A single origin of Batesian mimicry among hybridizing populations of admiral butterflies (Limenitis arthemis) rejects an evolutionary reversion to the ancestral phenotype

Wesley K. Savage* and Sean P. Mullen

Department of Biological Sciences, Lehigh University, B217 Iacocca Hall, 111 Research Drive, Bethlehem, PA 18015–4732, USA

Batesian mimicry is a fundamental example of adaptive phenotypic evolution driven by strong natural selection. Given the potentially dramatic impacts of selection on individual fitness, it is important to understand the conditions under which mimicry is maintained versus lost. Although much empirical and theoretical work has been devoted to the maintenance of Batesian mimicry, there are no conclusive examples of its loss in natural populations. Recently, it has been proposed that non-mimetic populations of the polytypic Limenitis arthemis species complex represent an evolutionary loss of Batesian mimicry, and a reversion to the ancestral phenotype. Here, we evaluate this conclusion using segregating amplified fragment length polymorphism markers to investigate the history and fate of mimicry among forms of the L. arthemis complex and closely related Nearctic Limenitis species. In contrast to the previous finding, our results support a single origin of mimicry within the L. arthemis complex and the retention of the ancestral white-banded form in non-mimetic populations. Our finding is based on a genome-wide sampling approach to phylogeny reconstruction that highlights the challenges associated with inferring the evolutionary relationships among recently diverged species or populations (i.e. incomplete lineage sorting, introgressive hybridization and/or selection).

Keywords: wing pattern evolution; mimicry; amplified fragment length polymorphism; Limenitis; phylogeny; gene flow

1. INTRODUCTION

Batesian mimicry is a classic example of adaptation because the mimic gains a direct fitness advantage (e.g. reduced predation) due to its phenotypic resemblance to the protected model. Although the relationship between a Batesian mimic and its model can vary both temporally (Waldbauer 1988) and spatially (Ritland & Brower 2000; Ries & Mullen 2008), in the simplest iteration, the fitness of a palatable Batesian mimic is dependent on the frequency of a chemically defended, and typically aposmotic, model (Bates 1862; Fisher 1930; reviewed by Mallet & Joron 1999).

Under frequency-dependent selection (Fisher 1930; Ayala & Campbell 1974), protection from predation for Batesian mimics is expected to break down (i) in the absence of the model (Wallace 1870; Waldbauer & Sternburg 1987; Waldbauer 1988; Pfennig et al. 2001, 2007) or (ii) when the mimic becomes abundant relative to the model (Fisher 1930; Brower & Brower 1962; Huheey 1964; Oaten et al. 1975; Getty 1985; Mallet & Joron 1999; Harper & Pfennig 2007). Although numerous laboratory (Brower 1960; Nonacs 1985; Lindström et al. 2004; Rowland et al. 2007) and theoretical (Huheey 1976, 1988; Ruxton et al. 2004; Mappes et al. 2005) studies support these predictions, empirical examples of mimetic breakdown in natural populations are relatively rare (but see Harper & Pfennig 2007, 2008; Joron 2008).

One putative example of mimetic breakdown is represented by the phenotypic transition zone between mimetic and non-mimetic admiral butterfly populations, which comprise the polytypic Limenitis arthemis complex. White-banded populations of these butterflies (Limenitis arthemis arthemis and Limenitis arthemis rubrofasciata) range widely from northeastern to northwestern North America, as far west as Alaska (Scott 1986), and closely resemble the ancestral wing pattern phenotype characteristic of the entire genus (Mullen 2006). By contrast, red-spotted purples (Limenitis arthemis arizonensis and Limenitis arthemis astyanax) are allopatrically distributed in southwestern North America and the southeastern United States, and are well-known Batesian mimics of the chemically defended Pipevine swallowtail (Battus philenor, Brower & Brower 1962; Platt et al. 1971; Platt 1975; Mullen 2006). A recent study on the evolutionary relationships among the North American representatives of this genus has tested the hypothesis that an adaptive Batesian mimicry phenotype was lost in white-banded populations of the L. arthemis species complex (Prudic & Oliver 2008). The results of this study posit that selection for mimicry breaks down outside the range of the model (Batts), and consequently that white-banded populations of L. arthemis are the result of an evolutionary reversal by loss of the mimetic phenotype (figure 1).

If true, this finding represents the only documented example (in butterflies) of a reversion to the ancestral phenotype because of selection against mimicry. Although several recent studies have found evidence for a
breakdown of selection favouring Batesian mimicry phenotypes in allopatry (with respect to the model; Harper & Pfennig 2007, 2008; Pfennig et al. 2007), none of these studies explicitly addressed the origin of the allopatric mimic populations. Allopatric Batesian mimics could arise via several mechanisms, including range expansion of the mimic, range contraction of the model or gene flow into non-mimetic allopatric populations. Thus, although there is clear evidence that selection for mimicry breaks down outside of the range of the Battus model in this system (Prudic & Oliver 2008; Ries & Mullen 2008), there are several reasons to question the conclusion that the absence of mimicry in white-banded populations represents an evolutionary loss of mimicry.

First, if the absence of mimicry in white-banded populations of *L. a. arthemis* was the result of population expansion to regions outside the range of the model, then there is no *a priori* reason to expect evidence of genetic differentiation between mimetic and non-mimetic populations except for those regions of the genome controlling wing pattern. This is not the case because significant population structure corresponding to wing pattern phenotypes exists in this complex (Mullen 2006; Mullen et al. 2008; Prudic & Oliver 2008), consistent with the hypothesis that the hybrid zone between mimetic (*L. a. astyanax*) and non-mimetic (*L. a. arthemis*) subspecies is the result of secondary contact between two evolutionarily distinct lineages that diverged in allopatry (Mullen et al. 2008). Second, significant levels of both contemporary and historical gene flow have occurred between these two wing pattern forms, indicating that, although partially isolated, for many loci individual gene tree topologies will not accurately represent the true ‘species’ history (Pamilo & Nei 1988; Maddison 1997; Nichols 2001). Third, despite its wide use in phylogeny reconstruction, mitochondrial DNA (mtDNA) is notoriously unreliable when used to infer the history of a phenotypic trait among recently diverged hybridizing taxa (Ballard & Whitlock 2004). Finally, the conclusion that mimicry was lost in this complex was largely based on (i) a species tree estimated from mtDNA, and (ii) the assumption that no mtDNA introgression occurs between mimetic and non-mimetic phenotypes (figure 1a; Prudic & Oliver 2008). However, an equally parsimonious interpretation is that mimicry evolved independently to give rise to both mimetic subspecies in this complex and was not lost in *L. a. arthemis* (figure 1b). Alternatively, it is possible that mimicry evolved once and the mimetic lineages diverged in allopatry (figure 1c).

Given that inferences about wing pattern evolution in this group so directly hinge upon a robust reconstruction of the relationships among the three subspecies, we used amplified fragment length polymorphism (AFLP) markers to assay genetic variation at a genome-wide scale. AFLP fingerprinting often provides resolution in molecular phylogenies where nucleotide sequence data fail to recover relationships in recently evolved taxa (Sullivan et al. 2004; Bonin et al. 2005; Mendelson & Shaw 2005; Mendelson et al. 2005; Althoff et al. 2007; Meudt & Clarke 2007; Holland et al. 2008). Here, we present a phylogeny of the North American species of *Limenitis*, which indicates that mimicry evolved once in the *L. arthemis* complex, and that the hybrid zone between mimetic and non-mimetic populations is the result of secondary contact between two distinct lineages, rather than the loss of mimicry due to range expansion and primary differentiation.

2. MATERIAL AND METHODS

(a) Taxon sampling

Because the species-level phylogeny of this genus has previously been published (Mullen 2006), this study focused specifically on the unresolved relationships among mimetic and white-banded phenotypes within the polytypic *L. arthemis* species complex. Previous work also indicates ongoing gene flow among subspecies within this group (Mullen et al. 2008), so we restricted our phylogenetic sampling of the two eastern subspecies (*L. a. arthemis* and

Proc. R. Soc. B (2009)
Table 1. Specimen information for eight wing pattern races of *Limenitis* examined in this study. (Listed by taxon is the sample size (*N*), collection locality, collection date, phenotype and the pattern of mimicry (with the model species).)

<table>
<thead>
<tr>
<th>species</th>
<th><em>N</em></th>
<th>locality</th>
<th>date</th>
<th>wing pattern</th>
<th>mimicry pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>L. arthemis</em> group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>L. a. arizonensis</em></td>
<td>3</td>
<td>Arizona, USA</td>
<td>2001</td>
<td>black</td>
<td>Batesian (<em>Battus philenor</em>)</td>
</tr>
<tr>
<td><em>L. a. arthemis</em></td>
<td>4</td>
<td>New Brunswick, CA</td>
<td>2004</td>
<td>white-banded</td>
<td>non-mimetic</td>
</tr>
<tr>
<td><em>L. a. astyanax</em></td>
<td>4</td>
<td>Mississippi, USA</td>
<td>2001</td>
<td>black</td>
<td>Batesian (<em>Battus philenor</em>)</td>
</tr>
<tr>
<td><em>Limenitis lorquini</em></td>
<td>4</td>
<td>British Columbia, CA</td>
<td>2002</td>
<td>white-banded, orange tip</td>
<td>Batesian (<em>Adelphia</em>)</td>
</tr>
<tr>
<td><em>Limenitis weidemeyerii</em></td>
<td>3</td>
<td>South Dakota, USA</td>
<td>2003</td>
<td>white-banded</td>
<td>non-mimetic</td>
</tr>
<tr>
<td><em>Limenitis archippus lahontani</em></td>
<td>4</td>
<td>Nebraska, USA</td>
<td>2003</td>
<td>orange</td>
<td>Müllerian (<em>Danaus</em>)</td>
</tr>
<tr>
<td><em>Limenitis populi</em></td>
<td>2</td>
<td>Sweden</td>
<td>2002</td>
<td>white-banded</td>
<td>non-mimetic</td>
</tr>
<tr>
<td><em>Limenitis reducta</em></td>
<td>2</td>
<td>France</td>
<td>2002</td>
<td>white-banded</td>
<td>non-mimetic</td>
</tr>
</tbody>
</table>

Table 2. Summary statistics derived from eight AFLP primer pair combinations for eight *Limenitis* species. (The number of fragments and the genetic differentiation estimates (*F*_ST) are listed along with the percentage of polymorphic bands for each primer pair combination. The percentage of polymorphic bands is also listed for each taxon at each primer pair combination.)

<table>
<thead>
<tr>
<th>EcoRI-MseI primers</th>
<th>5-FAM-5′</th>
<th>NED-5′</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACT-CTC</td>
<td>ACT-CTG</td>
</tr>
<tr>
<td>number of fragments</td>
<td>342</td>
<td>277</td>
</tr>
<tr>
<td><em>F</em>_ST</td>
<td>0.12</td>
<td>0.15</td>
</tr>
<tr>
<td>percent polymorphic</td>
<td>86.3</td>
<td>79.1</td>
</tr>
</tbody>
</table>

*L. a. astyanax* to populations well outside the region where the two subspecies are known to hybridize (*Mullen et al. 2008; Ries & Mullen 2008*); the third subspecies, *L. a. arizonensis*, is reproductively isolated de facto due to its range allopatry.

In total, we included eight *Limenitis* taxa, three of which are contained in the *L. arthemis* complex. Each taxon is represented by at least two specimens from a single locality. The total dataset consists of 24 *Limenitis* specimens, out of which 20 represent six North American wing pattern races (i.e. species and subspecies) and four Palaearctic representatives of admiral butterflies (*Limenitis*). The four Palaearctic specimens were used for Nearctic–Palaearctic out-group comparison/character polarization. Our geographical sampling and total sample sizes are restricted because three recent works have thoroughly established the weak geographical population structure of each wing pattern race, and that sampling large numbers of individuals from throughout the range does not improve resolution (*Mullen 2006; Mullen et al. 2008; Prudic & Oliver 2008*). In this study, we are conducting a genome-wide analysis of mimicry evolution, not of phylogeographic relationships. Specimens, locality information and wing pattern phenotypes are listed in table 1.

(b) AFLP methods

Genomic DNA was isolated from flight muscle tissue using a commercially available extraction kit (DNeasy Tissue Kit, Qiagen). DNA concentrations were standardized by spectrophotometry for subsequent molecular applications. We used the AFLP Plant Mapping Kit (Applied Biosystems) and a standard protocol (*Vos et al. 1995; Bonin et al. 2005*) to genotype the 24 specimens with eight EcoRI–MseI selective primer combinations (table 2). The basic steps were performed as follows: (i) restriction enzyme digestion of genomic DNA and ligation with EcoRI and MseI adaptors, (ii) dilution of restriction–ligation products and preselective amplification with primers containing single nucleotide additions, and (iii) selective amplification with three nucleotide additions to the fluorescently labelled 5′ oligonucleotides. The AFLP fragments were sized with ROX-500 (~250) standard and capillary electrophoresis on an ABI Prism 3730 DNA analyser (Applied Biosystems), and scored by the absence/presence of peaks using GENEMAPPER software v. 3.7 (Applied Biosystems). Fragments were scored according to GENEMAPPER settings suggested by *Holland et al. (2008)*, using bin widths of 0.5 but raising the minimum relative fluorescent units value to 100 to exclude incorrectly assigned peaks that result from background noise and stutter. Fragments were scored in the 100–400 bp range; raising the minimum size threshold was recently suggested as a step that minimizes homoplasy that is more prevalent in smaller fragments (*Althoff et al. 2007*).
(c) Data analysis

The presence–absence of peaks for each AFLP primer pair was converted into a binary data matrix (0 or 1) for all individuals. We used AFLP-SURV v. 1.0 (Vekemans 2002) to calculate genetic diversity and basic summary statistics for each of the eight taxa. Genetic diversity was estimated with Zhivotovsky’s (1999) Bayesian method of computation with a non-uniform prior distribution.

We analysed the binary matrix with four different phylogenetic methods: tree reconstructions using distance; neighbor-joining (NJ); and parsimony were calculated in PAUP* v. 4.0b10 (Swofford 2000); and Bayesian inference (BI) was performed using MrBayes v. 3.1.2 (Ronquist & Huelsenbeck 2003). The restriction-site distance measure of Nei & Li (1979) was used for computing distance and NJ trees with the minimum evolution objective function. Clade support was evaluated by 5000 bootstrap pseudoreplicates for both distance (full heuristic search) and NJ methods, both under the Nei & Li (1979) distance measure. Parsimony estimation was performed under default assumptions and settings for the binary character state matrix, and clade support was also estimated with 5000 bootstrap pseudoreplicates.

BI was conducted assuming the F81-like restriction site model implemented in MrBayes with the noobsencesites coding bias correction (Ronquist et al. 2005; Koopman et al. 2008). The two-state model implemented in MrBayes understates the complex genetic processes of AFLP evolution, and thus is less likely than distance and NJ methods to provide accurate inference of phylogeny (Althoff et al. 2007; Luo et al. 2007; Holland et al. 2008). A new Bayesian model using AFLP marker lengths was recently described by Luo et al. (2007), but it has yet to overcome the computational challenges associated with large datasets (it takes weeks to analyse a small dataset; Luo et al. 2007; Koopman et al. 2008).

The Dirichlet prior for the state frequencies set to (7.3, 1.0) to analyse a small dataset; Luo (2007), but it has yet to overcome the computational challenges associated with large datasets (it takes weeks to analyse a small dataset; Luo et al. 2007; Koopman et al. 2008).

BI was conducted assuming the F81-like restriction site model implemented in MrBayes with the noobsencesites coding bias correction (Ronquist et al. 2005; Koopman et al. 2008). The two-state model implemented in MrBayes understates the complex genetic processes of AFLP evolution, and thus is less likely than distance and NJ methods to provide accurate inference of phylogeny (Althoff et al. 2007; Luo et al. 2007; Holland et al. 2008). A new Bayesian model using AFLP marker lengths was recently described by Luo et al. (2007), but it has yet to overcome the computational challenges associated with large datasets (it takes weeks to analyse a small dataset; Luo et al. 2007; Koopman et al. 2008).

BI was conducted assuming the F81-like restriction site model implemented in MrBayes with the noobsencesites coding bias correction (Ronquist et al. 2005; Koopman et al. 2008). The two-state model implemented in MrBayes understates the complex genetic processes of AFLP evolution, and thus is less likely than distance and NJ methods to provide accurate inference of phylogeny (Althoff et al. 2007; Luo et al. 2007; Holland et al. 2008). A new Bayesian model using AFLP marker lengths was recently described by Luo et al. (2007), but it has yet to overcome the computational challenges associated with large datasets (it takes weeks to analyse a small dataset; Luo et al. 2007; Koopman et al. 2008).

BI was conducted assuming the F81-like restriction site model implemented in MrBayes with the noobsencesites coding bias correction (Ronquist et al. 2005; Koopman et al. 2008). The two-state model implemented in MrBayes understates the complex genetic processes of AFLP evolution, and thus is less likely than distance and NJ methods to provide accurate inference of phylogeny (Althoff et al. 2007; Luo et al. 2007; Holland et al. 2008). A new Bayesian model using AFLP marker lengths was recently described by Luo et al. (2007), but it has yet to overcome the computational challenges associated with large datasets (it takes weeks to analyse a small dataset; Luo et al. 2007; Koopman et al. 2008).

3. RESULTS

(a) AFLP variability

The eight AFLP primer pairs yielded 1875 segregating fragments out of 2262 total amplified bands; for each primer pair, the number of amplified fragments ranged from 214 to 342. The total number of polymorphic fragments per taxon ranged from a low of 356 in Limenitis reducta to a high of 682 in Limenitis lorquini (table 2). Total gene diversity for the complete taxon set was 0.133, and overall $F_{ST}$ was 0.184. Nei’s genetic distances ranged from 0.008 to 0.012 within the L. arthemis complex, with the maximum genetic distance between Limenitis weidemeyerii and L. reducta (0.063; data not shown). Primer pair and per taxon results of AFLP are summarized in table 2. The AFLP distances are, as expected, much greater between Limenitis taxa, but are similar in relative differences to the uncorrected $p$-distances for mitochondrial and nuclear sequence data between taxa (table 3; Mullen 2006).

(b) AFLP phylogeny

All four phylogenetic methods (distance, NJ, parsimony and BI) based on the AFLP data recovered the same general topology illustrated in figure 2. The major result of these methods indicates that (i) named Limenitis species and subspecies all form monophyletic lineages, and (ii) the two mimetic lineages of L. arthemis (L. a. astyanax and L. a. arizonensis) form a clade that is sister to the white-banded form (L. a. arthemis). For the first of these results, nodes are strongly supported under all four methods of phylogenetic reconstruction with high bootstrap (distance, NJ and parsimony) and clade credibility values (BI). The second result is moderately supported under both distance-based methods, which are widely used methods with dominant data-like AFLPs (e.g. Mendelson et al. 2005), but poor support under the parsimony and BI analyses. However, both of these methods did recover the same topological relationship of the L. arthemis complex: (L. a. arthemis (L. a. astyanax and L. a. arizonensis)), and the weaker support is not unexpected given the poor fit of these two models for AFLP data (Althoff et al. 2007; Luo et al. 2007; Koopman et al. 2008) and the evidence for hybridization between the two eastern wing pattern races (Mullen et al. 2008).

Figure 2 illustrates the general AFLP-based topology of the Limenitis taxa sampled here, as well as the lack of resolution at deeper phylogenetic depths among Limenitis lineages. While distance and NJ methods resolved the same topologies and similar bootstrap support values among lineages within the Nearctic Limenitis group, none of the four phylogenetic methods resolved the Nearctic Limenitis as a monophyletic clade (yellow node, figure 2); however, this relationship is well supported by previous phylogenetic analyses that used mitochondrial and nuclear DNA sequence data (see Mullen 2006; Prudic & Oliver 2008). Similarly, our AFLP phylogeny did not resolve the relationship between the Palaearctic taxa (Limenitis populi and L. reducta) and the Nearctic Limenitis taxa, but previous work by Mullen (2006) and Prudic & Oliver (2008) suggests that L. populi shared an ancestor with the Nearctic species more recently than L. reducta. Among the Nearctic taxa, each species is recovered as a strongly supported monophyletic lineage, but the relationships among L. lorquini, L. weidemeyerii, L. archippus and L. arthemis were unresolved. This is consistent with previous efforts to reconstruct the relationships among the Nearctic taxa (Mullen 2006; Prudic & Oliver 2008), which each failed to strongly resolve the pattern of speciation in this group.

In general, nodes supported by distance and NJ bootstrap pseudoreplicates based on the AFLP data had lower support in parsimony and Bayesian methods. Specifically, these latter two methods both recovered the focal mimetic clade (L. a. astyanax and L. a. arizonensis), but with lower clade credibility and bootstrap values (0.64 and 56%, respectively) relative to the higher bootstrap support values obtained by distance and NJ analyses (86 and 77%, respectively). Thus, two phylogenetic methods recover moderate support for a mimetic clade of L. arthemis, and two provide weaker inference but still recover the same topology.
4. DISCUSSION

Recently, it has been proposed that the absence of mimicry among white-banded populations of L. arthemis represents the evolutionary loss of mimicry, which would be the first documented example of a reversion to an ancestral wing pattern phenotype (Prudic & Oliver 2008). If true, this conclusion has important implications for the maintenance of Batesian mimicry and the flexibility of the genome to respond to differential environmental pressures (i.e. selection). However, as Barton & Hewitt (1985) first pointed out, it is important to distinguish between the history of a phenotypic trait (e.g. mimicry) and the history of populations exhibiting that trait. Because the former will often be inaccessible due to the confounding effects of reticulate population histories and stochastic evolutionary and ecological processes, the resolution of how many times mimicry arose and/or was lost in the L. arthemis complex remains elusive.

The results of this study, based on a large dataset consisting of a genome-wide sampling of nuclear markers, indicate that the two mimetic (L. a. arizonensis and L. a. astyanax) subspecies in the L. arthemis complex are a monophyletic lineage, which is a sister group to the white-banded, non-mimetic subspecies (L. a. arthemis; figure 2). Thus, we find a singular origin of mimicry in the L. arthemis clade (L. a. arizonensis and L. a. astyanax) and retention of the ancestral white-banded phenotype (L. a. arthemis), where the model species (B. philenor) does not occur. However, consistent with previous attempts to reconstruct the phylogenetic relationships in this genus (Mullen 2006; Mullen et al. 2008; Prudic & Oliver 2008), we were unable to resolve the basal relationships among the Nearctic species (figure 2).

Although concerns about AFLPs have been raised, recent work indicates that homoplasy can be significantly reduced by implementing strict scoring parameters (Althoff et al. 2007; Holland et al. 2008). In the case of this study, the genome size of Limenitis is approximately 380 Mb (S. P. Mullen 2009, unpublished data), placing it well within the range where the average fragment homology is quite high across species (89%; Althoff et al. 2007). In addition, the advantage of employing a large number of genome-wide markers for phylogeny reconstruction is that it avoids many of the problems associated with single-locus inference among recently diverged taxa, as demonstrated by an increasing number of recent studies (Kardolus et al. 1998; Albertson et al. 1999; Baayen et al. 2000; Bakkeren et al. 2000; Ganter & Lopes 2000; Hodkinson et al. 2000; van Raamsdonk et al. 2000; Giannasi et al. 2001; Parsons & Shaw 2001; Buntjer et al. 2002; Allender et al. 2003; Després et al. 2003; Brouat et al. 2004; Sullivan et al. 2004; Bonin et al. 2005; Mendelson & Shaw 2005; Mendelson et al. 2005). In fact, given the expectation that hybridization among mimetic and non-mimetic populations should lead to substantial introgression of neutral markers, the nodal support for a single origin of mimicry is surprisingly high, a result that is inconsistent with the conclusion that mimicry was lost in white-banded populations. Therefore, it is likely that incomplete lineage sorting and/or introgression (Funk & Omland 2003; Anderson et al. 2009) have confounded previous efforts to recover the history of wing pattern evolution in this group.

Although examples of mitochondrial introgression are common (e.g. Krosby & Rohwer 2009; Shaw 2002), an alternative interpretation of the mitochondrial data presented in Mullen (2006), Mullen et al. (2008) and Prudic & Oliver (2008) is that southeastern populations of mimetic L. a. astyanax historically came into contact with a lineage of white-banded L. a. arthemis in sympathy with Battus, causing a selective sweep of alleles underlying the mimetic phenotype. Results from this study and previous work (Mullen 2006; Mullen et al. 2008; Prudic & Oliver 2008) support the following history for wing pattern evolution in the L. arthemis complex. First, the ancestral white-banded form gave rise to a mimetic lineage because of strong selection in the range of the codistributed model, B. philenor. Pleistocene vicariance of the mimetic lineage gave rise to allopatric populations of mimetic L. a. arizonensis in southwestern North America and mimetic L. a. astyanax in the southeast. Subsequently, hybridization between white-banded and mimetic populations would have resulted in sympatric mimetic forms fixed for different mtDNA haplotypes (e.g. IB and II) because of selection for the mimicry phenotype in the range of the model (figure 3). Non-mimetic populations of L. a. arthemis could have persisted in northwestern refugia, corresponding to mitochondrial haplotype IA, and expanded their range to give rise to the current zone of secondary hybridization between the two wing pattern phenotypes. Indeed, a history of episodic contact between mimetic and non-mimetic lineages is highly consistent with both the findings presented by Mullen et al. (2008) and Prudic & Oliver (2008), and also explains how mimetic phenotypes at the range limit of the model often possess non-mimetic mitochondrial haplotypes, as well as the lack of genealogical exclusivity for each phenotype.
and the highly asymmetrical gene flow between these two forms (*L. a. astyanax*--*L. a. arthemis*; Mullen et al. 2008). Ultimately, conclusive resolution of the history of wing pattern evolution in this group will almost certainly require identification of the specific locus/loci underlying the trait for mimicry.

**The causes and consequences of allopatric Batesian mimicry**

Given the dramatic potential consequences of frequency-dependent selection on Batesian mimicry phenotypes with respect to both individual fitness and the potential origin of reproductive barriers among populations, it is essential to understand the conditions under which mimicry is expected to evolve and be maintained. Here, we focused on elucidating the evolutionary origins and fate of a Batesian mimic phenotype among hybridizing populations of the *L. arthemis* butterfly species complex, and provided compelling evidence against an evolutionary loss of mimicry by reversion to the ancestral phenotype in this system. Batesian mimicry is a classic example of adaptive phenotypic evolution but, as yet, we have only an incomplete understanding of the conditions that maintain or break down the mimetic relationship between palatable Batesian mimics and unpalatable models. This is particularly true with respect to allopatric Batesian mimics that...
exist outside the geographical range of their respective models (Fisher 1930; Harper & Pfennig 2007, 2008; Pfennig et al. 2007). Specifically, it is unclear to what extent allopatric mimicry exists across diverse taxonomic groups, and how often allopatric mimics arise as a result of range expansion, range contraction of the model and/or through gene flow into non-mimetic populations.

The results of this study mean that we still have no unequivocal example of the loss of mimicry in natural populations (but see Harper & Pfennig 2008; reviewed in Joron 2008), and although it may well occur, results from a comprehensive analysis of mimicry evolution in Papilio butterflies strongly indicate otherwise (Kunte in review). Our results also caution against inferring character evolution using gene trees for recently diverged taxa, which are unlikely to resolve the true species phylogeny. Although recent advances in Bayesian and coalescent methodologies have improved our ability to recover phylogenetic relationships among closely related taxa with incongruent gene trees (Edwards & Beerli 2000; Maddison & Knowles 2006; Edwards et al. 2007; Liu & Pearl 2007), it is still true that gene trees and species trees are often not the same (Pamilo & Nei 1988; Maddison 1997; Edwards & Beerli 2000; Nichols 2001).

We are grateful to B. Evans, S.-P. Quek and M. Kronforst for their technical and laboratory assistance. D. W. Pfennig, M. Kronforst, M. Hassel and three anonymous reviewers offered detailed comments and suggestions that greatly improved the manuscript. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by Microsoft Corporation.

REFERENCES


Kunze, K. In review. The diversity and evolution of Batesian mimicry in Papilio swallowtail butterflies.


Proc. R. Soc. B (2009)


Veckemans, X. 2002 AFLP-SURV version 1.0. Distributed by the author, Laboratoire Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.


