Received Date: 30-Jan-2014
Revised Date : 14-Jun-2014
Accepted Date: 16-Jun-2014
Article type : Original Article

Phylogeography of Heliconius cydno *and its closest relatives: disentangling their origin and diversification*

Carlos F. Arias[1,2], Camilo Salazar[2,3], Claudia Rosales[2], Marcus R. Kronforst[5], Mauricio Linares[3], Eldredge Bermingham[2,4] and W. Owen McMillan[2]

1 Department of Biology, McGill University, 1205 Ave. Dr. Penfield H3A 1B1, Montreal, QC.

2 Smithsonian Tropical Research Institute. Apartado 0843-03092, Panamá, República de Panamá.

3 Biology program, Faculty of Natural Sciences and Mathematics, Universidad del Rosario, Carrera 24 No. 63c-69, Bogotá DC, 111221, Colombia.

4 Patricia and Phillip Frost Museum of Science, 3280 South Miami Avenue, Miami, FL.

5 Department of Ecology & Evolution, University of Chicago, 1101 E. 57th Street, Chicago, IL 60637, USA.

Keywords: *Heliconius cydno*, phylogeography, AFLPs, mtDNA, adaptive radiation, gene flow, genetic drift, Pliocene-Pleistocene boundary

Corresponding author: Carlos F. Arias. Smithsonian Tropical Research Institute. Apartado 0843-03092, Panamá, República de Panamá.

solracarias@gmail.com

Running title: *Heliconius* cydno phylogeography

Abstract

The origins of the extraordinary diversity within the Neotropics have long fascinated biologists and naturalists. Yet, the underlying factors that have given rise to this diversity remain controversial. To test the relative importance of Quaternary climatic change and Neogene tectonic and paleogeographic reorganizations in the generation of biodiversity, we examine intraspecific variation across the *Heliconius cydno* radiation and compare this variation to that within the closely related *H. melpomene* and *H. timareta* radiations. Our data, which consist of both mtDNA and genome scan data from nearly 2250 AFLP loci, reveal a complex history of differentiation and admixture at different geographic scales. Both mtDNA and AFLP phylogenies suggest that *H. timareta* and *H. cydno* are probably geographic extremes of the same radiation that likely diverged from *H. melpomene* prior to the Pliocene-Pleistocene boundary, consistent with hypotheses of diversification that rely on geological events in the Pliocene. The MtDNA suggest that this radiation originated in Central America or the Northwestern region of South America, with a subsequent colonization of the eastern and western slopes of the Andes. Our genome-scan data indicate significant admixture among sympatric *H. cydno/H.timareta* and *H. melpomene* populations across the extensive geographic ranges of the two radiations. Within *H. cydno*, both mtDNA and AFLP data indicate significant population structure at local scales, with strong genetic differences even among adjacent *H. cydno* color pattern races. These genetic patterns highlight the importance of past geoclimatic events, intraspecific gene flow, and local population differentiation in the origin and establishment of new adaptive forms.

Introduction

The Neotropics contain some of the most biodiverse habitats on earth and the mechanisms responsible for this diversity have been the topic of scientific discussions dating back 150 years. The region has had a complex geological and climatic history and, over the past four decades, a number of hypotheses have tied diversification in this region to major tectonic and/or climatic events (Rull 2011). For example, one of the first models of high neotropical diversity proposed that climatic fluctuations during the Pleistocene triggered diversification by isolating populations in forest refugia (Haffer 1969). A second hypothesis argued that the uplift of the Andes and its acceleration during the Pliocene (Gregory-Wodzicki 2000; Hoorn et al. 2010) prompted diversification and speciation by both isolating populations and creating an intricate topology with a range of new ecological conditions (Bush 1994; Elias et al. 2009). A third mechanism suggested that the final closure of the Isthmus of Panama at ~3 million years ago (MYA) changed climate and created a land corridor that joined previously isolated plant and animal communities, likewise creating novel ecological opportunities and promoting speciation (The Great American Biotic Interchange; Coates and Obando 1996; Kirby et al. 2008). Despite the potential for these different geoclimatic changes to promote speciation, the relative importance of each remains controversial (Rull 2011).

The study of recent radiations provides a unique opportunity to understand and test how phenotypic variation arises and accumulates in the Neotropics. In this respect, *Heliconius* butterflies provide an excellent system to study the relative importance of Quaternary climatic change and Neogene tectonic and paleogeographic reorganizations in the generation of biodiversity. With over 400 distinct color pattern varieties among its 43 species, the group is a striking example of a modern adaptive radiation (Emsley 1965; Turner et al. 1979; Turner 1981). Most species converge on a handful of common color patterns, generating mimicry rings which coexist locally (Mallet and Gilbert 1995). Geographic variation in these mimicry rings creates a complex and colorful tapestry where distantly related species often look identical and closely related species or races can look strikingly different (Turner 1976; Mallet and Gilbert 1995). The locally close resemblance between distantly related species (i.e. between *Dismorphia sp., Ithominae*

sp. and *Heliconius sp.*) led to the original hypothesis of mimicry (Bates 1862) and provided Darwin with some of the most visually appealing examples of natural selection and its importance in speciation. Subsequently, we have learned a great deal about the importance of color as both intra- (Papageorgis 1975; Brown 1981; Mallet and Gilbert 1995; Mallet et al. 1998; Mallet et al. 2007) and interspecific signals (Mallet et al. 1998; Jiggins et al. 2001; Naisbit et al. 2001; Merrill et al. 2011; Merrill et al. 2012).

Nonetheless, the evolutionary processes that generate this phenotypic variation remain puzzling. At the core of this puzzle there is a paradox. Natural selection can explain why wing patterns of different *Heliconius* species should converge—strong selection against rare color patterns promotes mimicry (Müller 1879). Natural selection can also explain the maintenance of existing wing pattern diversity—strong frequency-dependent selection removes non-mimetic individuals creating sharp transition zones between divergent phenotypes (Mallet et al. 1998). However, natural selection cannot easily explain the origin of new phenotypes in *Heliconius*. This is a complex paradox—the selection that stabilizes existing patterns is the same force that eliminates novel forms, yet pattern divergence is frequent (Mallet and Gilbert 1995; Turner and Mallet 1996; Joron and Mallet 1998).

Attempts to explain this paradox have often relied on explanations where climatic and/or geological events isolate populations promoting divergence. Some of the earliest theories hypothesized that Pleistocene climatic fluctuations separated previously widespread populations into isolated forest refuges (Brown et al. 1974; Sheppard et al. 1985; Turner and Mallet 1996). During periods of isolation, wing patterns diverged by a combination of biotic (and possibly abiotic) forces leading to the geographical patchwork of mimetic patterns observed today (Turner and Mallet 1996). This scenario, termed the 'refugium/biotic' drift model, is appealing because it can explain both pattern divergence within a species and pattern convergence between species. However, the idea that Pleistocene climatic changes created a series of isolated forest patches across the Neotropics has received little support as a phenomenon promoting diversification (see Knapp and Mallet 2003; Flanagan et al. 2004; Whinnett et al. 2005; Hines et al. 2011;

Patel et al. 2011; Bennett et al. 2012; Ribas et al. 2012) (but see Brower 1994, 1996b). An alternative hypothesis suggests that adaptive divergence in wing color patterns can occur in the absence of complete isolation (Mallet 2010). This mechanism presupposes dispersal limitation and fine-scale population structure, something that has only recently been experimentally explored in *Heliconius* (Kronforst and Gilbert 2008) but may be particularly important in the establishment and spread of new phenotypes in this genus (Mallet 2010; Hines et al. 2011).

Most previous studies attempting to reconstruct the history of diversification among *Heliconius* species have focused on *Heliconius erato* and *H. melpomene* (see Sheppard et al. 1985; Brower 1994, 1996b; Flanagan et al. 2004; Hines et al. 2011). The two species are distantly related, yet have undergone identical color pattern radiations, such that the range of each species is composed of a nearly identical patchwork of wing pattern races. Previous molecular work on the two radiations has demonstrated strong geographic partitioning in extant genetic variation at most loci (Brower 1994; Flanagan et al. 2004; Quek et al. 2010; Hill et al. 2013). Genetic patterns in the two co-mimics differ both in terms of the levels of standing variation and the relationship among major biogeographic groups. Despite the strong phenotypic concordance, this genetic discordance has been interpreted as demonstrating that the two species did not co-diversify and it is argued that *H. erato* diversified first, and *H. melpomene* "adverged" on existing wing pattern variation (Flanagan et al. 2004; Quek et al. 2010). More recent genome-wide studies paint an even more complex history with evidence for pervasive gene flow among closely related *Heliconius* species (The *Heliconius* genome consortium 2012; Kronforst et al. 2013; Martin et al. 2013) (but see Brower 2011; Brower 2013 for altenative explanations). These studies are changing our ideas about how adaptive variation arises (see; Kronforst et al. 2012) and argues that to better understand the geographic mosaic of pattern variation within *Heliconius* we need to examine variation in more species from across the larger radiation.

Here we reconstruct the history of diversification in *Heliconius cydno*. Our objective is to provide a general understanding of relative timing of diversification in this group within

the context of the broader *H. melpomene* and *H. timareta* radiations. Unlike *H. erato*, *H. cydno* is very closely related to both *H. melpomene* and *H. timareta*. The three species form a closely related complex with a partially overlapping distribution on the North Andes slopes of South and Central America. Both *H. cydno* and *H. timareta* are broadly sympatric with *H. melpomene,* whereas *H. timareta and H. cydno* are parapatric*.* The geographic range of *Heliconius cydno* extends along the western slopes of the Andes from Ecuador and up into Central America as far as southern Mexico. It is also found in both the Cauca and Magdelena valleys in Colombia and on the eastern slope of the Andes in Venezuela and Northern Colombia (Figure 1)*. Heliconius timareta* is a geographic replacement of *H. cydno* found on the eastern edge of the Andes from Colombia to Peru and several new races have been recently identified across this region (Brower 1996a; Giraldo et al. 2008; Mérot et al. 2013). Unlike *H. cydno*, which is iridescent blue/black with white or yellow markings and typically mimics *H. sapho* and *H. eleuchia* (see Figure 1), *Heliconius timareta* is typically red and yellow and often falls into the same mimicry ring as *H. melpomene* and *H. erato* (Figure 1).

Here, we use mtDNA "barcode" gene cytochrome oxidase (*Co1*) and amplified fragment length polymorphisms (AFLPs) to construct a comprehensive phylogenetic and biogeographic hypothesis for *H. cydno* diversification. In particular, we test whether diversification between *H. cydno* and its close relatives *H. melpomene* and *H. timareta* is correlated with major glacial cycles in the Pleistocene or whether diversification occurred earlier in the Pliocene consistent with the final Andean uplift and/or the closure of the isthmus of Panama. Recent climate models suggest that forest refugia were not likely before ~1MYA (Ravelo et al. 2004; Ribas et al. 2012). Thus, if diversification was driven by isolation in forest refugia, we expect a young age (between 0-1 MYA) for *H. cydno/ H. timareta/H. melpomene* radiation. In contrast, if the two major geological events that occurred in this region, the final uplift of the North Andes (Gregory-Wodzicki 2000; Hoorn et al. 2010) and the closure of the Isthmus of Panama (Coates and Obando 1996) played an important role in the diversification of the group, we expect an older origin for this diversification. Finally, we use population genetic analyses to investigate the importance of other mechanisms, such as hybridization and dispersal

limitation, in the diversification of the group. Our genetic data provide a broader understanding of this radiation and a context to explore the importance of geoclimatic processes, local population differentiation, and intraspecific hybridization in the evolution of diversity in the Neotropics.

Methods

Sampling and Molecular Data.

Three hundred and ten adult butterflies were collected by the authors or provided by colleagues from several locations (Figure 1; Table 4SM). Our sampling included 186 individuals of *Heliconius cydno*, representing 13 out of the 14 known geographic races under Lamas' recent classification (Lamas et al. 2004). Geographic races differ with respect to the size, shape, and distribution of white and yellow pattern elements. To place the *H. cydno* radiation into a comprehensive phylogenetic framework, we also examined the genetic variation in a subset of 10 *H. melpomene* races from both sympatric and allopatric populations (n=84 individuals), as well as, *H. heurippa* (n= 10 individuals) and 3 allopatric races of *H. timareta* (n=30 individuals) (Figure 1). Wings were removed and bodies preserved in a 10% dimethyl sulphoxide (DMSO) solution (Dawson et al. 1998). All *H. cydno* and *H. timareta* (*H. t. florencia* and *H. t.* subsp. nov) specimens from Colombian localities were stored in the collection of M. Linares at Universidad del Rosario in Colombia while *H. melpomene* and other *H. timareta* specimens were stored in collections at the Smithsonian Tropical Research Institute (see Table 4SM). Genomic DNA was extracted from thoracic tissue using a Qiagen DNeasy Kit, according to the manufacturer's protocol. Individuals were scored for two different molecular markers, mtDNA and AFLPs. We sequenced the arthropod "barcode" region, specifically the 5' half of the mitochondrial gene cytochrome oxidase subunit 1 [*Co1*, 684 base pairs (bp)]. PCR was performed using the primers and amplification conditions of Elias et al. (2007). Sequences from the related genera *Eueides* (3 individuals) and *Heliconius numata* (8 individuals) were downloaded from GeneBank (Table 4SM) and used as outgroups. All sequences were edited and aligned using Geneious Pro v5.5.4. A subset of 194 individuals, between 4-8 individuals of each geographic race, was examined for anonymous nuclear variation across the genome

using AFLPs (Vos et al. 1995). AFLP data was generated using the AFLP Core Reagent Kit from Invitrogen and following the manufacturer's protocol. Eight selective primer combinations were used to generate fragments: EcoRI_CC-FAM/MseI_CTA, EcoRI_CG_FAM/MseI_CAC, EcoRI_CG_FAM/MseI_CAT, EcoRI_CG_FAM/MseI_CAG, EcoRI_CC_FAM/MseI_CAA, EcoRI_CG_FAM/MseI_CTG, EcoRI_CC_FAM/MseI_CAT and EcoRI_CG_FAM/MseI_CAA. Reaction products were electrophoresed on an ABI 3730XL DNA analyzer (Applied Biosystems) with GeneScan Liz-600 size standard (Applied Biosystems). Fragments were sized and scored using ABI GeneMapper v4.0 (PE Applied Biosystems). Only those ranging from 50 to 500 (bp) in length were scored, and within-project normalization was enabled. The advanced peak detection algorithm was used, with light smoothing turned on and all other settings left at defaults settings. We established a bin width of 1.0 and the other scoring parameters were left at default settings. Each locus was visually inspected, and those with a weak or noisy signal were checked by eye or removed. To confirm repeatability and detect run-to-run variation, eight individuals per taxon were independently genotyped twice for all primer combinations (i.e. separate PCR reactions, but same starting DNA).

Phylogenetic and Population Genetic Analyses.

A neighbor-joining tree was constructed for the AFLP markers in PAUP*4.0b10 (Swofford 2000). A major limitation for the use of AFLP data in the reconstruction of phylogenetic hypotheses resides in the degree of homoplasy present in the data and, in particular, the different ways by which taxa can share absences (Meudt and Clarke 2007). We used Nei-Li distances to build our phylogenetic hypothesis since this method reduces the influence of homoplasy by relying only on shared presences as opposed to shared absences. Branch support was estimated with 1,000 bootstrap replicates. Phylogenetic analyses for the "barcode" region were conducted using Bayesian Inference (BI). We employed the GTR+I+G model of nucleotide substitution, which was selected by the Akaike Information Criterion (AIC) using ModelTest v3.8 (Posada and Crandall 1998). BI analyses were conducted in BEAST 1.6.2 (Drummond and Rambaut 2007). These analyses can take into account phylogenetic uncertainty in both the

calibration date and the reconstruction of ancestral range states. Here, we estimated the date of the splits in the tree and the ancestral reconstruction of discrete geographic states on an unknown phylogeny. In the former, we assumed a rate of 1.5% pairwise divergence per million years (similar to Quek et al. 2004; Quek et al. 2010; Hill et al. 2013). For the latter, we established ten discrete geographic states, concordant with the major geographic features in the distribution of *H. cydno, H. timareta* and *H. melpomene*: Northern Magdalena valley (MGN), Southern Magdalena valley (MGS), Northern Cauca valley (CVN), Southern Cauca valley (CVS), Pacific-West Andes (PW), Central America (CA), North-East Andes (EAN), South-East Andes (EAS), French Guiana (FG) and Brazil (BA). The analysis was run under a strict molecular clock in combination with a Yule speciation process, since it has been found that the *Co1* exhibits the most clock like rate among commonly used mtDNA markers (Beltrán et al. 2002). A Continuous-Time Markov Chain model (CTMC) was used to reconstruct discrete geographic states. Finally, we used the Bayesian stochastic search variable selection (BSSVS) that builds a Bayes factor, a test that helps to determine the most parsimonious description of the phylogeographic diffusion process (Lemey et al. 2009). Mixing properties and a stationary distribution of the MCMC were assessed by visual inspection of the parameter trend plots and by verifying that the Effective Sample Size (ESS) was higher than 200. To ensure that the Markov chains mixed sufficiently, two independent runs with 40,000,000 steps were performed for each run. The final analysis was based on genealogies sampled every 8,000 steps from the two independent runs.

Population genetic analyses were conducted for both *mt*DNA and AFLP loci on the complete data set. Population-level estimates were compared between color pattern races and populations concordant with the major geographic regions in the distribution of *H. cydno, H. timareta* and *H. melpomene* (see above). ARLEQUIN v3.5 software (Schneider et al. 2000) was used to evaluate population nucleotide diversity (π , average number of pair-wise differences per sequence), population differentiation, and gene flow. We estimated overall levels of genetic differentiation by calculating the estimator Θ for Wright's F_{ST} (Weir and Cockerham 1984). We tested the null hypothesis of no differentiation by permuting genotypes between populations (10^5 permutations).

Furthermore, we estimated F_{ST} between different races and among different geographic regions. We also used an analysis of molecular variance (AMOVA) with 10^4 permutations (Excoffier et al. 1992) to test how genetic variation was structured 'among geographic regions', 'among races within geographic regions', and 'within races', using a model that contained populations on different geographic regions and individuals of all races for each taxon. Locus-by-locus AMOVAS were also run over the AFLP data, with groupings by taxon (*H. cydno*, *H. timareta/H. heurippa* and *H. melpomene*) and geographic location. We further investigated the effect of geographic distance on genetic divergence by performing a Mantel test using the R package *vegan* (Mantel 1967; Oksanen et al. 2013). Mean F_{ST} values across all populations for both AFLP and mtDNA were mined from our ARLEQUIN analysis and geographic distances were estimated with the *geo.dist* function from the R script *geodist* (Available on: http://www.plantevolution.org/en/downloads.html Accessed 2013 August 12, Joli 2008). *Geodist* generates a matrix of geographic distances from latitude and longitude data by taking into account the curvature of the earth. We performed the Mantel test by taxon (*H. cydno*, *H. timareta/H. heurippa* and *H. melpomene*). Furthermore, we evaluated the effect of isolation by distance within the valleys and regions by grouping *H. cydno* races in three major biogeographic categories (West of the Andes: *H. pachinus*, *H. c. galanthus*, *H. c. zelinde* and *H. c. alithea*; Magdalena valley: *H. c. cydn*o, *H. c. hermogenes*, *H. c. lisethae* and *H. c. wanningeri*; and Cauca valley: *H. c. cydnides* and *H. c. weymeri;* the comparison for the East Andes *H. cydno* races was not performed since there are only two populations). The significance of the Mantel test was assessed after 10,000 permutations. Additionally, historical demographic parameters were calculated for the *Co1* data. Tajima's D (Tajima 1989) and Fu's F_S test (Fu 1997) were estimated to examine departures from a neutral model of evolution. For each case, we ran 10,000 coalescent simulations. Significant D values can be due to selective effects, population expansion, bottleneck, or heterogeneity of mutation rates. Fu´s F_S, on the other hand, is very sensitive to population demographic expansions, which generally lead to large negative F_S values (Fu 1997).

A Bayesian model–based clustering algorithm was implemented in the program STRUCTURE v2.1.4 (Pritchard et al. 2000) on our AFLP loci. We estimated admixture among individuals in 2 comparisons: 1) the complete data collection, *H. cydno*, *H. melpomene* and *H. timareta*/*H. heurippa*) and 2) the color pattern races of *H. cydno* alone. Both analyses were run under an admixture model (individuals may have mixed ancestry) with the option of correlated allele frequencies between populations, as suggested by the authors (Pritchard et al. 2000). This model assumes that frequencies in the different populations are likely to be similar, probably due to migration or shared ancestry (Falush et al. 2003). We determined the number of ancestral clusters, *K*, by comparing the likelihood ratios in ten independent runs for *K* values between 1 and 30 for the first comparison (all data) and between 1 and 15 for the second (*H. cydno* races). Each run consisted of 10^5 iterations, after a short burn-in period of 10^4 iterations. The best estimate of *K* was calculated using the *ad hoc* statistic ΔK (Evanno et al. 2005). We also ran a Discriminant Analysis of Principal Components (DAPC), a methodological approach that relies on data transformation using Principal Component analysis (PCA) as a prior step to a Discriminant Analysis (DA). DA partitions genetic variation into a between group and a within-group component, and attempts to find groups that minimize the latter. This analysis was implemented in the *adegenet* package for the R software (http://adegenet.r-forge.r-project.org/; Jombart 2010). The number of clusters was assessed using the function *find.clusters*, which runs successive K-means clustering with increasing number of clusters (*K*). We covered a wide range of possible clusters (as in STRUCTURE) for the same two comparisons explained above. In all analyses, 100 principal components were retained in the data transformation step. We used AIC to assess the best-supported model, and therefore the number and nature of clusters.

Results

Phylogenetic and population genetic analysis

We examined mtDNA (n= 303) and AFLP (n=194) variation from across the *H. cydno*/*H. melpomene* radiations, including a number of *H. timareta* subspecies. Nucleotide diversity for mtDNA differed between the species, with a higher nucleotide diversity for *H. melpomene* (π = 0.01974) and similar estimates for *H. cydno* and *H. timareta* (π = 0.01075 and π = 0.01131 respectively). Phylogenetic analyses (BI) for the *Co1* gene showed that *H. cydno* and *H. melpomene* were reciprocally monophyletic. There was 100% support in posterior probability (BI) for species monophyly (Figure 2a). A basal lineage of *H. melpomene* occurs on the eastern slopes of the Andes, with *H. melpomene* races from Brazil and French Guiana branching later and following a westward splinting trend (Figure 2a). For *H. cydno*, three well supported clades were largely concordant with major geographic regions: a) one formed by Northeastern Andes (EAN) races, including some haplotypes from *H. c. cydno* and *H. c. wanningeri*, both from Northern Magdalena valley (MGN); b) a second, that combines Pacific-West races (PW), *H. c. chioneus* from Central America (CA), and *H. c. weymeri* from the southern Cauca valley (CVS); and c) a third, including northern Magdalena valley (MGN) races and *H. c. cydnides* from the northern Cauca valley (CVN; Figure 2a).

Within the backdrop of the *H. melpomene* and *H. cydno* radiations, the south Colombian/North Ecuadorian *H. timareta* individuals that we sampled were paraphyletic, falling in two different places in the phylogeny. One *H. timareta* lineage has moderately high support as sister to the *H. cydno* clade. This group contains a geographically mixed collection of individuals including all *H. timareta* individuals (n=11) collected from Ecuador, nearly all *H. pachinus* collected in Costa Rica (except one; n=7), and a portion of *H. t. florencia* and *H. t.* subsp. nov from Southern Colombia (n=12). The remaining *H. timareta* lineage fell in an apparently derived lineage with individuals of *H. heurippa* and *H. cydno* collected from Central America and the Magdalena valley (see Figure 2a). Although closely related, we observed no shared mitochondrial haplotypes between *H. cydno* and *H. timareta*.

We identified 2257 polymorphic AFLP loci between these taxa. These data showed stronger resolution of the three radiations, clustering them by species designation, geography and racial designation. As seen in mtDNA, the AFLP tree clearly distinguishes *H. melpomene* and *H. cyndo* (Figure 2b). However unlike the mtDNA tree, the *H. melpomene* basal lineages occur farther to the east with more derived lineages splitting in a westward trend (Figure 2b), a pattern observed in previous AFLP trees of this radiation (Quek et al. 2010). Similarly, the AFLPs showed a third well-supported group (Figure 2b) that contained all but two (*H. timareta* subsp. nov from Colombia and *H. t. florencia* from Colombia) *H. timareta* specimens (Figure 2b).

Over the entire dataset, the best estimate for the number of distinct populations (*K*) was three, which perfectly corresponded to the three radiations (Figure 3). Interestingly, a clear signal of admixture was detected between sympatric populations of *H. melpomene* and *H. cydno/H. timareta*. For instance, *H. melpomene* populations from the western Andes showed an admixture pattern with *H. cydno* (CA, PW, EAN; blue cluster; Figure 3), whereas *H. melpomene* populations from eastern Andes were admixed with *H. timareta* populations (EAS; green cluster; Figure 3). Moreover, two *H. timareta* individuals showed strong signatures of admixture with more than 50% of their genome shared with *H. melpomene* (one *H. t. florencia* and one *H. t.* subsp. nov). Finally, *Heliconius heurippa*, which has been proposed as a hybrid species between *H. melpomene* and *H. cydno*, fell within the *H. timareta* lineage, consistent with Nadeau *et al*. (2013), suggesting that *H. heurippa* is one of several cases where *H. timareta* has gained color pattern elements from *H. melpomene* (Pardo-Diaz et al. 2012) (but see Brower 2011; Brower 2013).

There was strong genetic structure in both mtDNA and AFLP markers across all *H. cydno, H. melpomene and H. timareta* populations (Figure 1SM). In particular, all *H. cydno* races were genetically structured from each other, even at the *Co1* locus where a number of haplotypes were shared between them (Figure 1SM). Regional F_{ST} values varied from 0.008 to 0.955 for mtDNA and 0.051 to 0.435 for AFLPs (Table 1SM). Hierarchical analyses (AMOVA) on both mtDNA and AFLPs in *H. cydno* showed that

nearly 37.73% (mtDNA) and 87.56% (AFLP) of the total variation was due to variation within races, whereas variation among races within geographic regions accounted for 44.75% (mtDNA) and 8.47% (AFLP) of the total, with only 17.52% (mtDNA) and 3.97% (AFLP) of the variation occurring between geographic regions (Table 2SM). In *H. timareta*, similar to *H. cydno,* genetic variation at both mtDNA and AFLP data was better explained by variation among races within geographic regions (mtDNA:~76%; AFLP: ~12%) and within races (mtDNA:~23%; AFLP: ~87%), with a small proportion of the variation explained by geographic regions (mtDNA:~1%; AFLP: ~1%; Table 2SM). These results suggest that differences among *H. cydno* and *H. timareta* races within geographical regions are as important as differences observed between geographic regions, but both had a minor contribution to overall genetic variance. In contrast, AMOVA analysis on *H. melpomene* showed a discordant pattern between mtDNA and AFLP markers. mtDNA data suggested that nearly 87% of the total variation was between geographic regions, whereas AFLP data suggested that just 12% occurs between the regions, with close to 72% of the variation due to variation within races (Table 2SM). These results suggest differences in current and past patterns of gene flow between *H. melpomene* populations. Consistent with these results, a locus-by-locus AMOVA found that 268 loci (11.8%) were significantly associated with taxon (*H. cydno*, *H. timareta* and *H. melpomene*) but not with geography, and 54 loci (2.3%) were significantly associated with geography but not taxa. This implies that there are specific regions of the genome that maintain differences between *H. cydno, H.timareta* and *H. melpomene.* However, the strong association of some loci with geography and not taxa suggests that some regions of the genome are shared between species at specific geographic locations, most likely due to current gene flow. In fact, there was strong correlation between geographic distance and genetic divergence between *H. melpomene* populations at both AFLPs and mtDNA ($r_{M,\ mtDNA}$=0.536, $r_{M,\ AFLP}$=0.811, $P < 0.05$; Figure 2SM). In contrast, *H. timareta* populations did not showed significant association between geographic and genetic distance ($r_{M,\ mtDNA}$=0.634, $r_{M,\ AFLP}$=0.558, $P > 0.05$; Figure 2SM), while *H. cydno* populations showed significant isolation by distance (IBD) in mtDNA ($r_{M,\ mtDNA}$=0.415, $p<0.05$; Figure 2SM), but not AFLP data ($r_{M,\ AFLP}$=0.104, $p>0.05$; Figure 2SM). A separate analyses of IBD for *H. cydno* races by

geographic region (West Andes, MG, and CV) showed no clear correlation between geographic distance and genetic divergence on the West Andes ($r_{M, mtDNA}$=0.464, $r_{M, AFLP}$=0.077, P > 0.05; Figure 3SM) and MG regions ($r_{M, mtDNA}$=0.592, $r_{M, AFLP}$=0.216, p>0.05; Figure 3SM), but strong IBD was detected in the CV region with a more continuous sampling strategy ($r_{M, mtDNA}$=0.387, $r_{M, AFLP}$=0.516, P > 0.05; Figure 3SM) (Arias et al. 2012).

Phylogenetic reconstruction of the AFLP data clustered individuals of *H. cydno* races largely into phenotypic groups and by geography, but without clear resolution between them (Figure 2b). Assignment tests of these races showed clear correspondence with geography. The admixture pattern in the discriminant analysis of principal components (DAPC) gave a best estimate of 13 for *K*. Seven of the clusters (C1, C5, C6, C9, C10 C11, and C12; Figure 4) were perfectly assigned to phenotypic color pattern races (*H. pachinus*, *H. c. zelinde, H. c. alithea*, *H. c. hermogenes*, *H. c. lisethae*, *H. c. wanningeri,* and *H. c. cordula*; Figure 4). Whereas, the other six clusters were composed of individuals from different races: two of the clusters (C7 and C8) included individuals from *H. cydnides* and *H. c. weymeri*; the C2 cluster was formed by a combination of *H. c. cydno*, *H. c. zelinde, H. c. galanthus* and *H. c. chioneus* individuals; cluster C4 contained individuals from *H.c. zelinde* and *H. c. cydno*; cluster C3 was formed by a group of *H. c. galanthus* and *H. c. chioneus* samples; and finally the cluster C13 comprised specimens from *H. c. weymeri*, *H. c. cydno*, *H. c. cordula* and *H. barinasensis* (Figure 4). The STRUCTURE analysis produced nearly identical results, and identified eleven groups (Figure 4SM). These clusters were sorted by phenotype, with eight clusters corresponding to eight different *H. cydno* color pattern races. The only exceptions were: a) *H. c. chioneus*, *H. c. galanthus* and *H. c. cydno* that form one cluster (Pale-Green cluster, Figure 4SM) and b) *H. c. cydnides* and *H. c. weymeri* that were not distinguishable (Pale-Blue cluster; Figure 4SM).

Divergence time estimates and demographic analysis

We estimated the time of divergence between *H. melpomene* and *H. cydno/H. timareta* based on a general insect molecular clock (similar to Quek et al. 2010; Hill et al. 2013)

that assumed a rate of 1.5% pairwise divergence per million years (Quek et al. 2004). Using this calibration and based on our mtDNA data, the mean pairwise divergence among *H. melpomene* and *H. cydno/H.timareta* clade was 3.2%, which places the minimum age for the origin of the group near the Pliocene-Pleistocene boundary, at least ~2.1 million years ago (2.5-1.7 MYA). Likewise, *H. melpomene* started diversifying at least 2 MYA (3%, 2.4-1.6 MYA), while the first detectable split in *H. cydno/H.timareta* clade occurred 1.3 MYA (1.9%, 1.6-1 MYA). Haplotype diversity was similar between *H. cydno/H.timareta* and *H. melpomene*, with 21.8% and 18.3% unique haplotypes, respectively. Within *H. cydno* races, the highest nucleotide diversity (π) was for *H. c. cydno* followed by *H. c. cydnides* and *H. c. weymeri*. The lowest π was for *H. c. hermogenes* followed by *H. c. lisethae* and *H. c. barinasensis* (Table 3SM). We further annotated the tree branches with the most probable range states of their descendent nodes by color labeling following a Bayesian approach. This strategy suggested that the *H. cydno/H.timareta* clade likely originated in the northwestern part of South America [Maximum State Credibility (MSC)= MG 32% + CVN 22% + CA 16% = 70%; Figure 2a], while the Northeastern Andes showed a MSC of 22%. The remaining regions (PW, CVS, MGS) revealed MSC levels of < 5% each (Figure 2a). Notably, the northern parts of the Magdalena and Cauca valley, and Central America have the highest nucleotide diversities (π) (Table 3SM).

Demographic history assessed through two different neutrality tests, Tajima's D (D_T) and Fu's F_S, yielded slightly similar results. Variation within most regions and races was largely consistent with neutral processes (Table 3SM). For *H. cydno* in two geographic regions, North East Andes (EAN) and the Pacific West Andes (PW), we observed high and significantly negative F_S values, with some races within these regions also showing negative values for this statistic (Table 3SM). The difference in these two regions was reflected in the shape of their mismatch distributions. The CA, MG and CV regions presented wide (9-12 mutations), irregular and multimodal mismatch distributions (Figure 5SM), consistent with old and stationary populations. In contrast, EAN and PW populations showed narrow (3-7 mutations), smooth and unimodal distributions (Figure 5SM), consistent with recently expanding populations. In the case of *H. melpomene*, the

Pacific West Andes (PW) and the South East Andes (EAS) region, exhibited high and significant negative F_S values, where *H. m. melpomene* from Colombia was the only race in the EAS region displaying a negative and significant value for this statistic (Table 3SM). Mismatch distributions were generally congruent with differences observed between the regions, with CA, EAS and FG showing wide (10-15 mutations) and multimodal distributions, consistent with stable and old populations. However, PW presented a narrow (4 mutations) and unimodal distribution, as expected in a recent expanding population (Figure 5SM). Finally, in *H. timareta* none of the geographic regions or races presented significant negative F_S values. In fact, the mismatch distribution for the EAN region was wide (15 mutations) and bimodal, consistent with old, stable populations (Table 3SM, Figure 5SM).

Discussion

We present the first comprehensive genetic examination of the *Heliconius cydno* radiation and place our emerging genetic patterns within the context of the two closely related species, *H. melpomene* and *H. timareta.* Overall, our phylogeographic and demographic analyses suggest a complex interplay among interspecific gene flow, geological processes and local population differentiation within the context of strong natural selection that have been playing out over the past ~2 million years.

Interspecific gene flow and species boundaries among recent Heliconius radiations

Hybridization among members of the *melpomene*/*cydno*/*silvaniform* (MCS) clade is known to occur (Mallet et al. 2007). This hybridization is increasingly recognized as an important source of evolutionary novelty by providing the genetic raw material for both accelerated adaptation and speciation. Indeed, *H. timareta*, which mimics *H. melpomene* in many places across its range, was able to phenotypically track *H. melpomene* through the adaptive introgression of *H. melpomene* color pattern alleles (Pardo-Diaz et al. 2012; The *Heliconius* genome consortium 2012; but see Brower 2011, 2013 and Mallet blog: http://www.heliconius.org/2013/introgression-browers-

criticisms-part-i/ and http://www.heliconius.org/2013/introgression-browers-criticisms-part-ii/). Moreover, hybridization can promote speciation due to the dual role that color patterns play in mimicry and mating behavior in *Heliconius* (Jiggins et al. 2001; Jiggins et al. 2004; Mavárez et al. 2006; Melo et al. 2009). In one of the best examples of hybrid trait speciation (see Jiggins et al. 2008), *H. heurippa* has been proposed to have arisen via hybridization between *H. cydno* and *H. melpomene* (Mavárez et al. 2006; Salazar et al. 2008). The genomic region that controls red color pattern variation introgressed from *H. melpomene* (Salazar et al. 2010; Pardo-Diaz et al. 2012) and strong mating discrimination between *H. melpomene*, *H. cydno* and lab-created hybrids with a *H. heurippa* wing color pattern (red plus yellow bands) is in part caused by phenotypic differences among them (Melo et al. 2009; Salazar et al. 2010; Pardo-Diaz et al. 2012). Our data suggest that *H. heurippa* is more likely a northeastern race of *H. timareta*. Although *H. heurippa*'s mtDNA haplotype falls on a derived branch of the *H. cydno*/*H. timareta* mtDNA lineage, our AFLP data clearly place *H. heurippa* within the *H. timareta* diversification. This placement of *H. heurippa* was also found in a recent restriction associated DNA (RAD) sequencing analysis, which surveyed more nucleotide sites but fewer races of *H. cydno* and *H. timareta* (Nadeau et al. 2013). In both studies, *H. heurippa* clusters with races of *H. timareta*, to the exclusion of both *H. cydno* and *H. melpomene*. Thus, in this case, adaptive introgression from a red banded form of *H. melpomene* into a *H. timareta* race with a phenotype very similar to *H. cydno*, perhaps the newly discovered and geographically proximal *H. t.* subsp. nov form, likely led to the origin of *H. heurippa*. This possibility was suggested five years ago when the *H. timareta* subsp. nov. race was discovered and is presently in the process of being tested experimentally in Colombia (Linares, unpublished). As has been pointed out previously (Nadeau et al. 2013), this event did not lead to genomic admixture on a large scale. Rather most of the nuclear variation from *H. timareta*, save that around the color pattern region, seems to have been retained (Nadeau et al. 2013). Genetic differences between *H. heurippa* and *H. timareta* are greater than those observed among *H. timareta* races (Figure 1SM), a pattern that suggests some level of reproductive isolation.

There is clear genetic evidence for hybridization and admixture among the three species. Although our AFLP data distinguished the three species radiations, there were two *H. timareta* individuals that showed strong signatures of admixture with more than 50% of their genome shared with *H. melpomene*– one *H. t. florencia* and one *H. t.* subsp. nov. Sympatric *H. melpomene* and *H. t. florencia* are perfect mimics and it is possible that the *H. t. florencia* individual was incorrectly identified. However, this individual possessed a *H. timareta* mtDNA haplotype and was more likely a F_1 hybrid from a mating between a female *H. t. florencia* with a *H. melpomene* male. The *H. t.* subsp. nov individual, in contrast, is easily distinguishable from the co-occurring *H. melpomene* race based on its wing color pattern. This individual had a *H. melpomene* mtDNA haplotype, but largely possessed a *H. t.* subsp. nov wing pattern phenotype. In this case, knowledge of the genetic basis of color pattern suggests that this individual was not an F_1, but rather some later generation hybrid. There are strong pre- and postzygotic reproductive barriers between *H. timareta* and *H. melpomene* (Sanchez *et al.* in prep), but when matings occur, F_1 males are fertile, viable and capable of backcrossing with either parental type (Sanchez *et al.* in prep).

In addition to the above individual examples, there was also general evidence for recurrent gene flow among sympatric *Heliconius melpomene* and *H. cydno/timareta* populations. The genome scan data suggest that *H. melpomene* races [*H. m. melpomene* (Colombia), *H. m. vulcanus* and *H. m. rosina*] sympatric with *H. cydno* have as much as 20% of their genome coming from the *H. cydno* gene pool (Figure 3). *Heliconius melpomene* races sympatric with *H. timareta* [*H. m. ecuadorensis, H. m. malleti* and *H. m. plesseni*] in the eastern Andes slopes showed a similar pattern of admixture (~20%). In contrast, none of the three allopatric races of *H. melpomene* [*H. m. nanna*, *H. m. thelxiopeia* and *H. m. melpomene* (French Guiana)] demonstrate any evidence for admixture with either *H. timareta* or *H. cydno* (Figure 3). These data are consistent with recent explorations of genomic admixture based on RAD and whole genome sequencing data (see Kronforst et al. 2013; Martin et al. 2013; Nadeau et al. 2013). Similar to our AFLP data, these studies showed that sympatric species pairs of *H. melpomene* and *H. cydno/ H. timareta* had patterns of admixture consistent with

substantial contemporary gene flow. Despite the ongoing gene flow, the genetic data largely cluster sympatric taxa by recognized species boundaries based on morphology, behavior and ecology.

Center of origin and diversification

Even within the context of substantial hybridization, our genetic data, particularly the mtDNA data, retain useful information about the potential origin and timing of diversification of the *H. cydno* radiation. Overall the level of mtDNA variation within the *H. cydno*/*H. timareta* group is less than that seen across the *H. melpomene* radiation. Given the allopatric distribution of *H. cydno* and *H. timareta*, it is tempting to speculate that they are geographic extensions of the same radiation. Our data and a recent maximum likelihood phylogeny constructed using aligned RAD sequence data (Nadeau et al. 2013) provide strong support for a distinct monophyletic *H. cydno/timareta/heurippa* lineage (Figure 2 and 3) (but see Brower 1996a, b; Brower 2011). Viewed in this light, the mtDNA data indicate a complicated relationship that is difficult to reconcile based on geographic proximity or a simple colonization history. In particular, the *H. timareta* individuals cluster in two different regions of the broader *H. cydno*/*H. timareta* mtDNA lineage (Figure 2). Most fall on a basal mtDNA branch with the geographically distant *H. pachinus*. *Heliconius pachinus* is an incipient species from the Pacific coast of Costa Rica, which is historically known to hybridize with *H. cydno galanthus* (also from Costa Rica; comparably see Figure 2 and 4a; Kronforst et al. 2006). The remaining individuals of *H. timareta* cluster in a derived position of the mtDNA genealogy with *H. heurippa, H. cydno galanthus* from Central America, one individual of *H. pachinus,* and *H. cydno wanningeri* from the upper Magdalena valley. In this case, the proximal position of the bulk of *H. pachinus* individuals on our mtDNA genealogy suggests that the lineage evolved in Central America and spread down the eastern slopes of the Andes and into the Magdalena and Cauca valleys, across to the western slopes of the Andes and ultimately recolonized Central America, perhaps displacing the ancestral *H. pachinus*. Along the way, the lineage that we now recognize as *H. timareta* acquired color pattern alleles through introgression and began mimicking *H. melpomene*; whereas, *H. cydno* largely tracked phenotypic variation in distantly

related *H. sapho* and *H. eleuchia*. The exact order of colonization is not clear from the present mtDNA data, as there is little support for the internal branches within the mtDNA haplotype tree and because contemporary gene flow is likely obscuring historical patterns. This scenario is also supported by the ancestral area reconstruction and by extant levels of variation. The mtDNA tree suggests a 70% posterior probability for the northwestern part of South America and Central America, relative to a 22% posterior probability for a Northeastern Andes origin. In addition, the mismatch distribution analysis found broad and multimodal curves for Central America, the Magdalena and Cauca valleys, consistent with older and stationary populations. In contrast, the Northeastern Andes and the Pacific-West region showed mismatch distributions consistent with recent colonization and population expansion. A similar pattern of origination northwest of the Andes and later southward expansion has been proposed for *H. erato* based on mtDNA and nuclear variation (Quek et al. 2010; Hill et al. 2013). In the *H. erato* case, and as seen in our data, the presence of ancient and diverse mtDNA haplotypes in Central America supports this origin (Hill et al. 2013). However, the northwest Andes or Central American origin of the *H. cydno/H. timareta* radiation is not obviously supported by our AFLP data. AFLP data clustered *H. cydno* and *H. timareta* races mainly by phenotype and secondarily by geographic region, but the relationships between races and/or geographic regions are not well supported (Figure 2b). In general, our AFLP data are more likely revealing contemporary patterns of gene flow and do not provide the genealogical resolution needed to tease apart different colonization scenarios. Thus, locus-by-locus AMOVAs found that some loci were highly correlated with geography but not taxonomy, consistent with current gene flow between populations at the local scale.

Time of diversification

Heliconius melpomene and H. cydno/*H. timareta* populations began to diverge around the Pliocene-Pleistocene boundary, ~2.1 million years ago (2.5-1.7 MYA), and the first detectable split within *H. melpomene* occurred around the same period. In contrast, *H. cydno/H. timareta* most likely started diversifying during the Pleistocene [at least ~1.3 MYA [(1.6 -1 MYA)]. Earlier studies, also using mtDNA and other nuclear markers,

proposed that the differentiation between *H. melpomene* and *H. cydno* occurred earlier within the Pleistocene (1.5 MYA; Beltrán et al. 2002). Our study differs from those studies in that we used a larger sample of races within *H. melpomene* and *H. cydno* and a different estimate of divergence per lineage per time. Here we used a insect molecular clock (similar to Quek et al. 2010; Hill et al. 2013) that assumes a rate of 1.5% pairwise divergence per million years (Quek et al. 2004). More recently, Hoyal and Charleston (2012), using the Simonsen et al. (2011) Papilioninae butterfly calibration and a multilocus species-tree approach, estimated that the divergence between *H. cydno* and *H. melpomene* occurred during the Pleistocene (around ~ 0.6 MYA). However, species-tree estimation methods in the presence of ongoing gene flow can make some splitting events appear much more recent than they actually are. This is because species-tree methods assume that the incomplete lineage sorting observed is not caused by current gene flow (Knowles 2009; McCormack et al. 2009) - an assumption that is almost certainly violated in *Heliconius* butterflies.

If our time estimates for the split between *H. cydno/H. timareta* and *H. melpomene* (~2.1 MYA) are roughly correct, diversification may be tied to geoclimatic events associated with the Pliocene-Pleistocene boundary. Two events, the closure of the Isthmus of Panama (~3 MYA; Coates and Obando 1996; Kirby et al. 2008) and the final uplifting of the North Andes (2-5 MYA; Hoorn et al. 1995; Gregory-Wodzicki 2000; Ochoa et al. 2012) are particularly noteworthy. The former likely increased connectivity between Central America and South America, whereas the emergence of the Andes created new habitats that could promote diversification. Different lines of evidence suggest an important role of both events. First, the placement of *H. pachinus* from Central America as an ancient lineage in the *H. cydno/H. timareta* diversification suggests that the lineage may have evolved in Central America and spread into South America (see above). Moreover, the other two races in Central America appeared more closely related to races from different geographic basins. In our analyses, *Heliconius c. galanthus* was more closely related to eastern forms of *H. cydno*, while *H. c. chioneus* clustered with Pacific races. This observed pattern suggests a complex pattern of colonization and re-colonization of Central America, perhaps facilitated by the closure of

the Isthmus of Panama. Additionally, palynological data suggest a change in plant composition from tropical lowland to a montane forest type during the late Pliocene in the northern Andes (Van der Hammen et al. 1973; Hooghiemstra and Ran 1994). Presently there are clear altitudinal differences between the distribution of *H. melpomene* and *H. timareta*/*H. cydno*. Races of *H. melpomene* are more associated with tropical lowland habitats (0 to 1000 m), while races of *H. timareta*/*H. cydno* occupy higher altitudinal ranges in the understories of forests in the Andes (between 500 to 2000 m). Coincident diversification with the rise of the Andes has been observed in other animal groups (Jiggins et al. 2006; Wahlberg and Freitas 2007; Silva-Brandão et al. 2008; Elias et al. 2009; Santos et al. 2009; Hoorn et al. 2010; Chaves et al. 2011). For example, Elias et al. (2009) found that the origin and diversification of two closely related genera of Ithomiine butterflies was correlated with elevation range. Similarly, the riodinid butterfly *Ithomiola* shows a clear pattern of parapatric speciation across an elevation gradient, where old basal species are distributed in the lowlands and younger derived taxa showed highland distributions (Hall 2005). This is exactly the pattern that we observe and our results suggest a mixed role of both the closing of the Isthmus of Panama and the final rise of the Andes in the evolution of *H. cydno/H. timareta* complex. However, more comprehensive genomic sampling will be required to more accurately estimate branching order, colonization history, and the timing of differentiation.

Genetic differentiation among Heliconius populations

Within *Heliconius* species, our data suggested strong differences among races at local scales. Although major mtDNA lineages are spread widely, there are strong genetic differences among adjacent races. Similarly, AFLP data revealed widespread gene flow among geographic regions but significant local population differentiation. Notably, in the Magdalena valley there are four different color pattern races of *H. cydno* in a range of less than ~500 Km (Figure 1), with no clear geographic barriers or environmental gradients separating them. Yet, both STRUCTURE and DAPC assigned individuals to four well-defined clusters, implying prominent population differentiation at a very small geographic scale. The clustering we observe could be the result of discrete sampling

from a population exhibiting continuous variation (see Pritchard et al. 2000). Consistent with this expectation, continuous sampling of individuals of *H. cydno* along the Cauca valley showed strong isolation by distance (Figure 3SM; Arias et al. 2012). A similar pattern of strong isolation by distance was observed in several different *Heliconius* species across a geographic transect in Costa Rica (Kronforst and Gilbert 2008). This is exactly the type of population structure that would promote local differentiation (see Mallet 2010).

The coupling of strong population structure and interspecific hybridization provides a potent source for the origin and spread of new adaptive variation. As our data suggest, these two local processes are strongly entwined with past geoclimatic events. A similar interplay appears to underlie diversification of African rift lake cichlids (Loh et al. 2013) and marine and freshwater stickleback populations (Schluter and Conte 2009). In the former case, there is strong genetic differentiation among ecologically divergent forms within isolate rift lakes. River populations, in contrast, showed admixed genomes, suggesting an important role for these populations as a conduit for functionally important genetic variation that facilitated rapid radiations within newly formed lakes (Loh et al. 2013). For sticklebacks, Schluter and Conte (2009) proposed a model for parallel evolution in freshwater adapted forms that relies on recurrent gene flow between freshwater and marine populations. This gene flows allows freshwater adapted alleles to be maintained in low frequency in marine populations, which similar to the cichlid fish example facilitates rapid adaption to novel freshwater environments created by geoclimatic events. These findings, together with our study, underscore the interplay between historical and contemporary processes in the generation of new adaptive forms. In particular, these case studies highlight the importance of hybridization in the formation and spread of new variation (sensu Rius and Darling 2014). A better understanding of how these processes (historical and contemporary) are playing out across the genomes of recent adaptive radiations is needed and will likely require higher resolution data, a reality given growing genomic resources.

Acknowledgements

We thank C. Jiggins and members of the McMillan and Bermingham Laboratory for useful comments on early versions of the manuscript. We are grateful to R. Gillespie and three anonymous reviewers for valuable comments on the manuscript. We acknowledge the Smithsonian Tropical Research Institute (STRI), where the laboratory work was carried out and funded. This project was also funded by NSF grant (DEB-640279) and IOS (1052541) to WOM. C.F. Arias was supported by fellowships from Colfuturo, the Levinson Family, Max Binz Family, NEO-STRI and McGill University. C. Salazar received support from Universidad del Rosario. M. Linares was funded by Fondo FIUR from Universidad del Rosario. M. Kronforst was funded by NSF grant (DEB-1316037). For the genetic access permit number RGE0027-LAM3483, we thank the Ministerio de Ambiente, Vivienda y Desarrollo Territorial (Colombia).

References

Arias, C. F., C. Rosales, C. Salazar, J. Castaño, E. Bermingham, M. Linares, and W. O. McMillan. 2012. Sharp genetic discontinuity across a unimodal *Heliconius* hybrid zone. Molecular Ecology 21:5778-5794.

Bates, H. 1862. Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. Transactions of the Linnean Society of London 23:495-566.

Beltrán, M., C. Jiggins, V. Bull, W. McMillan, E. Bermingham, and J. Mallet. 2002. Phylogenetic discordance at the species boundary: gene genealogies in *Heliconius* butterflies. Molecular Biology and Evolution 19:2176-2190.

Bennett, K., S. Bhagwat, and K. Willis. 2012. Neotropical refugia. The Holocene 22:1207-1214.

Brower, A. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences, USA 91:6491-6495.

Brower, A. 1996a. A new mimetic species of *Heliconius* (Lepidoptera: Nymphalidae), from Sourtheastern Colombia, revealed by analysis or mitochondrial DNA sequences. Zool. J. Linn. Soc. 116:317-332.

Brower, A. 1996b. Parallel race formation and the evolution of mimicry in *Heliconius* butterflies: a phylogenetic hypothesis from mitochondrial DNA sequences. Evolution 50:195-221.

Brower, A. V. Z. 2011. Hybrid speciation in Heliconius butterflies? A review and critique of the evidence. Genetica 139:589-609.

Brower, A. V. Z. 2013. Introgression of wing pattern alleles and speciation via homoploid hybridization in Heliconius butterflies: a review of evidence from the genome. Proceedings of the Royal Society B: Biological Sciences 280.

Brown, K., P. Sheppard, and J. Turner. 1974. Quaternary refugia in tropical America: evidence from race formation in *Heliconius* butterflies. Proceedings of the Royal Society of London, B 187:369-378.

Brown, K. S. 1981. The biology of *Heliconius* and related genera. Annual Review of Entomology 26:427-456.

Bush, M. 1994. Amazonian speciation: a necessarily complex model. Journal of Biogeography 21:5-17.

Chaves, J. A., J. T. Weir, and T. B. Smith. 2011. Diversification in *Adelomyia* hummingbirds follows Andean uplift. Molecular Ecology 20:4564-4576.

Coates, A. and J. Obando. 1996. The geological evolution of the Central American Isthmus. Pp. 21–56 *in* J. Jackson, A. Budd, and A. Coates, eds. Evolution and environment in tropical America. The University of Chicago Press., Chicago.

Dawson, M. N., K. A. Raskoff, and D. K. Jacobs. 1998. Field preservation of marine invertebrate tissue for DNA analyses. Molecular Marine Biology and Biotechnology 7:145-152.

Drummond, A. and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

Elias, M., R. Hill, K. Willmott, K. Dasmahapatra, A. Brower, J. Mallet, and C. Jiggins. 2007. Limited performance of DNA barcoding in a diverse community of tropical butterflies. Proceedings of the Royal Society of London, B 274:2881-2889.

Elias, M., M. Joron, K. Willmott, K. Silva-Brandao, V. Kaiser, C. Arias, L. Gomez Piñerez, S. Uribe, A. Brower, A. Freitas, and C. Jiggins. 2009. Out of the Andes: patterns of diversification in clearwing butterflies. Molecular Ecology 18:1716-1729.

Emsley, M. G. 1965. Speciation in *Heliconius* (Lep., Nymphalidae): morphology and geographic distribution. Zoologica, New York 50:191-254.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. Molecular Ecology 14:2611-2620.

Excoffier, L., P. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Falush, D., M. Stephens, and J. Pritchard. 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Flanagan, N., A. Tobler, A. Davison, O. Pybus, D. Kapan, S. Planas, M. Linares, D. Heckel, and W. McMillan. 2004. Historical demography of Müllerian mimicry in the neotropical *Heliconius* butterflies. Proceedings of the National Academy of Sciences of the United States of America 101:9704.

Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and backgroud selection. Genetics 147:915-925.

Giraldo, N., C. Salazar, C. Jiggins, E. Bermingham, and M. Linares. 2008. Two sisters in the same dress: *Heliconius* cryptic species. BMC Evolutionary Biology 8:324.

Gregory-Wodzicki, K. 2000. Uplift history of the Central and Northern Andes: a review. Geological Society of America Bulletin 112:1091.

Haffer, J. 1969. Speciation in amazonian forest birds. Science 165:131-137.

Hall, J. 2005. Montane speciation patterns in *Ithomiola* butterflies (Lepidoptera: Riodinidae): are they consistently moving up in the world? Proceedings of the Royal Society B: Biological Sciences 272:2457-2466.

Hill, R. I., L. E. Gilbert, and M. R. Kronforst. 2013. Cryptic genetic and wing pattern diversity in a mimetic *Heliconius* butterfly. Molecular Ecology 22:2760-2770.

Hines, H. M., B. A. Counterman, R. Papa, P. A. D. Moura, M. Z. Cardoso, M. Linares, J. Mallet, R. D. Reed, C. D. Jiggins, M. R. Kronforst, and W. O. Mcmillan. 2011. Wing patterning gene redefines the mimetic history of *Heliconius* butterflies. Proceedings of the National Academy of Sciences 108:19666-19671.

Hooghiemstra, H. and E. T. H. Ran. 1994. Late and Middle Pleistocene climatic change and forest development in Colombia; pollen record Funza II (2e158 m core interval). Pp. 211-246 *in* R. Bonnefille, and H. Hooghiemstra, eds. Pollen and Climate. Elsevier, Amsterdam, Netherlands.

Hoorn, C., J. Guerrero, G. A. Sarmiento, and M. A. Lorente. 1995. Andean tectonics as a cause for changing drainage patterns in Miocene northern South America. Geology 23:237-240.

Hoorn, C., F. P. Wesselingh, H. ter Steege, M. A. Bermudez, A. Mora, J. Sevink, I. Sanmartín, A. Sanchez-Meseguer, C. L. Anderson, J. P. Figueiredo, C. Jaramillo, D. Riff, F. R. Negri, H. Hooghiemstra, J. Lundberg, T. Stadler, T. Särkinen, and A. Antonelli. 2010. Amazonia through time: Andean uplift, climate change, landscape evolution, and biodiversity. Science 330:927-931.

Hoyal, J. and M. Charleston. 2012. Phylogenetic codivergence supports coevolution of mimetic *Heliconius* butterflies. PLoS ONE 7:e36464.

Jiggins, C., C. Estrada, and A. Rodrigues. 2004. Mimicry and the evolution of pre-mating isolation in *Heliconius melpomene*. Journal of Evolutionary Biology 17:680-691.

Jiggins, C., R. Mallarino, K. Willmott, and E. Bermingham. 2006. The phylogenetic pattern of speciation and wing pattern change in neotropical Ithomia butterflies (Lepidoptera: Nymphalidae). Evolution 60:1454-1466.

Jiggins, C., R. Naisbit, R. Coe, and J. Mallet. 2001. Reproductive isolation caused by colour pattern mimicry. Nature 411:302-305.

Jiggins, C., C. Salazar, M. Linares, and J. Mavarez. 2008. Hybrid trait speciation and *Heliconius* butterflies. Philosophical Transactions of the Royal Society B: Biological Sciences 363:3047.

Jombart. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11:94.

Joron, M. and J. Mallet. 1998. Diversity in mimicry: paradox or paradigm? Trends in Ecology & Evolution 13:461-466.

Kirby, M. X., D. S. Jones, and B. J. MacFadden. 2008. Lower Miocene stratigraphy along the Panama canal and its bearing on the Central American peninsula. PLoS ONE 3:e2791.

Knapp, S. and J. Mallet. 2003. Refuting Refugia? Science 300:71-72.

Knowles, L. L. 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. Systematic Biology 58:463-467.

Kronforst, M. and L. Gilbert. 2008. The population genetics of mimetic diversity in *Heliconius* butterflies. Proceedings of the Royal Society B: Biological Sciences 275:493-500.

Kronforst, M., L. Young, L. Blume, and L. Gilbert. 2006. Multilocus analyses of admixture and introgression among hybridizing *Heliconius* butterflies. Evolution 60:1254-1268.

Kronforst, M. R., G. S. Barsh, A. Kopp, J. Mallet, A. Monteiro, S. P. Mullen, M. Protas, E. B. Rosenblum, C. J. Schneider, and H. E. Hoekstra. 2012. Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. Pigment Cell & Melanoma Research 25:411-433.

Kronforst, M. R., M. E. B. Hansen, N. G. Crawford, J. R. Gallant, W. Zhang, R. J. Kulathinal, D. D. Kapan, and S. P. Mullen. 2013. Hybridization reveals the evolving genomic architecture of speciation. Cell Reports 5: 666–677.

Lamas, G., C. Callaghan, M. M. Casagrande, O. Mielke, T. Pyrcz, R. Robbins, and A. Viloria. 2004. Atlas of Neotropical Lepidoptera. Checklist: Part 4A. Hesperioidea—Papilionoidea. Association for Tropical Lepidoptera / Scientific Publishers, Gainesville, Florida.

Lemey, P., A. Rambaut, A. J. Drummond, and M. A. Suchard. 2009. Bayesian phylogeography finds its roots. PLoS Comput Biol 5:e1000520.

Loh, Y.-H. E., E. Bezault, F. M. Muenzel, R. B. Roberts, R. Swofford, M. Barluenga, C. E. Kidd, A. E. Howe, F. Di Palma, K. Lindblad-Toh, J. Hey, O. Seehausen, W. Salzburger, T. D. Kocher, and J. T. Streelman. 2013. Origins of Shared Genetic Variation in African Cichlids. Molecular Biology and Evolution 30:906-917.

Mallet, J. 2010. Shift happens! Shifting balance and the evolution of diversity in warning colour and mimicry. Ecological Entomology 35:90-104.

Mallet, J., M. Beltrán, W. Neukirchen, and M. Linares. 2007. Natural hybridization in Heliconiine butterflies: the species boundary as a continuum. BMC Evolutionary Biology 7:28.

Mallet, J. and L. Gilbert. 1995. Why are there so many mimicry rings? Correlations between habitat, behaviour and mimicry in *Heliconius* butterflies. Biological Journal of the Linnean Society 55:159-180.

Mallet, J., W. McMillan, and C. Jiggins. 1998. Mimicry and warning colour at the boundary between races and species. Pp. 390-403 *in* D. Howard, and S. Berlocher, eds. Endless Forms: Species and Speciation.

Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27:209-220.

Martin, S. H., K. K. Dasmahapatra, N. J. Nadeau, C. Salazar, J. R. Walters, F. Simpson, M. Blaxter, A. Manica, J. Mallet, and C. D. Jiggins. 2013. Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Research 23:1817-1828.

Mavárez, J., C. Salazar, E. Bermingham, C. Salcedo, C. Jiggins, and M. Linares. 2006. Speciation by hybridization in *Heliconius* butterflies. Nature 441:868-871.

McCormack, J. E., H. Huang, and L. L. Knowles. 2009. Maximum likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. Systematic Biology 58:501-508.

Melo, M., C. Salazar, C. Jiggins, and M. Linares. 2009. Assortative mating preferences among hybrids offers a route to hybrid speciation. Evolution 63:1660-1665.

Mérot, C., J. Mavárez, A. Evin, K. K. Dasmahapatra, J. Mallet, G. Lamas, and M. Joron. 2013. Genetic differentiation without mimicry shift in a pair of hybridizing *Heliconius* species (Lepidoptera: Nymphalidae). Biological Journal of the Linnean Society 109:830-847.

Merrill, R. M., Z. Gompert, L. M. Dembeck, M. R. Kronforst, W. O. McMillan, and C. D. Jiggins. 2011. Mate preference across the speciation continuum in a clade of mimetic butterflies. Evolution 65:1489–1500.

Merrill, R. M., R. W. R. Wallbank, V. Bull, P. C. A. Salazar, J. Mallet, M. Stevens, and C. D. Jiggins. 2012. Disruptive ecological selection on a mating cue. Proceedings of the Royal Society B: Biological Sciences 279:4907-4913.

Meudt, H. M. and A. C. Clarke. 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends in Plant Science 12:106-117.

Müller, F. 1879. *IItuna* and *IThyridia*; a remarkable case of mimicry in butterflies. Transactions of the Entomological Society of London 1879:xx-xxix.

Nadeau, N. J., S. H. Martin, K. M. Kozak, C. Salazar, K. K. Dasmahapatra, J. W. Davey, S. W. Baxter, M. L. Blaxter, J. Mallet, and C. D. Jiggins. 2013. Genome-wide patterns of divergence and gene flow across a butterfly radiation. Molecular Ecology 22:814-826.

Naisbit, R., C. Jiggins, and J. Mallet. 2001. Disruptive sexual selection against hybrids contributes to speciation between *Heliconius cydno* and *Heliconius melpomene*. Proceedings of the Royal Society of London Series B-Biological Sciences 268:1849-1854.

Ochoa, D., C. Hoorn, C. Jaramillo, G. Bayona, M. Parra, and F. De la Parra. 2012. The final phase of tropical lowland conditions in the axial zone of the Eastern Cordillera of Colombia: Evidence from three palynological records. Journal of South American Earth Sciences 39:157-169.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2013. vegan: Community Ecology Package.

Papageorgis, C. 1975. Mimicry in neotropical butterflies. American Scientist 63:522-532.

Pardo-Diaz, C., C. Salazar, S. W. Baxter, C. Merot, W. Figueiredo-Ready, M. Joron, W. O. McMillan, and C. D. Jiggins. 2012. Adaptive Introgression across Species Boundaries in *Heliconius* Butterflies. PLoS Genet 8:e1002752.

Patel, S., J. D. Weckstein, J. S. L. Patané, J. M. Bates, and A. Aleixo. 2011. Temporal and spatial diversification of *Pteroglossus araçaris* (AVES: Ramphastidae) in the neotropics: Constant rate of diversification does not support an increase in radiation during the Pleistocene. Molecular Phylogenetics and Evolution 58:105-115.

Posada, D. and K. A. Crandall. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14:817-818.

Pritchard, J., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Quek, S.-P., B. Counterman, P. Albuquerque de Moura, M. Cardoso, C. Marshall, W. McMillan, and M. Kronforst. 2010. Dissecting comimetic radiations in *Heliconius* reveals divergent histories of convergent butterflies. Proceedings of the National Academy of Sciences 107:7365-7370.

Quek, S.-P., S. J. Davies, T. Itino, and N. E. Pierce. 2004. Codiversification in an ant-plant mutualism: stem texture and the evolution of host use in *Crematogaster* (Formicidae: Myrmicinae) inhabitants of *Macaranga* (Euphorbiaceae). Evolution 58:554-570.

Ravelo, A. C., D. H. Andreasen, M. Lyle, A. Olivarez Lyle, and M. W. Wara. 2004. Regional climate shifts caused by gradual global cooling in the Pliocene epoch. Nature 429:263-267.

Ribas, C. C., A. Aleixo, A. C. R. Nogueira, C. Y. Miyaki, and J. Cracraft. 2012. A palaeobiogeographic model for biotic diversification within Amazonia over the past three million years. Proceedings of the Royal Society B: Biological Sciences 279:681-689.

Rius, M. and J. A. Darling. 2014. How important is intraspecific genetic admixture to the success of colonising populations? Trends in Ecology & Evolution 29:233-242.

Rull, V. 2011. Neotropical biodiversity: timing and potential drivers. Trends in Ecology & Evolution 26:508-513.

Salazar, C., S. Baxter, C. Pardo-Diaz, G. Wu, A. Surridge, M. Linares, E. Bermingham, and C. Jiggins. 2010. Genetic evidence for hybrid trait speciation in *Heliconius* butterflies. PLoS Genetics 6:e1000930.

Salazar, C., C. Jiggins, J. Taylor, M. Kronforst, and M. Linares. 2008. Gene flow and the genealogical history of *Heliconius heurippa*. BMC Evolutionary Biology 8:132.

Santos, J., L. Coloma, K. Summers, J. Caldwell, R. Ree, and D. Cannatella. 2009. Amazonian amphibian diversity is primarily derived from late Miocene Andean lineages. PLoS BIOLOGY 7:1000056.

Schluter, D. and G. L. Conte. 2009. Genetics and ecological speciation. Proceedings of the National Academy of Sciences.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin Ver.2000: A software for population genetics data analysis. Genetics and Biometry Laboratory. University of Geneva. Switzerland.

Sheppard, P., J. Turner , K. Brown, W. Benson, and M. Singer. 1985. Genetics and the evolution of Muellerian mimicry in *Heliconius* butterflies. Philosophical Transactions of the Royal Society B: Biological Sciences 308:433-613.

Silva-Brandão, K. L., N. Wahlberg, R. B. Francini, A. M. L. Azeredo-Espin, K. S. Brown Jr, M. Paluch, D. C. Lees, and A. V. L. Freitas. 2008. Phylogenetic relationships of butterflies of the tribe Acraeini (Lepidoptera, Nymphalidae, Heliconiinae) and the evolution of host plant use. Molecular Phylogenetics and Evolution 46:515-531.

Simonsen, T. J., E. V. Zakharov, M. Djernaes, A. M. Cotton, R. I. Vane-Wright, and F. A. H. Sperling. 2011. Phylogenetics and divergence times of Papilioninae (Lepidoptera) with special reference to the enigmatic genera Teinopalpus and Meandrusa. Cladistics 27:113-137.

Swofford, D. 2000. PAUP*. Phylogenetic analysis using parsimony (*and other methods).

Tajima, F. 1989. Stadistical method for testing the neutral mutation hypothesis by DNA polymorhism. Genetics 123:585-595.

The *Heliconius* genome consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94-98.

Turner, J. 1976. Adaptive radiation and convergence in subdivisions of the butterfly genus *Heliconius* (Lepidoptera: Nymphalidae). Zoological Journal of the Linnean Society 58:297-308.

Turner, J. and J. Mallet. 1996. Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting bance. Philosophical Transactions of the Royal Society B: Biological Sciences 351:835-845.

Turner, J. R. G. 1981. Adaptation and evolution in *Heliconius*: a defense of neo-Darwinism. Annual Review of Ecology and Systematics 12:99-121.

Turner, J. R. G., M. S. Johnson, and W. F. Eanes. 1979. Contrasted modes of evolution in the same genome: Allozymes and adaptive change in *Heliconius*. Proceedings of the National Academy of Sciences USA 76:1924-1928.

Van der Hammen, T., J. H. Werner, and H. Van Dommelen. 1973. Palynological record of the upheaval of the Northern Andes: A study of the Pliocene and Lower Quaternary of the Colombian Eastern Cordillera and the early evolution of its High-Andean biota. Review of Palaeobotany and Palynology 16.

Vos, P., R. Hogers, M. Bleeker, M. Rijans, T. Van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23:4407-4414.

Wahlberg, N. and A. V. L. Freitas. 2007. Colonization of and radiation in South America by butterflies in the subtribe Phyciodina (Lepidoptera: Nymphalidae). Molecular Phylogenetics and Evolution 44:1257-1272.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Whinnett, A., M. Zimmermann, K. Willmot, N. Herrera, R. Mallarino, F. Simpson, M. Joron, G. Lamas, and J. Mallet. 2005. Strikingly variable divergence times inferred across an Amazonian butterfly 'suture zone'. Proceedings of the Royal Society B: Biological Sciences 272:2525-2533.

Data access

Co1 "barcode" sequences have been deposited in GenBank (see Table 4SM). AFLP data, *Co1* alignment and additional data files have been submitted to DRYAD entry doi: 10.5061/dryad.b5d6b.

Author Contributions Box

C.F.A, E.B and W.O.M designed the project. C.F.A and C.R generated all molecular data and C.F.A, C. S and W.O.M analyzed the data. M.R.K and M.L contributed specimens and aided in interpretation of the results. C.F.A and W.O.M wrote the manuscript with input from C.S, M.L, M.R.K and E.B.

Figure 1. *Heliconius cydno* radiation and closely related taxa. The range of the three radiations, *H. cydno*, *H. melpomene*, and *H. timareta* encompasses much of tropical Central and South America. The map presents sample localities for *H. cydno*, *H. melpomene*, *H. timareta* and *H. heurippa*. Pink *H. cydno* races, Blue *H. melpomene* races, and Green *H. timareta* and *H. heurippa*. An asterisk denotes the only *H. cydno* races not collected in this study: *H. c. gadouae* distributed on the Northeast slopes of the Andes.

Figure 2. Phylogenetic trees for *Heliconius cydno* and related species. a) Bayesian phylogenetic reconstruction for the "barcode" mtDNA region. Branches on the tree are color coded with the maximum credibility state reconstruction from BEAST v1.6. For deeper nodes, pies are presented with the posterior probabilities for each state at that particular node. Numbers over branches represent posterior probabilities for branches with a probability greater than 50%. b) Neighbor-joining phylogenetic reconstruction for AFLP markers. Numbers over branches represent bootstrap support for branches with a probability greater than 50%. Color code follows the geographic regions previously defined, Yellow, Northeastern Andes (EAN); Green, Central America (CA); Dark Brown, Northern Magdalena valley (MGN); Pale Brown, Southern Magdalena valley (MGS); Bright Blue, Northern Cauca valley (CVN); Pale Blue, Southern Cauca valley (CVS); Pink, Pacific West Andes (PW); Purple, French Guiana (FG); Dark Blue, Brazil (BA); and Orange, Southeastern Andes (EAS). The EAS clade represents races from the Southeastern Andes and is formed by individuals of *H. m. malleti*, *H. m. ecuadorensis*, *H. m. plesseni* and *H. m. bellula* from the Colombian East Andes. Location of *H. melpomene* and *H. timareta* races are presented in Figure 1.

Figure 3. Population assignment tests for *H. melpomene*, *H. cydno* and *H. timareta*. Bar plots showing Bayesian assignment probabilities for three clusters from the software Structure 2.1.4. Each vertical bar corresponds to one individual. The proportion of color on the bar represents an individual's assignment probability to the different clusters. Black-horizontal bars on the top and bottom of figure show phenotypic classification. Gray-horizontal bars on the bottom of figure follows the geographic regions: Central America (CA), Pacific West Andes (PW), Cauca valley (CV), Magdalena valley (MG), Northeastern Andes (EAN), Southeastern Andes (EAS), French Guiana (FG) and Brazil (BA).

Figure 4. Discriminant Analysis of Principal Components (DAPC) for *H. cydno* races. The squares show the number of individuals assigned to the different genotypic clusters and its phenotypic classification using the function *find.clusters* (R package: *adegenet*; Jombart *et al.* 2010). The color of the squares corresponds to the geographic regions displayed on the map: Yellow, Northeastern Andes (EAN); Green, Central America (CA); Dark Brown, Northern Magdalena valley (MGN); Pale Brown, Southern Magdalena valley (MGS); Bright Blue, Northern Cauca valley (CVN); Pale Blue, Southern Cauca valley (CVS); and Pink, Pacific West Andes (PW).