Molecular Phylogenetics and Evolution 48 (2008) 481–490

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Recent colonization and radiation of North American *Lycaeides* (*Plebejus*) inferred from mtDNA

Zachariah Gompert [a,b,*], James A. Fordyce [c], Matthew L. Forister [d], Chris C. Nice [a]

[a] Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX 78666, USA
[b] Department of Botany, University of Wyoming, 1000 E. University Avenue, Laramie, WY 82071, USA
[c] Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA
[d] Department of Natural Resources and Environmental Science, University of Nevada, Reno, NV 89512, USA

ARTICLE INFO

Article history:
Received 5 September 2007
Revised 24 April 2008
Accepted 24 April 2008
Available online 1 May 2008

Keywords:
Coalescent
Mitochondrial DNA
Lycaeides
Bayesian skyline plot
Pleistocene
Holarctic biogeography

ABSTRACT

North American *Lycaeides* populations exhibit remarkable variation in ecological, morphological, and behavioral characters, as well as an established history of introgressive hybridization. We examined mitochondrial DNA variation from 55 Eurasian and North American *Lycaeides* populations using molecular phylogenetics and coalescent-based methods in order to clarify the evolutionary and demographic history of this polytypic group. Specifically we addressed the following questions: (1) Do mitochondrial alleles sampled from North America form a monophyletic group, which would be expected if North American *Lycaeides* were descended from a single Eurasian ancestor? (2) When did *Lycaeides* colonize North America? and (3) What is the demographic history of North American *Lycaeides* since their colonization? Bayesian maximum likelihood methods identified three major mitochondrial lineages for *Lycaeides*; each lineage contained haplotypes sampled from both Eurasia and North America. This suggests a complex colonization history for *Lycaeides*, which likely involved multiple founding lineages. Coalescent-based analyses placed the colonization of North America by Eurasian *Lycaeides* sometime during or after the late Pliocene. This was followed by a sudden increase in population size of more than an order of magnitude for the North American population of *Lycaeides* approximately 100,000–150,000 years before the present. These mitochondrial data, in conjunction with data from previous ecological, morphological, and behavioral studies, suggest that the diversity observed in *Lycaeides* in North America is the result of a recent evolutionary radiation, which may have been facilitated, in part, by hybridization.

1. Introduction

Gene genealogies can be used to infer historical evolutionary and demographic processes operating within species or groups of closely related species (Hein et al., 2005). For example, the effects of past geological events (e.g. Pleistocene climatic cycles) on the distribution and abundance of species have been inferred increasingly through the analysis of molecular gene genealogies (e.g. Avise and Walker, 1998; Knowles, 2001; Griswold and Baker, 2002; Pinho et al., 2007; Spellman et al., 2007). Coalescent-based analyses have been especially useful for the inference of historical processes, as these analyses utilize the information present in DNA sequence variation (Hein et al., 2005). Data on historical processes are necessary for understanding current evolutionary and ecological patterns and processes. Here we report the results of a large-scale survey of mitochondrial DNA (mtDNA) variation in the polytypic *Lycaeides* butterfly species group. This is a diverse genus with a complex evolutionary history, including one of the few documented cases of homoploid hybrid speciation in animals (Nice and Shapiro, 1999; Nice et al., 2005; Gompert et al., 2006a,b).

Butterflies of the genus *Lycaeides* are members of the family Lycaenidae, a family notable for its species richness, diversity of natural histories, and extreme local specialization (Downey and Dunn, 1964; Grimaldi and Engel, 2005). *Lycaeides* is a circumpolar, holarctic genus (Nice and Shapiro, 1999). Nabokov (1943, 1949) revised the classification of North American *Lycaeides* based on qualitative differences in wing pattern and quantitative differences in the size and shape of the male genitalia. He recognized two species: *L. idas* (Linnaeus) and *L. melissa* (W.H. Edwards). The former has a holarctic distribution, while that latter is a North American endemic. Gompert et al. (2006b) identified a third, yet unnamed, North American *Lycaeides* species, which is a homoploid hybrid species distributed in the alpine habitat of the Sierra Nevada. This species originated via hybridization between *L. idas* and *L. melissa* (Gompert et al., 2006b). *Lycaeides idas* and *L. melissa* are broadly sympatric in several regions of North America (Fig. 1).

* Corresponding author. Address: Department of Botany, University of Wyoming, 1000 E. University Avenue, Laramie, WY 82071, USA. Fax: +1 307 766 2851.
E-mail address: zgompert@uwyo.edu (Z. Gompert).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.04.035

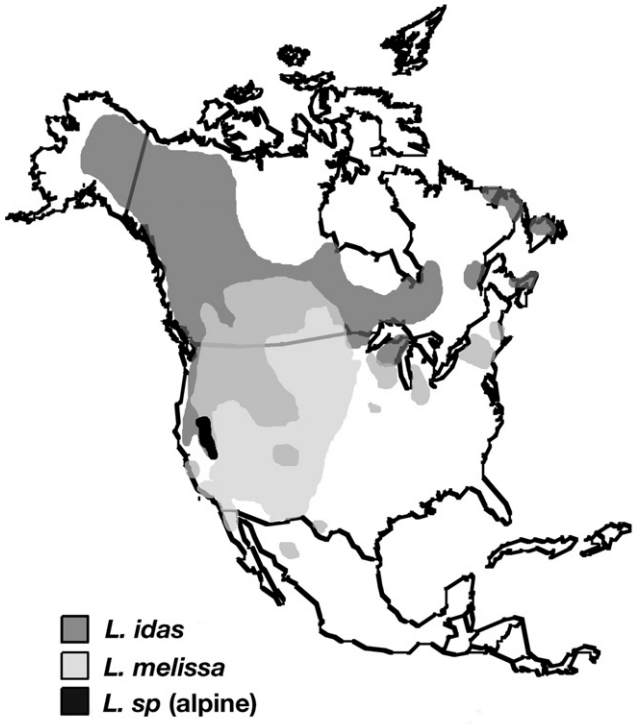

Fig. 1. Range map for North American *Lycaeides*. The approximate ranges of *L. idas*, *L. melissa*, and the unnamed hybrid species are shown in dark gray, light gray, and black, respectively. The range maps for *L. idas* and *L. melissa* follow Nabokov (1949),and Stanford and Opler (1996). The range map of the hybrid species is based on field-work conducted by Z.G., J.A.F., and C.C.N. during the summer of 2006.

Approximately 17 subspecies have been described based on differences in wing pattern, male genitalia, habitat specificity, and larval host plant use (Nabokov, 1949; Scott, 1986; Opler, 1992; Lane and Weller, 1994). Two of these subspecies have attracted the attention of conservation biologists. The Karner blue butterfly (*L. m. samuelis*) is listed in the United States as an endangered species (US Fish and Wildlife Service, 1992, 2003); the Lotis blue (*L. i. lotis*) has the same status, but is likely extinct (Arnold, 1993). *Lycaeides* was previously placed within the genus *Plebejus*, and is still referred to as *Plebejus* by some authors (e.g. Scott, 1986).

North American *Lycaeides* are diverse and exhibit remarkable ecological specialization among taxa and populations. This diversity suggests that North American *Lycaeides* could be thought of as an evolutionary radiation. While *Lycaeides* at a specific locality are generally monophagous, there is considerable variation in the specific host-plant species used by geographically separated populations (Scott, 1986; Brock and Kaufman, 2003). This variation in host-plant use is coupled with inter-population variation in female host-plant fidelity and host-plant preference (Nice et al., 2002; Gompert et al., 2006b). Populations also differ in the size and placement of wing pattern elements (Fordyce et al., 2002). Some of these population-level differences in wing pattern elements are discernable by male *Lycaeides* and play a role in mate recognition (Fordyce et al., 2002). Population-level variation in egg morphology (Forister et al., 2006) and male genital morphology (Nice and Shapiro, 1999; Nice et al., 2005; Lucas et al., 2008) has also been documented throughout North America; at present, an adaptive role for this morphological variation has not been demonstrated.

Nice et al. (2005) examined geographic patterns of mtDNA (AT-rich region) variation in *Lycaeides* across parts of North America in order to elucidate the evolutionary history of this ecologically, behaviorally, and morphologically diverse group. They found that mitochondrial variation was partitioned geographically and only loosely followed taxonomic boundaries. Specifically, three mitochondrial clades were detected (Nice et al., 2005). Several populations at the geographical boundaries of identified clades displayed discordant patterns of mtDNA and morphological variation. For example, populations of *L. m. samuelis* (the endangered Karner blue butterfly) in the western portion of their range were morphologically indistinguishable from those in the eastern portion of their range, but were assigned to a different mtDNA clade than the eastern *L. m. samuelis* populations. Nice et al. (2005) suggested that these patterns were indicative of isolation of *Lycaeides* in at least three refugia during Pleistocene glacial maxima followed by post-Pleistocene range expansion, secondary contact, and introgressive hybridization.

The hypothesis of introgressive hybridization at two of these contact zones has been tested using nuclear genetic markers. Gompert et al. (2006a) determined that *L. m. samuelis* populations in the western portion of their range possess mtDNA haplotypes from a different mtDNA clade than the eastern populations because of mitochondrial introgression. Examination of genomic divergence between the eastern and western *L. m. samuelis* populations using AFLP markers revealed that mitochondrial introgression had occurred from *L. m. melissa* to the western *L. m. samuelis* populations with minimal to no nuclear introgression (Gompert et al., 2006a). A second contact zone occurs in the Sierra Nevada. *Lycaeides* populations in the alpine (i.e. above tree-line) in the Sierra Nevada possess wing pattern elements indicative of *L. melissa*, *L. idas*-like mtDNA haplotypes, and male genitalic morphology intermediate to these two taxa (Nice et al., 2005; Gompert et al., 2006b). Based on extensive molecular data (i.e. AFLP markers, nuclear sequence data, mtDNA sequence data, and microsatellite markers), Gompert et al. (2006b) concluded this discordant pattern resulted from hybridization between *L. idas* and *L. melissa* and that the alpine-associated populations represent a species of hybrid origin. Thus, North American *Lycaeides* are highly specialized and have a history of hybridization, which, in at least one case, resulted in the origin of a species.

To better understand the potentially complex and dynamic processes that have influenced and continue to act on North American *Lycaeides*, data regarding the evolutionary and demographic history of this genus are required. Particularly we require information concerning the colonization and subsequent diversification of *Lycaeides* in North America. We do not know whether the specialized forms that exist in North America began to differentiate in Eurasia, nor do we know if the hybridizing *L. idas* and *L. melissa* of North America are each others' closest relatives (*L. idas* also occurs in Eurasia, and the Eurasian species *L. argyrognomon* is thought to be closely related to *L. idas* and *L. melissa* (Nice et al., 2005)). Here we examine mtDNA data using molecular phylogenetics and coalescent-based models to address the following questions regarding the evolutionary and demographic history of North American *Lycaeides*: (1) Are North American *Lycaeides* the result of colonization from Eurasia by a single ancestor, or did multiple differentiated forms colonize North America? (2) When did colonization of North America occur? (3) What is the demographic history of North American *Lycaeides* since their colonization?

2. Methods

2.1. Collection and DNA extraction

Adult *Lycaeides* were obtained from 55 populations from across North America, Europe, and Asia (Table 1). North American taxa sampled included *Lycaeides idas*, *L. melissa*, the hybrid species from the Sierra Nevada, and the Warners Mountain entity. Eurasian taxa from which samples were obtained include *L. idas* and *L. argyrogn-*

Table 1
Population data

Locality #	Taxa	Locality	Latitude	Longitude	mtDNA (n)
1	*L. i. alaskensis*	Koyukuk River, AK	67° 27′ 10″N	150° 03′ 30″W	h63(3)
2	*L. i. anna*	Donner Pass, CA	39° 18′ 53″N	120° 20′ 57″W	h18(1) h19(1) h20(5) h21(3)
3		Leek Springs, CA	38° 37′ 59″N	120° 14′ 24″W	h18(8) h29(2)
4		Trap Creek, CA	39° 22′ 43″N	120° 40′ 27″W	h20(5) h21(1) h43(1) h44(1) h45(2)
5		Yuba Gap, CA	39° 29′ 24″N	120° 35′ 39″W	H20(8) h48(1) h49(1)
6	*L. i. azureus*	Indian Valley, CA	40° 45′ 33″N	122° 58′ 25″W	h23(4) h24(2) h25(1) h26(3)
7	*L.i. degener*[a]	Lerida, Spain	Not available	Not available	h30(1)
8	*L. i. idas*[a]	S. Urals, Russia	Not available	Not available	h56(1)
9	*L. i. magnagraeca*[a]	Galich Mt., Macedonia	Not available	Not available	h51(2) h76(1)
10	*L. i. ricei*	Cave Lake, CA	41° 58′ 46″N	120° 12′ 25″W	h01(10)
11		Deadfall Mdw., CA	41° 14′ 11″N	122° 57′ 38″W	h16(6) h17(4)
12		Soda Mt., OR	42° 05′ 19″N	122° 28′ 58″W	h01(4) h27(1)
13		Josephine Co., OR	42° 25′ 35″N	123° 27′ 41″W	h28(4)
14		Marble Mts., CA	41° 49′ 40″N	122° 44′ 52″W	h31(2) h32(1) h33(7)
15		Mt. Ashland, OR	42° 04′ 52″N	122° 43′ 16″W	h01(10)
16	*L. i. scudderii*	Alberta, Canada	50° 42′ 37″N	114° 37′ 42″W	h01(5)
17	*L. i. sublivens*	Mineral Creek, CO	37° 41′ 00″N	107° 41′ 00″W	h68(2)
18		San Juan Co., CO	37° 56′ 00″N	107° 34′ 17″W	h01(5) h38(1) h64(1)
19	*L. m. anneta*	Alta, UT	40° 35′ 33″N	111° 37′ 27″W	h01(2) h46(8)
20		Teton Mts., WY	43° 51′ 13″N	110° 45′ 53″W	h01(3) h06(2)
21	*L. m. inyoensis*	Big Pine, CA	37° 10′ 23″N	118° 16′ 44″W	h01(10)
22		Manzanar Airfield, CA	36° 44′ 19″N	118° 08′ 03″W	h01(7)
23	*L. m. melissa*	Beckwourth Pass, CA	39° 47′ 35″N	120° 06′ 38″W	h01(1) h02(8)
24		Brandon, SD	43° 36′ 29″N	96° 34′ 39″W	h01(5) h04(2) h05(1) h06(1) h07(1)
25		Cedarville, CA	41° 31′ 47″N	120° 10′ 11″W	h01(1) h08(8) h09(1)
26		Elmore Co., ID	42° 53′ 40″N	116° 06′ 30″W	h01(1) h02(1) h09(1)
27		Garderville, CA	38° 48′ 54″N	119° 46′ 44″W	h01(3) h02(7)
28		Granite Mts., NV	40° 53′ 18″N	119° 33′ 60″W	h67(3)
29		Gowanda, CO	40° 18′ 00″N	104° 50′ 00″W	h01(1)
30		Montague, CA	41° 46′ 21″N	122° 28′ 38″W	h01(9) h34(1)
31		Pitkin Co., CO	39° 10′ 00″N	106° 47′ 00″W	h69(4)
32		Sierravalley, CA	39° 37′ 48″N	120° 21′ 40″W	h01(10)
33		Spring Creek, SD	44° 25′ 04″N	100° 24′ 50″W	h01(8) h04(1) h39(1)
34		Verdi, NV	39° 03′ 01″N	119°55′ 48″W	h02(10) h47(1)
35	*L. m. samuelis*	Fish Lake, WI	45° 44′ 14″N	92° 46′ 43″W	h01(5)
36		Fort McCoy, WI	43° 47′ 59″N	90° 49′ 59″W	h01(5)
37		Indiana Dunes, IN	41° 40′ 07″N	87° 03′ 04″W	h22(5)
38		Necedah, WI	44° 04′ 00″N	90° 11′ 20″W	h01(5)
39		Saratoga, NY	43° 03′ 24″N	73° 48′ 47″W	h22(5)
40	*Lycaeides* Alpine sp.	Bodie Hills, CA	38° 16′ 43″N	119° 05′ 40″W	h03(2)
41		County Line Hill, CA	37° 27′ 51″N	118° 11′ 35″W	h02(3) h10(1) h11(2) h12(2) h13(2)
42		Carson Pass, CA	38° 42′ 47″N	120° 01′ 17″W	h03(8) h14(1) h15(1)
43		Jeff Davis Pk., CA	38° 38′ 20″N	119° 54′ 05″W	h03(4) h42(4) h50(2)
44		Lake Emma, CA	38° 16′ 54″N	119° 28′ 59″W	h03(10)
45		Mt. Rose, NV	39° 19′ 21″N	119° 55′ 48″W	h03(1) h35(9)
46		Piute Pass, CA	37° 14′ 11″N	118° 40′ 08″W	h03(1)
47		Sonora Pass, CA	38° 19′ 54″N	119° 38′ 02″W	h03(6) h40(3) h41(1)
48		South Fork, CA	37° 12′ 38″N	118° 34′ 07″W	h03(8), h42(1)
49		Sweetwater Mts., CA	38° 27′ 03″N	119° 20′ 04″W	h03(9) h40(1)
50		Tioga Crest, CA	37° 57′ 57″N	119° 15′ 25″W	h03(6) h42(4)
51		Wassuk Mts., NV	38° 38′ 54″N	118° 49′ 30″W	h03(1)
52	*Lycaeides* Warner Mts. entity	Eagle Pk., CA	41° 15′ 38″N	120° 13′ 11″W	h01(10)
53	*L. a. argyrognomon*[a]	Slovakia	Not available	Not available	h02(1) h36(4) h56(3) h62(1)
54	*L. a. jautica*[a]	Magadan, Russia	Not available	Not available	h57(3)
55	*P. a. argus*[a]	Atali, Russia	Not available	Not available	h58(2)
56		Russia	Not available	Not available	h65(1)
57		Plastovek, Slovakia	Not available	Not available	h59(2) h60(1)
58	*P. a. orientalis*[a]	Raec, Macedonia	Not available	Not available	h59(4) h75(1)
59	*P. icarioides*	Sweetwater Mts., CA	38° 27′ 03″N	119° 20′ 04″W	h74(1)
60	*P. saepiolus*	Tioga Crest, CA	37° 57′ 57″N	119° 15′ 25″W	h71(1) h72(1)
61	*P. shasta*	Tioga Crest, CA	37° 57′ 57″N	119° 15′ 25″W	h73(3)
62	*P. sephirus magnificus*[a]	Babuna mt., Macedonia	Not available	Not available	h53(1)

Locality number, nominal taxa, sample locality, latitude, longitude, and mtDNA haplotypes are given for each population.
[a] Eurasian.

omon. Both males and females were collected from most *Lycaeides* populations; males were collected from *L. m. samuelis* populations in accord with USFWS permit PRT842392. Additional specimens were collected from several species in the genus *Plebejus*, which are closely related to *Lycaeides* (Kandul et al., 2004). This includes *P. saepiolus*, *P. shasta*, and *P. icarioides* from North America and *P. argus* and *P. sephirus* from Eurasia. Taxonomic identification of collected specimens was made based on morphological (wing patterns and male genitalic variation), ecological, and behavioral characters as well as geographical data. DNA was isolated following the methods of Hillis et al. (1996) and Brookes et al. (1997).

2.2. Molecular methods

We sequenced portions of the mitochondrial genes cytochrome oxidase *c* subunit I (COI) and cytochrome oxidase *c* subunit II (COII)

for 1–11 individuals from each sampled population (Table 1). 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 and Sperling, 1999). This yielded fragments of approximately 450 and 550 base pairs for COI and COII, respectively. Fluorescently labeled dideoxy terminators were used for single stranded sequencing reactions for both COI and COII. Labeled amplicons were separated and visualized using a Beckman 8800 automated sequencer (Beckman Coulter Inc.). Sequences were aligned using Sequencher 4.2.2. or by eye. Sixty-nine unique haplotypes were identified (GenBank Accession Nos. EU409321–EU409362) (Table 1).

2.3. Phylogeny reconstruction and Bayesian hypothesis tests

To examine the relationship among the sampled haplotypes from North American, European, and Asian *Lycaeides* populations, gene genealogies for the combined data from COI and COII were constructed using Bayesian maximum likelihood methods. *Plebejus sephirus* was set as the outgroup. We evaluated three alternative partition schemes for the sequence data: (1) a single partition, (2) separate partitions for each gene region, and (3) separate partitions for each gene region and codon position. Sequence evolution models for each partition were selected using Modeltest 3.7 based on Akaike Information Criterion (AIC) (Posada and Crandall, 1998). When the data were placed in a single partition, the HKY+I+γ model was selected, while the GTR+I+γ model was selected for both COI and COII when the data were partitioned by gene region. When the data were partitioned by gene region and codon position the models selected were TrN+I, F81, TIM+γ, HKY+I, TrN+I, and TIM for COI 1st codon position, COI 2nd codon position, COI 3rd codon position, COII 1st codon position, COII 2nd codon position, and COII 3rd codon position, respectively.

Bayesian phylogenetic analyses were conducted using MrBayes ver 3.1.2 (Huelsenbeck and Ronquist, 2001) for each of our three data partition models. For each model, Markov Chain Monte Carlo was performed with one cold chain and three hot chains run for 8×10^6 generations with a burnin of 5×10^5 generations. Two independent runs were conducted for each model, and the standard deviation of the split frequencies was monitored to ensure convergence upon the stationary distribution. We used Bayes factors (see Ronquist et al., 2005) to compare the posterior odds of our preferred Bayesian gene genealogy with the data partitioned by gene region and codon position to both our preferred Bayesian gene genealogy with the data in a single partition and with the data partitioned by gene region. Bayes factors were calculated by taking the difference between the marginal log likelihood values of the model with data partitioned by gene region and codon position, M_1, and each of the models with fewer data partitions, M_0 (see Nylander et al., 2004).

To explicitly test the hypothesis that North American *Lycaeides* are derived from a single colonization event from Eurasia, and thus have a single old world ancestor, we used Bayes factors (Ronquist et al., 2005) to compare the posterior odds of our preferred Bayesian gene genealogy with no topological constraints to genealogies with either North American *Lycaeides* constrained to be monophyletic or North American and Eurasian *Lycaeides* constrained to be reciprocally monophyletic. In both cases we used a model with the data partitioned by gene region and codon position (see Section 3). If North America was colonized by a single Eurasian species, then North American *Lycaeides* should be monophyletic. If this occurred sufficiently long ago, then North American and Eurasian *Lycaeides* should be reciprocally monophyletic. A single haplotype (h02) was shared between North America and Eurasian populations (see Section 3). Two copies of this haplotype were included in these analyses, so the haplotype could be constrained to be in both the North American and Eurasian clades. Gene genealogies for the constrained and unconstrained trees were produced as described in the preceding paragraph using MrBayes ver 3.1.2 (Huelsenbeck and Ronquist, 2001). Bayes factors were calculated by taking the difference between the marginal log likelihood values of the unconstrained topology, M_1, and the constrained topology, M_0 (see Nylander et al., 2004; Brandley et al., 2005; Hedin and Bond, 2006).

To determine if our preferred Bayesian gene genealogy, which was obtained using a model that partitioned the sequence data by gene region and codon position (see Section 3), conformed to a molecular clock, we ran an additional Bayesian phylogenetic analysis using MrBayes ver 3.1.2 (Huelsenbeck and Ronquist, 2001). This analysis was performed with a strict molecular clock enforced and an exponential prior for branch lengths. We then compared the posterior odds of our preferred Bayesian gene genealogy with that of the constrained Bayesian gene geology using a Bayes factor. We did this by taking the difference between the marginal log likelihood values of the model with the molecular clock enforced, M_1, and the unconstrained model, M_0 (Nylander et al., 2004). The constrained model was selected as M_1, because it resulted in a greater marginal likelihood than the unconstrained model.

2.4. Estimation of TMRCA

We used the software BEAST v1.4.2 (Drummond and Rambaut, 2006) to estimate the time to the most recent ancestor (TMRCA) for the extant mtDNA variation for North American and Eurasian haplotypes. As both North American and Eurasian haplotypes occur in all three major *Lycaeides* mitochondrial lineages (see Section 3), the TMRCA for North American mtDNA haplotypes should be the same as that for both North American and Eurasian mtDNA haplotypes. Further, as a single haplotype is shared between North American and Eurasian populations, the TMRCA provides an estimate of the upper bounds on the divergence between North American and Eurasian *Lycaeides*. TMRCA was estimated under Bayesian skyline plot and constant population size demographic models (see below and Section 3). All sampled *Lycaeides* haplotypes were included in this analysis. The SRD06 model of sequence evolution was used; the SRD06 model has fewer parameters than the GTR+I+γ model, but has been shown to provide a better fit for protein coding sequence data (Shapiro et al., 2006). The analysis was implemented under a relaxed clock with the rate for each branch drawn from a lognormal distribution. The Markov Chain Monte Carlo algorithm was iterated for 5×10^6 generations with a burn-in of 5×10^5 generations. Two independent runs were conducted to ensure convergence upon the stationary distribution. Effective sample size (ESS) was examined for all parameters to determine if the MCMC chain was run for enough generations to obtain sufficient independent samples for all parameters from the posterior distribution. The program TRACER v1.3 (Rambaut and Drummond, 2004) was used to compile and visualize the results from BEAST v1.4.2.

2.5. Estimation of population growth

Several mitochondrial haplotypes are shared among North American *Lycaeides* and all three major mitochondrial lineages contain mitochondrial haplotypes from multiple *Lycaeides* species (see Section 3); these patterns are likely due to substantial mitochondrial gene flow among North American *Lycaeides* (Gompert et al., 2006a, 2006b). Therefore, using molecular variation to make demographic inferences for each species separately could be problematic. Instead, we have chosen to treat North American *Lycaeides* as a single interbreeding unit when estimating demographic parameters. In particular, we employ a Bayesian skyline plot to estimate θ ($2 \times$ effective population size $\times$ mutation rate) through

time for North American *Lycaeides*. Bayesian skyline plots (see below) are generally used to estimate the demographic parameter θ through time for a given interbreeding population or species (Drummond et al., 2005; Crandall et al., 2008). We believe this is justified given the pattern of variation for mtDNA observed in North American *Lycaeides*. While our analysis may inflate our estimate of θ, it should have little influence on the rate of change in θ through time, which is the relevant parameter for this study.

If *Lycaeides* has expanded its range, diversified, and/or invaded new habitats, following its colonization of North America, there should be evidence of population growth in North American *Lycaeides*. To examine the demographic history of North American *Lycaeides* we generated a Bayesian skyline plot (Drummond et al., 2005). The Bayesian skyline plot provides a coalescent-based estimate of the demographic parameter θ ($2 \times$ effective population size $\times$ mutation rate), through time (Strimmer and Pybus, 2001; Drummond et al., 2005). This method does not require a pre-specified parametric demographic model. Furthermore, the Bayesian skyline plot, unlike the generalized skyline plot, provides credibility intervals on θ that take into account both phylogenetic and coalescent uncertainty (Drummond et al., 2005). The software BEAST v1.4.2 (Drummond and Rambaut, 2006) was used to generate a Bayesian skyline plot. For this analysis only mtDNA haplotypes that were sampled in North America were included; this includes haplotypes found both in North America and Eurasia. The SRD06 model of sequence evolution was used (Shapiro et al., 2006). The analysis was implemented under a relaxed clock with the rate for each branch drawn from a lognormal distribution. The Markov Chain Monte Carlo algorithm was iterated for 2×10^7 generations with a burn-in of 2×10^6 generations. Two independent runs were conducted to ensure convergence upon the stationary distribution. Effective sample size (ESS) was examined for all parameters to determine if the MCMC chain was run for enough generations to obtain sufficient independent samples for all parameters from the posterior distribution. The program TRACER v1.3 (Rambaut and Drummond, 2004) was used to compile and visualize the results from BEAST v1.4.2. We re-ran this analysis with all Eurasian *Lycaeides* haplotypes included. Including the Eurasian haplotypes did not alter the results of the analysis (results not shown).

To further examine the demographic history of North American *Lycaeides*, we compared the fit of five parametric demographic models to our mtDNA data. These models were as follows: constant population size, exponential growth, logistic growth, expansion growth, and piecewise constant population size with a single change in population size (see Pybus and Rambaut, 2002 for model details). Likelihoods for each model were estimated using the software Genie v3.0 (Pybus and Rambaut, 2002). The ultrametric maximum likelihood tree used for these analyses was generated using Paup 4.0b10 (Swofford, 2002) under the TrN+I+γ sequence evolution model, which was selected based on AIC results from ModelTest 3.7 when the *Lycaeides* mtDNA data was not divided into partitions (Posada and Crandall, 1998). A molecular clock was enforced for the maximum likelihood tree estimation. Likelihoods for the five demographic models were estimated by Genie v3.0 (Pybus and Rambaut, 2002) using the Powell algorithm. Model likelihoods were compared using corrected Akaike Information Criterion (AIC_C).

3. Results

3.1. Phylogeny reconstruction

The model with data partitioned by gene region and codon position resulted in an increase in 308.23 and 249.61 likelihood units relative to the models with a single partition and with the data partitioned by gene region, respectively (Table 2). Following the criteria of Kass and Raftery (1995), this magnitude of difference indicates that we should favor the model with data partitioned by gene region and codon position and that we have "decisive" evidence against the alternative models with fewer data partitions. "Strong" to "decisive" evidence against a hypotheses based on Bayes factors provides a similar level of support for rejecting a null hypotheses as a *p*-value of 0.001 (Goodman, 1999). Thus, we present our phylogenetic results based on the model with sequence data partitioned by gene region and codon position.

Bayesian maximum likelihood analysis strongly supported the monophyly of the sampled *Lycaeides* mitochondrial haplotypes (posterior probability 1.00; Fig. 2). Haplotypes sampled from the Eurasian species *P. argus* formed a monophyletic clade sister to *Lycaeides* (Fig. 2). Thus, based on this analysis, the genus *Plebejus* appears to be paraphyletic with respect to *Lycaeides*. Within *Lycaeides* three differentiated mitochondrial lineages were detected (Fig. 2). However these lineages did not correspond to nominal species (*L. idas*, *L. melissa*, the alpine hybrid species, and *L. argyrognomon*) or geographic regions (North America, Asia, and Europe). Lineage I consisted of individuals sampled from *L. i. anna*, *L. i. azureus*, *L. i. ricei*, and the alpine species in North America, as well as an *L. i. degener* individual from Spain. Lineage II included individuals collected from *L. m. melissa* populations in North America and both *L. i. idas* and *L. argyrognomon* populations from Asia. Within this clade haplotype h02 was shared between North American *L. m. melissa* and Asian *L. argyrognomon*, and haplotype h56 was shared between Asian *L. i. idas* and *L. argyrognomon*. Lineage III consisted of individuals sampled from *L. i. alaskensis*, *L. i. ricei*, *L. i. sublivens*, *L. m. annetta*, *L. m. inyoensis*, *L. m. melissa*, and *L. m. samuelis* populations from North America, as well as *L. argyrognomon*, *L. a. jautica*, and *L. i. magnagraeca* populations from Asia. These patterns suggest recent gene flow between Eurasian and North American taxa, and thus, suggest a recent origin for the variation observed in North America.

There were several North American populations with haplotypes from two major evolutionary lineages. The *L. m. melissa* populations at Beckwourth Pass, CA, Elmore Co., ID, and Gardnerville, CA possessed haplotypes from both lineage II and lineage III. The *L. i. ricei* population at Marble Mts., CA had haplotypes from lineages I and III, and the alpine population at County Line Hill in the White Mountains had haplotypes from lineages I and II (Table 1 and Fig. 2).

Bayes factors were used to compare the posterior odds of our unconstrained Bayesian mitochondrial genealogy to genealogies with: (1) North American *Lycaeides* constrained to be monophyletic and (2) North American and Eurasian *Lycaeides* constrained to be reciprocally monophyletic. The hypothesized unconstrained genealogy resulted in an increase of 105.87 (North American *Lycaeides* monophyletic) and 107.06 (North American and Eurasian *Lycaeides* reciprocally monophyletic) likelihood units compared to the constrained genealogies (Table 2). Based on the criteria of Kass and Raftery (1995) this magnitude of difference suggests that we should favor the unconstrained genealogy and consider the evidence against the constrained topologies as "decisive" for both the hypothesis of North American monophyly and the hypothesis of reciprocal monophyly.

Enforcing a strict molecular clock resulted in an increase in 37.03 likelihood units relative to a model without a clock enforced (Table 2). When calculating marginal likelihoods it is not uncommon for a constrained model to have a higher likelihood than an unconstrained model, as the unconstrained model may not sample as tightly around the maximum likelihood parameter values and thus result in a lower harmonic mean of the likelihood values than the constrained model (Nylander et al., 2004). Based on this magnitude of difference in the marginal likelihoods of the models we

Table 2
Results from Bayes factor analyses

	Marginal likelihood (M_1) Ln$f(X\|M_1)$	Marginal likelihood (M_0) Ln$f(X\|M_0)$	Bayes factor Ln(B_{10})	Evidence against M_0
Gene region and codon position vs. all data	−2706.89	−3015.12	308.23	"Decisive"
Gene region and codon position vs. gene region	−2706.89	−2956.50	249.61	"Decisive"
NA monophyletic	−2706.89	−2812.76	105.87	"Decisive"
Reciprocal monophyly	−2706.89	−2813.95	107.06	"Decisive"
Clock-like evolution	−2669.86	−2706.89	37.03	"Decisive"

The model with the greater likelihood is denoted M_1 and the model with the lesser likelihood is denoted M_0. Complete descriptions of these models are given in the text.

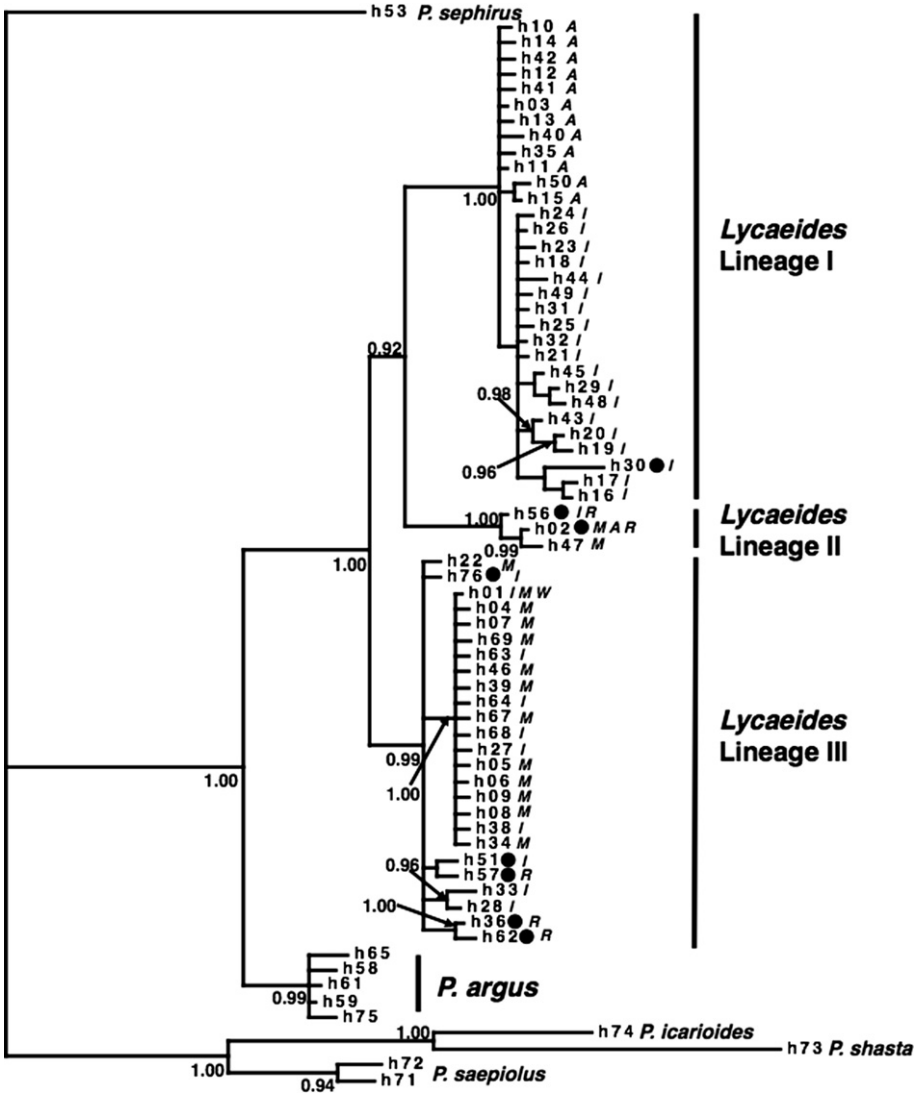

Fig. 2. Bayesian maximum likelihood genealogy based on 915 bp of COI and COII. Branches are labeled with haplotypes (*Lycaeides*) and species name. Species names for haplotypes sampled from *Lycaeides* are given as abbreviations following the haplotype numbers: *L. idas* (*I*), *L. melissa* (*M*), *L. argyrognomon* (*R*), *L.* sp. (alpine) (*A*), and the Warners' Mountain entity (*W*). Haplotypes marked with black dots were sampled in Eurasia or Eurasia and North America (h02 only). Numbers represent Bayesian posterior probability values. See Table 1 for sampling localities.

should favor the model with clock-like evolution and consider the evidence against the unconstrained model "decisive".

3.2. Estimation of TMRCA

Estimates of TMRCA were identical for the North American mtDNA haplotypes and all *Lycaeides* haplotypes. The TMRCA under the constant population size demographic model was 0.03257 substitutions per site (95% highest posterior density (95% HPD): 0.01695–0.05287 substitutions per site). This corresponds to 2,960,909 million years before the present (mybp) (95% HPD: 1,540,909–4,806,364 mybp) based on the general arthropod mitochondrial clock (1.1×10^{-8} substitutions per site per year) (Brower, 1994), or 4,175,641 mybp (95% HPD: 2,173,077–6,778,205 mybp) following the slower *Papilio* (Lepidoptera: Papilionidae) mtDNA molecular clock (7.8×10^{-9} substitutions per site per year proposed by Zakharov et al., 2004). The TMRCA under the Bayesian skyline plot demographic model was 0.02572 substi-

tutions per site (95% HPD: 0.01464–0.03982 substitutions per site), which corresponds to 2,338,182 mybp (95% HPD: 1,330,909–3,620,000 mybp) following the general arthropod mitochondrial clock (Brower, 1994), or 3,297,436 mybp (95% HPD: 1,876,923–5,105,128 mybp) following the *Papilio* (Lepidoptera: Papilionidae) mtDNA molecular clock (Zakharov et al., 2004). These estimates place the TMRCA for *Lycaeides* mtDNA variation, and thus a rough estimate of the upper bounds for the divergence time between North American and Eurasian *Lycaeides*, in the Pliocene or even possibly the early Pleistocene.

3.3. Estimation of population growth

The coalescent-based Bayesian skyline plot shows evidence of θ remaining fairly constant for most of the history of the extant mitochondrial variation in North American *Lycaeides* (Fig. 3). However, this trend is punctuated by a recent and drastic increase in θ of more than an order of magnitude, from approximately 0.01 to about 0.50. Note that our estimate of θ is for all North American *Lycaeides* species combined, and assumes North American *Lycaeides* function as a single interbreeding group with respect to the mitochondrial genome. Based on the general arthropod mtDNA molecular clock of 1.1×10^{-8} substitutions per site per year (Brower, 1994), this period of population growth occurred approximately 100,000 years before the present (ybp) (Fig. 3). The slower *Papilio* (Lepidoptera: Papilionidae) mtDNA molecular clock (7.8×10^{-9}) of Zakharov et al. (2004) places this period of growth at approximately 150,000 ybp. Thus, both of these clocks date this period of population growth to the boundary of the late and middle Pleistocene.

In contrast to the results from the Bayesian skyline plot analysis, the constant population size model was selected based on AIC_C among the five parametric demographic models examined using Genie v3.0 (Table 3). The maximum likelihood estimate of θ under the constant population size model was 0.02916 (95% confidence interval: 0.02243–0.03887). This estimate of θ is intermediate between the pre- and post-growth estimates of θ from the Bayesian skyline plot analysis.

4. Discussion

Our results are consistent with the hypothesis that North American *Lycaeides* are not derived from a single Eurasian ancestor. The mitochondrial haplotypes sampled from North American *Lycaeides* do not form a distinct monophyletic lineage (Fig. 2). Conversely, both Eurasian and North American mitochondrial haplotypes occur in all three major *Lycaeides* lineages. This pattern would be expected if divergence among North American *Lycaeides* began in Eurasia, with different Eurasian taxa giving rise to different North American taxa. For example, it is possible that Eurasian *L. idas* is the ancestor of North American *L. idas*, while Eurasian *L. argyrognomon* is the ancestor of North American *L. melissa*. However, this simple interpretation is complicated by the fact that the recognized taxonomic entities in both North America and Eurasia do not correspond to any of the three major mitochondrial lineages recovered by phylogenetic analysis (Fig. 2). Instead, all described *Lycaeides* species possess mtDNA haplotypes from at least two of these three lineages. Indeed, several populations (e.g. Gardnerville, NV (*L. m. melissa*) and County Line Hill, CA (alpine hybrid species)) possess mitochondrial haplotypes from more than one mitochondrial lineage.

However, there are several possible explanations for this pattern. Different North American *Lycaeides* taxa may have distinct Eurasian ancestors, however, intra-continental relationships based on mtDNA may have become obscured as a result of introgressive hybridization in both Eurasia and North America following

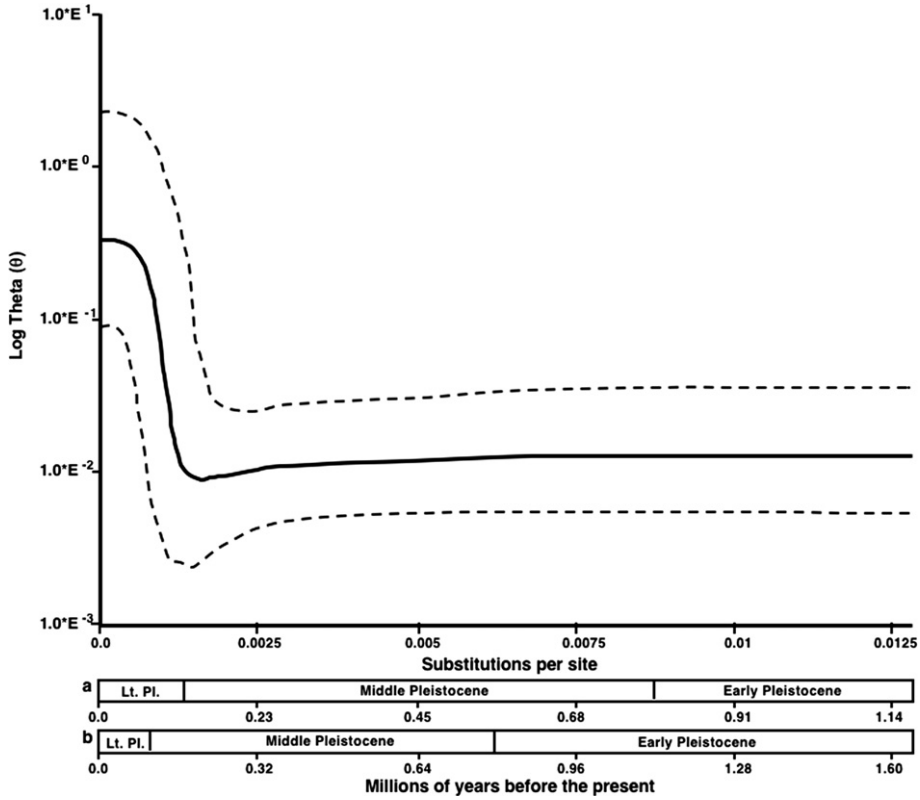

Fig. 3. Bayesian skyline plot for North American *Lycaeides* based on mtDNA sequence data. The solid line gives the mean estimate of the log of θ through time; the dotted lines delimit the 95% HPD of the log of θ through time. The time scales shown are based on the general arthropod mitochondrial clock (1.1×10^{-8} substitutions per site per year) (Brower, 1994)[a], and the *Papilio* mtDNA molecular clock (7.8×10^{-9} substitutions per site per year) (Zakharov et al., 2004)[b].

Table 3
AIC_C model comparison

Model	Equation	Ln likelihood	*K*	AIC_C	Delta
Constant	$\theta(t) = \theta(0)$	−402.702	1	807.484	0.000
Piecewise	$\theta(t) = \{\theta(0)\}$ if $t < z$ $\{\theta(0)f\}$ otherwise	−402.702	3	811.904	4.420
Exponential	$\theta(t) = \theta(0)\exp(-\rho t)$	−406.072	2	816.389	8.905
Logistic	$\theta(t) = \theta(0)\,((1+c)/(1+c*\exp(-\rho t)))$	−406.024	3	818.548	11.064
Expansion	$\theta(t) = \theta(0)\,(\alpha + ((1-\alpha)\exp(-\rho t)))$	−442.404	3	891.308	83.824

ρ is the intrinsic rate of growth divided by the mutation rate, c is the logistic shape parameter, α is the population size at time $t = \infty$, z is the time of population size change divided by the mutation rate, t is time.

colonization of North America. This scenario is quite likely, as substantial evidence exists that introgressive hybridization has occurred in North American *Lycaeides* at contact zones (Gompert et al., 2006a,b).

Alternatively, it is possible that the Eurasian ancestor(s) of North American *Lycaeides* may have been polymorphic for haplotypes from all three mtDNA lineages when North America was colonized (see Nice et al., 2005). If this is true the ancestral polymorphism persists to the present in both Eurasia and North America. These two scenarios are not mutually exclusive, and it is quite possible that some combination of ancestral polymorphism and introgressive hybridization is responsible for the pattern of mtDNA variation observed in *Lycaeides* (e.g. Funk and Omland, 2003). While sister species status for North American *L. idas* and *L. melissa* cannot be completely ruled out based on these results, our mitochondrial data provide no support for this relationship.

As most members of the *Plebejus*/*Lycaeides* clade are endemic to Eurasia, a Eurasian origin for North American *Lycaeides* has generally been assumed (e.g. Nice et al., 2005). This assertion is supported by the fact that the Eurasian species *P. argus* is placed sister among the sampled taxa to both North American and Eurasian *Lycaeides* (Fig. 2), which suggests the Eurasian species *P. argus* and *Lycaeides* share a recent common ancestor.

Colonization of North America by *Lycaeides* occurred relatively recently. Haplotype h02 is shared among North American *L. m. melissa* and Eurasian *L. a. argyrognomon* (Table 1). This inter-continental haplotype sharing is not the result of current gene flow, but instead must be due to retained ancestral polymorphism, as it is unlikely that butterfly populations separated by a large expanse of ocean are exchanging individuals at present (average dispersal distance for *Lycaeides* has been estimated at 500 m (King, 1998)). As a mitochondrial haplotype is shared between North American and Eurasian *Lycaeides*, the MRCA of the extant *Lycaeides* mitochondrial haplotypes must have existed prior to the point in time when North American *Lycaeides* became isolated from Eurasian *Lycaeides*. Thus, our TMRCA estimates, which range from 2,338,182 to 4,175,641 ybp, indicate that *Lycaeides* colonized North America and became isolated from *Lycaeides* in Eurasia no longer ago than the late Pliocene. Note the TMRCA dates provide a rough estimate of the upper bounds for the colonization and isolation of *Lycaeides* in North America; this event could have occurred more recently. These dates are in general concordance with colonization having occurred from north-eastern Eurasia via a Beringian land bridge (Pielou, 1992). However, as these dates are based on molecular clocks that may not be applicable to North American *Lycaeides*, these dates should be interpreted with caution.

We have conflicting results concerning the demographic history of North American *Lycaeides*. The Bayesian skyline plot indicated a rapid increase in population size approximately 100,000–150,000 ybp (Fig. 3). In contrast, the AIC_C model selection procedure indicated that a demographic model of constant population size best explained our data (Table 3). This discrepancy may result from the fact that the population increase identified in the Bayesian skyline plot analysis occurred very recently, which means that the population size was constant for most of the history of the extant mitochondrial variation in *Lycaeides* (Fig. 3). Thus, the mitochondrial data can be explained fairly well by a model of constant population size. However, the population size increase indicated by the Bayesian skyline plot analysis likely provides a better representation of the demographic history of North American *Lycaeides* than the parametric model of constant population size. The Bayesian skyline plot model, unlike the parametric models tested, is not constrained to construct a simple demographic history, but instead allows more realistic fluctuations in population size through time. Moreover, the biological importance of the population size increase inferred from the Bayesian skyline plot is not diminished by the relatively long duration that the population size remained fairly constant (Fig. 3). Thus, we interpret our demographic results on the basis of the Bayesian skyline plot model.

Our estimate of θ, and thus population size, is for North American *Lycaeides* as a whole. Consequently our Bayesian skyline plot results do not imply an increase in population size for any specific species or population, but instead for North American *Lycaeides* in total. There are several possible explanations for the recent and drastic population size increase demonstrated in North American *Lycaeides*. The population size increase could be the result of an increase in local density at the population level. However, this explanation, at least by itself, seems unlikely, as there is no evidence that such an increase in local density has occurred, nor is there a clear biological reason why such an increase would be expected. Colonization of North America and subsequent range expansion across the continent could, and probably does, account for at least some of this increase in population size. Range expansion, both due to the initial colonization of North America and as Pleistocene glaciers receded, would increase the number of North American *Lycaeides* populations and thus the total effective population size for North American *Lycaeides*. Nice et al. (2005) found evidence of post-Pleistocene range expansion in North American *Lycaeides* based on sequence variation in the AT-rich region of the mitochondrial genome. We have less evidence that the colonization of North America by Eurasian *Lycaeides* is the cause of the inferred population size increase. Mitochondrial genetic diversity is lower in Canada and Alaska than in much of the continental USA, which suggests that this population size increase did not occur at the time of colonization (Table 1). However, this pattern might be an artifact of our limited sampling in northern North America compared to sampling in the contiguous United States. Additional sampling of *Lycaeides* from northern North America and Eurasia would be necessary to determine whether the inferred population size increase occurred at the time North America was colonized.

Evolutionary diversification in terms of adaptation to novel habitats and/or host-plants may have played a role in this population size increase as well (e.g. Orr and Smith, 1998). Diversification via an increase in specialization among populations would have increased the number of distinct populations that could be supported in a given geographic area, which in turn would have increased the total population size of North American *Lycaeides* (e.g. Schluter, 1996, 2000). This scenario is consistent with the

remarkable degree of specialization in terms of host-plant use and preference (Nice et al., 2002; Gompert et al., 2006b), as well as other ecologically relevant characters (e.g. egg adhesion (Fordyce and Nice, 2003)), observed in North American *Lycaeides* today. For example, the alpine-associated hybrid *Lycaeides* species has come to occupy a habitat not used by other *Lycaeides* in the region, and its presence in the alpine habitat is facilitated by its strong host plant preference coupled with its unique lack of egg adhesion (Fordyce and Nice, 2003; Gompert et al., 2006b). Thus, the recent population size increase detected in North American *Lycaeides* is most likely due to a combination of range expansion and evolutionary diversification.

The extent that hybridization has influenced the evolutionary diversification and by extension increase in population size of North American *Lycaeides* is unclear. There is substantial evidence that introgressive hybridization has occurred among differentiated populations of North American *Lycaeides* (Gompert et al., 2006a, 2006b). Seehausen (2004) suggested that adaptive radiations might be facilitated by hybridization, as hybridization can provide genetic variation and create novel combinations of alleles, which can facilitate further diversification and specialization. Hybridization has led to an increase in diversification in *Lycaeides* via the production of the unnamed alpine-associated hybrid species (Gompert et al., 2006b). Hybridization could have contributed to diversification in *Lycaeides* in other ways, for example, by providing the genetic variation necessary for host plant switches (e.g. Lewontin and Birch, 1966; Rieseberg et al., 1999; Schwarz et al., 2007). However, at present, we have no evidence to support this hypothesis. Nevertheless, such a role for hybridization in the diversification of North American *Lycaeides* remains plausible.

The results obtained in this study were based solely on mtDNA variation. Sole reliance on mtDNA for molecular systematics, defining population structure, estimating population parameters, and conservation genetics is known to be problematic due to issues associated with introgressive hybridization and ancestral polymorphism (Funk and Omland, 2003; Gompert et al., 2006a; Forister et al., 2008). However, the results from this study, and particularly the way in which these results were interpreted, should be minimally affected by these problems. We have made no attempt to define taxonomic boundaries based on our mtDNA data, and our interpretation of these results does not necessitate that our mitochondrial genealogy is representative of the species phylogeny for *Lycaeides*.

Our use of a molecular clock is partially justified, as we found evidence that our Bayesian tree conformed to a model of clock-like evolution (Table 2). Even so, dates from molecular clocks should be considered cautiously, due to the stochastic nature of the mutation process. This is particularly true for our study, as the molecular clocks used in this study were not calibrated for *Lycaeides*. Therefore, it is unclear whether these clocks are applicable to the evolution of mtDNA variation in North American *Lycaeides*. We have attempted to minimize this problem by employing two different plausible molecular clocks, even so, the dates presented in this study should be treated as rough estimates. Despite this uncertainty in the dating, the results of this analysis are consistent with the hypothesis of a recent population size increase in North American *Lycaeides*.

5. Conclusions

Lycaeides in North America are remarkably diverse and locally specialized in terms of ecological, morphological, and behavioral characters. Given the relatively short period of time that *Lycaeides* has occupied North America and its recent population size expansion in North America, it is likely that much of this diversity arose rapidly on the North American continent. It is clear that hybridization has contributed to this diversification via the formation of a hybrid species (Gompert et al., 2006b), however, the role of hybridization in the diversification of North American *Lycaeides* remains to be fully examined. Regardless of the part played by hybridization, this scenario suggests that North American *Lycaeides* represents a recent evolutionary radiation.

Acknowledgments

We wish to thank the following people who assisted with specimen collection: A.M. Shapiro, S. Fuller, M. Amaral, T. McCabe, J. Oliver, K. Prudic, E. Runquist, P. Opler, G. Austin, K. Anderson, D. Underwood, D. Ganser, G. Gelembiuk, N. Anthony, G. Marrone, M. Zubrik, S. Davkov. The Museum of Entomology at the University of Wisconsin and the Bohart Museum of Entomology at the University of California, Davis also kindly provided specimens. Thanks to Lauren Lucas and an anonymous reviewer for providing comments on an earlier version of this manuscript. This research was funded by a NSF graduate research fellowship to Z.G., the University of Tennessee, and a Research Enhancement Grant from Texas State University to C.C.N.

References

Arnold, R.A., 1993. The Lotis Blue, *Lycaeides idas lotis* (Lintner). In: New, T.R., (Ed.), Conservation Biology of the Lycaenidae (Butterflies). Occasional Paper of the IUCN Species Survival Commission No. 8, pp. 143–144.

Avise, J.C., Walker, D., 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proceedings R. Soc. B 265, 1707–1712.

Brandley, M.C., Schmitz, A., Reeder, T.W., 2005. Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scincid lizards. Systematic Biology 54, 373–390.

Brock, J.P., Kaufman, K., 2003. Butterflies of North America. Houghton Mifflin, Boston, MA.

Brookes, M.I., Graneau, Y.A., King, P., Rose, O.C., Thomas, C.D., Mallet, J.L.B., 1997. Genetic analysis of founder bottlenecks in the rare British butterfly *Plebejus argus*. Conservation Biology 11, 648–661.

Brower, A.V.Z., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius-Erato* inferred from patterns of mitochondrial-DNA evolution. Proceedings of the National Academy of Sciences of the United States of America 91, 6491–6495.

Caterino, M.S., Sperling, F.A.H., 1999. *Papilio* phylogeny based on mitochondrial cytochrome oxidase I and II genes. Molecular Phylogenetics and Evolution 11, 122–137.

Crandall, E.D., Frey, M.A., Grosberg, R.K., Barber, P.H., 2008. Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. Molecular Ecology 17, 611–626.

Downey, J.C., Dunn, D.B., 1964. Variation in Lycaenid butterfly *Plebejus icarioides* III. Additional data on food-plant specificity. Ecology 45, 172–178.

Drummond, A.J., Rambaut, A., 2006. BEAST v1.4. Available from: <http://beast.bio.ed.ac.uk/>.

Drummond, A.J., Rambaut, A., Shapiro, B., Pybus, O.G., 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22, 1185–1192.

Fordyce, J.A., Nice, C.C., Forister, M.L., Shapiro, A.M., 2002. The significance of wing pattern diversity in the Lycaenidae: mate recognition by two recently diverged species. Journal of Evolutionary Biology 15, 871–879.

Fordyce, J.A., Nice, C.C., 2003. Variation in butterfly egg adhesion: adaptation to local host plant senescence characteristics? Ecology Letters 6, 23–27.

Forister, M.L., Fordyce, J.A., Nice, C.C., Gompert, Z., Shapiro, A.M., 2006. Egg morphology varies among populations and habitats along a suture zone in the *Lycaeides idas-melissa* species complex (Lepidoptera: Lycaenidae). Annals of the Entomological Society of America 99, 933–937.

Forister, M.L., Nice, C.C., Fordyce, J.A., Gompert, Z., Shapiro, A.M., 2008. Considering evolutionary processes in the use of single-locus genetic data for conservation with examples from the Lepidoptera. Journal of Insect Conservation 12, 37–51.

Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology Evolution and Systematics 34, 397–423.

Gompert, Z., Nice, C.C., Fordyce, J.A., Forister, M.L., Shapiro, A.M., 2006a. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. Molecular Ecology 15, 1759–1768.

Gompert, Z., Fordyce, J.A., Forister, M.L., Shapiro, A.M., Nice, C.C., 2006b. Homoploid hybrid speciation in an extreme habitat. Science 314, 1923–1925.

Goodman, G.N., 1999. Toward evidence-based medical statistics. 2: the Bayes factor. Annals of Internal Medicine 130, 1005–1013.

Grimaldi, D., Engel, M.S., 2005. Evolution of the Insects. Cambridge University Press, Cambridge, UK.
Griswold, C.K., Baker, A.J., 2002. Time to the most recent common ancestor and divergence times of populations of common chaffinches (*Fringilla coelebs*) in Europe and North Africa: insights into Pleistocene refugia and current levels of migration. Evolution 56, 143–153.
Hedin, M., Bond, J.E., 2006. Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear genes (18S and 28S): conflict and agreement with the current system of classification. Molecular Phylogenetics and Evolution 41, 454–471.
Hein, J., Schierup, M.H., Wiuf, C., 2005. Gene Genealogies, Variation and Evolution: A Primer in Coalescent Theory. Oxford University Press, USA.
Hillis, D.M., Mable, B.K., Larson, A., Davis, S.K., Zimmer, E.A., 1996. Nucleic acids IV: sequencing and cloning. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer, Sunderland, MA, pp. 321–381.
Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754–755.
Kandul, N.P., Coleman, J.W.S., Lukhtanov, V.A., Dantchenko, D.A., Sekercioglu, C., Haig, D., Pierce, N.E., 2004. Phylogeny of *Agrodiaetus* Hubner 1822 (Lepidoptera: Lycaenidae) inferred from mtDNA sequences of COI and COII, and nuclear sequences of EF1-α: karyotype diversification and species radiation. Systematic Biology 53, 1–21.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of American Statistical Association 90, 773–795.
King, R.S., 1998. Dispersal of karner blue butterflies (*Lycaeides melissa samuelis* Nabokov) at Necedah National Wildlife Refuge. Transactions of the Wisconsin Academy of Sciences, Arts and Letters 86, 101–110.
Knowles, L.L., 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. Molecular Ecology 10, 691–701.
Lane, C.P., Weller, S.J., 1994. A review of *Lycaeides* Hubner and Karner blue butterfly taxonomy. Minnesota Agricultural Experiment Station Miscellaneous Publication Series 84, 5–21.
Lewontin, R.C., Birch, L.C., 1966. Hybridization as a source of variation for adaptation to new environments. Evolution 20, 315–336.
Lucas, L.K., Fordyce, J.A., Nice, C.C., 2008. Patterns of genitalic morphology around suture zones in North American *Lycaeides* (Lepidoptera: Lycaenidae): implications for taxonomy and historical biogeography. Annals of the Entomological Society of America 101, 172–180.
Nabokov, V., 1943. The Nearctic forms of *Lycaeides* Hub. (Lycaenidae, Lepidoptera). Psyche 50, 87–99.
Nabokov, V., 1949. The nearctic members of the genus *Lycaeides* Hubner (Lycaenidae, Lepidoptera). Bulletin of the Museum of Comparative Zoology 101, 479–541.
Nice, C.C., Anthony, N., Gelembiuk, G., Raterman, D., ffrench-Constant, R., 2005. The history and geography of diversification within the butterfly genus *Lycaeides* in North America. Molecular Ecology 14, 1741–1754.
Nice, C.C., Fordyce, J.A., Shapiro, A.M., ffrench-Constant, R., 2002. Lack of evidence for reproductive isolation among ecologically specialized Lycaeinid butterflies. Ecological Entomology 27, 702–712.
Nice, C.C., Shapiro, A.M., 1999. Molecular and morphological divergence in the butterfly genus *Lycaeides* (Lepidoptera: Lycaenidae) in North America: evidence of recent speciation. Journal of Evolutionary Biology 12, 936–950.
Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53, 47–67.
Opler, P.A., 1992. A Field Guide to Eastern Butterflies. Houghton Mifflin Co., Boston, New York, London.
Orr, M.R., Smith, T.B., 1998. Ecology and speciation. Trends in Ecology and Evolution 13, 502–506.
Pielou, E.C., 1992. After the Ice Age: The Return of Life to Glaciated North America. University of Chicago Press, Chicago, IL.
Pinho, C., Harris, D.J., Ferrand, N., 2007. Contrasting patterns of population subdivision and historical demography in three western Mediterranean lizard species inferred from mitochondrial DNA variation. Molecular Ecology 16, 1191–1205.
Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Pybus, O.G., Rambaut, A., 2002. GENIE: estimating demographic history from molecular phylogenies. Bioinformatics 18, 1404–1405.
Rambaut, A., Drummond, A.J., 2004. Tracer v1.3. Available from: <http://beast.bio.ed.ac.uk/Tracer>.
Rieseberg, L.H., Archer, M.A., Wayne, R.K., 1999. Transgressive segregation, adaptation and speciation. Heredity 83, 363–372.
Ronquist, F., Huelsenbeck, J.P., van der Mark, P., 2005. MrBayes 3.1 Manual, Draft 5/26/2005, online at: <http://mrbayes.csit.fsu.edu/manual.php/>.
Schluter, D., 1996. Ecological speciation in postglacial fishes. Philosophical Transactions of the Royal Society of London, Series B 351, 807–814.
Schluter, D., 2000. Ecological character displacement in adaptive radiation. American Naturalist 156 (Suppl.), S4–S16.
Schwarz, D., Shoemaker, K.D., Botteri, N.L., McPheron, B.A., 2007. A novel preference for an invasive plant as a mechanism for animal hybrid speciation. Evolution 61, 245–256.
Scott, J.A., 1986. The Butterflies of North America, a Natural History and Field Guide. Stanford University Press, Stanford, CA.
Seehausen, O., 2004. Hybridization and adaptive radiation. Trends in Ecology & Evolution 19, 198–207.
Shapiro, B., Rambaut, A., Drummond, A.J., 2006. Choosing appropriate substitution models for the phylogenetic analysis of protein-coding sequences. Molecular Biology and Evolution 22, 7–9.
Simon, C., Frati, F., Beckenbach, A.T., 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. Annals of the Entomological Society of America 87, 651–701.
Spellman, G.M., Riddle, B.R., Klicka, J., 2007. Phylogeography of the mountain chickadee (*Poecile gambeli*): diversification, introgression, and expansion in response to Quaternary climate change. Molecular Ecology 16, 1055–1068.
Stanford, R.E., Opler, P.A., 1996. Atlas of Western USA Butterflies, Including Adjacent Parts of Canada and Mexico. VCH, Colorado.
Strimmer, K., Pybus, O.G., 2001. Exploring the demographic history of DNA sequences using the generalized skyline plot. Molecular Biology and Evolution 18, 2298–2305.
Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (and other methods). Ver. 4.0b10. Sinauer, Sunderland, MA.
US Fish and Wildlife Service, 1992. Endangered and threatened wildlife and plants: determination of endangered status for the Karner blue butterfly. Proposed Rule. Federal Register 57, 2241–2246.
US Fish and Wildlife Service, 2003. Final recovery plan for the Karner blue butterfly (*Lycaeides melissa samuelis*). US Fish and Wildlife Service, Fort Snelling, MN.
Zakharov, E.V., Caterino, M.S., Sperling, F.A.H., 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). Systematic Biology 53, 193–215.