Evolution of mimicry genes in butterflies: a case study of Heliconius

Research

http://dx.doi.org/10.1098/rsbl.2013.0503

Received: 31 May 2013
Accepted: 24 June 2013

Subject Areas:
- evolution

Keywords:
- Heliconius, hybridization, introgression, mimicry

Author for correspondence:
Marcus R. Kronforst
e-mail: mkronforst@uchicago.edu


Evolutionary biology

Do Heliconius butterfly species exchange mimicry alleles?

Joel Smith and Marcus R. Kronforst

Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637, USA

Hybridization has the potential to transfer beneficial alleles across species boundaries, and there are a growing number of examples in which this has apparently occurred. Recent studies suggest that Heliconius butterflies have transferred wing pattern mimicry alleles between species via hybridization, but ancestral polymorphism could also produce a signature of shared ancestry around mimicry genes. To distinguish between these alternative hypotheses, we measured DNA sequence divergence around putatively introgressed mimicry loci and compared this with the rest of the genome. Our results reveal that putatively introgressed regions show strongly reduced sequence divergence between co-mimetic species, suggesting that their divergence times are younger than the rest of the genome. This is consistent with introgression and not ancestral variation. We further show that this signature of introgression occurs at sites throughout the genome, not just around mimicry genes.

1. Introduction

Genetic exchange between species is gaining ground as a potentially important source of variation for adaptation and speciation [1–3]. A recent paper by the Heliconius Genome Consortium [4] demonstrated that Heliconius butterfly species with similar mimetic wing patterns exhibit shared ancestry around wing patterning loci, suggesting that hybridization and introgression have transferred mimicry alleles from one species to another. The results of a companion study further support this conclusion [5]. However, there are two entirely distinct phenomena that could produce a signal of shared ancestry around mimicry loci; introgression (figure 1a) and shared ancestral polymorphism (figure 1b). While trans-species polymorphisms owing to shared ancestral variation are widespread in nature [6], this alternative has not been carefully considered in the case of Heliconius mimicry.

The D-statistic used by the HGC to infer introgression is capable of distinguishing biased allele sharing from random sorting of ancestral polymorphism [7]. This test statistic uses nucleotide sites that fall into two categories: ‘ABBA’ and ‘BABA’, where each letter refers to an allele in each of four taxa. ‘A’ refers to the outgroup, or ancestral allele, and ‘B’ refers to the derived allele (see electronic supplementary material). ABBA and BABA configurations should occur with equal frequency in cases where there has been no admixture between taxa and ancestral populations mate randomly. When tallied across the genome, skew towards ABBA or BABA sites would indicate a systematic bias in allele sharing between taxa. In cases where allele sharing is enriched between sympatric taxa, introgression is a potential explanation.

However, ancestral polymorphism could interact with selection to give the false appearance of introgression in a mimicry system. If an ancestral species were polymorphic for two mimetic wing patterns, which it then passed to each of two descendant species, subsequent directional selection for local mimicry could result in co-occurring sister species with shared mimicry phenotypes controlled by homologous haplotypes. As an example, polymorphism in the ancestor could be the product of co-occurring mimicry models or co-mimics,
a phenomenon well known to promote intraspecific mimetic polymorphism [8,9]. Subsequent divergent selection could then simply be a product of shifting ranges for one or both of the models/co-mimics. This scenario is analogous to well-studied molecular trans-species polymorphisms such as MHC alleles [6] and the ABO blood group [10], except that variation is locally monomorphic today as a result of selection for mimicry.

This alternative explanation could account for a variety of HGC results, such as localized elevation of the D-statistic around mimicry loci and gene trees that group taxa by mimicry phenotype as opposed to species (but it is unlikely to explain the genome-wide elevation in D noted between sympatric species [4]). The fact that ancestral polymorphism has not been considered as a potential explanation for shared variation around mimicry genes has been controversial, with subsequent published work [11] and blog discussions (http://gcbias.org/2012/05/23/journal-tea-may-21st/) calling for clarification. Heliconius butterflies tend to be locally monomorphic so it is difficult to imagine how a wing pattern polymorphism could be maintained through a speciation event. However, there are multiple examples of local polymorphism in Heliconius [12–14], including one of the focal species here, Heliconius timareta, which has as many as four coexisting phenotypes in Ecuador [14]. Here, we specifically test the hypothesis that shared mimicry phenotypes are due to ancestral variation, as opposed to introgression, by examining DNA sequence divergence in putatively introgressed regions of the genome.

Figure 1. Sequence divergence distinguishes between ancestral polymorphism and introgression as the source of shared haplotypes between species. (a) Introgression predicts young divergence times between shared haplotypes resulting in (b) reduced sequence divergence between species, as compared with the genomic background. (c) Ancestral polymorphism predicts older divergence times resulting in (d) greater sequence divergence between species.

2. Material and methods
We reanalysed the HGC data, focusing on DNA sequence divergence between species. The data are deposited in the European Nucleotide Archive under accession nos. ERP000993 and ERP000991.

Our approach considers the impact of introgression versus ancestral polymorphism on sequence divergence between shared alleles, as compared with the genomic background. Introgression should result in recent splitting of alleles, and low sequence divergence, compared with the genomic background (figure 1c). By contrast, ancestral polymorphism should result in more ancient splitting of alleles and greater sequence divergence (figure 1d). Note that a third potential explanation, convergent molecular evolution, was excluded by HGC analyses [4] but would result in mimicry allele divergence times that match the genomic background.

The HGC data consist of targeted resequencing data around two mimicry loci (B/D and N/Yb) for two Heliconius melpomene/
H. timareta comparisons, co-mimetic Heliconius melpomene amaryllis and Heliconius timareta ssp. nov. in Peru and Heliconius melpomene aglaope and Heliconius timareta florencia in Peru and Colombia (see electronic supplementary material). There are also targeted resequencing data for outgroup taxa as well as genome-wide RAD data for all taxa except H. t. florencia. For our analyses, we focused on sympatric H. m. amaryllis and H. timareta in Peru, using H. m. aglaope as an allopatric comparison, because these are the samples for which both resequencing and RAD data exist. Each ingroup taxon consists of four individuals in the resequencing datasets and five individuals in the RAD dataset. The silvaniform outgroup includes one individual from each of the following taxa: Heliconius hecale, Heliconius numata silvana, Heliconius ethilla, Heliconius pardalinus sergestus and Heliconius pardalus ssp. nov. Because the D-statistic depends on identifying derived alleles that are shared between taxa, the inclusion of multiple taxa in the outgroup provides additional confidence in identifying ancestral alleles.

To estimate allele frequencies, we used biallelic sites having a GATK [15] quality score greater than 30 (99.9% accuracy). In accordance with the HGCI analysis, only sites with at least 50% of the genotypes available for each of the four groups were used. Alleles were polarized with respect to the outgroup major allele, and sites with an outgroup frequency of 50% were excluded. After filtering, the B/D, N/Yb and RAD datasets contained 55847, 79549 and 142869 SNPs, respectively.

We calculated Patterson’s D-statistic in non-overlapping 5 kbp windows across the B/D (717 kbp) and the N/Yb (1.15 Mbp) scaffolds from the H. melpomene reference genome. We used the following four group comparison to calculate D: H. m. aglaope, H. m. amaryllis, H. t. ssp. nov., silvaniform outgroup (see electronic supplementary material). Positive D outlier windows are indicative of allele sharing between sympatric H. m. amaryllis and H. timareta in Peru. We used a similar approach for the rest of the genome, calculating D and divergence based on RAD data. For the genome-wide analysis we excluded the B/D and N/Yb regions.

Finally, we compared mean DNA sequence divergence (d_{XY}) between the top 10% D outliers (positive values) and the remaining windows in each region. To estimate d_{XY}, we calculated divergence for each SNP based on allele frequencies, summed across SNPs, and divided by total bp in the window (see electronic supplementary material). Note that for RAD data, sequenced loci were clustered into windows based on variants being within 5 kbp of each other, and divergence was calculated by dividing by total bp in the window. For this reason, divergence estimates in figure 2c are artificially low but internally consistent. We used Welch’s t-test to compare divergence between outlier and background intervals.

3. Results

When we considered the putatively introgressed regions around two wing patterning loci, the B/D locus (which controls red patterning) and the N/Yb locus (which controls yellow patterning), we found distinct signals of reduced divergence consistent with introgression (figure 2a,b). In addition, results from the HGCI suggested that introgression may also occur in parts of the genome other than these two wing patterning loci [4]. We scanned the genome and identified widespread signatures of elevated D consistent with genome-wide introgression between H. m. amaryllis and H. timareta in Peru. When we examined sequence divergence in these additional regions, excluding the B/D and N/Yb loci, we again found reduced divergence indicative of introgression (figure 2c).

The results reveal an additional pattern consistent with introgression. The HGCI analyses suggest that mimicry introgression may have also occurred between H. m. aglaope and H. timareta in Colombia [4]. Furthermore, directionality of introgression is suspected to be from H. melpomene into H. timareta [5]. If so, then our comparison between H. m. aglaope and Peruvian timareta, in putatively introgressed regions of the genome (middle white bars in figure 2), effectively should be a comparison between aglaope and amaryllis. The fact that this level of divergence is similar to H. m. aglaope and H. m. amaryllis background divergence (right-most bars in figure 2) lends additional support to the introgression hypothesis.

![Figure 2. Sequence divergence is reduced between co-mimetic Heliconius melpomene amaryllis and Heliconius timareta at (a) the B/D mimicry locus, (b) the N/Yb mimicry locus and (c) genome-wide regions of elevated Patterson’s D-statistic. ***p < 0.001, **p < 0.01, *p < 0.05, #p > 0.05.](http://rsbl.royalsocietypublishing.org/)}
4. Discussion

Hybridization is widespread among Heliconius species [16], and previous work suggests that this may result in interspecific gene flow ([17–19]; but see [11,20]). This, combined with examples of mimicry between species known to hybridize [21,22], makes the mimicry introgression hypothesis appealing [23]. However, the alternative hypothesis of trans-species polymorphism owing to shared ancestral variation must also be considered. This alternative explanation is relevant when it comes to Heliconius wing patterns, because H. melpomene and H. timareta are closely related, having diverged 1–1.5 Ma [19], and both species are polymorphic for the same wing pattern variation, suggesting their common ancestor may have been so as well.

Despite these features that make ancestral polymorphism plausible, our results ultimately support the hypothesis that introgression has moved mimicry alleles between Heliconius species [4,5,23]. By comparing two alternative hypotheses, we are able to substantiate this previously tentative conclusion. Furthermore, given the widespread signatures of introgression found across the genome, it is quite likely that interspecific gene flow has impacted aspects of Heliconius biology beyond wing patterning.

We thank Kanchon Dasmathapatra and Jim Mallet for sharing data and for their comments on a previous version of the manuscript. We also thank Wei Zhang and Grace Lee for assistance with the analysis and reviewers for comments on previous versions of the manuscript. This research was supported by NSF grant no. DEB-1316037 to M.R.K.

References