Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly

ZACHARIAH GOMPERT,*CHRIS C. NICE,*JAMES A. FORDYCE,†MATTHEW L. FORISTER‡ and ARTHUR M. SHAPIRO§

*Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX 78666, USA, †Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA, ‡Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook, NY 11794, USA, §Section of Evolution and Ecology, University of California, Davis, CA 95616, USA

Abstract

The federally endangered North American Karner blue butterfly (Lycaeides melissa samuelis) and the closely related Melissa blue butterfly (L. m. melissa) can be distinguished based on life history and morphology. Western populations of L. m. samuelis share mitochondrial haplotypes with L. m. melissa populations, while eastern populations of L. m. samuelis have divergent haplotypes. Here we test two hypotheses concerning the presence of L. m. melissa mitochondrial haplotypes in western L. m. samuelis populations: (i) mitochondrial introgression has occurred from L. m. melissa populations into western L. m. samuelis populations, or (ii) western populations of the nominal L. m. samuelis are more closely related to L. m. melissa than to eastern L. m. samuelis populations, yet are phenotypically similar to the latter. A Bayesian algorithm was used to cluster 190 L. melissa individuals based on 143 informative amplified fragment length polymorphism (AFLP) loci. This method clearly differentiated L. m. samuelis and L. m. melissa. Thus, genomic divergence was greater between western L. m. samuelis populations and L. m. melissa populations than it was between western and eastern populations of L. m. samuelis. This supports the hypothesis that the presence of L. m. melissa mitochondrial haplotypes in western L. m. samuelis populations is the result of mitochondrial introgression. These data provide valuable information for conservation and management plans for the endangered L. m. samuelis, and illustrate the risks of using data from a single locus for diagnosing significant units of biodiversity for conservation.

Keywords: AFLP, conservation genetics, DNA barcoding, genomic divergence, Lycaeides melissa samuelis, mitochondrial introgression

Received 5 October 2005; revision accepted 14 December 2005

Introduction

The geographic distribution of genetic variation within and among taxa provides information on historical and contemporary demographic and evolutionary processes (Avise 1994, 2000; Knowles 2000, 2001). This information can also inform conservation efforts, both in terms of identifying units for conservation and designing management plans (Moritz 1994; Meffe & Carroll 1997; Primack 2004). The quest to identify appropriate biological units for conservation has a long history (Crandall et al. 2000). At present, consensus has not been reached on how to best delineate units for conservation (Crandall et al. 2000; Moritz 2002). Defining units for conservation based on any single character, whether mitochondrial sequence data (e.g. Hebert et al. 2003) or morphology, may be problematic. Multiple characters must be examined and the processes that influence those characters must be understood to accurately delineate units for conservation. Here we examine patterns of genetic variation based on both mitochondrial DNA (mtDNA) and amplified fragment length polymorphism (AFLP) (Vos et al. 1995) markers to test alternative hypotheses concerning the history and current status of the North
American endangered Karner blue butterfly (*Lycaeides melissa samuelis*).

The Karner blue butterfly (*L. m. samuelis*) and its close relative, the Melissa blue butterfly (*L. m. melissa*), are small butterflies belonging to the family Lycaenidae. *L. m. samuelis* has experienced a 99% range-wide decline in population size over the past century, most of which occurred in the last 25 years (US Fish and Wildlife Service 1992). This decline led to the listing of *L. m. samuelis* as an endangered species in the United States in 1992 (US Fish and Wildlife Service 1992, 2003). Remnant populations of *L. m. samuelis* are restricted to oak savannahs, pine prairies, and lakeshore sand dunes in Minnesota, Wisconsin, Indiana, Michigan, New York, and New Hampshire (Scott 1986; US Fish and Wildlife Service 1992, 2003). The closely related *L. m. melissa* is not considered endangered or threatened and is found in dry prairies and alfalfa fields over a large expanse of western North America, from Minnesota to California (Lane & Weller 1994; Brock & Kaufman 2003). Both *L. m. samuelis* and *L. m. melissa* use papilionaceous legumes (Fabaceae) as larval host plants (Scott 1986; Brock & Kaufman 2003). However, *L. m. samuelis* uses only wild lupine (*Lupinus perennis*), while *L. m. melissa* uses a number of legume genera including *Astragalus*, *Melicope*, *Glycyrrhiza*, and *Lupinus* — but not *Lupinus perennis* (Scott 1986; US Fish and Wildlife Service 1992; Lane & Weller 1994). *L. m. samuelis* populations are bivoltine, while *L. m. melissa* populations are variable but generally have more than two generations per year (US Fish and Wildlife Service 1992; Nice & Shapiro 1999). These two butterflies also differ morphologically, both in wing patterns (Nabakov 1949; Opler & Krizek 1984; Lane & Weller 1994) and in the size and shape of the male genitalia (Nabakov 1949; Lane & Weller 1994; C. C. Nice, unpublished).

Nice et al. (2005) examined the geographic distribution of genetic variation for the AT-rich region of the mitochondrial genome in North American *Lycaeides*. Western populations of *L. m. samuelis* (i.e. populations in the state of Wisconsin) shared haplotypes with *L. m. melissa* populations in western North America; in fact, there were no haplotypes in the Wisconsin *L. m. samuelis* populations that were not shared with *L. m. melissa* populations. In contrast, eastern *L. m. samuelis* populations (i.e. populations east of Lake Michigan) contained different haplotypes not found in any other *Lycaeides* populations (Nice et al. 2005). Thus there is discord between the traditional taxonomic boundary between *L. m. samuelis* and *L. m. melissa* based on morphological and ecological characteristics (Nabakov 1949; Lane & Weller 1994) and between the patterns of genetic variation observed for mtDNA. Packer et al. (1998) surveyed allozyme variation in one *L. m. melissa* population from Minnesota and two *L. m. samuelis* populations, one from Wisconsin and one from New York. They found low levels of genetic divergence and concluded that *L. m. samuelis* and *L. m. melissa* were not clearly differentiated (Packer et al. 1998).

Phylogeographic evidence suggests that *L. m. melissa* and *L. m. samuelis* populations were confined to different glacial refugia during the Pleistocene, and that they may have experienced secondary contact following post-Pleistocene range expansion (Nice et al. 2005). A similar phylogeographic boundary has been observed in other organisms and is attributed to Pleistocene refugia southeast and southwest of Lake Michigan (Austin et al. 2002). Secondary contact may have facilitated gene exchange between *L. m. samuelis* and *L. m. melissa* in which case Lake Michigan may have served as a geographic barrier against mitochondrial introgression into the eastern *L. m. samuelis* populations. Alternatively, populations in Wisconsin that are nominally *L. m. samuelis* may be more closely related to *L. m. melissa* populations than to *L. m. samuelis* populations east of Lake Michigan. This may be because *L. m. samuelis* is paraphyletic, or the ecological and morphological similarity of western *L. m. samuelis* populations to eastern *L. m. samuelis* populations may be the result of convergent evolution under similar selective pressures. Multiple studies have suggested that lineages of lycaenids diversify rapidly and respond to natural selection acting on ecological traits (Nice & Shapiro 1999; Nice et al. 2002; Fordyce & Nice 2003a).

Host-associated selection, in particular, could be expected to produce convergent patterns in populations that do not share an immediate common ancestor (Shapiro 1991; Nice & Shapiro 2001). For example, molecular data and ecological studies suggest that host plant use has driven convergent evolution of adult phenology and wing patterns in populations of the nominal butterfly species *Mitoura muiri* in the Coast Range and Sierra Nevada of California (Nice & Shapiro 2001; Forister 2004).

The two scenarios presented above have different implications for the management and conservation of *L. m. samuelis*. If Wisconsin *L. m. samuelis* populations possess *L. m. melissa* mitochondrial haplotypes as the result of mitochondrial introgression, then all *L. m. samuelis* populations can continue to be managed as a single unit. However, if *L. m. samuelis* populations on opposite sides of Lake Michigan are not each other's closest relatives, then it may be necessary to manage *L. m. samuelis* populations east and west of Lake Michigan as separate units. In particular, if the latter scenario is correct, it is important that translocations do not cross Lake Michigan. This concern is pertinent, as translocations have been proposed for reintroduction of *L. m. samuelis* individuals into areas where populations no longer exist and for supplementing current populations (US Fish and Wildlife Service 2003).

The two hypothesized scenarios of the biogeographic history of *L. m. melissa* in North America can be distinguished by examining the overall pattern of relatedness among *L. m. melissa* populations based on the nuclear genome. Two clear predictions can be made. If Wisconsin *L. m. samuelis* populations have *L. m. melissa* mitochondrial haplotypes as
a result of mitochondrial introgression, the nuclear genome of Wisconsin \( L. \text{m. samuelis} \) individuals should be more similar to the nuclear genome of \( L. \text{m. samuelis} \) individuals east of Lake Michigan than to the nuclear genome of \( L. \text{m. melissa} \) individuals (e.g. Funk & Omland 2003). Conversely, if the Wisconsin \( L. \text{m. samuelis} \) populations are more closely related to \( L. \text{m. melissa} \) populations than to \( L. \text{m. samuelis} \) populations east of Lake Michigan, the nuclear genome of Wisconsin \( L. \text{m. samuelis} \) individuals should be more similar to the nuclear genome of \( L. \text{m. melissa} \) individuals. In this case, patterns of variation observed in mtDNA and nuclear markers would both conflict with the current taxonomy.

In order to accurately assess overall genomic divergence, a large number of presumed neutral nuclear markers are needed, as individual gene genealogies are subject to stochastic events and take time to reflect the true population or species phylogeny (Funk & Omland 2003; Machado & Hey 2003). The AFLP technique (Vos et al. 1995) is an ideal choice for such an undertaking for a number of reasons. This technique can generate a large number of nuclear markers (> 100) in a short amount of time with only a modest start up cost (Bensch & Akesson 2005). This technique is especially useful in non-model organisms as no prior knowledge of the genome is required (Bensch & Akesson 2005). AFLP markers have been used successfully to detect genetic structure (e.g. Reineke et al. 1999; Wang et al. 2003; Mock et al. 2004; Irwin et al. 2005) and to identify cases of introgression (e.g. Sullivan et al. 2004) in wild populations.

Here we use data from mtDNA sequences and AFLP markers to test two alternative hypotheses regarding the biogeographic history of the endangered \( L. \text{m. samuelis} \): (i) mitochondrial introgression from \( L. \text{m. melissa} \) populations into Wisconsin \( L. \text{m. samuelis} \) populations has led to the presence of \( L. \text{m. melissa} \) mitochondrial haplotypes in the Wisconsin populations of \( L. \text{m. samuelis} \), or (ii) Wisconsin \( L. \text{m. samuelis} \) populations are more closely related to \( L. \text{m. melissa} \) populations than to \( L. \text{m. samuelis} \) populations east of Lake Michigan.

**Methods**

**Sample collection**

Butterflies were collected from five \( Lycaeides \text{melissa samuelis} \) populations and three \( Lycaeides \text{melissa melissa} \) populations (Fig. 1B, Table 1). Both males and females were collected

![Fig. 1](image)

**Fig. 1** Mitochondrial DNA haplotype network and population map. (A) Mitochondrial DNA haplotype network showing the single most parsimonious haplotype network for the three haplotypes identified. Each circle represents a haplotype. Black squares represent missing haplotypes. Haplotype C is separated from haplotypes A and B by six and seven mutations, respectively. (B) Population map. Dark grey shading marks approximate range of \( Lycaeides \text{melissa melissa} \) and light grey shading marks the approximate range of \( Lycaeides \text{melissa samuelis} \). Population abbreviations are given in Table 1. Diamonds represent populations either fixed for mtDNA haplotype A or with both haplotypes A and B (which differ by a single base pair), circles represent populations fixed for haplotype C. Empty shapes represent populations with a high probability of assignment to cluster 1 based on AFLP data, filled shapes represent populations with a high probability of assignment to cluster 2 (Fig. 3). The pattern of molecular variation is discordant between mtDNA data and AFLP markers.
from *L. m. melissa* populations, while only males were collected from *L. m. samuelis* populations (with the exception of two females collected at Saratoga Springs, NY) in accordance with USFWS permit PRT842392. DNA was isolated from about 0.5 g of thoracic tissue following the methods of Hillis et al. (1996) and Brookes et al. (1997).

**Mitochondrial DNA**

We sequenced portions of the mitochondrial gene cytochrome oxidase c subunit 1 (COI) and cytochrome oxidase subunit II (COII) for five individuals from each of the eight populations. Polymerase chain reaction (PCR) was performed using the primer pair C1-J-1751/C1-N-2191 for COI (Simon et al. 1994) and Pierre/Eva for COII (Caterino & Sperling 1999). This yielded fragments of approximately 450 and 550 base pairs (bp) for COI and COII, respectively. Fluorescently labelled dideoxy terminators were used for single-stranded sequencing reactions for both COI and COII according to Applied Biosystems specifications. Labelled amplicons were separated and visualized on 6% denaturing polyacrylamide gels, using an ABI PRISM 377 DNA sequencer (Applied Biosystems). GeneScan was used to visualize AFLP bands, which were sized by comparison to a size standard ladder (ROX standard, Applied Biosystems) added to each lane. Bands with low peak heights (less than 150 relative fluorescent units) were not scored. Bands that were present in less than 5% of the individuals surveyed were not included for subsequent analysis. Because almost all *L. m. samuelis* individuals collected were male, a single AFLP marker that appeared to be sex-linked was also excluded from all further analyses. AFLP banding patterns were highly reproducible. Twenty arbitrarily chosen individuals underwent a second amplification. For the twenty

<table>
<thead>
<tr>
<th>ID</th>
<th>Population</th>
<th>Taxon</th>
<th>mtDNA Haplotypes (no. of individuals)</th>
<th>P (cluster 1)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>Sierraville, CA</td>
<td><em>L. m. melissa</em></td>
<td>A(5)</td>
<td>0.992</td>
<td>27</td>
</tr>
<tr>
<td>SC</td>
<td>Spring Creek, SD</td>
<td><em>L. m. melissa</em></td>
<td>A(5)</td>
<td>0.968</td>
<td>28</td>
</tr>
<tr>
<td>BS</td>
<td>Brandon, SD</td>
<td><em>L. m. melissa</em></td>
<td>A(4), B(1)</td>
<td>0.935</td>
<td>24</td>
</tr>
<tr>
<td>FL</td>
<td>Fish Lake, WI</td>
<td><em>L. m. samuelis</em></td>
<td>A(5)</td>
<td>0.121</td>
<td>20</td>
</tr>
<tr>
<td>FMC</td>
<td>Fort McCoy, WI</td>
<td><em>L. m. samuelis</em></td>
<td>A(5)</td>
<td>0.040</td>
<td>19</td>
</tr>
<tr>
<td>NEC</td>
<td>Necedah, WI</td>
<td><em>L. m. samuelis</em></td>
<td>C(5)</td>
<td>0.032</td>
<td>23</td>
</tr>
<tr>
<td>IN</td>
<td>Indiana Dunes, IN</td>
<td><em>L. m. samuelis</em></td>
<td>C(5)</td>
<td>0.005</td>
<td>22</td>
</tr>
<tr>
<td>SS</td>
<td>Saratoga Springs, NY</td>
<td><em>L. m. samuelis</em></td>
<td>C(5)</td>
<td>0.044</td>
<td>27</td>
</tr>
</tbody>
</table>

**Amplified fragment length polymorphism markers**

In order to estimate overall genomic divergence and diversity within and between *L. m. samuelis* and *L. m. melissa*, AFLP marker profiles were produced for 19–28 individuals from each of the eight populations (190 individuals in total), following a modified version of the procedures introduced in Vos et al. (1995). AFLP markers were generated using three selective primer pairs: EcoRI-ACA and MseI-CTTG, EcoRI-ACA and MseI-CTTA, EcoRI-AGT and MseI-CTTA. Amplicons were separated and visualized on 6% denaturing polyacrylamide gels, using an ABI PRISM 377 DNA sequencer (Applied Biosystems). GeneScan was used to visualize AFLP bands, which were sized by comparison to a size standard ladder (ROX standard, Applied Biosystems) added to each lane. Bands with low peak heights (less than 150 relative fluorescent units) were not scored. Bands that were present in less than 5% of the individuals surveyed were not included for subsequent analysis. Because almost all *L. m. samuelis* individuals collected were male, a single AFLP marker that appeared to be sex-linked was also excluded from all further analyses. AFLP banding patterns were highly reproducible. Twenty arbitrarily chosen individuals underwent a second amplification. For the twenty
individuals, 95.5% of scored bands were detected in both amplifications.

The program structure (Pritchard et al. 2000) was used to cluster individuals based on their AFLP banding profiles. structure employs a model-based Bayesian clustering algorithm to assign individuals probabilistically to clusters to minimize deviations from linkage equilibrium. The admixture model was run for 500 000 generations with an initial burn-in of 50 000 generations. Prior information regarding the population or taxon from which an individual was sampled was ignored. structure was also used to estimate the number of clusters (k) that best explained the data. The method of Evanno et al. (2005) was used to infer k. This procedure identifies the appropriate number of clusters using the ad hoc statistic \( \Delta k \), which is based on the second order rate of change in the log probability of the data between successive values of k. Evanno et al. (2005) demonstrated that this method is able to detect the appropriate number of clusters for simulated data sets under a number of gene exchange models. It is not possible to evaluate \( \Delta k \) for \( k = 1 \) (Evanno et al. 2005). We explored the probability of the data for 2–9 clusters. Ten simulations were run for each k, multiple runs of the same k produced highly consistent individual assignment probabilities. Multiple runs for each k were used to compute the variance in structure estimates of the log probability of data for each k. These variance estimates were used in the calculation of \( \Delta k \) as described by Evanno et al. (2005).

**Results**

**Mitochondrial DNA**

Sequences were obtained for 410 bp of COI and 510 bp of COII for all 40 individuals examined (GenBank Accessions DQ234691–DQ234697). A conflicting phylogenetic signal between these gene regions was not detected using a partition homogeneity test (\( P = 1.000 \)), thus COI and COII sequences were combined for all analyses. Three unique haplotypes were detected for the combined sequence data (Table 1). A single most parsimonious haplotype network was produced (Fig. 1A). Haplotypes A and B differed by a single base, while these haplotypes differed from haplotype C by six or seven bases, respectively. The Sierraville, CA and Spring Creek, SD Lycaeides melissa melissa populations were fixed for haplotype A (Fig. 1B). A single individual from the Brandon, SD L. m. melissa population had haplotype B, while the other four individuals had haplotype A (Fig. 1B). All three Wisconsin L. m. samuelis populations (Fish Lake, Fort McCoy, and Necedah) were also fixed for haplotype A; however, the Indiana Dunes, IN and Saratoga Springs, NY L. m. samuelis populations were fixed for haplotype C (Fig. 1B). Sequence divergence between L. melissa and Wisconsin L. m. samuelis populations (haplotypes A and B) and L. m. samuelis populations east of Lake Michigan (haplotype C) was 0.65–0.76%. Based on data from COI and COII, the Wisconsin L. m. samuelis populations are indistinguishable from the L. m. melissa populations.

AMOVA partitioned approximately 12% of the total genetic variation for COI and COII between subspecies (\( \Phi_{CT} = 11.64, P < 0.001 \), Table 2A). SAMOVA was able to partition approximately 99% of the total genetic variation for COI and COII between the following two regional groups: (i) all three L. m. melissa populations and the Wisconsin L. m. samuelis populations and (ii) the L. m. samuelis east of Lake Michigan (\( \Phi_{CT} = 99.16, P < 0.001 \), Table 2B). The groups identified by SAMOVA explained an additional 87% of the total genetic variation beyond that explained by groups based on subspecies identification.

<table>
<thead>
<tr>
<th>A. Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance component</th>
<th>% of total</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among subspecies</td>
<td>1</td>
<td>9.111</td>
<td>0.162</td>
<td>11.64</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Among populations/within subspecies</td>
<td>6</td>
<td>36.391</td>
<td>1.208</td>
<td>86.56</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Within populations</td>
<td>32</td>
<td>0.802</td>
<td>0.025</td>
<td>1.80</td>
<td>0.096</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B. Source of variation</th>
<th>d.f.</th>
<th>Sum of squares</th>
<th>Variance component</th>
<th>% of total</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Among regional groups</td>
<td>1</td>
<td>45.335</td>
<td>3.020</td>
<td>99.16</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Among populations/within groups</td>
<td>6</td>
<td>0.167</td>
<td>0.001</td>
<td>0.02</td>
<td>1.000</td>
</tr>
<tr>
<td>Within populations</td>
<td>32</td>
<td>0.802</td>
<td>0.025</td>
<td>0.82</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Table 2 AMOVA for mitochondrial gene regions in COI and COII. (A) AMOVA for populations grouped according to subspecific identifications based on morphological and ecological differences. (B) AMOVA for populations grouped according to regions identified by SAMOVA to maximize \( \Phi_{CT} \).
Amplified fragment length polymorphism markers

The three primer pairs generated a total of 143 AFLP bands ranging in size from 71 to 481 bp, all of which were polymorphic among all individuals. All three primer pairs produced similar numbers of AFLP bands. A total of 130 (90.91%) bands were polymorphic within *L. m. melissa* and a total of 124 (86.71%) bands were polymorphic within *L. m. samuelis*. Twenty AFLP bands were present exclusively in *L. m. melissa* populations and 11 AFLP bands were found only in *L. m. samuelis* populations.

Two clusters best explained the AFLP data (Fig. 2). Under the admixture model an individual’s assignment probability to each cluster can be interpreted as the proportion of that individual’s genome that originated in each cluster. *L. m. melissa* individuals were assigned with high probability to one cluster (cluster 1), and no *L. m. melissa* individuals had an assignment probability to cluster 1 less than 0.645 (Figs 1B and 3). Nearly all *L. m. samuelis* individuals were assigned with high probability to another cluster (cluster 2), and no *L. m. samuelis* individuals had an assignment probability to cluster 2 less than 0.455 (Figs 1B and 3). This includes the Wisconsin *L. m. samuelis* populations that were grouped with *L. m. melissa* based on mtDNA. The mean assignment probability of *L. m. melissa* populations to cluster 1 ranged from 0.935 at Brandon, SD to 0.992 at Sierraville, CA (Table 1). The mean assignment probability of *L. m. samuelis* populations to cluster 1, which equals one minus their mean assignment probability to cluster 2, ranged from 0.005 at Indiana Dunes, IN to 0.121 at Fish Lake, WI (Table 1). The lowest assignment probability to cluster 1 for a *L. m. melissa* population (Brandon, SD) and the highest assignment probability to cluster 1 for a *L. m. samuelis* population (Fish Lake, WI) occurred nearest the boundary between these taxa. However, even at these locations AFLP markers clearly distinguish between *L. m. melissa* and *L. m. samuelis* individuals (Fig. 3). AFLP data, which provides a metric of genomic divergence, support the nominal taxonomic boundary between these taxa, which was based on ecological and morphological data (Nabakov 1949; Lane & Weller 1994).

**Discussion**

**Phylogeographic history of Lycaeides melissa samuelis**

Mitochondrial DNA (COI and COII) and AFLP markers identified different boundaries between *Lycaeides melissa samuelis* and *Lycaeides melissa melissa*. All three *L. m. melissa* populations and the Wisconsin *L. m. samuelis* populations were fixed, or nearly fixed (as in Brandon, SD), for the same mitochondrial haplotype (haplotype A), while *L. m. samuelis*...
populations east of Lake Michigan were fixed for a different divergent haplotype (haplotype C). Thus, COI and COII mitochondrial DNA data partitions these populations into two groups: (i) L. m. melissa and Wisconsin L. m. samuelis and (ii) L. m. samuelis east of Lake Michigan, which are separated by 0.65–0.76% sequence divergence. This degree of sequence divergence is typical of other butterfly subspecies (e.g., Nice & Shapiro 2001; Fordyce & Nice 2003b). The geographic pattern of genetic variation for COI and COII is very similar to the pattern identified by Nice et al. (2005) based on the AT-rich region of the mitochondrial genome. There is an apparent phylogeographic boundary between mitochondrial clades at or near Lake Michigan.

Unlike the mitochondrial data, AFLP data provided no evidence for a genetic boundary near Lake Michigan. Bayesian clustering of individuals based on AFLP marker profiles partitioned individuals into two clusters: one that consisted of L. m. samuelis individuals and one that consisted of L. m. melissa individuals. This pattern is in accord with patterns of variation in habitat and host-plant use (Scott 1986; US Fish and Wildlife Service 1992; Lane & Weller 1994; Brock & Kaufman 2003), phenology (US Fish and Wildlife Service 1992; Nice & Shapiro 1999), wing morphology (Nabakov 1949; Lane & Weller 1994), male genitalic morphology (Nabakov 1949; Lane & Weller 1994), and allozyme data (Packer et al. 1998) and thus corresponds to the pattern expected based on taxonomic designations.

The incongruent patterns of genetic variation observed in mtDNA and nuclear AFLP markers support the hypothesis that the presence of mitochondrial haplotypes in the Wisconsin L. m. samuelis populations that are identical to haplotypes found in L. m. melissa populations is the result of mitochondrial introgression from L. m. melissa populations into the Wisconsin L. m. samuelis populations (Fig. 1B). This mitochondrial introgression appears to have progressed as far as Lake Michigan. However, we cannot rule out the possibility of ancestral polymorphism. For example, the L. m. melissa lineage may have become fixed for one mitochondrial variant while L. m. samuelis continued to be polymorphic, until selective sweeps or genetic drift fixed different mitochondrial haplotypes in the eastern and western portions of their range. This scenario implies that there has been insufficient time for significant sequence divergence to accumulate between L. m. melissa and western L. m. samuelis. These possibilities seem unlikely given homogeneity in terms of habitat, host plant use, morphology, and the AFLP data presented here, over the entire range of L. m. samuelis.

Despite extensive mitochondrial introgression from L. m. melissa into the Wisconsin L. m. samuelis populations, there has been little nuclear introgression. This lack of nuclear introgression is evidenced by the fact that there are only six individuals with moderate assignment probabilities to both cluster 1 and cluster 2, most of which are from Fish Lake, WI (Fig. 3). Many more individuals would be expected to have moderate assignment probabilities to both clusters if nuclear introgression were prevalent. There are two likely explanations for the lack of nuclear introgression in combination with widespread mitochondrial introgression. First, natural selection against L. m. melissa × L. m. samuelis hybrids and backcrosses may be sufficiently strong to limit nuclear introgression, while still allowing for neutral mitochondrial alleles to pass almost freely from L. m. melissa populations to the Wisconsin L. m. samuelis populations. It is not uncommon to see unidirectional introgression in such cases (Chan & Levin 2005). This would provide evidence that at least some of the morphological and/or ecological differences between L. m. melissa and L. m. samuelis are important reproductive isolating barriers involved in maintaining the boundary between these taxa. Dissimilarity in wing pattern and/or male genitalia structure between L. m. melissa and L. m. samuelis may preclude hybrid and backcross individuals from mating. There is evidence that wing pattern is important for mate recognition and preference for other Lycaeides populations (Fordyce et al. 2002). Such prezygotic barriers are especially permeable to introgression of maternally inherited genes (Chan & Levin 2005). Additionally, differences in habitat and host-plant use between L. m. melissa and L. m. samuelis may reduce the fitness of individuals of mixed descent in either of the parental habitats. A second explanation for the lack of nuclear introgression between L. m. samuelis and L. m. melissa populations despite substantial mitochondrial introgression is a mitochondrial selective sweep. Because animal mitochondrial genomes usually do not undergo recombination (but see Eyre-Walker et al. 1999), a selective advantage for the L. m. melissa mitochondrial genome at a single locus may have been sufficient to drive a selective sweep of the entire mitochondrial genome. Such non-neutral variation in mitochondrial alleles has been postulated to explain other phylogeographic patterns (Levin 2000; Brumfield et al. 2001). At present we are unable to discriminate between these two possibilities. It would be possible to detect a mitochondrial selective sweep by comparing effective population size estimated from a number of nuclear gene sequences to an estimate based on mtDNA for the Wisconsin L. m. samuelis populations. A significantly lower effective population size estimate for mtDNA than for nuclear DNA would be indicative of a selective sweep (Galtier et al. 2000). However, at present nuclear sequence data from several loci is not available for L. m. samuelis.

Conservation implications

We conclude, based on our data and the available morphological and ecological data, that L. m. samuelis is a unique
entity, distinct from L. m. melissa. This study finds little evidence for separate origins of the L. m. samuelis populations on different sides of Lake Michigan. As a result, our data do not suggest the need to treat populations east and west of Lake Michigan as separate units for conservation and management purposes. This does not mean that we can say for certain that translocations between different L. m. samuelis populations could take place without negative consequences, as population level local adaptation may still be present within L. m. samuelis, which could lead to reduced fitness of interpopulation hybrids and potentially lower the mean fitness of the recipient population of the translocation (i.e. outbreeding depression). Further investigation is needed prior to undertaking interpopulation translocations. However, it is clear in this case that the evolutionary history of the mitochondrial genome is not indicative of the history of the nuclear genome, which means that the mtDNA data do not accurately reflect the evolutionary relationships of this group.

The findings of this study highlight a potential problem regarding the recent trend to rely primarily on DNA sequence data, especially from the mitochondrial genome, to identify units of biodiversity (e.g. Moritz 1994; Holland & Hadfield 2002; Hebert et al. 2003, 2004; Tautz et al. 2003). This trend has met with a number of criticisms (e.g. Will & Rubinoff 2004; Prendini 2005; Wheeler 2005; Will et al. 2005). As stated by some of these critics, data from a single locus such as mtDNA should be used with caution. In this case, mtDNA incorrectly identifies the Wisconsin populations of the endangered species L. m. samuelis as populations of the widespread L. m. melissa. This is a case in which DNA systematics would fail to identify the appropriate units of biodiversity for conservation purposes. Such techniques would not support the conservation status of the Wisconsin L. m. samuelis populations, which is clearly warranted based on the strong correlations between patterns of genomic divergence, morphological characters and ecological data. This does not mean that mtDNA data should be ignored in general, as mtDNA has been used effectively to identify units for conservation (e.g. Holland & Hadfield 2002); however, we recommend obtaining corroborating evidence from nuclear markers to support conclusions from mtDNA.

Acknowledgements

We thank N. Anthony, G. Gelembiuk, R. ffrench-Constant, S. Fuller and M. Amaral for assistance in the field. J. Ott, B. Fitzpatrick, M. Forstner, B. Holland, an anonymous reviewer and the Texas State Evolution, Ecology and Behaviour discussion group provided helpful comments and discussion. C. Carnes and the USFWs assisted with logistical support and permitting. This work was funded by a Research Enhancement Grant from Texas State University to C. Nice.

References


© 2006 Blackwell Publishing Ltd, Molecular Ecology, 15, 1759–1768


Zachariah Gompert is a Masters student at Texas State University with research interests in evolutionary biology, speciation, and genetics. Chris Nice is Assistant Professor at Texas State University with research interests in evolutionary and ecological genetics. James Fordyce is Assistant Professor at University of Tennessee with research interests in factors that promote population differentiation and plant insect interactions. Matt Forister, a postdoctoral research associate at the State University of New York, Stony Brook, is interested in speciation, specialization, and the genetics of host use in herbivorous insects. Arthur M. Shapiro is a professor at the University of California, Davis. He studies the biogeography, evolution, and phenology of butterflies in western North America and the Andean-Patagonian region.