

# Genome-Wide Regulatory Adaptation Shapes Population-Level Genomic Landscapes in *Heliconius*

James J. Lewis<sup>\*,1</sup> and Robert D. Reed<sup>1</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY

\*Corresponding author: E-mail: [jjl336@cornell.edu](mailto:jjl336@cornell.edu).

Associate editor: Patricia Wittkopp

## Abstract

**Cis-regulatory evolution is an important engine of organismal diversification. Although recent studies have looked at genomic patterns of regulatory evolution between species, we still have a poor understanding of the magnitude and nature of regulatory variation within species. Here, we examine the evolution of regulatory element activity over wing development in three *Heliconius erato* butterfly populations to determine how regulatory variation is associated with population structure. We show that intraspecific divergence in chromatin accessibility and regulatory activity is abundant, and that regulatory variants are spatially clustered in the genome. Regions with strong population structure are highly enriched for regulatory variants, and enrichment patterns are associated with developmental stage and gene expression. We also found that variable regulatory elements are particularly enriched in species-specific genomic regions and long interspersed nuclear elements. Our findings suggest that genome-wide selection on chromatin accessibility and regulatory activity is an important force driving patterns of genomic divergence within *Heliconius* species. This work also provides a resource for the study of gene regulatory evolution in *H. erato* and other heliconiine butterflies.**

**Key words:** evolutionary epigenetics, gene regulatory evolution, evolution and development, evolutionary genomics.

## Introduction

*Cis*-regulatory variants have been widely implicated as causal elements in numerous ecological, morphological, behavioral, and sexual adaptations in a wide range of eukaryotes (Wray 2007; Wittkopp and Kalay 2012; Albert and Kruglyak 2015). Unfortunately, we still have a poor understanding of how regulatory variation is structured within and between populations, and how this variation is connected to other factors such as genomic variation, introgression, local selection, and development, which has recently been shown to be important for regulatory macroevolution (Lewis et al. 2016). Furthermore, our limited knowledge of population-level regulatory variation largely fails to incorporate chromatin accessibility alongside other markers of regulatory activity (Li et al. 2011). Here we use *Heliconius erato* clade butterflies, which have a well-understood population structure (Flanagan et al. 2004), as a model to characterize the relationship between variation in regulatory activity and intraspecific genomic divergence.

*Heliconius* butterflies are well known for their Müllerian mimicry, where multiple species converge on locally adapted wing phenotypes as a shared warning signal to predators, producing population boundaries driven by wing phenotype (Supple et al. 2015). *Heliconius erato* radiated throughout South and Central America ~2–4 Ma, rapidly diversifying into dozens of morphs (Flanagan et al. 2004; Joron et al. 2006; Hoyal Cuthill and Charleston 2012). We study three populations of *H. erato* clade butterflies—*H. erato petiverana*,

*H. erato lativitta*, and the incipient species *H. himera*—to determine how population boundaries and limited introgression have driven localized regulatory adaptation between allopatric and parapatric populations of *Heliconius* butterflies. *Heliconius e. petiverana* inhabits much of Central America, while *H. e. lativitta* is found in the western Amazon Basin and *H. himera* inhabits more arid, high-elevation locales in the western Andean regions of Ecuador and Peru (Jiggins et al. 1996; Supple et al. 2015). Direct introgression occurs between *H. e. lativitta* and *H. himera* in three narrow hybrid zones in Andean valleys (Jiggins et al. 1996), while indirect introgression potentially occurs between *H. e. petiverana* and the two South American populations via hybridization chains incorporating geographically interposed races of *H. erato*.

In this study, we address the question of how chromatin state and regulatory activity are associated with intraspecific genomic divergence in the *H. erato* clade. Focusing on developing wing tissue we observed widespread, population-level variation between parapatric and allopatric populations of the *H. erato* clade, and found that regulatory loci show very different patterns of accessibility and activity variation between populations and developmental stages. Importantly, we also found a strong association between population structure and spatial clustering of regulatory variation in the genome, thus suggesting that regional variation in regulatory activity is an important force underlying genomic divergence and local adaptation.

## Results and Discussion

### Population-Level Variation in Regulatory Accessibility and Activity

There are few published population-level analyses of regulatory function (e.g., Kasowski et al. 2013), thus our first goal was simply to characterize the nature and magnitude of regulatory variation in the *H. erato* clade. To do this we assayed chromatin accessibility using ATAC-seq (Buenrostro et al. 2013), and regulatory activity using ChIP-seq for H3K4me3 and H3K27ac histone modifications, which tend to associate with active promoter and enhancer regions (Kharchenko et al. 2011). We generated regulatory profiles for three populations of *H. erato* clade butterflies for mid- and late-pupal stage forewings and hindwings (fig. 1A–C). All assays were highly consistent between replicates and tissues (fig. 1D–E) and passed quality assessment, including exceeding ENCODE standards (fig. 1F–H and supplementary figs. S1–S3, Supplementary Material online), with the vast majority of called peaks being reproducible (supplementary fig. S2, Supplementary Material online). ChIP-seq peak calls were consistent in number with previous examples from vertebrate tissue samples and population studies (Smagulova et al. 2011; Kasowski et al. 2013). We found that transposase accessible sites (TASs) outnumbered histone marks by an average ~4:1 in a given tissue and stage, and across all tissue and stages >90% of histone marks overlapped a TAS (fig. 1F). We observed a high degree of regulatory element reutilization between developmental stages, with 93% of TASs, 68% of H4K4me3, and 60% of H3K27ac marks retained over wing development (supplementary fig. S3, Supplementary Material online). TASs and also H3K4me3 marks, which are enriched at transcription start site (TSS) proximal elements at genes and noncoding RNAs, at active enhancer loci (Boyle, Davis, et al. 2008; Kharchenko et al. 2011), and recently were found to mark sites of enhancer RNA transcription (Henriques et al. 2018), were notably more likely to be shared across developmental stages. This is consistent with previously observed patterns of accessibility and activity of regulatory loci (Li et al. 2011) and a tendency toward tissue and developmental stage specificity of H3K27ac marks, which are often associated with TSS-distal enhancer activity (Kharchenko et al. 2011; Nord et al. 2013).

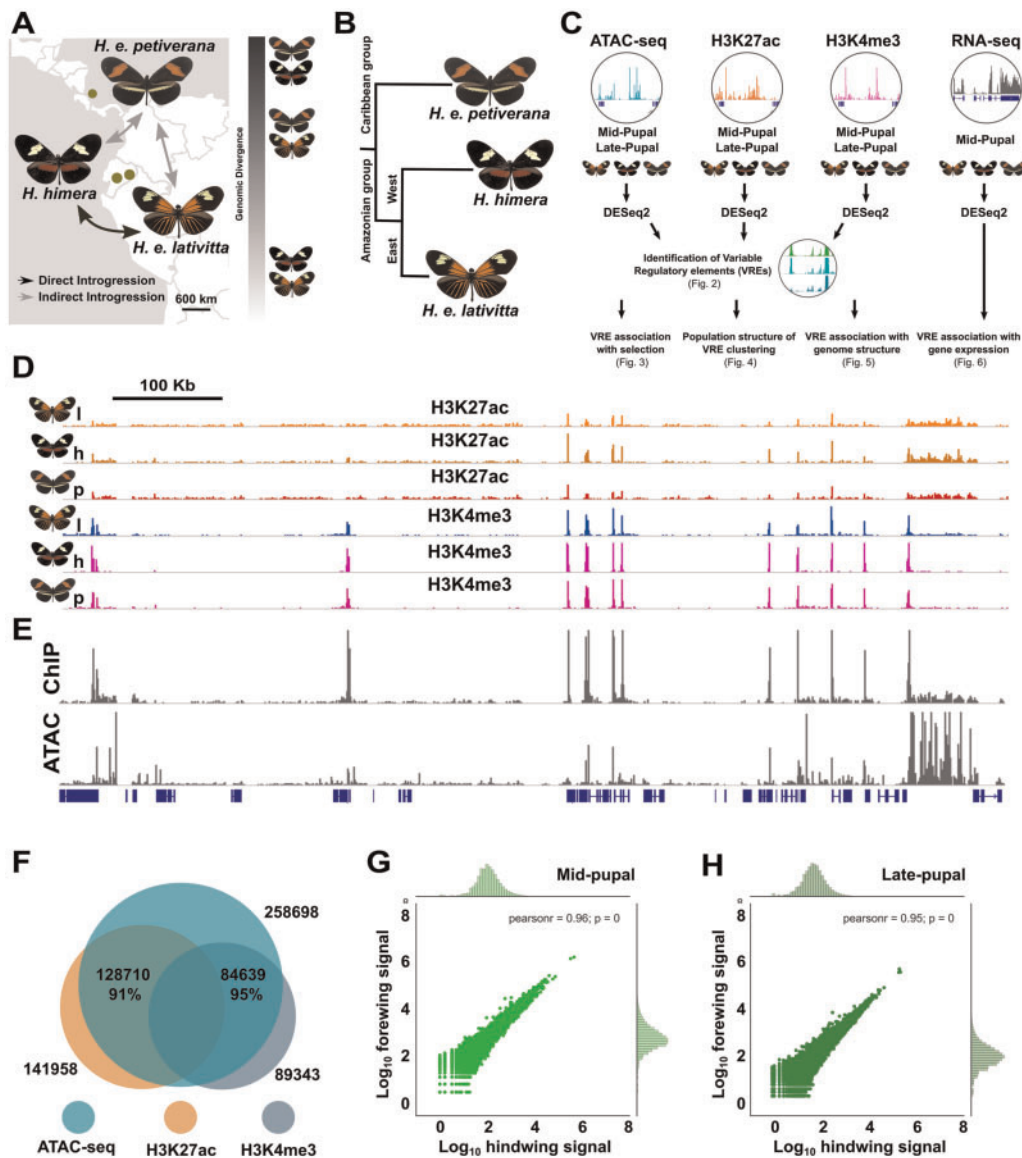
To identify population-level divergence in wing tissue regulatory activity between distinct *H. erato* populations, we used pairwise comparisons to characterize variable regulatory elements (VREs, e.g., fig. 2A–C, supplementary figs. S4–S6 and table S3, Supplementary Material online), using DESeq2 (Love et al. 2014). VREs were abundant throughout the genome, with TASs showing consistently higher levels of variation than histone marks, and a trend toward higher levels of variation later in development (fig. 2D). Similar to previous study of regulatory element macroevolution (Lewis et al. 2016), we found that earlier stage VREs were more likely to be found in regulatory loci active in both developmental stages, and that this was more likely for H3K4me3 and ATAC-seq marked loci than H3K27ac marks (supplementary fig. S3, Supplementary Material online). Median fold change

between populations for each VRE type ranged between 2- and 2.5-fold (fig. 2E). VREs were spatially distributed roughly as expected with variable H3K4me3 marks more TSS-proximal relative to variable H3K27ac marks, while variable TASs were distributed approximately equal to the combined distributions of both histone marks (fig. 2F). Interestingly, VREs were significantly clustered ( $P < 0.01$ , two-sample Kolmogorov–Smirnov test) within the genome for all three assays at both developmental stages (fig. 2G and supplementary fig. S8, Supplementary Material online), indicating that clusters of multiple VREs are likely coregulated. To provide a reference point for our regulatory assays, we performed mRNA-seq assays for mid-pupal wing tissue for each population. Of the annotated genes in the *H. erato* reference assembly (Lewis et al. 2016), about 15–30% were differentially expressed in pairwise population comparisons, while in total 40% of genes showed variation between one or more populations (fig. 2D). This expression variation is similar in magnitude to the combined regulatory divergence between populations, and suggests that much of the observed population VRE structure is likely associated with transcriptional divergence.

Although comparisons across studies must be treated with caution, it is interesting to note that regulatory variation in *H. erato* is on par with estimates of regulatory divergence between vertebrate species isolated by more than 10 My (Villar et al. 2015). Furthermore, TAS variability, for which there is currently no comparable data set, was much higher over all comparisons (supplementary fig. S9, Supplementary Material online) and more closely resembles divergences seen in transcription factor binding site turnover in *Drosophila* (Arnold et al. 2014), though this may in part reflect greater sensitivity of ATAC-seq compared with ChIP-seq assays for histone modifications. In sum, our initial comparison of regulatory elements between *H. erato* clade populations reveals striking levels of variation between populations and raises the question of how this variation might be structured across populations and genomes.

### VREs Underlie Local Adaptation

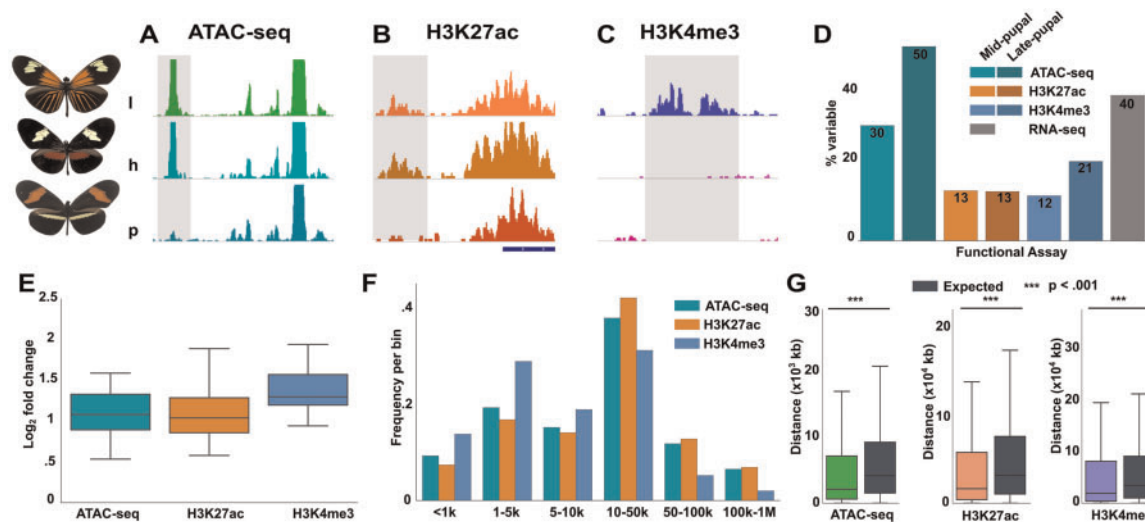
We next sought to determine to what extent VREs are associated with genomic divergence indicative of local adaptation. We used  $F_{st}$  analysis, a statistic comparing intrapopulation with interpopulation nucleotide variation, to identify genomic loci displaying elevated levels of population-specific nucleotide divergence—frequently evidence of natural selection. In total, 7.7 million variants were called using whole-genome resequencing data from four *H. e. lativitta*, six *H. himera*, and five *H. e. petiverana* individuals, and  $F_{st}$  was calculated for pairwise population comparisons in 5-kb bins (supplementary fig. S10, Supplementary Material online). Mean genome-wide  $F_{st}$  values for pairwise comparisons (fig. 3A, marked “G”) were consistent with previously observed divergences between parapatric and allopatric races in *Heliconius* (Martin et al. 2013; Supple et al. 2015). We identified the top 5% of 5-kb bins ranked by  $F_{st}$  as “outlier bins” (fig. 3A, marked “O”), which were distributed across 138 of the 142 scaffolds. Because population divergence is unlikely to be solely due



**Fig. 1.** *Heliconius* populations, experimental design, and data overview. (A) Three populations were used in this study: *Heliconius erato petiverana* (Costa Rica), *H. e. lativitta* (Ecuador), and incipient species *H. himera* (Ecuador). Black arrow indicates direct introgression, and gray arrows indicate indirect introgression. Approximate divergence for each population pair derived from Van Belleghem et al. (2017) based on nucleotide sequence. (B) Pruned phylogenetic tree based on Van Belleghem et al. (2017) shows evolutionary relationships between the study populations. (C) Experimental design for this study. (D) ChIP-seq tracks for mid-pupal forewings showing normalized read depth for H3K27ac (top, oranges) and H3K4me3 (bottom, purples) across ~900 kb of chromosome 19. p, *H. e. petiverana*; l, *H. e. lativitta*; h, *H. himera*. (E) Enrichment profiles for combined H3K27ac and H3K4me3 mid-pupal histone ChIP-seq (top) and ATAC-seq (bottom) tracks show a high degree of similarity between ATAC-seq and ChIP-seq normalized read depth across all loci. This is confirmed in (F), with 91% of H3K27ac and 95% of H3K4me3 ChIP-seq peaks overlapping one or more assayed ATAC-seq peaks by at least 1 bp (aggregate of all samples and stages shown). Correlation plots of  $\text{Log}_{10}$  signal in forewings and hindwings of *H. e. lativitta* show a high degree of similarity between wing tissues at both mid-pupal (G) and late-pupal (H) stages.

to divergence in wing tissue-associated regions of the genome, we separated active from inactive high  $F_{st}$  outlier loci, with a locus labeled “active” if a regulatory mark (variable or not) was present. We then looked at the relationship between VREs and active outlier bins in population comparisons corresponding to the pairwise  $F_{st}$  scans (fig. 3B). We found that variable TASs were present in ~50–85% of active outlier bins, with mid-pupal wings showing a weaker association with regions of high  $F_{st}$ —likely due in part to the overall decreased number of regulatory loci in mid-pupal wings

relative to late-pupal wings (fig. 3B). This trend held for both histone marks as well, and showed the same pattern as observed for the presence of VREs in all outlier bins (supplementary fig. S11, Supplementary Material online). In most cases, the increased presence of late-pupal VREs in outlier bins was highly significant ( $P < 0.01$ , Fisher’s exact test). Regardless of variation between developmental stages, the fact that >50% of 5-kb outlier bins contain VREs suggests that wing development regulatory elements appear to be under strong genome-wide selection in *H. erato*.



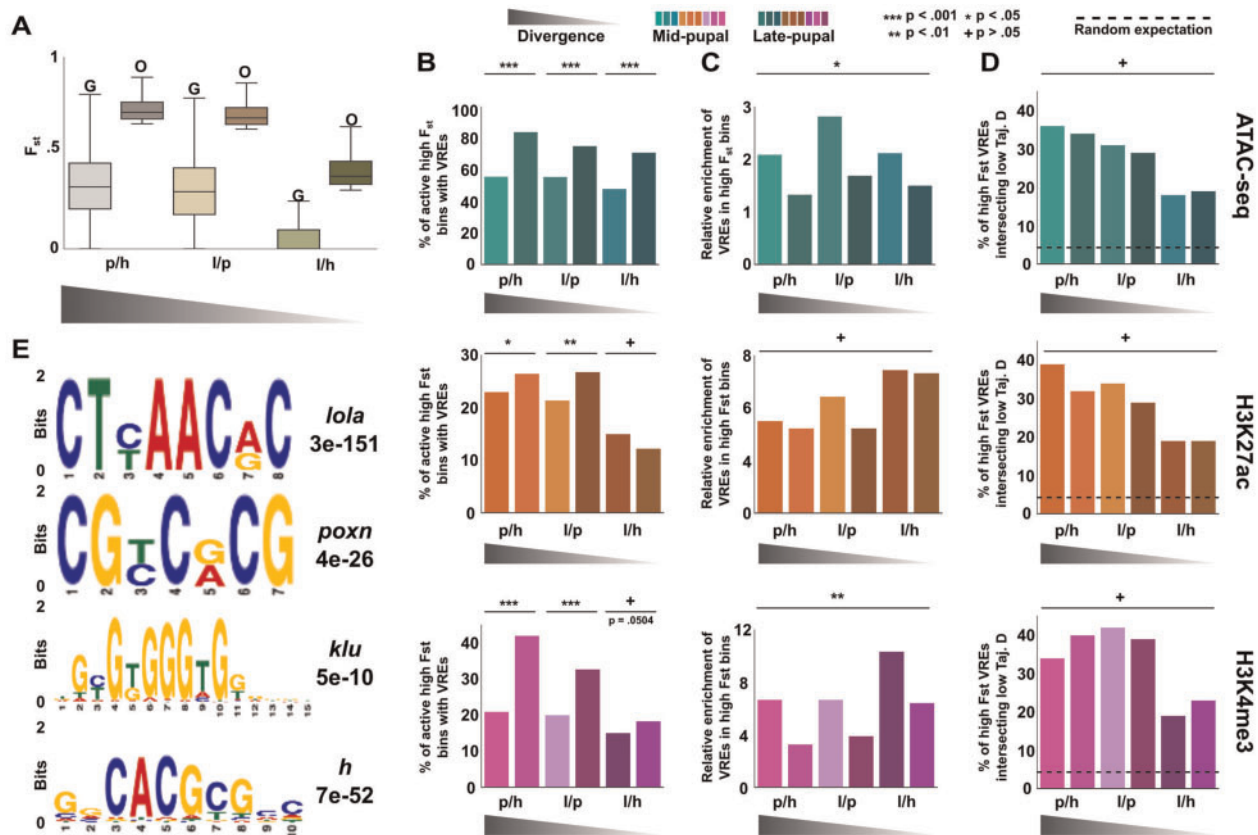
**Fig. 2.** Variation in regulatory loci between three populations of *Heliconius erato*. Examples of VREs displaying population-level variability in normalized read depth for (A) ATAC-seq, (B) H3K27ac ChIP-seq, and (C) H3K4me3 ChIP-seq. (D) Percent of all unique regulatory marks displaying variability between one or more populations by assay. (E) Boxplots showing fold change between population signals for all VRE comparisons. Outliers removed due to very high fold change in the top 1% of data sets representing complete gain or loss of peaks between populations. (F) Distance of VREs from the nearest annotated TSS for ATAC-seq, H3K27ac, and H3K4me3 marked loci. (G) Distance to the next closest VRE for VREs in mid-pupal *H. e. lativitta* wing tissue compared with the same peak sets randomly shuffled throughout the genome (dark blue boxplots). Significance calculated by two-sample Kolmogorov–Smirnov test. For all panels: p, *H. e. petiverana*; l, *H. e. lativitta*; h, *H. himera*.

Interestingly, relative VRE enrichment after adjusting for active genome size within regions of high  $F_{st}$  showed an opposite trend between mid-pupal and late-pupal stages in all three assays for all population pairs (fig. 3C): In all cases, mid-pupal VREs were more enriched in outlier bins than were late-pupal VREs and this was significant for TAS and H3K4me3 VREs (paired Student’s  $t$ -test,  $P < 0.05$ ). Variable H3K4me3 marks displayed the greatest difference of enrichment in these genomic regions, with mid-pupal marks often showing about double the enrichment observed in late-pupal loci. These observations show that VREs with evidence of variable histone modification enrichment, as opposed to a change chromatin accessibility alone, are more likely to be the focus of processes driving population genomic divergence, such as natural selection, and that this difference is overall more pronounced at the mid-pupal patterning stage than during the late-pupal stage associated with structural development. We thus speculate that while late-pupal wing tissues display an overall greater population-level variability in regulatory signal, this is due to relaxed selection on late-acting loci driving structural development.

To verify that VREs associated with population divergence were the focus of directional selective pressure and not a consequence of demography or some other mechanism, we used Tajima’s  $D$ , calculated in 2-kb bins across the genome, to test for enrichment of high  $F_{st}$  VRE loci in regions of low Tajima’s  $D$ . Although Tajima’s  $D$  will likely miss most cases of soft or partial selective sweeps, it provides some indication of the role natural selection plays in VRE-associated genomic divergence. Genome-wide Tajima’s  $D$  distributions were mostly positive, suggesting a fairly recent population decline similar to that observed in *H. melpomene* (Martin et al. 2016)

and indicating that regions of low Tajima’s  $D$  are not due to population expansion (supplementary fig. S12A, Supplementary Material online, labeled “G”). The bottom 2.5% of Tajima’s  $D$  bins, which had average values of between  $\sim -0.75$  and  $-1$  (supplementary fig. S12A, Supplementary Material online, labeled “O”), were selected as genomic regions of low Tajima’s  $D$ . In 30–40% of both pairwise comparison incorporating the most divergent population, *H. e. petiverana*, VRE-associated regions of high  $F_{st}$  were also regions of low Tajima’s  $D$  in at least one of the two populations compared (fig. 3D). This was 6–9 times greater than would be expected by chance alone, and strongly suggests that these regions are the focus of a moderate-to-high degree of directional selection. Interestingly, the two more closely related populations showed a lesser degree of enrichment in low Tajima’s  $D$  loci ( $\sim 20\%$ , 4.5 times greater than expected), suggesting that very recent, soft, or partial selective sweeps around VREs that are undetected by Tajima’s  $D$  play an important role in the early stages of population divergence.

To determine the genome-wide importance of VREs as a target of natural selection, we tested whether VREs for each assay type were more often found in negative-valued Tajima’s  $D$  loci. VREs were on average 175–215% more likely to appear in genomic regions with negative Tajima’s  $D$  values than expected, with H3K4me3 VREs showing a significantly greater frequency than TAS VREs ( $P < 0.01$ , Student’s  $t$ -test) and a borderline significant increase over variable H3K27ac marks ( $P = 0.063$ , Student’s  $t$ -test, supplementary fig. S12B, Supplementary Material online). To better understand how these VREs are targeted by natural selection, we performed a linear regression to test for an association between peak magnitude, determined by normalized read counts, and



**Fig. 3.** Genome-wide evidence of selection around VREs. (A) Genome-wide  $F_{st}$  for all 5-kb bins (marked “G”) and for the top 5% outlier bins (marked “O”) for all three population comparisons shows much greater population structure in outlier bins. (B) Percent of active outlier bins (showing any regulatory activity) with VREs for all three population comparisons and assays shows significant overrepresentation in outlier-associated late-pupal wings relative to mid-pupal wings. Significance determined by Fisher’s exact test. (C) Fold enrichment of VREs in outlier bins relative to active outlier genome size shows greater concentration of high  $F_{st}$  VREs in mid-pupal wings over late-pupal wings. (D) Percent of high  $F_{st}$  VREs that overlap with the bottom 2.5% of Tajima’s  $D$  bins in either population shows large enrichment of loci under strong directional selection relative to a random expectation. Significance of differences between mid-pupal and late-pupal samples in both (C) and (D) calculated by Student’s  $t$ -test. (E) Selected enriched transcription factor binding motifs within TAS variants in mid-pupal wings suggest selection on structural, olfactory, and neural phenotypes (see [supplementary table S1, Supplementary Material](#) online). For all panels: p, *Heliconius erato petiverana*; l, *H. e. lativitta*; h, *H. himera*.

population divergence. Surprisingly, our regression showed no apparent association between peak magnitude and genomic divergence for any of the three assays ([supplementary fig. S12C, Supplementary Material](#) online). The same analysis performed on VREs in negative Tajima’s  $D$  bins did show a slight, but significant, negative association with Tajima’s  $D$  score (average Pearson  $r$  was  $\sim 0.09$ ), suggesting that while peak magnitude is a significant variable in determining selective potential, this accounts for  $< 1\%$  of the variance in selection ([supplementary fig. S12D, Supplementary Material](#) online).

Although our data are not ideally suited to identify within-population variation, we did find that up to 3% of our TAS elements displayed significant within-population variation (false discovery rate [FDR] 10%, [supplementary fig. S13A, Supplementary Material](#) online) when intrapopulation signals at each peak are given a z-score. As we would expect, these regions showed significantly higher Tajima’s  $D$  distributions than the whole genome for all three populations (five of six comparisons

were significant, two-sample Kolmogorov–Smirnov test,  $P < 0.001$ ), indicating that balancing selection is maintaining diversity at these loci ([supplementary fig. S13B, Supplementary Material](#) online).

Considered with the analyses described above, our observations of VREs in *H. erato* show that regulatory variation in all three populations is strongly associated with genomic signatures of selection and that adaptation is likely driving divergence in a range of wing development processes at different developmental timepoints. Indeed, analysis of population-specific variable TASs using MEME-CHIP (Ma et al. 2014) found enrichment for binding site motifs for factors associated with development of scale, olfactory, and neural cells ([fig. 3E](#) and [supplementary table S1, Supplementary Material](#) online), confirming that many features of wing development are simultaneously under selection. Importantly, observing regulatory evolution at multiple developmental timepoints appears to highlight changes in selective pressures affecting different aspects of tissue development.

### Population Structure of Regulatory Evolution

We next wanted to determine whether population divergence predicts distribution of VREs between populations. To address this, we identified all VREs that were shared between two populations, but not a third. VRE assignment to each of the three populations was, in general, evenly distributed between the three populations, suggesting relatively equal gene regulatory divergence for all branches within the *H. erato* clade (fig. 4A). For all three regulatory assays, shared VREs appeared to track overall with phylogenetic divergence at both developmental stages assayed, where the parapatric populations shared a much higher number of VREs than seen in comparisons between allopatric populations (fig. 4B–D). Interestingly, the histone marks showed a particularly strong pattern of VRE segregation, where ~66% of variable histone marks were shared between the parapatric populations, while only ~1% were shared in allopatric comparisons. This important observation shows that chromatin accessibility and histone marks show very different patterns of evolution at the population level, where after ~1–2 My of divergence ancestral variation is widely maintained in chromatin accessibility, while differences in actual regulatory activity have become strongly subdivided geographically.

We then wanted to assess to what extent regulatory divergence might occur in the same genomic regions through independent regulatory changes in two or more populations. To test for this, we first observed the number of VREs in each population with one or more additional nonoverlapping VREs in a second population within 5 kb of the initial variant, which we refer to as clustered VREs. Clustered VREs were surprisingly common, and showed a significant trend toward increased clustering as phylogenetic distance decreased (fig. 4E–G), suggesting that geographic proximity and shared recent ancestry have driven adaptation at the same loci between closely related species pairs. Interestingly, a slight increase in VRE clustering was seen between mid-pupal and late-pupal developmental stages, though to a lesser degree than observed in analysis of  $F_{st}$  outliers, suggesting that parallel divergence is not constrained to regions of relaxed purifying selection.

To better understand how clustered variable regulatory sites were evolving between populations, we assigned clustered sites into two categories: 1) sites indicating presumptive epistatic coevolution, where multiple gains of VREs appear in close proximity in a pair of populations, representing concerted gain or loss of multiple VREs (e.g., fig. 4H, top), and 2) sites displaying independent, parallel regulatory evolution, where populations share independent VREs in the same region, suggesting parallel selection on the same regulatory apparatus (e.g., fig. 4H, bottom). We were surprised to find strongly opposing trends for these two categories across all types of VREs, where VREs shared between the most closely related populations were overwhelmingly examples of epistatic coevolution (fig. 4I–K), while parallel evolution prevailed as population divergence increased (fig. 4L–N). This further supports our proposal that after ~1–2 My of

evolution, ancestral regulatory variation is largely sorted out of populations and shared variation in regulatory activity is often due to parallelism or convergence.

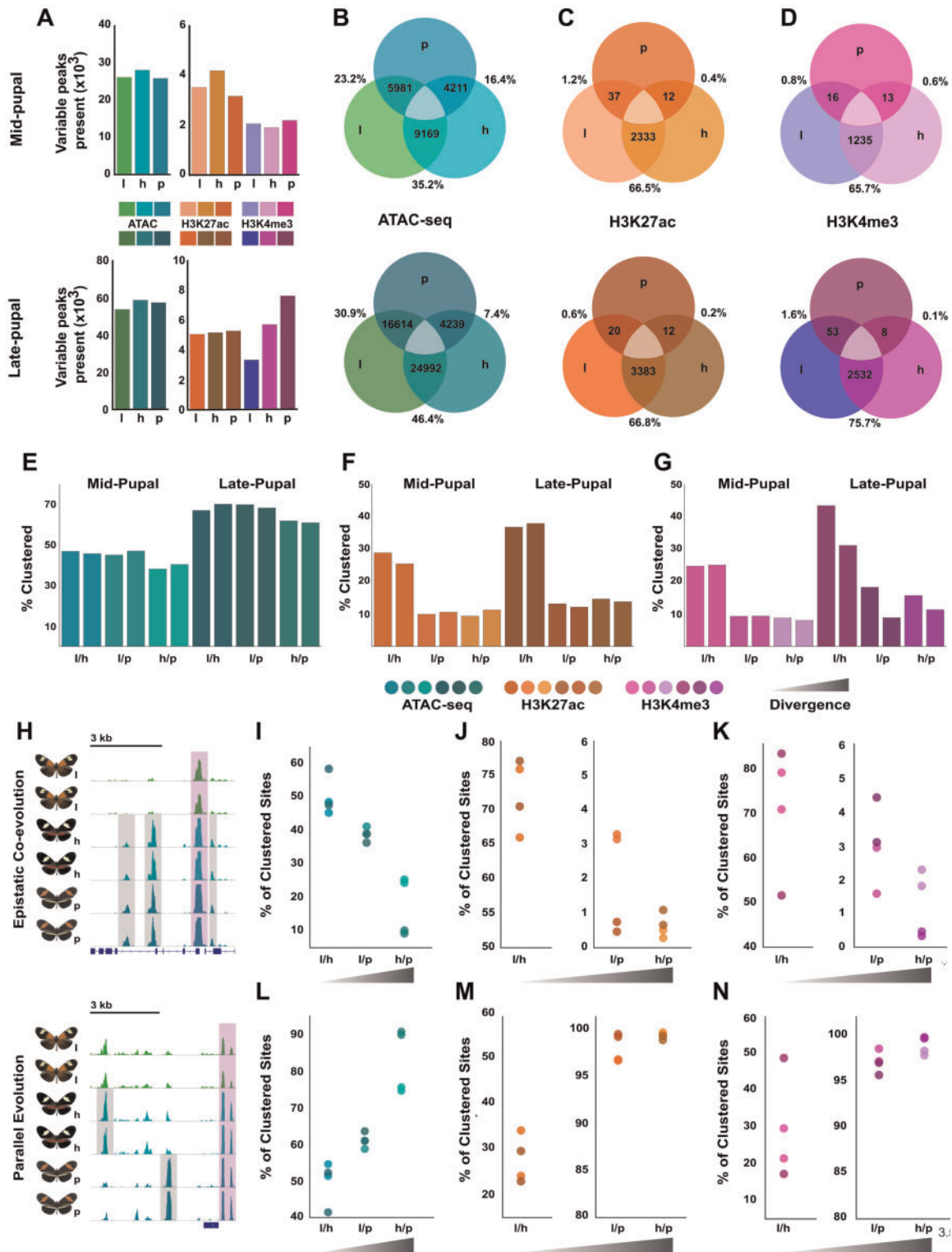
To determine whether introgression was potentially driving the shared VREs between *H. himera* and *H. e. lativitta*, we calculated absolute genomic divergence,  $D_{xy}$ , in 5-kb bins for the genome. We again set the bottom 5% of bins as representative of regions with a high degree of introgression between the two populations. To get a benchmark, we first analyzed all shared VREs between *H. himera* and *H. e. lativitta*, which were ~1.25- to 1.75-fold enriched in low  $D_{xy}$  loci than expected (supplementary fig. S14A, Supplementary Material online). VRE clusters between *H. himera* and *H. e. lativitta* showing evidence of epistatic coevolution were further enriched in low  $D_{xy}$  loci by an average of ~1.5-fold, suggesting that introgression has played an important role in maintaining clade-level regulatory divergence (supplementary fig. S14B, Supplementary Material online). As we would expect, a corresponding analysis of clustered VREs between divergent pairs showing parallel evolution was moderately enriched in regions of high  $F_{st}$  (supplementary fig. S14C, Supplementary Material online).

Holistically, we find that clustered VREs are a prevalent feature of regulatory evolution, and that regulatory variation tied to closely related populations often manifests as concerted covariation of arrays of regulatory elements, while in more distantly related populations, independent changes in regulatory regions are observed. We interpret this to indicate that local selection pressures favor the same specific regulatory changes over short periods of evolutionary time. This is followed by adjustment of the same transcriptional networks via independent regulatory mutations as populations (and potentially species) continue to diverge in allopatry.

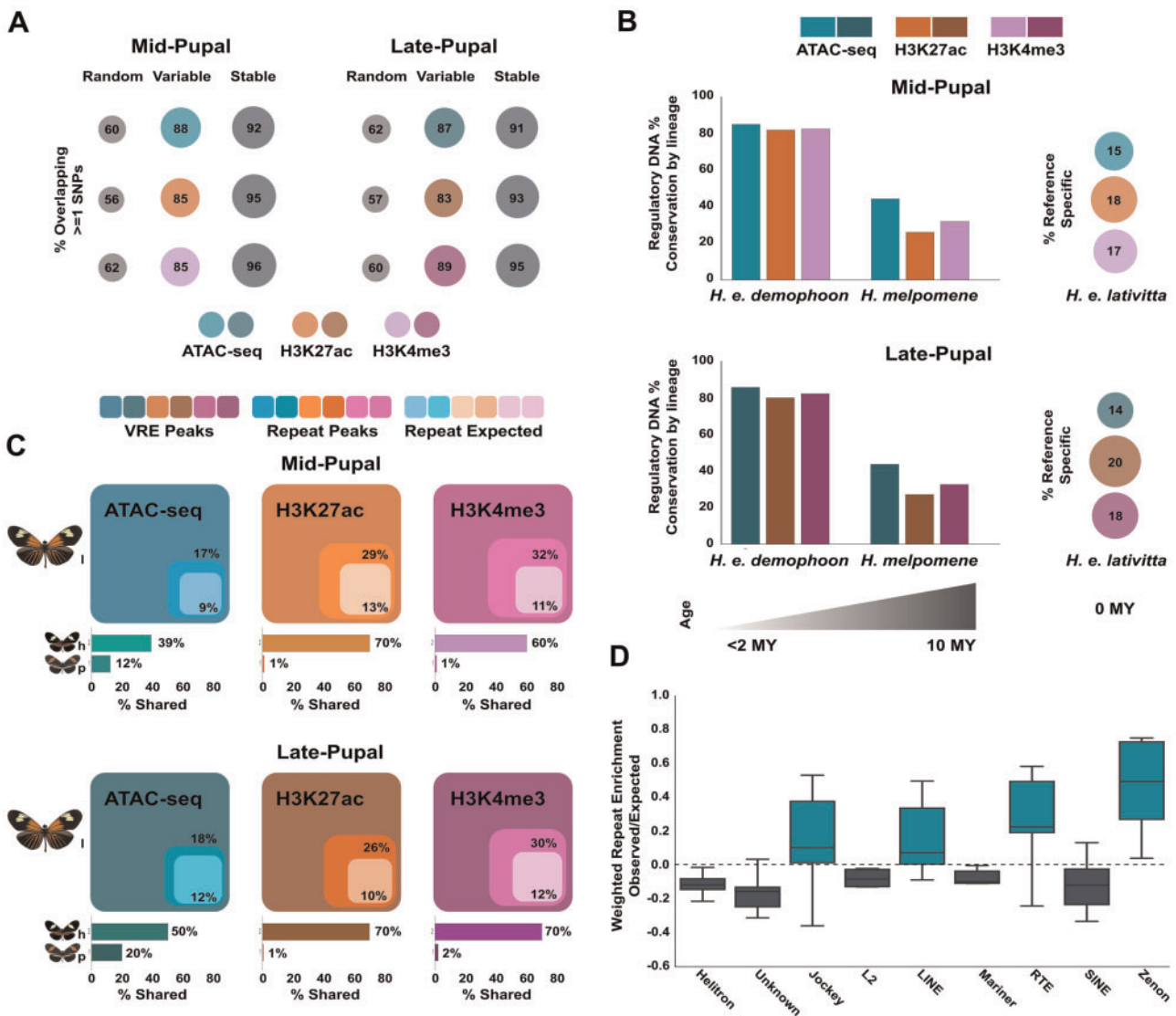
### Genomic Mechanisms of Regulatory Evolution in *Heliconius*

Previous studies have associated regulatory element turnover with several genomic features, such as repetitive sequences, nucleotide polymorphism, and DNA sequence age. We thus sought to investigate the mechanistic basis of regulatory evolution in *H. erato*, first seeking to understand the relationship between sequence polymorphism and regulatory element change. Approximately 85–90% of VREs overlapped one or more single nucleotide polymorphisms (SNPs) from our resequencing data (fig. 5A), potentially accounting for a majority of regulatory changes (but excluding the effect of large insertions and deletions). This was substantially greater than expected for sites of the same size distributed randomly throughout the genome, but surprisingly, slightly less than observed across all nonvariable (labeled “stable”) regulatory loci. Importantly, this demonstrates that while SNPs may be important drivers of regulatory evolution, most nucleotide polymorphism observed at regulatory loci is unlikely to be functionally relevant in a given stage or tissue type and suggests that VRE loci are younger than stable regulatory elements.

To ask whether novel or conserved DNA content is associated with regulatory variation, we identified conserved



**FIG. 4.** Population structure of VRE clusters. (A) VREs in each population and developmental stage, assigned by the population showing greatest signal in pairwise population comparisons for each assay. Patterns of shared VREs between populations for (B) ATAC-seq, and (C) H3K27ac and (D) H3K4me3 ChIP-seq loci show phylogenetic distance as a major determinant of shared VREs. Gray areas indicate that the union set does not exist. Percent values show percent of elements relative to the population with the fewest assigned VREs. Percent of VREs clustering between populations, defined as nonoverlapping VREs with one or more additional VREs in another population  $<5$  kb away, in mid-pupal and late-pupal developmental stages for (E) ATAC-seq, (F) H3K27ac, and (G) H3K4me3 ChIP-seq assays. Double bars are shown for each population pair, corresponding to the order on the pair label. Clustered VREs were categorized by evidence of either epistatic coevolution (H, top) or parallel evolution between populations (H, bottom). Gray outline shows distal clustered VREs and purple marks the putative promoter, which appears to track with changes in neighboring distal activity. The percent of clustered sites showing evidence of epistatic coevolution decreases greatly with increased phylogenetic divergence for ATAC-seq (I), and both ChIP-seq marks (J, K). The opposite trend was observed for parallel evolution in all three data types, which becomes prevalent as population divergence increases (L–N). For all panels: p, *Heliconius erato petiverana*; i, *H. e. lativitta*; h, *H. himera*.



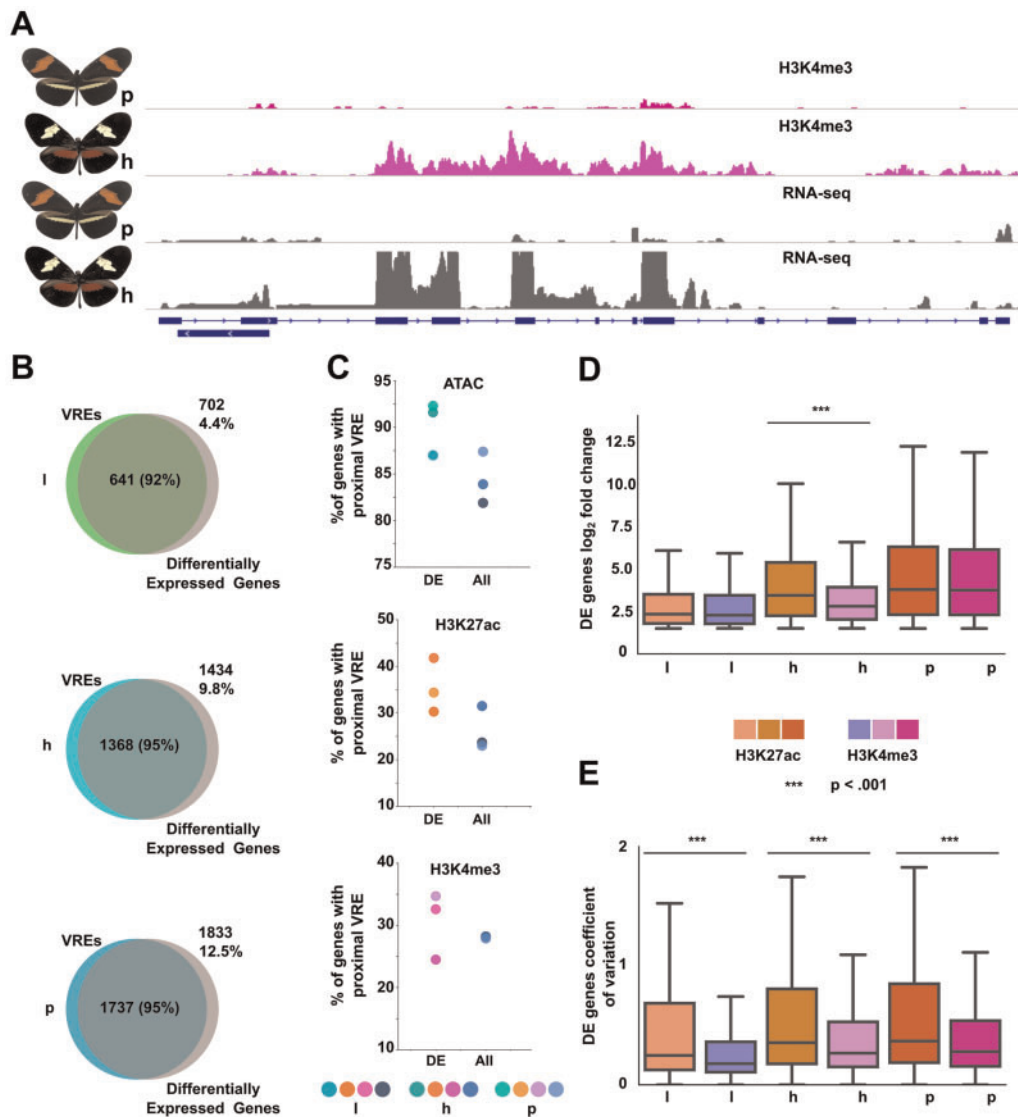
**Fig. 5.** Genomic features associated with regulatory variation in *Heliconius*. (A) Fraction of variable and nonvariable (“stable”) regulatory loci overlapping one or more SNPs. “Random” represents randomly selected genomic sequences overlapping one or more SNPs. Stable loci show a significant increase in SNPs relative to variable regulatory loci (chi-square test). (B) Conservation of VRE DNA within the species (*Heliconius erato demophoon*) and genus (*H. melpomene*) shows that most VREs are found in DNA conserved within species, but often not conserved between species. (C) Enrichment of VREs in repetitive sequences. VREs are significantly enriched in repetitive sequences, especially for variable histone marks. Bar graphs show the trend of conserved repeat associated VREs between recently diverged populations. (D) Four long interspersed nuclear element types were found to be enriched over expected frequency for VREs in the reference population. For all panels: p, *H. e. petiverana*; l, *H. e. lativitta*; h, *H. himera*.

regulatory sequences (Lewis et al. 2016) to determine the degree to which VREs were found in DNA conserved at the species level (<2 My, *H. e. demophoon* reference) and the genus level (<10 My, *H. melpomene* reference) (fig 5B). The great majority of regulatory changes, >80%, occurred in DNA conserved at the species level and showed no substantial differences associated with any of the regulatory assays, contrary to macroevolutionary studies of regulatory element change between vertebrate species (Villar et al. 2015). At the genus level, roughly 30–40% of variable regulatory DNA was conserved, with a slight decrease in conservation of variable H3K27ac marks, confirming the youth of variable regulatory DNA as indicated by our SNP analysis. This was a

substantially greater turnover of variable regulatory DNA compared with previous results showing a relatively high conservation of all regulatory loci within the genus, though variable loci remain more conserved than randomly selected sites (Lewis et al. 2016). Thus, our results support a model of intraspecific regulatory evolution driven mostly by changes in DNA that has evolved rapidly after the speciation process.

We next asked to what extent VREs were associated with repetitive sequences, which are known to evolve rapidly and likely make up a large fraction of the novel DNA at the species level. Similar to previous studies (Villar et al. 2015; Trizzino et al. 2017), we found that VREs are enriched for repetitive sequences across all three assays with no substantial effect of





**FIG. 6.** Regulatory variants associate with two types of gene expression changes. (A) Example of variable regulatory activity associated with differences in gene expression between *Heliconius erato petiverana* and *H. himera* on chromosome 15 (normalized read depth shown). (B) Venn diagrams showing the number of highly differentially expressed genes in mid-pupal wing tissue genes with one or more VREs within 25 kb of the TSS. Intersections show that 82–95% of highly differentially expressed genes have a proximal VRE. (C) The majority of VREs proximal to differentially expressed genes are TAS elements, followed by H3K4me3 and H3K27ac variants, and VREs show greater association with highly differentially expressed genes compared with all genes. H3K27ac and H3K4me3 are associated with distinct changes in gene expression, with (D) H3K27ac variants associated with greater expression fold change. (E) H3K4me3 VRE-associated highly differentially expressed genes show reduced expression-normalized variance (coefficient of variation) within populations. Significance calculated by two-sample Kolmogorov–Smirnov test. For all panels: p, *H. e. petiverana*; i, *H. e. lativitta*; h, *H. himera*.

developmental stage (fig. 5C). A large fraction of repeat-associated VREs were shared between the parapatric Ecuadorian populations, yet were mostly absent from the more diverged Panamanian population. Analysis of each major repeat class in the *Heliconius* repeat database (Lavoie et al. 2013) using observed/expected occurrences weighted by repeat class frequency found four major repeat groups significantly associated with variable regulatory DNA (fig. 5D). Interestingly, all four classes were long interspersed nuclear elements, which have been shown to have rapid turnover in *Heliconius* butterflies (Lavoie et al. 2013) and imply a potentially important functional role for these presumed deleterious sequences. Analysis of repetitive sequence overlap

variation between the three populations found no evidence of population-specific change in repeat element utilization (supplementary fig. S15, Supplementary Material online), leading us to believe that our results are representative of the species.

#### VREs Are Associated with Transcriptional Divergence

Finally, we investigated that the role regulatory divergence might play in driving transcriptional divergence between the three populations. Visual observation of our data indicated that strong transcriptional variation was frequently accompanied by regulatory divergence (e.g., fig. 6A). To investigate the role of *cis*-regulatory evolution in major

gene expression divergence between populations, we first classified our differentially expressed genes as “highly” and “weakly” differentially expressed based on read depth and the degree of divergence, with a minimum  $\log_2$  fold change of 1.5 for highly differentially expressed genes (supplementary fig. S16A, Supplementary Material online). As expected the majority of genes were weakly or not differentially expressed, which tracks similarly to our observation of variation in regulatory elements. We then identified all VREs at the mid-pupal stage between population pairs within 25 kb of a highly differentially expressed gene TSS, termed TSS-proximate VREs. Although *cis*-acting loci often fall outside of this window, previous literature has shown that regulatory loci are frequently within 25 kb (Ghavi-Helm et al. 2014) and assigning more distal VREs is challenging without further data. This was supported by our own analysis, which showed a higher frequency of VREs within 25 kb of differentially expressed TSSs, and no association was found between the distance of a VRE and the change in gene expression out to 100 kb (supplementary fig. S16B, Supplementary Material online). One or more TSS-proximate VREs were present for 92–95% of highly differentially expressed genes in all population comparisons, suggesting that assayed VREs contribute to the majority of transcriptional divergence between populations (fig. 6B). We then tested the relative abundance of each VRE type within 25 kb of highly differentially expressed genes. Chromatin accessibility proved the most frequent indicator of transcriptional divergence, with ~92–95% of highly differentially expressed genes being associated with one or more TAS variants (fig. 6C). Variable H3K27ac and H3K4me3 marks were proximal to roughly 25–42% of highly differentially expressed genes, with variable H3K27ac marks displaying a slight enrichment over H3K27ac marks around highly differentially expressed genes. For all three regulatory assays, the distribution of VREs around highly differentially expressed genes was higher than observed for all protein-coding genes.

We were interested in the role of the assayed histone modifications in gene expression evolution, so we investigated that the potential impact VREs for each histone mark had on gene expression. We first tested whether the effect of VREs of each type might have on absolute change in gene expression. We found an increase in  $\log_2$  fold change in highly differentially expressed genes with one or more associated H3K27ac variants compared with genes with associated H3K4me3 variants, with a significant difference between the two sets of VREs in *H. himera* ( $P < 0.001$ , two-sample Kolmogorov–Smirnov test), indicating that H3K27ac is stronger predictor of overall gene expression divergence (fig. 6D). Interestingly, this comparison became highly significant for all three populations when all differentially expressed genes were considered (supplementary fig. S17A, Supplementary Material online), leading us to suspect that H3K27ac VREs were both the stronger predictor of gene expression change and were more likely to associate with highly differentially expressed genes. This was confirmed when we compared the frequency of proximal histone VREs for both highly and all differentially expressed genes, with H3K27ac showing a notably higher association with highly differentially expressed genes while

H3K4me3 shows a correspondingly weaker association (supplementary fig. S17B, Supplementary Material online).

We next tested whether either histone mark was more associated with a change in variance for VREs associated with highly differentially expressed genes. In this case, H3K4me3 showed a significantly lower coefficient of variation for associated highly differentially expressed genes ( $P < 0.001$ , two-sample Kolmogorov–Smirnov test, fig. 6E), leading us to speculate that change in H3K4me3 signal may play a role in the evolution of gene expression variance (i.e., fine tuning), while playing a lesser role in expression magnitude. This hypothesis is supported by previous literature demonstrating that H3K27ac signal correlates best with gene expression within cell lines, while H3K4me3 plays a direct role in fine tuning the degree to which RNA-polymerase II is recruited to TSSs (Karlič et al. 2010; Lauberth et al. 2013). Taking into account our prior observations of selection on the two histone variants (fig. 3C and supplementary fig. S12B, Supplementary Material online), evidence indicates that fine tuning of gene expression divergence between populations, associated with H3K4me3 variation, is often under greater selective pressure than is absolute divergence.

## Conclusion

Understanding the origins of diversification via adaptation to local selective pressures is one of the primary goals of evolutionary population biology. Recent population genetic work in *H. erato* has largely focused on three major mimicry-related color pattern loci and has portrayed these regions as “hotspots” of genomic adaptation set in a mostly free-flowing genomic landscape (Counterterman et al. 2010; Supple et al. 2015). Interestingly, our results suggest that gene expression, chromatin accessibility, and active regulatory elements are all under genome-wide divergent selection between regional morphs of the *H. erato* clade—even between hybridizing populations. We found evidence of functional regulatory variation associated with both population structure and developmental stage, indicating a highly complex adaptive landscape, with stronger selection during earlier-stage developmental patterning, but greater variability in the late-stage developmental period associated with the structural maturation of scale cells. This result has important implications for how inferences are drawn from studies of regulatory evolution, with evidence pointing to the developmental characteristics of model tissues and organisms used in regulatory evolution as a driving force behind many observed patterns. We also found that regulatory assay choice can affect observations of regulatory evolution, with our data indicating the strongest selection—and subsequently the greatest adaptive impact—on histone modifications indicative of regulatory activity. Moreover, we find evidence that transcriptional divergence associated with regulatory evolution is induced via distinct mechanisms altering both magnitude and variance of transcription. Our findings highlight the need for further comparative functional genomic study at the population level to refine our understanding of how selection and adaptation intertwine with complex regulatory architectures to

determine species-level genomic landscapes. Finally, our results suggest that the *H. erato* genome is likely under much greater selection throughout than would be predicted solely by hybrid zone-focused studies of wing color patterns, and suggest that we should reassess “island”-centric models of adaptation and population divergence (Storz 2005; Cruickshank and Hahn 2014; Wolf and Ellegren 2017).

## Materials and Methods

### *Heliconius* Stocks and Tissue Sampling

All samples of *H. e. lativitta*, *H. e. petiverana*, and *H. himera* were taken from laboratory colonies at Cornell University derived from individuals collected from Ecuador (*H. e. lativitta* and *H. himera*) and Costa Rica (*H. e. petiverana*). mid-pupal samples were collected from individuals reared for 72 h (3 days) at  $\sim 30^{\circ}\text{C}$ , and were phenotyped for the emergence of early wing scale buds. Late-pupal samples were collected as “ommochrome” stage pupae,  $\sim 7$  days postpupation at  $30^{\circ}\text{C}$ , and were phenotyped for ommochrome pigment deposition in the wing and eyes without any signs of melanin pigmentation. To ensure similar staging between samples, both time and morphology were used to determine stage.

### ChIP-seq, Input Control, and ATAC-seq Sample Preparation

ChIP-seq was performed as previously described (Lewis et al. 2016), with minor modifications. We performed two biological replicates of each assay, including ChIP-seq input controls, for a combined set of 24 ATAC-seq, and 72 ChIP-seq and control experiments. For each population, developmental stage, tissue, and biological replicate, wing pairs (left and right) from four to six individuals were fixed for 5 min with 1% freshly prepared formaldehyde, quenched for 5 min with 1 M glycine solution to a final concentration of 0.125 M, rinsed with two washes of cold phosphate-buffered saline, then combined prior to tissue homogenization. Postextraction nuclear samples were incubated for  $\sim 12$  and 13 min with 0.5  $\mu\text{l}$  micrococcal nuclease at  $37^{\circ}\text{C}$  before adding ethylenediaminetetraacetic acid to quench digestion. Digested nuclear preps were then split into three aliquots for immunoprecipitation and input control prep. Chromatin immunoprecipitation was performed with antibodies to H3K27ac and H3K4me3 (Abcam: ab4729 and ab8580), using 3–5  $\mu\text{g}$  of digested chromatin per immunoprecipitation. Libraries were prepared with the NEB DNA Ultra library prep kit using  $\sim 40$  ng of input, and amplified for 14 cycles prior to agarose gel size selection.

ATAC-seq was performed as described by Buenrostro et al. (2013), with minor modifications. Tissue dissection was performed as previously described (Lewis et al. 2016). For all samples, nuclear extractions were performed on freshly dissected ( $< 15$  min) wing pairs from a single individual, which were dounce homogenized in sucrose buffer using pestle B (mid-pupal samples) or pestles A and B (late-pupal samples). Nuclei were counted using a hemocytometer, and  $\sim 400,000$  nuclei were isolated for each ATAC-seq library prep. Libraries

were amplified for ten cycles, and size selected on an agarose gel for fragments between 35 and 1,000 bp.

Sequencing for ChIP-seq, input control, and ATAC-seq libraries was done on a NextSeq 500 at Cornell University using  $2 \times 37$  bp paired-end (PE) reads. ChIP-seq, input control, and ATAC-seq libraries were sequenced to a minimum depth of 20, 30, and 50M PE reads, respectively (supplementary table S2, Supplementary Material online).

### Read Alignment and Peak Calling

Read alignment and filtering for ChIP-seq, input control, and ATAC-seq samples were performed as previously described (Lewis et al. 2016). Briefly, raw sequence reads were aligned using Bowtie2 (Langmead and Salzberg 2012) against the *H. e. lativitta* v1.0 assembly and filtered for uniquely mapping pairs and quality threshold of “20” with a custom python script. Duplicate read alignments were removed with picardtools 2.1.1 “MarkDuplicates.” ChIP-seq peaks were called using MACS2 (Zhang et al. 2008) for each biological replicate using combined input from both input control replicates to avoid variation in enrichment profiles between replicates due to minor differences in MNase digestion and library size selection. ATAC-seq peaks were called using F-seq (Boyle, Guinney, et al. 2008). All peaks called by the peak calling software were included in our analysis.

### Quality Control Analysis of ChIP-seq, ATAC-seq, and RNA-seq Assays

Visual observation suggested a high degree of similarity between ATAC-seq and ChIP-seq data sets. Median FRiP scores were 24.6% for H3K27ac ChIP-seq, 38.2% for H3K4me3 ChIP-seq, and 81.7% for ATAC-seq samples. Peak call lengths for both ChIP-seq data sets began at one nucleosome, with a median length of two nucleosomes. Median fold enrichment at ChIP-seq peaks was distributed as expected, with  $\sim 3.5$ -fold change over input, extending to over 20-fold enrichment in larger peaks. We observed a high degree of similarity between forewing and hindwing samples of the same data type, population, and developmental stage. Average Pearson correlation of signal intensity at annotated regulatory loci between forewing and hindwing samples for each data type and developmental stage were 0.90, 0.83, and 0.91 for ATAC-seq, and H3K27ac and H3K4me3 ChIP-seq assays, respectively. To further verify similarity between wing tissues, we called differences between wing tissues using DESeq2 (Love et al. 2014) in our Day 3 ATAC-seq data, which showed the greatest signal-to-noise ratio. In all three population comparisons only two genomic loci showed significant differences, *abd-a* and *ubx* (supplementary fig. S2, Supplementary Material online), both of which are HOX genes known to determine hindwing identity (Wagner 2007). Although gene and noncoding RNA annotation remains a challenge for nonmodel genomes, the median fraction of RNA-seq reads aligning to gene bodies or untranslated regions (UTRs) was 67.7%. As UTRs were not well annotated in the original annotation file, for this analysis, UTRs were defined as 500 bp upstream and downstream of the first and last annotated exons. Manual observation of RNA alignments found numerous unannotated

regions of conserved alignment, suggesting that some portion of the remaining RNA-seq reads align to unannotated exons and noncoding RNAs. All RNA-seq samples fell within the range of accepted samples from the Epigenomics Roadmap for percent of reads aligning to exons and introns (Roadmap Epigenomics Consortium 2015).

### Multiple Testing Correction

All false discovery rates were set at  $FDR = 0.1$  unless otherwise stated.

### Analysis of Population Variation in ChIP-seq and ATAC-seq Peaks

Population-specific data sets were generated as follows: Peak calls from biological replicates for each tissue, developmental stage sample, and data type were combined and merged using bedtools (Quinlan and Hall 2010) with duplicate peaks removed if they overlapped by 147 bp, or 1 nucleosome, (ChIP-seq) or 50 bp (ATAC-seq). Population-level peak sets for each data type tissue, and developmental stage, were merged as described above to produce species-wide tissue and developmental stage-specific peak sets for subsequent comparison between populations and to determine total unique peak sets for each assay and developmental stage. Counts of reads falling into merged peak files for each assay type and developmental stage were determined using bedtools “genomecov” function for each sample. Wing tissue counts were then grouped by population, assay type, and developmental stage, after which DESeq2 was used to normalize read counts and calculate differentially enriched regulatory regions in pairwise comparisons for all three assays and both developmental stages. These differentially enriched loci were considered “variable.” The total fraction of each assay considered “variable” was determined by merging variable peaks from each population and determining what fraction this number was out of all peaks from a developmental stage tested.

Analysis of population-level variable peak sets was performed using bedtools and linux command line utilities. The bedtools “closest” function was used for all proximity analyses, and bedtools “intersect” was used for analysis of overlapping regulatory variants. Linux command line utility “shuf” was used in combination with other functions to determine expected distributions for all random distribution analyses. Linux command line utilities “awk” and “grep” were used in combination with bedtools and shell scripts for data parsing, to assign regulatory variants to specific populations, and to automate most analyses. Python scripts were used for some data parsing and statistical analyses. The Scipy python library (Jones et al. 2001) was used for Fisher’s exact tests, two-sample Kolmogorov–Smirnov tests, and *t*-tests.

### RNA-seq Sample Preparation

RNA-seq was performed on mid-pupal forewings and hindwings as previously described (Lewis et al. 2016), with three biological replicates for each tissue from each population (for a total of 18 samples). Samples were sequenced on a

NextSeq 500 to a minimum depth of 18M PE reads (supplementary table S2, Supplementary Material online).

### Differential Gene Expression Analysis

RNA-seq data were aligned to the reference assembly using Tophat2 (Kim et al. 2013). Read counts for each annotated gene were determined using bedtools “multicov” function, and differentially expressed genes were identified using DESeq2 (Love et al. 2014) as prescribed by the authors, with an adjusted *P*-value cutoff of “0.01.”

### Whole-Genome Resequencing Sample Preparation

DNA was extracted from 4 *H. e. lativitta*, 6 *H. himera*, and 5 *H. e. petiverana* samples with a DNeasy kit following the manufacturer’s guidelines. Sequencing libraries were prepared using the Nextera DNA Library Prep kit following manufacturer’s guidelines and sequence on a NextSeq 500 at  $2 \times 37$  bp PE reads. Each sample was sequenced to a minimum of  $2 \times$  coverage (supplementary table S2, Supplementary Material online).

### SNP Calling, $F_{st}$ , and Tajima’s *D* Analysis

Whole-genome sequencing samples were aligned to the reference assembly using Bowtie2 (Langmead and Salzberg 2012). Aligned read files were sorted using samtools, followed by duplicate read marking and read group addition using picardtools 2.1.1 “MarkDuplicates” and “AddOrReplaceReadGroups” functions. Raw variant VCF files were produced for each sample using GATK (DePristo et al. 2011) “HaplotypeCaller.” Joint genotyping was performed using GATK “GenotypeGVCFs” with “-stand-emit-conf” set to “30.” To remove variants with hypercoverage, low coverage, low quality, and strand-biased variant calls, the joint genotype variant file was filtered using GATK “VariantFiltration” with the following setting: “-filterExpression “DP < 5 || DP > 500 || QD < 2.0 || FS > 60 || MQ < 20.0 || MQRankSum < -12.5 || ReadPosRankSum < -8.0.” This process removed ~250,000 variants.

$F_{st}$  analysis was performed for pairwise population comparisons using 5-kb bins across the genome with VCFtools (Danecek et al. 2011) “-weir-fst-pop” function with “-fst-window-size” and “-fst-window-step” set to “5000.” All subsequent  $F_{st}$  analysis was performed using the unweighted “Mean Fst” column.

Identification of top 5% outlier bins and analysis of regulatory variants in outlier bins were performed using bedtools “intersect” and linux command line utilities. High  $F_{st}$  outliers were labeled “active” for a given assay and developmental stage if they intersected a regulatory element, whether variable or not. To control for changes in regulatory element numbers between assays, stages, and populations, relative enrichment of high  $F_{st}$ -associated VREs was calculated as the fraction of VREs intersecting active high  $F_{st}$  loci divided by the fraction of the active genome comprised by these bins for each assay, population, and developmental stage. Tajima’s *D* analysis was performed using the VCFtools “-TajimaD” function with a bin size of “2,000” bp. The bottom 2.5% of Tajima’s *D* bins by score were set as outliers similar to the  $F_{st}$

analysis. Intersection of high  $F_{st}$  VREs with low Tajima's  $D$  bins and overall enrichment of VREs in negative-valued Tajima's  $D$  bins were performed using bedtools "intersect." A bin size of 2,000 bp for Tajima's  $D$  was chosen due to the expectation of linkage decay at around 1 kb. To verify that our results were robust to bin size, we confirmed our increase over expected of high  $F_{st}$  loci in low Tajima's  $D$  bins using 5-kb bins, despite greatly exceeding the expected linkage distance and potentially introducing additional recombination sites (supplementary fig. S18, Supplementary Material online).

### Transcription Factor Binding Site Motif Analysis

A custom Python script was used to extract 300 bp around the center of each variable ATAC-seq peak for annotated regulatory loci sets. To determine motif enrichment for each class, MEME-CHIP (Ma et al. 2014) was run in allowing multiple motifs per peak and searching for up to 15 motifs.

### Analysis of VRE Cluster Evolution

VREs were categorized as clustered between populations if a variable regulatory site was present in a second population (query population) within 5 kb of a variant in the original, subject population. Element distance was determined using bedtools "closest," excluding overlapping elements. Bedtools "intersect" was used to distinguish between clustered loci categorized as "epistatic coevolution" clusters and "parallel evolution" clusters, with intersecting clustered VREs being given the label "epistatic coevolution." Reciprocal analysis of clustered variable elements was performed to account for slight differences in results depending on the chosen "subject" population in pairwise comparisons.

### Analysis of Within-Population Variation

To detect loci displaying within-population variation, we devised a statistical test based on deriving z-scores that could be FDR corrected to provide a measure of statistically significant sites that vary within populations. Normalized read counts of ATAC-seq samples from DESeq2 for each population and developmental stage were used to calculate the dispersion index (variance of normalized read counts divided by the mean normalized read counts) for each peak within populations. The mean of all dispersion indexes was next subtracted from the dispersion index for each peak. We then divided this value by the standard deviation of the dispersion index for all peaks to get a z-score for each peak locus. We converted the resulting z-scores to  $P$ -values and applied a 10% FDR correction to our data to identify within-population variation for each population and developmental stage.

### $D_{xy}$ and Introgression Analysis

$D_{xy}$  was calculated for 5-kb bins across the genome using a previously developed pipeline (Martin et al. 2016) for each population pair. The five percent of genomic bins with the lowest  $D_{xy}$  scores were then taken as "low  $D_{xy}$ " bins for analysis of introgressing loci. Intersection of VREs with low  $D_{xy}$  bins was performed using the bedtools "intersect" function.

### SNP Enrichment Analysis and Regulatory DNA Evolution

SNP enrichment was determined using bedtools "intersect" and the filtered SNP set used to determine  $F_{st}$  values. Random loci were selected using bedtools "shuffle" on the peak sets for each regulatory data type and developmental stage. Division of VREs by population had no observed effect on SNP enrichment frequency, so the complete set of VREs for each data type and developmental stage were used for variable element enrichment analysis. Regulatory loci not found significantly different between any population pair were deemed "stable" for this analysis.

Regulatory DNA conservation was determined using a reciprocal best-hit BLAST algorithm previously developed for identifying conserved regulatory regions in Lepidoptera. Species-level conservation was determined via alignment to the *H. e. demophoon* (an *erato* race from panama) reference assembly (Van Belleghem et al. 2017), and the most recent *H. melpomene* reference assembly (version 2) was used for determining genus-level conservation. Regulatory DNA not aligning to the *H. e. demophoon* reference was deemed "reference specific," which accounts for DNA novel to *H. e. lativitta*, slight differences in the reference assemblies, and regions without a one-to-one reciprocal relationship.

### Repetitive Element Enrichment in Variable Regulatory Loci

Repetitive DNA elements were annotated using RepeatMasker and the *Heliconius* repeat database (Lavoie et al. 2013). To determine repeat element enrichment in VREs and avoid spurious overlap with repeats, we followed the approach described by Lewis et al. (2016) using bedtools "intersect" to identify regulatory elements overlapping repetitive sequences by 50% or more. Enrichment was determined using VREs for the reference assembly population for each regulatory data type and developmental stage to avoid potential bias due to genome evolution. Expected repeat enrichment was determined by using the entire peak set corresponding to each variable peak set. Analysis of shared repeat-associated VREs was performed with bedtools "intersect" and the VRE sets for each nonreference population. To estimate the importance of specific repeat groups in regulatory evolution, we used a weighted observed over expected model. For each repeat group, we took the percent of observed VREs overlapping repeats of that group divided by the percent observed in the entire peak set. To determine the relative significance of each repeat group on regulatory evolution, the base score was then weighted slightly by dividing by  $(1 - \text{percent of all repeats composed by the query group})$ , which had a minor effect on repeat enrichment. Comparison of repeat utilization divergence between populations was performed by subtracting the percent utilization for each repeat group in the nonreference population from the percent utilization in the reference.

### Differentially Expressed Gene Set and Regulatory Variant Analysis

Highly differentially expressed genes for all three populations were determined by thresholding on the total set of

differentially expressed genes, taking only the subset of genes with a log<sub>2</sub> fold change of 1.5 or greater and mean normalized read counts of 10 or greater. The differentially expressed gene set was identified as described above. Regression analysis on highly differentially expressed genes and VREs was performed using the bedtools “closest” function to extract loci and the Python SciPy library to perform regressions.

The set of highly differentially expressed genes with VREs within 25 kb of the TSS were identified using bedtools “closest” and linux command line tools. Command line utilities were used to determine regulatory variant types associated with differentially expressed genes using regulatory variant sets corresponding to the same population comparisons used for differential gene analysis. Differentially expressed gene sets for each population comparison with associated H3K27ac and H3K4me3 variants were separated into groups corresponding to the associated histone variant, allowing for overlap between groups when TSS-proximal VREs existed for both histone marks. Analysis of variance was performed using the coefficient of variation of each mark to control for the effect of changes in expression level on sample variance. Command line tools and Python scripts were used to assign genes to the three regulatory variant assays and to parse log<sub>2</sub> fold change and standard error values for subsequent analysis of histone modifications.

### Analysis of Reproducibility between Reference Genomes

To ensure that our results were not a product of reference genome choice, all data sets were realigned to an alternate *H. erato* reference assembly derived from *H. e. demophoon* (v1), a race closely related to *H. e. petiverana*, and select analyses were replicated (supplementary fig. S19, Supplementary Material online). Overall regulatory variability was highly similar to that observed using the *H. e. lativitta* reference, and in a direct comparison, 75% of alignable VREs from the *H. e. lativitta* assembly were also called as VREs in *H. e. demophoon*. For this analysis, a VRE from *H. e. lativitta* was considered reproducible if it had a BLAST hit against the *H. e. demophoon* VRE DNA sequence set with an *e*-value of “e-10” or less. Reanalysis of genomic divergence showed that a high percentage of active high  $F_{st}$  bins were also captured by our VREs and these remained highly enriched for low Tajima’s *D* values in *H. e. demophoon*. The only noticeable effect was a slight loss of clade-specific results associated with *H. e. lativitta* and *H. himera*, such as the increased similarity between parapatric and allopatric comparisons in our Tajima’s *D* analysis, which was expected.

### Data Availability

ChIP-seq, ATAC-seq, RNA-seq, and whole-genome resequencing data are available for download at GEO: GSE105080, GSE109889, GSE111022, and SRA: PRJNA499127, and for interactive searching and browsing at butterflygenome.org

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

We thank Andy Clark, Charles Danko, Amy McCune, Philipp Messer, and members of the Reed Lab for their helpful comments on the project and manuscript. We also thank three anonymous reviewers for their helpful suggestions. This work was supported by National Science Foundation (Grant Nos. DEB-1354318 and DEB-1546049). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## References

- Albert FW, Kruglyak L. 2015. The role of regulatory variation in complex traits and disease. *Nat Rev Genet.* 16(4): 197–212.
- Arnold CD, Gerlach D, Spies D, Matts JA, Sytnikova YA, Pagani M, Lau NC, Stark A. 2014. Quantitative genome-wide enhancer activity maps for five *Drosophila* species show functional enhancer conservation and turnover during cis-regulatory evolution. *Nat Genet.* 46(7): 685–692.
- Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* 132(2): 311–322.
- Boyle AP, Guinney J, Crawford GE, Furey TS. 2008. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* 24(21): 2537–2538.
- Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.* 10(12): 1213–1218.
- Counterman BA, Araujo-Perez F, Hines HM, Baxter SW, Morrison CM, Lindstrom DP, Papa R, Ferguson L, Joron M, French-Constant RH, et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in *Heliconius erato*. *PLoS Genet.* 6(2): e1000796.
- Cruickshank TE, Hahn MW. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol.* 23(13): 3133–3157.
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics* 27(15): 2156–2158.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet.* 43(5): 491–498.
- Flanagan NS, Tobler A, Davison A, Pybus OG, Kapan DD, Planas S, Linares M, Heckel D, McMillan WO. 2004. Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. *Proc Natl Acad Sci U S A.* 101(26): 9704–9709.
- Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W, Furlong EEM. 2014. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512(7512): 96–100.
- Henriques T, Scruggs BS, Inouye MO, Muse GW, Williams LH, Burkholder AB, Lavender CA, Fargo DC, Adelman K. 2018. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev.* 32(1):26–41.
- Hoyal Cuthill J, Charleston M. 2012. Phylogenetic codivergence supports coevolution of mimetic *Heliconius* butterflies. *PLoS One* 7(5): e36464.
- Jiggins CD, Macmillan WO, Neukirchen W, Mallet J. 1996. What can hybrid zones tell us about speciation? The case of *Heliconius erato* and *H. himera* (Lepidoptera: Nymphalidae). *Biol J Linn Soc.* 59:221–242.

- Jones E, Oliphant E, Peterson P, et al. SciPy: Open Source Scientific Tools for Python, 2001, <http://www.scipy.org/>, last accessed November 27, 2018.
- Joron M, Jiggins CD, Papanicolaou A, McMillan WO. 2006. *Heliconius* wing patterns: an evo-devo model for understanding phenotypic diversity. *Heredity* 97(3): 157–167.
- Karlič R, Chung H-R, Lasserre J, Vlahoviček K, Vingron M. 2010. Histone modification levels are predictive for gene expression. *Proc Natl Acad Sci U S A*. 107:2926–2931.
- Kasowski M, Kyriazopoulou-Panagiotopoulou S, Grubert F, Zaugg JB, Kundaje A, Liu Y, Boyle AP, Zhang QC, Zakharia F, Spacek DV, et al. 2013. Extensive variation in chromatin states across humans. *Science* 342(6159): 750–752.
- Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, Sabo PJ, Larschan E, Gorchakov AA, Gu T, et al. 2011. Comprehensive analysis of the chromatin landscape in *Drosophila*. *Nature* 471(7339): 480–485.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14(4): R36.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9(4): 357–359.
- Laubert SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 152(5): 1021–1036.
- Lavoie CA, Platt RN, Novick PA, Counterman BA, Ray DA. 2013. Transposable element evolution in *Heliconius* suggests genome diversity within Lepidoptera. *Mob DNA*. 4:1–10.
- Lewis JJ, van der Burg Karin RL, Mazo-Vargas A, Reed RD. 2016. ChIP-Seq-annotated *Heliconius erato* genome highlights patterns of cis-regulatory evolution in Lepidoptera. *Cell Rep.* 16(11): 2855–2863.
- Li XY, Thomas S, Sabo PJ, Eisen MB, Stamatoyannopoulos JA, Biggin MD. 2011. The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome Biol.* 12(4): R34.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15(12): 550.
- Ma W, Noble WS, Bailey TL. 2014. Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc.* 9(6): 1428–1450.
- Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, Blaxter M, Manica A, Mallet J, Jiggins CD. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* 23(11): 1817–1828.
- Martin SH, Möst M, Palmer WJ, Salazar C, McMillan WO, Jiggins FM, Jiggins CD. 2016. Natural selection and genetic diversity in the butterfly *Heliconius melpomene*. *Genetics* 203(1): 525–541.
- Nord AS, Blow MJ, Attanasio C, Akiyama JA, Holt A, Hosseini R, Phouanenavong S, Plajzer-Frick I, Shoukry M, Afzal V, et al. 2013. Rapid and pervasive changes in genome-wide enhancer usage during mammalian development. *Cell* 155(7): 1521–1531.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6): 841–842.
- Roadmap Epigenomics Consortium, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, Kheradpour P, Zhang Z, Wang J, et al. 2015. Integrative analysis of 111 reference human epigenomes. *Nature* 518: 317.
- Smagulova F, Gregoretti IV, Brick K, Khil P, Camerini-Otero RD, Petukhova GV. 2011. Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* 472(7343): 375–378.
- Storz JF. 2005. INVITED REVIEW: Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol.* 14(3): 671–688.
- Supple MA, Papa R, Hines HM, McMillan WO, Counterman BA. 2015. Divergence with gene flow across a speciation continuum of *Heliconius* butterflies. *BMC Evol Biol.* 15:204.
- Trizzino M, Park Y, Holsbach-Beltrame M, Aracena K, Mika K, Caliskan M, Perry GH, Lynch VJ, Brown CD. 2017. Transposable elements are the primary source of novelty in primate gene regulation. *Genome Res.* 27(10): 1623–1633.
- Van Belleghem SM, Rastas P, Papanicolaou A, Martin SH, Arias CF, Supple MA, Hanly JJ, Mallet J, Lewis JJ, Hines HM, et al. 2017. Complex modular architecture around a simple toolkit of wing pattern genes. *Nat Ecol Evol.* 1(3):52.
- Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M, Park TJ, Deaville R, Erichsen JT, Jasinska AJ, et al. 2015. Enhancer evolution across 20 mammalian species. *Cell* 160(3): 554–566.
- Wagner GP. 2007. The developmental genetics of homology. *Nat Rev Genet.* 8(6): 473–479.
- Wittkopp PJ, Kalay G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet.* 13(1): 59–69.
- Wolf JBW, Ellegren H. 2017. Making sense of genomic islands of differentiation in light of speciation. *Nat Rev Genet.* 18(2): 87.
- Wray GA. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet.* 8(3): 206–216.
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nussbaum C, Myers RM, Brown M, Li W, et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:1–9.