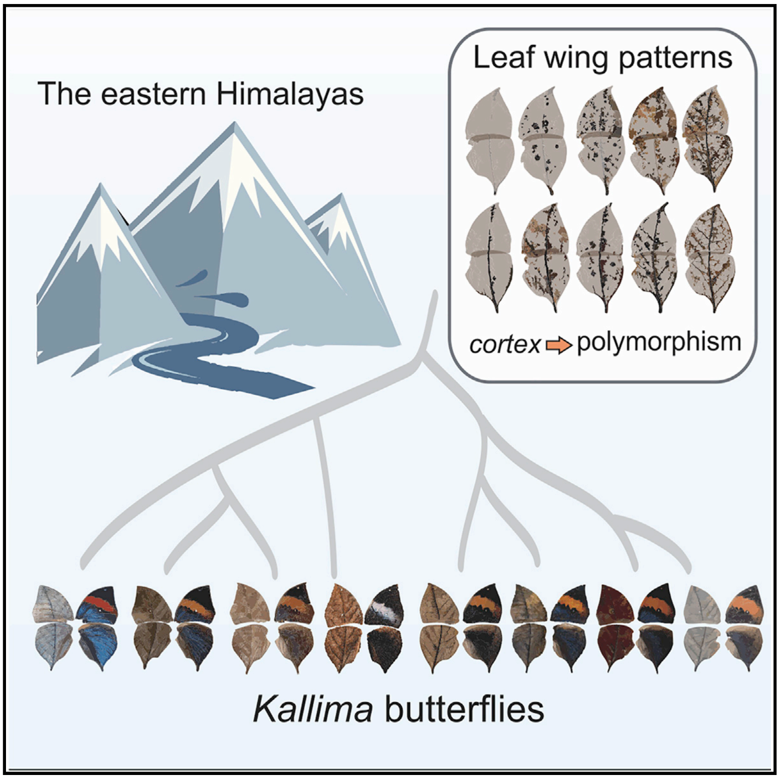
Cell

The evolution and diversification of oakleaf butterflies

Graphical abstract



Highlights

- The eastern Himalayas are a center of diversification in *Kallima* evolution
- A single gene, *cortex*, is involved in controlling leaf wing polymorphism
- Developmental constraints can be reconciled with evolvability
- Complex evolutionary innovation is driven by geographic changes and natural selection

Authors

Shuting Wang, Dequn Teng, Xueyan Li, ..., Li Zhang, Wen Wang, Wei Zhang

Correspondence

weizhangvv@pku.edu.cn

In brief

Macro- and micro-evolutionary analyses reveal that biogeography drove the diversification of oakleaf butterflies within the Himalayan mountain ecosystem and natural selection produced complex, polymorphic adaptations.

Wang et al., 2022, Cell *185*, 3138–3152
August 18, 2022 © 2022 Elsevier Inc.
https://doi.org/10.1016/j.cell.2022.06.042

CellPress

CellPress

Cell

Article

The evolution and diversification of oakleaf butterflies

Shuting Wang,[1,2,19] Dequn Teng,[1,19] Xueyan Li,[3,19] Peiwen Yang,[1] Wa Da,[4] Yiming Zhang,[1,2] Yubo Zhang,[1,2] Guichun Liu,[3] Xinshuang Zhang,[5] Wenting Wan,[3] Zhiwei Dong,[3] Donghui Wang,[1,6] Shun Huang,[1] Zhisheng Jiang,[2] Qingyi Wang,[2] David J. Lohman,[7,8,9] Yongjie Wu,[10] Linlin Zhang,[11,12] Fenghai Jia,[13] Erica Westerman,[14] Li Zhang,[5] Wen Wang,[3,15,16] and Wei Zhang[1,2,17,18,20,*]

[1]State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing 100871, China
[2]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China
[3]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China
[4]Tibet Plateau Institute of Biology, Lhasa, Tibet 850001, China
[5]Chinese Institute for Brain Research, Beijing 100871, China
[6]National Teaching Center for Experimental Biology, Peking University, Beijing 100871, China
[7]Biology Department, City College of New York, City University of New York, New York, NY 10031, USA
[8]Ph.D. Program in Biology, Graduate Center, City University of New York, New York, NY 10016, USA
[9]Entomology Section, National Museum of Natural History, Manila 1000, Philippines
[10]Key Laboratory of Bio-Resources and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China
[11]CAS and Shandong Province Key Laboratory of Experimental Marine Biology, Center for Ocean Mega-Science, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China
[12]Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266071, China
[13]Jiangxi University of Traditional Chinese Medicine, Nanchang 330004, China
[14]Department of Biological Sciences, University of Arkansas, Fayetteville, AR 72701, USA
[15]School of Ecology and Environment, Northwestern Polytechnical University, Xi’an 710072, China
[16]Center for Excellence in Animal Evolution and Genetics, Kunming 650223, China
[17]Institute of Ecology, Peking University, Beijing 100871, China
[18]Institute for Tibetan Plateau Research, Peking University, Beijing 100871, China
[19]These authors contributed equally
[20]Lead contact
*Correspondence: weizhangvv@pku.edu.cn

https://doi.org/10.1016/j.cell.2022.06.042

SUMMARY

Oakleaf butterflies in the genus *Kallima* have a polymorphic wing phenotype, enabling these insects to masquerade as dead leaves. This iconic example of protective resemblance provides an interesting evolutionary paradigm that can be employed to study biodiversity. We integrated multi-omic data analyses and functional validation to infer the evolutionary history of *Kallima* species and investigate the genetic basis of their variable leaf wing patterns. We find that *Kallima* butterflies diversified in the eastern Himalayas and dispersed to East and Southeast Asia. Moreover, we find that leaf wing polymorphism is controlled by the wing patterning gene *cortex*, which has been maintained in *Kallima* by long-term balancing selection. Our results provide macroevolutionary and microevolutionary insights into a model species originating from a mountain ecosystem.

INTRODUCTION

Mountain ecosystems harbor extraordinarily diverse terrestrial biodiversity in which dramatic elevational gradients may have promoted the formation of microenvironments and accelerated diversification (Rahbek et al., 2019a). For example, geographic changes resulting from orogenic activity, combined with climate change, caused the Qinghai-Tibetan Plateau (QTP) uplift and produced diverse habitats for a variety of organisms (Favre et al., 2015). Montane biota and its relationship to lowland biodiversity hotspots have been topics of intense research interest, in which relevant studies have generally been conducted above the species level (Rahbek et al., 2019b). Here, we focus on cryptic oakleaf butterflies of the genus *Kallima* (Nymphalidae: Nymphalinae), whose members are phenotypically diverse and distributed in montane and lowland biodiversity hotspots (Küppers, 2015a; Myers et al., 2000; Tsukada, 1985). The highly variable wing patterns and pan-Asian distribution of this group make it uniquely suited to study how geographic changes and natural selection shape phenotypic and species diversity.

Because leaves are ubiquitous in nature, leaf masquerade mimicry exists widely throughout the animal kingdom as a special form of plant resemblance (Komárek, 2003), as observed in Amazon leaffish (*Monocirrhus polyacanthus*), Malayan leaf frog (*Megophrys nasuta*), leaf insects (Phylliidae), and leaf-mimic katydids (various Tetigoniidae). *Kallima* butterflies are among the most well-known and striking examples of leaf mimics, regarded by Alfred Russel Wallace as ‘‘the most wonderful and undoubted case of protective resemblance in a butterfly’’ (Wallace, 1889). When the wings of *Kallima* butterflies are folded back, they look like dead, brown leaves, including wing pattern elements resembling the leaf midrib, secondary veins, petiole, and variable markings resembling decay. The extreme, cryptic wing patterns of these species seem to provide strong evidence for natural selection (Beddard, 1892; Morgan, 1903), although many authors have denied that complex mimetic phenotypes can evolve gradually (Goldschmidt, 1945; Punnett, 1915; Watson, 1936). Therefore, *Kallima* butterflies have drawn considerable attention in multiple biological disciplines including taxonomy, phylogeography, physiology, and morphology, over the past few decades (Kamalanathan and Mohanraj, 2012; Küppers, 2015a, 2015b; Nakamura, 2014; Shirozu and Nakanishi, 1984; Tang et al., 2013; Zhou et al., 2013), and we recently reported a reference genome assembly of *Kallima inachus* (Yang et al., 2020). There is increasing evidence that the main characteristics of the leaf-like wing pattern tend to be produced by gradual selection rather than saltation (large, sudden mutational changes) (Alexander, 2019; Suffert, 1927; Suzuki et al., 2014, 2019). These studies have laid the foundation for exploration of the genomic basis of the various wing pattern elements that together constitute the mimetic phenotype.

In the present study, we performed a holistic analysis to characterize the genetic and species diversity of *Kallima* butterflies. At the macroevolutionary level, we traced the origin of the genus *Kallima* by sequencing the genomes of 105 butterflies from 21 nymphalid genera. Phylogenetic and population genetic analyses suggest that *Kallima* butterflies diversified on the Asian mainland, with several lineages dispersing to areas that are currently islands to the south and east. Subsequent demographic and biogeographical modeling efforts indicate that the eastern Himalayas are a center of *Kallima* diversification. At the microevolutionary level, to dissect the genomic basis of leaf wing polymorphism, we focused on *K. inachus* and performed genome-wide association studies (GWASs) by sequencing the genomes of 78 *K. inachus* butterflies with polymorphic phenotypes. A combination of evidence obtained from *de novo* genome assembly, gene expression, chromatin interaction analyses, and CRISPR-Cas9 genome editing led to the identification of a known wing pattern regulator, the *cortex* gene (Nadeau et al., 2016; Van’t Hof et al., 2016; van der Burg et al., 2020), that plays a major role in controlling leaf wing polymorphism, likely by regulating a potential downstream gene, *snake*, involved in butterfly pigmentation and color switching (Nishikawa et al., 2013; Yoda et al., 2020). Furthermore, through population genetic analyses, we show that leaf mimicry is maintained by balancing selection, possibly as a type of negative frequency-dependent selection, providing a rare and remarkable example with clear survival value supporting the role of long-term balancing selection in adaptive evolution. In summary, we explore the evolution of *Kallima* butterflies and provide genetic evidence for the morphogenesis of leaf mimicry, which serves as a complex, charismatic model for understanding microevolution, macroevolution, adaptation, and Darwinian gradualism.

RESULTS

Evolution and demography of *Kallima* butterflies

To characterize the evolutionary history of leaf mimicry in *Kallima* butterflies, we collected 36 samples from the genus *Kallima* including specimens of *Kallima knyvettii*, *Kallima limborgii amplirufa*, *Kallima paralekta*, *Kallima incognita*, two subspecies of *Kallima alicia*, and two subspecies of *K. inachus*, from 11 geographic locations in East and Southeast Asia (Figure 1A; Table S1). Notably, the sampling locations of three *Kallima* species overlapped in Medog, southeastern Tibet, suggesting that the eastern Himalayan region could represent either a hotspot of diversification or a glacial refuge. To test this hypothesis and disentangle the relationships among *Kallima* species and populations, we analyzed whole-genome sequencing data from 105 butterfly samples from 21 nymphalid genera and constructed phylogenies using genome-wide SNPs (Figure 1B; Figures S1A and S1B). The *Kallima* species formed a monophyletic group, but *Kallima* and the other leaf-mimicking species that we sampled formed a polyphyletic assemblage. These species included *Doleschallia bisaltide* (autumn leaf butterfly), *Kallimoides rumia* (African leaf butterfly), and *Junonia cymodoce* (blue leaf butterfly), suggesting multiple origins of leaf masquerade mimicry in Nymphalidae. Delving deeper into the genus *Kallima*, both principal component analysis (PCA) and ADMIXTURE analysis (Alexander et al., 2009) indicated the existence of lineages with distinct genetic structure roughly corresponding to recognized species (Figures 1C–1E; Figure S1C). We calculated the fixation index (F_{ST}) across the genome for *Kallima* butterflies and observed that the F_{ST} values increased at increasing levels in the taxonomic hierarchy: populations, subspecies, and species (Table S2). Furthermore, we estimated historical population sizes using pairwise sequentially Markovian coalescent (PSMC) and SMC++ models (Li and Durbin, 2011; Terhorst et al., 2017). Unlike other mainland *Kallima* butterflies, *K. knyvettii* has consistently had a small effective population size (Figure 2A; Figures S2A and S2B). Both *K. limborgii amplirufa* and *K. paralekta* had less nucleotide diversity (π) than other mainland species/subspecies (Table S2). To further investigate population dynamics of these species, we performed genome-wide demographic inference using the Generalized Phylogenetic Coalescent Sampler (G-PhoCS) (Gronau et al., 2011) and Treemix (Pickrell and Pritchard, 2012). Both population genetic analyses yielded evidence supporting extensive, directional gene flow from Medog in southeastern Tibet to insular populations of *K. alicia alicia* on Hainan and *K. paralekta* on Java, respectively; from the Medog population of *K. incognita* to *K. paralekta* on Java and *K. l. amplirufa* on the Thai-Malay

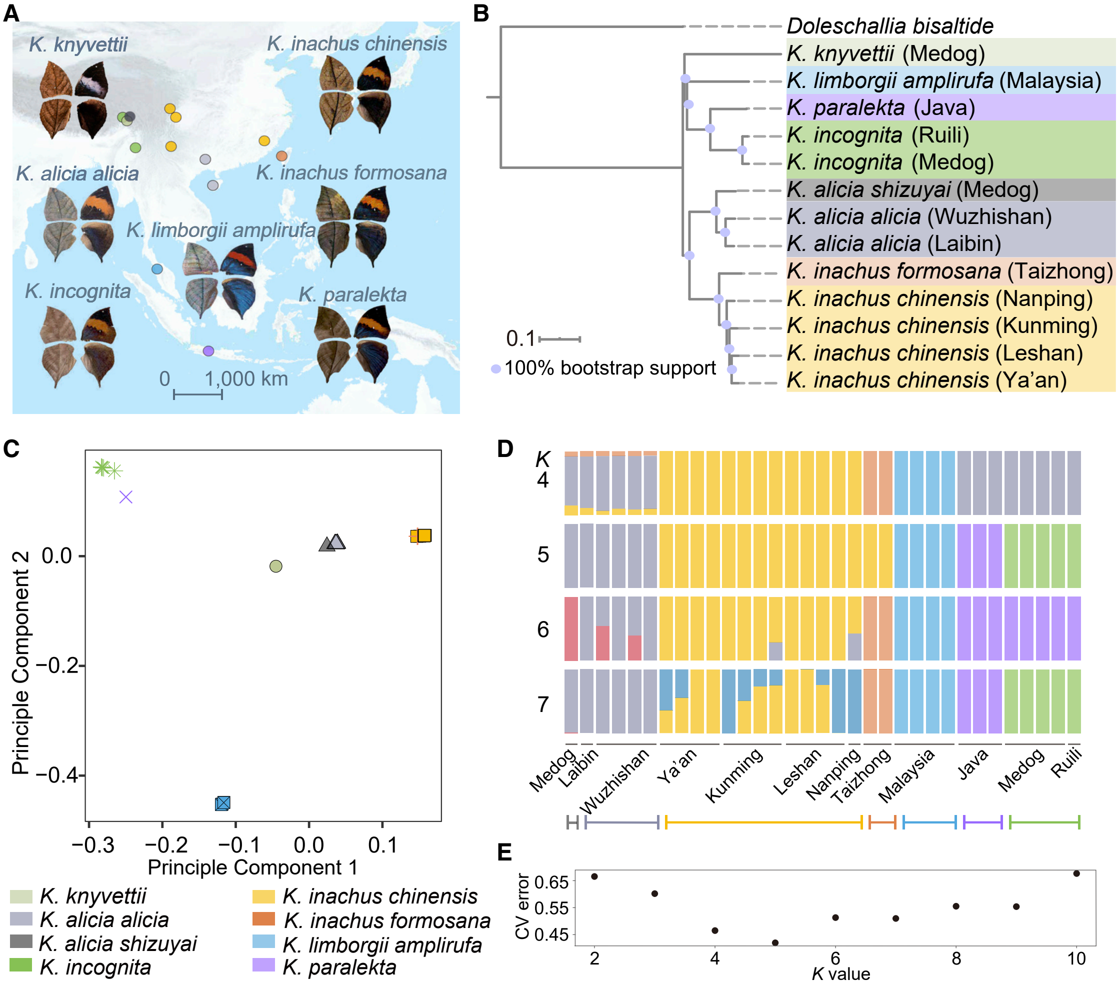


Figure 1. The geographic distribution, genome-wide phylogeny, and genetic structure of *Kallima* butterflies

(A) Collection locations of the *Kallima* samples we analyzed are shown in different colors, along with images of male dorsal and ventral wing patterns.

(B) A rooted maximum-likelihood phylogeny based on the genome-wide SNP data of 14 representative samples.

(C) Principal component analysis (PCA) shows separate groups of 34 *Kallima* samples from different species.

(D) Population structure analysis showing the genetic structure of 33 *Kallima* samples, with the colored fractions representing the genetic composition with *K* clusters.

(E) Cross-validation (CV) error plot for population structure analysis showing that *K* = 5 is the likely number of true genetic populations according to cross-validation (CV) error.

See also Figure S1 and Tables S1 and S2 in the supplemental information.

Peninsula, and from the nominate subspecies on the mainland to the island subspecies of *K. inachus formosana* on Taiwan (Figures 2B–2D; Figures S2D and S3; Table S3). However, little gene flow was detected between the mainland populations of *K. incognita* and *K. inachus chinensis* (Figure S3). These results suggest dispersal from the Asian mainland to peripheral islands following the earlier divergence of *K. incognita* and *K. inachus*. The inferred divergence times of *K. i. formosana* and *K. paralekta* in the mid-Pleistocene (Figures 2C and 2D) corresponded to a period of intense global sea level fluctuations that continued until the end of the Pleistocene (Voris, 2000; Woodruff, 2010). The continental islands of Hainan and Taiwan were connected to the mainland by land bridges, and Java, Sumatra, and Borneo were connected to the Thai-Malay peninsula in a single landmass known as Sundaland (Cannon et al., 2009; Lohman et al., 2011). Thus, it is likely that many *Kallima* taxa that are currently restricted to islands were formerly part of large, panmictic populations that became isolated when sea levels rose. We then evaluated the suitability of the climate for *Kallima* butterflies using the maximum entropy model (Maxent) (Phillips et al., 2006). The results of our niche model support the notion that the eastern Himalayan area has been climatically suitable for *Kallima* butterflies since the last interglacial (LIG) and that land bridges that formed during the last glacial maximum (LGM) facilitated secondary contact and gene flow among *Kallima* species/subspecies (Figure 2E; Figure S2C). In

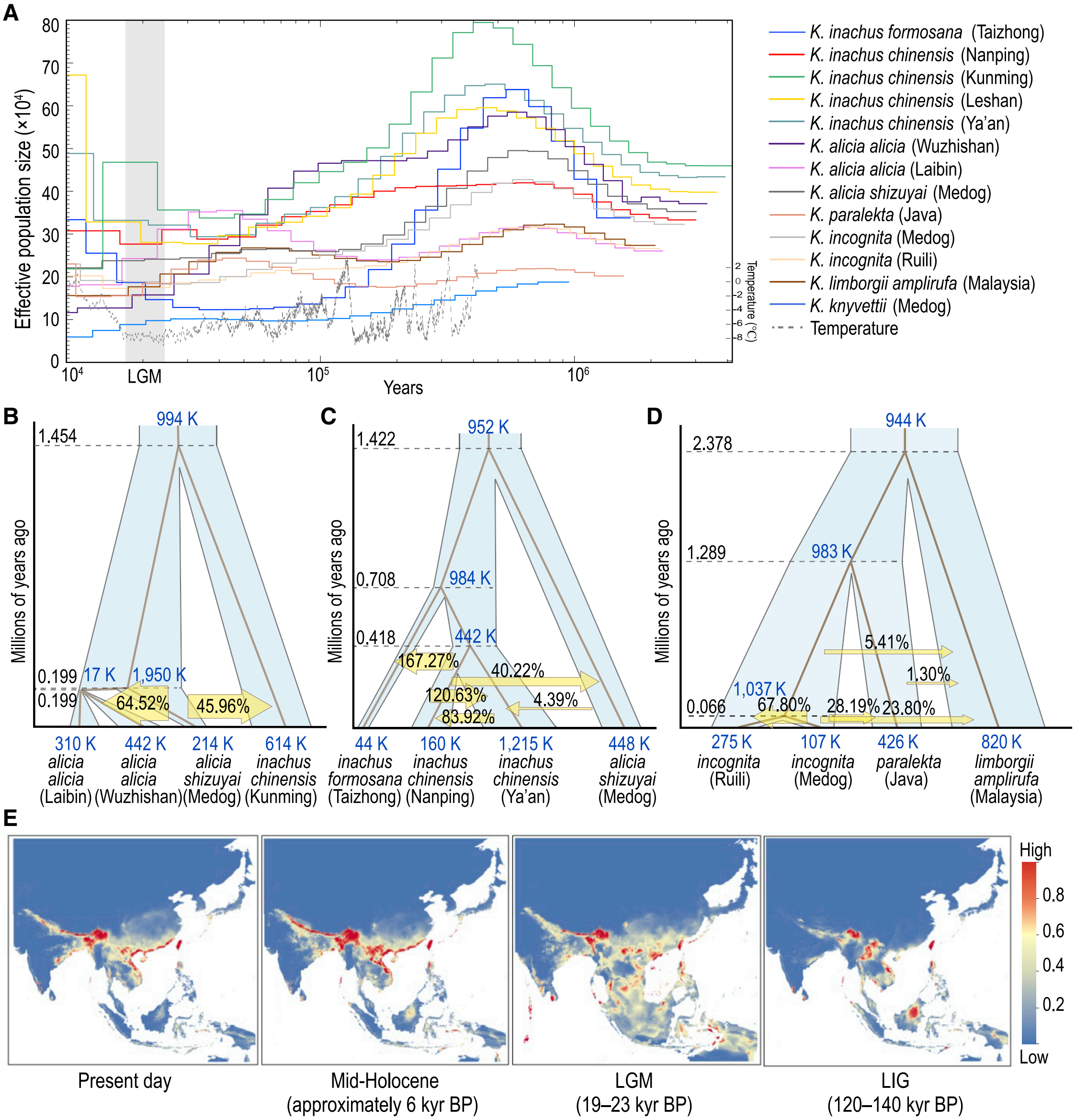


Figure 2. Demographic history of *Kallima* butterflies

(A) Historical effective population sizes inferred from 13 representative *Kallima* samples using PSMC, assuming a mutation rate of $\mu = 3 \times 10^{-9}$ and an average generation time of $g = 0.5$ years.

(B–D) Divergence times, effective population sizes, and migration rates inferred among *Kallima* butterflies using G-PhoCS. The dashed lines indicate divergence times, and the branch widths are proportional to the population sizes.

(E) Ecological niche modeling performed for *Kallima* butterflies using Maxent, where the suitable habitat is predicted beginning from the last interglacial (LIG) period. The color scheme indicates the probability of occurrence from high (red) to low (blue).

See also Figures S2 and S3 and Tables S2 and S3 in the supplemental information.

summary, these results suggest that *Kallima* butterflies likely differentiated in the eastern Himalayas and subsequently dispersed to areas that are now islands.

Genetic basis of leaf wing polymorphism

Since *Kallima* butterflies have polymorphic leaf wing patterns, we characterized the ventral wing morphology of *K. i. chinensis* into discrete wing forms (Figure 3A; Figures S4A and S4B). To assess phenotypic segregation, we established several crosses of *K. i. chinensis* with different wing forms and were fascinated to discover ten discrete leaf wing forms in total, which were likely determined by combinations of five alleles (*P*, *V*, *S*, *R*, and *M*, for *plain*, *veined*, *scrambled*, *rippled*, and *moldy*, respectively, describing the wing patterns) located at a single Mendelian locus (Figures 3A and 3B). To map the leaf mimicry locus genetically, we sequenced genomes of 78 *K. i. chinensis*

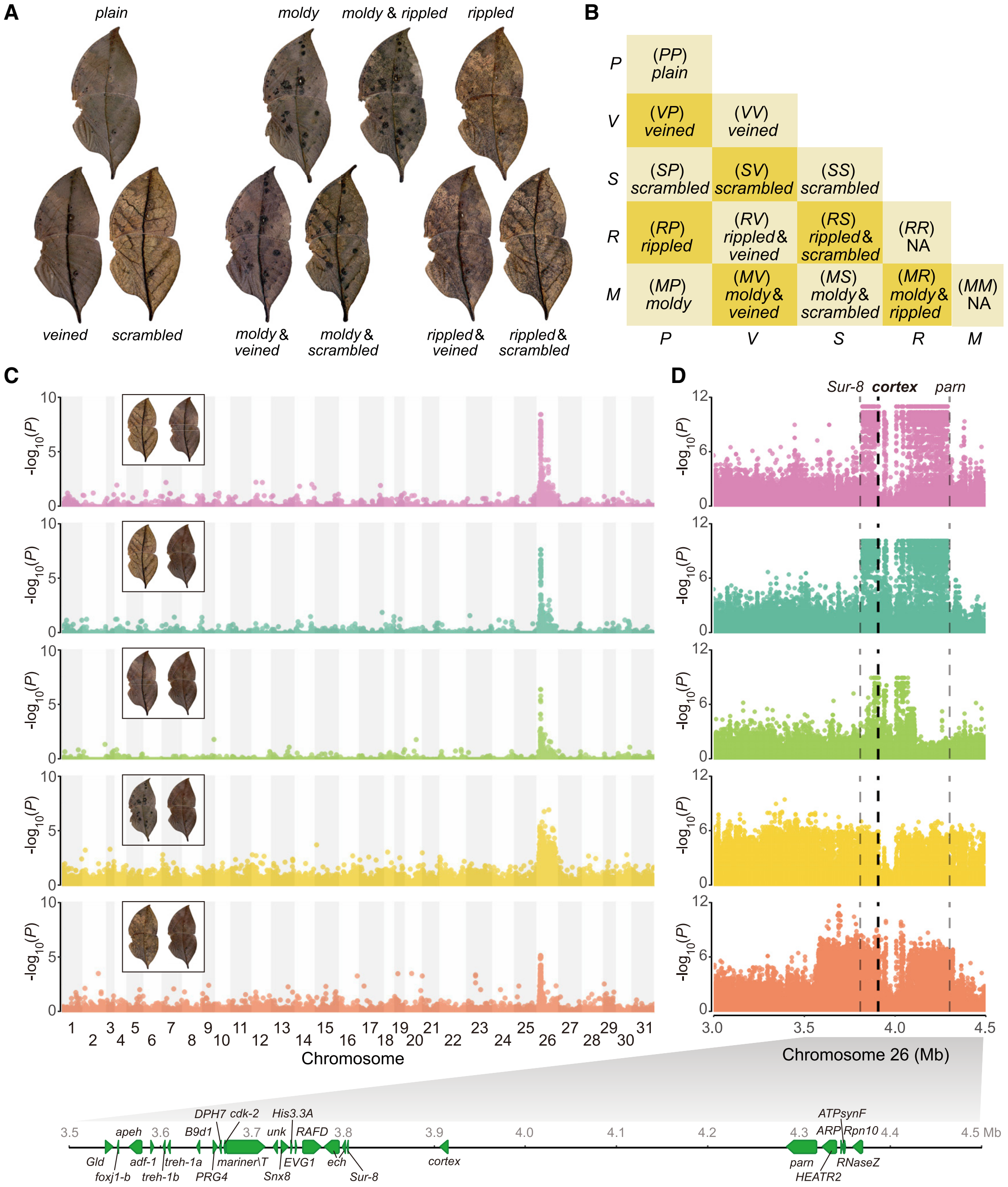


Figure 3. A major locus controlling leaf wing polymorphism in *K. inachus*

(A) Ten ventral wing forms demonstrating polymorphism.

(B) Five contributing *cortex* alleles and their dominant epistasis related to leaf wing patterns.

(C) Multiple genome-wide association studies performed on *veined* versus *scrambled* forms, *plain* versus *scrambled* forms, *plain* versus *veined* forms, *moldy* versus *plain* forms, and *rippled* versus *plain* forms. The leaf wing locus localized to a major peak of false discovery rate (FDR)-adjusted p values on chromosome 26.

(D) The *cortex* gene is located in the major peak on chromosome 26.

See also Figures S4 and S6 and Table S1 in the supplemental information.

specimens from China and conducted GWASs to compare multiple forms (Table S1). The top-associated GWAS hit was a striking, major peak on chromosome 26 (Figures 3C and 3D; Figures S4C–S4G). Moreover, we noted that *cortex* was the only gene located in the top-associated region identified in the genomic comparisons of multiple forms (Figure 3D; Data S1).

The gene is known to control wing and body color patterns in silkworms, Batesian mimicry polymorphism in *Papilio clytia*, Müllerian mimicry in *Heliconius* butterflies, industrial melanism in British peppered moths, and seasonal wing coloration in *Junonia coenia* (Ito et al., 2016; Nadeau et al., 2016; Van't Hof et al., 2016; van der Burg et al., 2020; VanKuren et al., 2019). Therefore, *cortex* is particularly important for lepidopteran wing morphogenesis, and variants of this gene have repeatedly experienced natural selection. Notably, these associated hits all presented strong but different patterns of linkage disequilibrium (LD) within the *cortex* region (Figure 4A); thus, we hypothesized that multiple mechanisms may independently lead to the reduction or suppression of recombination among different *cortex* haplotypes. To test this hypothesis, we performed *de novo* genome assembly for four *K. i. chinensis* samples and obtained four individual genomes containing five complete *cortex* haplotypes (Figure 4B; Table S1). Indeed, we identified two chromosomal inversions spanning haplotypes *V* and *R* with distinct lengths and boundaries, corresponding to different patterns of associated hits (Figure 4B; Figure S5A). Given that the *cortex* genes of two other nymphalid genera, *Heliconius* and *Junonia*, have a conserved distal transcription start site (TSS) specific to wing development (Hanly et al., 2022; Nadeau et al., 2016; van der Burg et al., 2020), we also annotated the distal TSS and additional exons based on available RNA-seq data and sequence homology in genomic DNA for haplotypes *V* and *P* (Figure 4B). The inversion observed in *V* contained only *cortex* and its predicted proximal promoter and exon regions based on RNA-seq data, but the predicted distal promoter region based on homology was excluded and showed the opposite orientation, whereas the inversion found in *R* contained both *cortex* and multiple flanking genes, such as two *trehalase* genes (*Treh-1a* and *Treh-1b*) (Data S1). Since *Treh-1a* is required for the development of the red autumn morph of *Junonia coenia* (van der Burg et al., 2020), these *trehalase* genes might also be involved in the expression of phenotype *R*. Moreover, a phylogenetic analysis of the five haplotypes resulted in separate clusters of *V* and *R* grouped with *M* and *S*, respectively, suggesting the involvement of independent chromosomal rearrangement across the *cortex* region (Figures S5B and S5C). We also addressed the reduced linkage in *S* using high-throughput chromosome conformation capture (Hi-C) based on one individual with an *SS* genotype, which revealed that the *cortex* region was a putative topologically associating domain (TAD) including only the *cortex* gene (Figure 4C; Figure S5E; Data S1). A related observation was that this putative TAD was not detected in the Hi-C data generated using several samples including *P*, *S*, and *R* haplotypes. Moreover, we demonstrated that the *cortex* gene in *M* had a pronounced signature of positive selection (branch-site model, $p < 0.01$) in the form of a nonsynonymous substitution (K96H) (empirical Bayes analysis, $p > 95\%$) (Figures S5B–S5D). To further investigate the reduced recombination between *cortex* haplotypes, we calculated the proportion of transposable elements (TEs) among the five haplotypes and observed that more long interspersed nuclear elements (LINEs) and long terminal repeat (LTR) elements accumulated in the four derived haplotypes than in the original haplotype *P* (Figure S5F). The accumulation of LINEs and LTR elements mainly accounted for the differences in the proportion of TEs observed between *cortex* haplotypes. In particular, more LINEs accumulated near the breakpoint regions of the inversions in *V* (21.99%) and *R* (21.47%) than in the homologous regions in *P* (12.62% and 15.4%), indicating that TEs may contribute to the reduction in recombination between *cortex* haplotypes. We also performed functional validation of *cortex* using CRISPR-Cas9-mediated genome editing. Loss of *cortex* function in mosaic knockouts (mKOs) led to the disruption of leaf wing phenotypes (Figures 4D, S5G, and S5H). Scale pigmentation on the dorsal and ventral wings of *cortex* mKOs was faded around the lateral vein region, and the pattern elements mimicking the leaf midrib vein were blurred and darkened, which was consistent with the effects of *cortex* mKO observed in multiple *Heliconius* species, *Danaus plexippus*, and *Junonia coenia* (Livraghi et al., 2021; van der Burg, 2020) and suggested a conserved function of *cortex* in regulating butterfly scale cell development. In sum, our results provide clear evidence that the *cortex* gene plays a key role in controlling leaf mimicry, which is shaped by natural selection, and both structural and nonstructural genetic mechanisms have contributed to maintaining leaf wing polymorphism.

To investigate the development of leaf wing patterns in *K. i. chinensis*, we explored the expression of *cortex* via quantitative reverse transcription PCR (qRT-PCR) and evaluated transcriptome-wide gene expression via RNA-seq in *veined* (*VP*) and *plain* (*PP*) forms during a series of wing developmental stages (Figures 5A–5C; Table S1; Data S1). In both forms, *cortex* was expressed at low levels based on bulk RNA-seq data. However, we did identify a number of known Hox and wing patterning genes that were differentially expressed in our samples. Notably, *ultrabithorax*, which is important for determining hindwing identity in many species of Lepidoptera (Matsuoka and Monteiro, 2021; Tendolka et al., 2021), was also upregulated in the hindwings of *K. i. chinensis* (Figure 5B). Subsequently, we examined the expression of *cortex* in qRT-PCR comparisons, which showed similar and variable expression in *plain* and *veined* individuals in the fifth-instar larval stage and the prepupal stage, likely owing to its much lower expression level than the reference gene *ef1*α (Figure 5C). However, we observed significant upregulation and wing-biased expression of *cortex* in *veined* individuals relative to *plain* individuals 1 day after pupation, followed by a decreased but still significantly high expression level 3 days after pupation (Figure 5C). In addition, the *in situ* hybridization of *cortex* in fifth-instar larval-stage wings showed spatial correspondence to the margin region of the adult wing in the *plain* (*PP*) form (Figure S6A). Our results demonstrate the spatiotemporal regulation of *cortex* expression across developmental stages, tissues, and forms.

According to the observed expression profiles of *cortex*, we characterized the differentially expressed genes between body and wing tissues, forewings and hindwings, and *veined* and *plain* forms in focal developmental stages, which showed different expression patterns between body and wing tissues and generally upregulated expression patterns in hindwings relative to forewings (Figure 5B; Data S1). Notably, the comparison of *veined* (*VP*) and *plain* (*PP*) hindwing discs on days 1 and 3 after pupation identified a significantly upregulated gene, *snake* (Figure 5B; Data S1). Since *snake* is involved in swallowtail butterfly

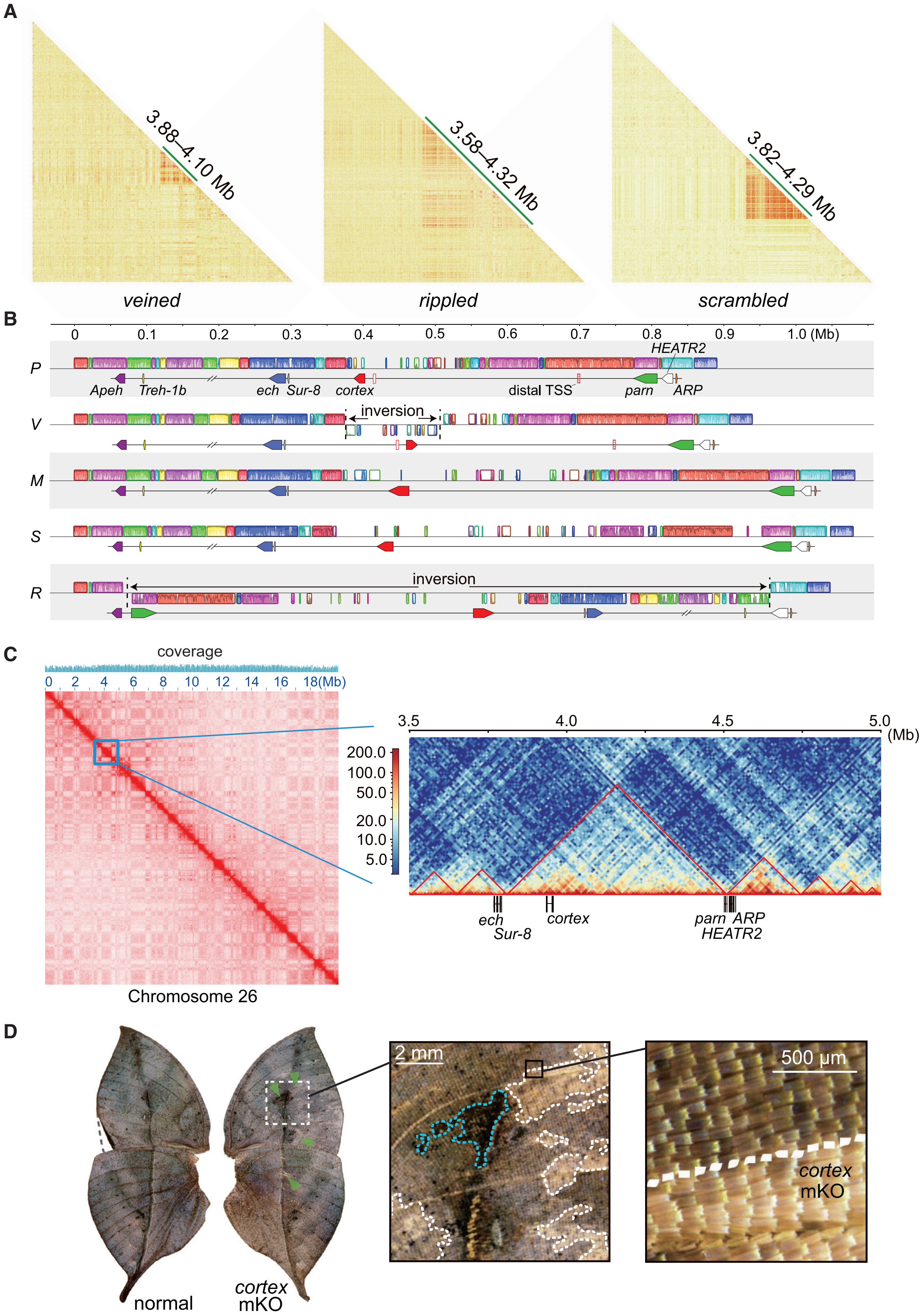


Figure 4. Suppressed recombination in *cortex* haplotypes and functional validation in *K. inachus*

(A) Linkage disequilibrium (LD), measured as r^2, is plotted and the elevated LD is observed around the *cortex* region in the samples of *veined*, *rippled*, and *scrambled* forms.

(legend continued on next page)

pupal and wing pigmentation (Nishikawa et al., 2013; Yoda et al., 2020) and shows expression patterns correlated with those of *cortex*, we consider it a putative downstream target in the leaf wing developmental pathway. We also examined the expression of genes adjacent to *cortex* and identified multiple genes that were upregulated in pupal wings relative to the abdomen, such as *B9d1*, *PRG4*, *DPH7*, *His3.3A*, and *parn* (Data S1). Many of these genes were involved in the long inversion region of *R*, suggesting their potential roles in forming the *rippled* pattern. Moreover, we noted that another wing pattern gene, *WntA*, (Martin et al., 2012; Mazo-Vargas et al., 2017) was upregulated 1 day after pupation and showed similar patterns of expression in *veined* (*VP*) and *plain* (*PP*) forms, suggesting that it may play a conserved role in wing development of different forms (Figures 5D and 5E; Data S1).

Evolution of leaf wing polymorphism

Leaf wing polymorphism originated prior to diversification of the genus *Kallima* and has been maintained throughout its evolutionary history (Figures 6A–6C). Based on the dominance and recessiveness as well as the divergence times of different haplotypes (Figures 3A, 3B, and 6A–6C), haplotypes *M* and *S* were likely derived from the original, recessive haplotype *P*, whereas independent chromosomal inversions led to haplotypes *V* and *R*. Therefore, we hypothesized that the frequencies of the polymorphic forms, such as trans-specific polymorphisms, which are maintained much longer than other mimicry polymorphisms, have been stabilized by long-term balancing selection (Joron and Mallet, 1998; Mallet and Joron, 1999). Nevertheless, the theoretical explanation for this hypothesis may be similar to the explanation for polymorphism in Batesian mimicry, in that imperfect leaf mimicry may be improved by suppressing the recombination of different leaf mimetic haplotypes (Booker et al., 2015; Charlesworth and Charlesworth, 1975). We examined the nucleotide variation in *cortex* and neighboring genes by comparing different haplotypes using Hudson-Kreitman-Aguadé (HKA) tests (Hudson et al., 1987), which consistently revealed a significant excess of nucleotide polymorphism in *cortex* relative to neighboring genes in comparisons among *M*, *P*, *S*, and *V*. This excess polymorphism suggests the effect of balancing selection on *cortex*, which may be an important driving force for maintaining leaf wing polymorphism in *Kallima* butterflies (Table S4). One exception was the increased polymorphism of both *cortex* and *parn* in comparison with *R*, revealing strong LD and genetic hitchhiking due to the long span of the chromosomal inversion found in *R* (Figure 4B; Table S4). We also observed varying intron lengths in *cortex* relative to the stable intron lengths observed in neighboring genes, indicating the existence of intron variants in *cortex* and a possible role of alternative splicing or expression regulation in promoting leaf wing polymorphism (Figure S6B).

To further dissect the type of balancing selection affecting *Kallima*, we counted phenotype and genotype frequencies and selection coefficients in a local population of *K. i. chinensis* and observed a large proportion of *P*, *V*, and *S* haplotypes and a small proportion of recently differentiated *R* and *M* haplotypes in this population (Figure S6C; Data S2). In addition, the inferred *cortex* sequence network had five distinct clades, supporting the stabilization of polymorphism after early origination (Figure S6D). Since *R* and *M* only existed in heterozygotes in both field samples and butterflies obtained from lab crosses (Figure 3B; Figure S6C), we hypothesized that homozygote disadvantage limited *R* and *M* to low frequencies with respect to the major role of frequency-dependent selection in maintaining leaf wing polymorphism, similar to a recently reported scenario of inversion polymorphism and mutation load in *Heliconius numata* (Jay et al., 2021). We further explored the driving forces of leaf wing polymorphism by simulating evolutionary scenarios with different population genetic parameters, and a model that showed a good fit to the observed data was an outbred population subject to negative frequency-dependent selection with different fitness levels assigned to different forms (Figure S6E; Data S2). These results point to an important role of long-term balancing selection in maintaining *Kallima* leaf wing polymorphism, with both negative frequency-dependent selection and homozygote disadvantage acting on different haplotypes.

DISCUSSION

Since the discovery of homeobox genes (Laughon and Scott, 1984; McGinnis et al., 1984; Scott and Weiner, 1984), the roles of genes in the evolution of morphological diversity have been investigated in multiple systems with the aim of interpreting and synthesizing developmental constraints and evolvability (Carroll, 2008; Gilbert et al., 1996; Hoekstra and Coyne, 2007). Notably, multiple studies on color patterns have made great strides in revealing the underlying genetics of phenotypic diversity, such as inversion polymorphisms of the *P* locus controlling mimicry wing patterns in *Heliconius numata* (Joron et al., 2011), the *agouti-related peptide 2* gene involved in switching stripe patterns in African cichlid fishes (Kratochwil et al., 2018), and the transcription factor-encoding *Pannier* gene, leading to melanic pattern polymorphisms in the harlequin ladybird (Ando et al., 2018; Gautier et al., 2018). In the present study, we used diverse leaf wing butterflies as a system to address long-standing questions concerning the evolution and maintenance of complex, mimetic wing patterning and polymorphisms. By characterizing the evolution and underlying genetic basis of leaf

(B) Alignment of the *cortex* region among five haplotypes suggests two different chromosomal inversions in the *V* and *R* haplotypes, where locally collinear blocks are shown in the same color. The predicted distal transcription start site (TSS) and additional exons of *cortex* are annotated and labeled in the *P* and *V* haplotypes based on RNA-seq data (blank rectangle) or sequence homology (dashed rectangle).

(C) Heatmaps showing the normalized average interaction frequencies of all putative topologically associating domains (TADs) on chromosome 26 as well as a putative TAD containing *cortex* and neighboring genes in haplotype *S*.

(D) Whole wings showing *cortex* loss-of-function effects caused by CRISPR-Cas9 genome editing on the ventral wing surfaces. *Cortex* mosaic knockout (mKO) results in faded scale pigmentation around the lateral vein region and a blurred and darkened leaf midrib on the ventral side (indicated by green arrows).

See also Figure S5, Tables S1 and S4, and Data S1 in the supplemental information.

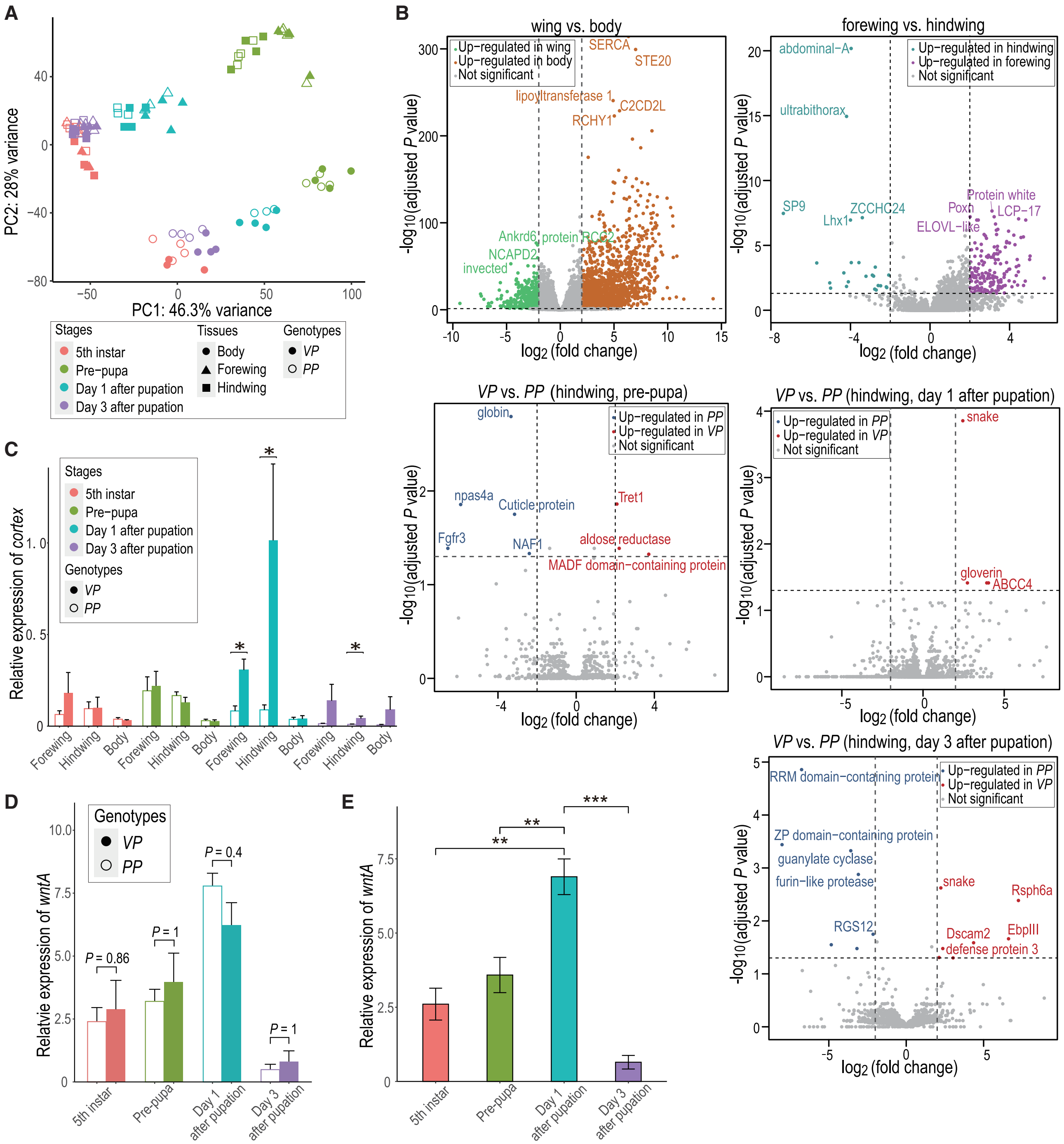


Figure 5. Gene expression patterns in *K. inachus*

(A) Principal component analysis showing that individual samples form clusters according to the developmental stage and tissue, regardless of the *cortex* genotype.

(B) Volcano plots showing differentially expressed genes between wing and body tissues and between forewing and hindwing tissues at all developmental stages and between the *VP* and *PP* genotypes in the prepupal stage 1 and 3 days after pupation.

(C) Spatiotemporal expression patterns of *cortex* were examined by qRT-PCR, revealing significantly elevated expression in the *VP* forewing relative to the *PP* forewing on day 1 after pupation and in the *VP* hindwing relative to the *PP* hindwing on day 3 after pupation. n = 3 for samples with the *VP* genotype in the fifth-instar larval stage and n = 4 for the other samples. * indicates p < 0.05, Wilcoxon test.

(legend continued on next page)

resemblance in *Kallima* butterflies, we determined that the evolution of leaf wing polymorphism was a gradual process in which toolkit genes were involved in producing these diverse wing patterns. In particular, *cortex* plays a complex role in controlling leaf mimicry diversity. There is clear evidence of phenotypic diversity driven by natural selection acting on chromosomal inversions and dimensional chromatin interactions at the DNA level. In addition, there are differences in the spatiotemporal regulation of gene expression at the RNA level. These findings demonstrate that the wing pattern gene *cortex* has been repeatedly employed in generating adaptive wing patterns in Lepidoptera and, thus, further illustrate the genetic basis and evolvability of wing development under the highly conserved nymphalid ground plan (Jiggins et al., 2017; Oliver et al., 2012; Van Belleghem et al., 2017).

In addition to resynthesizing developmental constraints and evolvability in butterfly wings, the trans-specific leaf wing polymorphism provides extraordinary evidence of plant resemblance phenotypes subject to long-term balancing selection, whereas most other mimetic butterflies exhibit intraspecific balanced polymorphism causing them to resemble toxic species, such as the rich local polymorphism observed in Müllerian mimetic *Heliconius numata* (Jay et al., 2021; Joron et al., 2011) and Batesian mimetic *Elymnias hypermnestra* (Ruttenberg et al., 2021). In contrast to leaf masquerade species, which, in the case of *Kallima*, have resembled leaves for millions of years aided by balanced polymorphism, Müllerian and Batesian mimetic butterflies rapidly evolve different wing patterns. For example, the butterflies of the Batesian mimetic *Papilio polytes* species group share an ancestral polymorphism, but two of the three species in the species group have secondarily lost this polymorphism (Zhang et al., 2017). Owing to the generally transient nature of balancing selection caused by the creation of genetic load, the majority of available relevant evidence comes from intraspecific studies (Hedrick, 2007). Balanced polymorphism shared by multiple species is relatively unusual but has been documented in genes associated with host-pathogen interactions, vertebrate major histocompatibility complex (MHC) genes, and genes involved in the responses to biotic and abiotic stresses in *Arabidopsis thaliana* and related species, among others (Fijarczyk and Babik, 2015; Wu et al., 2017); this phenomenon is difficult to resolve without clear functional information. Thus, functional evidence of long-term balanced wing polymorphism in *Kallima* butterflies provides rare insight into the nature of selection maintaining polymorphisms, shedding light on these forces driving evolutionary innovation and genetic diversity.

Furthermore, we explored how species and genetic diversity has been shaped by climate change and the Earth's history at macro and micro scales. Our findings indicate the eastern Himalayan region where the rapid uplift of the QTP occurred (Coleman and Hodges, 1995; Royden et al., 2008) as a center of *Kallima* butterfly diversification. We integrated microevolutionary and macroevolutionary approaches with evidence from genomic, demographic, and biogeographic analyses to trace the origin and evolution of *Kallima* butterflies. Diversification coincided with the uplift period of the QTP during the late Miocene and early Pliocene (Figures 6A and 6B; Li and Fang, 1999; Royden et al., 2008; Zheng et al., 2000). The orogeny of the plateau formed a dramatic elevational gradient and diversified habitats for butterflies in southeastern Tibet and northwestern Yunnan (Favre et al., 2015), creating opportunities for isolation, habitat specialization, and divergence. Today, this region supports a diverse and highly endemic fauna, which allowed us to identify several *Kallima* species in southeastern Tibet and detect directional gene flow from this region. However, some rare *Kallima* species with limited geographic ranges remain unsampled; therefore, although this study offers insight into the conserved gene regulatory networks and spectacular phenotypic diversity of *Kallima* butterflies, many questions about the evolution of these astounding, leaf-mimicking butterflies remain to be answered more than 150 years after they were first observed by Alfred Russel Wallace.

Limitations of the study

Due to low frequencies of *K. i. chinensis* with rare leaf wing forms in nature, we did not rear and collect sufficient individuals, which prevented us from delving into the detailed functional basis of all leaf-wing patterns, especially the potential functional elements in multiple *cortex* haplotypes and their regulatory features. Additionally, given the limited genomic DNA quality of specimens from *K. alicia*, *K. incognita*, *K. paralekta*, and *K. limborgii*, we did not perform *de novo* assembly of *cortex* haplotypes from these species. We were also unable to sample *K. albofasciata*, *K. horsfieldii*, or *K. spiridiva*. Therefore, additional *Kallima* specimens will be needed in follow-up studies to adequately address the above issues.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Butterfly rearing and collection
- METHOD DETAILS
  - Sample preparation and sequencing
  - Genotype calling
  - Phylogenetic analyses

(D and E) Spatiotemporal expression patterns of *WntA* were analyzed for each *PP* or *VP* genotype (D) or for a combination of *PP* and *VP* (E) in the hindwing tissues using RNA-seq. The samples with *PP* and *VP* genotypes show similar patterns of expression (D) that increase significantly on day 1 after pupation (E). ** indicates adjusted $p < 0.01$ and *** indicates adjusted $p < 0.001$ (Wilcoxon test, normalized using the trimmed mean of the M values).
(C–E) Data are presented as the mean (SEM).
See also Figure S6, Table S1, and Data S1 in the supplemental information.

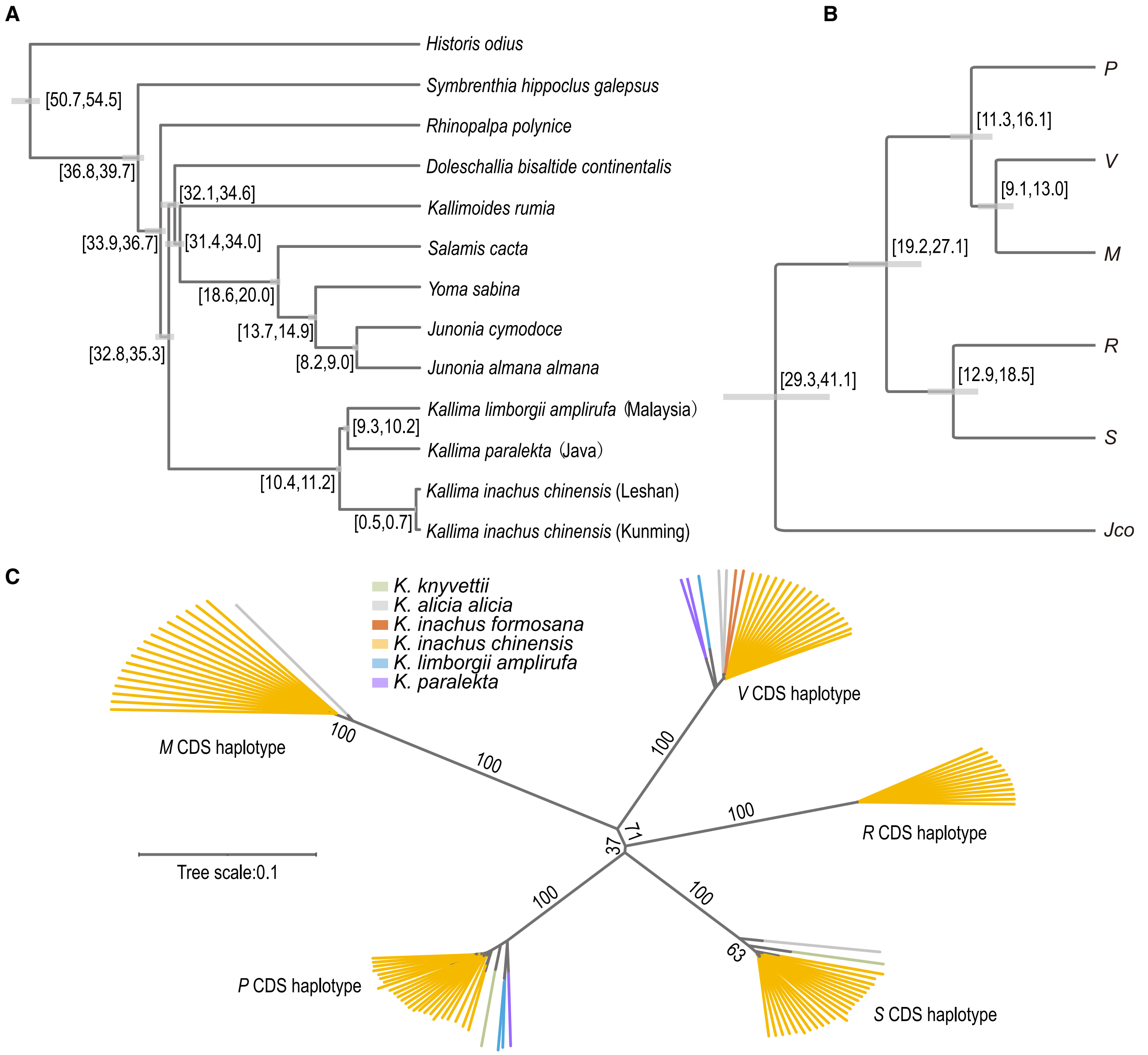


Figure 6. Maximum-likelihood phylogenies of five *cortex* haplotypes and coding sequence (CDS) regions in *Kallima* butterflies and of genome-wide SNP data in nymphalid butterflies

(A) Genome-wide SNP data were used to construct a maximum-likelihood phylogenetic tree for *Kallima* and other nymphalid butterflies. Divergence times are calibrated according to the divergence times of *Kallima* and *Junonia* (between 31.0033 and 44.768 mya), *Kallimoides* and *Junonia* (between 27.8345 and 40.7399 mya), and *Historis* and *Kallima* (between 37.827 and 54.7998 mya).

(B) A 50-kb region of the alignment of *cortex* haplotypes in *K. inachus* with strong linkage disequilibrium was used to construct a maximum-likelihood phylogenetic tree with *Junonia coenia* (*Jco*) as the outgroup. Divergence times are calibrated based on the divergence of *Kallima* and *Junonia* (between 31.0033 and 44.768 mya).

(C) A maximum-likelihood phylogeny was constructed based on phased *cortex* CDSs extracted from individuals of *K. inachus*, *K. alicia*, *K. limborgii*, *K. paralekta*, and *K. knyvettii*. Sequences with the same *cortex* haplotype form a well-supported monophyletic clade, indicating ancestral polymorphism and differentiation of *cortex* haplotypes. The numbers indicate bootstrap support values.

See also Figure S6, Table S4, and Data S2 in the supplemental information.

○ Population structure inference
○ Population genetic analyses
○ Demographic analyses
○ Habitat suitability modeling
○ Genome assembly and synteny mapping
○ Hi-C library data analysis
○ Annotation of transposable elements and promoter regions

  - ○ RT–qPCR analyses
  - ○ CRISPR/Cas9-mediated genome editing
  - ○ Transcriptome analyses and assembly
  - ○ *In situ* hybridization
  - ○ Population genetics simulations
  - ○ Genotyping
- ● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.06.042.

ACKNOWLEDGMENTS

We thank the Medog Forestry Bureau (China), Medog National Natural Reserve (China), Ya'an Forestry Bureau (China), Fengtongzhai National Natural Reserve (China), Wuzhishan National Natural Reserve (China), Ruili Forestry Bureau (China), Qirong Chen, Jianglang Wu, Shuai Zhan, Ningning Wu, and Lixia Yu for assistance with sample collection. We thank Han Yang, Lei Cao, Jiaxin Ni, Bingyuan Zhou, Fengyuan Li, Hua Zeng, Yanchen Jiang, Xiujia Sun, Ningning Wu, Shuhui Wang, Zixuan Zhang, Tian Lu, Wei Lu, Sufan Yuan, Shuyu He, Weifan Lv, Xingwei Wang, Zi'ang Liu, and Xinlu Ren for assistance with butterfly rearing and helpful discussions. We thank Cheng Li, Yuting Liu, Chao Zhang, and Qingyi Zhou for their assistance with data analysis and helpful discussions. We thank Jinwu He, Qian Shu, and Jiaxin Ni for their help in gene editing or *in situ* hybridization experiments. We thank Manyuan Long, Hongya Gu, and Fuwen Wei for helpful discussions. We thank the National Center for Protein Sciences at Peking University for assistance with qRT-PCR, and the Computing Platforms of the Center for Life Sciences and the School of Life Sciences for assistance with computation. This project was supported by grants from the National Natural Science Foundation of China (32170420 and 31871271 to W.Z. and 32070482 to X.L.), the Beijing Natural Science Foundation (JQ19021 to W.Z.), the Peking-Tsinghua Center for Life Science, the State Key Laboratory of Protein and Plant Gene Research, Qidong-SLS Innovation Fund to W.Z., the Second Tibetan Plateau Scientific Expedition and Research program (2019QZKK05010603) to W.D., the Major Special Program of the Tibet Autonomous Region (XZ201701-GA-05) to W.D., the Chinese Institute for Brain Research to L.Z., US NSF grant DEB-1541557 to D.J.L., and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB42000000) to L.Z.

AUTHOR CONTRIBUTIONS

S.W., D.T., P.Y., Yiming Zhang, Yubo Zhang, and W.Z. carried out experiments, analyzed the data, reared butterflies, and collected samples. X.L., G.L., W. Wan, Z.D., D.J.L., D.W., Z.J., Q.W., and Linlin Zhang carried out experiments and analyzed the data. W.D., F.J., Y.W., and E.W. helped design the project and collect samples and interpreted the results. S.H. reared butterflies and collected samples. X.L. and W. Wang helped design the project, provided sequencing data, and interpreted the results. X.Z. and Li Zhang prepared samples for sequencing and interpreted the results. S.W., D.T., and P.Y. contributed to drafting the manuscript. W.Z. conceived and designed the study and wrote the manuscript with input from all coauthors. All authors proofread and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 9, 2021
Revised: January 20, 2022
Accepted: June 22, 2022
Published: August 1, 2022

REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

Alexander, V.N. (2019). The mechanism for mimicry: instant biosemiotic selection or gradual Darwinian fine-tuning selection? Biosemiotics *12*, 39–55.

Ando, T., Matsuda, T., Goto, K., Hara, K., Ito, A., Hirata, J., Yatomi, J., Kajitani, R., Okuno, M., Yamaguchi, K., et al. (2018). Repeated inversions within a pannier intron drive diversification of intraspecific colour patterns of ladybird beetles. Nat. Commun. *9*, 3843.

Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics *21*, 263–265.

Bassett, A.R., Tibbit, C., Ponting, C.P., and Liu, J.L. (2013). Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. Cell Rep. *4*, 220–228.

Beddard, F.E. (1892). Animal Coloration: an Account of the Principal Facts and Theories Relating to the Colours and Markings of Animals (S. Sonnenschein & Company).

Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi-C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268–276.

Bigelow, N.H., Brubaker, L.B., Edwards, M.E., Harrison, S.P., Prentice, I.C., Anderson, P.M., Andreev, A.A., Bartlein, P.J., Christensen, T.R., Cramer, W., et al. (2003). Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. J. Geophys. Res. *108*. https://doi.org/10.1029/2002JD002558.

Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., et al. (2004). Aligning multiple genomic sequences with the threaded blockset aligner. Genome Res. *14*, 708–715.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Booker, T., Ness, R.W., and Charlesworth, D. (2015). Molecular evolution: breakthroughs and mysteries in Batesian mimicry. Curr. Biol. *25*, R506–R508.

Brown, J.L., Bennett, J.R., and French, C.M. (2017). SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ *5*, e4095.

Cannon, C.H., Morley, R.J., and Bush, A.B.G. (2009). The current refugial rainforests of Sundaland are unrepresentative of their biogeographic past and highly vulnerable to disturbance. Proc. Natl. Acad. Sci. USA *106*, 11188–11193.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell *134*, 25–36.

Charlesworth, D., and Charlesworth, B. (1975). Theoretical genetics of Batesian mimicry II. Evolution of supergenes. J. Theor. Biol. *55*, 305–324.

Chazot, N., Wahlberg, N., Freitas, A.V.L., Mitter, C., Labandeira, C., Sohn, J.C., Sahoo, R.K., Seraphim, N., de Jong, R., and Heikkilä, M. (2019). Priors and posteriors in Bayesian timing of divergence analyses: the age of butterflies revisited. Syst. Biol. *68*, 797–813.

Clement, M., Posada, D., and Crandall, K.A. (2000). TCS: a computer program to estimate gene genealogies. Mol. Ecol. *9*, 1657–1659.

Coleman, M., and Hodges, K. (1995). Evidence for Tibetan Plateau uplift before 14 Myr ago from a new minimum age for east–west extension. Nature *374*, 49–52.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Darling, A.E., Mau, B., and Perna, N.T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One *5*, e11147.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. Bioinformatics *34*, 2666–2669.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol. Biol. Evol. *29*, 1969–1973.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016a). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101.

Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016b). Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. *3*, 95–98.

Favre, A., Päckert, M., Pauls, S.U., Jähnig, S.C., Uhl, D., Michalak, I., and Muellner-Riehl, A.N. (2015). The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. Biol. Rev. Camb. Philos. Soc. *90*, 236–253.

Fijarczyk, A., and Babik, W. (2015). Detecting balancing selection in genomes: limits and prospects. Mol. Ecol. *24*, 3529–3545.

Gautier, M., Yamaguchi, J., Foucaud, J., Loiseau, A., Ausset, A., Facon, B., Gschloessl, B., Lagnel, J., Loire, E., Parrinello, H., et al. (2018). The genomic basis of color pattern polymorphism in the harlequin ladybird. Curr. Biol. *28*, 3296–3302.e7.

Gilbert, S.F., Opitz, J.M., and Raff, R.A. (1996). Resynthesizing evolutionary and developmental biology. Dev. Biol. *173*, 357–372.

Goldschmidt, R.B. (1945). Mimetic polymorphism, a controversial chapter of Darwinism. Q. Rev. Biol. *20*, 147–164.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652.

Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G., and Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. *43*, 1031–1034.

Guindon, S. (2010). Bayesian estimation of divergence times from large sequence alignments. Mol. Biol. Evol. *27*, 1768–1781.

Hanly, J.J., Livraghi, L., Heryanto, C., McMillan, W.O., Jiggins, C.D., Gilbert, L.E., and Martin, A. (2022). A large deletion at the cortex locus eliminates butterfly wing patterning. G3(Bethesda) *12*. jkac021.

Harris, R.S. (2007). Improved pairwise alignment of genomic DNA. Ph.D. Thesis (The Pennsylvania State University).

Harrison, S.P., Yu, G., Takahara, H., and Prentice, I.C. (2001). Palaeovegetation: diversity of temperate plants in east Asia. Nature *413*, 129–130.

Hedrick, P.W. (2007). Balancing selection. Curr. Biol. *17*, R230–R231.

Hoekstra, H.E., and Coyne, J.A. (2007). The locus of evolution: evo devo and the genetics of adaptation. Evolution *61*, 995–1016.

Hu, J., Fan, J.P., Sun, Z.Y., and Liu, S.L. (2020). NextPolish: a fast and efficient genome polishing tool for long-read assembly. Bioinformatics *36*, 2253–2255.

Hudson, R.R., Kreitman, M., and Aguadé, M. (1987). A test of neutral molecular evolution based on nucleotide data. Genetics *116*, 153–159.

Ito, K., Katsuma, S., Kuwazaki, S., Jouraku, A., Fujimoto, T., Sahara, K., Yasukochi, Y., Yamamoto, K., Tabunoki, H., Yokoyama, T., et al. (2016). Mapping and recombination analysis of two moth colour mutations, Black moth and Wild wing spot, in the silkworm *Bombyx mori*. Heredity *116*, 52–59.

Jay, P., Chouteau, M., Whibley, A., Bastide, H., Parrinello, H., Llaurens, V., and Joron, M. (2021). Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. Nat. Genet. *53*, 288–293.

Jiggins, C.D., Wallbank, R.W., and Hanly, J.J. (2017). Waiting in the wings: what can we learn about gene co-option from the diversification of butterfly wing patterns? Philos. Trans. R. Soc. Lond. B Biol. Sci. *372*, 20150485.

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., et al. (2011). Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature *477*, 203–206.

Joron, M., and Mallet, J.L. (1998). Diversity in mimicry: paradox or paradigm? Trends Ecol. Evol. *13*, 461–466.

Kamalanathan, V., and Mohanraj, P. (2012). The life cycle and immature stages of *Kallima albofasciata*, the endemic Oakleaf, in the Andaman Islands (Indian Ocean, Bay of Bengal). J. Insect Sci. *12*, 66.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649.

Keightley, P.D., Pinharanda, A., Ness, R.W., Simpson, F., Dasmahapatra, K.K., Mallet, J., Davey, J.W., and Jiggins, C.D. (2015). Estimation of the spontaneous mutation rate in *Heliconius melpomene*. Mol. Biol. Evol. *32*, 239–243.

Komárek, S. (2003). Mimicry, Aposematism and Related Phenomena: Mimetism in Nature and the History of its Study (Lincom Europa).

Kosakovsky Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., and Frost, S.D.W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. *23*, 1891–1901.

Kratochwil, C.F., Liang, Y., Gerwin, J., Woltering, J.M., Urban, S., Henning, F., Machado-Schiaffino, G., Hulsey, C.D., and Meyer, A. (2018). Agouti-related peptide 2 facilitates convergent evolution of stripe patterns across cichlid fish radiations. Science *362*, 457–460.

Kruse, K., Hug, C.B., Hernández-Rodríguez, B., and Vaquerizas, J.M. (2016). TADtool: visual parameter identification for TAD-calling algorithms. Bioinformatics *32*, 3190–3192.

Küppers, P.V. (2015a). Butterflies of the world, Part 44: Nymphalidae XXV: *Kallima* (Verlag Goecke & Evers).

Küppers, P.V. (2015b). In Butterflies of the world, Supplement 25: The Leaf Butterflies of the Genus Kallima Doubleday, *Vol 1849* (Verlag Goecke & Evers).

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laughon, A., and Scott, M.P. (1984). Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature *310*, 25–31.

Leigh, J.W., and Bryant, D. (2015). POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. *6*, 1110–1116.

Letunic, I., and Bork, P. (2011). Interactive Tree of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. *39*, W475–W478.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinform. *12*, 323.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature *475*, 493–496.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, J.J., and Fang, X.M. (1999). Uplift of the Tibetan Plateau and environmental changes. Chin. Sci. Bull. *44*, 2117–2124.

Li, X., Fan, D., Zhang, W., Liu, G., Zhang, L., Zhao, L., Fang, X., Chen, L., Dong, Y., Chen, Y., et al. (2015). Outbred genome sequencing and CRISPR/Cas9 gene editing in butterflies. Nat. Commun. *6*, 8212.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289–293.

Livraghi, L., Hanly, J.J., Van Bellghem, S.M., Montejo-Kovacevich, G., van der Heijden, E.S., Loh, L.S., Ren, A., Warren, I.A., Lewis, J.J., Concha, C., et al.

(2021). *Cortex* cis-regulatory switches establish scale colour identity and pattern diversity in *Heliconius*. eLife *10*, e68549.

Lohman, D.J., de Bruyn, M., Page, T., von Rintelen, K., Hall, R., Ng, P.K.L., Shih, H.T., Carvalho, G.R., and von Rintelen, T. (2011). Biogeography of the Indo-Australian archipelago. Annu. Rev. Ecol. Evol. Syst. *42*, 205–226.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., et al. (2015). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. Nat. Methods *12*, 71–78.

Mallet, J., and Joron, M. (1999). Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. Annu. Rev. Ecol. Syst. *30*, 201–233.

Martin, A., Papa, R., Nadeau, N.J., Hill, R.I., Counterman, B.A., Halder, G., Jiggins, C.D., Kronforst, M.R., Long, A.D., McMillan, W.O., et al. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a *Wnt* ligand. Proc. Natl. Acad. Sci. USA *109*, 12632–12637.

Matsuoka, Y., and Monteiro, A. (2021). Hox genes are essential for the development of eyespots in *Bicyclus anynana* butterflies. Genetics *217*, 1–9.

Mazo-Vargas, A., Concha, C., Livraghi, L., Massardo, D., Wallbank, R.W.R., Zhang, L., Papador, J.D., Martinez-Najera, D., Jiggins, C.D., Kronforst, M.R., et al. (2017). Macroevolutionary shifts of *WntA* function potentiate butterfly wing-pattern diversity. Proc. Natl. Acad. Sci. USA *114*, 10701–10706.

McGinnis, W., Garber, R.L., Wirz, J., Kuroiwa, A., and Gehring, W.J. (1984). A homologous protein-coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. Cell *37*, 403–408.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.

Morgan, T.H. (1903). Evolution and Adaptation (MacMillan).

Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A., and Kent, J. (2000). Biodiversity hotspots for conservation priorities. Nature *403*, 853–858.

Nadeau, N.J., Pardo-Diaz, C., Whibley, A., Supple, M.A., Saenko, S.V., Wallbank, R.W., Wu, G.C., Maroja, L., Ferguson, L., Hanly, J.J., et al. (2016). The gene *cortex* controls mimicry and crypsis in butterflies and moths. Nature *534*, 106–110.

Nagano, T., Lubling, Y., Yaffe, E., Wingett, S.W., Dean, W., Tanay, A., and Fraser, P. (2015). Single-cell Hi-C for genome-wide detection of chromatin interactions that occur simultaneously in a single cell. Nat. Protoc. *10*, 1986–2003.

Nakamura, N. (2014). Distribution of *Kallima inachus* (Doyere,[1840]) and related species (Lepidoptera, Nymphalidae) in Indochina and adjacent regions with status alteration of *Kallima inachus alicia* Joicey & Talbot, 1921. Butterflies (Teinopalpus) *66*, 22–39.

Nishikawa, H., Iga, M., Yamaguchi, J., Saito, K., Kataoka, H., Suzuki, Y., Sugano, S., and Fujiwara, H. (2013). Molecular basis of wing coloration in a Batesian mimic butterfly, *Papilio polytes*. Sci. Rep. *3*, 3184.

Oliver, J.C., Tong, X.L., Gall, L.F., Piel, W.H., and Monteiro, A. (2012). A single origin for nymphalid butterfly eyespots followed by widespread loss of associated gene expression. PLoS Genet. *8*, e1002893.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., et al. (2001). Terrestrial ecoregions of the worlds: a new map of life on Earth. BioScience *51*, 933–938.

Phillips, S.J., Anderson, R.P., and Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. Ecol. Modell. *190*, 231–259.

Pickett, E.J., Harrison, S.P., Hope, G., Harle, K., Dodson, J.R., Peter Kershaw, A., Colin Prentice, I., Backhouse, J., Colhoun, E.A., D'Costa, D., et al. (2004). Pollen-based reconstructions of biome distributions for Australia, Southeast Asia and the Pacific (SEAPAC region) at 0, 6000 and 18,000 14C yr BP. J. Biogeogr. *31*, 1381–1444.

Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLOS Genet. *8*, e1002967.

Prentice, I.C., Jolly, D., Afanas'eva, N.B., Ager, T.A., Anderson, K., Anderson, P.M., Andrieu, V., Andreev, A.A., Ballouche, A., Bartlein, P.J., et al. (2000). Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. J. Biogeogr. *27*, 507–519.

Punnett, R.C. (1915). Mimicry in Butterflies (University of Cambridge).

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

Rahbek, C., Borregaard, M.K., Antonelli, A., Colwell, R.K., Holt, B.G., Nogues-Bravo, D., Rasmussen, C.M.Ø., Richardson, K., Rosing, M.T., Whittaker, R.J., et al. (2019a). Building mountain biodiversity: geological and evolutionary processes. Science *365*, 1114–1119.

Rahbek, C., Borregaard, M.K., Colwell, R.K., Dalsgaard, B., Holt, B.G., Morueta-Holme, N., Nogues-Bravo, D., Whittaker, R.J., and Fjeldså, J. (2019b). Humboldt's enigma: what causes global patterns of mountain biodiversity? Science *365*, 1108–1113.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. *67*, 901–904.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Ray, N., and Adams, J.M. (2001). A GIS-based vegetation map of the world at the Last Glacial Maximum (25,000–15,000 BP). Internet Archaeol. *11*. https://doi.org/10.11141/ia.11.2.

Reed, R.D., and Serfas, M.S. (2004). Butterfly wing pattern evolution is associated with changes in a Notch/Distal-less temporal pattern formation process. Curr. Biol. *14*, 1159–1166.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Royden, L.H., Burchfiel, B.C., and van der Hilst, R.D. (2008). The geological evolution of the Tibetan Plateau. Science *321*, 1054–1058.

Ruttenberg, D.M., VanKuren, N.W., Nallu, S., Yen, S.H., Peggie, D., Lohman, D.J., and Kronforst, M.R. (2021). The evolution and genetics of sexually dimorphic 'dual' mimicry in the butterfly *Elymnias hypermnestra*. Proc. Biol. Sci. *288*, 20202192.

Scott, M.P., and Weiner, A.J. (1984). Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of *Drosophila*. Proc. Natl. Acad. Sci. USA *81*, 4115–4119.

Shin, J., Blay, S., Graham, J., and McNeney, B. (2006). LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. J. Stat. Softw. *16*, 1–10.

Shirozu, T., and Nakanishi, A. (1984). A revision of the genus *Kallima* Doubleday (Lepidoptera, Nymphalidae): I. Generic classification. Lepid. Sci. *34*, 97–110.

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics *22*, 2688–2690.

Suffert, F. (1927). Zur vergleichende Analyse der schmetterlingszeichnung. Biol. Zbl *47*, 385–413.

Suzuki, T.K., Tomita, S., and Sezutsu, H. (2014). Gradual and contingent evolutionary emergence of leaf mimicry in butterfly wing patterns. BMC Evol. Biol. *14*, 229.

Suzuki, T.K., Tomita, S., and Sezutsu, H. (2019). Multicomponent structures in camouflage and mimicry in butterfly wing patterns. J. Morphol. *280*, 149–166.

Tang, Y., Zhou, C., Chen, X., and Zheng, H. (2013). Foraging behavior of the dead leaf butterfly, *Kallima inachus*. J. Insect Sci. *13*, 58.

Tendolkar, A., Pomerantz, A.F., Heryanto, C., Shirk, P.D., Patel, N.H., and Martin, A. (2021). *Ultrabithorax* is a micromanager of hindwing identity in butterflies and moths. Front. Ecol. Evol. *9*, 161.

Teng, D.Q., Li, F.Y., and Zhang, W. (2021). Using comprehensive machine-learning models to classify complex morphological characters. Ecol. Evol. *11*, 10421–10431.

Terhorst, J., Kamm, J.A., and Song, Y.S. (2017). Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat. Genet. *49*, 303–309.

Tsukada, E. (1985). Butterflies of the South East Asian Islands. In Nymphalidae (1), *Vol 4* (Plapac Co., Ltd.).

Van Belleghem, S.M., Rastas, P., Papanicolaou, A., Martin, S.H., Arias, C.F., Supple, M.A., Hanly, J.J., Mallet, J., Lewis, J.J., Hines, H.M., et al. (2017). Complex modular architecture around a simple toolkit of wing pattern genes. Nat. Ecol. Evol. *1*, 52.

van der Burg, K.R.L., Lewis, J.J., Brack, B.J., Fandino, R.A., Mazo-Vargas, A., and Reed, R.D. (2020). Genomic architecture of a genetically assimilated seasonal color pattern. Science *370*, 721–725.

van der Burg, K.R.L., Lewis, J.J., Martin, A., Nijhout, H.F., Danko, C.G., and Reed, R.D. (2019). Contrasting roles of transcription factors spineless and EcR in the highly dynamic chromatin landscape of butterfly wing metamorphosis. Cell Rep. *27*, 1027–1038.e3.

VanKuren, N.W., Massardo, D., Nallu, S., and Kronforst, M.R. (2019). Butterfly mimicry polymorphisms highlight phylogenetic limits of gene reuse in the evolution of diverse adaptations. Mol. Biol. Evol. *36*, 2842–2853.

Van't Hof, A.E., Campagne, P., Rigden, D.J., Yung, C.J., Lingley, J., Quail, M.A., Hall, N., Darby, A.C., and Saccheri, I.J. (2016). The industrial melanism mutation in British peppered moths is a transposable element. Nature *534*, 102–105.

Villanea, F.A., Safi, K.N., and Busch, J.W. (2015). A general model of negative frequency dependent selection explains global patterns of human ABO polymorphism. PLoS One *10*, e0125003.

Voris, H.K. (2000). Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and time durations. J. Biogeogr. *27*, 1153–1167.

Wallace, A.R. (1889). Darwinism: an Exploitation of the Theory of Natural Selection with Some of its Applications (MacMillan).

Wan, W.T., Dong, Z.W., Ren, Y.D., Yang, J., Pan, X.Y., He, J.W., Chang, Z., Liu, W., Liu, G.C., Zhao, R.P., et al. (2021). Chromatin accessibility profiling provides insights into larval cuticle color and adult longevity in butterflies. Zool. Res. *42*, 614–619.

Watson, D.M.S. (1936). A discussion on the present state of theory of natural selection. Proc. R. Soc. Lond. B *121*, 43–73.

Weaver, S., Shank, S.D., Spielman, S.J., Li, M., Muse, S.V., and Kosakovsky Pond, S.L. (2018). DataMonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Mol. Biol. Evol. *35*, 773–777.

Wolff, J., Rabbani, L., Gilsbach, R., Richard, G., Manke, T., Backofen, R., and Grüning, B.A. (2020). Galaxy HiCExplorer 3: a web server for reproducible Hi-C, capture Hi-C and single-cell Hi-C data analysis, quality control and visualization. Nucleic Acids Res. *48*, W177–W184.

Woodruff, D.S. (2010). Biogeography and conservation in Southeast Asia: how 2.7 million years of repeated environmental fluctuations affect today's patterns and the future of the remaining refugial-phase biodiversity. Biodivers. Conserv. *19*, 919–941.

Wu, Q., Han, T.S., Chen, X., Chen, J.F., Zou, Y.P., Li, Z.W., Xu, Y.C., and Guo, Y.L. (2017). Long-term balancing selection contributes to adaptation in *Arabidopsis* and its relatives. Genome Biol. *18*, 217.

Xie, S., Shen, B., Zhang, C., Huang, X., and Zhang, Y. (2014). sgRNAcas9: a software package for designing CRISPR sgRNA and evaluating potential off-target cleavage sites. PLoS One *9*, e100448.

Yang, J., Wan, W., Xie, M., Mao, J., Dong, Z., Lu, S., He, J., Xie, F., Liu, G., Dai, X., et al. (2020). Chromosome-level reference genome assembly and gene editing of the dead-leaf butterfly *Kallima inachus*. Mol. Ecol. Resour. *20*, 1080–1092.

Yoda, S., Otaguro, E., Nobuta, M., and Fujiwara, H. (2020). Molecular mechanisms underlying pupal protective color switch in *Papilio polytes* butterflies. Front. Ecol. Evol. *8*, 1–14.

Zhang, J., Nielsen, R., and Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. *22*, 2472–2479.

Zhang, L., and Reed, R.D. (2017). A practical guide to CRISPR/Cas9 genome editing in Lepidoptera. In Diversity and evolution of butterfly Wing patterns (Springer), pp. 155–172.

Zhang, W., Dasmahapatra, K.K., Mallet, J., Moreira, G.R., and Kronforst, M.R. (2016). Genome-wide introgression among distantly related *Heliconius* butterfly species. Genome Biol. *17*, 25.

Zhang, W., Westerman, E., Nitzany, E., Palmer, S., and Kronforst, M.R. (2017). Tracing the origin and evolution of supergene mimicry in butterflies. Nat. Commun. *8*, 1269.

Zheng, H.B., Powell, C.M., An, Z.S., Zhou, J., and Dong, G.R. (2000). Pliocene uplift of the northern Tibetan Plateau. Geology *28*, 715–718.

Zhou, C., Chen, X., and He, R. (2013). COII phylogeography reveals surprising divergencies within the cryptic butterfly *Kallima inachus* (Doyére, 1840) (Lepidoptera: Nymphalidae: Kallimini) in southeastern Asia. Pan-Pac. Entomol. *88*, 381–398.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-digoxigenin alkaline phosphatase, Fab fragments	Roche Applied Science	11093274910, RRID: AB_2734716
Biological samples		
Adult butterfly thoracic tissues, see Table S1	This paper	N/A
Pupal butterfly wing disc tissues, see Table S1	This paper	N/A
Pupal butterfly body tissues, See Table S1	This paper	N/A
Larval butterfly wing disc tissues	This paper	N/A
Chemicals, peptides, and recombinant proteins		
RNA*later*^{TM} Stabilization Solution	Thermo Fisher Scientific	AM7020
TRIzol^{TM} Reagent	Invitrogen	15596026
EnGen® Spy Cas9 NLS	New England Biolabs	M0646M
Digoxigenin-11-UTP	Roche Applied Science	11209256910
BM Purple AP Substrate	Roche Applied Science	11442074001
Critical commercial assays		
KAPA Hyper Prep Kit	Roche Applied Science	KK8504
Ligation Sequencing Kit	Oxford Nanopore Technologies	SQK-LSK109
KAPA RNA Hyper Library Preparation Kit	Roche Applied Science	KK8560
TURBO DNA-free™ Kit	Invitrogen	AM1907
HiScript cDNA Synthesis Kit	Vazyme Biotech	R211-01
TB Green *Premix Ex Taq*^{TM} II (Tli RNaseH Plus)	Takara Bio	RR820Q
RiboMAX™ Large Scale RNA Production Systems-SP6 and T7	Promega	P1280, P1300
MAXIscript T7 kit	Life Technology	AM1314
TransDirect® Animal Tissue PCR Kit	Trans	AD201-01
HiPure Insect DNA Kit	Magen	D3129-03
Deposited data		
Illumina paired-end raw reads	This paper; NCBI Sequence Archive (SRA)	PRJNA698415
Illumina paired-end raw reads	This paper; NCBI Sequence Archive (SRA)	PRJNA698433
Illumina paired-end raw reads	This paper; NCBI Sequence Archive (SRA)	PRJNA533504
de novo genome assemblies	This paper; Dryad data	https://doi.org/10.5061/dryad.3xsj3txhj
Experimental models: Organisms/strains		
Kallima inachus (lab reared)	Leshan Shawan Xiong's Butterfly Garden	N/A
Oligonucleotides		
Forward primer for *cortex*: AGTGGGGAGAGGGTCGTAG	This paper	N/A
Reverse primer for *cortex*: CCTCGTCCAAGTAAACGGGAA	This paper	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Forward primer for *ef1α*: GCTGAGCGTGAACGTGGTAT	This paper	N/A
Reverse primer for *ef1α*: GACGATGAGCACGGCGCAAT	This paper	N/A
Forward primer for the sgRNA target region: CTATTTGGCATAAGCAGGT	This paper	N/A
Reverse primer for the sgRNA target region: GATTCTAGTTGTCTCACTCCA	This paper	N/A
Recombinant DNA		
pEASY®-T1 Cloning Vector	TransGen Biotech	CT101-01
T-Vector pMD19	Takara	3271
Software and algorithms		
Trimmomatic v0.38	Bolger et al., 2014	http://www.usadellab.org/cms/?page=trimmomatic
Bowtie2-2.3.4.3	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
SAMtools v1.9	Li et al., 2009	http://www.htslib.org/
Picard v1.96	Broad Institute	http://broadinstitute.github.io/picard/
GATK v3.7	McKenna et al., 2010	https://gatk.broadinstitute.org
RaxML	Stamatakis, 2006	https://cme.h-its.org/exelixis/web/software/raxml/index.html
iTOL	Letunic and Bork, 2011	https://itol.embl.de/
VCFtools	Danecek et al., 2011	https://vcftools.github.io/index.html
LASTZ 1.04.03	Harris, 2007	http://www.bx.psu.edu/miller_lab/
Multiz v11.2	Blanchette et al., 2004	http://www.bx.psu.edu/miller_lab/
PhyTime	Guindon, 2010	https://www.stat.auckland.ac.nz/~guindon/phytime/
TCS Networks in PopART	Clement et al., 2000; Leigh and Bryant, 2015	http://popart.otago.ac.nz/
PLINK v1.90b6.7	Purcell et al., 2007	http://zzz.bwh.harvard.edu/plink/
ADMIXTURE	Alexander et al., 2009	http://dalexander.github.io/admixture/index.html
Haploview v4.2	Barrett et al., 2005	https://www.broadinstitute.org/haploview/haploview
Ldheatmap	Shin et al., 2006	https://sfustatgen.github.io/LDheatmap/index.html
IGV v2.5.2	Robinson et al., 2011	https://software.broadinstitute.org/software/igv/
GARD on the Datamonkey webserver	Pond et al., 2006; Weaver et al., 2018	http://www.datamonkey.org/GARD/
PAML	Zhang et al., 2005	http://abacus.gene.ucl.ac.uk/software/paml.html
SVMorph	Teng et al., 2021	https://github.com/TDQ233/SVMorph
PSMC	Li and Durbin, 2011	https://github.com/lh3/psmc
SMC++	Terhorst et al., 2017	https://github.com/popgenmethods/smcpp
G-PhoCS	Gronau et al., 2011	http://compgen.cshl.edu/GPhoCS/
Treemix	Pickrell and Pritchard, 2012	https://bitbucket.org/nygcresearch/treemix/wiki/Home
Maxent	Phillips et al., 2006	http://biodiversityinformatics.amnh.org/open_source/maxent/
SDMtoolbox v2.4	Brown et al., 2017	http://www.sdmtoolbox.org/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
ArcMap 10.7	Esri	https://desktop.arcgis.com/en/arcmap/
Nanofit v2.3.0	De Coster et al., 2018	https://github.com/wdecoster/nanofilt
NextDenovo v2.2-beta.0	Nextomics	https://github.com/Nextomics/NextDenovo
NextPolish v1.1.0	Hu et al., 2020	https://github.com/Nextomics/NextPolish
Geneious v9.1.4	Kearse et al., 2012	https://www.geneious.com/
progressiveMauve	Darling et al., 2010	http://darlinglab.org/mauve/user-guide/progressivemauve.html
BWA	Li and Durbin, 2009	http://bio-bwa.sourceforge.net/
Juicer	Durand et al., 2016b	https://github.com/aidenlab
Juicebox	Durand et al., 2016a	https://github.com/aidenlab
HiCExplorer	Wolff et al., 2020	https://github.com/deeptools/HiCExplorer
TADtool	Kruse et al., 2016	https://github.com/vaquerizaslab
BLASTX	NCBI	https://blast.ncbi.nlm.nih.gov/Blast.cgi
sgRNAcas9 v3.0	Xie et al., 2014	https://sourceforge.net/projects/sgrnacas9/
STAR v2.7.1a	Dobin et al., 2013	https://github.com/alexdobin/STAR
RSEM v1.3.1	Li and Dewey, 2011	https://deweylab.github.io/RSEM
DESeq2	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html
Trinity v2.8.5	Grabherr et al., 2011	https://github.com/trinityrnaseq/trinityrnaseq
RepeatModeler v1.0.11	Robert Hubley & Arian Smit	http://www.repeatmasker.org/RepeatModeler/
RepeatMasker v4.1.1	Smit, AFA, Hubley, R & Green, P.	http://repeatmasker.org/
Custom codes	This paper	https://github.com/TDQ233/Kallima-inachus; https://github.com/wstree/kallima-project
Other		
Nikon SMZ18 stereomicroscope	Nikon	N/A
HP LaserJet Pro M227fdw scanner	HP	N/A
Nikon DS-Ri2 digital camera	Nikon	N/A
Canon 70D digital camera	Canon	N/A
Illumina HiSeq Xten platform	Illumina	N/A
Nanopore promethION platform	Oxford Nanopore Technologies	N/A
Roche LightCycler 96 detection system	Roche Applied Science	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Wei Zhang (weizhangvv@pku.edu.cn).

Materials availability

All materials generated in this study are available from the lead contact upon request.

Data and code availability

Sequence data are available from the NCBI Short-Read Archive (SRA) database (accession numbers PRJNA698415, PRJNA698433, and PRJNA533504). The *de novo* genome assemblies are available on Dryad (https://doi.org/10.5061/dryad.3xsj3txhj). The custom codes are available on GitHub (https://github.com/TDQ233/Kallima-inachus, https://github.com/wstree/kallima-project).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Butterfly rearing and collection

Livestock to start colonies of lab-reared *K. inachus* butterflies was supplied by Leshan Shawan Xiong's Butterfly Garden (Leshan, China), which has a breeding license. Colonies were maintained at 26–28°C under approximately 70% relative humidity and a 14:10 light:dark cycle. Adults were fed fruit juice or a mixture of honey and alcohol. Larvae of *K. inachus* were raised on the host plant *Strobilanthes cusia* (Acanthaceae). Female and male individuals were selected to mate at a ratio of 1:1 with access to sufficient food and host plants for oviposition. Butterfly samples were collected in the field with the permission and assistance of Medog Forestry Bureau (China), Medog National Natural Reserve (China), Ya'an Forestry Bureau (China), Fengtongzhai National Natural Reserve (China), Wuzhishan National Natural Reserve (China), and Ruili Forestry Bureau (China) or supplied by Leshan Shawan Xiong's Butterfly Garden (Leshan, China), Changde Xiaoyejia Arts & Crafts Co. Ltd. (Changde, China), Wuhan Diezhongdie Creative Culture Co. Ltd. (Wuhan, China), and Shanghai Qiuyu Biotechnology Co. Ltd. (Shanghai, China) with breeding, collecting or import licenses.

METHOD DETAILS

Sample preparation and sequencing

For Illumina sequencing, genomic DNA from adult butterfly thoracic tissues was isolated using a phenol–chloroform DNA extraction protocol. Illumina paired-end libraries were constructed using the KAPA Hyper Prep Kit and sequenced on the Illumina HiSeq Xten platform at Novogene. Raw reads were demultiplexed using sample-specific barcodes. For Nanopore sequencing, genomic DNA was isolated from adult butterfly thoracic tissues from individuals representing the *veined*, *scrambled*, *rippled*, and *moldy* wing forms. The long-read Nanopore sequencing library was prepared using a SQK-LSK109 Ligation Sequencing Kit and sequenced on the Nanopore promethION platform at GrandOmics. For Hi-C sequencing, genomic DNA was isolated from adult butterfly thoracic tissue from a homozygous *scrambled* individual. Library preparation was performed principally following previously described methodologies (Belton et al., 2012; Lieberman-Aiden et al., 2009; Ma et al., 2015; Nagano et al., 2015), and the library was sequenced on the Illumina HiSeq Xten platform at Novogene. Additional Hi-C data from several different individuals were downloaded from the NCBI Short-Read Archive (SRA) database (accession number PRJNA533504). For RNA-seq and RT–qPCR sampling, individuals with different wing forms were collected in the late fifth-instar larval stage, the prepupal stage, 24 h after pupation, and 72 h after pupation. For each developmental stage, we collected four individuals with the *PP* genotype and four individuals with the *VP* genotype. Wing discs were dissected in PBS and stored in RNA*later*^{TM} Stabilization Solution at -80°C. RNA was extracted using TRIzol^{TM} Reagent, and the Illumina paired-end libraries were constructed using the KAPA RNA Hyper Library Preparation Kit and sequenced on the Illumina HiSeq Xten platform at Novogene.

Genotype calling

Low-quality Illumina sequencing data were filtered from raw reads using Trimmomatic v0.38 (Bolger et al., 2014), and the filtered reads were mapped to the reference genome of *K. inachus* (Yang et al., 2020) using Bowtie2-2.3.4.3 (Langmead and Salzberg, 2012) with the parameter–very-sensitive-local. The aligned results were converted from SAM format to BAM format using SAMtools v1.9 (Li et al., 2009), and were reordered, sorted, and duplicate marked using Picard-tools v1.96 (http://broadinstitute.github.io/picard/). RealignerTargetCreator and IndelRealigner (McKenna et al., 2010) in GATK 3.7 were used to realign indels, and UnifiedGenotyper6 (McKenna et al., 2010) in GATK3.7 was used to call SNPs according to the following parameters: heterozygosity 0.05, stand_call_conf 50.0, stand_emit_conf 10.0, and dcov 250. We filtered the sites for downstream analyses according to the parameters maf 0.05 and Qual 50 or Qual 30 using VCFtools v0.1.17 (Danecek et al., 2011).

Phylogenetic analyses

We constructed maximum-likelihood phylogenetic trees using RAxML (Stamatakis, 2006) with the GTRGAMMA model and 100 bootstrap replicates based on genome-wide SNPs, coding sequence (CDS) regions, *de novo*-assembled *cortex* haplotypes, and phased *cortex* CDS regions. For the tree of CDS regions, concatenated CDSs of each individual were generated and extracted using FastaAlternateReferenceMaker (McKenna et al., 2010) in GATK 3.7. For the tree of *cortex* haplotypes, the outgroup *cortex* haplotype of *Junonia coenia* was extracted from the reference genome of *J. coenia* v1.0 (van der Burg et al., 2019), and all the *cortex* haplotypes were aligned using LASTZ 1.04.03 (Harris, 2007) with the *plain* haplotype as a reference according to the parameter –format=maf –chain. Subsequently, pairwise alignments were merged into a multiple sequence alignment using the roast script of Multiz v11.2 (Blanchette et al., 2004) with a guided topology (((*M*, *P*), (*S*, (*R*, *V*))), *Jco*). The multiple sequence alignment was transformed into FASTA format using the maf2fasta script of Multiz v11.2 (Blanchette et al., 2004). The *cortex* haplotype alignment was divided into partitioned sequence alignments in a 50-kb sliding window, and the 50 kb alignment that included the *cortex* gene and displayed strong linkage disequilibrium was used to construct the tree of *cortex* haplotypes. We also estimated the divergence times of the genome-wide and *cortex* haplotype-based phylogenies using PhyTime (Guindon, 2010), with the nodes calibrated according to the divergence times between *Kallima* and *Junonia* (between 31.0033 and 44.768 Mya), *Kallimoides* and *Junonia* (between 27.8345 and 40.7399 Mya), and *Historis* and *Kallima* (between 37.827 and 54.7998 Mya) (Chazot et al., 2019). The PhyTime results were further processed using TreeAnnotator (Drummond et al., 2012). All tree images were generated using iTOL (Letunic and Bork, 2011). We also double-checked phased sequences by PCR using the genotyping primers (see below in the CRISPR/Cas9-mediated genome editing section) and constructed the *cortex* haplotype network using the TCS method implemented in PopART (Clement et al., 2000; Leigh and Bryant, 2015).

Population structure inference

We converted VCF files to BED files using VCFtools v0.1.17 (Danecek et al., 2011) and PLINK v1.90b6.7 (Purcell et al., 2007), and filtered the data according to the parameters geno 0.20 and maf 0.05 in PLINK v1.90b6.7 (Purcell et al., 2007). Principal component analysis (PCA) was performed using PLINK v1.90b6.7 (Purcell et al., 2007). Proportions of individual ancestry were inferred using ADMIXTURE (Alexander et al., 2009) by defining genetic clusters from $K = 2$ to $K = 10$.

Population genetic analyses

The genome-wide patterns of genetic differentiation (F_{ST} values) and nucleotide diversity (π values) were calculated using VCFtools v0.1.17 (Danecek et al., 2011) with a block size of 50 kb, and the standard error was calculated using a jackknife approach. For local regions, a block size of 500 bp was used, and the standard error was calculated using a moving block bootstrap approach (Zhang et al., 2016). Pairwise linkage disequilibrium (LD) was estimated by calculating the square of the correlation coefficient (r^2) using VCFtools v0.1.17 (Danecek et al., 2011) according to the parameter –plink –thin 200. LD heatmaps were calculated using Haploview v4.2 (Barrett et al., 2005) according to the parameter minMAF 0.10 and were visualized using LDheatmap (Shin et al., 2006). Prior to the detection of positive selection, we extracted *cortex* CDS regions by visualizing BAM files in IGV v2.5.2 (Robinson et al., 2011), and the raw reads involving *cortex* CDS regions were extracted and employed to manually assemble the complete CDS using Geneious v9.1.4 (Kearse et al., 2012). Subsequently, we detected recombination blocks in the *cortex* CDS alignment using GARD (Pond et al., 2006) on the Datamonkey webserver (Weaver et al., 2018), which yielded three putative recombination blocks. Positive selection signals in specific lineages were detected with partitioned sequence alignments using the branch-site model implemented in the CodeML program within PAML (Zhang et al., 2005). We performed HKA tests (Hudson et al., 1987) to compare the level of *cortex* polymorphism to the polymorphism of neighboring genes using sequence data extracted from the reference genome of *Vanessa tameamea* v1.0 (GenBank: PEHJ00000000.1) for the calculation of interspecific divergence. Genome-wide association studies were performed for individuals presenting different forms. The different forms were identified and classified using SVMorph (Teng et al., 2021). The VCF file was first compressed using VCFtools v0.1.17 (Danecek et al., 2011) according to the parameter –thin 200, and genome-wide association analysis was then performed using PLINK v1.90b6.7 (Purcell et al., 2007) according to the following parameters: –assoc fisher –allow-no-sex –adjust –allow-extra-chr. The false-discovery rate corrected *P* values were calculated using PLINK v1.90b6.7 (Purcell et al., 2007). Genome-wide F_{ST} values were also calculated between individuals with different forms in a block size of 50 kb using VCFtools v0.1.17 (Danecek et al., 2011), and the F_{ST} values were standardized into *Z* scores.

Demographic analyses

We used the pairwise sequentially Markovian coalescent (PSMC) model (Li and Durbin, 2011) and SMC++ (Terhorst et al., 2017) to infer historical effective population sizes. For PSMC analyses, we generated the diploid consensus sequence of each sample using SAMtools (Li et al., 2009) and masked bases with a depth lower than a third of the average depth or greater than twice the average depth. The parameters were set as follows: -N25 -t15 -r5 -p "4+25*2+4+6". For SMC++ analyses, we employed population data to infer effective population sizes in recent years and applied "-c 100000" to mask the uncalled region. For all the PSMC and SMC++ analyses, we assumed an average mutation rate of $\mu = 3 \times 10^{-9}$ (Keightley et al., 2015) and an average generation time of $g = 0.5$ years (Yang et al., 2020).

Demographic parameters, such as population sizes, divergence times, and migration rates, were inferred using the Generalized Phylogenetic Coalescent Sampler (G-PhoCS ver1.2.3) software (Gronau et al., 2011) with the default settings for a gamma distribution. To do so, we chose individuals of each population with good sequencing coverage and performed filtering steps for the reference genome according to Table S3, which yielded 5,638 putative neutral loci in total. Considering the genetic distance in the genus *Kallima*, we performed three analyses, focusing on populations and subspecies in *K. inachus*, populations and subspecies in *K. alicia*, and populations in *K. incognita* and *K. paralekta*. We estimated possible migration bands twice in multiple independent tests and calibrated the raw estimates by assuming an average mutation rate of $\mu = 3 \times 10^{-9}$ (Keightley et al., 2015) and an average generation time of $g = 0.5$ year (Yang et al., 2020). The lower migration rates above 0.003 in two replicates were considered significant, and these significant migration bands were tested together in full-model analyses. To determine the number of MCMC iterations, we checked the convergence by visually monitoring the likelihood plots in Tracer v1.7 (Rambaut et al., 2018) to ensure that the MCMC chains reached an equilibrium distribution. We conducted G-PhoCS analysis with 200,000 MCMC iterations. Migration events were also inferred using TreeMix (Pickrell and Pritchard, 2012) by assuming zero to eight migration edges.

Habitat suitability modeling

We used the maximum entropy model (Maxent) (Phillips et al., 2006) to evaluate habitat suitability for *Kallima* butterflies. We downloaded the presence data of *K. inachus* from the Global Biodiversity Information Facility (GBIF) (https://www.gbif.org/), to which we added our own sample data. The presence data were spatially rarified using SDMtoolbox v2.4 (Brown et al., 2017). We downloaded 19 bioclimatic variables of the 2.5 arc-minute scale (CCSM4) from WorldClim (http://www.worldclim.org/) for the projection of the

LGM and mid-Holocene periods, and we downscaled the original LIG 30 arc-second data to the 2.5 arc-minute resolution scale to synchronize the resolution. Subsequently, we masked these climate layers to the area that was reasonably surveyed and removed highly correlated variables with a Pearson's correlation coefficient greater than or equal to 0.8. Data formats were transformed using ArcMap 10.7 (https://desktop.arcgis.com/en/arcmap). Ten replicates with cross-validation were performed for each analysis. Vegetation maps and Maxent projection results were overlaid in ArcMap 10.7 (http://www.bridge.bris.ac.uk/resources/Databases/BIOMES_data; Bigelow et al., 2003; Harrison et al., 2001; Olson et al., 2001; Pickett et al., 2004; Prentice et al., 2000; Ray and Adams, 2001).

Genome assembly and synteny mapping

Low-quality Nanopore sequencing data were filtered from raw reads using Nanofit v2.3.0 (De Coster et al., 2018). *De novo* genome assembly was performed using NextDenovo v2.2-beta.0 (https://github.com/Nextomics/NextDenovo), and base errors were further corrected with Illumina sequencing data using NextPolish v1.1.0 (Hu et al., 2020). Reference-guided genome assembly of the *cortex* region of *K. inachus* was performed using Geneious v9.1.4 (Kearse et al., 2012) by combining the *de novo*-assembled *plain* haplotype and the reference genome of *K. inachus* (Yang et al., 2020). To assemble the *cortex* haplotypes of heterozygous samples, we did not perform an Illumina polishing step for Nanopore sequencing data because it could lead to the fusion of different haplotypes and introduce more mistakes relative to the base errors occurring during Nanopore sequencing. To test the reliability of our unpolished genome assemblies, we aligned Illumina sequencing data from multiple individuals to the *de novo*-assembled *cortex* haplotypes. The results were highly consistent between the *cortex* haplotypes and the Illumina sequencing reads with the corresponding genotypes. In addition, we performed haplotype alignment based on two assembled scaffolds of haplotype *P* extracted from different individuals and observed complete collinearity and a high degree of base identity between the two scaffolds (Figure S4G), suggesting that the assembled *cortex* scaffolds obtained from the heterozygous individuals were reliable. Five *cortex* haplotypes (*plain*, *veined*, *scrambled*, *rippled*, and *moldy*) were extracted from *de novo*-assembled genomes and aligned to detect rearrangement breakpoints using progressiveMauve (Darling et al., 2010) with the default parameters.

Hi-C library data analysis

Hi-C raw reads were aligned to the reference-guided genome assembly of the *scrambled* haplotype using BWA (Li and Durbin, 2009) following the Juicer pipeline (Durand et al., 2016b). Normalized Hi-C contact matrices were then generated at several resolutions in *.hic* format, and multiple tools, such as the head algorithm (Rao et al., 2014) in Juicer software juicer_tools.1.7.6, the hicFindTADs algorithm in HiCExplorer software (Wolff et al., 2020), and TADtool (Kruse et al., 2016), were employed for TAD identification. We regarded TAD regions detected by all three tools as putative TADs and manually annotated these regions using BLASTX searches against NCBI's nr protein database (https://blast.ncbi.nlm.nih.gov/Blast.cgi). For the Arrowhead algorithm, the following parameters were used: -k KR, -r 25000, –ignore_sparsity. For the hicFindTADs algorithm, Hi-C contact matrices were first converted into *.cool* format using hicConvertFormat and TADs were then called with a step size of 40,000 bp. For TADtool, Hi-C contact matrices were first converted into sparse matrix format with a resolution value of 40 kb, and the insulation score was then calculated to identify TADs. The Hi-C heatmap and putative TADs were visualized using Juicebox (Durand et al., 2016a) and hicPlotTADs in HiCExplorer software (Wolff et al., 2020).

Annotation of transposable elements and promoter regions

A transposable element library of *K. inachus* was constructed based on the reference genome assembly using RepeatModeler v1.0.11 (http://www.repeatmasker.org/RepeatModeler/). The distribution of transposable elements among different *cortex* haplotypes was calculated using RepeatMasker v4.1.1 (http://www.repeatmasker.org) with the default parameters and the TE library generated by RepeatModeler. The promoter regions were annotated based on the published results of chromatin accessibility profiling in *K. inachus* (Wan et al., 2021).

RT–qPCR analyses

Total RNA was diluted to 200 ng/μL and treated with the TURBO DNA-free™ Kit (Invitrogen) to remove the remaining genomic DNA. cDNA was generated from total RNA using the HiScript cDNA Synthesis Kit (Vazyme). RT–qPCR was performed in a 20-μL system, with 50 ng diluted cDNA and forward and reverse primers at 0.4 μM. Primers were designed to span the *cortex* intron (F: AGTGGGGAGAGGGTCGTAG, R: CCTCGTCCAAGTAAACGGGAA), and the target products were smaller than 200 bp. We also analyzed the expression of the reference gene *ef1α* for normalization (F: GCTGAGCGTGAACGTGGTAT, R: GACGATGAGCACGGCGCAAT). The reactions were run on a Roche LightCycler 96 detection system using TB Green Premix *Ex Taq* II (Tli RNaseH Plus) (Takara).

CRISPR/Cas9-mediated genome editing

We designed single-guide RNAs (sgRNAs) and evaluated their specificity using sgRNAcas9 v3.0 (Xie et al., 2014) to search genomic regions against the reference genome of *K. inachus* (Yang et al., 2020) for $GN_{19}NGG$ or $N_{20}NGG$ sequences. We performed CRISPR/Cas9 mutagenesis assays following previously described methods of CRISPR/Cas9 genome editing in Lepidoptera (Bassett et al., 2013; Li et al., 2015; Zhang and Reed, 2017). We generated sgRNAs by *in vitro* transcription using the MAXIscript T7 kit (Life

Technology, Cat No. AM1314) or RiboMAX™ Large-Scale RNA Production Systems (SP6 and T7) (Promega, Cat No. P1280, P1300). Subsequently, we performed the microinjection of sgRNAs and EnGen® Spy Cas9 NLS (NEB, Cat No. M0646M) to collected butterfly eggs between 0.5 h and 1 h or between 1 h and 2 h after oviposition. Injected eggs were placed in transparent plastic boxes until hatching and were fed host leaves (*S. cusia*) until the third-instar stage. The injected larvae were transferred to the host plant *S. cusia* until emergence. The average hatching rate with the deleterious *cortex* KOs was 30.69%. For genotyping, we extracted genomic DNA from individual wing tissues showing mutant phenotypes using the TransDirect® Animal Tissue PCR Kit (Trans, Cat No. AD201-01) or a phenol–chloroform protocol and performed PCR using genotyping primers flanking the sgRNA target region (CT-CRISPR-F: CTATTTGGCATAAGCAGGT, CT-CRISPR-R: GATTCTAGTTGTCTCACTCCA). Mosaic mutants were identified by comparing the sizes of the obtained PCR products via standard agarose gel electrophoresis and by comparing sequences via Sanger sequencing. To identify small indels in the sgRNA target regions, PCR products were cloned into the T-Vector pMD19 (Takara, Cat No. 3271) or the *pEASY* vector (TransGen Biotech) and sequenced. Wing phenotypes were photographed using a scanner (HP LaserJet Pro M227fdw) or a digital camera (Canon 70D or Nikon DS-Ri2) under a microscope (Nikon SMZ18).

Transcriptome analyses and assembly

Low-quality Illumina reads were filtered from the raw sequence data using Trimmomatic v0.38 (Bolger et al., 2014), and the filtered reads were mapped to the reference genome of *K. inachus* (Yang et al., 2020) using STAR v2.7.1a (Dobin et al., 2013). Differentially expressed genes were called using RSEM v1.3.1 (Li and Dewey, 2011) and DESeq2 (Love et al., 2014), and the results were FDR adjusted. *De novo* transcriptome assembly was performed using Trinity v2.8.5 (Grabherr et al., 2011) with default parameters.

In situ hybridization

In situ hybridization (ISH) assay for *cortex* was performed principally following a previously described procedure with at least three biological replicates (Martin et al., 2012). ISH templates were amplified by PCR using primers for *cortex* (CT-F: ATTTAGGTG ACACTATAGAATAGCCAGCTGTGATGGATCTGT, CT-R: AATTAATACGACTCACTATAGGGAGGGTCAAACTTGCAGCACT), and a T7 or SP6 promoter sequence was attached. For sense and antisense riboprobe synthesis, *in vitro* transcription was performed using RiboMAX™ Large-Scale RNA Production Systems (T7 and SP6) (Promega) supplemented with Digoxigenin-11-UTP (Roche Applied Science). The sense probes were used as the negative control, which showed no effect (Figure S6A). Given that the fifth-instar larval stage is the developmental time point associated with the expression of genes involved in the production of complex wing patterns, such as eyespots in multiple nymphalid butterflies (Oliver et al., 2012; Reed and Serfas, 2004), individuals were sampled in the fifth-instar larval stage, and their wing discs were dissected in phosphate-buffered saline (PBS). Tissues were fixed in 4% paraformaldehyde (PFA) in PBS for 30 min, washed in cold phosphate-buffered saline containing 0.01% Tween 20 (PBST), and treated with 25 μg/mL proteinase K. Wing discs were freed from their peripodial membrane and postfixed in 4% PFA for 20 min. After washing in cold PBST, wing discs were gradually transferred to standard hybridization buffer, prehybridized for 1 h at 64°C, and incubated in standard hybridization buffer containing a riboprobe at 30 ng/mL for 40 h at 64°C. For secondary detection of the riboprobe, wing discs were blocked in Tris-buffered saline containing 0.01% Tween 20 (TBST) and 1 g/L BSA and subsequently incubated with a 1:4000 dilution of anti-digoxigenin alkaline phosphatase Fab fragments (Roche) at 4°C for 3 h or overnight. Wing discs were washed once in cold BSA buffer, washed ten times in cold TBST and finally stained with BM Purple (Roche) for 4–8 h. For tissue clearing, wing discs were gradually placed in ethanol and xylene, and the redundant tissue was carefully removed. For imaging, wing discs were transferred back to PBST, mounted in PBS containing 60% glycerol, and imaged using a digital camera (Nikon DS-Ri2) under a microscope (Nikon SMZ18).

Population genetics simulations

To investigate the driving forces leading to the current frequencies of *cortex* haplotypes, we developed a population genetics model adapted from Zhang et al. (2017) and Villanea et al. (2015) by incorporating negative frequency-dependent selection (NFDS). In this model, the fitness (w) of a genotype was negatively correlated with its frequency (f), given the strength of NFDS denoted as z, and the additional beneficial or deleterious effects of a genotype were denoted as a. Our model also comprised a random sampling procedure to simulate the perturbation caused by genetic drift. Therefore, the fitness of each genotype could be calculated as follows:

$$w_{PP} = a_{PP}(1 - zf_{PP}) \quad \text{(Equation 1)}$$

$$w_{SS} = a_{SS}(1 - zf_{SV} - zf_{SP} - zf_{SS}) \quad \text{(Equation 2)}$$

$$w_{SP} = a_{SP}(1 - zf_{SV} - zf_{SP} - zf_{SS}) \quad \text{(Equation 3)}$$

$$w_{SV} = a_{SV}(1 - zf_{SV} - zf_{SP} - zf_{SS}) \quad \text{(Equation 4)}$$

$$w_{VP} = a_{VP}(1 - zf_{VV} - zf_{VP}) \quad \text{(Equation 5)}$$

$$w_{VV} = a_{VV}(1 - zf_{VV} - zf_{VP}) \quad \text{(Equation 6)}$$

$$w_{MP} = a_{MP}(1 - zf_{MP}) \quad \text{(Equation 7)}$$

$$w_{MS} = a_{MS}(1 - zf_{MS}) \quad \text{(Equation 8)}$$

$$w_{MV} = a_{MV}(1 - zf_{MV}) \quad \text{(Equation 9)}$$

$$w_{MM} = 0 \quad \text{(Equation 10)}$$

$$w_{RP} = a_{RP}(1 - zf_{RP}) \quad \text{(Equation 11)}$$

$$w_{RS} = a_{RS}(1 - zf_{RS}) \quad \text{(Equation 12)}$$

$$w_{RV} = a_{RV}(1 - zf_{RV}) \quad \text{(Equation 13)}$$

$$w_{MP} = a_{MP}(1 - zf_{MP}) \quad \text{(Equation 14)}$$

$$w_{RR} = 0 \quad \text{(Equation 15)}$$

$$w_{MR} = a_{MR}(1 - zf_{MR}) \quad \text{(Equation 16)}$$

with P, S, V, M, and R denoting five *cortex* haplotypes. Given that we did not observe individuals with MM or RR genotypes and we only obtained limited individuals with the MR genotype from lab crossings, we set the fitness of MM or RR genotypes as null. We used p, s, v, m, and r to denote the initial frequencies of haplotypes P, S, V, M, and R, respectively. Therefore, the changes in the frequencies of the five haplotypes (p', s', v', m', and r') and the mean fitness w_{mean} could be calculated as follows:

$$p' = \frac{(p^2 w_{PP} + spw_{SP} + vpw_{VP} + mpw_{MP} + rpw_{RP})}{w_{mean}} \quad \text{(Equation 17)}$$

$$s' = \frac{(s^2 w_{SS} + spw_{SP} + svw_{SV} + msw_{MS} + rsw_{RS})}{w_{mean}} \quad \text{(Equation 18)}$$

$$v' = \frac{(v^2 w_{VV} + vpw_{VP} + svw_{SV} + mvw_{MV} + rvw_{RV})}{w_{mean}} \quad \text{(Equation 19)}$$

$$m' = \frac{(mpw_{MP} + msw_{MS} + mvw_{MV} + mrw_{MR})}{w_{mean}} \quad \text{(Equation 20)}$$

$$r' = \frac{(rpm_{RP} + rsw_{RS} + rvw_{RV} + mrw_{MR})}{w_{mean}} \quad \text{(Equation 21)}$$

$$w_{mean} = \left(p^2 w_{PP} + s^2 w_{SS} + v^2 w_{VV} + 2spw_{SP} + 2vpw_{VP} + 2mpw_{MP} + 2rpw_{RP} + 2svw_{SV} + 2msw_{MS} + 2rsw_{RS} + 2mvw_{MV} + 2rvw_{RV} + 2mrw_{MR}\right) \quad \text{(Equation 22)}$$

We first performed preliminary tests including only two original haplotypes, *P* and *S*, to repeat the modeling procedure presented by Zhang et al. (2017) by testing *z* values ranging from 0 to 0.9 to simulate scenarios from neutral drift to strong selection; *Ne* values ranging from 100 to 100,000 to simulate the effects of genetic drift; *a* values ranging from 1 to 5 to simulate different levels of advantage; and different sets of initial allele frequencies. Each test included ten separate runs. The preliminary tests indicated that the combination of *z* and *a* determined the state of equilibrium (Data S2). Therefore, we chose a moderate strength of NFDS ($z = 0.2$), a relatively large population size according to the estimated population size (Ne = 10,000), and a sufficient generation time to fix a derived haplotype (generations = 400) for formal tests. According to the phylogenetic results (Figures 6A–6C), we described four separate phases of the evolution of leaf wing polymorphism, assuming that chromosomal inversion originated later in each clade. Phase 1 included two original haplotypes, *P* and *S*, assuming *S* to be a derived haplotype according to dominance-recessiveness; Phase 2 involved the origin of *R* due to chromosomal inversion occurring in *S* and included *P*, *S*, and *R*; Phase 3 involved the origin of *M* and included *P*, *S*, *R*, and *M*; and Phase 4 involved the origin of *V* due to chromosomal inversion occurring in *M* and included *P*, *S*, *R*, *M*, and *V*. Each phase included multiple tests, and each test included ten separate runs in which the specified haplotypes reached equilibrium. The results of the phase-by-phase simulation showed that NFDS is the main driving force of the maintenance of the five haplotypes, in combination with the additional benefits or disadvantages of different genotypes (Figure S6; Data S2).

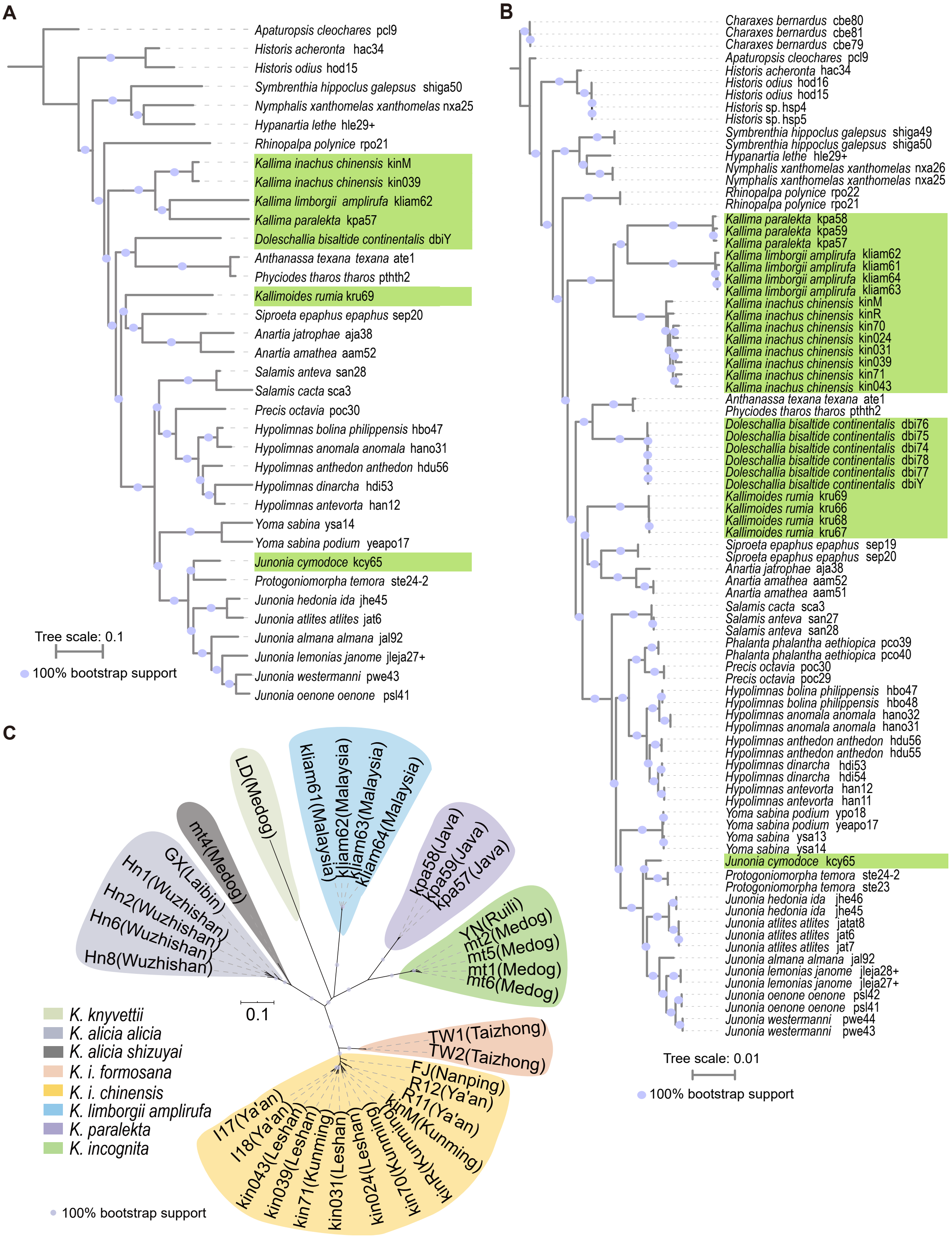
Genotyping

We extracted genomic DNA from butterfly leg tissues using a HiPure Insect DNA Kit (D3109-02). We identified different *cortex* haplotypes by amplifying genomic DNA fragments using PCR and performing Sanger sequencing mainly using universal primers designed for exon 6, including CT-exon6F (AGGCT-GTTCATGGGGAAATG) and CT-exon6R (CCACGTTAGTAACATTACGACAGC). We also designed haplotype-specific primers to validate the genotyping results, including CT-V-F (TTCGCGGTATACAAGGC CAG) and CT-V-R (GCAATTCCTTTGACGAGGCG) for haplotype *V*, CT-E2-F (CGTAATCGTGTTGATAGATTCGTAG) and CT-E3-R (GCTGTAAGTAGG-TAAGTCTAGGA) for haplotype *P*, and CT-R-F (GTATGCAAGGCCAGGGTCTAA) and CT-R-R (GTATGCAA GGCCAGGGTCTAA) for haplotype *R*.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical details and methods as well as sample sizes are indicated in the figure legends or methods. Statistical analyses were performed using R, Excel, PLINK, DEseq2, or custom codes. Data are presented as mean ± SEM and are noted in the figure legends and tables. For phylogenetic analyses, a bootstrap approach was used and bootstrap support values are noted in the figures. For population genetic analyses, standard errors were calculated using a moving block bootstrap approach or jackknife approach. A chi-squared test, Fisher's exact test, Student's *t*-test, or Wilcoxon rank-sum test was used to determine statistical significance, and significance is denoted as * ($P < 0.05$), ** ($P < 0.01$), or *** ($P < 0.001$) in the figures and tables. For the genome-wide association studies and differential expression analyses, *P* values were FDR adjusted using the Benjamini-Hochberg method.

Supplemental figures



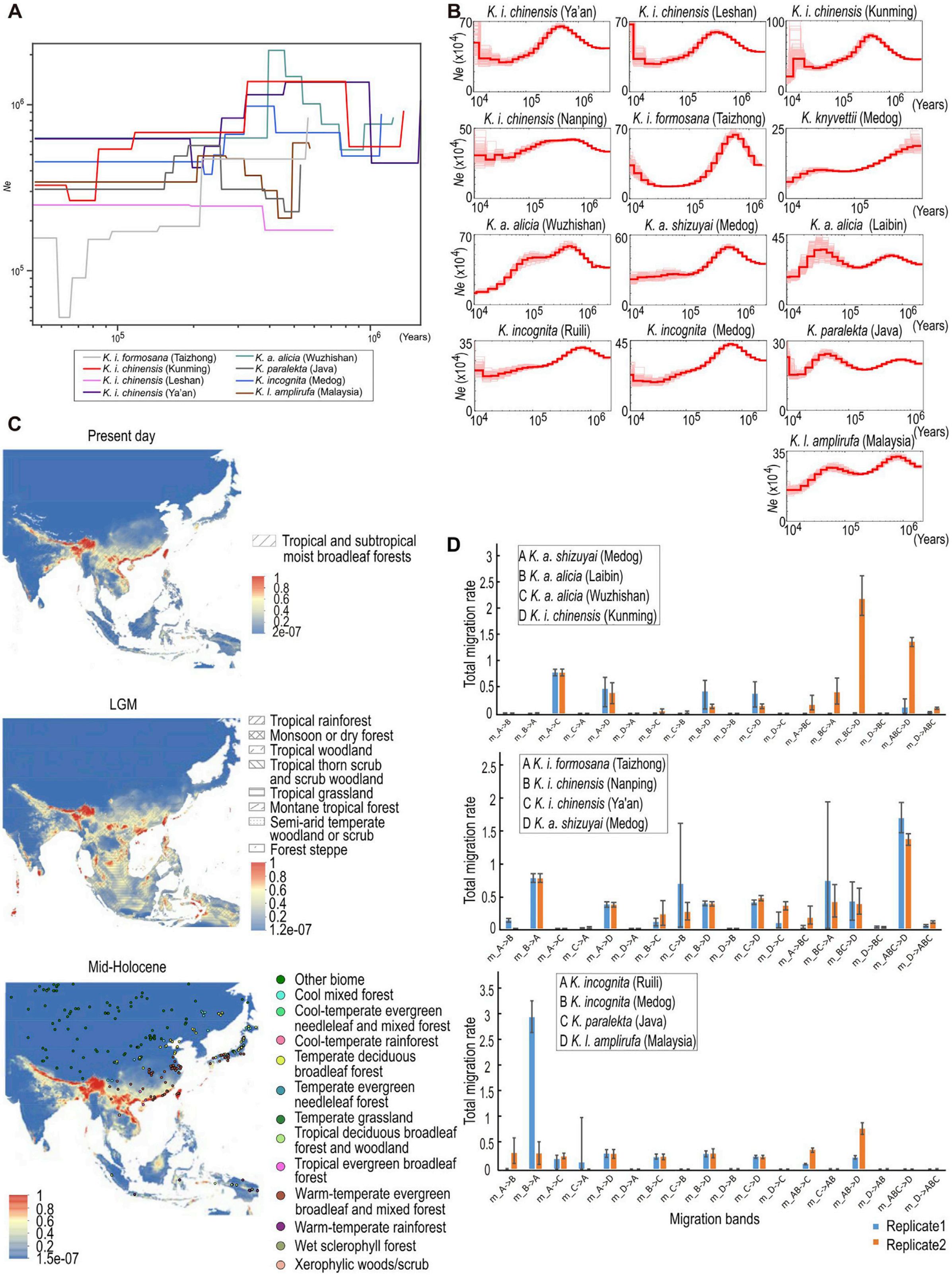
(legend on next page)

Figure S1. Phylogenies of nymphalid butterflies, related to Figure 1

(A) A maximum-likelihood phylogenetic tree of 35 species constructed from genome-wide SNP data.

(B) A maximum-likelihood phylogenetic tree of 84 samples constructed from 5,638 putatively neutral loci. Leaf-mimicking butterflies are highlighted in green in (A) and (B).

(C) A maximum-likelihood phylogeny constructed for 34 *Kallima* samples based on concatenated genome-wide CDS.



(legend on next page)

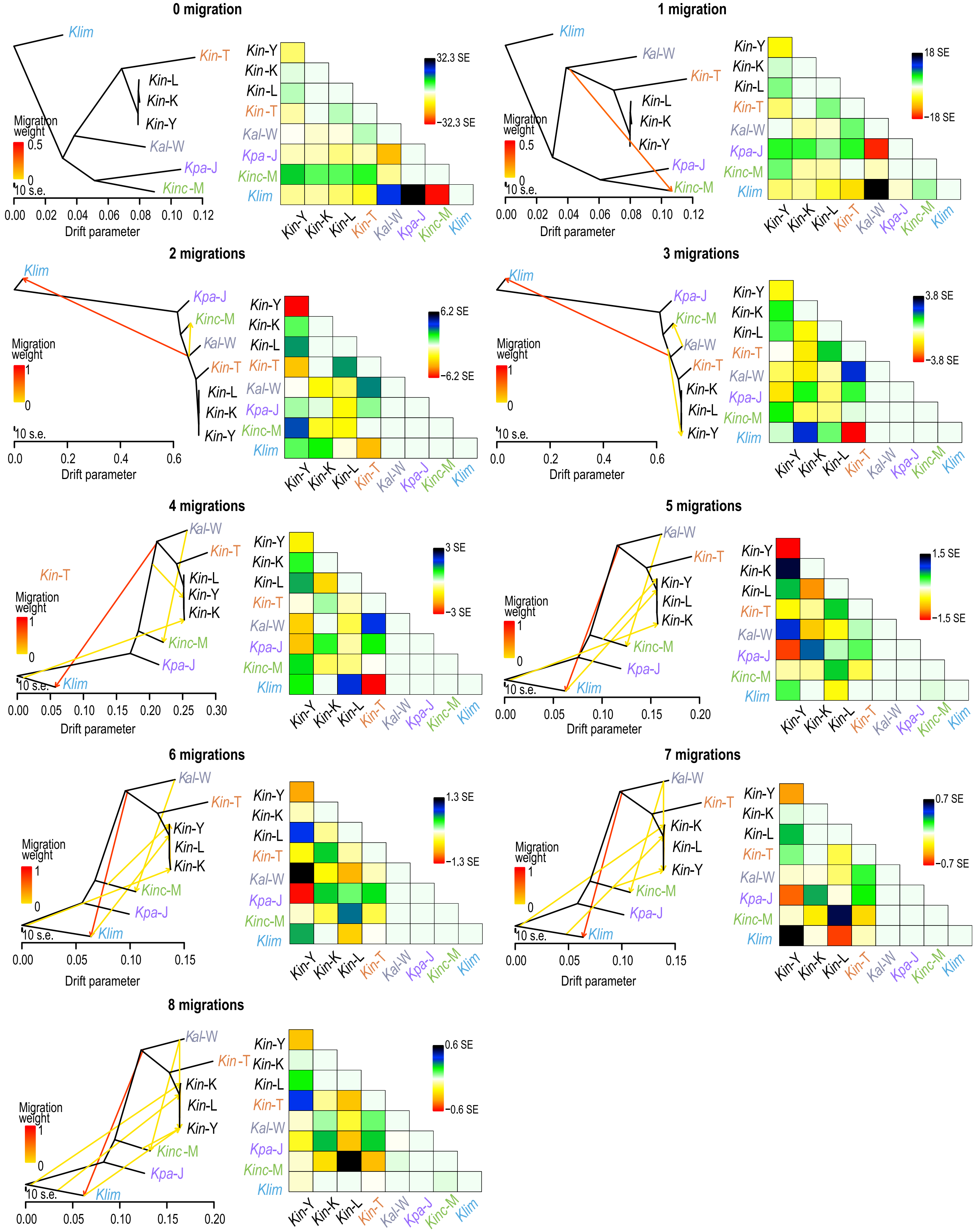
Figure S2. Demographic history inference and ecological niche modeling for *Kallima* butterflies, related to Figure 2

(A) Historical effective population sizes inferred from population data of eight *Kallima* individuals using SMC++ with a mutation rate of 3×10^{-9} and an average generation time of 0.5 years. The population sizes of most species have remained stable for several hundred thousand years.

(B) Historical effective population sizes inferred with PSMC from 13 individual genomes of different *Kallima* species, including *K. i. chinensis* from Ya'an, *K. i. chinensis* from Leshan, *K. i. chinensis* from Kunming, *K. i. chinensis* from Nanping, *K. i. formosana* from Taizhong, *K. a. alicia* from Wuzhishan, *K. a. shizuyai* from Medog, *K. a. alicia* from Laibin, *K. incognita* from Ruili, *K. incognita* from Medog, *K. paralekta* from Java, *K. l. amplirufa* from Malaysia, and *K. knyvettii* from Medog, where the pink shadow represents 100 bootstrap replicates.

(C) Habitat information and ecological niche models of *Kallima* butterflies. Ecological niche modeling was performed using Maxent, and suitable habitat was predicted with an overlay of vegetation information. The color scheme indicates the probability of occurrence from high (red) to low (blue). The current predicted habitat is covered by tropical and subtropical moist broadleaf forests. The habitat of *Kallima* butterflies was fragmented during the LGM. During the mid-Holocene, warm temperate rainforest and tropical evergreen broadleaf forest may provide a suitable ecological habitat.

(D) All possible migration bands were tested in two repeated analyses using G-PhoCS. Migration bands were tested among populations of *K. alicia* subspecies and *K. i. chinensis*, among populations of *K. i. chinensis*, *K. i. formosana*, and *K. a. shizuyai*, and among populations of *K. incognita*, *K. paralekta*, and *K. l. amplirufa*. Each bar represents the 95% highest posterior density of the mean migration rate.



0 migration

1 migration

2 migrations

3 migrations

4 migrations

5 migrations

6 migrations

7 migrations

8 migrations

Figure S3. Inference of phylogenies and migration among *Kallima* species using Treemix, related to Figure 2

In each case, a maximum likelihood tree was inferred among *K. inachus*, *K. alicia*, *K. incognita*, and *K. paralekta* species or populations with the outgroup species *K. limborgii amplirufa* based on genome-wide allele frequency data. We set the linkage disequilibrium threshold to 0.5 for each 50 kb window in the input data and assumed zero to eight migration edges. The residual fit matrix was also generated in each case by dividing the residual covariance of pairwise populations by the average standard error using Treemix. Residuals above zero indicate populations that are closely related and could be candidates for admixture events. *Kin*-Y: *K. i. chinensis* (Ya'an); *Kin*-K: *K. i. chinensis* (Kunming); *Kin*-L: *K. i. chinensis* (Leshan); *Kin*-T: *K. i. formosana* (Taizhong); *Kal*-W: *K. a. alicia* (Wuzhishan); *Klim*: *K. l. amplirufa*; *Kinc*-M: *K. incognita* (Medog); *Kpa*-J: *K. paralekta* (Java).

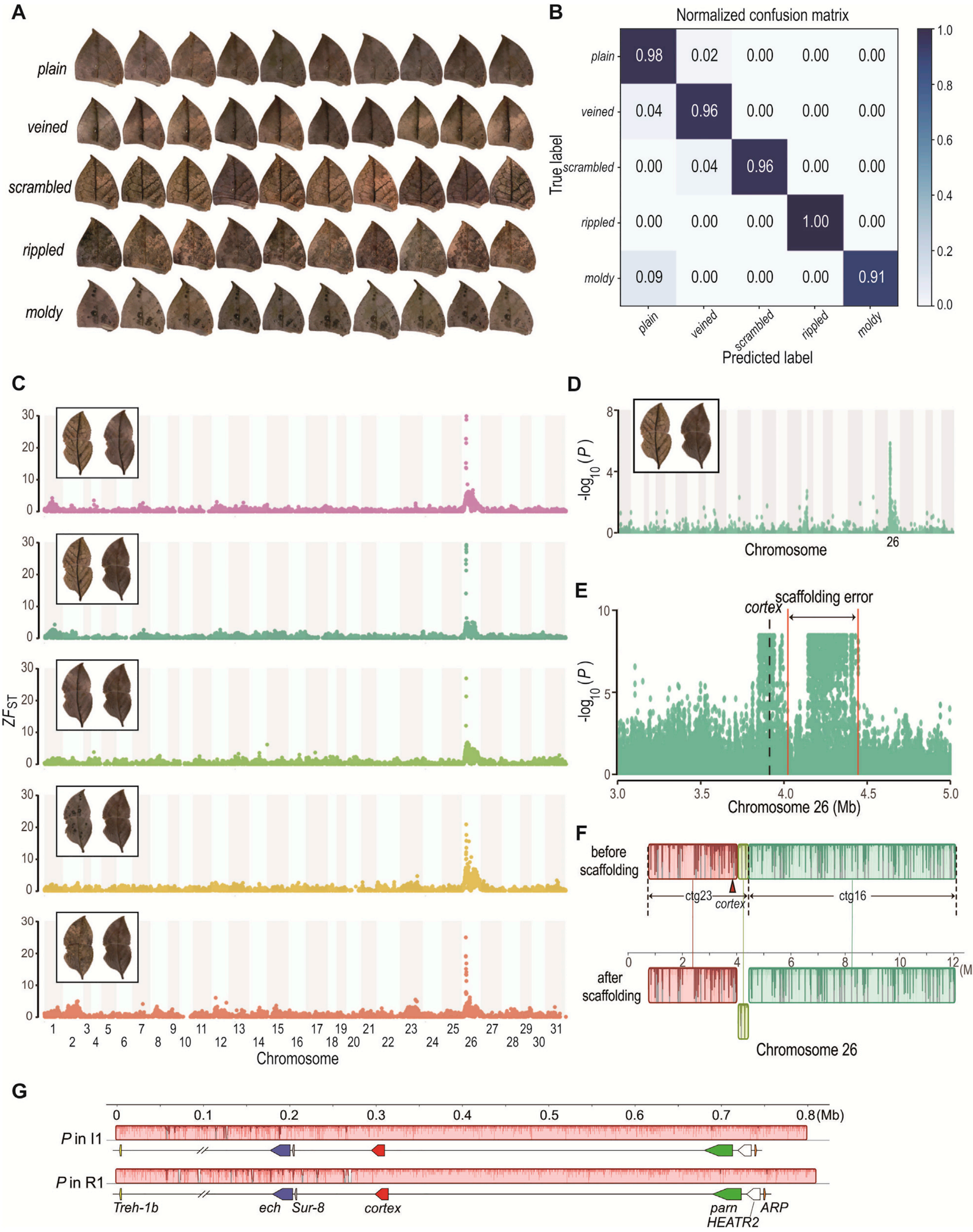


Figure S4. Genome-wide characterization of the leaf wing polymorphism and scaffolding of *cortex* haplotypes, related to Figure 3

(A) Forewing images of *plain*, *veined*, *rippled*, *moldy*, and *scrambled* forms were used to train a support vector machine (SVM) model for GWAS sample classification.

(B) A normalized confusion matrix was calculated to evaluate the accuracy of the SVM classification, indicating good performance.

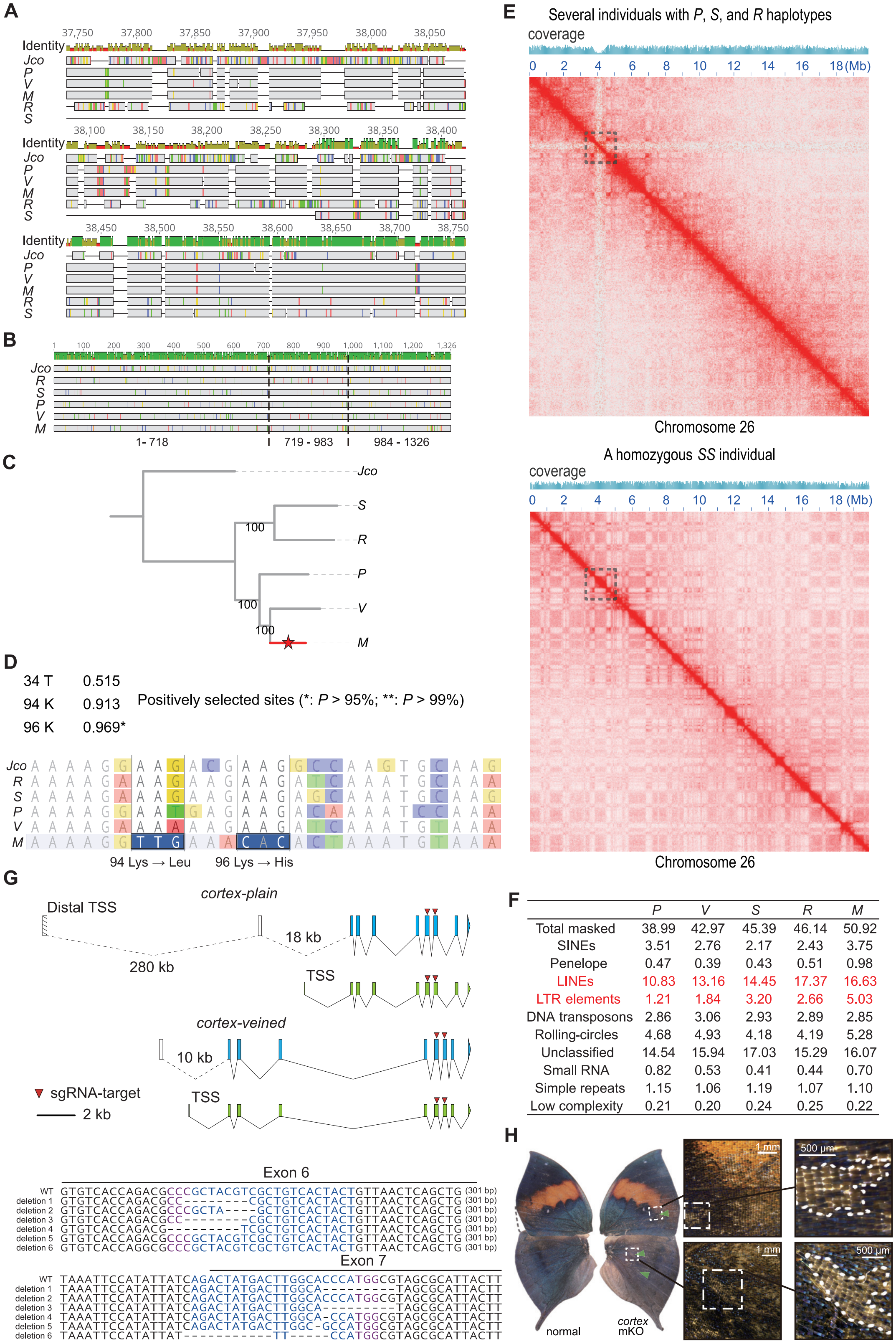
(legend continued on next page)

(C) Genome-wide F_{ST} values were calculated based on the full-genome sequences of *veined* versus *scrambled* forms (red), *plain* versus *scrambled* forms (dark green), *plain* versus *veined* forms (light green), *moldy* versus *plain* forms (yellow), and *rippled* versus *plain* forms (orange) in 50-kb sliding windows and are standardized into *Z* scores. The major peaks of ZF_{ST} values are all located on chromosome 26, consistent with the GWAS results.

(D and E) The results of the original GWAS analyses performed for *plain* versus *scrambled* forms using an uncorrected reference genome. The largest peak in the original GWAS result is also located on chromosome 26 (D) but contains a scaffolding error of approximately 4.02–4.45 Mb in the uncorrected reference genome (E).

(F) The Hi-C scaffolding process separated contig 23 into two fragments and concatenated them in a false order of inversion, which led to the scaffolding error.

(G) Validation of *de-novo*-assembled *cortex* haplotype scaffolds. Two assembled scaffolds of the *plain* (*P*) haplotype were extracted from the *de novo* assemblies of two heterozygous individuals I1 (*VP*) and R1 (*RP*). The alignment shows complete collinearity and a high degree of base identity between them.



	P	*V*	*S*	*R*	*M*
Total masked	38.99	42.97	45.39	46.14	50.92
SINEs	3.51	2.76	2.17	2.43	3.75
Penelope	0.47	0.39	0.43	0.51	0.98
LINEs	10.83	13.16	14.45	17.37	16.63
LTR elements	1.21	1.84	3.20	2.66	5.03
DNA transposons	2.86	3.06	2.93	2.89	2.85
Rolling-circles	4.68	4.93	4.18	4.19	5.28
Unclassified	14.54	15.94	17.03	15.29	16.07
Small RNA	0.82	0.53	0.41	0.44	0.70
Simple repeats	1.15	1.06	1.19	1.07	1.10
Low complexity	0.21	0.20	0.24	0.25	0.22

(legend on next page)

Figure S5. Alignments and signatures of *cortex* haplotypes, related to Figure 4

(A) The *de-novo*-assembled *plain* (*P*), *veined* (*V*), *moldy* (*M*), *rippled* (*R*), and *scrambled* (*S*) *cortex* haplotypes were aligned, with the orthologous haplotype of *Junonia coenia* as the outgroup (*Jco*). The alignment shows various sequence identities in different regions of the *cortex* haplotypes.

(B) The *cortex* CDS regions of *Kallima inachus* (*R*, *S*, *P*, *V*, and *M*) and *Junonia coenia* (*Jco*) were aligned, and potential recombination breakpoints were inferred using GARD (Genetic Algorithm for Recombination Detection).

(C) Positive selection analyses of different branches were performed using the PAML (Phylogenetic Analysis by Maximum Likelihood) branch-site model, and haplotype *M* shows a putative signature of positive selection.

(D) A nonsynonymous substitution (K96H) in the CDS region of *M* is significantly subject to positive selection according to the branch-site model.

(E) Two Hi-C heatmaps on chromosome 26 were constructed using Hi-C data generated from several individuals with *P*, *S*, and *R* haplotypes and one homozygous (*SS*) individual against an assembled reference genome containing the *S* haplotype. The putative TAD-containing *cortex* and neighboring genes are labeled with a gray dotted square in each map. The heatmap generated from mixed individuals shows lower alignment coverage in the *cortex* region.

(F) Percentages of repetitive sequences in *cortex* haplotypes.

(G) The gene structure of *cortex* related to the *P* and *V* haplotypes. The transcripts with proximal promoters are shown in green rectangles, and the transcripts with the distal promoter and additional exons based on RNA-seq data (blank rectangle) or sequence homology (dashed rectangle) are shown in blue rectangles. The predicted distal promoter of the *V* haplotype is not shown because it is outside the inversion region and in the opposite orientation. The sgRNA target sites employed for the somatic mutagenesis of *cortex* are indicated by red arrows. Indels were identified by PCR, molecular cloning, and Sanger sequencing. *Cortex* sgRNA targets are highlighted in blue and PAM sequences are highlighted in purple. TSS, transcription start site.

(H) We observed several mutant phenotypes in *cortex*-mKO mutants including different levels of scale pigmentation fading (four mutants) and irregular eye spots (three mutants). An additional dorsal pattern is shown for whole wings of a *cortex*-mKO mutant with faded scale pigmentation on the dorsal side (indicated by green arrows).

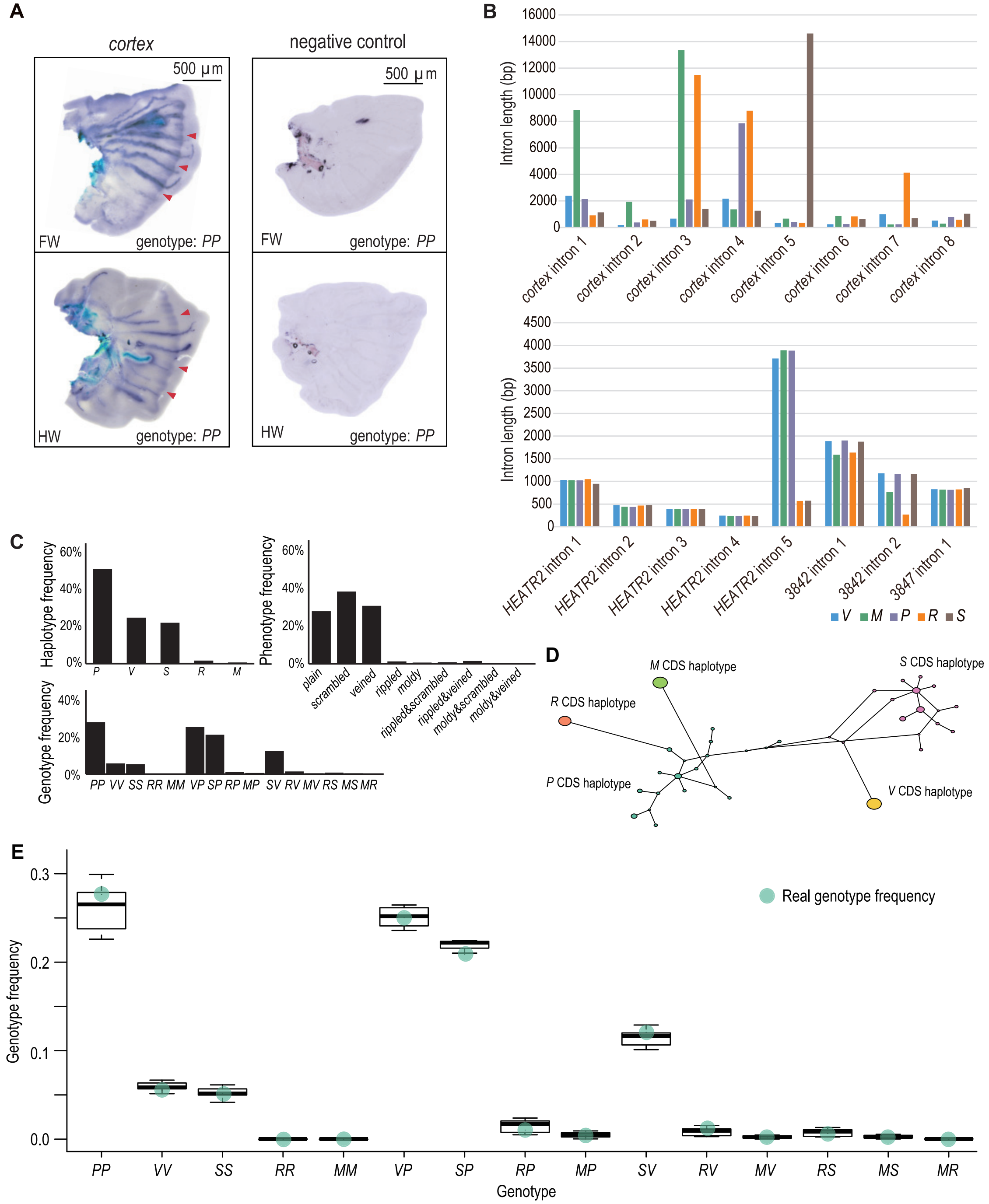


Figure S6. Tracing the evolution of leaf wing polymorphism, related to Figures 3, 5, and 6

(A) *In situ* hybridization of *cortex* and the negative control in fifth-instar larval-stage wing discs of individuals with the *PP* genotype, showing a consistent spatial correlation with the patterns of adult leaf wings in the marginal region under the framework of the nymphalid ground plan (indicated by red arrows).

(B) The distribution of intron lengths in the *cortex* gene and neighboring genes. The intron lengths of *cortex* and neighboring genes were calculated in five haplotypes, with high variation observed in *cortex* and low variation observed in the neighboring genes, *HEATR2* and *3842*.

(C) *cortex* haplotype frequencies, phenotype frequencies, and genotype frequencies calculated in a natural population of *K. inachus*.

(D) Gene network of phased *cortex* CDS regions in the genus *Kallima*.

(E) The real genotype frequencies and the distribution of genotype frequencies in the best-fit population genetic simulation.

See also Data S2 in the supplemental information.