Sciencexpress Report

Homoploid Hybrid Speciation in an Extreme Habitat

Zachariah Gompert,[1]* James A. Fordyce,[2] Matthew L. Forister,[3] Arthur M. Shapiro,[4] Chris C. Nice[1]

[1]Department of Biology, Population and Conservation Biology Program, Texas State University, San Marcos, TX 78666, USA. [2]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA. [3]Department of Natural Resources and Environmental Science, University of Nevada, Reno, NV 89512, USA. [4]Section of Evolution and Ecology, University of California, Davis, CA 95616, USA.

*To whom correspondence should be addressed. E-mail: zg1002@txstate.edu

Theory predicts that homoploid hybrid speciation, which is hybrid speciation without a change in chromosome number, is facilitated by adaptation to a novel or extreme habitat. Using molecular and ecological data, we show that the alpine-adapted butterflies in the genus *Lycaeides* are the product of hybrid speciation. The alpine populations possess a mosaic genome derived from both *L. melissa* and *L. idas* and are differentiated from, and younger than, their putative parental species. As predicted, adaptive traits may allow for persistence in the environmentally extreme alpine habitat and reproductively isolate these populations from their parental species.

Homoploid hybrid speciation is characterized by hybridization between parental species that results in a derivative hybrid species without a change in chromosome number (*1*–*6*). A growing list of possible examples, e.g., African cichlids (*7*), cyprinid fishes (*8*), *Rhagoletis* fruit flies (*9*), *Heliconius* butterflies (*10*), and swallowtail butterflies (*11*), suggest that homoploid hybrid speciation in animals may be more common than previously thought. Models predict that ecological isolation spurs homoploid hybrid speciation, especially when the hybrids invade novel or extreme habitats (*12*). Colonization of a novel habitat by an incipient hybrid species may allow it to avoid introgression and competition with the parental species (*4*, *12*). Although these predictions have been borne out in plants (*13*), no examples of homoploid hybrid speciation in animals have involved adaptation to a novel habitat, although a switch to a novel host plant species has been documented (*9*).

The ecologically, morphologically, and behaviorally distinct species *L. melissa* and *L. idas* (*14*–*17*) have come into secondary contact in the Sierra Nevada of western North America (*18*) (Fig. 1). *Lycaeides melissa* populations occur in Great Basin habitats on the east side of the Sierra Nevada, while *L. idas* populations occupy wet meadows at mid-elevation on the west slope of these mountains. Unnamed populations of *Lycaeides* occur in alpine habitat above the tree-line of the Sierra Nevada; an environmentally extreme habitat not occupied by *L. melissa* nor *L. idas*. The alpine habitat is characterized by a short growing season and severe fluctuations in ambient temperature and relative humidity on a daily and seasonal basis (*19*). These alpine butterflies have male genitalia that are intermediate in size and shape compared to *L. melissa* and *L. idas* (*18*), with wing pattern elements that are qualitatively similar to those of *L. melissa* (*14*). However, analyses of mitochondrial DNA (mtDNA) variation shows that the alpine populations' haplotypes share a more recent common ancestor with haplotypes of *L. idas* than those of *L. melissa* (fig. S1) (*20*). These discordant patterns suggest that hybridization may have played a role in the evolutionary history of alpine *Lycaeides* populations.

If the alpine *Lycaeides* populations are a hybrid species they should possess a genome that is a blend of alleles derived from both *L. melissa* and *L. idas*. We tested this using a large multilocus genomic dataset, consisting of 128 Amplified Fragment Length Polymorphism (AFLP) markers, three microsatellite markers (*Msat201*, *Msat4*, and *MsatZ12-1*), and sequence data from three nuclear genes (*Nuc1*, *Nuc3*, and *Ef1*α) and two mitochondrial genes (*COI* and *COII*) (*20*). To assess the overall genomic composition of the alpine *Lycaeides* populations we used the Bayesian program STRUCTURE version 2.1 to cluster *L. melissa*, *L. idas*, and alpine individuals on the basis of their multilocus genotypes (*20*, *21*) under the assumption that the data represented two separate populations (K = 2). Individuals from *L. melissa* and *L. idas* clustered to different groups with high probability, while alpine *Lycaeides* individuals were assigned to both groups with moderate probability (Fig. 2A and table S1). This pattern is inconsistent with a bifurcating mode of speciation, where alpine butterflies originated from a single parental species, and suggests that the alpine genome is a mosaic of the two species. In further support of this hypothesis, five AFLP fragments were shared between *L. melissa* and *L. idas* to the exclusion of the alpine populations, while the alpine populations shared 12 unique alleles with *L. melissa* and 16

unique alleles with *L. idas*. Additionally, in *L. melissa* and *L. idas* different alleles were fixed at the *Nuc1* locus (fig. S2 and table S1), while the alpine populations shared three *Nuc1* alleles with *L. melissa* and three *Nuc1* alleles with *L. idas*.

Although, the mosaic genome of the alpine populations is consistent with homoploid hybrid speciation, a similar pattern could have arisen if the alpine populations have continuous gene flow with *L. melissa* and/or *L. idas*. If so, the alpine populations should not be genetically differentiated from *L. melissa* and *L. idas* and there should be evidence of gene flow with these species. When STRUCTURE (*21*), was run under the assumption that the data represented three separate populations (K=3), *L. melissa* and *L. idas* individuals were still assigned to their respective clusters, but the alpine *Lycaeides* individuals were assigned to a distinct, third cluster (Fig. 2B and table S1) (*20*). Additionally, alpine populations were fixed for unique alleles at the mitochondrial genes *COI* and *COII*, as well as the nuclear gene *Nuc3* (figs. S1 and S3 and table S1). These data and examination of pairwise F_{ST} (*20*) (table S2), suggest that alpine populations are differentiated from *L. melissa* and *L. idas*. We did not detect excess heterozygosity or deviations from linkage equilibrium for any microsatellite markers or nuclear gene sequences (*20*); such deviations would indicate ongoing gene flow between the alpine populations and either *L. melissa* or *L. idas*. We used the program NewHybrids, which employs the Bayesian assignment algorithm of Anderson and Thompson (*22*), to assess the probability that gene flow occurs between *L. melissa* and/or *L. idas* and the alpine populations (*20*). No individuals were identified that might be considered F_1's produced from crosses between *L. melissa* or *L. idas* and the alpine populations (fig. S4). Thus, we conclude that the alpine populations are genetically differentiated from *L. melissa* and *L. idas* and are not exchanging genes with either.

The genetic patterns documented above could have occurred if *L. melissa*, *L. idas*, and the alpine populations all arose rapidly from a single ancestral species distributed along a geographic cline with the alpine populations originating from the center of the cline. This scenario is unlikely for several reasons. Phylogeographic data suggest that the current distribution of *L. melissa* and *L. idas* is the result of post-Pleistocene range expansion and secondary contact, and thus does not reflect the distribution of the ancestor of these species (*18*). The alpine populations also have a more recent origin than either *L. melissa* and *L. idas*. We calculated a coalescent-based estimate of the time to the most recent common ancestor (TMRCA) for mitochondrial variation for each of the three putative species: the alpine populations, *L. melissa*, and *L idas* (*20*). The estimated TMRCA for the alpine populations, 442,579 years before present (ybp), is substantially younger than that of either *L. melissa* or *L. idas,* 1,902,995 ybp and 1,267,885 ybp, respectively (*20*). Furthermore, pairwise estimates of τ (species divergence time x mutation rate) based on nuclear and mitochondrial sequence data were approximately four times greater for the divergence of *L. melissa* and *L. idas* (0.006576, 95% CI 0.002823-0.009855) than for the divergence of the alpine populations and either *L. melissa* (0.001318, 95% CI 0.000638-0.002233) or *L. idas* (0.001468, 95% CI 0.000763-002454) (*20*) (fig. S5). Thus, *L. melissa*, *L. idas*, and the alpine populations did not arise rapidly from a single ancestral species.

Homoploid hybrid speciation is more likely when a hybrid species colonizes a novel habitat (*5*, *6*, *12*, *13*) such as the alpine habitat occupied by *Lycaeides*. Reproductive isolation between the hybrid and the parental taxa may be maintained by behavioral and ecological adaptations to the alpine habitat specifically associated with the alpine host plant. Females from the alpine populations have near perfect host fidelity for their host plant, the perennial alpine endemic, *Astragalus whitneyi* (Fig. 3) (*20*). In fact, alpine females have stronger host fidelity than has been recorded in other *Lycaeides* populations. Because males and females of *Lycaeides* locate mates and copulate on or near their larval host plants (*15*), strong fidelity with *A. whitneyi* may serve as a strong pre-zygotic barrier to gene flow between alpine *Lycaeides* populations and their putative parental species. In addition, host fidelity is coupled with a unique lack of egg adhesion in the alpine populations. While *Lycaeides* females from non-alpine populations "glue" their eggs to their host plant when they oviposit, the alpine populations' eggs fall off the plant after oviposition and remain near the site of new plant growth in the following spring (*16)*. Because *Lycaeides* overwinter as diapausing eggs (*16*) and the senesced, above ground biomass of the alpine host plant is blown away by strong winds in the winter (*16*), any eggs attached to the alpine host plant would be carried far from the site of new host plant growth, resulting in likely death of neonate larvae. Any females from non-alpine populations that oviposit on the alpine host plant would suffer a major reduction in fitness. Together, these alpine-associated adaptive traits, strong host fidelity for an alpine endemic host plant and the loss of egg adhesion, may act as an effective ecological barrier to gene flow.

Two other mechanisms may also contribute to reproductive isolation. Color pattern differences on the underside of the wings operate as species recognition cues, isolating the alpine populations from *L. idas* (*14*). Differences in male genital morphology have also been documented (*18*) and may operate in a similar manner to limit gene flow, although this was not explicitly tested in this study. Thus, morphological characters and adaptation to an extreme, novel habitat may create reproductive isolation between the hybrid species and its parental species.

References and Notes

1. J. A. Coyne, H. A. Orr, *Speciation* (Sinauer Associates, 2004).
2. V. Grant, *Plant Speciation* (Columbia University Press, New York, 1981).
3. M. L. Arnold, *Natural Hybridization and Evolution* (Oxford University Press, Oxford, 1997).
4. T. E. Dowling, C. L. Secor, *Annual Review of Ecology and Systematics* **28**, 593-619 (1997).
5. L. H. Rieseberg, *Annual Review of Ecology and Systematics* **28**, 359-389 (1997).
6. B. L. Gross, L. H. Rieseberg, *Journal of Heredity* **96**, 241-252 (May, 2005).
7. R. Schelly, W. Salzburger, S. Koblmuller, N. Duftner, C. Sturmbauer, *Molecular Phylogenetics and Evolution* **38**, 426-438 (Feb, 2006).
8. B. D. Demarais, W. L. Minckley, *Copeia*, 697-703 (Aug 18, 1992).
9 D. Schwarz, B. M. Matta, N. L. Shakir-Botteri, B. A. McPheron, *Nature* **436**, 546-549 (Jul 28, 2005).
10. J. Mavarez *et al.*, *Nature* **441**, 868-871 (Jun 15, 2006).
11. J. M. Scriber, G. J. Ording, *Entomologia Experimentalis et Applicata* **115**, 247-263 (Apr, 2005).
12. C. A. Buerkle, R. J. Morris, M. A. Asmussen, L. H. Rieseberg, *Heredity* **84**, 441-451 (Apr, 2000).
13. L. H. Rieseberg *et al.*, *Science* **301**, 1211-1216 (Aug 29, 2003).
14. J. A. Fordyce, C. C. Nice, M. L. Forister, A. M. Shapiro, *Journal of Evolutionary Biology* **15**, 871-879 (Sep, 2002).
15. C. C. Nice, J. A. Fordyce, A. M. Shapiro, R. ffrench-Constant, *Ecological Entomology* **27**, 702-712 (2002).
16. J. A. Fordyce, C. C. Nice, *Ecology Letters* **6**, 23-27 (2003).
17. M. L. Forister, J. A. Fordyce, C. C. Nice, Z. Gompert, A. M. Shapiro, *Annals of the Entomological Society of America* **99**, 933-937 (2006).
18. C. C. Nice, N. Anthony, G. Gelembiuk, D. Raterman, R. ffrench-Constant, *Molecular Ecology* **14**, 1741-1754 (May, 2005).
19. M. V. Lomolino, B. R. Riddle, J. H. Brown, *Biogeography* (Sinauer Associates, Sunderland, Massachusetts, ed. Third, 2006).
20. Materials and methods are available as supporting material on *Science* Online.
21. J. K. Pritchard, M. Stephens, P. Donnelly, *Genetics* **155**, 945-959 (Jun, 2000).
22. E. C. Anderson, E. A. Thompson, *Genetics* **160**, 1217-1229 (Mar, 2002).
23. V. Nabokov, *Bulletin of the Museum of Comparative Zoology* **101**, 479-541 (1949).
24. R. E. Stanford, P. A. Opler, *Atlas of western USA butterflies, including adjacent parts of Canada and Mexico.* (Denver and Fort Collins, Colorado, 1996).
25. J. A. Scott, *The butterflies of North America, a natural history and field guide* (Stanford University Press, Stanford, California, 1986).
26. This research was funded by a NSF graduate research fellowship to Z. G., the University of Tennessee, a Research Enhancement Grant from Texas State University to C.C.N., and National Science Foundation Grant DEB-9306721 to A.M.S. Thanks to D. Bolnick, B. Fitzparick, S. Gavrilets, C. Jiggins, J. Ott, the EEB discussion group at Texas State University, and two anonymous reviewers for comments on an earlier version of this manuscript, and Allison Stephenson for assistance in the lab. Genbank accession numbers for sequence data are EF 090312-EF090397.

Supporting Online Material

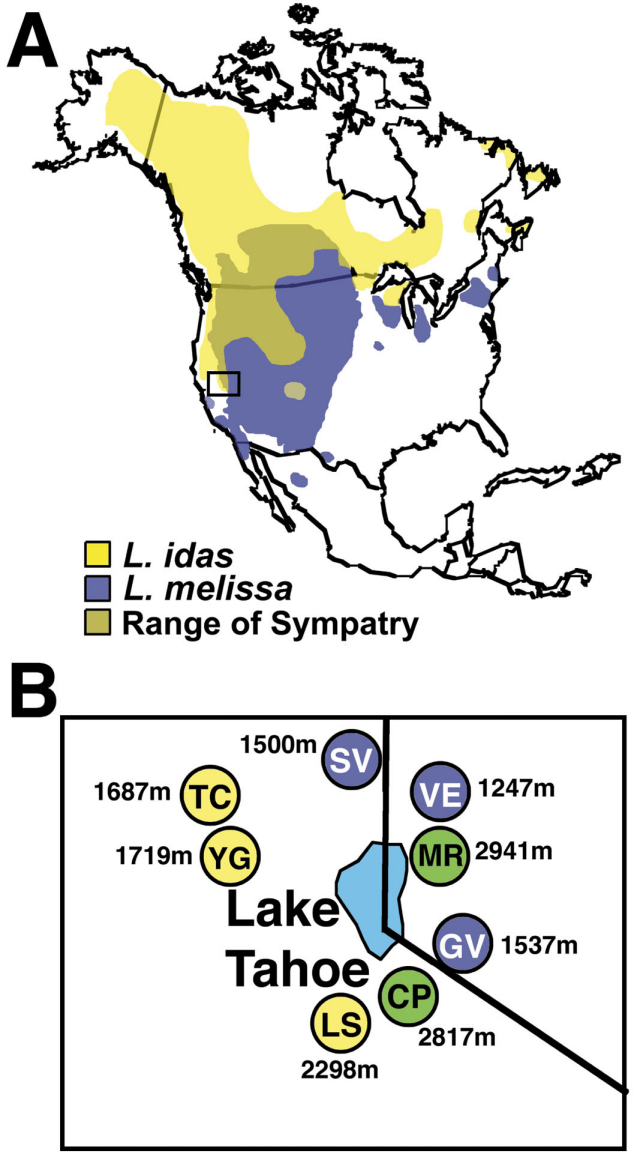
www.sciencemag.org/cgi/content/full/1135875/DC1
Materials and Methods
SOM Text
Figs. S1 to S10
Tables S1 to S4
References

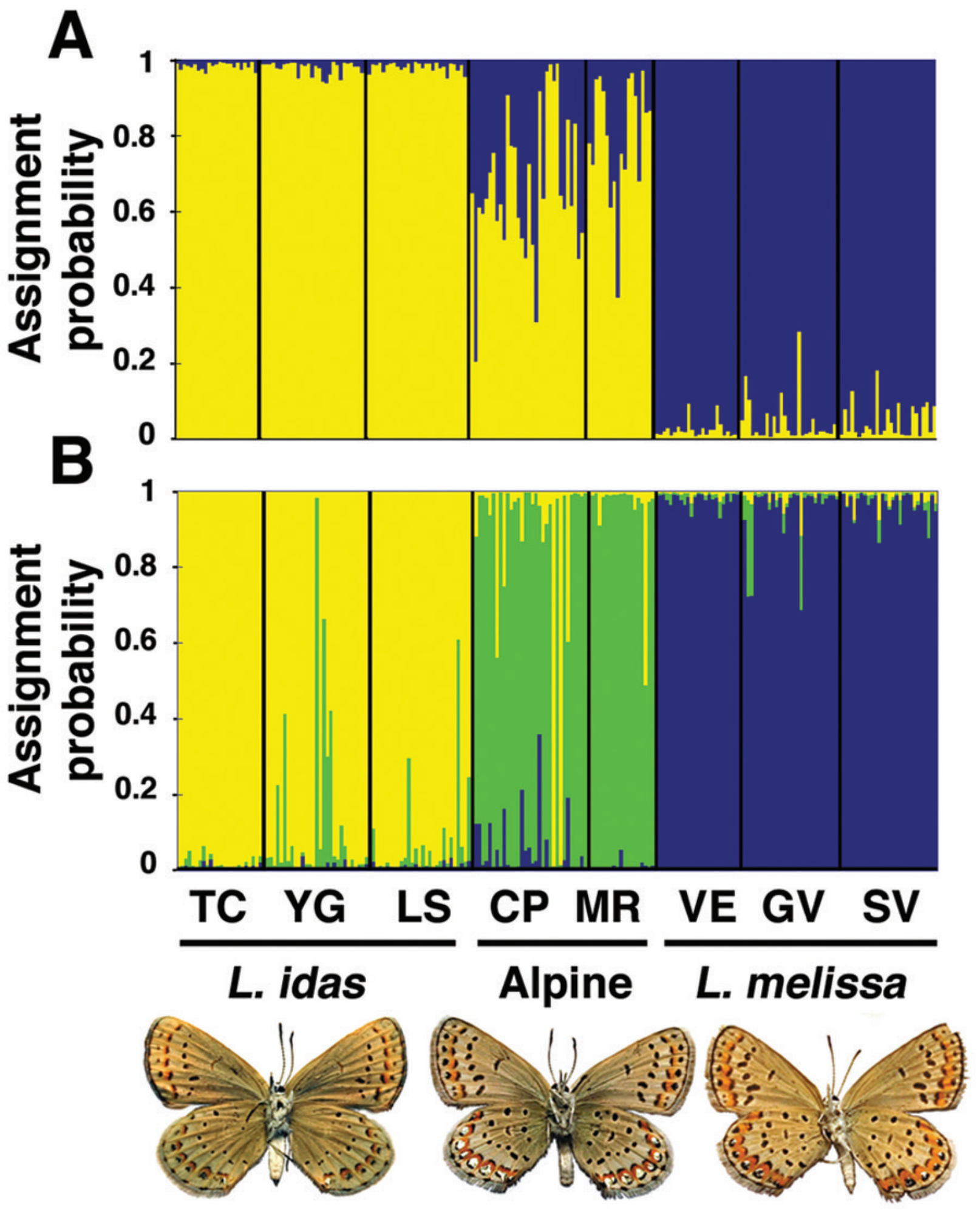
3 October 2006; accepted 16 November 2006
Published online 30 November 2006;
10.1126/science.1135875
Include this information when citing this paper.

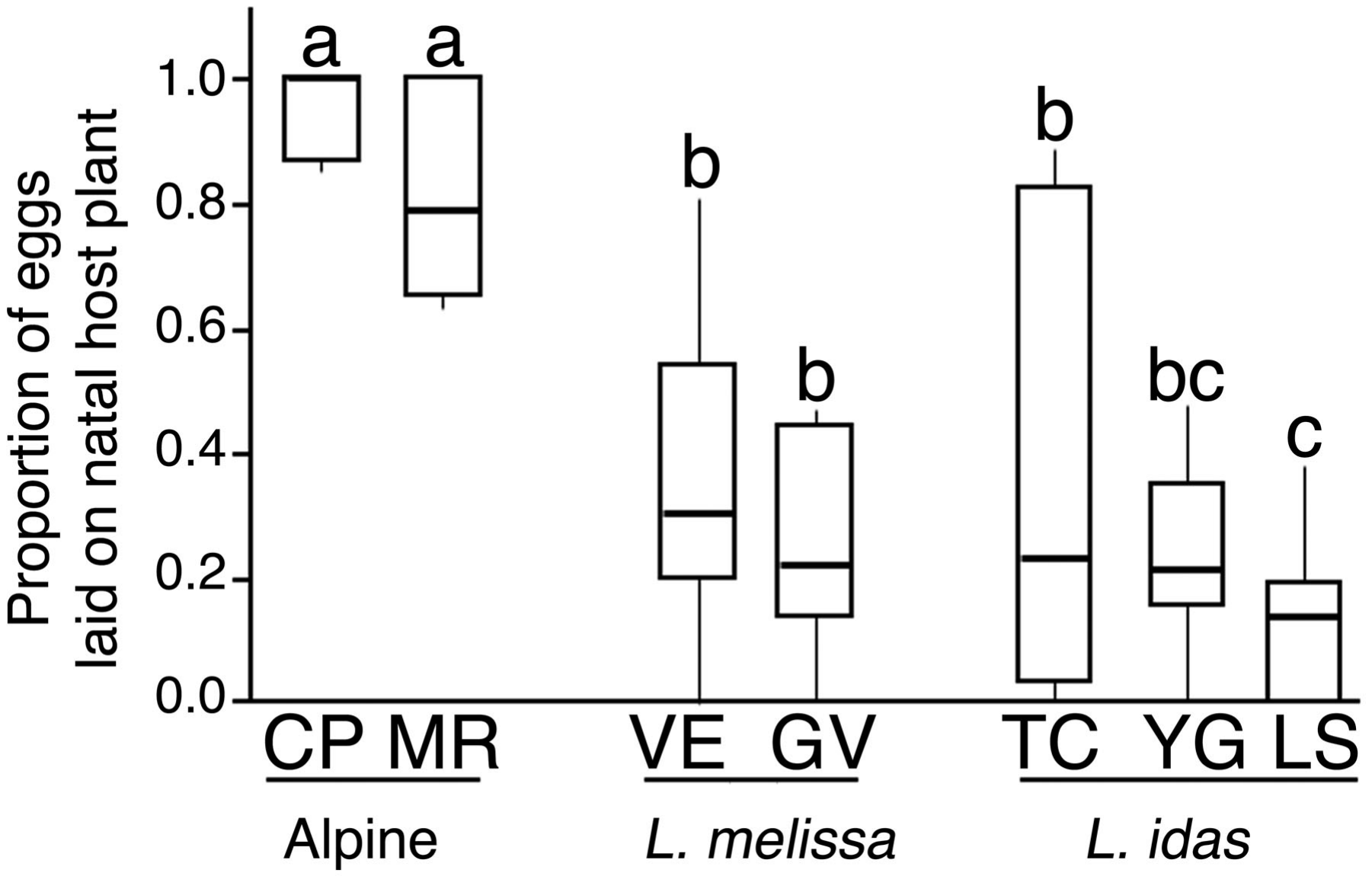
Fig. 1. Approximate range of *L. melissa* and *L. idas* in North America (**A**) and sampling localities for this study (**B**). The range map follows Nabokov (*23*), Stanford and Opler (*24*) and Scott (*25*). *Lycaeides idas* is shown in yellow, *L. melissa* is shown in blue, and regions of sympatry are shown in gray. The box denotes the focal region for this study. Sampling localities for *L.idas*, *L. melissa*, and alpine *Lycaeides* are shown in yellow, blue, and green respectively. (CP = Carson Pass, MR = Mt. Rose, VE = Verdi, GV = Gardnerville, TC = Trap Creek, YG = Yuba Gap, LS = Leek Springs).

Fig. 2. Bar plots showing Bayesian assignment probabilities from the software STRUCTURE 2.1 (21) for two (**A**) and three (**B**) clusters (*20*). Each vertical bar corresponds to one individual. The proportion of each bar that is yellow, blue, and green represents an individuals assignment probability to clusters one, two, and three respectively. See table S1 for mean population assignment probabilities. (CP = Carson Pass, MR = Mt. Rose, VE = Verdi, GV = Gardnerville, SV = Sierraville, TC = Trap Creek, YG = Yuba Gap, LS = Leek Springs)

Fig. 3. Natal host plant fidelity from seven focal populations. Box plots show the median proportion of eggs laid on the natal host plant for each population. Kruskal- Wallis test indicates significant differences in natal host plant preference among populations ($T = 46.67$; $p < 0.0001$). Different letters indicate differences in strength of preference for natal host among populations ($\alpha < 0.05$). Natal host plants are listed in table S1. (CP = Carson Pass, MR = Mt. Rose, VE = Verdi, GV = Gardnerville, SV = Sierraville, TC = Trap Creek, YG = Yuba Gap, LS = Leek Springs).


A
L. idas
L. melissa
Range of Sympatry
B
1500m
SV
1687m
TC
VE
1247m
1719m
YG
MR
2941m
Lake
Tahoe
GV
1537m
CP
LS
2817m
2298m


A
Assignment probability
1
0.8
0.6
0.4
0.2
0
B
Assignment probability
1
0.8
0.6
0.4
0.2
0
TC YG LS CP MR VE GV SV
L. idas
Alpine
L. melissa


Proportion of eggs laid on natal host plant
1.0
0.8
0.6
0.4
0.2
0.0
a
a
b
b
b
bc
c
CP
MR
VE
GV
TC
YG
LS
Alpine
L. melissa
L. idas