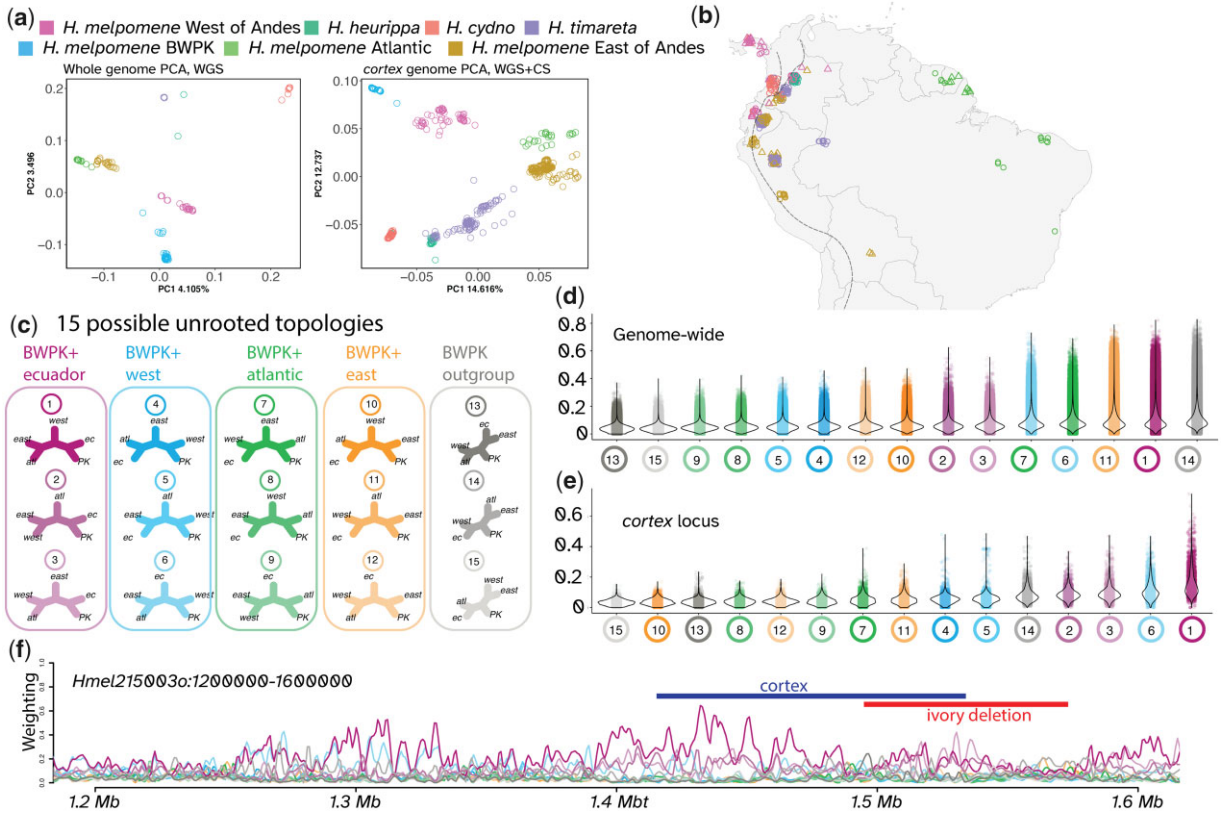

Fig. 5. Phylogenetic clustering and weighting imply a mixed origin for *H. melpomene* BWPK. Clustering by PCA from both whole-genome data (a) and from the *cortex* locus using additional selective sequencing data (b) show similar geographic clustering to that previously reported by Martin *et al.* (2019); c) with the addition of separation between *H. melpomene* from East of the Andes and *H. melpomene* from French Guiana, Suriname, and Brazil (here termed "Atlantic"). *Heliconius m.* BWPK form a distinct cluster, closest to the West cluster of *H. melpomene*. Topology weighting was performed using TWISST (Martin and Van Belleghem 2017), with 5 defined groups (West of Andes, Ecuador, East of Andes, Atlantic, and BWPK) (c). d) Genome-wide, the most common topologies were 14 (BWPK as outgroup), 11 (BWPK clustered with East of Andes), and 1 (BWPK clustered with Ecuador). In contrast to this in (e), TWISST results at the *cortex* locus showed that topology 1 was most common, supporting an Ecuadorian origin for the *cortex* variants in *H. m. BWPK*.

Demographic origins of the *ivory* allele

The *H. m.* BWPK stock was kept in insectaries for 30 years (ca. 400 generations) before producing the *ivory* mutation. The precise history and ancestry of the stock is not recorded, but we do know that the original stock was generated with a mix of butterflies from multiple locations after the addition of butterflies from the vicinity of Tinalandia, Ecuador in the 1980s. The first pale PK (i.e. heterozygotes for *ivory*) appeared in the PK culture at BW after 2013 and were shared with LEG for study in 2014–2015. We sought to determine if the *ivory* deletion or associated variation originated from this introduction event from Tinalandia.

Whole genome PCA clustering of a geographic spread of samples of *H. melpomene*, as well as sister taxa *H. cydno* and *H. timareta*, replicates the geographic and species clustering previously described by Martin *et al.* (2019), with the addition of a distinct "Atlantic" group within *H. melpomene* (Supplementary Table 1). Consistent with a large contribution from Ecuadorian ancestry, *H. m.* BWPK cluster together near *H. melpomene* from the west of the Andes. Similarly, PCA of just the *cortex* locus, using many more individuals generated by selective sequencing by Moest *et al.* (2020) (Supplementary Table 4), also places BWPK closest to western *H. melpomene*.

Given the mixed ancestry of *H. m.* BWPK, we examined heterogeneity in genome-wide ancestry to determine if variation around the ivory locus in particular was of Ecuadorian origin. We built phylogenetic trees in windows of 50 SNPs across the genome and using topology weighting to determine the closest neighbor of *H. m.* BWPK at all genomic positions. Of the 15 possible topologies, the most frequent topology places BWPK as an outgroup to the 4 other *H. melpomene* taxa. The second most common topology groups BWPK with *H. melpomene* from western Ecuador (*H. m. vulcanus* and *H. m. plesseni*), and the third most common places BWPK with *H. melpomene* from the East of the Andes. In contrast, sliding windows across the *cortex* locus (including the selective sequencing data) show the most common tree topology here places BWPK sister to the western Ecuador taxon. This result is supportive of *H. m.* BWPK having an admixed genome, but with the allele at *cortex* having originated in Ecuador. Importantly, the *ivory* deletion was not detected in any of 458 wild individuals, indicating it was likely a de novo mutation within the BWPK stock, though it is possible it occurs at very low frequencies in the wild.

Discussion

Using whole-genome sequencing of a set of related individuals, we were able to map the *ivory* mutation to a large deletion that removed the distal promoter of the wing patterning gene *cortex*. Cortex was first identified as a switch between melanic and nonmelanic patterns in both *Heliconius* and the peppered moth *Biston betularia*, and has later been mapped as a switch between melanic and nonmelanic scale types in a number of other species, including 3 other geometrid moths (Van't Hof *et al.* 2019), and *Papilio* butterflies (VanKuren *et al.* 2019), and also has a role in wing pattern polyphenism in *Junonia coenia* (van der Burg *et al.* 2020). This is another example where the same gene has been the target of

natural selection in a wild population and of artificial selection in a captive population.

The list of genes that have been identified as the targets of selection in butterfly wing pattern variation includes transcription factors like Optix and Bab (Reed *et al.* 2011; Ficarrotta *et al.* 2021), signaling pathway components like WntA (Martin *et al.* 2012), or terminal effector genes like Yellow (Martin *et al.* 2020), all of which have molecular functions that support a role in developmental patterning or pigment synthesis (McMillan *et al.* 2020). Cortex, on the other hand, has no clear functional association with color pattern or cell differentiation, and yet has been mapped in a very broad phylogenetic spread of species, suggesting its function may be ancestral in the Lepidoptera. The gene does not appear to have a role in *Drosophila* wing development or patterning (Nadeau *et al.* 2016), limiting our ability to make inferences about its molecular function and interactions. In identifying this large deletion, this study provides novel insight into the genetics of this hotspot locus, which will aid future characterization of its developmental function as a switch gene for color patterning.

The loss of all black and red scales in the *ivory* mutant supports a model in which white/yellow scales are the "default" state during scale cell differentiation, with scale cell precursors failing to differentiate into red or black cell types. This model was initially proposed based on observation and ultrastructure and pigmentation of scales (Gilbert *et al.* 1987), as well as on the analysis of pattern homologies of the Heliconiini (Nijhout and Wray 1988). It is likely, given the repeated involvement of this locus in wing pattern evolution, that this function of cortex will prove to be conserved throughout Lepidoptera.

Structural variation and mutations

Large deletions have repeatedly been observed in cases of both natural variation and domestication, with 29 deletions affecting regulatory regions which are larger than 1 kb listed on GepheBase (Courtier-Orgogozo *et al.* 2020). Recurrent deletions of a *pitx1* enhancer in sticklebacks, up to 8 kb in length, have caused convergent pelvic reduction (Chan *et al.* 2010), and the deletion of a 60.7-kb regulatory region at the *AR* gene in humans, which removes enhancers present in chimpanzee and conserved in other mammals, leads to loss of penile spines (McLean *et al.* 2011). Similarly large deletions have also been confirmed in the genetic basis of domestication phenotypes, including a 44-kb deletion, also at *pitx1*, causing feathered feet in chickens (Domyan *et al.* 2016), and a 141-kb deletion upstream of *agouti/ASIP* in Japanese quail (Nadeau *et al.* 2008).

The large deletion that causes ivory is a form of structural variation (SV). Other structural variants affecting this locus have been found in wild populations of *Heliconius*; multiple inversions have formed a wing pattern supergene in *Heliconius numata*, and an independent inversion, similar in size and position, has been found in multiple species of the *erato* clade (Joron *et al.* 2011; Edelman *et al.* 2019). It is possible that this genomic region will prove to be prone to a higher-than-average level of SV, or, conversely, that the region is more tolerant to SV than other parts of the genome.

Regulatory consequences of the *ivory* mutation

Large deletions have the potential to remove a large section of *cis*-regulatory sequence, as well as transcribed sequences. The *ivory* deletion causes the loss of 1 of 2 promoters. As many as 40% of developmentally expressed genes in *Drosophila* have 2 promoters which can cause distinct regulatory programs (Bhardwaj *et al.* 2019), and multiple promoters are also common in human genes (Singer *et al.* 2008). This both increases the complexity of gene regulatory interactions and increases the number of transcript isoforms per gene—for example, an alternate promoter in the *Drosophila* gene *Zfh1* creates an isoform that has a shorter 5′UTR which is missing miRNA seed sites on its 5′UTR, permitting differential degradation of the mRNA by *miR-8* (Boukhatmi and Bray 2018). We determined that the *ivory* deletion contains 1 annotated promoter of *cortex*, but that in some contexts during development, another TSS is used, much closer to the translation start site. Alternate promoter usage is likely to be common and widespread in animals.

We expect that the use of 2 separate, context-specific promoters at *cortex* will create a hierarchical regulatory organization, where CREs used in different tissues or at different times can loop to different promoters. Additionally, transcripts from each of the 2 promoters contain a different arrangement of 5′ noncoding exons, giving them different UTRs, possibly leading to differential translational regulation or degradation of the mRNA.

The *ivory* deletion does not include protein-coding sequence, which led us to hypothesize that by removing just 1 promoter, *cortex* would still be expressed in tissues that use the alternate promoter, and that therefore *ivory* mutants would bypass the pleiotropic, highly lethal effects observed in CRISPR experiments that target coding exons of *cortex* (Livraghi *et al.* 2021). However, crispants with deletions in the distal promoter in *H. erato* did not appear to be free of these pleiotropic effects, and exhibited high lethality and comparatively small clone sizes, dissimilar to the *H. melpomene* [*ivory* −/−] mutants. This may be due to inherent differences in the regulatory function of the distal TSS between these 2 species, or to the existence of other functional sequences within the deletion which might be necessary for generating the pale phenotypes. Specifically, the 78-kb deletion removes part of the 3′UTR of the neighboring gene *parn* as well as 2 miRNAs (Surridge *et al.* 2011), and it is also possible that a large deletion affects other cortex *cis*-regulatory elements or affect 3D chromatin interactions on a large portion of the chromosome. Additionally, the *ivory* butterflies carry a set of fixed SNPs adjacent to the deletion, which could contribute to the phenotype. In summary, the complete loss of black and red scales caused by the *ivory* deletion and by CRISPR-Cas9 mutagenesis in 2 *Heliconius* species indicates that either the distal promoter or another functional sequence within the 5′ noncoding region of *cortex* are necessary for black and red scale development.

Heterozygote advantage of deleterious mutations in captivity

The *cortex* mutant phenotype has been artificially selected by a breeder in the heterozygous state, when unusually widespread yellow and white patterns emerged in the stock. The heterozygote carriers of the *ivory* mutation ([*ivory* WT/−]) do not appear to have any reduction in fitness or vigor. In contrast, homozygous *ivory* butterflies of either sex are unable to fly (Supplementary Movie 1), limiting their ability to feed and reproduce—we have not observed any successful matings involving an [*ivory* −/−] butterfly. With a homozygous fitness of zero, those recessive effects make the *ivory* mutant allele functionally equivalent to a recessive lethal allele with a heterozygous phenotype under artificial selection such as in *PAX3* overo horses (28). This was previously formalized by Hedrick with models for lethal and "near-lethal" mutations with heterozygous advantage, and shown to be only possible under conditions of intense selective breeding (Hedrick 2015). To further probe the empirical evidence for this phenomenon, we compiled cases of heterozygote advantage under artificial selection from the literature using a database of gene-to-phenotype relationships (Courtier-Orgogozo *et al.* 2020), and found 39 analogous gene-to-trait relationships spanning domesticated mammals and birds (Table 2), 15 of which are recessive-

Table 2. Known cases of heterozygote advantage in animal breeding (see Gephebase for references[a]).

Gene	Species	Heterozygote advantage (HA) trait for breeding	Detrimental recessive effect	Nb of HA alleles	Mutation types
ACAN	Cattle	Short stature	Lethal	1	Frameshift (4-bp ins)
				1	New start codon (SNP)
ACAN	Horse	Short stature	Lethal	1	Loss of 6 a.a. (18-bp del)
				1	Frameshift (1-bp del)
ALX1	Cat	Brachycephaly	Frontonasal dysplasia	1	Loss of 4 a.a. (12-bp del)
BBS9	Pig	Fast growth	Lethal (misexpressed *BMPER*)	1	Large SV (212-kb del)
BMP15	Sheep	Increased female fecundity	Female infertility	5	Missense (SNP)
				2	Nonsense (SNP)
				1	Frameshift (1-bp ins)
				1	Frameshift (17-bp del)
				1	Cis-regulatory (SNP)
cortex[b]	*Heliconius*	White patterns	Flightless	1	Large SV (78-kb del)
EDNRB	Horse	White patterns	Lethal	1	Missense (SNP)
FGF3/4/19 cluster	Dog	Dorsal hair ridge	Dermoid sinuses	1	Large SV (133-kb duplication)
FGFR3	Sheep	Increased bone length	Chondrodysplasia	1	Missense (SNP)
FOXI3	Dog	Hairlessness	Ectodermal dysplasia	1	Frameshift (7-bp duplication)
GDF9	Sheep	Increased female fecundity	Female infertility	5	Missense (SNP)
HMGA2	Rabbit	Short stature	Lethal	1	Large SV (12.1-kb del)
HOXC10	Chicken	Head crest	Cerebral Hernia—likely epistatic	1	Large SV (197-bp duplication)
IHH	Chicken	Short legs	Lethal	1	Large SV (11.9-kb del)
KIT	Camel	White patterns	Lethal	1	Frameshift (1-bp del)
KIT	Dog	White patterns	Lethal	1	Frameshift (1-bp ins)
KIT	Donkey	White patterns	Lethal—not proven	1	Missense (SNP)
				1	Splice Site mutation (SNP)
KIT	Fox	White patterns	Lethal	1	Missense (SNP)
				1	Splice Site mutation (SNP)
KIT	Horse	White patterns	Lethal	3	Frameshift (1-bp, 4-bp del)
				2	Large SV (1.3-kb, 1.9-kb del)
				2	Nonsense (SNP)
				1	In-Frame Del (54-bp)
				1	Splice Site mutation (SNP)
KIT	Pig	White patterns	Lethal—not proven	1	Splice Site mutation (SNP)
MITF	Buffalo	White patterns	Microphtalmia (inferred from mice)	1	Nonsense (SNP)
				1	Splice Site mutation (SNP)
MITF	Dog	White patterns	Deafness (risk increase[a])	1	Cis-regulatory (195-bp ins)
MITF	Horse	White patterns	Microphtalmia (inferred from mice)	2	Large SV (8.7-kb, 63-kb del)
				1	Frameshift (4-bp del)
				1	Missense (SNP)
MITF	Quail	White patterns	Small size and slow growth	1	Frameshift (2-bp del)
Mlana	Pigeon	White patterns	Eye defects	4	CNV including 3 other genes
MNR2 locus	Chicken	Head comb	Male infertility	1	Large SV (7.4-Mb inversion)
MRC2	Cattle	High muscle	Musculoskeletal defects	1	Frameshift (2-bp del)
				1	Missense (SNP)
MSTN[c,d]	Dog	Racing performance	Nonconform appearance	1	Frameshift (2-bp del)
PAX3	Horse	White patterns	Lethal	2	Missense (SNP)
PMEL17	Dog	White patterns	Auditory and ocular defects	6	Splice Site mutation (TE insertions)
PMEL17	Horse	White patterns	Ocular defects	1	Missense (SNP)
RNASEH2Ba locus	Cattle	Milk yield (causal gene unknown)	Lethal (due to *RNASEH2Ba* KO)	1	Large SV (660-kb deletion)
RYR1	Pig	Meat yield and quality	Hyperthermia	1	Missense (SNP)
PRLR/SPEF2	Pig	Increased female fecundity	Sperm defects (due to *SPEF2* KO)	1	Cis-regulatory (TE insertion)
SPINT1	Gecko	White patterns	Metastases	1	Unknown
STX17	Horse	White patterns	Melanoma	1	Cis-regulatory (4.6-kb duplication)
TBXT (Brachyury)	Cat	Short or no tail	Lethal	4	Frameshift (1-bp del; 14-bp ins)
TBXT (Brachyury)	Dog	Short or no tail	Lethal	1	Missense (SNP)
TRPM1	Horse	White patterns	Ocular defects	1	Cis-regulatory (TE insertion)
TRPV4	Cat	Short ears	Chondrodysplasia	1	Missense (SNP)

[a] The corresponding entries and references can be retrieved using a search for "@HeterozygoteAdvantage" or for the gene and species names at www.gephebase.org.
[b] This study.
[c] Recommendations to outcross are recent in the affected stock.
[d] *MSTN* loss-of-function alleles exist in other bred species. Homozygotes sometimes requires assisted birthing, but are viable and artificially selected.

lethal. Large SVs such as the *ivory* deletion accounted for 12 out of 83 derived alleles in this data set, suggesting that the recessive deleterious effects of macromutations occasionally provide heterozygous states of interest to artificial selection. Of note, 17 out the 40 cases of captive heterozygote advantage involve selection for depigmentation traits, highlighting the trend among breeders and fanciers to select for conspicuous variants (Table 2). Finally, cases of heterozygous advantage also occur in the wild, such as in ruff birds where a large inversion haplotype provides male coloration and behavioral traits that are maintained by sexual selection in spite of being recessive lethal (Küpper *et al.* 2016). In wild populations of the polymorphic butterfly *H. numata*, *cortex* itself is situated in the center of an inversion polymorphism under balancing selection (Maisonneuve *et al.* 2021). Further work will be needed to determine the subgene level elements that are within the *ivory* deletion and mediate homozygous inviability.

Conclusion

By using autozygosity mapping and association on a comparatively small pool of individuals, we were able to identify a structural mutation involved with butterfly wing pattern. This approach may prove fruitful in other studies of butterfly wing patterning or in the identification of de novo mutants, especially if combined with recent developments in whole-genome sequencing from dried museum specimens (Cong *et al.* 2021; Grewe *et al.* 2021). This could allow the mapping of other cases like *Hindsight* in *J. coenia*, or *pseudozorro* in *Parnassius apollo* (Weatherbee *et al.* 1999; Pierrat and Descimon 2011). The identification and characterization of spontaneous mutants has provided very valuable insights into the genetics of development in model organisms. Further whole-genome sequencing of deleterious mutations in butterflies will help to identify parts of the genome that are functionally required for normal development, which will assist in a more complete understanding of their evolution and development.

Data availability

Whole-genome sequences are accessible on the NCBI SRA under the Bioproject accession number PRJNA610063.

Supplemental material is available at G3 online.

Acknowledgments

We thank Ron Boender of Butterfly World, FL for generously providing living specimens for study by the Gilbert Lab at UT Austin. Thanks to greenhouse technicians T. Freiburger, N. Fogel, I. Terry, and E. Reese for keeping plant and butterfly cultures healthy in the Austin facility during the study, and G. Julian at the University of Cambridge for expert stock maintenance of *H. erato*. We also thank the GWU HPC team for computing infrastructure (MacLachlan *et al.* 2020), and S.M. Van Belleghem for comments on an earlier draft.

Funding

This work was supported by National Science Foundation awards IOS-2110534 to AM and IOS-2110532 to WOM, a George Washington University University Facilitating Fund research grant to AM and JJH, Biotechnology and Biological Sciences Resarch Council grant BB/R007500 to CDJ and LL, and past National Science Foundation grants for greenhouse facilities as well as the Worthington Foundation Endowment to LEG.

Conflicts of interest

None declared.

Literature cited

Andersson L. Domestic animals as models for biomedical research. Ups J Med Sci. 2016;121(1):1–11. doi:10.3109/03009734.2015.1091522.

Andersson L, Georges M. Domestic-animal genomics: deciphering the genetics of complex traits. Nat Rev Genet. 2004;5(3):202–212. doi:10.1038/nrg1294.

Andrade P, Pinho C, Pérez I de Lanuza G, Afonso S, Brejcha J, Rubin C-J, Wallerman O, Pereira P, Sabatino SJ, Bellati A, *et al.* Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. Proc Natl Acad Sci USA. 2019;116(12):5633–5642. doi:10.1073/pnas.1820320116.

Berry SD, Davis SR, Beattie EM, Thomas NL, Burrett AK, Ward HE, Stanfield AM, Biswas M, Ankersmit-Udy AE, Oxley PE, *et al.* Mutation in bovine beta-carotene oxygenase 2 affects milk color. Genetics. 2009;182(3):923–926. doi:10.1534/genetics.109.101741.

Bhardwaj V, Semplicio G, Erdogdu NU, Manke T, Akhtar A. MAPCap allows high-resolution detection and differential expression analysis of transcription start sites. Nat Commun. 2019;10:3219.doi:10.1038/s41467-019–11115-x.

Bonilla C, Boxill L-A, Donald SAM, Williams T, Sylvester N, Parra EJ, Dios S, Norton HL, Shriver MD, Kittles RA. The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. Hum Genet. 2005;116(5):402–406. doi:10.1007/s00439-004–1251-2.

Bosse M, Megens H-J, Derks MFL, Cara ÁMR, de Groenen MAM. Deleterious alleles in the context of domestication, inbreeding, and selection. Evol Appl. 2019;12(1):6–17. doi:10.1111/eva.12691.

Boukhatmi H, Bray S. A population of adult satellite-like cells in *Drosophila* is maintained through a switch in RNA-isoforms. eLife. 2018;7:e35954.doi:10.7554/eLife.35954.

Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. Am J Hum Genet. 2007;81(5):1084–1097. doi:10.1086/521987.

Bruders R, Hollebeke HV, Osborne EJ, Kronenberg Z, Maclary E, Yandell M, Shapiro MD. A copy number variant is associated with a spectrum of pigmentation patterns in the rock pigeon (*Columba livia*). PLoS Genet. 2020;16(5):e1008274.doi:10.1371/journal.pgen.1008274.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):421.doi:10.1186/1471–2105-10–421.

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

Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, *et al.* Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a Pitx1 enhancer. Science. 2010;327(5963):302–305. doi:10.1126/science.1182213.

Cieslak M, Reissmann M, Hofreiter M, Ludwig A. Colours of domestication. Biol Rev. 2011;86(4):885–899. doi:10.1111/j.1469-185X.2011.00177.x.

Cong Q, Shen J, Zhang J, Li W, Kinch LN, Calhoun JV, Warren AD, Grishin NV. Genomics reveals the origins of historical specimens. Mol Biol Evol. 2021;38(5):2166–2176. doi:10.1093/molbev/msab013.

Courtier-Orgogozo V, Arnoult L, Prigent SR, Wiltgen S, Martin A. Gephebase, a database of genotype-phenotype relationships for natural and domesticated variation in Eukaryotes. Nucleic Acids Res. 2020;48(D1):D696–D703. doi:10.1093/nar/gkz796.

Courtier-Orgogozo V, Martin A. The coding loci of evolution and domestication: current knowledge and implications for bio-inspired genome editing. J Exp Biol. 2020;223:jeb208934. doi:10.1242/jeb.208934.

Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin AV, Hughes DST, Ferguson LC, Martin SH, *et al.*; The Heliconius Genome Consortium. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature. 2012;487:94–98. doi:10.1038/nature11041.

Domyan ET, Kronenberg Z, Infante CR, Vickrey AI, Stringham SA, Bruders R, Guernsey MW, Park S, Payne J, Beckstead RB, *et al.* Molecular shifts in limb identity underlie development of feathered feet in two domestic avian species. eLife. 2016;5:e12115.doi:10.7554/eLife.12115.

Domyan ET, Shapiro MD. Pigeonetics takes flight: evolution, development, and genetics of intraspecific variation. Dev Biol. 2017;427(2):241–250. doi:10.1016/j.ydbio.2016.11.008.

Driscoll CA, Macdonald DW, O'Brien SJ. From wild animals to domestic pets, an evolutionary view of domestication. Proc Natl Acad Sci USA. 2009;106(Suppl 1):9971–9978. doi:10.1073/pnas.0901586106.

Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-Accinelli G, Belleghem SMV, Patterson N, Neafsey DE, *et al.* Genomic architecture and introgression shape a butterfly radiation. Science. 2019;366(6465):594–599. doi:10.1126/science.aaw2090.

Ficarrotta V, Hanly JJ, Loh LS, Francescutti CM, Ren A, Tunström K, Wheat CW, Porter AH, Counterman BA, Martin A. A genetic switch for male UV-iridescence in an incipient species pair of sulphur butterflies. Proceedings of the National Academy of Sciences. 2022 Jan 18;119(3).

Fontanesi L, Forestier L, Allain D, Scotti E, Beretti F, Deretz-Picoulet S, Pecchioli E, Vernesi C, Robinson TJ, Malaney JL, *et al.* Characterization of the rabbit agouti signaling protein (ASIP) gene: transcripts and phylogenetic analyses and identification of the causative mutation of the nonagouti black coat colour. Genomics. 2010;95(3):166–175. doi:10.1016/j.ygeno.2009.11.003.

Futahashi R, Osanai-Futahashi M. Pigments in insects. In: H Hashimoto, M Goda, R Futahashi, R Kelsh, T Akiyama, editors. Pigments, Pigment Cells and Pigment Patterns. Singapore: Springer; 2021. p. 3–43. doi:10.1007/978–981-16–1490-3_1.

Gilbert LE, Forrest HS, Schultz TD, Harvey DJ. Correlations of ultrastructure and pigmentation suggest how genes control development of wing scales of *Heliconius* butterflies. J Res Lepidoptera USA. 1987;26:141–160.

Giovannini A, Laura M, Nesi B, Savona M, Cardi T. Genes and genome editing tools for breeding desirable phenotypes in ornamentals. Plant Cell Rep. 2021;40(3):461–478. doi:10.1007/s00299-020–02632-x.

Gratten J, Pilkington JG, Brown EA, Beraldi D, Pemberton JM, Slate J. The genetic basis of recessive self-colour pattern in a wild sheep population. Heredity. 2010;104(2):206–214. doi:10.1038/hdy.2009.105.

Grewe F, Kronforst MR, Pierce NE, Moreau CS. n.d. Museum genomics reveals the Xerces blue butterfly (*Glaucopsyche xerces*) was a distinct species driven to extinction. Biol Lett. 2021;17(7):20210123.doi:10.1098/rsbl.2021.0123.

Guo L, Bloom J, Sykes S, Huang E, Kashif Z, Pham E, Ho K, Alcaraz A, Xiao XG, Duarte-Vogel S, *et al.* Genetics of white color and iridophoroma in "Lemon Frost" leopard geckos. PLoS Genet. 2021;17(6):e1009580.doi:10.1371/journal.pgen.1009580.

Hanly JJ, Wallbank RWR, McMillan WO, Jiggins CD. Conservation and flexibility in the gene regulatory landscape of heliconiine butterfly wings. EvoDevo. 2019;10(1):15.doi:10.1186/s13227-019–0127-4.

Hedrick PW. Heterozygote advantage: the effect of artificial selection in livestock and pets. J Hered. 2015;106(2):141–154. doi:10.1093/jhered/esu070.

Jiggins CD, Wallbank RWR, Hanly JJ. Waiting in the wings: what can we learn about gene co-option from the diversification of butterfly wing patterns? Philos Phil Trans R Soc B. 2017;372(1713):20150485.doi:10.1098/rstb.2015.0485.

Jones MR, Mills LS, Alves PC, Callahan CM, Alves JM, Lafferty DJR, Jiggins FM, Jensen JD, Melo-Ferreira J, Good JM. Adaptive introgression underlies polymorphic seasonal camouflage in snowshoe hares. Science. 2018;360(6395):1355–1358. doi:10.1126/science.aar5273.

Joron M, Frezal L, Jones RT, Chamberlain NL, Lee SF, Haag CR, Whibley A, Becuwe M, Baxter SW, Ferguson L, *et al.* Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature. 2011;477(7363):203–206. doi:10.1038/nature10341.

Junker J, Rick JA, McIntyre PB, Kimirei I, Sweke EA, Mosille JB, Wehrli B, Dinkel C, Mwaiko S, Seehausen O, *et al.* Structural genomic variation leads to genetic differentiation in Lake Tanganyika's sardines. Mol Ecol. 2020;29(17):3277–3298. doi:10.1111/mec.15559.

Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–3066. doi:10.1093/nar/gkf436.

Küpper C, Stocks M, Risse JE, Dos Remedios N, Farrell LL, McRae SB, Morgan TC, Karlionova N, Pinchuk P, Verkuil YI, *et al.* A supergene determines highly divergent male reproductive morphs in the ruff. Nat Genet. 2016;48(1):79–83. doi:10.1038/ng.3443.

Kuriyama T, Murakami A, Brandley M, Hasegawa M. Blue, black, and stripes: evolution and development of color production and pattern formation in lizards and snakes. Front Ecol Evol. 2020;8:232.doi:10.3389/fevo.2020.00232.

Langevin M, Synkova H, Jancuskova T, Pekova S. Merle phenotypes in dogs – SILV SINE insertions from Mc to Mh. PLoS One. 2018;13(9):e0198536.doi:10.1371/journal.pone.0198536.

Lewis JJ, Geltman RC, Pollak PC, Rondem KE, Belleghem SMV, Hubisz MJ, Munn PR, Zhang L, Benson C, Mazo-Vargas A, *et al.* Parallel evolution of ancient, pleiotropic enhancers underlies butterfly wing pattern mimicry. Proc Natl Acad Sci USA. 2019;116(48):24174–24183. doi:10.1073/pnas.1907068116.

Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv 13033997 Q-Bio; 2013. https://arxiv.org/abs/1303.3997.

Livraghi L, Hanly JJ, Belleghem SMV, Montejo-Kovacevich G, Heijden ESM, van der Loh LS, Ren A, Warren IA, Lewis JJ, Concha C, López, *et al.* Cortex cis-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*. ELife. 2021 Jul 19;10:e68549.

Livraghi L, Martin A, Gibbs M, Braak N, Arif S, Breuker CJ. Chapter Three - CRISPR/Cas9 as the key to unlocking the secrets of butterfly wing pattern development and its evolution. In: RH Ffrench-Constant, editor. Advances in Insect Physiology, Butterfly Wing

Patterns and Mimicry. Cambridge (MA): Academic Press; 2018. p. 85–115. doi:10.1016/bs.aiip.2017.11.001.

MacLachlan G, Hurlburt J, Suarez M, Wong KL, Burke W, Lewis T, Gallo A, Flidr J, Gabiam R, Nicholas J, *et al.* Building a shared resource HPC center across university schools and institutes: a case study. ArXiv 200313629 Cs; 2020.

Maisonneuve L, Chouteau M, Joron M, Llaurens V. Evolution and genetic architecture of disassortative mating at a locus under heterozygote advantage. Evolution. 2021;75(1):149–165. doi:10.1111/evo.14129.

Makino T, Rubin C-J, Carneiro M, Axelsson E, Andersson L, Webster MT. Elevated proportions of deleterious genetic variation in domestic animals and plants. Genome Biol Evol. 2018;10(1):276–290. doi:10.1093/gbe/evy004.

Marsden CD, Vecchyo DO-D, O'Brien DP, Taylor JF, Ramirez O, Vilà C, Marques-Bonet T, Schnabel RD, Wayne RK, Lohmueller KE. Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proc Natl Acad Sci USA. 2016;113(1):152–157. doi:10.1073/pnas.1512501113.

Martin A, Papa R, Nadeau NJ, Hill RI, Counterman BA, Halder G, Jiggins CD, Kronforst MR, Long AD, McMillan WO, *et al.* Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. Proc Natl Acad Sci USA. 2012;109(31):12632–12637. doi:10.1073/pnas.1204800109.

Martin SH, Davey JW, Salazar C, Jiggins CD. Recombination rate variation shapes barriers to introgression across butterfly genomes. PLoS Biol. 2019;17(2):e2006288.doi:10.1371/journal.pbio.2006288.

Martin SH, Singh KS, Gordon IJ, Omufwoko KS, Collins S, Warren IA, Munby H, Brattström O, Traut W, Martins DJ, *et al.* Whole-chromosome hitchhiking driven by a male-killing endosymbiont. PLOS Biol. 2020;18(2):e3000610.doi:10.1371/journal.pbio.3000610.

Martin SH, Van Belleghem SM. Exploring evolutionary relationships across the genome using topology weighting. Genetics. 2017;206(1):429–438. doi:10.1534/genetics.116.194720.

Mazo-Vargas A, Concha C, Livraghi L, Massardo D, Wallbank RWR, Zhang L, Papador JD, Martinez-Najera D, Jiggins CD, Kronforst MR, *et al.* Macroevolutionary shifts of WntA function potentiate butterfly wing-pattern diversity. Proc Natl Acad Sci USA. 2017;114(40):10701–10706. doi:10.1073/pnas.1708149114.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al.* The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297–1303. doi:10.1101/gr.107524.110.

McLean CY, Reno PL, Pollen AA, Bassan AI, Capellini TD, Guenther C, Indjeian VB, Lim X, Menke DB, Schaar BT, *et al.* Human-specific loss of regulatory DNA and the evolution of human-specific traits. Nature. 2011;471(7337):216–219. doi:10.1038/nature09774.

McMillan WO, Livraghi L, Concha C, Hanly JJ. From patterning genes to process: unraveling the gene regulatory networks that pattern *Heliconius* wings. Front Ecol Evol. 2020;8:221. doi:10.3389/fevo.2020.00221.

Mérot C, Oomen RA, Tigano A, Wellenreuther M. A roadmap for understanding the evolutionary significance of structural genomic variation. Trends Ecol Evol. 2020;35(7):561–572. doi:10.1016/j.tree.2020.03.002.

Metallinos DL, Bowling AT, Rine J. A missense mutation in the endothelin-B receptor gene is associated with Lethal White Foal Syndrome: an equine version of Hirschsprung disease. Mamm Genome. 1998;9(6):426–431. doi:10.1007/s003359900790.

Moest M, Belleghem SMV, James JE, Salazar C, Martin SH, Barker SL, Moreira GRP, Mérot C, Joron M, Nadeau NJ, *et al.* Selective sweeps on novel and introgressed variation shape mimicry loci in a butterfly adaptive radiation. PLoS Biol. 2020;18(2):e3000597.doi:10.1371/journal.pbio.3000597.

Morris J, Hanly JJ, Martin SH, Van Belleghem SM, Salazar C, Jiggins CD, Dasmahapatra KK. Deep convergence, shared ancestry, and evolutionary novelty in the genetic architecture of *Heliconius* mimicry. Genetics. 2020;216(3):765–780. doi:10.1534/genetics.120.303611.

Moyers BT, Morrell PL, McKay JK. Genetic costs of domestication and improvement. J Hered. 2018;109(2):103–116. doi:10.1093/jhered/esx069.

Mundy NI, Stapley J, Bennison C, Tucker R, Twyman H, Kim K-W, Burke T, Birkhead TR, Andersson S, Slate J. Red carotenoid coloration in the zebra finch is controlled by a cytochrome P450 gene cluster. Curr Biol. 2016;26(11):1435–1440. doi:10.1016/j.cub.2016.04.047.

Nadeau NJ, Minvielle F, Ito S, Inoue-Murayama M, Gourichon D, Follett SA, Burke T, Mundy NI. Characterization of Japanese quail yellow as a genomic deletion upstream of the avian homolog of the mammalian ASIP (agouti) gene. Genetics. 2008;178(2):777–786. doi:10.1534/genetics.107.077073.

Nadeau NJ, Pardo-Diaz C, Whibley A, Supple MA, Saenko SV, Wallbank RWR, Wu GC, Maroja L, Ferguson L, Hanly JJ, *et al.* The gene *cortex* controls mimicry and crypsis in butterflies and moths. Nature. 2016;534(7605):106–110. doi:10.1038/nature17961.

Nijhout HF, Wray GA. Homologies in the colour patterns of the genus *Heliconius* (Lepidoptera: Nymphalidae). Biol J Linn Soc. 1988;33(4):345–365. doi:10.1111/j.1095–8312.1988.tb00449.x.

Norris BJ, Whan VA. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. Genome Res. 2008;18(8):1282–1293. doi:10.1101/gr.072090.107.

Park K-I, Ishikawa N, Morita Y, Choi J-D, Hoshino A, Iida S. A bHLH regulatory gene in the common morning glory, Ipomoea purpurea, controls anthocyanin biosynthesis in flowers, proanthocyanidin and phytomelanin pigmentation in seeds, and seed trichome formation. Plant J. 2007;49(4):641–654. doi:10.1111/j.1365-313X.2006.02988.x.

Pierrat V, Descimon H. A new wing pattern mutant in the Apollo butterfly, *Parnassius apollo* (L. 1758) (Lepidoptera: Papilionidae). Ann Soc Entomol Fr NS. 2011;47(3–4):293–302. doi:10.1080/00379271.2011.10697722.

Pinharanda A, Rousselle M, Martin SH, Hanly JJ, Davey JW, Kumar S, Galtier N, Jiggins CD. Sexually dimorphic gene expression and transcriptome evolution provide mixed evidence for a fast-Z effect in *Heliconius*. J Evol Biol. 2019;32(3):194–204. doi:10.1111/jeb.13410.

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet. 2006;38(8):904–909. doi:10.1038/ng1847.

Price-Waldman R, Stoddard MC. Avian coloration genetics: recent advances and emerging questions. J Hered. 2021;112(5):395–416. doi:10.1093/jhered/esab015.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–575. doi:10.1086/519795.

Reed RD, Papa R, Martin A, Hines HM, Counterman BA, Pardo-Diaz C, Jiggins CD, Chamberlain NL, Kronforst MR, Chen R, *et al.* optix drives the repeated convergent evolution of butterfly wing pattern mimicry. Science. 2011;333(6046):1137–1141. doi:10.1126/science.1208227.

Rieder S, Taourit S, Mariat D, Langlois B, Guérin G. Mutations in the agouti (ASIP), the extension (MC1R), and the brown (TYRP1) loci

and their association to coat color phenotypes in horses (*Equus caballus*). Mamm Genome. 2001;12(6):450–455. doi: 10.1007/s003350020017.

Singer GA, Wu J, Yan P, Plass C, Huang TH, Davuluri RV. Genome-wide analysis of alternative promoters of human genes using a custom promoter tiling array. BMC Genomics. 2008;9(1):349.doi: 10.1186/1471–2164-9–349.

Spelt C, Quattrocchio F, Mol J, Koes R. ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. Plant Cell. 2002;14(9): 2121–2135. doi:10.1105/tpc.003772.

Strain GM, Clark LA, Wahl JM, Turner AE, Murphy KE. Prevalence of deafness in dogs heterozygous or homozygous for the merle allele. J Vet Intern Med. 2009;23(2):282–286. doi: 10.1111/j.1939-1676.2008.0257.x.

Surridge AK, Lopez-Gomollon S, Moxon S, Maroja LS, Rathjen T, Nadeau NJ, Dalmay T, Jiggins CD. Characterisation and expression of microRNAs in developing wings of the neotropical butterfly *Heliconius melpomene*. BMC Genomics. 2011;12(1):62.doi: 10.1186/1471–2164-12–62.

Toews DPL, Taylor SA, Vallender R, Brelsford A, Butcher BG, Messer PW, Lovette IJ. Plumage genes and little else distinguish the genomes of hybridizing warblers. Curr Biol. 2016;26(17): 2313–2318. doi:10.1016/j.cub.2016.06.034.

Van Belleghem SM, Lewis JJ, Rivera ES, Papa R. *Heliconius* butterflies: a window into the evolution and development of diversity. Curr Opin Genet Dev. 2021;69:72–81. doi:10.1016/j.gde.2021.01.010.

Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J, Lewis JJ, Hines HM, *et al.* Complex modular architecture around a simple toolkit of wing pattern genes. Nat Ecol Evol. 2017;1:1–12. doi:10.1038/s41559-016–0052.

van der Burg KRL, Lewis JJ, Brack BJ, Fandino RA, Mazo-Vargas A, Reed RD. Genomic architecture of a genetically assimilated seasonal color pattern. Science. 2020;370(6517):721–725. doi: 10.1126/science.aaz3017.

VanKuren NW, Massardo D, Nallu S, Kronforst MR. Butterfly mimicry polymorphisms highlight phylogenetic limits of gene reuse in the evolution of diverse adaptations. Mol Biol Evol. 2019;36(12): 2842–2853. 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. The industrial melanism mutation in British peppered moths is a transposable element. Nature. 2016;534(7605):102–105. doi:10.1038/nature17951.

van't Hof AE, Reynolds LA, Yung CJ, Cook LM, Saccheri IJ. Genetic convergence of industrial melanism in three geometrid moths. Biol Lett. 2019;15(10):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, *et al.* Evolutionary novelty in a butterfly wing pattern through enhancer shuffling. PLOS Biol. 2016;14(1):e1002353.doi: 10.1371/journal.pbio.1002353.

Weatherbee SD, Frederik Nijhout H, Grunert LW, Halder G, Galant R, Selegue J, Carroll S. Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. Curr Biol. 1999;9(3): 109–115. doi:10.1016/S0960-9822(99)80064-5.

Yu S, Wang G, Liao J. Association of a novel SNP in the ASIP gene with skin color in black-bone chicken. Anim Genet. 2019;50(3): 283–286. doi:10.1111/age.12768.

Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18(5):821–829. doi:10.1101/gr.074492.107.

Zhang L, Mazo-Vargas A, Reed RD. Single master regulatory gene coordinates the evolution and development of butterfly color and iridescence. Proc Natl Acad Sci USA. 2017;114(40):10707–10712. doi:10.1073/pnas.1709058114.

Communicating editor: B. Oliver