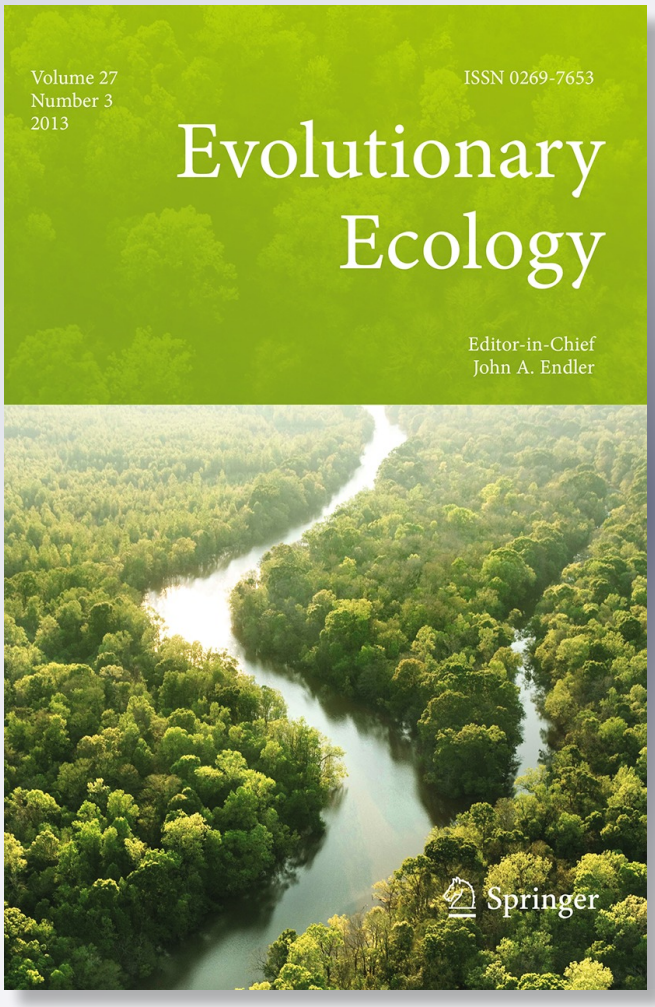
A role for both ecology and geography as mechanisms of genetic differentiation in specialized butterflies

Michelle H. Downey & Chris C. Nice

Evolutionary Ecology

ISSN 0269-7653
Volume 27
Number 3

Evol Ecol (2013) 27:565-578
DOI 10.1007/s10682-012-9626-7



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Author's personal copy

Evol Ecol (2013) 27:565–578
DOI 10.1007/s10682-012-9626-7

ORIGINAL PAPER

A role for both ecology and geography as mechanisms of genetic differentiation in specialized butterflies

Michelle H. Downey · Chris C. Nice

Received: 28 June 2012 / Accepted: 21 December 2012 / Published online: 4 January 2013
© Springer Science+Business Media Dordrecht 2013

Abstract An important mechanism of speciation for specialized phytophagous insects is host-associated differentiation, in which natural selection drives the evolution of reproductive isolation. Here we use molecular population genetics data to test the hypothesis that regional and local specialization on three alternate hosts restricts gene flow in the *Mitoura gryneus* species complex of butterflies. Over half of the variation in mitochondrial DNA sequences is explained by host plant association ($\phi_{CT} = 0.57$, $P = 0.002$) consistent with the hypothesis of host-associated divergence on the three hosts. AFLP analyses revealed the number of clusters of individuals was $K = 2$, with all individuals associated with one host grouping separately from all other host-associated individuals. Combined with previous experimental results, these findings present evidence of varying levels of differentiation among host associations and identify a role for both isolation in allopatry as well as ecological factors in limiting gene exchange. The *Mitoura* species complex includes multiple, differentiated lineages at varying stages of divergence, providing an opportunity to examine the multifarious mechanisms that generate biodiversity in phytophagous insects.

Keywords Ecological speciation · Host races · Intraspecific differentiation · *Mitoura* · Specialization

Electronic supplementary material The online version of this article (doi:10.1007/s10682-012-9626-7) contains supplementary material, which is available to authorized users.

M. H. Downey (✉)
Department of Ecology and Evolutionary Biology, Rice University, Houston, TX 77005, USA
e-mail: mhdowney@rice.edu

C. C. Nice
Department of Biology, Population and Conservation Biology Program, Texas State University-San Marcos, San Marcos, TX 78666, USA

Author's personal copy

Introduction

Recent studies of speciation have examined closely the influence of ecological factors on the process of genetic divergence among populations under varying degrees of spatial isolation and gene flow (Nosil et al. 2006; Thorpe et al. 2008; Berner et al. 2009). Ecological speciation theory proposes divergent natural selection as the mechanism that initiates and drives reproductive isolation between groups, either through acting directly on traits associated with reproduction or indirectly on genetically correlated traits (Schluter 2001; Rundle and Nosil 2005; Funk and Nosil 2008). These empirical studies and theoretical advances have focused attention on the relative roles of natural selection and genetic drift in the process of divergence among populations (Felsenstein 1980; Slatkin 1987; Funk 1998; Gavrilets et al. 1998; Lenormand 2002; Coyne and Orr 2004). The challenge of speciation studies is to characterize the balance between local adaptation and gene flow, and link proximate ecological processes to the larger phylogeographic history of the natural system of interest (Feder et al. 2005; Butlin et al. 2008; Raesaenen and Hendry 2008).

In order to untangle the complexity of speciation, it is helpful to examine a system of closely related taxa in various stages of speciation and for which different mechanisms are responsible for divergence. This facilitates study of the relative contributions of different reproductive isolating barriers under different spatial and ecological conditions (Tregenza 2002, Nosil et al. 2009; Via 2009). Plant–insect systems have frequently been examined to address speciation questions because phytophagous insects are often closely associated with their host plants, and host-associated divergence plays a central role in sympatric or ecologically-based speciation models (Drès and Mallet 2002; Funk et al. 2002).

The *Mitoura gryneus* species complex is well-suited for studying mechanisms of divergence because these butterflies are closely associated with their respective hosts (trees in the family Cupressaceae) and exhibit behaviors such as female oviposition preference and male lekking, which might influence assortative mating based on host plant use (Scott 1992; Forister 2004). The geographic distributions of *Mitoura* hosts occur, in varying degrees, from complete allopatry, to parapatry and sympatry; therefore, these butterflies provide the opportunity to study the potential interaction between geographic and ecological reproductive isolating barriers.

Population genetic (Nice and Shapiro 2001) and experimental (Forister 2004, 2005) work with *Mitoura* in northern California examined three nominal species (*Mitoura muiri*, *M. nelsoni*, and *M. siva*) that differ in host plant association, morphology, and phenology of flight time. Nice and Shapiro (2001) found genetic structure associated with some of these ecological factors, while taxonomic designations were not indicative of levels of population genetic divergence. Forister (2004, 2005) tested the hypothesis of host-associated differentiation and specialization among these same populations by measuring both female preference and larval performance. Preference and performance varied, with one host-associated population exhibiting a concordance of preference and performance for the natal host. For the other butterfly-host plant associations the relationship was more complex, with larval performance not always highest on the plant preferred by females. Despite differences in morphology (taxonomic designations are based primarily on morphological variation), the low levels of genetic differentiation and variation in adaptation to natal hosts suggest that evolutionary divergence in this group is incomplete and ongoing.

In contrast to the butterflies in the northwest, *Mitoura* in the south-central portion of North America are morphologically homogeneous and are considered one nominal taxon (*M. gryneus*). In addition, *Mitoura* within this region are associated with different species of trees within a single genus, *Juniperus*, that occur both allopatrically and sympatrically.

Author's personal copy

In order to better isolate the role of ecological specialization in driving divergence, these butterflies that do not differ in morphology or phenology, but do differ in host plant use, were examined. Experimental work with *M. gryneus* (Downey and Nice 2011) found variation in female oviposition preference and larval performance among different host-associated populations, indicating host race formation is occurring but is incomplete in some cases. The current study complements this previous work on *Mitoura* and tests the hypothesis that populations of *M. gryneus* associated with different hosts experience restricted gene flow, and exhibit significant population genetic divergence. Furthermore we ask, are the patterns of genetic differentiation coincident with ecological divergence (i.e., patterns of specialization in terms of preference and performance; Downey and Nice 2011) as would be predicted under the hypothesis of host-associated differentiation?

Materials and methods

Biology of *Mitoura gryneus* and host plants

The juniper hairstreak, *M. gryneus* (Family Lycaenidae), is part of a species complex that includes *M. muiri*, *M. nelsoni*, *M. siva*, *M. sweadneri*, *M. thornei* and others, all of which are considered by some taxonomists to be either separate species or subspecies of *M. gryneus* (Johnson 1981; Scott 1992). Host plant association is important in many of these taxonomic designations and all of the taxa in this complex use members of the Cupressaceae as hosts. Throughout much of the eastern United States, *M. gryneus* uses *Juniperus virginiana* as the sole host plant (with exceptions such as *M. sweadneri* associated with *J. silicicola* in Florida). In the western and northwestern regions of North America, *Mitoura* butterflies are more taxonomically diverse and are associated with a greater number of cupressaceous hosts in several genera including *Juniperus*, *Cupressus*, and *Calocedrus* (Johnson 1981; Nice and Shapiro 2001; Forister 2004, 2005).

Three host associations of *M. gryneus* on *Juniperus* species in Texas were examined in this study. Two host species are nearly entirely confined to Texas. Ashe juniper (*J. ashei*) occurs primarily in central Texas, and red-berried juniper (*J. pinchotii*) occurs primarily in western-northwestern Texas. Eastern red cedar (*J. virginiana*) occurs throughout eastern-northeastern regions of North America and into the eastern half of Texas (Adams 2008). *Mitoura* were sampled from areas where the hosts *J. ashei*, *J. pinchotii* and *J. virginiana* are allopatric, as well as where *J. pinchotii* and *J. ashei* are sympatric at various points across their range (Downey and Nice 2011, Fig. 1; see also Supporting information, Fig. S1). Hereafter, we will refer to these sampling populations of *M. gryneus* by host plant association (*J. ashei*, *J. pinchotii*, sympatric *J. ashei–J. pinchotii*, or *J. virginiana*).

Previous studies (Downey and Nice 2011) with butterflies from these populations examined female oviposition preference and larval performance to test the hypothesis that specialization on alternate hosts has led to host race formation and facilitated evolutionary divergence. Different host-associated populations exhibited different degrees of specialization. Females of *J. virginiana*-associated populations exhibited no preference for, but larvae were observed to have increased developmental efficiency on, their natal host. *J. ashei*-associated populations exhibited a concordance between preference and performance for their natal host, while *J. pinchotii*-associated populations preferred and exhibited equally high performance on *J. ashei* and *J. pinchotii*. Butterflies from localities where the hosts *J. ashei* and *J. pinchotii* were sympatric varied in the patterns of preference and performance.

Author's personal copy

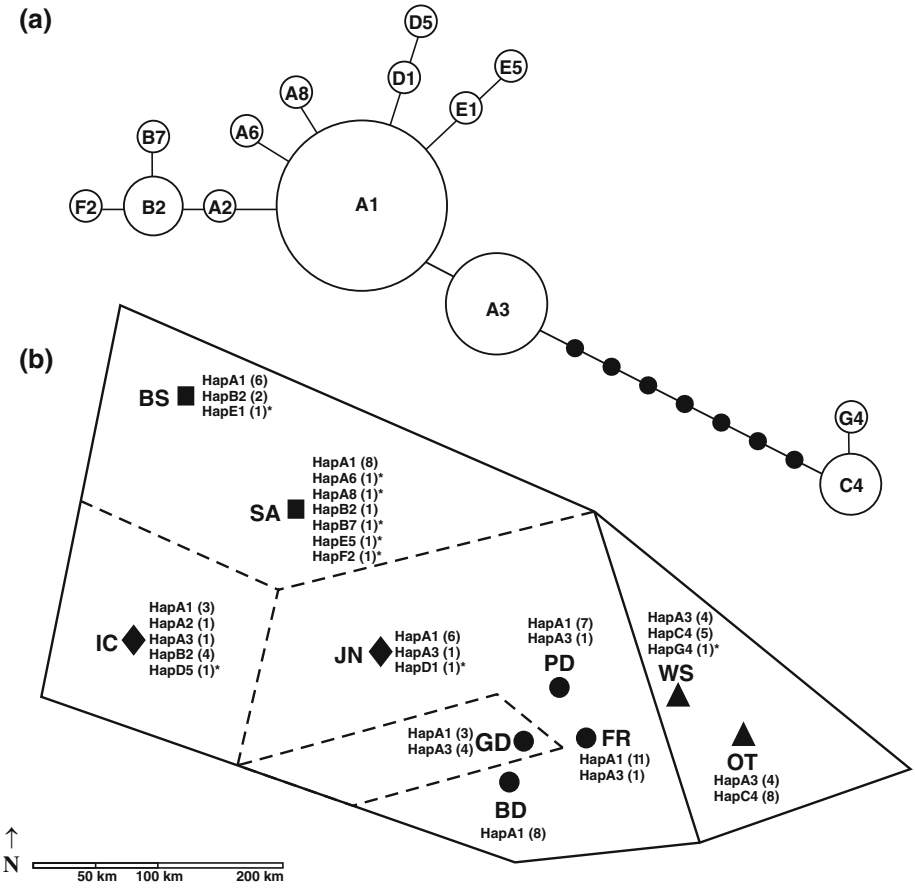


Fig. 1 Results of mtDNA analyses. **a** Haplotype network of combined COI-COII sequences. Connections are 1 nucleotide substitution difference. *Filled circles* are unsampled intermediate haplotypes. **b** Map of sampling locations indicating relative position of populations. Number of individuals with each haplotype in *parentheses*. Private alleles are indicated by *asterisks*. *Square symbols J. pinchotii* host association, *diamonds J. pinchotii* and *J. ashei*, *circles J. ashei*, and *triangles J. virginiana*. Population labels are the same as in Table 1. *Solid-line polygons* around populations are SAMOVA results when $K = 2$; *dashed line* indicates SAMOVA results (excluding *J. virginiana* populations) when $K = 4$

Molecular protocol for data collection

Adult specimens of *M. gryneus* were collected from ten sampling locations (hereafter "populations") during 2008–2009. Populations were the same as those used in earlier experimental work by Downey and Nice (2011), with the addition of one *J. ashei*-associated population (Bandera; Fig. 1; Table 1). Thoracic tissue from adult butterflies was used for genomic DNA extraction (head, abdomen, and wings stored as voucher specimens at −80 °C). Regions of mitochondrial DNA (cytochrome oxidase subunits I and II) were amplified and sequenced for a subset of 8–14 individuals per sampling population using the COI primer pairs RON (C1-J-1751) and NANCY (C1-N-2191) and the COII primer pairs PATRICK (C2-J-3038) and EVA (C2-N-3782) (Simon et al. 1994; Caterino and Sperling 1999). PCR products (see Supporting Information for protocol details) were purified using

Author's personal copy

Table 1 Collection information for *M. gryneus* specimens with mtDNA haplotype information, and per-population haplotype (gene) diversity (h) based on haplotype number and frequency

Population locality	Coordinates	Host plant	Haplotype (no. of individuals)	h
Big Spring (BS)	32° 14′59.74″ N 101° 29′02.75″ W	*J. pinchotii*	A1 (6), B2 (2), E1 (1)	0.56 ± 0.17
San Angelo (SA)	31° 29′02.75″ N 100° 29′48.92″ W	*J. pinchotii*	A1 (8), A6 (1), A8 (1), B2 (1), B7 (1), E5 (1), F2 (1)	0.69 ± 0.74
Independence Creek (IC)	30° 29′24.57″ N 101° 47′54.33″ W	*J. ashei, J . pinchotii*	A1 (3), A2 (1), A3 (1), B2 (4), D5 (1)	0.76 ± 0.11
Junction (JN)	30° 29′53.19″ N 99° 44′03.20″ W	*J. ashei, J. pinchotii*	A1 (6), A3 (1), D1 (1)	0.46 ± 0.20
Bandera Rd (BD)	29° 35′48.26″ N 98° 38′4.17″ W	*J. ashei*	A1 (8)	0.00 ± 0.00
Guadalupe (GD)	29° 53′9.87″ N 98° 32′4.49″ W	*J. ashei*	A1 (3), A3 (4)	0.57 ± 0.12
Pedernales (PD)	30° 16′39.18″ N 98° 15′23.73″ W	*J. ashei*	A1 (7), A3 (1)	0.25 ± 0.18
Freeman (FR)	29° 55′23.48″ N 98° 1′13.27″ W	*J. ashei*	A1 (11), A3 (1)	0.17 ± 0.13
Welsh (WS)	30° 13′58.19″ N 97° 15′41.53″ W	*J. virginiana*	A3 (4), C4 (5), G4 (1)	0.67 ± 0.10
Oak Thicket (OT)	29° 56′55.08″ N 96° 43′49.44″ W	*J. virginiana*	A3 (8), C4 (4)	0.48 ± 0.11

Promega Wizard SV Genomic DNA Purification System, followed by cycle sequencing using either a CEQ8800 Genetic Analysis System and protocols (Beckman Coulter) or the BigDye Terminator v3.1 Reaction Kit and ABI3730 DNA Analyzer (for the latter, sequencing was performed by the Nevada Genomics Center, www.ag.unr.edu/genomics). Sequence data was edited by eye and aligned using Geneious v5.0 (Drummond et al. 2010).

To obtain genome-wide estimates of differentiation from a large number of nuclear markers, AFLP data were also collected for 18–27 individuals per nine sampling populations (excluding the Bandera population; 184 individuals total) following a modification of the protocol described by Vos et al. (1995) (see also Gompert et al. 2006). Two selective primer pairs were used, a FAM-labeled EcoRI-ACA primer paired with MseI-CAGA and MseI-CAGC, respectively. Reaction products were mixed with a 500 MW LIZ size standard and analyzed on an ABI Prism 3730 DNA Analyzer (service provided by Nevada Genomics Center). Eight samples were run repeatedly on each 96-well plate to check that among-plate variability was minimized (≥95 % of detected markers were successfully scored in each independent amplification profile). Raw electropherograms were analyzed using ABI PeakScanner v.1.0 to detect presence or absence and size of bands for each individual (with "light peak smoothing"; all other settings were default). A binary matrix of presence (1) or absence (0) for AFLP bands was constructed using the automated scoring RawGeno package in R (R CRAN; Arrigo et al. 2009) with the following parameters: scoring range, 100–500 bp; maximum bin width, 2 bp; low intensity threshold, 50 rfu's; with 5 % low-frequency bins eliminated. Each selective primer pair data set was scored separately, and then concatenated for analyses.

Author's personal copy

mtDNA data analysis

Mitochondrial COI and COII sequence data were combined and analyzed as a single unit after a partition homogeneity test found no significant difference in phylogenetic signal between the two markers ($P = 0.23$) (PAUP v.4.0beta10, Swofford 2003). A statistical parsimony haplotype network was constructed using the program TCS v.1.21 (Clement et al. 2000) (Fig. 1a). Haplotype (gene) diversity, h (Nei 1987), was calculated for each population (Table 1). An analysis of molecular variance (AMOVA) was used to analyze population structure according to host plant association using uncorrected pairwise sequence differences (ARLEQUIN v.3.5; Excoffier et al. 2005). Three levels of genetic variance were examined: (1) within-population variance; (2) variance among populations, within groupings by host associations ("within-hosts"); and (3) variance among groupings of host associations ("among-hosts"). Significance of ϕ-statistics was assessed with 10,000 permutations of each level of the hierarchy. If "among-hosts" variation is significant (i.e., ϕ-statistic measuring differentiation among hosts is significantly greater than zero), populations using alternative hosts were considered to be differentiated, consistent with the hypothesis that host plant fidelity plays a role in maintaining reproductive isolation among *M. gryneus* using alternate hosts. Because the AMOVA examined single host-association groupings, the initial AMOVA excluded the Junction and Independence Creek populations, where *J. ashei* and *J. pinchotii* hosts are sympatric. The results from previous experimental work (Downey and Nice 2011) were used to determine how to include these populations in a subsequent AMOVA. Based on patterns of female preference and larval performance and the similarities with specific host associations, the Junction population was grouped with the *J. ashei*-associated populations, and the Independence Creek population with the *J. pinchotii*-associated populations. Due to a relatively greater degree of sequence divergence between *J. virginiana*-associated populations and all others (see "Results"; Fig. 1a), two additional AMOVAs were conducted. One AMOVA excluded the *J. virginiana*-associated populations to test for host-associated differentiation between the *J. ashei*- and *J. pinchotii*-associated populations. Another nested, hierarchical AMOVA was also conducted using conventional F-statistics calculated from haplotype frequencies across all populations and host associations, essentially discounting the relatively large sequence divergence between haplotypes found in *J. virginiana* populations and all others.

While the previous AMOVAs describe partitioning of genetic variation in terms of nested hierarchies, determined a priori, additional analyses were employed to test for other spatial relationships in the data. A spatial analysis of molecular variance (SAMOVA; Dupanloup et al. 2002) was used to assess the structure of populations in both spatial and genetic terms, grouping together geographically adjacent sampling populations that exhibit evidence of gene flow. The number of host population groups hypothesized ($K = 3$) was tested along with other possible numbers of groupings (from 1 through 9). SAMOVA assigns populations to groups using a simulated annealing procedure that detects areas of local maxima of genetic variance among or between populations; groups are geographically adjacent and genetically homogeneous (Dupanloup et al. 2002). If host-associated differentiation is occurring, and is the dominant phylogeographic pattern, $K = 3$ should show maximal among-group differentiation (i.e., the highest ϕ_{CT}) and correspond to the host-associated population groupings.

An alternative pattern of geographic genetic variation that may be detected is isolation by distance (IBD; Wright 1943), which is in contrast to the hypothesis of restricted gene flow between different host-associated populations. A Mantel test (Sokal and Rohlf 1995) was used to test for IBD by comparing a genetic distance matrix of pairwise ϕ_{ST}'s

Author's personal copy

(calculated between all pairs of populations from mtDNA data) to a geographic distance matrix. A strong signal of IBD indicates that other factors besides host plant association could be influencing the distribution of genetic variation. Lack of evidence of a pattern of IBD would be consistent with the hypothesis of host-associated differentiation.

AFLP data analysis

AFLP data were analyzed using STRUCTURE (Pritchard et al. 2000), a Bayesian clustering method that assigns individuals to groups (or clusters) that does not require a priori information. Ten replicates for each value of K number of clusters (from $K = 1$ through $K = 10$, the total number of sampling populations examined plus one) were conducted, with the mean ln L (log likelihood) and variance calculated. For each replicate, a burn-in of 50,000 with 500,000 Markov Chain Monte Carlo iterations was used. The model was run with admixture and correlated allele frequencies, allowing for recessive alleles (Falush et al. 2003) and alternate analyses both with and without population sampling location information was used (LOCPRIOR model; Hubisz et al. 2009). Although not explicitly spatial, the LOCPRIOR model in STRUCTURE optionally incorporates population location designation if there is a correlation between cluster assignment and population information. The value of K at which the mean ln L reached an asymptote was selected (Pritchard et al. 2000); this also corresponded with the value of K with the highest delta K statistic (Evanno et al. 2005).

Results

mtDNA haplotypes and analyses

A total of 828 bp of mtDNA sequence data (392 bp of COI and 436 bp of COII) were obtained, from which 14 unique haplotypes were identified. The AMOVA revealed that 56.6 % of the total genetic variation was explained when all populations were grouped according to host plant association ($df = 2$, $P = 0.002$; Table 2a). In addition, among-population variation within each host association was low and not statistically significant, indicating that populations within each host association were genetically similar (percent variation explained = 1.32 %, $df = 9$, $P = 0.43$; Table 2a). These proportions were relatively unchanged even when the host sympatric populations, Junction and Independence Creek, were excluded (with host-associated groupings explaining 62.4 % of the total variation, $df = 2$, $P = 0.005$; Table 2b). An examination of the haplotype network revealed substantial divergence (8 nucleotide substitutions) between haplotypes from *J. virginiana* populations and haplotypes from all other populations (Fig. 1a), raising the possibility that these AMOVA results were driven by this pattern. When considering only the *J. ashei* and *J. pinchotii*-associated populations in an AMOVA, 18.6 % of the variation was still explained by grouping populations by host association ($df = 1$, $P = 0.02$; Table 2c). An AMOVA conducted using conventional F-statistics revealed that 33.3 % of the genetic variation was due to grouping populations according to host plant association ($df = 2$, $P = 0.003$; Table 2d). Significant isolation by distance was not detected, regardless of whether all populations were included (grouping Junction with *J. ashei* and Independence Creek with *J. pinchotii* populations), Mantel test, $r = 0.06$, $P = 0.23$; or when *J. virginiana*-associated populations were excluded from the analysis, Mantel test,

Author's personal copy

Table 2 AMOVA results examining amount of variation explained when grouping sampling populations by host plant association

Source of variation	*df*	Sum of squares	Percentage of variation	*P* value
(a) All populations				
Among groups	2	81.54	**56.59**	**0.002**
Among populations within groups	9	7.79	1.32	0.43
Within populations	89	67.72	**42.09**	**<0.001**
(b) Excluding sympatric *J. ashei–J. pinchotii* populations				
Among groups	2	60.28	**60.02**	**0.038**
Among populations within groups	5	2.12	1.99	0.45
Within populations	71	57.79	**40.38**	**<0.001**
(c) Excluding *J. virginiana* populations				
Among groups	1	4.17	**18.63**	**0.018**
Among populations within groups	6	3.20	2.75	0.116
Within populations	69	27.72	**78.62**	**0.001**
(d) Using F-statistics				
Among groups	2	8.91	**33.30**	**0.003**
Among populations within groups	7	2.87	**4.68**	**0.039**
Within populations	88	21.01	**62.02**	**<0.001**

Significant results are in bold

$r = 0.15$, $P = 0.22$, again suggesting that host plant association influences population genetic structure.

When spatial information was included, the analysis using SAMOVA revealed that the number of groups (K) that explained the highest proportion of among-group variation was $K = 2$, in which *J. virginiana* populations grouped together and were separate from all other populations ($\phi_{CT} = 0.699$, $P = 0.02$; Figs. 1b, 2). *J. virginiana* populations were thus excluded from a subsequent SAMOVA to examine whether there was any substructure in the data among *J. ashei* and *J. pinchotii*-associated populations. The amount of variation explained by different values of K was similar, with $K = 4$ accounting for the highest proportion of variation ($\phi_{CT} = 0.235$, $P = 0.001$; Figs. 1b, 2). *J. ashei*-associated populations were differentiated from *J. pinchotii*-associated populations, although the *J. ashei*-associated Guadalupe population was a separate group. For the sympatric *J. ashei*-*J. pinchotii*-associated populations, Junction grouped with *J. ashei* populations and Independence Creek was a separate group (Fig. 1b).

AFLP analyses

The two AFLP selective primer pairs, EcoRI-ACA with MseI-CAGA and MseI-CAGC, yielded 223 and 234 variable fragments, respectively, for a total of 457 anonymous markers. Interpretation of STRUCTURE results for the combined data set, using both the lnL and the calculation of delta K, identified the number of clusters at $K = 2$ (this was consistent for both models that incorporated sampling location information (LOCPRIOR) and those that did not; the LOCPRIOR model results are presented herein). These clusters correspond with the distinct division observed in the haplotype network between

Author's personal copy

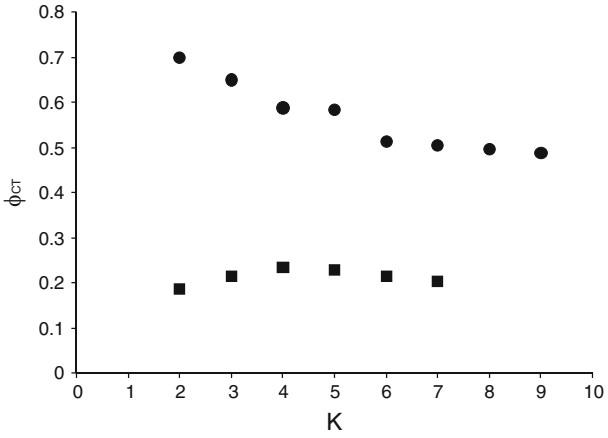


Fig. 2 SAMOVA results. Proportion of among-hosts variation (φ_{CT}) on the *y*-axis for each number of clusters, *K*, on the *x*-axis. *Circles* are results of SAMOVA including all populations; *squares* are results excluding *J. virginiana*-associated populations

J. virginiana-associated populations and all other populations (Fig. 1a), with individuals from *J. virginiana* populations clustering together and separately from all other individuals (Fig. 3). There was little evidence of other structure for values of *K* greater than two in the AFLP analysis (Fig. S2 in Supporting Information).

Discussion

For phytophagous insects that both mate and oviposit on their host plants, positive assortative mating in combination with selection resulting in increased fitness on natal hosts can generate reproductive isolation between different host-associated groups, and facilitate host race formation (Drès and Mallet 2002; Funk 1998; Via 1999). Host-associated differentiation should also be considered within the larger framework of historical biogeography, incorporating both the spatial and temporal context of population dynamics over time. This study examined patterns of genetic diversity at the scale of a group of potentially interbreeding populations of one nominal taxon, *M. gryneus*, testing host-associated differentiation as evidenced by restricted gene flow among populations corresponding with the three different host plant associations.

Patterns of neutral genetic divergence observed for mtDNA were found to be consistent with the hypothesis of host-associated differentiation. Host association was a good predictor of mtDNA sequence variation, and the results of the SAMOVA, as well as the lack of a pattern of IBD, further support the hypothesis of host-associated differentiation.

Differences at the nuclear (genomic) level are less clear. AFLP analysis using STRUCTURE supports restricted gene flow between *J. virginiana*-associated populations and all others ($K = 2$); however, there is a lack of clear structure for values of *K* greater than two, indicating much less differentiation between *J. ashei*- and *J. pinchotii*-associated populations.

Despite the clear differentiation between *J. ashei* and *J. virginiana* populations in both the mtDNA and nuclear data, introgression from *J. ashei* populations to *J. virginiana* populations may have occurred at some point in the past, given the presence of a shared mitochondrial haplotype (A3; Fig. 1a); however, this may also be an indication of shared ancestral polymorphism. In comparison, divergence between *J. ashei*- and *J. pinchotii*-associated populations appears to have occurred more recently, with shared haplotypes and

Author's personal copy

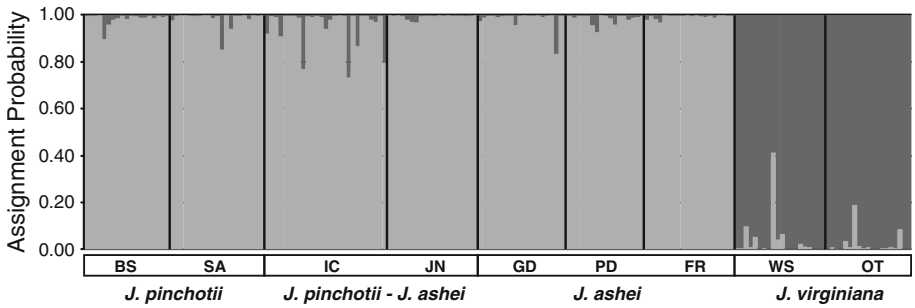


Fig. 3 STRUCTURE *bar plot* ($K = 2$). Population labels correspond with Table 1, Fig. 1. *Light grey* represents probability of assignment to Cluster 1; *dark grey* probability of assignment to Cluster 2

the lack of structure in the AFLP analysis suggesting recent or ongoing gene flow. However, *J. pinchotii* and sympatric *J. ashei*- *J.pinchotii*-associated populations (in particular San Angelo and Independence Creek) exhibit greater haplotype diversity in terms of both the number of haplotypes and the presence of private alleles compared to *J. virginiana* and *J. ashei*-associated populations, indicating some degree of differentiation from other host-associated populations.

When excluding *J. virginiana*-associated populations from the AMOVA, grouping populations by host association still explained some of the variation and IBD was again not observed, suggesting that host plant association is a better predictor of genetic differentiation between *J. ashei* and *J. pinchotii*-associated populations. However, analysis of nuclear data did not reflect the fine-scale structure revealed in the mtDNA data as evidenced by little support for clusters at K greater than two. It is not unusual, however, to find discrepancies between mitochondrial and nuclear markers (Chan and Levin 2005; Gompert et al. 2006). This may be due to a lack of resolution in the AFLP data in cases where divergence is recent and there is limited differentiation among populations (Waples and Gaggiotti 2006). In addition, differential dispersal of males and females might also exacerbate the difference between these marker types. Female philopatry with greater male dispersal can produce differentiation in maternally-inherited markers that would not be evident in data from biparentally-inherited markers (Ohshima and Yoshizawa 2010). Another possible explanation for the difference in patterns of population differentiation detected in nuclear versus mitochondrial markers is infection with bacterial endosymbionts such as *Wolbachia*, which are vertically transmitted and known to cause mating incompatibilities between infected and non-infected individuals and skew sex ratios in favor of females (Werren 1997). We tested individuals from our sampling populations and did not find evidence of *Wolbachia* infection (see Supporting Information online). In addition, we think that the patterns of discrepancy here are not necessarily what might be predicted of populations with *Wolbachia* infection, given that we might expect low mtDNA haplotype diversity within populations due to a selective sweep. We found high mitochondrial haplotype diversity both within and among populations (AMOVA results).

Comparing these patterns of population genetic differentiation with the experimental results of oviposition preference and larval performance (Downey and Nice 2011) provides a more complete picture of the evolutionary history of these butterflies and identifies possible mechanisms of divergence. For *J. virginiana*-associated populations, females did not exhibit significant preference for the natal host, although larvae had highest fitness when reared on *J. virginiana* (Downey and Nice 2011). Given eight nucleotide substitution

Author's personal copy

differences between the *J. virginiana*-associated mtDNA haplotypes and all others, it is likely that these populations experienced a relatively long period in allopatry with *J. virginiana* as the sole available host plant. Also congruent with observed genetic differentiation, the strongest division in larval performance occurs between *J. virginiana*- versus *J. pinchotii*- and *J. ashei*-associated populations. This suggests that local adaptation of *M. gryneus* on *J. virginiana*, as measured by increased larval fitness on the natal host, has evolved in allopatry. Without alternative hosts, *J. virginiana*-associated females might not have experienced selection resulting from ovipositing on the "wrong" host that would be necessary for the evolution of a strong preference for *J. virginiana*. The distinction between the *J. ashei* and *J. pinchotii* populations in terms of preference and performance varies, with *J. ashei*-associated butterflies exhibiting specialization on the natal host for both preference and performance, while *J. pinchotii*-associated populations did not distinguish between *J. ashei* and *J. pinchotii* in either preference or larval performance (although they clearly preferred these hosts over *J. virginiana*). The ambiguity in experimental preference and performance results are paralleled in the AFLP data, which failed to distinguish populations using *J. pinchotii* from populations using *J. ashei*.

Other factors might underlie the maintenance of a strong signal of specialization for *J. ashei* populations when compared to the other host associated populations. For example, because *J. ashei*-associated populations contain only a limited subset of haplotypes found also in western populations, and no private alleles, they might represent a more recent colonization from west to east. This pattern might indicate that adaptation to *J. ashei* is the derived condition, and *J. pinchotii* is the "ancestral" host (at least when considered within the range of this study). For *J. ashei* populations, the partially sympatric condition with *J. pinchotii* and proximity to *J. virginiana* hosts could mean that more than one mechanism driving differentiation is taking place, including an increased likelihood for disruptive selection in areas of primary contact (e.g., between *J. ashei* and *J. pinchotii*), or pre-zygotic isolating barriers developing in areas of secondary contact (e.g., between *J. ashei* and *J. virginiana*) as a consequence of reinforcing selection.

For populations associated with both *J. ashei* and *J.pinchotii* in sympatry, in which some degree of adaptation and preference for both hosts exists, gene flow between specialized *J. ashei*-populations and *J. pinchotii* populations could be the result of secondary contact. If the sympatric condition of host plants facilitates gene flow among populations of *Mitoura*, this could inhibit specialization and host race formation. Butterflies from the sympatric *J. ashei-J. pinchotii*-associated population at Independence Creek did not distinguish between these two hosts in preference trials, and larvae did not exhibit a clear fitness difference when reared on either host (Downey and Nice 2011). AFLP analyses indicate that gene flow might be taking place between the Independence Creek population and other *J. ashei* and *J.pinchotii* populations, which would break down any co-adapted gene complexes that increase fitness on one host over another. Preliminary reciprocal crosses between the host-associated populations conducted in the laboratory produced viable offspring and indicate that gene flow is a distinct possibility. Another factor that might inhibit divergence between these different host-associated populations is a failure of *Mitoura* to discriminate between *J. ashei* and *J. pinchotii*; both might be suitable hosts. More detailed study of areas where host species are sympatric is needed. Here we examined areas in western Texas where *J. ashei* and *J. pinchotii* are sympatric, but *J. ashei* is also sympatric with *J. virginiana* in northern Arkansas, and if *Mitoura* are found in this area, a comparsion with western populations could be enlightening.

The underlying genetic architecture for the traits of preference and performance may also influence how these traits respond to selection, and what might happen if host plants

Author's personal copy

are allopatric versus sympatric. Forister (2005) conducted crosses between two nominal species *Mitoura* that use alternate hosts, *M. nelsoni* and *M. muiri*, and found evidence of independent inheritance of adult preference and larval performance. These findings have implications when interpreting the results of studies of Texas populations of *Mitoura*. If genes underlying preference and performance are unlinked and can evolve independently, then local adaptation in allopatry (via enhanced larval performance) can take place without the evolution of preference, as observed with butterflies from *J. virginiana* populations (Downey and Nice 2011). Evidence for dominance in preference for one host over another could also help to explain the lack of two distinct host races between *J. ashei*- and *J. pinchotii*-associated populations. If secondary contact is taking place between *J. pinchotii*- and *J. ashei*-adapted butterflies in areas of host sympatry, and if preference for *J. ashei* is dominant (or codominant) to preference for *J. pinchotii*, then this could disrupt assortative mating and influence whether distinct host races can be maintained in a single population, and whether host formation will take place.

Ecological speciation can result from both local adaptation taking place in allopatry (Funk 1998) as well as from specialization on alternate resources in sympatry (Nosil et al. 2009). Both can work in concert to influence reproductive isolation among differently adapted populations and potentially facilitate speciation (Feder et al. 2003). At early stages of differentiation, there are many factors that can work to inhibit, as well as drive, divergence. For example, areas of host sympatry present an opportunity for gene flow to take place among populations of plant-feeding insects that might otherwise specialize on, or become locally adapted to, a single host. If female oviposition preference evolves, and nonrandom mating takes place on different hosts, then divergent natural selection may result in more than one host race. The underlying genetic architecture will also influence the outcome of specialization and host race formation, since independent inheritance of preference and performance traits means that preference might not always evolve in concert with local adaptation to a host. The *Mitoura* species complex offers opportunities to compare different stages of divergence at various phylogeographic scales in terms of the multiple factors (e.g., ecological specialization, adaptation in allopatry, and gene flow) that can drive or inhibit population genetic divergence and speciation.

Acknowledgments The authors would like to thank C. Campbell and L. Downey for field assistance, and N. Martin and J. Ott for reviewing earlier drafts of this manuscript. We appreciate the constructive comments from three anonymous reviewers. Funding support was provided by the Texas Academy of Science, the George Strait- Freeman Fellows Scholarship, and the Durrenberger Scholarship for Women in Science.

References

Adams RP (2008) *Juniperus* of Canada and the United States: taxonomy, key, and distribution. Phytologia 90:255–314

Arrigo N, Tuszynski JW, Ehrich D, Gerdes T, Alvarez N (2009) Evaluating the impact of scoring parameters on the structure of intra-specific genetic variation using RawGeno, an R package for automating AFLP scoring. BMC Bioinformatics 10:33–47

Berner D, Grandchamp A, Hendry AP (2009) Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. Evolution 63:1740–1753

Butlin RK, Galindo J, Grahame JW (2008) Sympatric, parapatric or allopatric: the most important way to classify speciation? Philos Trans R Soc B-Biol Sci 363:2997–3007

Caterino MS, Sperling FAH (1999) Papilio phylogeny based on mitochondrial cytochrome oxidase I and II genes. Mol Phylogenet Evol 11:122–137

Author's personal copy

Chan KMA, Levin SA (2005) Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. Evolution 59:720–729
Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9:1657–1659
Coyne J, Orr H (2004) Speciation. Sinauer Associates, New York
Downey MH, Nice CC (2011) Experimental evidence of host race formation in Mitoura butterflies (Lepidoptera: Lycaenidae). Oikos 120:1165–1174
Drès M, Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. Philos Trans R Soc B Biol Sci 357:471–492
Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Heled J, Kearse M, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2010) Geneious v5.1, available from http://www.geneious.com
Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define the genetic structure of populations. Mol Ecol 11:2571–2581
Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620
Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinformatics 1:47–50
Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164:1567–1587
Feder JL, Berlocher SH, Roethele JB et al (2003) Allopatric genetic origins for sympatric host-plant shifts and race formation in *Rhagoletis*. Proc Natl Acad Sci USA 100:10314–10319
Feder JL, Xie XF, Rull J et al (2005) Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis*. Proc Natl Acad Sci USA 102:6573–6580
Felsenstein J (1980) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? Evolution 35:124–138
Forister ML (2004) Oviposition preference and larval performance within a diverging lineage of lycaenid butterflies. Ecol Entomol 29:264–272
Forister ML (2005) Independent inheritance of preference and performance in hybrids between host races of Mitoura butterflies (Lepidoptera: Lycaenidae). Evolution 59:1149–1155
Funk DJ (1998) Isolating a role for selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. Evolution 52:1744–1759
Funk DJ, Nosil P (2008) Comparative analyses of ecological speciation. In: Tilmon KJ (ed) Specialization, speciation and radiation: the evolutionary biology of herbivorous insects. The University of California Press, Berkeley, pp 117–135
Funk DJ, Filchak KE, Feder JL (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116:251–267
Gavrilets S, Li H, Vose MD (1998) Rapid parapatric speciation on holey adaptive landscapes. Proc Biol Sci 265:1483–1489
Gompert Z, Nice CC, Fordyce JA, Forister ML, Shapiro AM (2006) Identifying units for conservation using molecular systematic: the cautionary tale of the Karner blue butterfly. Mol Ecol 15:1759–1768
Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. Mol Ecol Res 9:1322–1332
Johnson KD (1981) Revision of the Callophryina of the world with phylogenetic and biogeographic analyses (Lepidoptera: Lycaenidae). Dissertation, New York University
Lenormand T (2002) Gene flow and the limits to natural selection. Trends Ecol Evol 17:183–189
Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York
Nice CC, Shapiro AM (2001) Population genetic evidence of restricted gene flow between host races in the butterfly genus Mitoura (Lepidoptera : Lycaenidae). Ann Entomol Soc Am 94:257–267
Nosil P, Sandoval CP, Crespi BJ (2006) The evolution of host preference in allopatric vs. parapatric populations of *Timema cristinae* walking-sticks. J Evol Biol 19:929–942
Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Mol Ecol 18:375–402
Ohshima I, Yoshizawa K (2010) Differential introgression causes genealogical discordance in host races of *Acrocercops transecta*. Mol Ecol 19:2106–2119
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
Raesaenen K, Hendry AP (2008) Disentangling interactions between adaptive divergence and gene flow when ecology drives diversification. Ecol Lett 11:624–636
Rundle HD, Nosil P (2005) Ecological speciation. Ecol Lett 8:336–352
Schluter D (2001) Ecology and the origin of species. Trends Ecol Evol 16:372–380

Author's personal copy

Scott J (1992) The butterflies of North America: a natural history and field guide. Stanford University Press, Palo Alto

Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P (1994) Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Ann Entomol Soc Am 87:651–701

Slatkin M (1987) Gene flow and the geographic structure of natural populations. Science 236:787–792

Sokal RR, Rohlf FJ (1995) Biometry: the principles and practice of statistics in biological research. W.H. Freeman and Company, New York

Swofford DL (2003) PAUP*. Phylogenetic analysis using parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland

Thorpe RS, Surget-Groba Y, Johansson H (2008) The relative importance of ecology and geographic isolation for speciation in anoles. Philos Trans R Soc B Biol 363:3071–3081

Tregenza T (2002) Divergence and reproductive isolation in the early stages of speciation. Genetica 116:291–300

Via S (1999) Reproductive isolation between sympatric races of pea aphids I: gene flow restriction and habitat choice. Evolution 53:1446–1457

Via S (2009) Natural selection in action during speciation. Proc Natl Acad Sci USA 106:9939–9946

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol Ecol 15:1419–1439

Werren JH (1997) Biology of *Wolbachia*. Ann Rev Entomol 42:587–609

Wright S (1943) Isolation by distance. Genetics 28:114–138