POPULATION GENETIC STRUCTURE IN THE ACMON COMPLEX OF *PLEBEJUS* (LEPIDOPTERA: LYCAENIDAE) IN WESTERN NORTH AMERICA

THESIS

Presented to the Graduate Council
of Texas State University-San Marcos
in Partial Fulfillment
of the Requirements

for the Degree

Master of SCIENCE

by

Maurine K. Spencer, B.S.

San Marcos, Texas
May 2005

To Clovis and Richard

ACKNOWLEDGEMENTS

I would like to thank the members of my committee for their help and insight on this project. Chris Nice introduced me to MLB.com Gameday Audio, which allowed me to maintain my sanity while aligning sequences and dissecting butterfly genitalia. He also offered a great deal of his time sitting and just thinking through this complex of butterflies with me. Mike Forstner gave me the best birthday present ever while I was here by taking me out of the lab for a day and letting me help re-landscape his backyard. I was happy to learn I'm pretty dad-gum good with a talache. Jim Ott and I have shared wonderful conversations about our love of vegetable celebration festivals, allowing my mind a break from thinking about why Acmon complex DNA refuses to sequence.

Thanks also to JCO, KLP, A. Warren, K. Davenport, and M. Klein for collecting and donating Acmon complex specimens. Special thanks to Paul Opler, Art Shapiro, and Jim Fordyce for their valuable input on this project.

Thanks to my labmates in the Genetics Lab, who have kept me entertained while my rear end was plastered to that chair in front of the computer for so long during my analyses and the writing of this thesis.

And probably the person most responsible for the completion of this study, my husband, Clay, has reminded me that I have to sleep, kept clean clothes in my closet, kept food in my stomach and reminded me of my perspective. He is my greatest blessing.

This manuscript was submitted on May 2, 2005.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

POPULATION GENETIC STRUCTURE IN THE ACMON COMPLEX OF *PLEBEJUS* (LEPIDOPTERA: LYCAENIDAE) IN WESTERN NORTH AMERICA

by

Maurine K. Spencer, B. S.

Texas State University-San Marcos

May 2005

SUPERVISING PROFESSOR: CHRIS NICE

Population genetic and phylogeographic data were used to investigate on-going divergence in a complex of butterflies of the genus *Plebejus*. *Plebejus acmon acmon* and *P. lupini lupini* occur sympatrically in California and show no signs of intermediacy. Outside of California, subspecies (*P. acmon texanus* and *P. acmon lutzi*) possess intermediate forms of the traits used to distinguish *P. acmon* from *P. lupini*. Sequence data from the nuclear gene *wingless* and mitochondrial cytochrome oxidase subunit I (COI) for 39 populations of the Acmon complex were used to elucidate geographical patterns of population genetic variation. Results from with *wingless* show no

genetic structure in this nuclear gene, indicating that divergence within the species complex has occurred recently or is on-going. Examination of spatial genetic structure of the COI region indicates one large homogenous clade in California occurring west of the Sierra Nevada crest, as well as nine other morphologically and genetically heterogeneous groups occurring mostly east of the Sierran crest. Evidence exists for a Pleistocene refugium in California, with subsequent range expansion inhibited by the Sierra Nevada. At least one other Pleistocene refuge existed east of the Sierra Nevada from which post-Pleistocene range expansion occurred. The morphological and genetic variation of the eastern groups indicates a complex geographical pattern of localized differentiation.

INTRODUCTION

Studies of speciation seek to establish the point at which interbreeding or potentially interbreeding organisms became reproductively isolated (Dobzhansky 1937). Thus, barriers to gene flow delineate species boundaries. The core process of speciation is the formation and maintenance of reproductive isolation. A variety of mechanisms may cause or contribute to reproductive isolation. Often, multiple barriers to gene flow exist between species (Dobzhansky 1970, Grant 1985). Postzygotic isolating mechanisms, associated with a period of allopatry, have historically been championed as the main mechanisms of reproductive isolation (Mayr 1963, Allmon 1992), however, increasing support has accumulated for prezygotic isolation as the initial mechanism responsible for divergence leading to speciation (McNeilly and Antonovics 1968, Feder 1998, Schluter 1998, Reiseberg *et al.* 2004). Identifying the mechanisms responsible for speciation events are difficult because reproductive barriers continue to accumulate after gene flow has stopped (Irwin *et al.* 2001, Reiseberg *et al.* 2004). Additionally, if the duration of the speciation process precludes adequate study within a human lifetime, only bits and pieces of the entire speciation event are actually observable to any one scientist. Nonetheless, if speciation is seen as a continuation of the processes causing population divergence, then the speciation process may be studied by examining the mechanisms causing isolation among populations currently near the species boundary (Schemske 2000, Via 2002). Studies of incipient species may identify some forms of reproductive

barriers initially responsible for the reduction of gene flow. This can then allow investigators to extrapolate their findings toward understanding the complete speciation process without waiting for complete reproductive isolation to occur among all forms of the organism being studied.

Lycaenidae is the most species rich family of butterflies (Papilioniodea), representing as many as half of all butterfly species (Robbins 1982, Scott 1986). This abundance raises questions regarding the origin and maintenance of that diversity. Eight of the 17 endangered butterflies listed by the U.S. Fish and Wildlife Service (USFWS) are of the family Lycaenidae. This study focuses on the mechanisms creating the species diversity in this group and may provide insight on how best to protect and even encourage maintenance of that diversity.

Previous studies of some lycaenid taxa show evidence of recent and rapid differentiation (Nice and Shapiro 2001, Nice *et al.* 2005) that, coupled with the strong tendency for members of this group to be locally specialized, suggests the potential for incipient speciation. In western North America, some of the members of the genus *Plebejus* Kluk (Lycaenidae*)* form a complex of nominal species, hereafter referred to as the Acmon complex. In California, the two nominal species, *Plebejus acmon* Westwood, 1852 and *Plebejus lupini* Boisduval, 1869, have been described based on morphological, ecological and behavioral differences (Scott 1986, Opler 2003). Outside of California, subspecies within this complex possess intermediate forms of the traits used to distinguish *P. acmon* from *P. lupini* (Figure 1) (Scott 1986). The observed, seemingly incomplete differentiation between groups in the Acmon complex suggests that they may be incipient species.

Potential causes of divergence in the Acmon complex

Identification of the trait differences that may be responsible for influencing divergence within the group may provide insight to those processes driving the divergence. Potential prezygotic barriers to gene exchange within the complex include host fidelity, wing pattern diversity and differences in reproductive structures. These characteristics influence mate choice, contribute to the local specialization typical of Lycaenid butterflies (Fordyce *et al.* 2002, Nice *et al.* 2002) and may constitute barriers to gene flow, facilitating divergence and potentially, speciation.

Host fidelity, defined as the tendency of an insect to reproduce on the same host species that it used in earlier life-history stages, has been shown to function as an effective premating barrier (Bush 1969, Feder *et al.* 1994) and may, theoretically, act as a mechanism of sympatric speciation. If preference for differential hosts is genetically encoded, host fidelity is strong and mating occurs on or near the host plant, then genetically isolated populations may arise by assortative mating (Bush 1969, Craig *et al* 2001, Emelianov *et al.* 2001). The type of host fidelity observed in the Acmon complex, who mate on the host plant, suggests that these butterflies may exist as host races, defined as sympatric populations between which gene flow is restricted by differential host preference (Jaenike 1981, Dres and Mallet 2002).

The function of wing pattern diversity in influencing mate choice has been supported as an additional possible prezygotic isolating mechanism in butterflies (Vane-Wright and Boppré 1993, Jiggins *et al.* 2004). Subtle wing pattern diversity, particularly in wing spots and the aurorae (orange crescent-shaped patterns on the wing), has been

found to function as a mate-recognition signal in some populations of Lycaenid butterflies (Fordyce *et al.* 2002). A similar pattern of mate recognition based on wing pattern diversity may limit gene flow between members of the Acmon complex. Differences in the form of genitalia may also play a role in divergence due to the possibility of mechanical isolation between morphs (Shapiro and Porter 1989), although support for characterizing species based on genital morphology is weak at best (Porter and Shapiro 1990).

Given the previous work on these two species (Goodpasture 1973a and 1973b, Scott 1986, Glassberg 2001, Opler 2003), analysis of geographical patterns of population genetic variation is now warranted. This analysis seeks to reveal whether or not genetic boundaries are congruent with the nominal species boundaries. If boundaries are incongruent, the analysis will delineate units of evolutionary significance, as well as provide insight into the biogeographical history of these butterflies. Secondly, if genetic boundaries cannot be explained by current morphological, ecological and behavioral differences, we can then ask what is driving differentiation? A comprehensive phylogeographic analysis will provide a foundation for an investigation of the relationship between genetic differences and the previously documented phenotypic differences present within the complex. Thus the results may provide insight to the causes and/or general rules governing the evolution of reproductive isolation and suggest directions for experimental tests of the efficacy of isolating mechanisms.

Objectives

This study addresses four major objectives. A morphological analysis of male and female genitalia is first performed to quantify the morphological variation as well as verify the validity of current information, provided by Goodpasture (1973a and 1973b) and Scott (1986). Given the above morphological information, the hypothesis that genetic variation is partitioned along nominal species boundaries is tested using phylogenetic inference and an analysis of molecular variation (AMOVA) (Excoffier *et al.* 1992). This second objective will test the legitimacy of the nominal taxonomy as a descriptor of evolutionary relationships. The third objective is to describe the geographic patterns of genetic structure in the Acmon complex and subsequently detect genetic barriers shaping those patterns. This is accomplished using a spatial analysis of molecular variance (SAMOVA), which partitions populations into genetically homogenous and geographically cohesive groups (Dunaloup *et al.* 2002), allowing description of spatial pattern of the Acmon complex without groups defined *a priori.* The fourth objective is to assess and describe the relative contributions of historical biogeographical events and recent ecologic processes to genetic diversity in the Acmon complex. Nested clade analysis (NCA) is used to address the question of whether historical demographic processes have shaped the distribution of genetic variation and may provide a description of those specific mechanisms acting in the Acmon complex.

MATERIALS AND METHODS

Butterfly biology

Morphological characteristics typical of the two nominal species occur within the boundaries of the state of California where populations have been observed in sympatry or parapatry without any observable intermediacy (Figure 1) (Goodpasture 1973b, Opler 2003). Outside of California, it is very difficult to separate *P acmon* and *P. lupini* since there are so many populations with intermediate characters (Scott 1986). This situation suggests that divergence within the Acmon complex has occurred relatively recently and/or that reproductive isolation is incomplete between the two species.

Morphological variation seems to occur in a circular pattern around the range of this complex until we reach *P. a. acmon* and *P. lupini*, which show no evidence of hybridization or intermediacy (Figure 1). It is reported that these two species do not interbreed (Scott 1986). Variation in morphology occurs among the Acmon complex, with several *P. acmon* subspecies having morphological traits resembling those of *P lupini*. These subspecies (*P. a. lutzi* and *P. a. texanus*) with intermediate morphologies occur only outside of California (Scott 1986). *P a. acmon* occurs in California, west Nevada, and southwest Oregon and follows the traditional morphology considered typical of the *P. acmon* species. In east Oregon, southwest Idaho, and southeast Washington, subspecies *acmon* intergrades with *P. a. lutzi*. Subspecies *lutzi* occurs from northwest Oregon to British Columbia, southeast to northern Utah and northern Colorado following

the Rocky Mountains (Scott 1986, Tilden and Smith 1986). This subspecies is a morphological intermediate of the two species, with *lupini* wing patterns (wide border on the upper side of the forewing and edging on the upper hind wing), but the traditional genitalia of *acmon*. *Plebejus a. spangelatus* is a high elevation form of *P. a. lutzi* found in the mountains of Washington and Colorado. Subspecies *P. a texanus* inhabits much of the Southwest including central Mexico north to southern Arizona and New Mexico and northeast to Colorado and Nebraska. *Plebejus acmon texanus* intergrades to *P. a. lutzi* in south and central Colorado and intergrades to subspecies *P. a. acmon* in the east Mojave Desert of California. Its morphology displays the wing pattern of *P. lupini* as well; the genitalia is often midway between *P. acmon* and *P. lupini,* but at times more similar to the genitalia of *P. lupini* (Scott 1986).

Morphology.— Plebejus lupini is generally larger than *P. acmon* with a wingspan ranging from 2.2-2.9 cm. The wingspan of *P. acmon* may range from 2.0-2.9 cm, thus having on average a slightly smaller wingspan than *P. lupini.* The two nominal species are distinguishable by differences in wing pigment patterns of males (Figure 2). The upper side of the forewing of the *P. acmon* male is bright blue with a narrow border. The upper side of the hind wing has a pink marginal band with no inside edging. In contrast, males of *P. lupini* usually boast a duller shade of blue and has a broader border on the forewing. The marginal band of the hind wing is edged on the inside border with a dark line (Scott 1986). The upperside of the hindwings has an orange marginal band in *P. lupini* rather than the pink band observed in *P. acmon.* The underside of *P. lupini* wings has larger, more conspicuous black wing spots on both the fore and hind wings compared to *P. acmon.* The *P. acmon* female wing pattern consists of smaller spots within the

orange marginal band of the hind wings compared to *P. lupini* (Figure 2), although this is now considered to be an unreliable character and females cannot be readily distinguished on the basis of wing characters.

Further differences are displayed in the reproductive structures of the butterflies (Figure 2). The male uncus structure is another point of divergence between these two groups. The uncus functions as the primary mechanism of attachment during mating events, fitting into a pocket just beneath the ovipositor in females (Scott 1986). *Plebejus acmon* has a deep cleft, while the uncus of *P. lupini* is cleft for only a short distance. In females, the lamellae of *P. lupini* end in a bulbar shape. In *P. acmon*, lamellae are narrow with only a few irregular hardened areas (Goodpasture 1973a, Scott 1986).

Host plant ecology.—Plebejus acmon exists in a wide-range of habitats, including deserts areas, fields, prairies, weedy areas, and road edges. *Plebejus acmon* is usually found in lowlands and rarely found in alpine environments (Scott 1986). Adults often feed on Wild Buckwheat flowers (*Eriogonum spp.*), but also feed on the nectar of many other species. Larval host plants of *P acmon* include both members of the buckwheat family Polygonaceae (*Eriogonum* and *Polygonum* species) and members of the legume family Fabaceae (*Lotus spp.*, *Astragalus*, *Lupinus*, and *Melilotus).* There are ambiguities in the literature concerning host plant use among the *P. acmon* subspecies. Glassberg (2001) states that legume use by *P. acmon* is limited to California and southern Oregon, only within the range of the subspecies *acmon.* Others report that, based on an inability to locate buckwheat species in some areas where Acmon blues occur, *P. a. texanus* and *P. a. lutzi* also utilize both buckwheat and legume species (Scott 1986, Glassberg 2001). *Plebejus lupini* habitat consists of alpine slopes, mountain meadows, prairies, rocky

outcroppings, chaparral and sagebrush areas. *Plebejus lupini* larvae are found exclusively on members of *Eriogonum* species, largely different from those species utilized by *P. acmon.* In both species, the males patrol near the host plants to seek females, who wait on the host plants for males to court them. Mating, egg laying and caterpillar feeding all occur on the respective larval host plants in this complex (Scott 1986).

Specimen collection

A total of 109 individuals from the Acmon complex were obtained for this study representing the two nominal species and 3 acmon subspecies (*acmon, lutzi,* and *texana*) from 39 different sites (Table 1, Figure 3). One *Plebejus neurona* individual was included in the analysis due to its close relationship with the Acmon complex (Opler 2003). Two *P. emigdionis* individuals were included, representing a closely related, but distinct taxon to serve as an outgroup (Opler, pers. comm.). Samples included both fresh specimens preserved in alcohol or refrigeration and dried museum specimens. In addition, several sequences were obtained from potential outgroup taxa for the phylogenetic analyses, including *Lycaeides melissa* and *L. idas.* Relationships within the Lycaenid family have not yet been fully resolved so no substantiated sister species is available for outgroup use. These taxa have been chosen because they are known to be closely related to the Acmon complex, yet wing pattern and genitalic differences strongly suggest that they do not share a most recent common ancestor with the members of the Acmon complex.

Genitalic morphology

In order to verify and quantify the genitalic morphology described by Goodpasture (1973a and 1973b) and Scott (1986) in the Acmon complex, genitalia were dissected from 108 specimens and a series of measurements were taken using an ocular micrometer. The 108 specimens included all Acmon complex specimens (minus two individuals whose abdomens were lost) and the *P. neurona* specimen. The posterior section of the abdomen was removed and soaked in 5 M KOH at 90-95°C (15 minutes for males, 8 minutes for females). The sclerotized genitalic structures were removed from the abdomen under a dissecting microscope and mounted under a cover slip. For female specimens, the mating tube and the lamellae were each measured at their greatest width using an ocular micrometer (Figure 4). For male specimens, three measurements on each half of the uncus were taken: 1) the distance from the top of each half of the uncus to the top of the corresponding peniculus, 2) the distance from the top of each half of the uncus to the point where the halves converge, and 3) across the width of the pad-like structure on the uncus (Figure 4). The mean of each equivalent measurement taken for both right and left sides was used for subsequent analyses.

Measurements of 55 male specimens and 50 female specimens in all were taken. Female mating tubes and lamellae in the Acmon complex are not always sclerotized, therefore, due to the inexperience of the dissector, three female structures were unaccounted for.

A Principle Components Analysis (PCA) using a correlation matrix and a 95% confidence ellipse was used to determine whether the genitalic differences among taxa within the Acmon complex are statistically significant. Given the complex pattern of

male morphological variation outside of California coupled with a lack of morphological intermediacy within California (Goodpasture 1973b, Opler 2003), three separate Principal Component Analyses were performed for 1) individuals from California only, 2) individuals outside of California, and 3) individuals from the sampling area as a whole. These analyses were performed using JMP IN 5.0.1a software (SAS, Cary NC).

Female morphological analysis proved difficult since only one measurement could be taken for female lamellae and females could not be distinguished as having either *P. acmon* or *P lupini* wing types. Thus, normality of the distribution in the group as a whole was determined using the Shapiro-Wilk test of normality (Shapiro and Wilk 1965) to describe the data with the assumption that, if the population consisted of only one group, the data would be distributed normally. If the Acmon complex is composed of two (or more) groups, the null hypothesis of normal distribution would be rejected. This analysis was also performed using JMP IN 5.0.1a software (SAS, Cary NC).

Nuclear and mitochondrial DNA

Although mitochondrial sequences are most often used resolve intraspecific relationships, the nuclear genome was also sampled, providing a replicate across molecules to verify that the mitochondrial gene genealogy is reflective of the species history. It is possible that the mitochondrial molecule may contain a different history due to an enhanced sensitivity to introgression or processes like selective sweeps (Ballard and Whitlock 2004). Moreover, since mitochondrial DNA is maternally inherited, nuclear DNA may provide additional insight into an organism's evolutionary history.

A 318-bp protein coding *wingless* gene was amplified and sequenced. *Wingless* is a nuclear gene used to resolve generic and species level relationships because of its relatively rapid rate of substitution (Brower and Egan 1997, Campbell *et al.* 2000). This gene was sequenced for 32 individuals that represented all nominal partitions of the Acmon complex then discontinued due to lack of variation (see Results).

Mitochondrial DNA sequence has been extensively used to reconstruct phylogenetic relationships, test biogeographic hypotheses and explore phylogeographic relationships (Templeton *et al* 1995, de Brito *et al.* 2002, Paulo *et al.* 2002, Janzen *et al.* 2002). Mitochondrial DNA may be particularly useful in studies of incipient species because it is both maternally inherited and haploid, therefore having an effective population size ¼ that of nuclear DNA (Ballard and Whitlock 2004). In conjunction with the higher mutation rates characteristic of mtDNA (Moriyama and Powell 1997), new alleles may fix faster in mtDNA and allow finer resolution of evolutionary relationships among diverged taxa (Simon *et al* 1994).

A 393-bp region of the mtDNA cytochrome oxidase subunit 1 (COI) was sequenced for all individuals. COI is protein coding and has been previously used to examine among and within species variation in butterflies (e.g. Brower 1994, Sperling and Hickey 1994).

DNA extraction was performed using the Puregene extraction kit as directed in the manufacturer's instructions. Polymerase Chain Reactions (PCR) consisted of 26.5 µl ddH_2O, 10.0 µl buffer A, 4.0 µl $MgCl_2$, 1.0 µl dNTP, 1.0 µl DMSO, 2.5 µl of both forward and backward primers for each DNA sequence, 2µl of genomic DNA and 0.5 µl *Taq* polymerase. PCR amplification was run on a MJ Research ® PTC-100 Thermal

Cycler that was programmed for 35 cycles of denaturing at 94°C for 1 minute, annealing at 42-53°C for 1 minute and extension at 72°C for 2 minutes followed by a final 10 minute extension at 72°C. Samples were held at 4°C until electrophoresed in an agarose gel to determine the success of the amplification process. Annealing temperatures were optimized for different runs. The primers used to amplify and sequence the nuclear WG gene were WG1a (5'-ACC GTG AAG ACC TGC TGG ATG-3') and WG2a (5'-ACT ICG CAR CAC CAR TGG AAT GTR CA-3') (Brower and DeSalle 1994). The primers used to amplify and sequence the mitochondrial COI gene were Ron (C1-J-1751) (5'-GGA TCA CCT GAT ATA GCA TTC CC-3') and Nancy (C1-N-2191) (5'-CCC GGT AAA ATT AAA ATA TAA ACT TC-3') (Simon *et al.* 1994). Amplified products were cleaned using the Marligen Inc. (Ijamsville, MD) purification kit as directed by manufacturer's instructions. Cycle sequencing was performed in a 9 µl reaction consisting of 2.0 µl Big Dye v. 3.1, 3.5 µl ddH_2O, 0.5 µl of the appropriate primer, and 2-3 µl purified PCR product, as per manufacturer's instructions. Cycle sequencing products were cleaned by filtering through a sephadex G50 matrix (SIGMA). The cleaned products were then electrophoresed in a polyacrylamide gel on an ABI PRISM 377XL DNA sequencer (Applied Biosystems, Inc.).

The resulting nuclear and mitochondrial DNA sequences were aligned by ClustalW and confirmed visually in the SeqEd and Sequencher programs. Unique haplotypes were identified and imported into PAUP* v.4.0b10 (Swofford 2003) for phylogenetic analysis.

Phylogenetic analysis

Phylogenetic analyses were performed for each gene separately, including both neighbor-joining and Bayesian analysis.

A heuristic search by parsimony of 1,000 replicates was used to obtain the most parsimonious tree for the dataset. All characters were considered unordered and equally weighted and character state optimization was set to ACCTRAN. The tree-bisection reconnection (TBR) branch swapping option was used with starting trees obtained by stepwise addition with addition sequences beginning with a random seed and MulTrees option in effect. Branches were collapsed to create polytomies if branch lengths equaled zero. Appropriate individuals were set as outgroups and set as the monophyletic sister clade of the resulting ingroup. Character diagnostics and a 50% consensus tree were produced for the resulting phylogenies. A bootstrap analysis was performed on the dataset. One thousand replicates were run with half of the total characters resampled in each replicate. A heuristic search by simple addition was done with all other options set as formerly described. A 50% bootstrap consensus tree was produced.

In order to select the appropriate model of molecular evolution for each gene, MacClade 4.05 (Maddison and Maddison, 2000) was used to note the number of changes per codon position site, transition/transversion ratio and nucleotide composition.

Modeltest 3.06 PPC (Posada and Crandall, 1998) was used to determine which model of DNA evolution best fits the dataset. Modeltest uses log-likelihood scores to test for goodness of fit for a series of nested models in which the alternative hypothesis represents a special case of the null hypothesis. Akaike information criterion (AIC)

(Akaike 1974) were then used to weight the models scores. AIC corrects bias introduced by using the same data for estimating parameters and calculating log-likelihood scores.

Neighbor-joining analyses were completed with the original unweighted dataset under the HKY model of evolution; this model was chosen for both the mitochondrial and nuclear genes based upon the results of the Modeltest analysis. Bootstrap analyses were run for this model and a 50% bootstrap consensus tree was constructed.

A Bayesian analysis was conducted using the model determined most representative of the data by Modeltest. The program MrBayes (Huelsenbeck and Ronquist, 2001) was used to conduct the analysis. Parameters for the analysis included the likelihood model with the number of substitution types set to two in accordance with the HKY model of evolution. Additionally, among site variation was set with a proportion of sites being invariable while the rate for the remaining sites were drawn from a gamma distribution. *Lycaeides melissa* was set as the outgroup. A Markov chain Monte Carlo analysis of 1,000,000 cycles was performed with a sample frequency of ten. The results of the first 10,000 cycles were discarded and a consensus tree of all 90,000 remaining trees was produced.

Phylogeographic analyses

Analyses of molecular variation (AMOVA) (Excoffier *et al.* 1992) were performed to partition the total molecular variance along current nominal species boundaries. Wing morphology was used to declare male individuals to either *P. acmon* or *P. lupini*. Since *P. acmon* and *P. lupini* females are indistinguishable morphologically, females were assigned species designations according to the males present in the same

populations AMOVA analyses were performed using the program ARLEQUIN version 2.0 (Schneider *et al* 2000).

The spatial structure of the Acmon complex was investigated using spatial analyses of molecular variation (SAMOVA) as implemented by the program SAMOVA 1.0 (Dupanloup *et al.* 2002), which partitions populations into genetically homogenous and geographically cohesive groups. The SAMOVA approach is unique in that group structure is based on genetic data only and no *a priori* definitions of groups based on non-genetic traits (i.e. morphological, behavioural or ecological traits) are required (Dupanloup *et al* 2002). Analyses were performed with different group sizes (k) ranging from 1 to 13.

A nested clade analysis was employed to discriminate between historical events and contemporary processes (methods described in detail by Templeton, (1998); see also Templeton et al. 1995; Durand et al. 1999; Turner et al. 2000). Nested clade analysis is designed to distinguish contemporary population processes from haplotype distributions that have resulted from historical biogeographical events. A hierarchically nested haplotype network and the geographic position of every individual sampled were used to test the null hypothesis of no geographical association among haplotypes with exact permutation contingency analysis. Two distance measures were also calculated for each haplotype and nested clade. The clade distance, D_C, is the average distance of every individual possessing a haplotype in a particular clade to the geographic center of that clade. Nested clade distance, D_N, is the average distance of every individual possessing a haplotype in a particular clade to the geographic center of the next higher hierarchically nested clade. Permutations are then used to measure the distribution of these distances

under the null hypothesis of no geographical association. Joint analysis of significantly large or small (at the 5% significance level) D_C and/or D_N values is then used to distinguish among models of population structure and several historical events including past fragmentation and range expansion following the inference key available in Templeton et al. (1995) and Templeton (1998).

The nested clade analysis of *Plebejus* mtDNA COI haplotypes was performed as follows: Statistical algorithms of Templeton *et al.* (1992) were used to test whether the use of maximum parsimony was justified by computing the probability that a site difference between two randomly drawn haplotypes resulted from more than one mutational event (Hudson 1989). For this, Θ was estimated using equation 10.5 of Nei (1987) and used in equation 10 of Hudson (1989). Theta is equal to $4N\mu$ (N is the effective population size and μ is the mutation rate). Once parsimony is justified, the method of Templeton *et al.* (1992) is used to estimate a network(s) with parsimonious connections between haplotypes having probabilities ≥ 0.95. The maximum parsimony network configuration was determined using the TCS software (Clement et al. 2000). This maximum parsimony haplotype network of COI sequences was then nested into a series of hierarchically grouped clades following the nesting rules of Templeton et al. (1987). Tests of geographical association, calculation of D_C and D_N, and permutation tests were performed using GEODIS version 2.0 (Posada et al. 2000). Significance levels were estimated from 1000 random permutations of the data. Inferences about historical and contemporary population genetic processes were made following the inference key of Templeton (1998).

RESULTS

Morphological analyses

On the basis of variation in the uncus, PCA was unable to distinguish males of the nominal species *P. acmon* from *P. lupini*, nor identify other patterns of morphological variation within the group with confidence. However, some support was provided for possible boundaries of morphological variation within California populations. In the analysis of male specimens in California ($n = 32$) (Figure 5A), PCA recovered three principal components; only the first two components are reported because the third component had an eigenvalue less than 0.10 and explained less than 3.3% of the variation. The first principal component captures size differences primarily and the second appears to capture shape variation. The first component (PC1) explained 66.1% of the variation and had an eigenvalue of 1.98. The second component (PC2) explained 30.6% of the variation and had an eigenvalue of 0.92. Genitalic measurements had positive loadings on PC1, the largest in all analyses being measurement W (see Figure 4 for explanation of measurements), which had a loading of 0.687. Only loadings for measurement U were positive for PC2, which had a loading of 0.949. Males designated as *P. acmon* by wing pattern form a coherent group and appear distinct from *P. lupini* males. However the sample size for *P. lupini* is quite small ($n = 3$) and these results are considered preliminary.

Individuals designated *P. acmon* outside of California or in the group as whole are indistinguishable by size and shape from those individuals designated as *P. lupini* due to the overlap seen the plot. Analyses of male specimens both outside of California ($n = 23$) (Figure 5B) and all male specimens ($n = 55$) (Figure 5C) produced identical statistical results. PCA recovered three principal components with both groups; only the first two components are reported because the third component had an eigenvalue less than 0.19 and explained less than 6.3% of the variation. PC1 explained 64.8% of the variation and had an eigenvalue of 1.94. PC2 explained 28.9% of the variation and had an eigenvalue of 0.87. Again, all genitalic measurements had positive loadings on PC1, the largest in all analyses being measurement W (see Figure 4), which had a loading of 0.664 in the two analyses. Only loadings for measurement U were positive for PC2 at 0.936. Again, the first component appears to capture size differences and the second capturing shape variation.

Female genitalic analysis confirmed the variation in the lamellae as described by Goodpasture (1973a). Mean lamella width ($n = 49$) is 165.1 mm (SD = 78.0). The Shapiro-Wilk test found the variation was not distributed normally ($W = 0.86$, $p = <0.0001$) for all females considered as one group, suggesting patterns of morphological variation within the group (Figure 6).

Phylogenetic analyses – nuclear data

The skewness of the distribution of 10,000 randomly distributed trees was significant ($g^1 = -0.681$), indicating phylogenetic signal in the dataset (Hillis and Huelsenbeck 1992).

Of 32 *Plebejus* WG sequences, 11 unique haplotypes were identified. A total of 318 bp provided 32 variable sites and 12 of them were parsimony informative. The mean nucleotide frequencies were 19.8%, 34.6%, 34.6% and 10.9% for A, C, G, and T, respectively. The mean transistion to transversion ratio was 8.0 and the percentage of codon positional changes was 94.1% at the 3rd position, 5.8% at the 1st position and 2.9% at the 2nd position.

Preliminary data from the nuclear *wingless* gene indicated a lack of structured genetic variation for this gene in the Acmon complex. Neighbor-joining and Bayesian analyses were performed on the 32 individuals sequenced. Resulting phylogenies were almost completely unresolved due to lack of bootstrap support (Figure 7), although the two *Plebejus emigdionis* individuals formed a strongly supported monophyletic clade corresponding to the nominal classification based on morphology.

Lack of structured genetic variation in the *wingless* gene warranted investigation of the cause of this uniformity. Tajima's D statistic (Tajima 1987) and Fu's F_S (Fu 1997) were calculated to test for departures from mutation/drift or mutation/selection equilibrium in the *wingless* gene. Tajima's D was not significant at the 5% level (D = -1.365, P = 0.0744), nor was Fu's F_S (F_S = -2.338, P = 0.135), not rejecting the null hypothesis of neutral evolution at this gene.

Phylogenetic analyses - COI

Of 108 *Plebejus* COI sequences, 38 unique haplotypes were identified. A total of 390 bp provided 65 variable sites and 49 of them were parsimony informative. The mean nucleotide frequencies are 31.4%, 16.8%, 13.7% and 38.1% for A, C, G, and T,

respectively. The mean transistion to transversion ratio was 1.95 and the percentage of codon positional changes was 85.1% at the 3rd position, 12.8% at the 3rd position and 1.9% at the 2nd position. Sequence divergence (corrected by HKY) between haplotypes ranged from 0.24 to 3.78%. Haplotype diversity (h) in the COI region was calculated according to Nei (1987) for each population and ranged from 0 – 0.893, with a mean haplotype diversity (μ) of 0.367 (Table 1).

Phylogenetic analyses of the mtCOI dataset produced a slightly more resolved inference than the nuclear dataset, but again did not distinguish between the two nominal species. The neighbor-joining analysis and bootstrap resampling were performed on the 38 unique haplotypes (Figure 8). Bootstrap support was weak for all monophyletic clades detected within the Acmon complex.

The HKY+I model of DNA substitution was selected for Bayesian analysis by the AIC in ModelTest (Posada and Crandall, 1998). This model rejects the null hypotheses of equal base frequencies ($\pi_A = 0.317$, $\pi_C = 0.149$, $\pi_G = 0.139$, $\pi_T = 0.396$), equal transition to transversion rate (ti/tv = 4.481), and no invariable sites (I = 0.794). Bayesian analysis shows few internal nodes are well supported by the Bayesian posterior probabilities (P_B) (Figure 9). A large clade corresponding to low elevation localities from Oregon and California is well supported ($P_B = 100$). Two smaller clades from California are also well supported, one (haplotypes J, A, and K) corresponding to high elevation localities in the Eastern Sierra and White Mountains ($P_B = 98$) and the other (haplotypes F and Q) corresponding to localities in southern California ($P_B = 99$). All haplotypes found outside of California (except haplotype C, found in western Nevada) formed a moderately supported ($P_B = 74$) clade consisting of one monophyletic clade from Arizona

(P_B = 83) and one monophyletic clade from eastern Nevada, Wyoming and Colorado (P_B = 86).

AMOVA

Populations from which only females were sampled could not be confidently assigned to either *P acmon* or *P. lupini* thus, butterflies from only 27 of the 39 populations could be confidently assigned a species designation based on wing morphology (see Table 1); the 12 remaining populations were excluded from the analysis. Populations partitioned into two groups by nominal taxonomy explained 48.28% of the total genetic variance among species (ϕ_{CT} = 48.28, P = 0.00) (Table 2).

SAMOVA

As a comparison to the analysis partitioning total genetic variation between nominal species based on morphology, SAMOVA was employed to identify groups of populations that are phylogeographically homogenous. The SAMOVA algorithm was first used at k = 2 to determine whether it would partition the groups along the same nominal boundaries used in the previous AMOVA analysis. The groups identified by AMOVA and SAMOVA differed by 22% of their component populations (see Table 1). The SAMOVA partitioning accounted for 65.54% of the total genetic variance among groups (ϕ_{CT} = 65.54, P = 0.00) and appears to describe significantly more of the total variation than nominal taxonomic boundaries (Table 2).

The SAMOVA algorithm was then used to identify the number and composition of groups maximizing ϕ_{CT}. Ten groups accounted for 81.83% of the total genetic

variance among groups (ϕ_{CT} = 81.83, P = 0.00) (Figure 10, Table 2). These groups consisted of one large group in California mostly west of the Sierra Nevada, several population groups in eastern California, and several groups outside of California (Figure 11, Table 1).

Nested clade analysis

A statistical parsimony haplotype network (Templeton *et al* 1992) of the 38 haplotypes in the *Plebejus* dataset with 95% confidence consisted of four loops (representing ambiguities) that required breaking before nested cladograms could be constructed (Figure 12). Loops were broken based on predictions from the coalescent theory (Crandall and Templeton 1993, Smouse 1998, Posada and Crandall 2001). According to coalescent theory, it is expected that older alleles occur more frequently (Donnally and Tavaré 1986). Thus, frequent haplotypes, having more time to give rise to new haplotypes, are often interior nodes in a network with a greater number of mutational connections (Pfenninger and Posada 2002). Additionally, it is expected that singletons are more likely to remain connected to haplotypes from the same geographical location than to haplotypes from distant locations unless high levels of gene flow are occurring (Pfenninger and Posada 2002). Therefore, loops were broken to minimize the number of singletons connecting to other singletons in the network and maximize mutational connections on haplotypes with greater frequencies. When applicable, breaks were made with adjacent missing haplotypes before making them between observed haplotypes.

It is important to note that in the case of intraspecific phylogenies, loops may represent real relationships (Posada and Crandall 2001). In contrast to interspecific taxa, ancestral haplotypes are often extant and in many cases, more prevalent than derived haplotypes (Donnelly and Tavaré 1986, Crandall and Templeton 1993). As such, it is not unreasonable to imagine two different, but closely related ancestral haplotypes giving rise to the same derived haplotype.

The inference key of Templeton (2004) was used to make inferences for the four equally parsimonious topologies considered. Because there were only relatively minor differences resulting from these four nesting designs, results are presented only from the one most likely topology (according to coalescent predictions) is presented here with major differences in inference for the other topologies noted when relevant.

NCA (Figures 13-14, Table 3) detected 4 main 3-step clades. In Clade 3-4 and at the total cladogram level, geographical patterns compatible with contiguous range expansion (CRE) were detected. This was consistent with all other topologies except one, which detected significant geographical association at the total cladogram level, but with an inconclusive outcome. For Clades 2-1, 2-4 and 3-2 a pattern of restricted gene flow with isolation by distance was suggested (although in one of the three other topologies either Clade 2-1 or 3-2 did not show a significant pattern of isolation by distance).

DISCUSSION

Overall for the Acmon complex, analyses of morphological characters result in an unclear picture of the relationships among individuals in the Acmon group as a whole. Nominal species boundaries, identified historically by wing morphology and genitalic structures, thus seem to have little support. Although overall and spatial variation is present in both male and female genitalic structures in the Acmon complex, this variation is not concordant with distinct morphological boundaries for either gender. In addition, the taxonomy of this group is further contradicted by genetic boundaries identified by examination of spatial genetic structure, which differ substantially from those boundaries identified by current taxonomy. One large homogenous clade was identified in California, west of the Sierran crest (Figures 9 and 11). Nine additional genetically differentiated groups were detected east of the Sierran crest. These groups comprise a morphologically and genetically heterogeneous mosaic, mostly east of the Sierran crest.

Morphological variation of both male and female genitalic structures were found to be as described by Goodpasture (1973a, 1973b) and Scott (1986) though, results of PCA show no significant patterns of morphological variation. Some case may be made for possible structure within California. All *P. acmon* wing type males with the exception of one individual cluster on the basis of PCA. *P. lupini* individuals seem to be distinguishable from *P. acmon*, having higher scores on PC1 (size) (Figure 5A). Since all of the genitalic measurements were positively correlated, PC1 may be regarded as a size

component (Flessa and Bray 1977). Thus, *P. lupini* in California vary from *P. acmon* in size, with *P. lupini* being larger than *P. acmon.* The remaining components (including PC2) represent size-free shape components and account for (28.9 - 30.6%) of the variation in genitalic morphology. However, the nominal species are indistinguishable by shape due to considerable overlap exists between the two nominal species on PC2.

Results of AMOVA partitioned by nominal species boundaries explain relatively little of the total genetic variation among the sampled locations, suggesting that mtDNA variation is distributed along geographic boundaries more than nominal species boundaries. While the distribution in morphological variation is ambiguous for defining the boundaries within the Acmon complex, significant genetic boundaries appear to exist. SAMOVA results suggest a mosaic pattern of genetic structure in the Acmon complex consisting of ten groups, with one large homogeneous group located in mostly western California at low elevations and the rest of the groups being small and scattered at higher elevations throughout eastern California, and continuing to the easternmost part of their range.

Results of NCA (Figures 12-14, Table 3) suggest structure at the total cladogram level as a result of contiguous range expansion, defined as an expanding population front due to individual short-distance dispersal (Templeton *et al.* 1995). Contiguous range expansion can be inferred when some ancestral (interior) haplotypes are restricted to the ancestral region, while younger (tip) haplotypes are geographically widespread or located distantly from their ancestoral haplotypes (Cann *et al.* 1987, Templeton *et al.* 1995). Both the clade distance (D_C), measuring the geographical range of the clade, and nested clade distance (D_N), measuring how a clade is geographically distributed relative to its

sister clades, are significantly large for Clade 3-3 at the level of the total cladogram. Since Clade 3-3 represents a tip at this level, the pattern is representative of an expanding population by short distance dispersal. Additionally, Clade 3-4 shows a pattern of CRE as inferred by a significantly small D_C at the interior node of the clade (Clade 2-7), conforming to the predicted pattern of a restricted ancestral region with distantly located derived haplotypes under CRE.

Within California, restricted gene flow due to isolation by distance is inferred for Clades 2-1, 2-4, and 3-2. This predicts that older haplotypes will be more widespread throughout the range while younger haplotypes will tend to be more restrictive relative to the range of their immediate ancestor(s). Vagility in this complex is low thus, the pattern of isolation by distance is indicative of a stable system of populations connected by limited gene flow, whose presence in that region is well-established and in which there is equilibrium between drift and migration (Slatkin 1993). Populations in California exhibit this structure with geographically distant populations tending to also be genetically distant. For neutral variation, this restriction of gene flow is due to the range of the group exceeding the dispersal capabilities of the individual, thus population structure may result from population subdivision as predicted by Wright's stepping stone model of gene flow (Wright 1943).

SAMOVA identifies one large, coherent group consisting of 16 of the 26 populations sampled in the state (Figure 11, Group 1). This group (the Central California clade) is found at low elevations in the Central Valley and surrounding foothills. Both NJ and Bayesian analyses support this grouping and identify it as a monophyletic clade (Figures 8-9). There was little incongruence between the phylogenetic clades and Central

California clade, although NJ and Bayesian analyses include haplotype O (population 11) in the group (Figures 8 and 9). Additionally, the Central California clade is identified in the haplotype network as being separated by four missing haplotypes from the rest of the network (Figure 12). Haplotypes MM, O, and V are split from the rest of their closely related haplotypes and placed in Clade 3-2 as an artifact of nesting design and decisions relating to breaking the loops in this area. All remaining haplotypes in the Central California clade make up the total of Clade 3-1. Nonetheless, high levels of support from different analytical methods substantiate the Central California clade as a distinctive, independent lineage.

The Sierra Nevada seems to form a significant barrier to gene flow between the Central California clade and the rest of the sampled populations. This is a characteristic boundary identified in other phylogeographic studies of a variety of taxa (Calsbeek 2003), including butterflies (Forister *et al* 2004, Nice *et al.* 2002, 2005). All populations comprising the Central California clade occur west of the Sierra Nevada crest and at low elevations. The only exception to this is population 30. This population is still found at a low elevation, but on the east side of the Sierra crest. The Transverse Range does not seem serve as a significant barrier to gene flow in this complex, as populations in the Central California clade cross through these mountains (Figure 11). Thus, it is probable that population 30 is a result of this clade expanding through the Transverse Range and colonizing the eastern side of the Sierra. No haplotypes from the Central California are shared with populations to the east of the Sierran crest, strongly supporting the assertion of genetic isolation of this group from the rest of the Acmon complex. Haplotypes M and L (the most frequent and diversely distributed haplotypes in California) are closely

related to many recently derived haplotypes (tips on the haplotype network) that are geographically restricted to western California, a pattern indicative of restricted gene flow as discussed previously (Templeton 1998).

Morphologically and ecologically, the Central California clade is comprised of individuals displaying traditional *P. acmon* phenotype. Male wing patterns display the conventional bright blue wing color with pink marginal bands. The Central California clade is also homogenous in genitalic morphology, comprising most of the cluster observed in Figure 5A. Ecologically, its populations are located at low elevation, generally below 3000 feet and while *P. acmon* morphology is at times observed outside of this group, *P. lupini* morphology is not observed within it.

The nuclear *wingless* sequence was chosen to provide a replicate across molecules to substantiating the congruence of the mitochondrial gene genealogy and species history. Lack of divergence in this gene unfortunately does not allow this comparison. Lack of genetic structure in the *wingless* gene may be interpreted in several ways including 1) selection has reduced variation in *wingless*, 2) immigration is high enough to homogenize the populations at the level of the entire sampling area, or 3) the nuclear gene is evolving at a slower rate than the mitochondrial gene. Neither Tajima's nor Fu's tests rejected the null hypothesis of neutrality thus, there is no evidence of selection. Additionally, mark-recapture studies indicate dispersal rates for Lycaenids to be limited, on average less than 500m (King 1998) thus, it does not seem realistic that this level of dispersal could homogenize populations across such great distances. It is reasonable then to attribute lack of genetic structure in the nuclear *wingless* region to a slower rate of evolution, suggesting events shaping the variation in the Acmon complex are relatively recent.

The genetic patterns observed in the Acmon complex are likely a result of divergence in multiple Pleistocene refugia and the subsequent post-Pleistocene expansion in response to newly exposed suitable habitat at the retreat of the glaciers (Pielou 1991). Geographic and genetic structure in this complex seem to indicate two or more Pleistocene refugia, one located in California, west of the Sierra, and possibly two others of unknown localities corresponding to Clades 3-3 and 3-4.

Evidence for the California refugium is strong. A stable pattern of isolation by distance has been demonstrated in California, specifically for the Central California clade, indicating it is of relatively old lineage. This pattern is typical of areas serving as refugia during glacial maxima and has been demonstrated for a variety of taxa (Brant and Orti 2003, Kuchta and Tan 2005). The clade associated with this refugium is morphologically and ecologically homogenous and genetically isolated from the rest of the complex.

Populations east of the Sierra differ genetically, morphologically and ecologically from those associated with the Central California clade. If California represented the sole refugium for this complex, haplotypes of high frequencies (i. e. haplotypes M and L) would be expected to be prevalent along the path of dispersal to the eastern portions of the range as a consequence of the leading edge effect (Ibrahim *et al.* 1996) associated with contiguous range expansion. This is not the pattern given by the data.

SAMOVA describes populations east of the Sierra as small, geographically disjunct groups (Figure 11). Five of these are comprised of populations wholly or partially in California, but associated with higher elevations than those typical of the Central California clade. The remaining four groups defined by SAMOVA are fully

confined to sampled areas outside of California. These small groups are genetically and morphologically heterogeneous and there are no clearly observable boundaries separating these groups from each other. It is with these observations that the inference of multiple refugium is made. Unfortunately, the location of these hypothesized refugia may not be identified in this study due to limited sampling outside of California.

This group of populations presents a mosaic of morphological and genetic differentiation. A large amount of morphological and genetic variation exists between these groups, suggesting that differentiation may be occurring, but there has been insufficient time for evolutionary processes to clearly sort out the variation among recently expanded populations. Two hypotheses may explain the origin and pattern of this variation in light of known lycaenid biology and ecology. First, the mosaic pattern may be a product of drift acting to differentiate groups from one another. Members of the Acmon complex are small butterflies with low vagility. This, along with the potentially patchy distributions of host plants, may isolate populations enough for drift to significantly affect allele frequencies. An alternative hypothesis is that selection created and maintains differentiation between groups. Lycaenids often exhibit strong host plant fidelity associated with mating events occurring on or near the host plant (Nice and Shapiro 2001, Nice *et al.* 2002). Thus, gene flow may be restricted between groups associated with different host plants. Furthermore, this trend may be reinforced if host plant recognition and assortative mating are correlated.

The above hypotheses are not necessarily mutually exclusive. Lycaenids in general are a speciose group in part due to their tendency to locally specialize to their environment (see Fraser *et al.* 2002, Fordyce and Nice 2003). Many of these eastern

groups seem to be adapted to high elevations, which provide both geographic isolation and an environment well-suited for local adaptation. As a result, gene flow may be severely restricted.

Ecological and behavioral studies are needed to explore the potential force(s) driving the divergence observed both in the Sierra and to the east. Evolution in this group of butterflies seems to be progressing in a complicated manner and definable groups may not presently exist in this part of the complex's range.

Conclusions

The Acmon complex has historically been partitioned into two species groups by the morphological, ecological and behavioral differences between them. This variation, however, does not conform to distinct boundaries thus, an analysis of the geographical pattern of morphology and population genetic variation served as the next step in the investigation of the evolutionary history of this polytypic complex. Both morphological patterns of genitalic variation and geographic patterns of the mtDNA COI region observed in this study are discordant with the two traditionally recognized species. Instead, strong evidence for one large *P. acmon*-like group (the Central California clade) associated with low elevations west of the Sierran crest is observed. This group is genetically distinct from all other groups sampled, sharing no haplotypes with sampled locations to the east of the Sierra. The remaining populations seem to be divided into much smaller groups that create a mosaic pattern both morphologically and ecologically. These latter populations vary in host plant preference, wing pattern, and reproductive structures, reinforcing an environment that may restrict gene flow and drive divergence.

Some of these populations seem to be locally specialized to high elevations in the Sierra and surrounding mountain ranges, potentially facilitating the further differentiation of these groups. It is possible that the complicated genetic and phenotypic consequences result from multiple Pleistocene refugia, subsequent range expansion after the glacial retreat, and subsequent localized differentiation.

LITERATURE CITED

Akaike, H. 1974. A new look at statistical model identification. IEEE Transactions on Automatic Control 19:716-723.

Allmon, W. D. 1992. A causal analysis of stages in allopatric speciation. Pp. 219-257 *in* D. Futuyma and J. Antonovics, eds. Oxford surveys evolutionary biology. Oxford University Press.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13:729-744.

Brant, S. V., and G. Orti. 2003. Phylogeography of the Northern short-tailed shrew, *Blarina brevicauda* (Insectivora: Soricidae): past fragmentation and postglacial recolonization. Molecular Ecology 12:1435-1449.

Brower, A. 1994. Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). Molecular Phylogenetics and Evolution 3:159-174.

Brower, A. V. Z., and M. G. Egan. 1997. Cladistic analysis of *Heliconius* butterflies and relatives (Nymphalidae: Heliconiiti): a revised phylogenetic position for *Eueides* based on sequences from mtDNA and a nuclear gene. Proceedings of the Royal Society of London 264:969-977.

Bush, G. 1969. Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis*. Evolution 23:237-251.

Calsbeek, R., J. N. Thompson, and J. E. Richardson. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. Molecular Ecology 12:1021-1029.

Campbell, D. L., A. V. Z. Brower, and N. E. Pierce. 2000. Molecular evolution of the *wingless* gene and its implications for the phylogenetic placement of the butterfly family Riodinidae (Lepidoptera : Papilionoidea). Molecular Biology and Evolution 17:684-696.

Cann, R. L., M. Stoneking, and A. C. Wilson. 1987. Mitochondrial DNA and human evolution. Nature 325:31-36.

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657-1659.

Craig, T. P., J. D. Horner, and J. K. Itami. 2001. Genetics, experience, and host-plant preference in *Eurosta solidaginis*: implications for host shift and speciation. Evolution 55:773-782.

Crandall, K. A., and A. R. Templeton. 1993. Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. Genetics 134:959-969.

de Brito, R. A., M. H. Manfrin, and F. M. Sene. 2002. Mitochondrial DNA phylogeography of Brazilian populations of *Drosophila buzzatii*. Genetics and Molecular Biology 25:161-171.

Dobzhansky, T. 1937. Genetics and the origin of species. Columbia University Press, New York.

Dobzhansky, T. 1970. Genetics and the evolutionary process. Columbia University Press, New York.

Donnelly, P., and S. Tavaré. 1986. The ages of alleles and the coalescent. Adv. Appl. Prob. 18

Dres, M., and J. Mallet. 2002. Host races in plant-feeding insects and their importance in sympatric speciation. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 357:471-492.

Dupanloup, I., S. Schneider, and L. Excoffier. 2002. A simulated annealing approach to define the genetic structure of populations. Molecular Ecology 11:2571-2581.

Durand, J. D., A. R. Templeton, B. Guinand, A. Imsiridou, and Y. Bouvet. 1999. Nested clade and phylogeographic analyses of the chub, *Leuciscus cephalus* (Teleostei, Cyprinidae), in Greece: Implications for Balkan Peninsula biogeography. Molecular Phylogenetics and Evolution 13:566-580.

Emelianov, I., M. Dres, W. Baltensweiler, and J. Mallet. 2001. Host-induced assortative mating in host races of the larch budmoth. Evolution 55:2002-2010.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial-DNA restriction data. Genetics 131:479-491.

Feder, J. L. 1998. The apple maggot fly *Rhagoletis pomonella*: flies in the face of conventional wisdom about speciation. Pp. 130-144 *in* D. H. a. S. Berlocher, ed.

Endless forms: species and speciation. Oxford University Press, New York.

Feder, J. L., S. B. Opp, B. Wlazlo, K. Reynolds, W. Go, and S. Spisak. 1994. Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. Proceedings of the National Academy of Sciences of the United States of America 91:7990-7994.

Flessa, K. W., and R. G. Bray. 1977. On the measurement of size-independent morphological: an example using successive populations of a Devonian spiriferid brachiopod. Paleobiology 3:350-359.

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

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

Forister, M. L., J. A. Fordyce, and A. M. Shapiro. 2004. Geological barriers and restricted gene flow in the holarctic skipper *Hesperia comma* (Hesperiidae). Molecular Ecology 13

Fraser, A. M., T. Tregenza, N. Wedell, M. A. Elagara, and N. E. Pierce. 2002. Oviposition tests of ant preference in a myrmecophilous butterfly. Journal of Evolutionary Biology 15:861-870.

Fu, Y.-X. 1997. Statistical test of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915-925.

Glassberg, J. 2001. Butterflies through binoculars, The east. Oxford University Press, Oxford.

Goodpasture, C. 1973a. A new subspecies of *Plebejus acmon* (Lepidoptera: Lycaenidae). The Pan-Pacific Entomologist 49:149-159.

Goodpasture, C. 1973b. Biology and systematics of the *Plebejus* (*Icaricia*) *acmon* group (Lepidoptera: Lycaenidae) I. Review of the group. Journal of the Kansas Entomological Society 46:469-485.

Grant, V. 1985. The evolutionary process: a critical review of evolutionary theory. Columbia University Press, New York.

Hillis, D. M., and J. P. Huelsenbeck. 1992. Signal, noise and reliability in molecular phylogenetic analysis. Journal of Heredity 83:189-195.

Hudson, R. R. 1989. How often are polymorphic restriction sites due to a single

mutation? Theoretical Population Biology 36:23-33.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-755.

Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation by different forms of dispersal during range expansion. Heredity 77:282-291.

Irwin, D. E., S. Bensch, and T. D. Price. 2001. Speciation in a ring. Nature 409:333-337.

Jaenike, J. 1981. Criteria for ascertaining the existence of host races. American Naturalist 117:830-834.

Janzen, F. J., J. G. Krenz, T. S. Haselkorn, and E. D. Brodie. 2002. Molecular phylogeography of common garter snakes (*Thamnophis sirtalis*) in western North America: implications for regional historical forces. Molecular Ecology 11:1739-1751.

Jiggins, C. D., C. Estrada, and A. Rodrigues. 2004. Mimicry and the evolution of premating isolation in *Heliconius melpomene* Linnaeus. Journal of Evolutionary Biology 17:680-691.

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.

Kuchta, S. R., and A.-M. Tan. 2005. Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. Molecular Ecology 14:225-244.

Maddison, D. R., and W. P. Maddison. 2000. MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, Massachusetts.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge.

McNeilly, T., and J. Antonovics. 1968. Evolution in closely adjacent plant populations. IV. Barriers to gene flow. Heredity 23:205-218.

Moriyama, E., and J. Powell. 1997. Synonymous substitution rates in *Drosophila*: mitochondrial versus nuclear genes. Journal of Molecular Evolution 45:378-391.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY.

Nice, C. C., J. A. Fordyce, A. M. Shapiro, and R. Ffrench-Constant. 2002. Lack of

evidence for reproductive isolation among ecologically specialised lycaenid butterflies. Ecological Entomology 27:702-712.

Nice, C. C., N. Anthony, G. Gelembiuk, D. Raterman, and R. Ffrench-Constant. 2005. Phylogeography and quaternary history of the butterfly genus *Lycaeides*: inferences from morphology and mitochondrial DNA.

Nice, C. C., and A. M. Shapiro. 2001. Patterns of morphological, biochemical, and molecular evolution in the *Oeneis chryxus* complex (Lepidoptera : Satyridae): A test of historical biogeographical hypotheses. Molecular Phylogenetics and Evolution 20:111-123.

Opler, P. 2003. Fixation of type locality for *Lycaena acmon* Westwood and characterization of the species and its distribution. The Taxonomic Report of the International Lepidoptera Survey 4.

Paulo, O. S., W. C. Jordan, M. W. Bruford, and R. A. Nichols. 2002. Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian Lizard. Molecular Ecology 11:809-819.

Pfenninger, M., and D. Posada. 2002. Phylogeographic history of the land snail *Candidula unifasciata* (Helicellinae, Stylommatophora): fragmentation, corridor migration, and secondary contact. Evolution 56:1776-1788.

Pielou, E. C. 1991. After the ice age: the return of life to glaciated North America. University of Chicago Press, Chicago.

Porter, A. H., and A. M. Shapiro. 1990. Lock-and-Key Hypothesis: Lack of mechanical isolation in a butterfly (Lepidoptera: Pieridae) hybrid zone. Entomological Society of America 83:107-114.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

Posada, D., and K. A. Crandall. 2001. Intraspecific gene genealogies: trees grafting into networks. Trends in Ecology and Evolution 16:37-45.

Posada, D., K. A. Crandall, and A. R. Templeton. 2000. GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. Molecular Ecology 9:487-488.

Rieseberg, L., S. Church, and C. Morgan. 2004. Integration of populations and differentiation of species. New Phytologist 161:59-69.

Robbins, R. K. 1982. How many butterfly species? News of the Lepidopterist's Society 1982:40-41.

Schemske, D. 2000. Understanding the origin of species. Evolution 54:1069-1073.

Schluter, D. 1998. Ecological causes of speciation. Pp. 114-129 *in* D. H. a. S. Berlocher, ed. Endless forms: species and speciation. Oxford University Press, New York.

Schneider, S., J.-M. Kueffer, D. Roessli, and L. Excoffier. 1997. Arlequin ver. 1.1.: A software for population genetic data analysis. University of Geneva, Switzerland.

Scott, J. 1986. The butterflies. Pp. 157-504. The butterflies of North America: a natural history and field guide. Stanford University Press, Standford, CA.

Shapiro, A. M., and A. H. Porter. 1989. The lock-and-key hypothesis - Evolutionary and biosystematic interpretation of insect genitalia. Annual Review of Entomology 34:231-245.

Shapiro, S. S., and M. B. Wilk. 1965. An analysis of variance test for normality (complete sample). Biometrika 52:591-611.

Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compliation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87:651-701.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264-279.

Smouse, P. E. 1998. To tree or not to tree. Molecular Ecology 7:399-412.

Sperling, F. A. H., and D. A. Hickey. 1994. Mitochondrial DNA sequence variation in the spruce budworm species complex (Choristoneura: Lepidoptera). Molecular Biology and Evolution 11:656-665.

Stanford, R. E., and P. A. Opler. 1993. Atlas of western USA butterflies including adjacent parts of Canada and Mexico. Denver and Fort Collins, Colorado.

Swofford, D. 2003. PAUP: phylogenetic analysis using parsimony. Illinios Natural History Survey, Champaign.

Tajima, F. 1987. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Templeton, A. R. 1998. Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7:381-397.

Templeton, A. R. 2004. Statistical phylogeography: methods of evaluating and minimizing inference errors. Molecular Ecology 13:789-809.

Templeton, A. R., E. Boerwinkle, and C. F. Sing. 1987. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and an analysis of alcohol dehydrogenase activity in *Drosophila*. Genetics 117:343-351.

Templeton, A. R., K. A. Crandall, and C. F. Sing. 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. Genetics 132:619-633.

Templeton, A. R., E. Routman, and C. A. Phillips. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics 140:767-782.

Tilden, J., and A. Smith. 1986. The western butterflies: species accounts. Pp. 36-274. A field guide to western butterflies. Houghton Mifflin Company, Boston.

Turner, T. F., J. C. Trexler, J. L. Harris, and J. L. Haynes. 2000. Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. Genetics 154:777-785.

Vane-Wright, R., and M. Boppre. 1993. Visual and chemical signaling in butterflies: functional and phylogenetic perspectives. Philosophical Transactions of the Royal Society of London 340:197-205.

Via, S. 2002. The ecological genetics of speciation. American Naturalist 159:S1-S7.

Wright, S. 1943. Isolation by distance. Genetics 28:114-138.

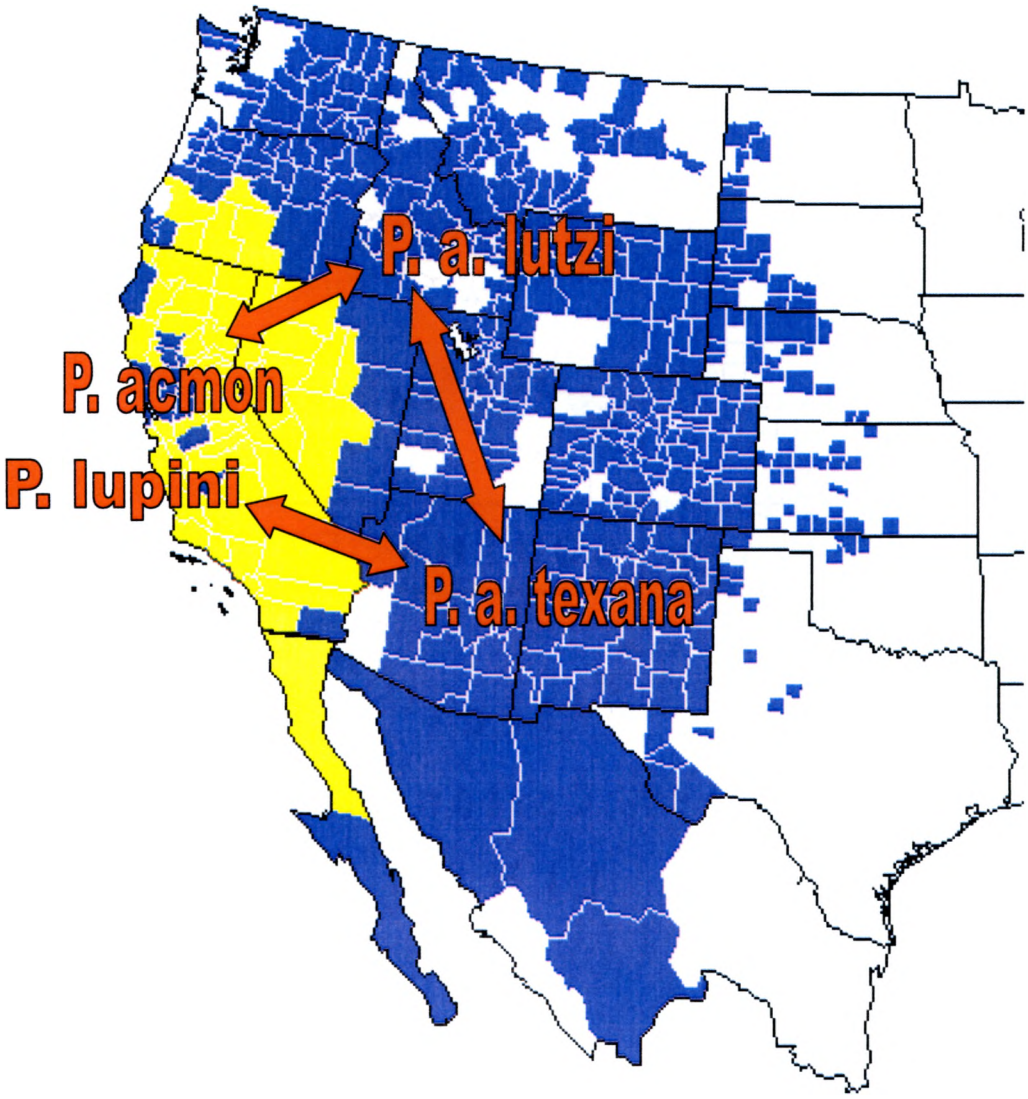

Figure 1. Distributions of *Plebejus lupini* and *P. acmon* in western North America (modified from information provided by Stanford and Opler 1993). Counties with records of *P. acmon* and *P. lupini* are represented in color. Counties where *P. acmon* only are found are represented in blue. Counties where *P. lupini* occurs with *P. acmon* are found are represented in yellow. The arrows between subspecies represent intergrading phenotypes and, in some cases, observed hybridization (Scott, 1986).

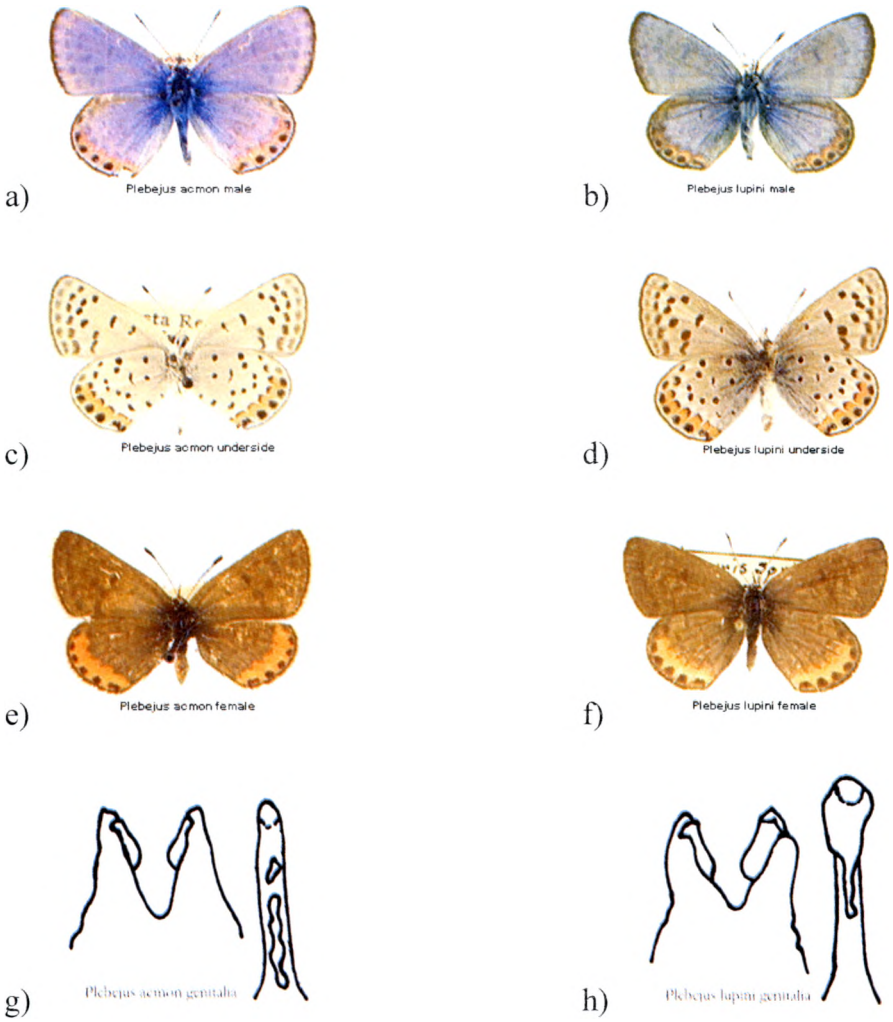

Figure 2. Comparison of the wing patterns and genitalia in *Plebejus lupini* and *P. acmon* (photos not to scale). The two nominal species are distinguished on the basis of these morphological differences. These specimens, photographed Robert W. Poole, are from the collection of the National Museum of Natural History, Smithsonian Institution. The photographs are the property of Neartica.com and used with their permission. Drawings of genitalia were reproduced from Scott, 1986.

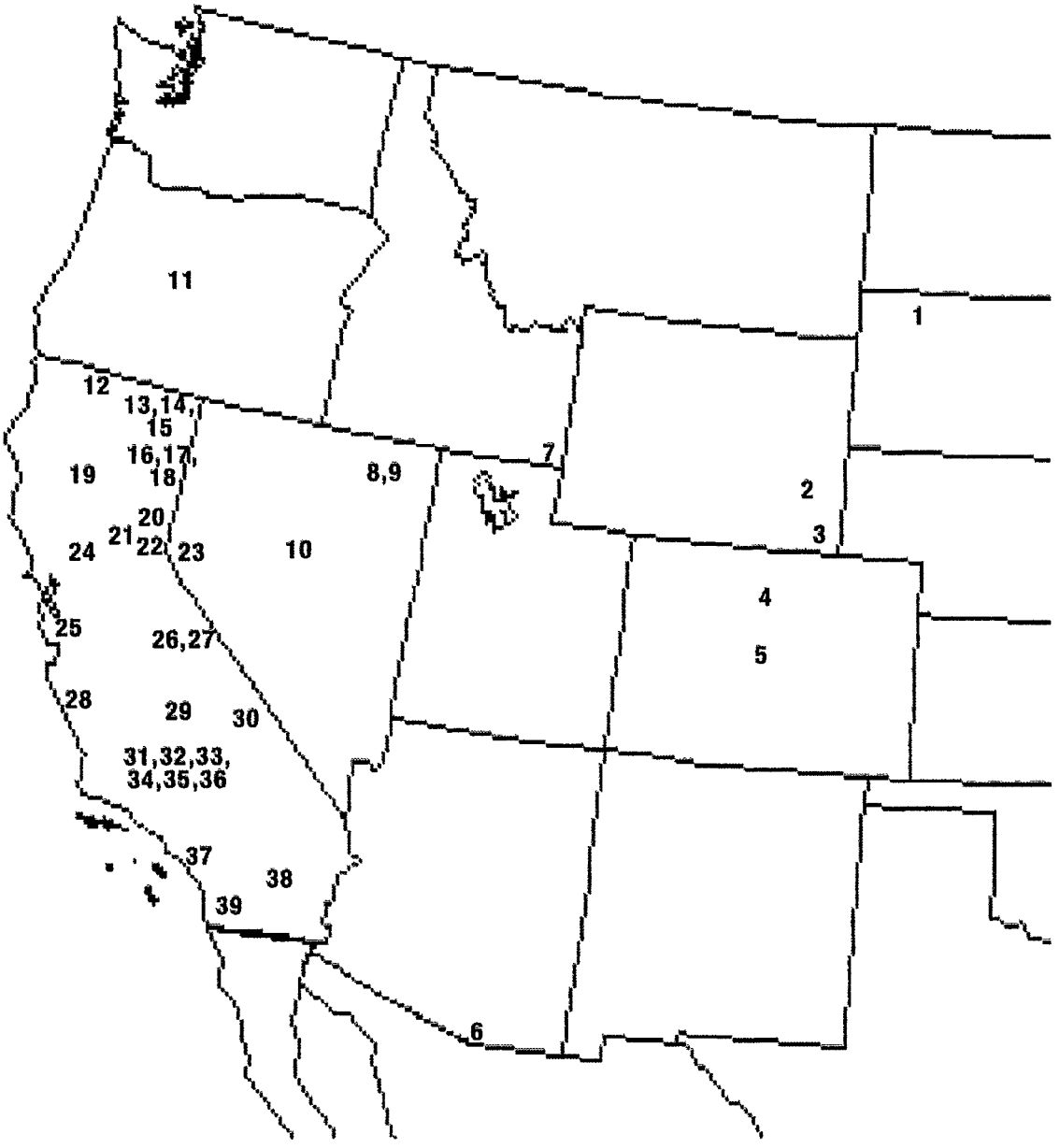

Figure 3. Population localities of *Plebejus* specimens analyzed in this study. See Table 1 for the locality information corresponding to the numbers on the map.

Table 1. Population data. The population numbers correspond with those from Figure 3. Those specimens with only a genus designation are female thus, unable to by identified by wing morphology. SAMOVA groups were determined using a spatial analysis of molecular variance for different groups sizes (k) (see text, Figure 11). h = haplotype diversity (Nei 1987).

Population	Nominal Taxonomic Designation	N	MtDNA Haplotypes (number of individuals)	SAMOVA (k=2)	SAMOVA (k=10)	h
1 Grand River National Grasslands, SD	*Plebejus lupini*	2	I(2)	2	10	0
2 Platte, WY	*Plebejus acmon*	1	S(1)	2	10	0
3 Chugwater, WY	*Plebejus*	1	I(1)	2	10	0
4 Dory Hill, CO	*Plebejus acmon*	3	N(3)	2	8	0
5 Horseshoe Mtn, CO	*Plebejus acmon*	2	B(2)	2	9	0
6 Lake Pena Blanca, AZ	*Plebejus lupini*	5	BB(3),JJ(1),KK(1)	2	9	0 7
7 Geneva, ID	*Plebejus lupini*	2	NN(1),Z(1)	2	2	1
8 Maggie Summit, NV	*Plebejus acmon*	2	EE(2)	2	10	0
9 Wildhorse Crossing, NV	*Plebejus lupini*	1	EE(1)	2	10	0
10 Austin Summit, NV	*Plebejus lupini*	1	P(1)	2	8	0
11 Dutchman Flat, OR	*Plebejus acmon*	3	O(3)	1	7	0
12 Shovel Creek, CA	*Plebejus acmon*	8	DD(1),FF(1),L(1),M(3),MM(1),V(1)	1	1	0 893
13 Alturas, CA	*Plebejus*	1	L(1)	1	1	0
14 Eagle Peak, CA	*Plebejus*	1	Z(1)	2	2	0
15 Cave Lake, CA	*Plebejus acmon*	1	M(1)	1	1	0
16 Hallelujah Junction, CA	*Plebejus*	2	R(2)	2	2	0
17 Lassen, CA	*Plebejus*	4	R(3),T(1)	2	2	0 378
18 North of Hallelujah Junction, CA	*Plebejus*	4	R(3),Z(1)	2	2	0 378
19 Government Flat, CA	*Plebejus acmon*	2	L(1),U(1)	1	1	1
20 Bassets, CA	*Plebejus lupini*	2	B(1),C(1)	2	3	1
21 Dutch Flat Powerhouse, CA	*Plebejus acmon*	3	M(1),V(2)	1	1	0 667
22 Iowa Hill Road, CA	*Plebejus acmon*	1	M(1)	1	1	0
23 Ash Canyon, NV	*Plebejus lupini*	1	C(1)	2	3	0
24 Stebbins Cold Canyon Reserve, CA	*Plebejus acmon*	1	V(1)	1	1	0
25 Palo Alto, CA	*Plebejus acmon*	4	L(2),M(2)	1	1	0 667
26 White Mountains, CA	*Plebejus acmon*	2	A(1),J(1)	2	4	1
27 Saddlebag Lake, CA	*Plebejus*	2	C(1),K(1)	2	4	1
28 Carmel Valley, CA	*Plebejus acmon*	8	AA(1),L(2),M(2),V(2),Y(1)	1	1	0 893
29 Kern Canyon, CA	*Plebejus*	3	L(1),M(1),W(1)	1	1	1
30 Silver Canyon, CA	*Plebejus*	1	M(1)	1	1	0
31 Mt Pinos, CA	*Plebejus acmon*	4	L(3),M(1)	1	1	0 5
32 Bald Mountain, CA	*Plebejus*	2	D(2)	2	5	0
33 Tehachapi Moutain Park, CA	*Plebejus acmon*	8	CC(1),L(3),LL(1),M(3)	1	1	0 786
34 Water Canyon Rd, Tehachapi Mtns, CA	*Plebejus acmon*	2	M(1),V(1)	1	1	1
35 East Onyx, CA	*Plebejus acmon*	2	F(2)	2	6	0
36 Wofford Heights, CA	*Plebejus neurona*	1	E(1)	2	5	0
37 Blue Jay Camp, CA	*Plebejus acmon*	11	GG(1),HH(1),II(1),L(5),M(2),V(1)	1	1	0 8
38 Lake Hemet, CA	*Plebejus acmon*	1	X(1)	1	1	0
39 Mt Laguna, CA	*Plebejus lupini*	1	Q(1)	2	6	0
40 West Onyx, CA	*Plebejus emigdionis*	2	G(1),H(1)			1
	Total=	108				μ = 0 367

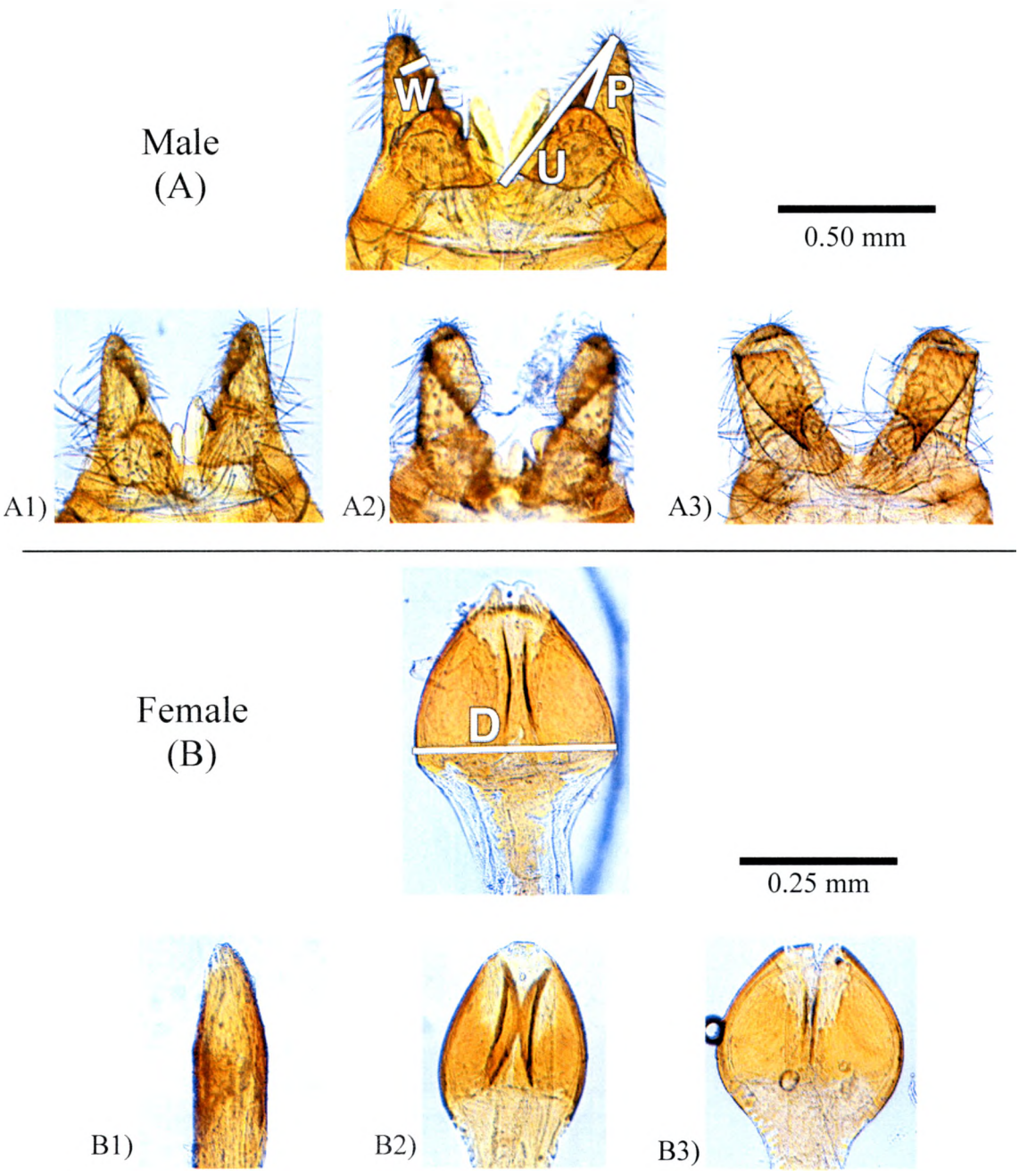

Figure 4. Measurements of *Plebejus* butterfly genitalia. Three measurements of each side of the male uncus (A) were taken: from the top of each uncus to the top of the corresponding peniculus (P), from the top of each uncus to the point where the right uncus and left uncus converge (U), and across the width of the pad-like structures on each uncus (W). (A1) Specimen from population 12. (A2) Specimen from population 11. (A3) Specimen from population 20. One measurement was taken at the greatest width of the female lamella (D). (B1) Specimen from population 37. (B2) Specimen from population 16. (B3) Specimen from population 20.

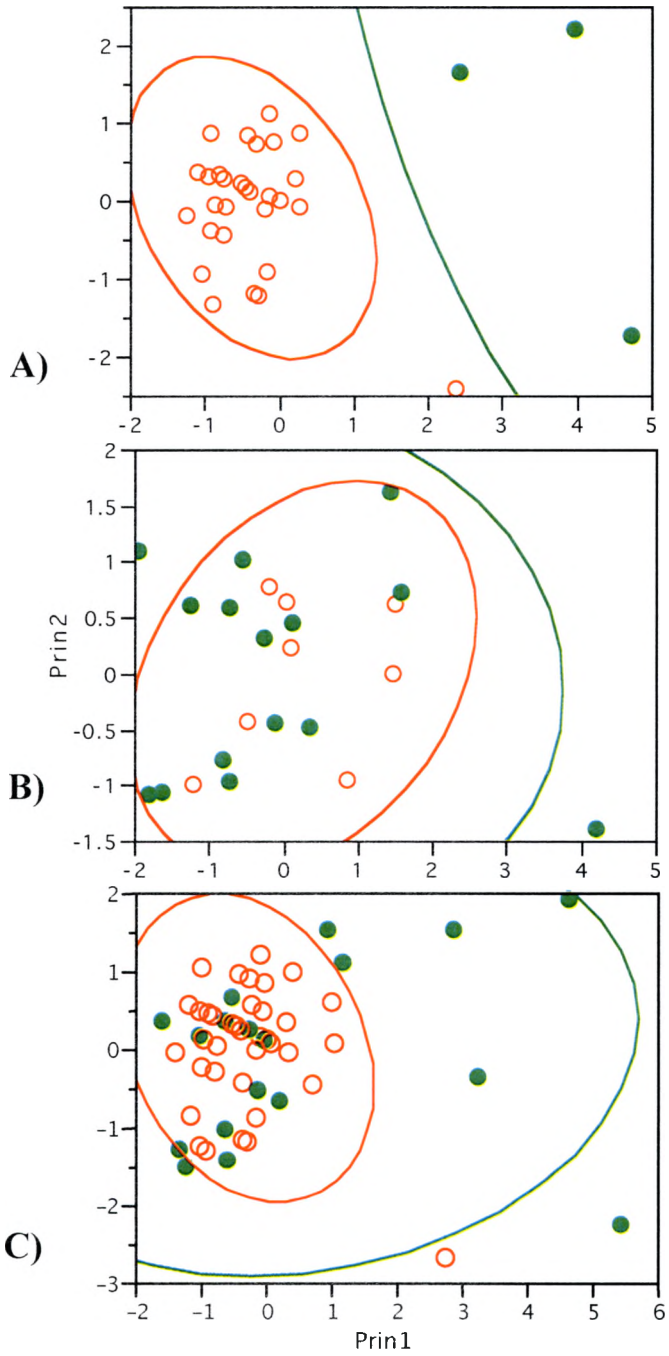

Figure 5. Ordinal plots with 95% confidence ellipse of the first two principal components in the analysis of three male genitalic measurements for the Acmon complex (A) in California (P = <0.0001), (B) outside of California (P = 0.8263), and (C) for the total area sampled (P = 0.0928). Wing morphology was used to identify individuals as *Plebejus acmon* (open circles) or *P. lupini* (closed circles). All measurements were positively correlated with principal component 1.

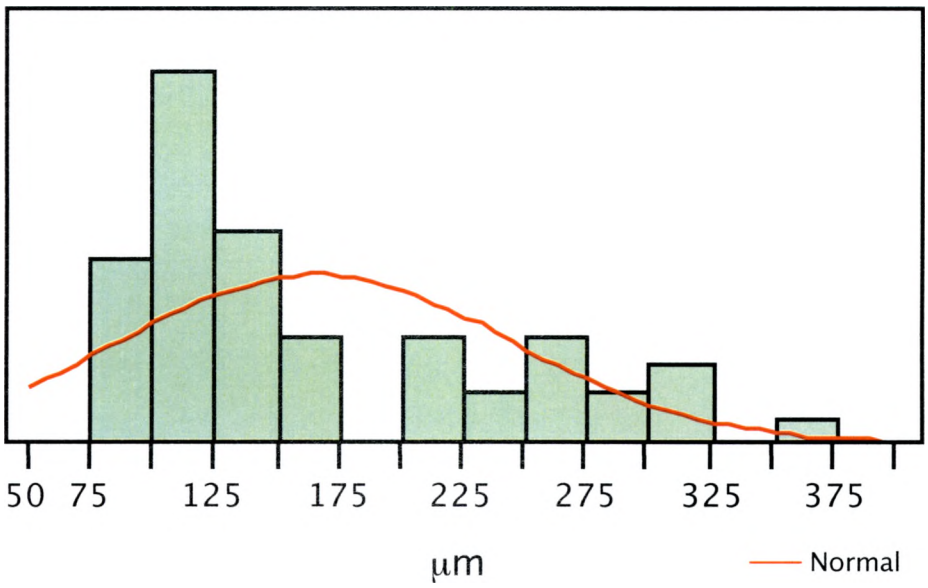

Figure 6. Distribution of female lamellae width fit with a normal curve of 49 individuals. Mean lamella width is 165.1 μm (SD = 78.0, SE = 11.1).

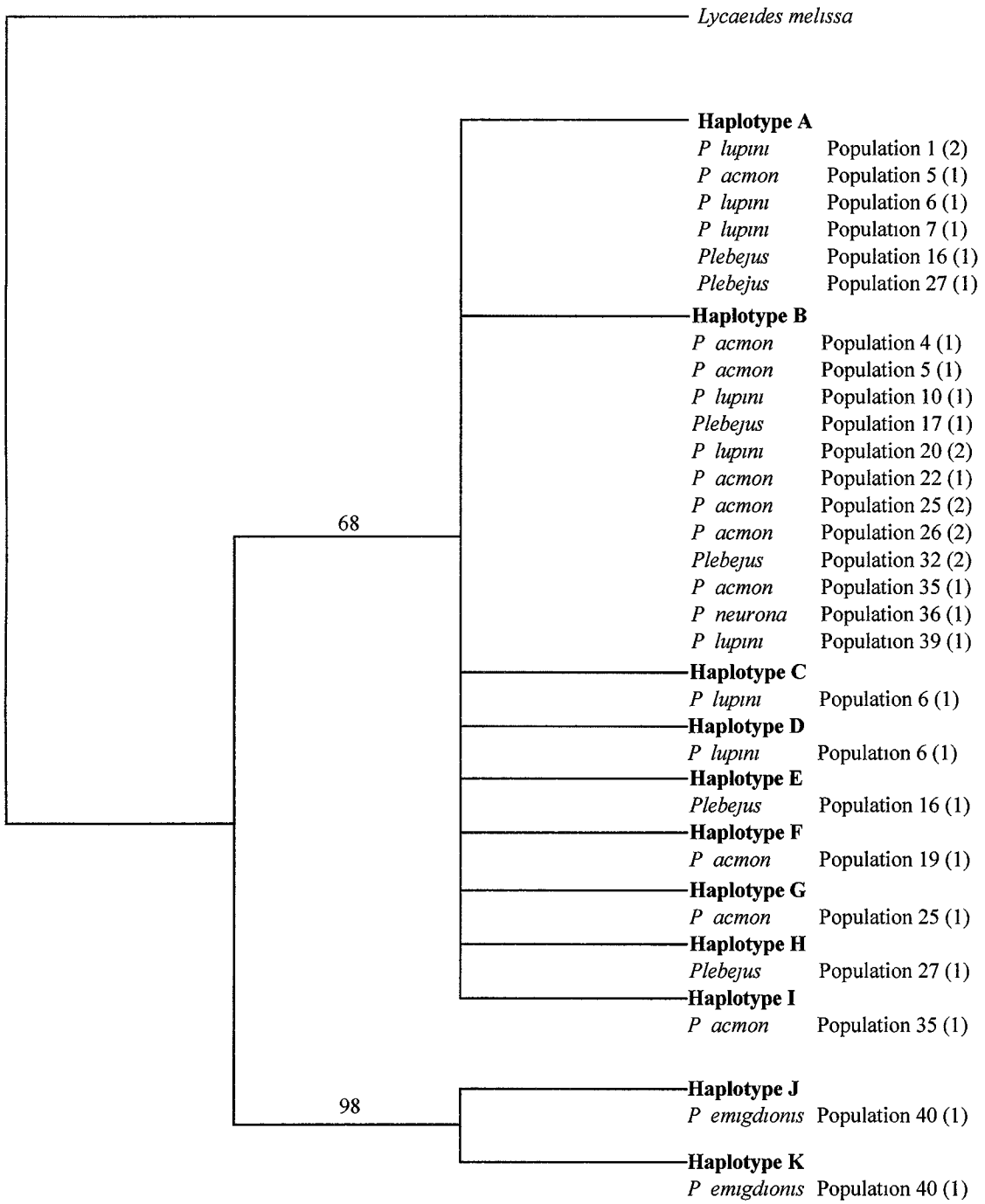

Figure 7. Neighbor-joining tree under HKY model of evolution for the nuclear *wingless* sequences of the Acmon complex. Bootstrap support indices greater than 50% are noted above the corresponding branch. Branches were collapsed when support was less than 50%. *Lycaeides melissa* was used as the outgroup. Nominal species identified only by genus are female and thus, unable to be identified by wing morphology. Population numbers corresponding to Table 1 are noted with sample sizes given in parentheses.

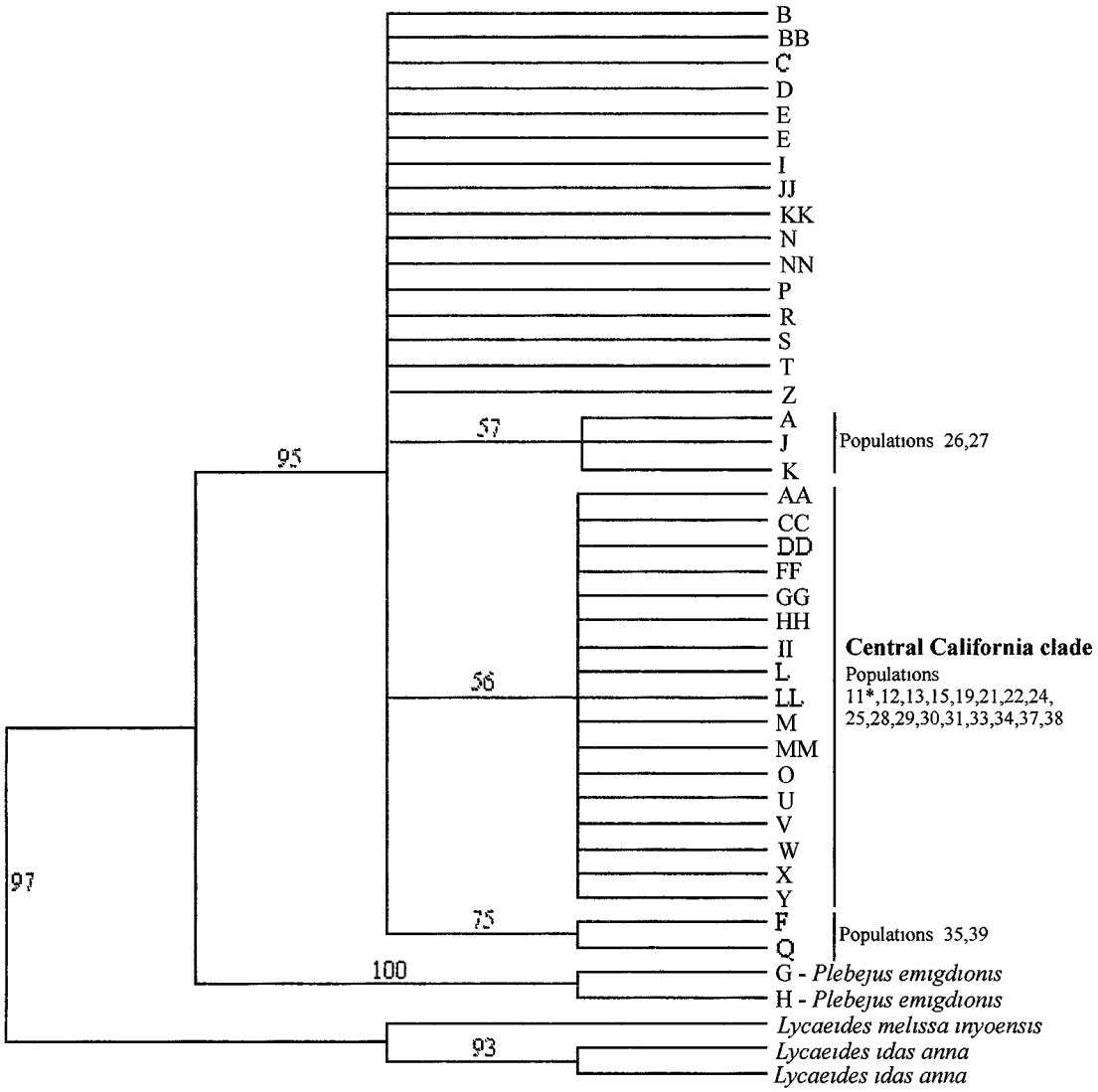

Figure 8. Neighbor-joining tree under HKY model of evolution for 38 mtCOI haplotypes found in the Acmon complex. Bootstrap support indices greater than 50% are noted above the corresponding branch. Branches with less than 50% bootstrap have been collapsed. *Population 11 is not included in the Central California clade by SAMOVA.

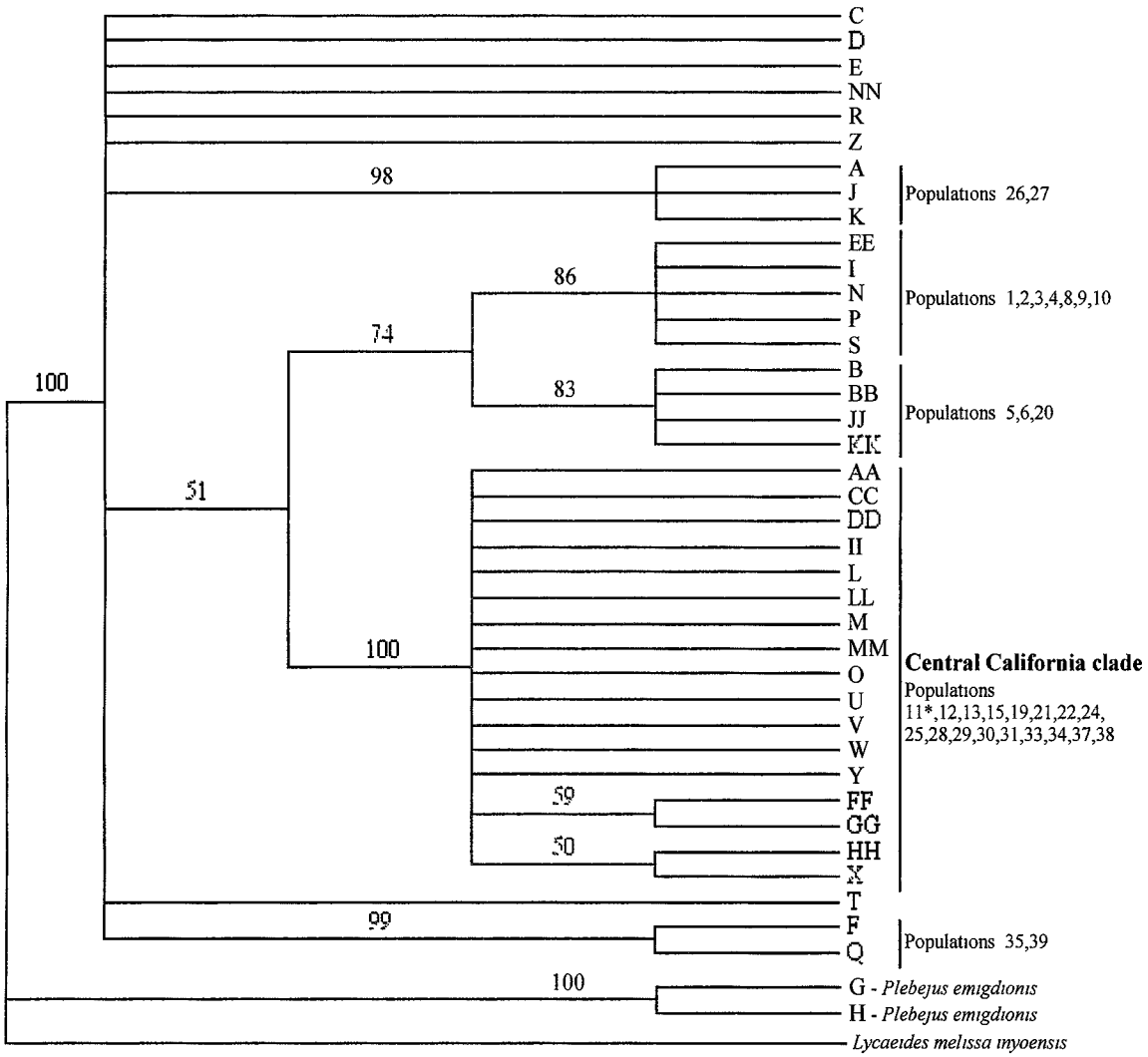

Figure 9. Bayesian posterior probability support indices for the 38 haplotypes in the *Plebejus* mitochondrial COI dataset. *Population 11 is not included in the Central California clade by SAMOVA.

Table 2. Nested analysis of molecular variance (AMOVA) for the 38 COI mtDNA haplotypes detected in the Acmon complex. Populations (described in Table 1) were grouped by current nominal species boundaries and by SAMOVA for k = 2 and k = 10 (see Table 1). d.f.: degrees of freedom, SSD: sums of squares, MSD: mean squared deviations, P-value: probability of obtaining a larger component of variance by chance under the null hypothesis that the variance component is zero (estimated by 1000 permutations).

Source of Variation	d f.	SSD	MSD	Variance	% of total	P-value
Nominal partitions						
Among species	1	55.64	55.64	2.04	48.28	0.00
Among populations/within species	24	126 89	5.29	1.61	38.04	0.00
Among individuals/within populations	53	30 68	0 58	0.58	13.68	<0.01
SAMOVA (k=2)						
Among groups	1	157.63	157.63	2.96	65.54	0.00
Among populations/within groups	35	111.35	3.18	0.94	20.73	0.00
Among individuals/within populations	69	42.82	0.62	0.62	13.74	0.00
SAMOVA (k=10)						
Among groups	9	243.23	27.03	3.35	81 83	0.00
Among populations/within groups	27	25.74	0.95	0.12	3.00	0 00
Among individuals/within populations	69	42.82	0 62	0.62	15 17	0.00

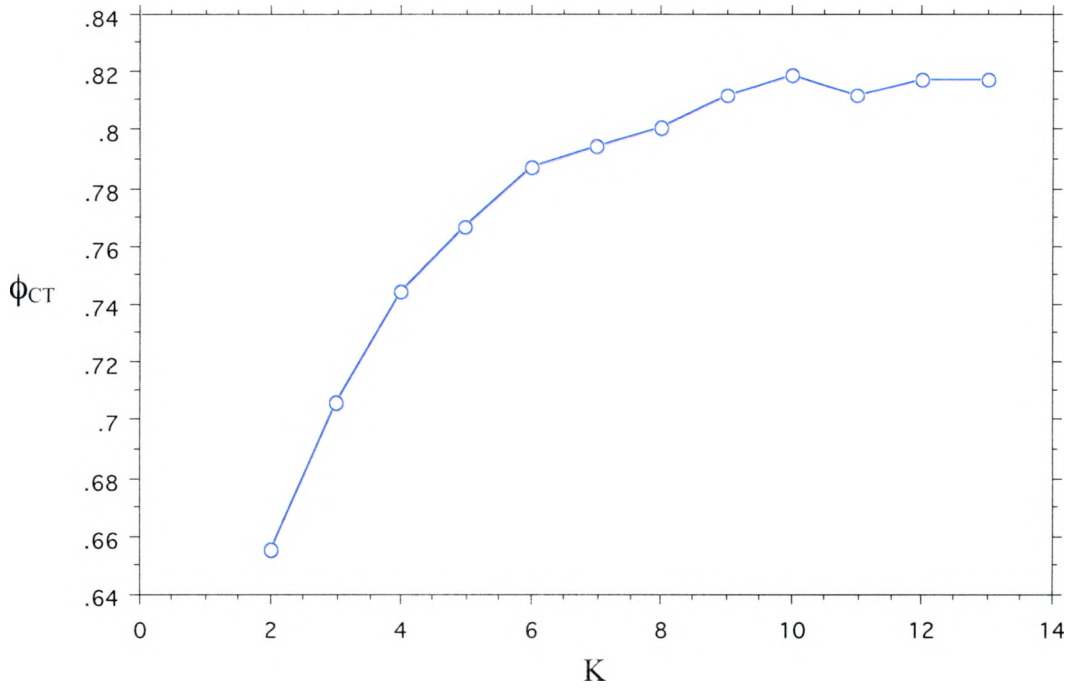

Figure 10. ϕ_{CT} values inferred by the SAMOVA algorithm for the 39 *Plebejus* populations divided into K groups. For each analysis, groups were defined with the constraint of genetic homogeneity and geographic proximity. Partitioning the populations into 10 groups maximized the allocation of variance among groups (81.83%).

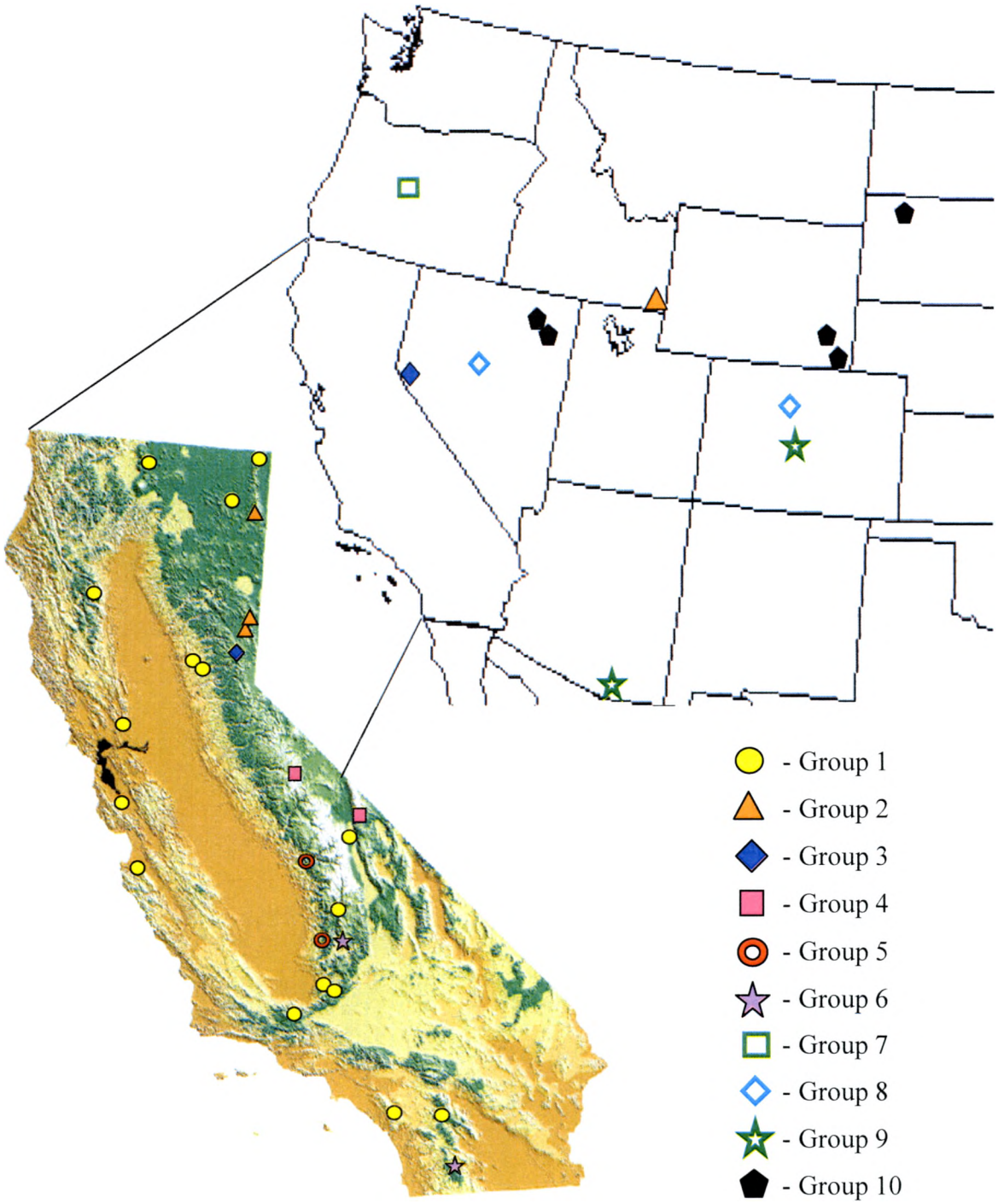

Figure 11. Groups partitioned by SAMOVA at k = 10, maximizing the percentage of the total variance explained by differences among groups. Groups are identified by their component populations in Table 1. Symbols differ for group number as indicated in the key.

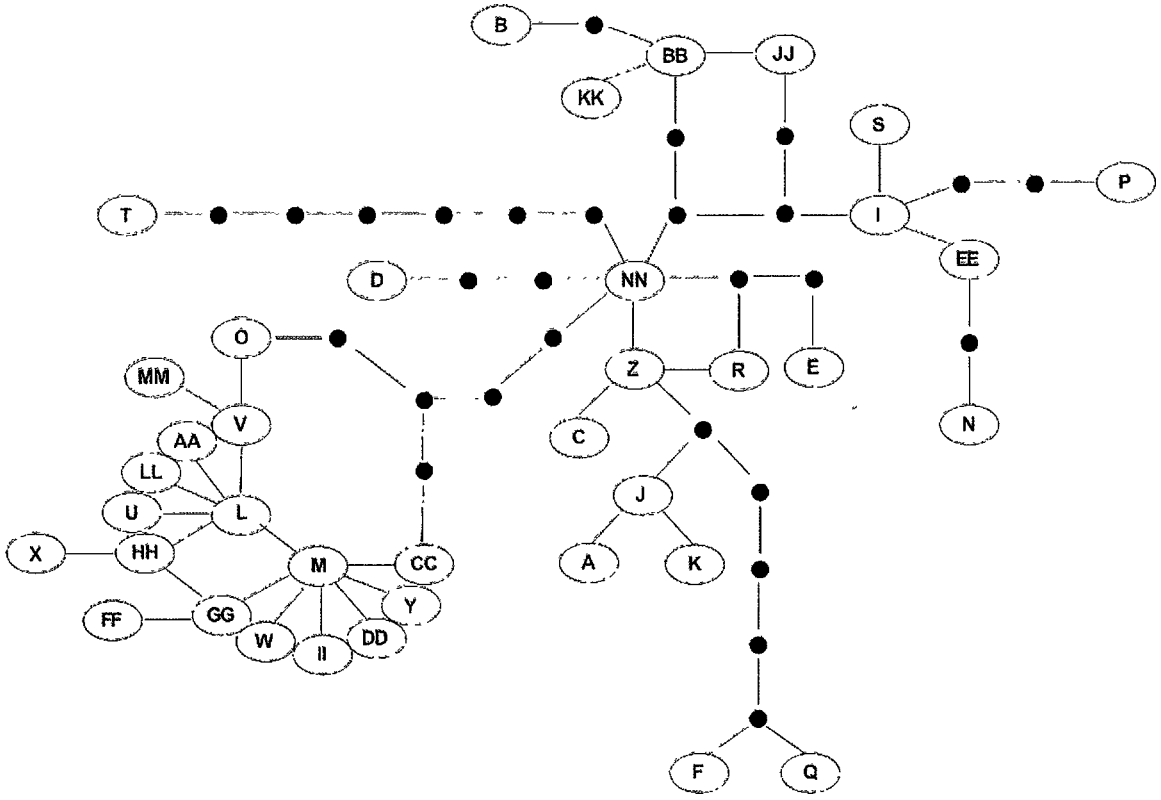

Figure 12. The estimated 95% haplotype network of the 38 mitochondrial COI haplotypes identified for the Acmon complex (*Plebejus* sp.) in western North America. Each line in the network represents one mutational change. The small dark circles represent hypothesized but undetected haplotypes missing from the network. The letters in the ovals correspond to those in Table 1.

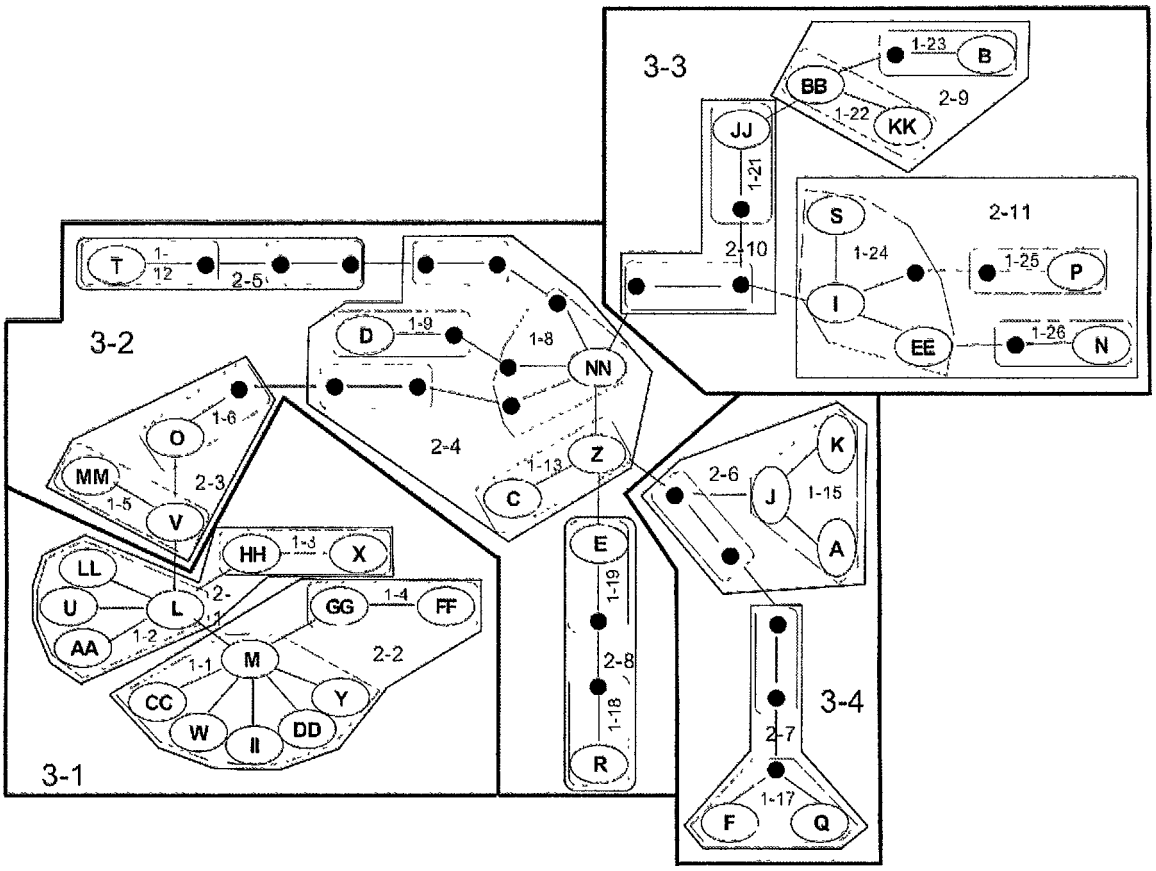

Figure 13. A nested statistical parsimony network of COI haplotypes in the Acmon complex (*Plebejus* sp). Ovals represent individual haplotypes connected by lines designating one mutational event. The small dark circles represent undetected haplotypes missing from the network. The letters in the ovals correspond to those in Table 1.

Haplotype					1-Step					2-Step					3-Step				
	D_C	D_N	I-T_C	I-T_N		D_C	D_N	I-T_C	I-T_N		D_C	D_N	I-T_C	I-T_N		D_C	D_N	I-T_C	I-T_N
AA	0	345																	
L	501	498	501	-55	1-2	496[L]	493[L]												
LL	0	428						460	253[L]	2-1	475	483							
U	0	886																	
HH	0	36	0	02	1-3	36	240[S]												
X	0	36																	
CC	0	547																	
DD	0	363																	
II	0	374	395	16	1-1	402	401						77	68	3-1	445	456		
M	395	396						79	38	2-2	397	414							
W	0	441																	
Y	0	179																	
FF	0	242	0	244	1-4	324	363												
GG	0	486																	
MM	0	315	295	-12	1-5	305[L]	305[L]												
V	295	303								2-3	205[S]	289							
O					1-6	0[S]	50[S]												
NN					1-8	0	363												
D					1-9	0	208[S]												
C	481	505	30	85				470[L]	302[L]	2-4	452	475[1]	159	50	3-2	454	442	-34	-53
Z	511	590			1-13	549	534												
T					1-12					2-5	0	794							
R					1-18	495	496	495	-183	2-6	189	200							
E					1-19	0	679												
BB					1-22	0[S]	266[S]	-	-	2-11	491	666							
KK								693[S]	450[S]										
B					1-23	693[L]	716[L]												
JJ					1-21					2-10	0	620	-618	-60	3-3	682[1]	646[1]		
I	951[L]	966[L]																	
S	0	486	800[L]	350	1-24	770	767												
EE	648	707						770[1]	224	2-9	700	689							
P					1-26	0	687												
N					1-25	0[S]	494												
A	0	69																	
J	0	69	0	-104	1-15					2-7	111[S]	273							
K	0	277											-68	67	3-1	236	332		
F	0	134	-	-	1-17					2 8	179	206							
Q	0	269																	

Figure 14. Results of the nested clade analysis of the Acmon complex (*Plebejus* sp.) COI mtDNA haplotypes. The nesting design with haplotypes and clade designations are given in Figure 13. D_C and D_N refer to average clade and nested clade distances, respectively. A superscript S signifies the distance measure as significantly small at the 5% level and a superscript L signifies the distance as significantly large. Shaded rows indicate interior clade status while rows not shaded indicate tip clade status.

Table 3. Historical events inferred from nested clade analysis in the Acmon complex (*Plebejus* sp) (see Figure 14) following the upgraded key of Templeton (1998). IBD is restricted gene flow with isolation by distance and CRE is contiguous range expansion.

Clade	Chain Event
2-1	1-2-3-4No → IBD
2-4	1-2-11-17-4No → IBD
3-2	1-2-3-4No → IBD
3-4	1-19-20-2-11-12No → CRE
Total cladogram	1-2-11-12No → CRE

APPENDIX

Reference information for individuals of the genus *Plebejus* collected for this study. All individuals are identified by a unique CCN number. If the mtCOI gene was successfully sequenced, haplotype is noted.

CCN #	Population #	mtDNA Haplotype	SAMOVA (k=2)	SAMOVA (k=10)	Species by wing morph	Sex	Collector
1	5	B	2	9	*acmon*	M	Opler
2	5	B	2	9	*acmon*	M	Opler
3	32	D	2	5		F	Opler
4	32	D	2	5		F	Opler
5	20	C	2	3	*lupini*	M	Opler
6	20	B	2	3	*lupini*	M	Opler
7	26	J	2	4	*acmon*	M	Opler
8	26	A	2	4	*acmon*		Opler
9	27	K	2	4		F	Opler
10	27	C	2	4		F	Opler
11	36	E	2	5	*neurona*	M	Opler
12	35	F	2	6	*acmon*	M	Opler
13	35	F	2	6	*acmon*	F	Opler
14	40	G			*emigdionis*	F	Opler
15	40	H			*emigdionis*	F	Opler
16	1	I	2	10	*lupini*	F	Opler
17	1	I	2	10	*lupini*	M	Opler
18	10	P	2	8	*lupini*	M	Opler
19	39	Q	2	6	*lupini*	M	M. Klein
20	16	R	2	2		F	Opler
21	16	R	2	2		F	Opler
22	25	M	1	1	*acmon*	M	Opler
23	25	L	1	1	*acmon*	M	Opler
24	25	M	1	1	*acmon*	F	Opler
25	25	L	1	1	*acmon*	M	Opler
26	4	N	2	8	*acmon*	F	Opler
27	4	N	2	8	*acmon*	M	Opler
28	4	N	2	8	*acmon*	M	Opler
29	11	O	1	7	*acmon*	M	A. Warren
30	11	O	1	7	*acmon*	M	A. Warren
31	11	O	1	7	*lupini*	M	A. Warren
36	3	I	2	10		F	Opler
37	19	L	1	1	*acmon*	F	JCO
38	22	M	1	1	*acmon*	M	JCO
39	13	L	1	1		F	JCO
40	2	S	2	10	*acmon*	M	Opler
41	17	R	2	2		F	Opler
42	17	R	2	2		F	Opler
43	17	T	2	2		F	Opler
44	17	R	2	2		F	Opler
45	19	U	1	1	*acmon*	M	JCO
193	8	EE	2	10	*acmon*	M	JCO
194	18	Z	2	2		F	Opler
195	18	R	2	2		F	Opler

CCN #	Population #	mtDNA Haplotype	SAMOVA (k=2)	SAMOVA (k=10)	Species by wing morph	Sex	Collector
196	18	R	2	2		F	Opler
197	18	R	2	2		F	Opler
201	8	EE	2	10	*lupini*	M	JCO
202	37	L	1	1		F	JCO
203	37	L	1	1	*acmon*	M	JCO
204	37		1	1		F	JCO
205	37	M	1	1		F	JCO
206	37	L	1	1	*acmon*	M	JCO
207	21	V	1	1	*acmon*	F	JCO
208	21	M	1	1	*acmon*	M	JCO
209	21	V	1	1	*acmon*	M	JCO
210	37	V	1	1		F	JCO
211	37	GG	1	1	*acmon*	M	JCO
212	37	L	1	1	*acmon*	M	JCO
213	37	M	1	1		F	JCO
214	37	L	1	1		F	JCO
215	37	HH	1	1		F	JCO
216	37	II	1	1		F	JCO
217	37		1	1		F	JCO
218	9	EE	2	10	*lupini*	M	JCO
219	30	M	1	1		F	KLP
220	34	V	1	1	*acmon*	F	JCO, KLP
221	34	M	1	1	*acmon*	M	JCO, KLP
222	34		1	1	*acmon*	F	JCO, KLP
223	7	NN	2	2	*lupini*	M	Opler
224	7		2	2	*lupini*	M	Opler
225	7	Z	2	2	*lupini*	M	Opler
226	6	BB	2	9	*lupini*	M	JCO
227	6	JJ	2	9	*lupini*	M	JCO
228	31	L	1	1	*acmon*	M	K. Davenport
229	31	L	1	1	*acmon*	M	K. Davenport
230	31	L	1	1	*acmon*	M	K. Davenport
231	31	M	1	1	*acmon*	M	K. Davenport
232	29	L	1	1		F	K. Davenport
233	29	W	1	1		F	K. Davenport
234	29	M	1	1		F	K. Davenport
235	6	KK	2	9	*lupini*	M	JCO
236	38	X	1	1	*acmon*	M	JCO
237	28	V	1	1	*acmon*	F	JCO, KLP
238	28	M	1	1	*acmon*	F	JCO, KLP
239	28	L	1	1	*acmon*	F	JCO, KLP
240	28	Y	1	1	*acmon*	M	JCO, KLP
301	28	M	1	1	*acmon*	M	JCO, KLP
302	28	AA	1	1	*acmon*	M	JCO, KLP

CCN #	Population #	mtDNA Haplotype	SAMOVA (k=2)	SAMOVA (k=10)	Species by wing morph	Sex	Collector
303	28	L	1	1	*acmon*	F	JCO, KLP
304	6	BB	2	9	*lupini*	M	KLP
305	6	BB	2	9	*lupini*	F	KLP
306	23	C	2	3	*lupini*	M	KLP
307	33	M	1	1	*acmon*	F	JCO, KLP
308	33	LL	1	1	*acmon*	F	JCO, KLP
309	33	M	1	1	*acmon*	F	JCO, KLP
310	33	L	1	1	*acmon*	F	JCO, KLP
311	33	M	1	1	*acmon*	M	JCO, KLP
312	33	CC	1	1	*acmon*	F	JCO, KLP
313	33	L	1	1	*acmon*	F	JCO, KLP
314	33	L	1	1	*acmon*	F	JCO, KLP
315	12	FF	1	1	*acmon*	F	JCO, KLP
316	12	M	1	1	*acmon*	M	JCO, KLP
317	12	DD	1	1	*acmon*	F	JCO, KLP
319	12	M	1	1	*acmon*	F	JCO, KLP
320	12	V	1	1	*acmon*	M	JCO, KLP
321	12	L	1	1	*acmon*	M	JCO, KLP
322	12	MM	1	1	*acmon*	M	JCO, KLP
323	12	M	1	1	*acmon*	M	JCO, KLP
324	24	V	1	1	*acmon*	M	CCN
325	28	V	1	1	*acmon*	F	JCO
326	14	Z	2	2		F	CCN
327	15	M	1	1	*acmon*	M	CCN

CCN #	County	State	Locality	Latitude	Longitude
1	Park	CO	Horseshoe Mtn	37°36'00"N	106°26'30"W
2	Park	CO	Horseshoe Mtn	37°36'00"N	106°26'30"W
3	Kern	CA	Bald Mtn	34°49'00"N	119°56'30"W
4	Kern	CA	Bald Mtn	34°49'00"N	119°56'30"W
5	Sierra	CA	Bassets	39°34'00"N	120°33'30"W
6	Sierra	CA	Bassets	39°34'00"N	120°33'30"W
7	Mono	CA	White Mtns	37°40'30"N	118 18 30"W
8	Mono	CA	White Mtns	37°40'30"N	118 18 30"W
9	Mono	CA	Saddlebag Lake	37°56'30"N	119°13'30"W
10	Mono	CA	Saddlebag Lake	37°56'30"N	119°13'30"W
11	Kern	CA	Wofford Heights	35 42 21 N	118 27 18 W
12	Kern	CA	E. Onyx	35 41 24 N	118 13 8 W
13	Kern	CA	E. Onyx	35 41 24 N	118 13 8 W
14	Kern	CA	W. Onyx	35 41 24 N	118 13 8 W
15	Kern	CA	W. Onyx	35 41 24 N	118 13 8 W
16	Perkins	SD	Grand River Natl Grasslands	45°56'30"N	102°10'30"W
17	Perkins	SD	Grand River Natl Grasslands	45°56'30"N	102°10'30"W
18	Lander	NV	Austin Summit	39°26'30"N	117 03 15"W
19	San Diego	CA	Mt. Laguna	32°48'00"N	116°25'30"W
20	Lassen	CA	Halleluia Junction	39°46'30"N	120°2'20"W
21	Lassen	CA	Halleluia Junction	39°46'30"N	120°2'20"W
22	Santa Clara	CA	Palo Alto	37°26'30"N	122°10'30"W
23	Santa Clara	CA	Palo Alto	37°26'30"N	122°10'30"W
24	Santa Clara	CA	Palo Alto	37°26'30"N	122°10'30"W
25	Santa Clara	CA	Palo Alto	37°26'30"N	122°10'30"W
26	Gilpin	CO	Dory Hill	39°49'00"N	105°15'00"W
27	Gilpin	CO	Dory Hill	39°49'00"N	105°15'00"W
28	Gilpin	CO	Dory Hill	39°49'00"N	105°15'00"W
29	Deschutes	OR	Dutchman Flat	44°33'30"N	120°40'30"W
30	Deschutes	OR	Dutchman Flat	44°33'30"N	120°40'30"W
31	Deschutes	OR	Dutchman Flat	44°33'30"N	120°40'30"W
36	Laramie	WY	5.5 mi. S of Chugwater	41°45'00"N	104°30'00"W
37	Tehama	CA	near Govt Flat	40° 01' 10"N	122° 21' 40"W
38	Placer	CA	Iowa Hill Rd	39°03'30"N	120°52'30"W
39	Modoc	CA	SE of Alturas	41°25'30"N	120°37'30"W
40	Platte	WY		42 36 37 N	104 57 00 W
41	Lassen	CA		39°46'30"N	120°2'20"W
42	Lassen	CA		39°46'30"N	120°2'20"W
43	Lassen	CA		39°46'30"N	120°2'20"W
44	Lassen	CA		39°46'30"N	120°2'20"W
45	Tehama	CA	near Govt Flat	40° 01' 10"N	122° 21' 40"W
193	Elko	NV	1 mi NE Maggie Summit	41 41 03 N	116 02 29 W
194	Lassen	CA	4mi N Hallelujah Junction	39°46'30"N	120°2'20"W
195	Lassen	CA	4mi N Hallelujah Junction	39°46'30"N	120°2'20"W

CCN #	County	State	Locality	Latitude	Longitude
196	Lassen	CA	4mi N Hallelujah Junction	39°46'30"N	120°2'20"W
197	Lassen	CA	4mi N Hallelujah Junction	39°46'30"N	120°2'20"W
201	Elko	NV	1 mi N Maggie Summit	41 41 03 N	116 02 29 W
202	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
203	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
204	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
205	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
206	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
207	Nevada	CA	Dutch Flat Powerhouse	39°011'30"N	120°48'30"W
208	Nevada	CA	Dutch Flat Powerhouse	39°011'30"N	120°48'30"W
209	Nevada	CA	Dutch Flat Powerhouse	39°011'30"N	120°48'30"W
210	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
211	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
212	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
213	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
214	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
215	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
216	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
217	Orange	CA	Blue Jay Camp, Santa Ana Mtns	33 39 11 N	117 27 06 W
218	Elko	NV	Wildhorse Xing, Owyhee River	41°41'30"N	115°55'30"W
219	Inyo	CA	Silver Canyon	37°24'16"N	118 18 46"W
220	Kern	CA	Water Canyon Rd, Tehachapi Mtns	35 04 49 N	118 29 46 W
221	Kern	CA	Water Canyon Rd, Tehachapi Mtns	35 04 49 N	118 29 46 W
222	Kern	CA	Water Canyon Rd, Tehachapi Mtns	35 04 49 N	118 29 46 W
223	Bear River	ID	3 mi N Geneva, 2.3 mi E of type	42°18'30"N	111°03'15"W
224	Bear River	ID	3 mi N Geneva, 2.3 mi E of type	42°18'30"N	111°03'15"W
225	Bear River	ID	3 mi N Geneva, 2.3 mi E of type	42°18'30"N	111°03'15"W
226	Santa Cruz	AZ	Lake Pena Blanca	31°25'30"N	111°03'15"W
227	Santa Cruz	AZ	Lake Pena Blanca	31°25'30"N	111°03'15"W
228	Kern	CA	Mt. Pinos	34° 49' N	119° 9' W
229	Kern	CA	Mt. Pinos	34° 49' N	119° 9' W
230	Kern	CA	Mt. Pinos	34° 49' N	119° 9' W
231	Kern	CA	Mt. Pinos	34° 49' N	119° 9' W
232	Tulare	CA	Kern Canyon, Corral Creek	36°18'30"N	118 26 30"W
233	Tulare	CA	Kern Canyon, Corral Creek	36°18'30"N	118 26 30"W
234	Tulare	CA	Kern Canyon, Corral Creek	36°18'30"N	118 26 30"W
235	Santa Cruz	AZ	Lake Pena Blanca	31°25'30"N	111°03'15"W
236	Riverside	CA	Lake Hemet	33°41'30"N	117°00'00"W
237	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
238	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
239	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
240	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
301	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
302	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W

CCN #	County	State	Locality	Latitude	Longitude
303	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
304	Santa Cruz	AZ	Lake Pena Blanca	31°25'30"N	111°03'15"W
305	Santa Cruz	AZ	Lake Pena Blanca	31°25'30"N	111°03'15"W
306	Carson City	NV	Ash Canyon, W. of Carson City	39 1762 N	119 8106 W
307	Kern	CA	Tehachapi Mtn Park	35 04 18 N	118 29 00 W
308	Kern	CA	Tehachapı Mtn Park	35 04 18 N	118 29 00 W
309	Kern	CA	Tehachapı Mtn Park	35 04 18 N	118 29 00 W
310	Kern	CA	Tehachapi Mtn Park	35 04 18 N	118 29 00 W
311	Kern	CA	Tehachapi Mtn Park	35 04 18 N	118 29 00 W
312	Kern	CA	Tehachapi Mtn Park	35 04 18 N	118 29 00 W
313	Kern	CA	Tehachapı Mtn Park	35 04 18 N	118 29 00 W
314	Kern	CA	Tehachapi Mtn Park	35 04 18 N	118 29 00 W
315	Siskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
316	Sıskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
317	Sıskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
319	Sıskıyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
320	Siskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
321	Siskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
322	Sıskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
323	Siskiyou	CA	Shovel Creek	41°58'19"N	122°12'02"W
324	Yolo	CA	Stebbins Cold Canyon Reserve	34° 06' N	118 39 00 W
325	Monterey	CA	0.2 mi E jct G16xTassajara Rd,	36°26'30"N	121°41'30"W
326	Modoc	CA	Warners EP	41°19'00"N	120°10'30"W
327	Modoc	CA	Cave Lake	41°58'49"N	120°12'29"W

VITA

Maurine Spencer was born in Atlanta, Georgia, on December 16, 1979 (Birthday presents may be shipped to the permanent address below). She is the daughter of Larry and Vicki McWilliams. After graduating from Colleyville Heritage High School in Colleyville, Texas, in May of 1998, she attended Abilene Christian University. There she earned a Bachelor of Science in Biology while pitching for the varsity softball team. Don't ask to bat against her, you'll only embarrass yourself. In August of 2002, she enrolled as a graduate student in the Biology Department at Texas State University-San Marcos. While at Texas State, she worked as an instructional assistant for Functional Biology and Genetics. A conscientious objector to summer school, Maurine took all summers off during this time. She and her husband, Clay Spencer, spent one summer managing rural private property for wildlife. The following summer Clay and Maurine led a dental mission trip to Togo, West Africa, for 6 weeks, where she successfully extracted her first wisdom tooth.

Permanent address: 4102 Beckley Court

Colleyville, Texas 76034

This thesis was typed by Maurine Spencer.