To the Graduate Council:

I am submitting herewith a dissertation written by Nadia Antoine Ayoub entitled "The role of history in the diversification of the funnel-web spiders, *Agelenopsis aperta,* and its congeners." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Ecology and Evolutionary Biology.

Susan E. Riechert
Major Professor

We have read this dissertation and
recommend its acceptance:

Christine R. Boake

Gary F. McCracken

Randall L. Small

Accepted for the Council:

Anne Mayhew
Vice Chancellor and Dean of Graduate Studies

(Original signatures are on file with official student records.)

The role of history in the diversification of the funnel-web spiders, *Agelenopsis aperta*, and its congeners.

A Dissertation
Presented for
The Doctor of Philosophy
Degree
University of Tennessee, Knoxville

Nadia Antoine Ayoub
December 2004

Acknowledgements

Numerous people helped to make my dissertation work possible. The guidance of my advisor, Susan Riechert, was essential; she introduced me to the genus *Agelenopsis* and her years of work on *A. aperta* set an ideal background to the work presented here. I am grateful to each of my committee members: Gary McCracken for allowing me to use lab space, Randall Small for an introduction to phylogenetics (and continual advice in this and other arenas), and Christine Boake for always being available to give advice in many aspects of graduate life. My fellow graduate students kindly let me bounce ideas off them for the last five years. I am grateful to the department of Ecology and Evolutionary Biology for monetary support in the form of a graduate teaching assistantship and summer research grants. Funding for stipend and research also came from an NSF graduate fellowship. Finally I want to thank all of my family for years of support (many more than spent in graduate school) and my husband, Charles Winder, for always believing in me even when I didn't believe in myself.

Abstract

Identifying the influence of history on the diversity of life is an important goal of evolutionary biology, whether it is an historical event that is extrinsic to the organism such as past climate change or an intrinsic event such as a single switch in behavior. Phylogenetic trees offer an indirect record of the history of diversification in a group of organisms and can be used in conjunction with other information to deduce the multiple roles of history affecting a particular group. Intra- and inter-specific phylogenetic relationships in the spider genus *Agelenopsis* Geibel (Araneae: Agelenidae) were analyzed using sequence data from two mitochondrial genes, cytochrome oxidase I (COI) and 16S ribosomal RNA for 12 out of 13 described species as well as a potential new species. These data were used to address a number of questions regarding biogeography, speciation, and evolution of habitat-use. Concerning biogeography, patterns and levels of sequence variation were consistent with the hypothesis that Pleistocene climate change affected geographic distribution and evolutionary diversification patterns within *A. aperta*. Speciation patterns were uncovered regarding species monophyly, geography of speciation, and the rejection of a putative hybrid species. Finally, a potential interaction of history and natural selection was identified when looking at the evolution of habitat use within and among *Agelenopsis* species. It was found that riparian living arose once in the genus with a single switch to arid living. Subsequently, three species of *Agelenopsis* displayed multiple switches back to riparian living. The findings of this dissertation point to a

multitude of evolutionary questions that should be addressed and suitability of *Agelenopsis* as a model system in evolutionary biology.

Table of Contents

List of Tables

List of Figures

Introduction

Identifying historical contributions to the evolution of biodiversity can be difficult because direct records of history such as fossil evidence do not exist for all groups or do not provide sufficient information to address particular questions. However, modern organisms carry a signature of history in their underlying makeup, including molecular, physiological, morphological, and behavioral characters. These characters can be used to develop phylogenetic trees of historical relationships among extant organisms. Such reconstruction can help us to understand the role of history in the diversification of life.

Among the numerous historical influences on diversity, geologic and climatic events are perhaps the most influential. Such events can create barriers to gene flow among populations, allowing those populations to evolve along independent trajectories (Mayr 1963). They also have important consequences on organisms' modern day geographic distributions, and may in large part explain the known diversity of distribution patterns. In combination with information on the geologic record, and past climate change, phylogenetic trees can potentially be used to identify the role of geologic and climate change in the diversification of life (Rosen 1978, Avise 1994).

Historical contingency, i.e. modern organisms display certain characters because their ancestors evolved these characters (Gould 1989), is often offered as an alternative explanation to the effects of natural selection. Although phylogenetic trees cannot directly measure selection, they can be used to evaluate whether similar characters found in different organisms evolved only once in a common ancestor

(historical contingency) or evolved multiple times independently (evidence for natural selection).

This dissertation takes a phylogenetic approach to address the history of diversification of a group of North American funnel-web spiders, *Agelenopsis* Geibel, with special attention to the desert spider, *A. aperta* Gertsch. *Agelenopsis aperta* is widely distributed throughout the deserts of southwestern North America but also uses riparian habitats within the arid landscape. Genetically-based behavioral differentiation has been demonstrated for populations occupying the two different habitat types (see review in Riechert 1999). The combined characteristics of large range size and habitat based variation make this species ideal for an analysis of the role of history in diversification with particular reference to discerning potential effects of climate, geology and historical contingency.

To meet the goals of this study, collections of *A. aperta* were made from populations inhabiting both arid and riparian habitats throughout its geographic range. Representatives of most other *Agelenopsis* species were also collected. Both intraspecific and interspecific relationships were reconstructed using mitochondrial DNA sequence variation. A hierarchical approach that spans the population-species interface is used to provide the best resolution for examining evolutionary processes that are thought to occur at a population level. This dissertation is divided into three parts: 1) biogeographic patterns within *A. aperta*, 2) speciation patterns within *Agelenopsis*, and 3) the evolution of habitat-use.

Part I considers the role of historical events in driving the geographic distribution and subsequent diversification patterns in *A. aperta*, with special reference

to the importance of mountain building versus Pleistocene climate change in the desert southwest United States and northern Mexico. Because *A. aperta* is widely distributed throughout the area, is ecologically confined to arid parts of the region, and exhibits low dispersal, it is ideal for looking for signatures of mountain building or climate change. I examine whether patterns and levels of genetic variation are more consistent with post-Pleistocene range expansions out of glacial refugia versus older subdivisions caused by mountain building.

Part II explores patterns of speciation in the genus *Agelenopsis* by examining phylogenetic relationships within and among *Agelenopsis* species. In particular, I assessed whether taxonomically defined species represented monophyletic groups. In addition, I evaluated the hypothesis that *A. aleenae* (Chamberlin and Ivie) is a hybrid between *A. aperta* and *A. spatula* Chamberlin and Ivie. In addition, I inferred relationships among species and examined geographic patterns of variation within species. I also assessed whether *Barronopsis*, once considered a subgenus of *Agelenopsis*, is monophyletic, and if so, what its relationship is to *Agelenopsis*.

Part III bridges the boundary between intraspecific and interspecific variation to address the possibility of an interaction between history and natural selection. The evolution of habitat-use is explored within three *Agelenopsis* species that inhabit two main habitat types, arid and riparian. This is set within the phylogeny of the genus to explore the number of times habitat use has switched in the genus and within each of the species that uses both riparian and arid habitats, *A. aperta*, *A. aleenae*, and *A. spatula*.

The conclusions following the three Parts of this dissertation summarize the findings and offer some insights into the significance of these findings. The conclusions also address a number of interesting questions raised by the work and suggest directions for future research.

Literature Cited

Avise, J. C., 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Gould, S. J. 1989. Wonderful Life: The Burgess Shale and the Nature of History. Norton, New York.

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

Riechert, S. E. 1999. Using behavioral ecotypes to study evolutionary processes. In: Geographic variation in behavior: Perspectives on evolutionary mechanisms (eds. Foster S, Endler J), pp. 3-32. Oxford Univ. Press, New York.

Rosen, D. E., 1978. Vicariant patterns an historical explanation in biogeography. Systematic Zoology, 27, 159-188.

Part I.

Molecular evidence for Pleistocene glacial cycles driving diversification of a North American desert spider, *Agelenopsis aperta.*

This Part is a slightly revised version of a paper by the same name published in Molecular Ecology:

Ayoub, N. A., Riechert, S. E. (2004) Molecular evidence for Pleistocene glacial cycles driving diversification of a North American desert spider, *Agelenopsis aperta*. Molecular Ecology. 13, 3453-3465.

The use of "we" in this Part refers to my co-author and myself. My primary contributions to this paper include 1) identifying the question regarding the effects of glacial cycles versus mountain building on species modern day distributions, 2) designing the methods to address this question, 3) collecting many of the specimens, 4) carrying out all molecular experiments, 5) analyzing the data, 6) gathering and interpreting the literature, and 7) most of the writing. …

Abstract

The influence of historical climatic versus geologic changes on species diversification patterns was investigated in a widely distributed North American desert spider, *Agelenopsis aperta* (Araneae: Agelenidae), with particular reference to Pleistocene glacial cycles and earlier patterns of mountain building. Levels of sequence divergence obtained from the mitochondrial gene, cytochrome oxidase I, dated to the Pleistocene, eliminating Rocky Mountain orogeny as a cause of diversification, as orogeny ended 4 million years ago. The results of phylogenetic and network analyses showed the presence of three geographically defined clades, which were consistent with the presence of at least three glacial refugia: 1) east of the Rocky Mountains, 2) between the Rocky Mountains and Sierra Nevadas, and 3) west of the Sierra Nevadas. In addition, populations within the Rocky Mountains exhibited significantly lower genetic diversity than populations east of the Rocky Mountains and the haplotypes found within the Rockies were a subset of eastern haplotypes. These patterns suggest that a post-Pleistocene range expansion occurred out of an eastern glacial refugium into the Rocky Mountains. Examination of phylogeographic studies of other North American desert taxa indicated that mountain building better explained diversification patterns for some taxa but Pleistocene climate change was more important for others, including *A. aperta*.

Introduction

Historical events are important determinants of geographic distributions and subsequent diversification patterns of species. As such, historical biogeography has been devoted to uncovering these events by examining the geographic distribution of evolutionary lineages. In the last twenty years, the field of phylogeography, the study of the geographic distribution of gene lineages (Avise *et al*. 1987; Avise 1994; Avise 2000) has added considerable insight into the effects of history on species distributions and diversification patterns. Biogeographic studies typically focus on a single historic event, though multiple historic events have the potential to impact the diversification of taxa. It is thus important to consider multiple historical events simultaneously to fully understand the role of Earth's history in driving diversification. Geological events, such as continental drift and the uplifting of mountains, have predictable consequences for evolutionary lineages because they often result in physical obstacles to gene flow. Such barriers lead to the formation of monophyletic groups on either side of the barrier (Rosen 1978; Avise 1994).

A less understood, but equally important type of historical event, is climate change. In contrast to geologic change, generalization about the effects of climate change on evolutionary lineages is difficult. This is because species respond differently to climate change (Faunmap Working Group 1996; Delcourt & Delcourt 1991; Bernatchez & Wilson 1998). For instance, since the last glacial maximum, range expansion or contraction of populations occurred at varying tempos and in different directions for individual mammal species (Faunmap Working Group 1996). Because

responses to climate change are species specific, biogeographic studies searching for general patterns have focused on geologic historical events, often ignoring the effects of climate change. However, more can be learned about the relative influences of geological and climatic events on diversification by forming hypotheses in light of a species' life history and ecology (e.g. Masta 2000; Knowles 2001) and then testing whether geology or climate change is more important in driving diversification.

In arid southwestern North America both geologic change in the form of mountain building and climate change in the form of Pleistocene glacial cycles could have affected the current geographic distributions of species. Consequently, biogeographic hypotheses for this region have focused on these two historical events. The region has a diverse geographic topology resulting from the orogeny (mountain building) of the Cordilleras, the complex of mountain ranges in western North America that includes the Rocky Mountains, the Sierra Nevadas, and the Sierra Madres Occidental and Oriental (Fig. A1, see Appendix at end of this Part for all figures and tables). Orogeny of some mountains began as early as 50 to 60 million years ago and possibly continued through the Pleistocene (see Fig. A2). Several phylogenetic studies indicate that mountain building is the major cause of diversification for southwestern North American desert taxa and that the effects of the more recent Pleistocene climate change have been insignificant (e.g. Riddle & Honeycutt 1990, Riddle 1995, Pook *et al*. 2000).

Nevertheless, fossil evidence from the southwestern United States indicates that climate change caused drastic changes in species distributions, at least during the last glacial maximum 18 to 40 thousand years ago. Both plant and arthropod species

experienced a general downward elevation shift such that high-elevation, cold-adapted species expanded their ranges downward, while low-elevation, arid-adapted species retreated to even lower elevations than they use today (VanDevender 1987; Elias & Nelson 1989; Betancourt 1990; Elias *et al*. 1992; Elias 1992). Presumably, prior glacial periods over the last 2.5 million years would have had similar effects. Whether these changing distribution patterns led to diversification of evolutionary lineages is still unclear (e.g. Riddle *et al*. 2000).

When glacial cycles cause changes in distribution, two different genetic patterns typically emerge, regional genetic structuring and/or decreased genetic diversity. Regional genetic structuring with eventual monophyly occurs when populations become isolated in allopatric glacial refugia during glacial maxima, (e.g. Petit *et al*. 2003; Zamudio & Savage 2003). Decreased genetic diversity is observed in populations that have expanded out of a glacial refugium into previously uninhabited areas (e.g. previously glaciated areas). Founder effects lead to lower genetic diversity in the newly colonized areas than evidenced in the parent population remaining in the glacial refugium (Ibrahim *et al*. 1996; Hewitt 1996).

The desert-adapted spider, *Agelenopsis aperta* Gertsch (Araneae, Agelenidae), has many characteristics that make it an ideal test species for the examination of geologic versus climatic influences on diversification patterns in southwestern North America. It is widely distributed and restricted to low-elevation arid habitats. Its ecology, life history, and patterns of ecotypic differentiation have been well studied (reviewed in Riechert 1999).

The large distribution of *A. aperta* and its restriction to low elevations means that mountain building could have caused vicariance among *A. aperta* populations. In addition, based on the ecological limits of *A. aperta* and the fossil record for other desert-adapted species, we can form predictions about how Pleistocene climate change might have affected the distribution of *A. aperta*. The modern day ecological limitation of *A. aperta* to low elevation desert areas and the general downward elevation shifts of plant and arthropod desert species during the last glacial maximum indicate that *A. aperta* could have been restricted to lower elevations during glacial maxima. Specifically, populations of this species might not have existed within the Rocky Mountains or in the pass *A. aperta* currently occupies between the southern end of the Rocky Mountains and the northern end of the Sierra Madres Occidental (collectively referred to as the Rockies region; Fig. A1). In addition, the historic (18 kya) presence of cold-adapted arthropods at the northeastern extent of *A. aperta's* range (near Denver, Colorado; Elias *et al.* 1992) indicates that *A. aperta* might also have been restricted to more southerly regions during glacial maxima.

In this paper we examined patterns and levels of genetic variation (mitochondrial DNA sequences) in *A. aperta* with the objective of assessing the relative importance of pre-Pleistocene mountain building versus Pleistocene climate change in shaping distribution and diversification patterns of *A. aperta.* Specifically, to what extent do we see patterns of genetic variation consistent with post-Pleistocene range expansions out of glacial refugia versus older subdivisions consistent with mountain building?

Under a scenario of Pleistocene climate change we would predict that northern portions of *A. aperta's* range and the Rockies region (see Fig. A1) would express low levels of genetic diversity; we expect this diversity to be a subset of the genetic variation seen in more southerly portions of the range. We might also expect to see evidence for population subdivision into allopatric glacial refugia (i.e. east and west of the Rockies region) but with levels of divergence dating to within the Pleistocene. Alternatively, if mountain building played a dominant role in shaping current patterns of genetic diversity we would expect to see two (or more) primary subdivisions of genetic variation — at least one group east and one group west of the Rocky Mountains — with levels of sequence divergence significantly predating the Pleistocene.

The phylogeographic study reported here is the first for an arthropod that inhabits most of southwestern North America. As such this study provides considerable insight into the biogeography of the region.

Methods

Study System

The funnel web spider, *Agelenopsis aperta*, ranges across the southwestern United States and northern Mexico (Fig. A1, Paison 1997). Unlike many spiders, *A. aperta* does not exhibit aerial ballooning (SER, 30 years of personal observation). It thus has limited dispersal ability through movement on the ground. Although

predominantly an arid-land spider, *A. aperta* uses a wide variety of habitats including continuous riparian, riparian patches surrounded by desert grassland, and desert-scrub habitats. Moving from central Texas to the east, *A. aperta* is replaced by *A. naevia*, a spider that is limited to more mesic habitats. To the north, cold-adapted *Agelenopsis* species replace *A. aperta* (Paison 1997). Our collections further indicate that *A. aperta* is restricted to areas below two thousand meters within its range (Fig. A1), with other Agelenid genera represented above this elevation. The limits on *A. aperta's* range could be due to physiological intolerance of colder winters or heightened precipitation/humidity to the north, east, and at high elevations. Alternately, competition with Agelenids specifically adapted to wetter, colder climates could also explain *A. aperta's* range limits.

Collections

We collected *Agelenopsis* individuals from 36 sites spanning most of the known range of *A. aperta* (Fig. A1, Table A1) between June 2000 and August 2003. Individuals were reared in the lab to maturity, at which time species determination was made using genitalic characters. Legs were removed and stored at -80°C to be used for DNA extractions; the remainder of each specimen was preserved in 75% ethanol with voucher specimens from the United States populations deposited at the Colorado Museum of Natural History (Denver, Colorado) and Mexican specimens at the National Collection of Acari and Arachnids (UNAM; Mexico City, Mexico).

DNA Extraction and Sequencing

We extracted total genomic DNA from frozen spider legs using a modified CTAB protocol and standard phenol-chloroform extraction (Shahjahan *et al*. 1995; Sambrook *et al*. 1987). A 700 base pair portion of the mitochondrial, protein coding gene, cytochrome oxidase I (COI) was amplified using the polymerase chain reaction (PCR) with the primers LCO1490 (Folmer *et al*. 1994) and C1-N-2191 (Simon *et al*. 1994). Amplifications were carried out in a final concentration of 0.083 U/μL *Taq* polymerase (Promega), 1X buffer (Promega), 0.83 ng/μL template DNA, 0.83 mM each primer, 0.10 mM each dNTP, and 1 mM $MgCl_2$. Reactions consisted of 2-min denaturation at 94ºC, followed by 30 cycles of denaturation for 1-min at 94ºC, annealing for 1-min at 53ºC, and elongation for 1-min at 72ºC, and ended with a 2-min elongation at 72ºC. Amplification reactions were cleaned using either ExoSap-IT™ (Amersham) or Qiagen's Min-Elute™ PCR purification kit. We then sequenced the amplified product one direction with the primer C1-N-2191 using the Big Dye Terminator Ready Reaction Kit™ (Applied Biosystems Incorporated). Sequencing reactions were sent to the University of Tennessee Molecular Biology Research Facility to be run on an ABI 3100 automated sequencer. All polymorphic nucleotide sites in the data set were checked against the original chromatograms for the possibility of sequencing errors. We excluded "noisy" bases at the beginning and end of each sequence to further ensure reliability.

We sequenced at least five individuals per sampling site except in the case of one site from which we only collected two *A. aperta* individuals (site 37). We

sequenced 10 or more individuals per sampling site whenever possible except for populations from the Rockies region. Preliminary results failed to find sequence variation within this region. Thus, we only sequenced 5 individuals from most sampling sites within the Rockies region.

Phylogenetic and Network Analyses

We performed Bayesian, maximum likelihood (ML), and maximum parsimony (MP) analyses to assess phylogenetic relationships among *A. aperta* haplotypes. We also included *A. aleenae* Chamberlin & Ivie and *A. spatula* Chamberlin & Ivie haplotypes as outgroups in all phylogenetic analyses. Both morphological characters (Paison 1997), and previous molecular phylogenetic analyses that included most *Agelenopsis* species (unpublished data) indicated that *A. aleenae* and *A. spatula* are the most likely sister species to *A. aperta*. We used PAUP*v4.0b10 (Swofford 2002) to search for MP and ML trees. We conducted heuristic searches with TBR (tree bisection reconnection) branch swapping and 10 replications of random stepwise addition. We used the same search settings for a parsimony bootstrap analysis with 100 replications.

We used the likelihood ratio test carried out in MODELTEST v3.06 (Posada & Crandall 1998) to determine which model of evolution best fit the sequence data. We found the data best fit the Tamura & Nei (1993) model of evolution with a correction for among site rate variation of alpha=0.14. Mutations were described by three substitution classes: transitions between pyrimidines, transitions between purines, and

transversions, in a ratio of 33:9:1. We thus used this model in Bayesian and ML analyses and in any other analyses requiring calculation of sequence divergence.

We ran four independent Bayesian analyses using MRBAYES v2.01 (Huelsenbeck & Ronquist 2001), each starting from a different random tree as suggested by Huelsenbeck *et al*. (2002). To adequately search the likelihood space we employed a Metropolis-coupled, Monte Carlo, Markov chain algorithm with four differently heated chains, using the default settings in MRBAYES. Three runs lasted 10 million generations (every 1000th tree saved) and one lasted 1 million generations (every 100th tree saved) for a total of 16 chains. We plotted likelihood scores versus generation to obtain a rough estimate of when the Bayesian sampling method reached stationarity. Likelihood scores plateaued after about 50,000 generations. Thus, we discarded the first 10% of trees from each run as a "burn-in" to ensure that we were sampling from a valid posterior distribution. We computed the 50% majority rule consensus tree for each run using PAUP* to give us the posterior probabilities for each node. We then compared the four trees to ensure that all runs were producing the same topology and posterior probabilities.

To test alternative tree topologies we used the Shimodaira-Hasegawa test (Shimodaira & Hasegawa 1999) carried out in PAUP* with RELL optimization and 1000 bootstrap replicates. To test if rates of evolution differed across branches we enforced a molecular clock on the ML tree in PAUP* and then used chi-squared test to determine if the clock-enforced tree was significantly less likely than the clock-relaxed tree.

Network analysis. — Because intraspecific data often do not conform to the assumption of bifurcating evolutionary lineages that is required by phylogenetic methods (Posada & Crandall 2001), we also created a network of *A. aperta* haplotypes. We used a statistical parsimony method (Templeton *et al*. 1992) carried out in TCS v1.13 (Clement *et al*. 2000) and a minimum spanning tree method (Rohlf 1973) carried out in ARLEQUIN ver 2.000 (Schneider *et al*. 2000) to create a network of *A. aperta* haplotypes.

To test if the three primary groups of haplotypes uncovered by phylogenetic and network analyses (see Results) represented a non-random geographic distribution of haplotypes we calculated the chi-square statistic used in nested clade analysis (NCA; Templeton *et al*. 1995). Significance was evaluated by 1000 random permutations of clades among sample sites. We carried out this test in GEODIS v2.0 (Posada *et al*. 2000).

Because haplotypes from two clades were found at a single site (site 33), we tested for secondary contacting using a method supplementary to NCA (Templeton 2001). This test makes use of the average pairwise geographic distance between geographical centers (APD) of the clades or haplotypes found at a particular sampling site. Under a simple model of isolation by distance, haplotypes and clades should have closely located centers and the APD should become smaller with increasing clade level. In contrast, when secondary contact occurs, the APD at that site should stay the same or increase with increasing clade level. We calculated APD for all sampling sites at each nesting level. We used the nesting procedures described in Templeton *et al*.

(1987) and Templeton & Sing (1993). We evaluated if the APD was significantly large by 1000 random permutations of haplotypes or clades among sampling sites.

To evaluate how genetic variation was partitioned among and within these phylogenetically defined geographic regions we used an analysis of molecular variance (AMOVA; Excoffier *et al.* 1992) carried out in ARLEQUIN. We partitioned genetic variation into three levels: 1) variation among sites within regions (F_{SC}), 2) variation among sites compared to the total variation (F_{ST}), and 3) variation among regions (F_{CT}).

Genetic Diversity, Population Equilibrium, and Selective Neutrality

We qualitatively examined the geographic distribution of haplotypes to determine whether northern portions of *A. aperta's* range and the Rockies region expressed a subset of the haplotypes found in more southerly parts of the range. To examine if genetic diversity decreased with latitude or longitude we calculated nucleotide diversity, the average number of pairwise substitutions per nucleotide (π). We then tested for a significant relationship between nucleotide diversity and latitude or longitude using simple linear regression carried out in JMP ver 5 (SAS Institute Inc).

We tested for equilibrium population sizes by computing Tajima's D (Tajima 1989) and Fu's *Fs* (Fu 1997) for clades uncovered by phylogenetic and network analyses. Significantly negative values for these tests indicate that populations either experienced a past population expansion or a selective sweep (i.e. rapid fixation of an

advantageous allele). Fu (1997) showed that *Fs* is especially good at detecting population expansion. We computed nucleotide diversity and carried out equilibrium tests in ARLEQUIN.

To ensure selective neutrality of the COI portion sequenced for this study, we applied the McDonald-Kreitman test (McDonald and Kreitman 1991) using DnaSP ver 3.53 (Rozas and Rozas 1999). This test does not assume panmixia, unlike other tests of neutrality, making it appropriate when population subdivision exists. The test is based on the expectation that, under neutrality, the ratio of fixed synonymous to non-synonymous substitutions between sister taxa will be equal to the ratio of polymorphic synonymous to non-synonymous substitutions. We performed this test using *A. oklahoma* Gertsch as the sister taxon because there were no fixed non-synonymous differences between *A. aperta* and the more likely sister taxa, *A. aleenae* and *A. spatula*.

Results

Sequence Characteristics

We obtained 535 base pairs of COI sequence data for each of the 474 *A. aperta* individuals sequenced. There were 51 unique haplotypes, which we deposited in GenBank (accession numbers AY676064 - AY676114). Among these 51 haplotypes, 78 nucleotide sites were variable and 52 were parsimony informative. We did not observe any stop codons and the large ratio of synonymous to non-synonymous

substitutions (70:12) fit with expectations for functional protein coding genes (Graur & Li 2000). For a breakdown of variable sites by codon position see Table A2. The McDonald-Kreitman test indicated that sequences were evolving in a neutral manner (G=0.38, p=0.54) and thus appropriate for making historical inferences.

Phylogenetic and Network Analyses

Phylogenetic and network-based analyses were consistent with the hypothesis that populations east and west of the Rockies and east and west of the Sierra Nevadas were isolated in the past.

Phylogenetic analyses. — Bayesian, MP, and ML analyses uncovered three primary geographic groupings of *A. aperta* haplotypes: (1) a clade restricted to west of the Sierra Nevadas with 100% Bayesian posterior probability and 100% parsimony bootstrap support (west of Sierra Nevadas clade); (2) a clade restricted to the area between the Rockies and the Sierra Nevadas with 100% Bayesian posterior probability and 96% bootstrap support (intermontane clade); and (3) a group of haplotypes distributed throughout the areas east of the Rockies region, the Rockies, and in three sites west of the Rockies (Fig. A3). This widespread group of haplotypes did not form a monophyletic group in the phylogenetic analysis but constituted a single sub-network (see below), which we will refer to as the "eastern clade".

In addition to the three primary groups described above, a sub-clade of the widespread "eastern" group of haplotypes was found throughout the Rockies region and in three populations west of the Rockies region (Fig. A3). We should note that this

Rockies "sub-clade" encompasses a slightly larger area than the Rockies "region", which we defined in the introduction and Figure A1.

We failed to reject a molecular clock (X^2=61, df=53, p=0.2), indicating that rates of evolution across haplotypes are roughly equivalent.

Network analysis. — The statistical parsimony method produced three sub-networks that corresponded to the same primary geographic groupings described above (Fig. A4). Haplotypes differing by up to nine mutations could be connected with 95% confidence that multiple substitutions did not occur at any particular nucleotide position. Statistical parsimony could not connect the three sub-networks but the minimum spanning network method connected both western clades to interior positions of the eastern clade. These three clades exhibited a non-random geographic distribution (X^2=865.8; p<0.0001)

Secondary contact was inferred for site 33, which contained haplotypes from both the eastern and the intermontane clades. The APD at site 33 steadily increased with increasing clade level (from haplotype to total cladogram: 249, 242, 266, 370, 553 km). At the level of the total cladogram, the APD for site 33 was significantly large (p<0.001). In contrast, all other sites' APD remained roughly constant or decreased with increasing clade level and all were zero at the level of the total cladogram (data not shown). The only other significantly large APD was for site 18 at the 3-step level (APD=445, p=0.001). This site contained haplotypes from the Rockies sub-clade and other eastern haplotypes (Fig. A3; Table A1).

AMOVA. — The results of AMOVA showed a significant amount of population structuring at all levels (include site 33 with eastern: F_{CT}=0.57; F_{ST}=0.84;

F_{SC}=0.63; p<0.001 for each; include site 33 with intermontane; F_{CT}=0.50; F_{ST}=0.82; F_{SC}=0.64; p<0.001 for each;). However, the phylogenetically defined regions accounted for the majority of the variation (57% if site 33 included with eastern).

Genetic Diversity and Population Equilibrium

Our results were consistent with a range expansion into the Rockies region, but not expansion northward. Sampling sites within the Rockies region (defined in Figure A1) exhibited a subset of haplotypes found east of the Rockies (Table A1; Fig. A3). A single haplotype (G) was fixed in the sites within the Rockies region (defined in Figure A1) except for one population at the eastern border of the Rockies (site 18). Conversely, northern sites contained unique haplotypes as well as haplotypes shared with southern sites. To exemplify this phenomenon we arbitrarily chose 35ºN as the cut off for which to compare the latitudinal distribution of haplotypes (see Fig. A4).

The plots of nucleotide diversity versus latitude and longitude further show evidence for range expansion in the Rockies but not northward. Nucleotide diversity did not show a strong overall relationship with longitude (r^2=0.082; slope=-0.0003; p=0.091), but did decrease sharply to zero at l06°W and then increased after 110°W, corresponding to sampling sites within the Rockies region (Fig. A5A). In contrast, nucleotide diversity only marginally decreased with increasing latitude (r^2=0.054; slope=-0.0003; p=0.17; Fig. A5B).

Additional evidence supporting range expansion in the Rockies were the values of Tajima's D and Fu's *Fs* for the Rockies sub-clade. Both were significantly

negative (D=-1.78, p=0.02 (beta distribution), 0.01 (simulation); *Fs*=-6.71, p<0.00001), indicating that this sub-clade was the result of population growth. Conversely, values for the three primary clades were not significant, indicating that they are the result of stable population sizes. We should reiterate that the Rockies sub-clade encompasses a larger area than the Rockies region, which would be consistent with range expansion beyond the Rockies region.

Discussion

Effects of Pleistocene Climate Change versus Mountain Building on A. aperta

The results of phylogenetic and network analyses, and patterns and levels of genetic diversity are consistent with both the hypotheses of past fragmentation and range expansion occurring in the history of *A. aperta*. While evidence for range expansion points to the importance of the Pleistocene in shaping current patterns of genetic variation in *A. aperta*, evidence for population fragmentation could be explained by either vicariance due to mountain building or population isolation in allopatric glacial refugia.

In order to distinguish the importance of climate change versus mountain building on population subdivision we must estimate divergence times among isolated populations of *A. aperta*. Dating divergence times for *A. aperta* is problematic in the absence of a fossil record to calibrate a molecular clock for *Agelenopsis*. However, using a global arthropod molecular clock estimate of 2.3% sequence divergence

between lineages per million years (Brower 1994) we can roughly date divergence of *A. aperta* haplotypes. Average pairwise sequence divergence among all *A. aperta* haplotypes was 2.9% (± 1.4%) and maximum sequence divergence between any two haplotypes was 6.4%. These levels of sequence divergence translate to intraspecific divergence times between 0.7 and 1.9 million years ago with a maximum divergence time for haplotypes of 2.9 million years ago. Because gene sequences diverge before populations diverge (Nei & Li 1979; Wilson *et al.* 1985; Edwards & Beerli 2000), we can be fairly confident that all population divergences of *A. aperta* occurred within the last 2.5 million years, well within the Pleistocene. This level of sequence divergence is too recent to be explained by Rocky Mountain orogeny, as uplift of this mountain ended approximately 4 million years ago. The *A. aperta* COI mutation rate would have to be at least half the rate described by Brower (1994) for Rocky Mountain orogeny to explain the observed levels of sequence divergence.

If *A. aperta* haplotypes diverged within the Pleistocene then the three primary clades (Fig. A3) observed in this study are likely to be the result of isolation to three allopatric glacial refugia: (1) the entire area east of the Rocky Mountains, (2) between the Rocky Mountains and the Sierra Nevadas, and (3) west of the Sierra Nevadas. We interpret these regions to constitute a minimal number of refugia but others may have existed within each of these three regions or in unsampled areas such as the Sonoran Desert (southwestern Arizona and northwestern Mexico). Indeed, the three phylogenetically defined regions do not represent panmictic populations as evidenced by the significant F_{SC} (differentiation among sites within regions). This within region population structuring may represent further fragmentation during glacial maxima but

with more contact during interglacial periods than probably occurred among regions. Particularly east of the Rockies, sample sites containing multiple haplotypes rarely form monophyletic groups and these haplotypes are typically quite divergent (Fig. A3). Furthermore, some haplotypes are shared across multiple sites. These patterns suggest that there is at least some ongoing gene flow among sites (Avise 2000). Isolation by distance, however, could also explain within region population structure without any reference to historical isolation in multiple glacial refugia. Given the low vagility of this spider, we doubt that *A. aperta* could ever have formed a panmictic population across its entire range even in the absence of physical barriers to gene flow (i.e. in this case we are talking about areas of unsuitable habitat).

Theoretically, introgression from related species could result in some of the observed divergent haplotypes. However, we have sequenced COI from multiple populations of 12 out of 13 described *Agelenopsis* species and *A. aperta* haplotypes form a monophyletic group that exhibit an average pairwise divergence of 10% from the most closely related species (unpublished data). This sampling includes populations that co-occur with *A. aperta*, strongly suggesting that introgression has not occurred.

One intriguing aspect of our phylogenetic results was the lack of monophyly of eastern haplotypes. This lack of monophyly could be explained by an ancient dispersal out of the east into the west followed by isolation into three glacial refugia. If an ancient dispersal event occurred, then the eastern portion of *A. aperta's* range should contain the most genetic diversity and should be paraphyletic with respect to the areas that are the result of dispersal. Our phylogenetic tree conforms to the expectation of

high diversity in the east but we were unable to reject monophyly of eastern haplotypes (Shimodaira-Hasegawa test; ∂=5.79, p=0.098). An alternative explanation for lack of monophyly could simply be incomplete lineage sorting (Neigel & Avise 1986).

In addition to the minimal three glacial refugia, populations found within the Rockies region and a few populations west of the Rockies are probably the result of a post-Pleistocene expansion. The fixation of a single haplotype (G) throughout these sampling sites (Fig. A3) and the significantly negative D and *Fs* values are consistent with range expansion. Thus, the geographic proximity of the Rockies sub-clade and the intermontane clade probably is due range expansion beyond just the Rockies and the co-occurrence of the two at site 33 most likely represents secondary contact between previously isolated groups. This expansion probably occurred out of an eastern glacial refugium because the haplotypes found in the Rockies form a sub-clade of the eastern clade and the sister haplotypes to the Rockies sub-clade are found in northern Mexico (site 6; see Fig. A3). However, we can't exclude the possibility that range expansion occurred out of an unsampled glacial refugium such as the Sonoran Desert. One piece of evidence that may support the past fragmentation of the Rockies sub-clade and other eastern populations is the significant APD at site 18, which exhibited haplotypes from both Rockies sub-clade and other eastern haplotypes. Alternatively, a selective sweep could account for these patterns. Data from unlinked loci would be needed to confirm the hypothesis of a demographic range expansion.

Effects of Pleistocene Climate Change versus Mountain Building on other Southwestern North American Taxa

Biogeographic hypotheses for southwestern North America either focus on vicariance events due to uplift of the Cordilleras during the Pliocene or population fragmentation Pleistocene glacial cycles. Previous biogeographic studies of this region have usually treated these two historical events as mutually exclusive when, in fact, both events have likely played a role in driving diversification of North American taxa. By examining both simultaneously we can gain a better appreciation for the role that history has played in shaping the evolution of species.

The hypothesis of vicariance due to the uplift of the Cordilleras has some support from other studies. Phylogenetic analyses of three mouse genera indicated that divergences among species restricted to a particular desert dated to the early Pliocene, about five to six million years ago (Riddle 1995). At the species level, one lizard, one snake and one mouse also showed early Pliocene divergence times between populations east and west of the Rockies (Orange *et al.*1999, Pook *et al*. 2000, and Lee *et al*. 1996). In California, a number of plant and animal taxa exhibit splits between populations east and west of the Sierra Nevadas; and north and south of the Transverse Range (Calsbeek *et al*. 2003; see Fig. A1 for location of Transverse Range). These divergences are variable, with some dating to the Pliocene and others to the mid Pleistocene.

The increased amplitude and duration of glacial cycles over the last 700 to 900 thousand years, known as the pluvial-interpluvial period in southwest North America,

could also have caused isolation between populations east and west of the Rockies by forcing desert species into isolated areas of available habitat. Zink *et al*. (2001) found that two out of five avian species examined exhibited differentiation between populations east and west of the Rockies. Levels of intraspecific sequence divergence corresponded to early to mid Pleistocene divergences. A similar pattern of sequence divergence was found for two mouse species (Riddle *et al*. 2000, Zink 2002). The genetic data presented in this study for *A. aperta* also indicate a Pleistocene divergence time for populations east and west of the Rockies.

Riddle *et al*. (2000) argued that such consistent splits between eastern and western populations were due to a constant physical barrier to gene flow by the Rockies, rather than any effects of glacial cycles. Because individual species should respond to climate change in an idiosyncratic manner, the impact of Pleistocene glacial cycles on diversification is not expected to be consistent across multiple taxa. Although similar splits occur between populations east and west of the Rockies across the taxa mentioned above, the exact location of the population divergences for each of the species are variable and the timings of divergences fall after the final uplift of the Rockies. Furthermore, three of the five avian species examined by Zink *et al*. (2001) did not show phylogenetic differentiation between eastern and western populations, but population genetic parameters indicated that two of these species experienced different population histories east and west of the Rockies. This type of variation in genetic patterns indicates that the Rockies probably did not form a consistent physical barrier to gene flow. Rather, variable responses to glacial cycles may better explain differentiation between populations east and west of the Rockies for some species.

The variation among taxa in the split between populations east and west of the Rockies might be due simply to low dispersal rather than as a response to Pleistocene climate change. Due to the stochastic nature of the coalescent process, low dispersal species can form geographically defined genetic groups even when the species is continuously distributed throughout its range (Irwin 2002). However, we do not think this is the case for *A. aperta*, because genetic groups correspond to plausible geographic barriers. Furthermore, even though the exact geographic location of splits between eastern and western populations varies for different species, the location is similar enough to indicate a common cause.

In contrast to the Rockies split, mid-Pleistocene divergence of taxa east and west of the Sierra Nevadas could be due to the final uplift of this mountain range. Some evidence indicates that uplift of the Sierra Nevadas has continued to the present time (Ruddiman *et al.* 1989). Thus, this mountain range could have formed a physical barrier to gene flow during the Pleistocene. However, recent evidence indicates that orogeny of the Sierra Nevadas occurred 50 to 60 million years ago (e.g. Wernicke *et al.* 1996; House *et al.* 1998; Wakabayashi & Sawyer 2001). If this is the case, then the presence of monophyletic groups west of the Sierra Nevadas for desert taxa may well be due to the presence of a glacial refugium rather than the Sierra Nevadas suddenly growing and forming a physical barrier to gene flow during the Pleistocene. Besides *A. aperta*, a rattlesnake (Pook *et al.* 2000) and two amphibian species (Calsbeek *et al.* 2003) exhibit divergences east and west of the Sierra Nevadas dating to the mid-Pleistocene.

Conclusions

The phylogeographic structure of the desert spider, *Agelenopsis aperta*, is consistent with the hypothesis of isolation to at least three allopatric refugia during the Pleistocene, followed by a post-Pleistocene range expansion. By examining the ecology of *A. aperta* in combination with the fossil record for other desert-adapted species, we were able to attribute diversification patterns to climate change rather than mountain building. Examination of other North American desert taxa indicates that uplift of the Cordilleras better explains diversification patterns for some taxa, but Pleistocene glacial cycles were probably more important for other taxa. At this point, too few North American desert species have been examined to fully understand the relative roles of mountain building versus Pleistocene climate change in driving distribution and diversification patterns of species found in this region. However, initial evidence indicates that both factors can be important and, consequently, future studies should consider both simultaneously. In addition, future studies should consider the life history and ecology of the species in question to resolve variation in Pleistocene effects on historical and modern day distribution patterns.

Literature Cited

Avise JC (1994) *Molecular Markers, Natural History and Evolution*, Chapman and Hall, New York.

Avise JC (2000) *Phylogeography: The History and Formation of Species*, Harvard University Press, Cambridge, MA.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel J, Reeb C, and Saunders N (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489-522.

Bernatchez L, and Wilson C (1998) Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology*, **7**, 431-452.

Betancourt J, VanDevender T, Martin P, eds (1990) *Packrat Middens: The Last 40,000 Years of Biotic Change*, The University of Arizona Press, Tucson, AZ.

Brower A (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconious erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Science USA, **99**, 6491-6495.

Calsbeek R, Thompson JN, and Richardson JE (2003) Patterns of molecular evolution and diversification in a biodiversity hotspot: the California floristic province. Molecular Ecology, **12**, 1021-1029.

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

Delcourt P, Delcourt H (1991) *Quaternary ecology: a paleoecological perspective*, Chapman and Hall, New York.

Edwards S, Beerli P (2000) Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution,* **54**, 1839-1854.

Elias S (1992) Insect fossil evidence of late Quaternary environments in the northern Chihuahuan desert of Texas and New Mexico: Comparisons with the paleobotanical record. *The Southwestern Naturalist*, **37**, 101-116.

Elias S and Nelson A (1989) Fossil invertebrate evidence for late Wisconsin environments at the Lamb Spring site, Colorado. *Plains Anthropologist,* **34**, 309-326.

Elias S, Mead J, Agenbroad L (1992) Late Quaternary arthropods from the Colorado Plateau, Arizona and Utah. *Great Basin Naturalist,* **52**, 59-67.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Faunmap Working Group. Graham RW, Lundelius EL, Graham MA, *et al.* (1996) Spatial response of mammals to late quaternary environmental fluctuations. *Science*, **272**, 1601-1606.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology,* **3**, 294-299.

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

Graur D, Li W (2000) *Fundamentals of Molecular Evolution, 2nd ed*., Sinauer Associates, Inc., Sunderland, MA.

Hewitt G (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linneaen Society,* **58**, 247-276.

House MA, Wernicke BP, Farley KA (1998) Dating topography of the Sierra Nevada, California, using apatite (U-TH)/He ages. *Science*, **396**, 66-69.

Huelsenbeck J and Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatic*s, **17**, 754-755.

Huelsenbeck J, Larget B, Miller R, and Ronquist F (2002) Potential applications and pitfalls of bayesian inference of phylogeny. *Systematic Biology,* **51**, 673-687.

Ibrahim KM, Nichols RA, Hewitt GM. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282-291.

Iriwn DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution,* **56**, 2383-2394.

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshopprers. *Molecular Ecology*, **10**, 691-701.

Lee T, Riddle B, Lee P (1996) Speciation in a desert pocket mouse (*Chaetodipus penicillatus* Woodhouse). *Journal of Mammalogy,* **77**, 58-68.

Masta SE (2000) Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): Recent vicariance or sky island populations? *Evolution*, **54**, 1699-1711.

McDonald J., Kreitman M. (1991) Adaptive protein evolution at the *adh* locus in Drosphila. *Nature*, **351**, 652-654.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science USA*, **76**, 5269-5273.

Neigel JE, Avise JC (1986) Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: *Evolutionary processes and theory* (eds Karlin S, Nevo E), pp. 515-534. Academic Press, New York.

Orange D, Riddle B, Nickle D (1999) Phylogeography of the wide-ranging desert lizard, *Gambelia wislizenii* (Crotaphytidae). *Copeia*, **1999**, 267-273.

Paison T (1997) A biogeographic review of the spider genus *Agelenopsis* (Araneae: Agelenidae). Masters Thesis, University of Tennessee, Knoxville. TN.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, *et al*. (2003) Glacial refugia: Hotspots but not melting pots of genetic diversity. *Science,* **300**, 1563-1565.

Pook C, Wuster W, Thorpe F (2000) Historical biogeography of the western rattlesnake (Serpentes: Viperidae: *Crotalus viridis*), inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution,* **15**, 269-282.

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology,* **9**, 487-488.

Posada D, Crandall KA (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution,* **16**, 37-45.

Riddle BR (1995) Molecular biogeography in the pocket mice (Perognathus and Chaetodipus) and grasshopper mice (Onychomys) - the late Cenozoic development of a North-American aridlands rodent guild. *Journal of Mammalogy,* **76**, 283-301.

Riddle BR, Honeycutt RL (1990) Historical biogeography in North American arid regions: An approach using mitochondrial-DNA phylogeny in grasshopper mice (Genus Onychomys). *Evolution* **44**, 1-15

Riddle BR, Hafner D, Alexander L (2000) Phylogeograhy and systematics of the *Peromyscus eremicus* species group and the historical biogeography of North American warm regional deserts. *Molecular Phylogenetics and Evolution,* **17**, 145-160.

Riechert SE (1999) Using behavioral ecotypes to study evolutionary processes. In: *Geographic variation in behavior: Perspectives on evolutionary mechanisms* (eds Foster S, Endler J), pp. 3-32. Oxford Univ. Press, New York.

Rohlf F.J (1973) Algorithm 76: Hierarchical clustering using the minimum spanning tree. *The Computer Journal,* **142**, 1357-1362.

Rosen DE (1978) Vicariant patterns an historical explanation in biogeography. *Systematic Zoology*, **27**, 159-188.

Rozas J, Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174-175.

Ruddiman WF, Prell WL, Raymo ME (1989) Late Cenozoic uplift in southern Asia and American west: Rationale for general circulation modeling experiments. *Journal of Geophysical Research,* **94**, 18379-18391.

Sambrook JE, Fritsch F, Maniatis T (1987) *Molecular cloning: A Laboratory Manual*, Cold Spring Harbor, New York.

Schneider S, Roessli D, Excoffier L (2000) *Arlequin ver. 3.000: A software for population genetics data analysis*, Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Shahjahan R, Hughes K, Leopold R, DeVault J (1995) Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. *Biotechniques*, **19**, 333-334.

Shimodaira H and Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution,* **16**, 1114-1116.

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

Swofford D (2002) PAUP*. *Phylogenetic analysis using parsimony (*and other methods). Version 4*, Sinauer Associates, Sunderland, MA.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512-526.

Tajima F (1989) Statitical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585-595.

Templeton AR (2001) Using phylogeographic analyses of gene trees to test species status and processes. *Molecular Ecology*, **10**, 779-791.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659-669.

Templeton AR, Boerwinkle E, Sing CF (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 AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endocuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619-633.

Templeton AR, Routman E, Phillips CA (1995) Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochonral DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics*, **140**, 767-782.

Van Devender T, Thompson R, Betancourt J (1987) Vegetation history of the deserts of southwestern North America; The nature and timing of the late Wisconsin – Holocene transition. In: *The Geology of North America, Volume K-3: North America and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman W, Wright H), pp. 323-352. The Geological Society of America, Boulder, CO.

Wakabayashi J, Sawyer TL (2001) Stream incision, tectonics, uplift, and evolution of topography of the Sierra Nevada, California. *The Journal of Geology,* **109**, 539-562.

Wernicke B, Clayton R, Ducea M, Jones CH, *et al.* (1996) Origin of high mountains in the continents: the southern Sierra Nevada. *Science*, 271, 190-193.

Williams D, Dunkerly D, DeDeckker P, Kersaw P, Chappell M (1998) *Quaternary Environments*, Arnold, London.

Wilson AC, Cann RL, Carr SM, George M, *et al.* (1985) Mitochondrial DNA and two perspectives on evolutionary genetics: *Biological Journal of the Linnaean Society*. **26**, 375–400.

Zamudio KR, Savage WK (2003) Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution*, **57**, 1631-1652.

Zink RM (2002) Methods in comparative phylogeography, and their application to studying evolution in the North American aridlands. *Integrative and Computational Biology,* **42**, 953-959.

Zink RM, Kessen AE, Line TV, Blackwell-Rago RC (2001) Comparative phylogeography of some aridland bird species. *Condor*, **103**, 1-10.

Acknowledgements

We would like to thank the Robert Shafer and Arturo Romo families for allowing us to collect spiders on their ranches, Tila Pérez for loaning us spiders from the Arachnid Collection at the National University of Mexico, and the State Parks and natural lands authorities of Arizona, New Mexico, Utah, Colorado and Texas for permitting collection on their lands. We would also like to thank Matt Anderson, Florencia Campón, Sara Duncan and Rogelio Macias for their help in sample collection. All molecular data collection was carried out in the lab of Gary McCracken at the University of Tennessee. David Posada kindly calculated APD values and evaluated their significance with his program GEOLOC. We are grateful to Chris Fleming for making elevation maps for southwestern North America. Funding for this project came from NSF grants to SER, an NSF graduate fellowship to NAA, and research grants to NAA from the Department of Ecology and Evolutionary Biology at the University of Tennessee. This manuscript was greatly improved by the comments of Marguerite Butler, James Fordyce, Richard Harrison, Thomas Near, Amy Russell, Randall Small, and all the students in EEB 611. It also benefited from the comments of two anonymous reviewers and the editor, Paul Sunnucks.

Appendix

Table A1. Collecting sites of *Agelenopsis aperta*. Numbers in parentheses are number of individuals from that site exhibiting a particular haplotype.

site #	Town/Park	Elevation (m)	Latitude (°N), Longitude (°W),	Haplotypes exhibited
1	Hye, TX	450	30.27, 98.34	R (25), F (6), D (5), C (4), B (3), CC (1)
2	San Angelo, TX	560	31.53, 100.54	U (13), R (12), V (3), A (1), P (1), BB (1)
3	Balmorhea,TX	970	30.97, 103.72	AA (24), G (3), CC (2), Z (1)
4	Just outside Big Bend National Park, TX	1235	29.65, 103.08	CC (15), P (8), Q (1), AA (1), U (1)
5	Amistad Nat'l Rec Area, TX	180	29.50, 100.91	D (10), DD (1)
6	Cuatro Cienegas, Coahuila	710	26.92, 102.05	M (9), N (1)
7	Lerdo, Durango	1160	25.52, 103.55	GG (8), EE (1), CC (1)
8	Plan de Ayupla, Zacatecas	1860	22.62, 102.78	O (7), EE (2), GG (1)
9	Tres Marias, Taumalipas	50	22.47, 98	FF (10)
10	Brantley Lake SP, NM	975	32.57, 103.38	G (6), K (1), L (1), Y (1)
11	Clovis, NM	1300	34.39, 103.28	G (10)
12	Ute Lake, NM	1160	35.37, 103.41	X (5)
13	Palo Duro Canyon SP, TX	630	34.96, 101.67	U (25), S (5)
14	Foss Lake State Park, OK	500	35.5, 99.17	T (8)
15	Colorado City, CO	1785	37.94, 104.85	II (7)
16	Avondale, CO	1390	38.224, 104.388	II (7), F (1), HH (1)
17	Sedalia, CO	1775	39.43, 104.95	II (8), S (2)
18	Nogal Mt, NM	1800	33.48, 105.8	AA (2), G (1), I (1)
19	Carrizozo, NM	1655	33.63, 105.87	G (15)
20	Rio Grande Nature Center, NM	1511	35.08, 106.65	G (5)
21	Deming, NM	1322	32.27, 107.75	G (5)
22	South Western Research Station, AZ	1655	31.91, 109.37	G (15)

Table A1. Continued.

site #	Town/Park	Elevation (m)	Latitude (°N), Longitude (°W),	Haplotypes exhibited
23	Dead Horse Ranch State Park, AZ	995	34.6, 112.1	G (5)
24	Colorado River State Park, CO	1400	39.05, 108.55	G (10)
25	San Diego, CA	15	32.73, 117.15	G (7)
26	New Port Beach, CA	15	33.57, 117.84	WE (5), WF (1)
27	North LA area, CA	15	34.07, 118.47	WG (7)
28	Big Pine, CA	1220	37.10, 118.3	WC (16), WH (7), WD (5), WB (1)
29	Warm Springs, NV	1650	38.18, 116.37	WI (10)
30	Flowell, UT	1400	39.0, 112.5	WN (10)
31	Santaquin, UT	1555	39.95, 111.78	WM (8), WN (2)
32	exit 62, Hwy 80, UT	1430	40.82, 112.9	WJ (4), WL (3), WI (2), WK (1)
33	Springdale, UT	1190	37.23, 112.95	G (14), WA (12), H (2), WO (1)
34	St. George, UT	880	37.10, 113.58	G (6), J (2)
35	near Don Pedro Reservoir, CA	435	37.82, 120.3	WP (10)
36	Bells Station, CA	15	37.04, 121.31	WQ (1), WR (1)

Table A2. Breakdown of variable sites, transitions (Ts), transversions (Tv), and base composition by codon position.

	# variable sites	# parsimony informative sites	Ts	Tv	base composition			
					A	C	G	T
1st position	7	3	5	2	0.25	0.11	0.3	0.34
2nd position	8	0	6	2	0.13	0.27	0.14	0.45
3rd position	63	49	59	8	0.27	0.02	0.15	0.57
TOTAL	78	52	60	12	0.22	0.13	0.2	0.45

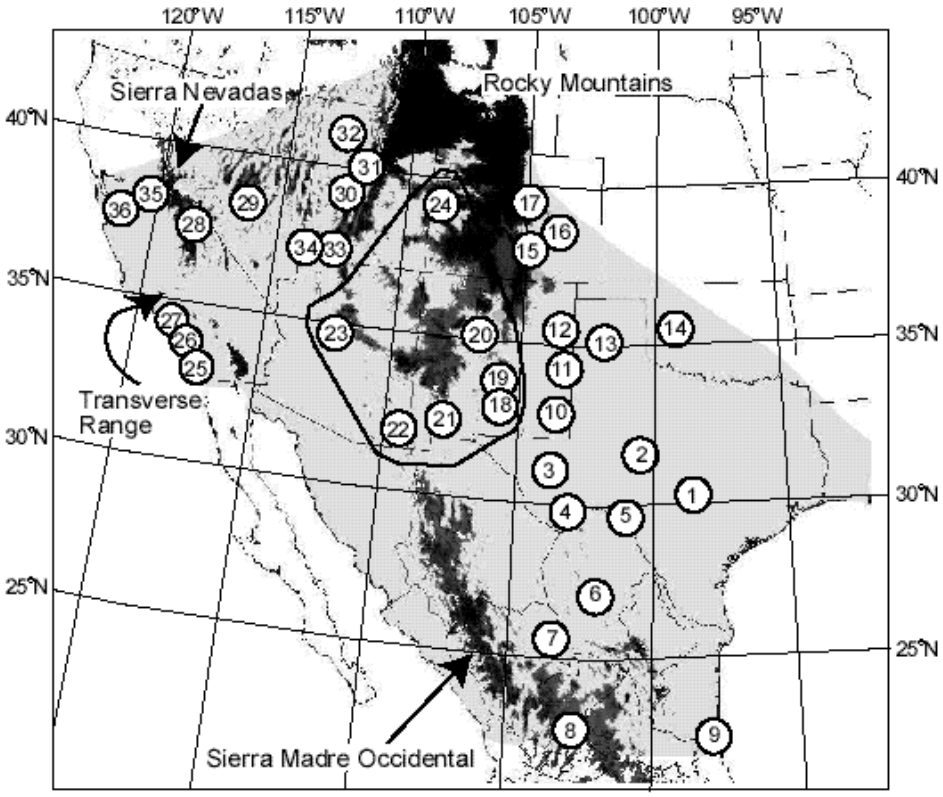


Figure A1. Sampling sites of *A. aperta*. Numbers correspond to sites in Table A1. The gray area represents the known range of *A. aperta*. Areas in black are above 2000 m. *Agelenopsis aperta* has never been collected above this elevation (Paison 1997; this study). The circled area is what we are calling the Rockies region and may have been unavailable to *A. aperta* during the Pleistocene.

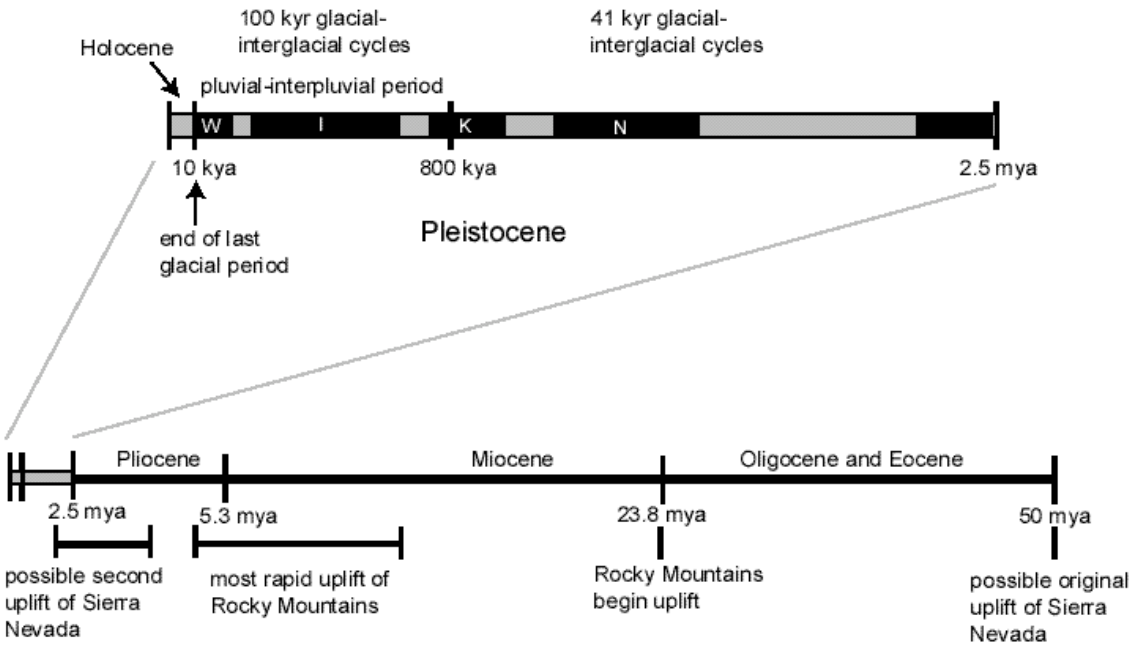


Figure A2. Time line of important geologic and climatic events in southwestern North America. Black boxes represent classically recognized glacial periods: W=Wisconsin, I=Illinoian, K=Kansan, N=Nebraskan. However, interglacial periods occurred within each of these time periods. Information on mountain building comes from Ruddiman *et al*. (1989), Wernicke *et al*. (1996), House *et al*. (1998), and Wakabayashi and Sawyer (2001). Pleistocene glacial cycle information comes from Williams *et al*. (1998). Mya=million years ago; kya=thousand years ago; kyr=thousand years.

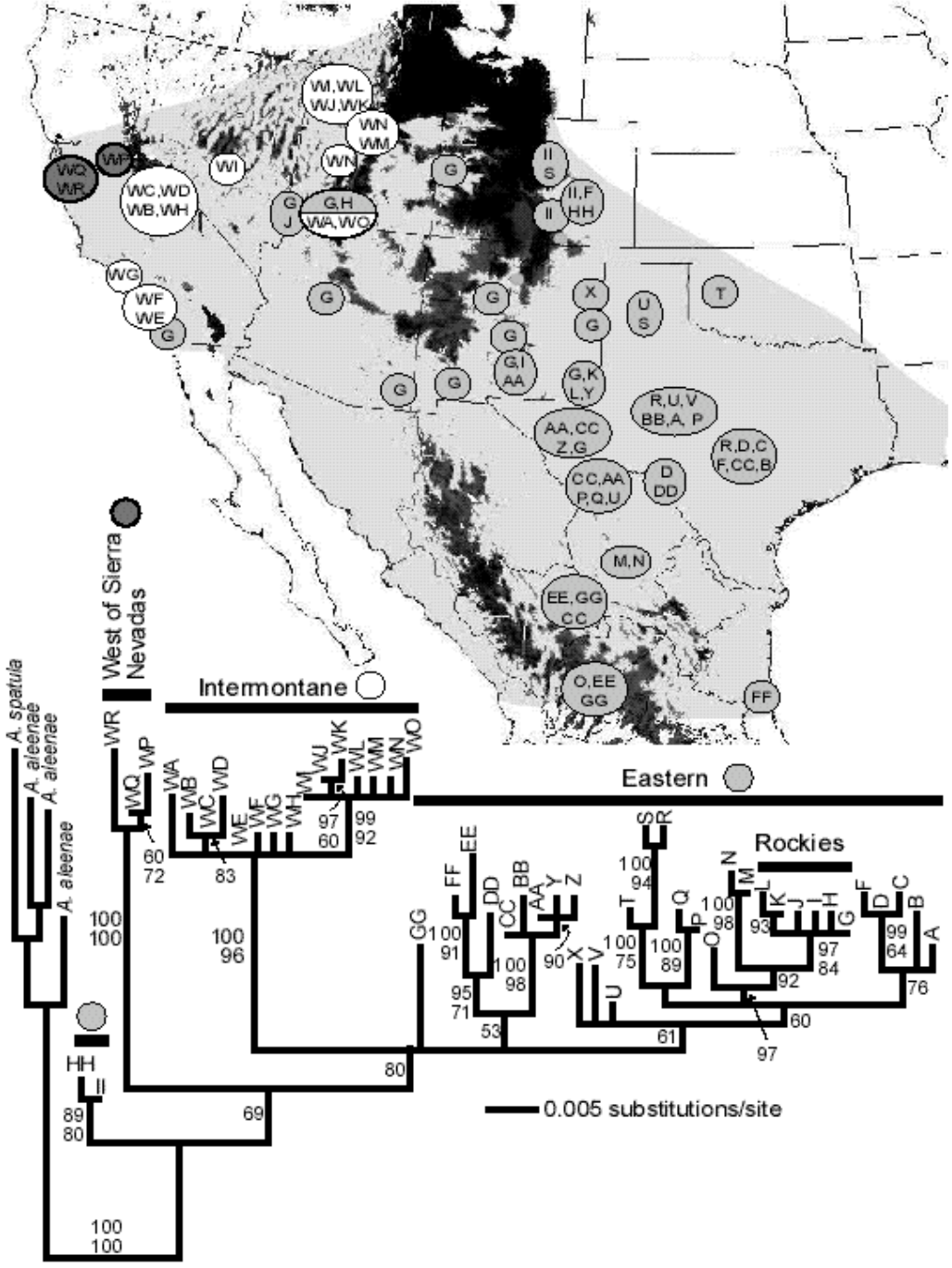


Figure A3. Geographic distribution of clades and haplotypes. Sites are color coded according to clade. The bicolored site contained haplotypes from two clades (but not in equal frequency). The tree shown below the map is the Bayesian phylogram; four independent runs produced the same topology and posterior probabilities. This tree is consistent with the strict consensus of 203 MP trees and the ML tree. Numbers beside branches represent the posterior probability (top) and the percent bootstrap support (bottom), if bootstrap support exceeded 50%.

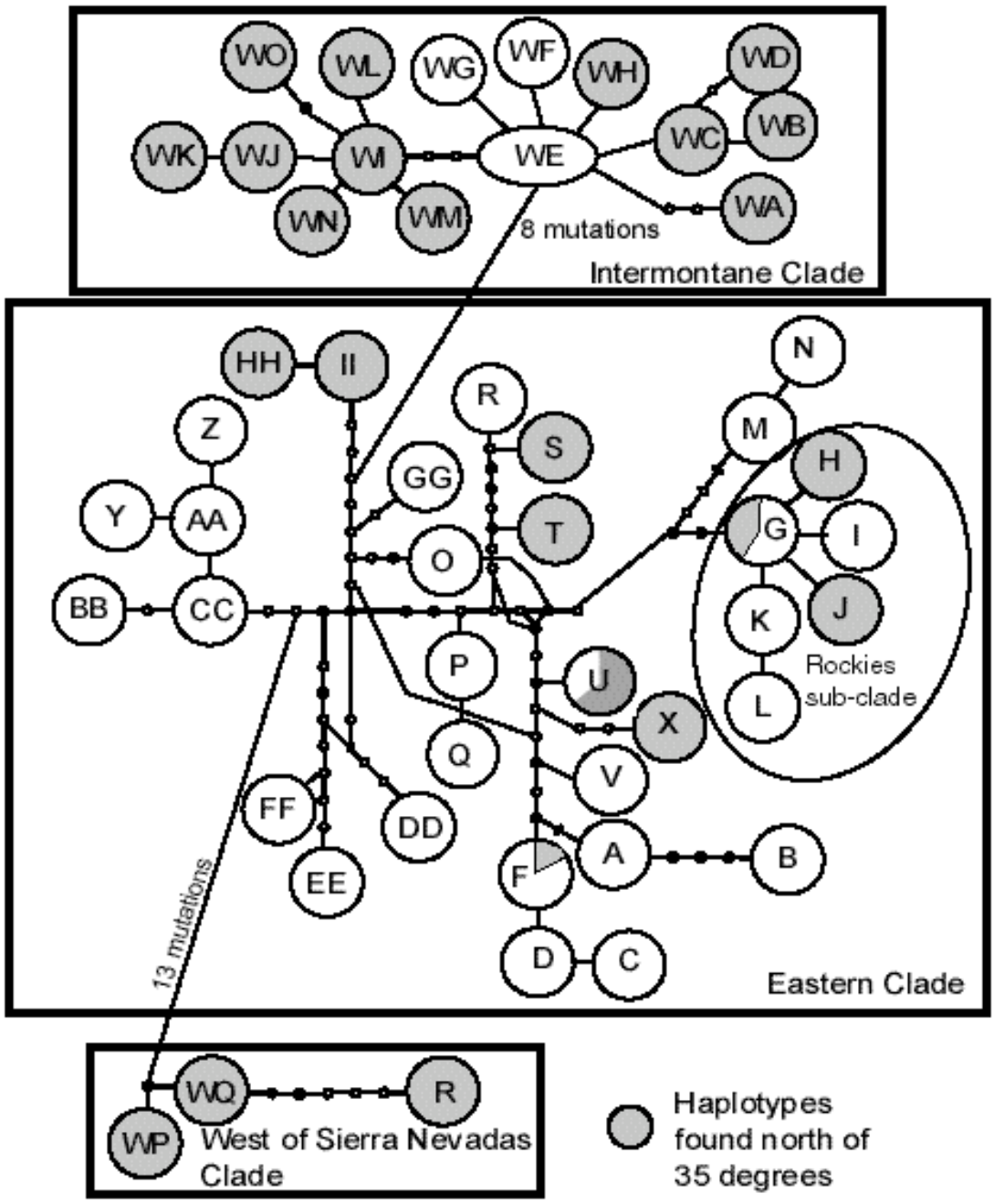


Figure A4. Minimum spanning network of observed *A. aperta* haplotypes. Sub-networks identified by statistical parsimony are boxed. Each line represents one mutational step unless indicated on the figure. Small circles represent missing haplotypes between observed haplotypes (large circles). Haplotypes found north of 35°N (arbitrarily chosen demarcation between north and south) are marked in gray according to frequency in which they occurred above or below 35°N.

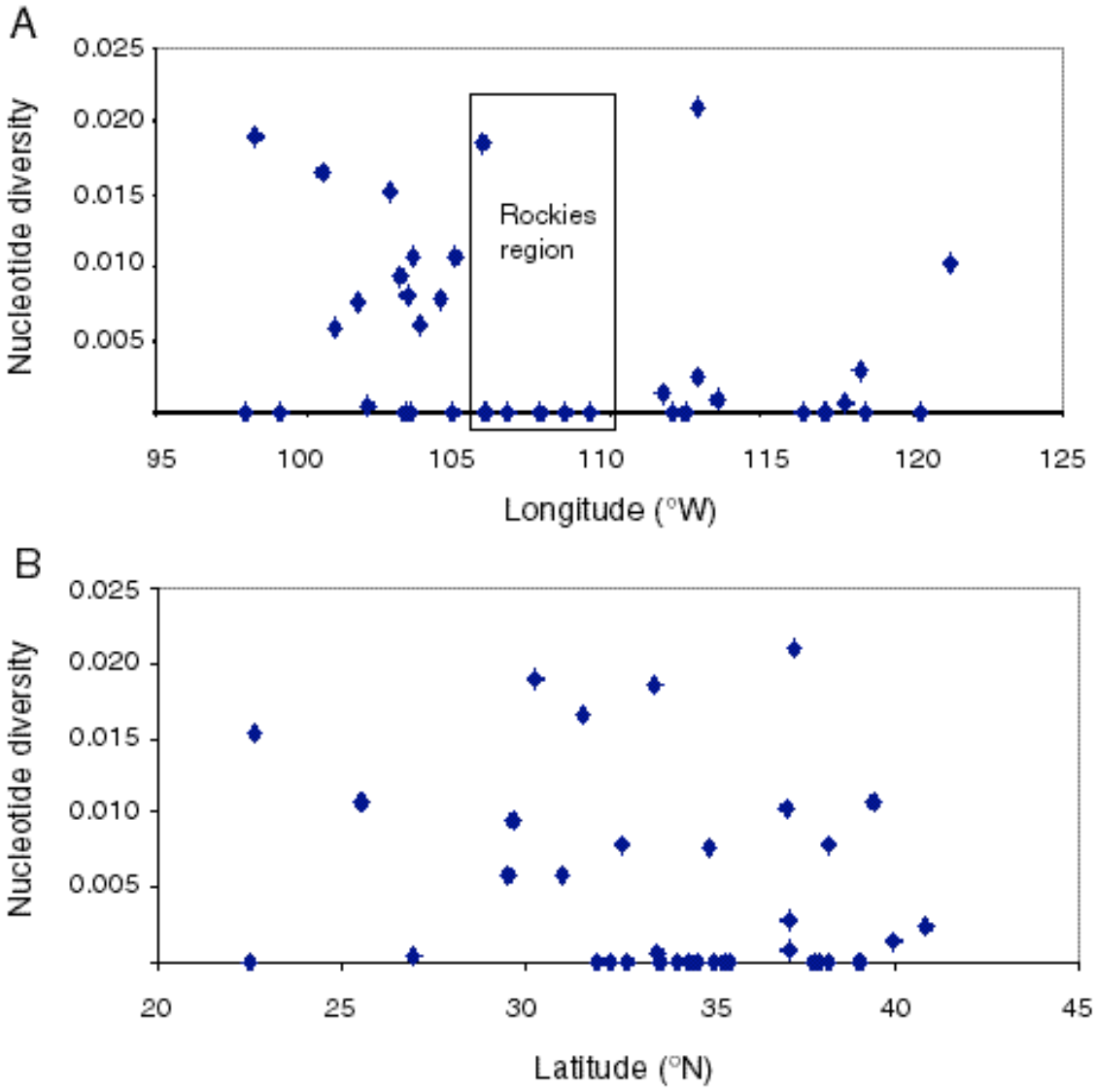


Figure A5. Plots of nucleotide diversity versus longitude (A) or latitude (B). Nucleotide diversity was estimated for each sampling site and is the average number of pairwise substitutions per nucleotide (π/# base pairs sequenced).

Part II

Speciation history of the North American funnel web spiders, *Agelenopsis* (Araneae: Agelenidae): Phylogenetic inferences at the population-species interface

This Part is a slightly revised version of a paper by the same name submitted to Molecular Phylogenetics and Evolution in July 2004. This paper is currently in review.

The use of "we" in this Part refers to my co-authors and myself. My primary contributions to this paper include 1) identifying the questions regarding speciation patterns, 2) collecting many of the specimens, 3) carrying out all molecular experiments, 4) carrying out phylogenetic analyses, 5) gathering and interpreting the literature, and 6) most of the writing. …

Abstract

Intra- and inter-specific relationships of 12 out of 13 described species as well as a potential new species in the spider genus *Agelenopsis* (Araneae: Agelenidae) were analyzed using sequence data from two mitochondrial genes, cytochrome oxidase I (COI) and 16S ribosomal RNA. We found that approximately half of the species examined formed well-supported (i.e. high bootstrap proportions and posterior probabilities) monophyletic groups, while the rest of the species were part of well-supported monophyletic species groups. Cases where species were not identified as being monophyletic could potentially be explained by poor taxonomy but we feel that incomplete sorting probably better explains lack of monophyly. We also examine the relationship between *Agelenopsis* and *Barronopsis*, which was once considered a sub-genus of *Agelenopsis*. We found the two genera to be reciprocally monophlyletic but more generic level sampling is needed to confirm an apparent sister relationship between the two.

Keywords: mitochondrial DNA, *Barronopsis*, Bayesian inference, data partitions, cytochrome oxidase I, COI, 16S ribosomal RNA, Agelenidae

Introduction

One of the major objectives of evolutionary biology is to understand the process of speciation. Interspecific phylogenies offer an indirect record of speciation events and can be used in combination with geographical or ecological data to explore the causes of speciation within a particular group (Barraclough and Nee, 2001). When intraspecific data is also available, we have the added insight of being able to explore speciation across a continuum of population divergence to species divergence (Avise, 1994, 2000; Templeton, 2001). This population-based approach to phylogenetic studies allows us to address a number of questions including the geographic basis of speciation, the incidence of hybridization between closely related species, and whether or not taxonomically defined species represent evolutionary units (i.e. monophyletic groups). The use of molecular data has become increasingly important in exploring this "population-species" interface (Templeton, 2001) because other character systems (i.e. morphological) often lack intraspecific variability (Avise *et al.*, 1987, 1994).

Many groups of spiders display little intraspecific morphological variation. In fact, even spider genera often contain so little morphological variation that species are described on the basis of only a few characters, typically male secondary sexual characters of the palps (organs that transfer sperm to females) and female genitalia (Eberhard, 1983). The advent of molecular methods has thus become increasingly useful in phylogenetic and phylogeographic studies of spiders (e.g. Huber *et al.*, 1993; Piel and Nutt, 1997; Hedin, 1997; Zehethofer and Sturmbauer, 1998; Tan *et al.*, 1999; Masta, 2000a; Piel and Nutt, 2000; Bond *et al.*, 2001; Hedin, 2001; Hedin and

Maddison, 2001a, b; Hedin and Wood, 2002; Johannesen, 2002; Smith and Bond, 2003; Croucher *et al*., 2004; Garb *et al*. 2004; Gillespie, 2004). However, only a few studies have attempted population sampling for most members of a genus (e.g. Hedin, 1997, 2001). Additional phylogenetic studies at the population-species interface are needed to identify general patterns of speciation in spiders.

An ideal group of spiders in which to examine patterns of speciation by exploring the population-species interface is a genus of funnel web spiders, *Agelenopsis* Geibel (Araneae: Agelenidae). *Agelenopsis* is widely distributed across North America and most species are abundant where they are found (Chamberlin and Ivie, 1941; personal observations by NAA and SER). Even though dispersal ability of these spiders should be restricted because they are not known to aerial disperse through ballooning (personal observations by NAA and SER), most species are still widely distributed (Paison, 1997; Platnick, 2004). The combination of having large ranges and restricted dispersal may lead to high levels of geographic differentiation within species. In addition, only a few morphological characters found in the male and female genitalia distinguish the 13 recognized species (Chamberlin and Ivie, 1941). Thus population-based sampling is feasible and molecular analysis is critical to clarifying the species and population relationships within this group.

The interspecific relationships of *Agelenopsis* have been little studied, with the only attempt at inferring historical relationships among species discussed in an unpublished Master's thesis (Paison, 1997). Paison (1997) made no attempt to use rigorous phylogenetic methods but rather constructed a branching diagram based on similarity of a few genitalic characters. One interesting hypothesis proposed, however,

was that one of the described species, *A. aleenae*, was simply a hybrid between two other species, *A. aperta* and *A. spatula*. This hypothesis was based on the observation that the male genitalia of *A. aleenae* is intermediate between the other two species (see Fig. A1) and that the single *A. aleenae* individual recorded was found in a region where the ranges of *A. aperta* and *A. spatula* overlap.

In Chamberlin and Ivie's (1941) review of North American Agelenidae, they not only summarized descriptions of species but also proposed that two species previously considered *Agelenopsis* should be segregated into a new sub-genus of *Agelenopsis*: *Barronopsis*. Later, Lehtinen (1967) elevated *Barronopsis* to generic level but the monophyly of these two genera has not been assessed.

The main objective of this paper is to uncover within and among species relationships of *Agelenopsis* to explore patterns of speciation. Within this overall objective we assess monophyly of described species with special attention to the status of *A. aleenae* (the putative hybrid species); infer relationships among species; and examine geographic patterns of variation. We also assess whether *Barronopsis* is monophyletic, and if so, what its relationship is to *Agelenopsis*.

Methods

Sampling

Taxon sampling for this study included 12 of the 13 described *Agelenopsis* species (Table A1). We also collected two *Agelenopsis* individuals that could not be

identified to any previously described species and probably constitute a new species, which we will refer to as "*A. novum*" throughout this paper. Most specimens were collected live in the field but specimens of *A. potteri* and *A. oregonensis* from Oregon, California, and Canada were sent to us preserved in 95% or 100% ethanol (see Table A1 for list of sampling localities). We included multiple individuals of each species from multiple populations when possible.

To evaluate the hypothesis that *A. aleenae* is a product of hybridization between *A. aperta* and *A. spatula* we collected *A. aperta*, *A. spatula,* and *A. aleenae* throughout their ranges (Fig. A2). Complete collections of *A. aperta* are described in Ayoub and Riechert (2004) and we include a subset of those data here. We collected 15-90 *Agelenopsis* individuals from each of nine sites throughout the known range of *A. spatula* (based on data from Paison 1997). Although *A. aleenae* had previously been known from only a single collection locality in New Mexico (Chamberlin and Ivie 1941, Platnick, 2004), five of the sites that we sampled within the range of *A. spatula* yielded *A. aleenae* individuals (Fig. A2) and only two of the sites yielded *A. spatula*. We also found *A. aleenae* at a site outside of the known range of *A. spatula* in central Texas (Hye, TX). *Agelenopsis aleenae* and *A. spatula* were never found in the same sampling site, but *A. aperta* was found in conjunction with *A. aleenae* in Hye, Texas and with *A. spatula* in Foss Lake, Oklahoma (see Table A1).

We also sampled two species of *Barronopsis*: *B. texana* and one that we could not identify. When Chamberlin and Ivie (1941) described *Barronopsis* they only placed two species in the subgenus. Subsequently, four additional species of *Barronopsis* have been described. Each species is described on the basis of

micromorphological characters of the male genitalia. Roth (1954) revised the subgenus but did not provide descriptions of females and no other work has since described North American *Barronopsis* females. We thus could not identify the female, *Barronopsis sp*. (see Table A1), to species but included it in our analyses to assess monophyly of *Barronopsis*. We further included specimens of two other agelenid genera to use as outgroups: *Hololena* and *Novolena*.

Spiders collected as juveniles were reared in our lab to maturity, at which time species determination was made using genitalic characters. Legs were removed and stored at -80°C until DNA extractions; the remainder of the specimen was preserved in 75% ethanol with voucher specimens deposited at the Arachnid Collection at California Academy, San Diego.

DNA Extraction and Sequencing

We extracted total genomic DNA from frozen spider legs using a modified CTAB protocol and standard phenol-chloroform extraction (Shahjahan *et al*., 1995; Sambrook *et al*., 1987). We sequenced two mitochondrial genes for this study: 1) the protein coding gene, cytochrome oxidase I (COI) and 2) the ribosomal RNA gene, 16S. To assess within species variation, we initially sequenced COI from multiple individuals and multiple populations of each species where possible (see Table A1). We further sequenced 16S from a subset of these individuals to obtain a better understanding of among species relationships.

Sequencing of COI involved PCR-amplifying a 700 base pair portion of COI with the primers LCO1490 (Folmer *et al.*, 1994) and C1-N-2191 (Simon *et al.*, 1994) and then direct sequencing with C1-N-2191. Sequencing of 16S involved amplifying a 500 base pair portion using the primers 16S_A and 16S_B2 (Tan *et al.*, 1999) and then direct sequencing with both primers. Amplifications and sequencing followed the procedures outlined in Ayoub and Riechert (2004). The only modification was a 55˚C annealing temperature for amplifying 16S.

COI sequences contained no internal length variation and were unambiguously aligned by eye. The 16S sequences included length variation among taxa and these sequences were aligned using the program ClustalW (Higgins *et al.*, 1994) with default gap opening and gap extension costs. A hypothetical secondary structure of 16S was determined by visually comparing the aligned agelenid sequences to the proposed secondary structures of multiple spider taxa (Masta, 2000b; Hedin and Maddison, 2001; and Smith and Bond, 2003).

Phylogenetic Analyses

All haplotypes (COI). – We initially evaluated within and among species relationships of COI haplotypes by performing maximum parsimony (MP) analysis on all observed COI haplotypes using PAUP* V4.0b10 (Swofford, 2002). Parsimony searches were heuristic with 10 replications of random stepwise addition sequences with TBR (tree bisection reconnection) branch swapping. We used the same search settings for parsimony bootstrap analysis with 100 replications.

Subset haplotypes (COI and 16S). – Because of the large number of COI haplotypes observed we did not attempt to search for maximum likelihood (ML) trees for the entire set of COI haplotypes. Instead we chose a representative sample of COI haplotypes from each species on which to search for ML trees (see Table A1 and Fig. A3 for subset used). We also evaluated relationships among the corresponding 16S haplotypes using both MP and ML searches. Gaps were treated as missing data. We used PAUP* to search for MP and ML trees for each gene separately. We conducted heuristic searches as was done for all COI haplotypes and used the same search settings for parsimony bootstrap analyses with 100 replications.

Combined data. – We evaluated congruence of COI and 16S data partitions in the context of parsimony with the incongruence length difference (ILD) test (Farris *et al.*, 1994) carried out in PAUP* with 100 replicates. Before beginning the permutations of the two data partitions we excluded invariant characters as suggested by Cunningham (1997). The results of the ILD test indicated that the COI and 16S data sets were homogeneous (p=0.81), thus we combined the two data sets. We searched for MP trees for the combined data using the same settings described above but with 1000 replications for the bootstrap analysis.

We did not attempt ML analysis of the combined data because we have no reason to believe that the two genes would have the same model of evolution or that a model of evolution calculated from the two combined would adequately represent either gene. Instead, we employed a Bayesian framework to explore the likelihood space of both genes simultaneously. We used MRBAYES v3.0 (Huelsenbeck and Ronquist, 2001), which is the only currently available phylogenetic software program

that will allow different likelihood models to be applied to different data partitions within the same analysis. We searched for trees using two types of data partitions. In the first we partitioned the data into two sets, by gene (COI and 16S). In the second we partitioned the data into five sets, by codon position (1st, 2nd, 3rd) for COI and by paired versus unpaired characters for 16S.

For both partition types we ran multiple independent Bayesian analyses each starting from a different random tree as suggested by Huelsenbeck *et al.* (2002). To adequately search all of the likelihood space, we employed a Metropolis-coupled, Monte Carlo, Markov chain (MCMC) algorithm with four differently heated chains, using the default settings in MRBAYES. For the partition by gene analysis we ran three independent searches for 10 million generations (every 500th tree saved). We found that likelihood scores reached a plateau after about 10,000 generations. However, initial Bayesian analyses run for 1 million generations produced variable posterior probabilities among independent runs for some nodes, indicating that the sampling method had not yet reached stationarity (although topology was the same among all runs). Thus, we discarded the first 10% (1 million generations) of trees from each run as "burn-in" to ensure that we were sampling from a valid posterior distribution. We computed the 50% majority rule consensus tree for each run using PAUP* to give us the posterior probabilities for each node.

The Bayesian analysis with five data partitions consisted of three independent searches for 1 million generations (every 100th tree saved). Likelihood scores reached a plateau after 1,000 generations, but we discarded the first 10% (100,000 generations) of trees as "burn-in". The posterior probabilities for each of the three runs were

similar, indicating that the MCMC procedure had reached stationarity. However, to make sure that a much longer sampling time did not change the topology or posterior probabilities we ran two more independent searches for 10 million generations. These runs were identical in topology and gave similar posterior probabilities to the runs that lasted 1-million generations.

Model fit. – To determine which model of evolution best fit the sequence data for Bayesian and likelihood analyses, we used MODELTEST v3.06 (Posada and Crandall 1998) to carry out the likelihood ratio test of progressively more complex models of evolution. We evaluated the model of evolution for each gene separately and for the five data partitions (three codon positions, paired, unpaired). For the Bayesian analyses we used the model given by MODELTEST but allowed MRBAYES to calculate optimal parameter values (the rate of change between bases, the proportion of invariant sites, and the gamma shape parameter) for each tree generated during the sampling procedure.

Results

Sequence characteristics

COI. — We obtained 635 bp of COI sequence data for 177 *Agelenopsis* individuals (including 95 *A. aperta* individuals from Ayoub and Riechert 2004), six *Barronopsis* individuals, and one each *Hololena* and *Novolena* individuals. For a complete list of sampling localities and COI sequences for *A. aperta* see Ayoub and

Riechert (2004). Among all agelenid haplotypes there were 192 variable characters of which 139 were parsimony informative. Among the *Agelenopsis* haplotypes 153 characters were variable of which 111 were parsimony informative. Third codon positions accounted for most of the variation in COI sequences with 90% of variable sites among *Agelenopsis* sequences located at third positions, 9% at first positions, and 1% at second positions. No indels or premature stop codons were observed.

16S. — We obtained 446 aligned bp (434-438 bp for individual sequences) of 16S sequence data for the 27 *Agelenopsis*, two *Barronopsis*, and one each *Hololena* and *Novolena* individuals sequenced. Gaps required for sequence alignment were never more than two bases long and alignment was unambiguous. Our proposed secondary structure for agelenid16S rRNA is presented in Figure A4. This secondary structure conforms well to other proposed spider 16S secondary structures (Masta, 2000b; Hedin and Maddison, 2001; Smith and Bond, 2003) and to *Drosophila yakuba* (Gutell and Fox 1988) secondary structure. Like other spiders, a hyper-variable region is located between positions 216 and 280 (see Fig. A4) and approximately 40% of the total number of parsimony informative characters are concentrated in this 65 bp region. Among all agelenids sequenced, 86 sites were variable, of which 61 were parsimony informative. Within *Agelenopsis,* only 51 sites were variable and 33 parsimony informative. Splitting 16S into paired (i.e. stems) and unpaired (i.e. loops) sites, unpaired sites accounted for 80% of the variable sites within *Agelenopsis*.

Phylogenetic analyses

COI. — Initial sequencing of *Agelenopsis* individuals produced 47 COI haplotypes (including 8 haplotypes from *A. aperta*; see Ayoub and Riechert (2004) for complete description of *A. aperta* COI haplotypes). Based on MP analyses of these haplotypes plus 4 *Barronopsis* haplotypes and *Hololena* and *Novolena*, we uncovered some monophyletic groupings from which to choose representatives for further analyses (Fig. A3). The results of the likelihood ratio test as carried out in MODELTEST chose the best fit model of evolution for this subset of sequences to be that of Tamura and Nei (TrN; 1993) including a proportion of invariant sites and a gamma shape correction for among-site rate variation (see Table A2 for details of parameters). MP and ML analyses of the representative sequences produced a fairly well resolved tree with most nodes having greater than 50% parsimony bootstrap support (Fig. A5). The only differences between the ML tree and the strict consensus of 12 MP trees were in relationships among *A. aperta* haplotypes and the relationships among *A. spatula* and *A. aleenae* haplotypes. These slight differences were not supported by greater than 50% parsimony bootstrap support.

16S. — The results of the likelihood ratio test of the best fit model of evolution for the 16S sequences was the General Time Reversible model (GTR; Rodriguez *et al.* 1990) including a proportion of invariant sites and a gamma shape correction for among-site rate variation (see Table A2 for details of parameters). Relationships among *Agelenopsis* haplotypes according to both MP and ML analyses (Fig. A5) were not well resolved, reflecting the paucity of variable characters within the genus. Most

of the variable characters were concentrated in separating a basal *A. utahana*/*A. oregonensis* species group from the rest of the *Agelenopses* (12 out of 33 parsimony informative characters). There were no conflicts between the ML tree and the strict consensus of 76,189 MP trees except that the ML tree was more resolved.

Combined data. — There were only a few differences between the topologies of the strict consensus of MP COI and 16S trees. Specifically, 16S placed *A. longistylus* as the first species to branch off after the basal *A. utahana*/*A. oregonesis* group followed by *A. oklahoma*. In contrast, the COI tree placed *A. longistylus* sister to an *A. oklahoma*/ *A. spatula*/ *A. aleenae*/ *A. aperta*/ *A. novum* group. However, these differences were not supported by greater than 50% parsimony bootstrap support for either gene. The only other difference was that COI indicated paraphyly of *A. utahana* with respect to *A. oregonensis* and the 16S data indicated monophyly of *A. utahana*. Thus, in the hopes of obtaining a more resolved tree by increasing the total number of characters, we combined the two genes. A strict consensus of 30 MP trees produced a topology very similar to the COI tree alone but with greater bootstrap support for some nodes (Fig. A6). The major differences between the MP trees from the combined analysis versus the COI data alone were in the placement of *A. longistylus* and *A. naevia* but these placements were not well-supported by either analysis. Thus, neither gene provided a clear assignment of these two species.

The Bayesian analyses produced similar results to the MP analysis. The resulting trees from the two types of data partitions were similar except that the tree produced by partitioning the data into five sets was more resolved and had higher posterior probabilities for some nodes than the tree produced by partitioning the data

by gene (Fig. A6). There were only a few differences between the Bayesian tree with five data partitions and the strict consensus of 60 MP trees. Specifically, the placement of *A. longistylus* differed between these placements had less than 50% parsimony bootstrap support and only 59% posterior probability. There were also slight differences in relationships among *A. aperta* haplotypes. And the MP trees indicated paraphyly of *A. utahana* with respect to *A. oregonensis* while the Bayesian trees indicated monophyly of *A. utahana* (Fig. A6).

Discussion

This paper presents the first phylogenetic hypotheses for the prominent agelendid genus, *Agelenopsis*. Although species relationships are not completely resolved by the data presented here, inferences can be made concerning the relationship between *Agelenopsis* and *Barronopsis*, species monophyly, geographic patterns of variation, and higher-level relationships among species.

Our results support reciprocal monophyly of *Agelenopsis* sensu strictu and *Barronopsis*. Our findings are consistent with both Chamberlin and Ivie's (1941) description of *Barronopsis* as a sub-genus of *Agelenopsis* and with Lehtinen's (1967) classification of *Barronopsis* as a separate genus. More generic level sampling is needed to verify if *Barronopsis* is in fact sister to *Agelenopsis* as appears to be the case from our phylogenetic results.

Species Monophyly

Our phylogenetic results supported monophyly of some *Agelenopsis* species but not others. Examples of species in which we had sufficient geographic sampling to evaluate monophyly and for which monophyly was well supported are *A. aperta*, *A. naevia*, *A. oklahoma,* and *A. longistylus*. *Agelenopsis kastoni* is probably also monophyletic but the two haplotypes observed came from a single population and, thus, we cannot reject the possibility that other populations of *A. kastoni* would prove to be more closely related to some other *Agelenopsis* species.

We also discovered two specimens that probably represent a new species, which we have referred to as "*A. novum*" in this paper. We were unable to identify these specimens to any previously described *Agelenopsis* species. In addition, the haplotype exhibited by these two individuals is fairly divergent from other *Agelenopsis* species (approximately 5% uncorrected COI+16S sequence divergence between *A. novum* and the most similar species, *A. aleenae*).

Other described *Agelenopsis* species were not found to represent monophyletic groups but are part of a monophyletic species group. For example, *A. aleenae* haplotypes are paraphyletic with respect to *A. spatula* but the two species combined represent a well-supported species group. We should mention, though, that the paraphyly of *A. aleenae* was not well supported by any analysis except the Bayesian analyses of both genes (Figs. A3, A5-A6). With the mitochondrial data presented here we were also able to reject the hypothesis that *A. aleenae* is simply a product of hybridization between *A. aperta* and *A. spatula*. Because mitochondrial DNA is

typically maternally inherited, a first generation hybrid individual should have mitochondrial haplotypes identical to one of the parent individuals. Our phylogenetic results show that *A. aleenae* exhibited multiple haplotypes that were distinct both from *A. aperta* haplotypes and from the single *A. spatula* haplotype observed. An hypothesis that is better supported by the mitochondrial data is that *A. aleenae* is in fact a distinct species and that *A. spatula* is an offshoot produced by peripatric speciation. Phylogenetic analyses show four geographically localized lineages within the *A. aleenae*/*A. spatula* clade (Figs. A3 and A7): a New Mexico *aleenae* (D), a central Texas *aleenae* (B), a Kansas *aleenae* (A), and a north Texas *spatula* (C). These data support at least four allopatric divergences from a common ancestor that probably looked like *A. aleenae* with development of the *A. spatula* morphology in a subset of the ancestral range of *A. aleenae*. This peripatric speciational mode which necessarily leads to paraphyletic species is becoming increasingly documented (Harrison 1998). In one of the few spider molecular analyses to consider intra- as well as inter-specific sequence variation, Hedin (1997) also concluded that peripatric speciation was the most likely explanation for paraphyly of two *Nesticus* species groups.

Agelenopsis utahana and *A. oregonensis* form another such group. Unlike the *spatula*/*aleenae* group, however, the *utahana*/*oregonensis* group does not form geographically distinct lineages. Even though samples of *A. utahana* came from such distant locations as Utah and New England (Fig. A2), we do not even see evidence of geographic structuring. Interestingly, though, one of the western COI haplotypes of *A. utahana* is more closely related to the *A. oregonensis* haplotype than to other *A. utahana* haplotypes (Figs. A3, A5). This pattern is suggestive of peripatric speciation

with *A. oregonensis* possibly formed from a once isolated western population of *A. utahana*.

Although *A. oregonensis* and *A. utahana* do have a few distinct characteristics of the male genitalia, they can be difficult to distinguish. The slight genitalic differences between the two species could thus represent variation within a single interbreeding species. However, paraphyly could simply be due to incomplete lineage sorting (i.e. insufficient time as passed for the two species to attain reciprocal monophyly; Neigel and Avise 1986). More sampling is needed of both these species across the entirety of their ranges to adequately quantify relationships between the two morphotypes and to verify that two distinct morphotypes exist (i.e. there is not a continuum of variation). In addition, extensive sampling in the region of overlap of these two species is needed to verify that hybridization is not playing a role in causing paraphyly of COI haplotypes.

A third species group is composed of *A. potteri, A. pensylvanica* and *A. emertoni*. Although males of these species are easy to distinguish based on genitalic characters (Fig. A1), the molecular divergence among haplotypes exhibited by these three species is quite small (average sequence divergence 0.8% for COI+16S). In addition, haplotypes do not form monopyletic groups (see Fig. A3). The ranges of *A. pensylvanica* and *A. emertoni* widely overlap with each other and partially overlap with *A. potteri* (Fig. A2). These three described morpho-species may actually represent a single interbreeding polymorphic species. Traditional ideas about spider genitalia would make this hypothesis seem unlikely; spider genitalia were thought to form a lock-and-key mechanism, such that any switch in male genitalic morphology

would immediately lead to reproductive isolation (Eberhard 1985). However, the lock-and-key mechanism is apparently not important for *Agelenopsis*; according to Gerring (1953) any male structure can mechanically inseminate any female structure. Even though the lock-and-key hypothesis is becoming more commonly disputed (see Eberhard 1985, Eberhard *et al.* 1998), most spider taxonomist consider genitalia to be monomorphic within species and Eberhard *et al*. (1998) found little variation in genitalic characters within six spider species. We are aware of only two documented cases of polymorphic spider genitalia (Huber and Gonzalez, 2001; Kaston, 1970) and a few cases of polymorphic insect genitalia (Kunze, 1959; Mound *et al*., 1998; Ulrich, 1988; Hausmann, 1999). The problem with documenting polymorphic genitalia, however, is that most spider species are described on the basis of genitalia such that any new genitalic character would be described as a new species without any knowledge of reproductive isolation. If these three *Agelenopsis* morpho-species turn out to represent a single inter-breeding species then the common use of genitalic characters as species identifying traits would have to be revised, at least for *Agelenopsis*.

Alternately, speciation may have occurred within the *A. potteri, A. pensylvanica, A. emertoni* group quite recently and there has been insufficient time for each species to attain monophyly (i.e. incomplete lineage sorting; Neigel and Avise 1986). Under this scenario, the overlap in ranges could be maintained by strong divergent sexual selection (i.e. Eberhard's (1985) theory) or through some micro-habitat differentiation. For instance, in Tennessee we collected *A. penslyvanica* from bushes in a suburban neighborhood and *A. emertoni* from leaf litter in a hard wood

forest. Detailed experiments on the biology and ecology of this species group are needed to determine what factors lead to morphological differentiation with little associated mitochondrial differentiation.

Geographic Patterns of Variation

An interesting overall pattern is that some species show fairly strong geographic structuring accompanied by deep molecular divergence, while others show very little geographic structure or molecular divergence. For instance, *A. longistylus* exhibits almost 9% COI sequence divergence between populations found only 152 km apart while *A. naevia* exhibits 0.2% sequence divergence between populations found 1230 km apart (Table A3). Other species exhibiting fairly deep molecular divergence and geographic structuring are *A. aperta* and the *A. aleenae*/ *A. spatula* species group with maximum levels of COI sequence divergence of 6.3% and 3.5% respectively. The only other species to approach this level are the *A. oregonensis*/ *A. utahana* species group with maximum COI sequence divergence of 3.1%. However, as discussed above, *A. oregonensis* and *A. utahana* show no evidence of geographic structuring, even though populations were sampled on opposite sides of the United States. Intriguingly, all the species with evidence of geographic structure are found in southwestern United States. Possibly these species have been able to maintain stable populations throughout Pleistocene climatic changes, while eastern and northern species were restricted to small glacial refugia. Ayoub and Riechert (2004) concluded that patterns of genetic structuring in *A. aperta* could probably be explained by the

presence of three glacial refugia, one of them encompassing the entire portion of *A. aperta's* range east of the Rocky Mountains. At this point, however, too little geographic sampling has been done of other *Agelenopsis* species to make conclusions about potential glacial refugia. We should mention that all the species for which we found geographic structure were thoroughly sampled throughout their geographic ranges. Extensive sampling of the other *Agelenopsis* species ranges could eventually turn up geographic structure that was undetected by our level of sampling.

Regardless of whether or not further sampling would expose geographic structure, the very low levels of divergence exhibited by haplotypes found 1000 to 4000 km apart indicates that many *Agelenopsis* species probably have much better dispersal abilities than predicted based on their dispersal mode (ground dispersing rather than aerial ballooning). For instance, *A. utahana* is found throughout northern United States, Canada and even Alaska. Most of this area was under ice sheets during the last glacial maxima, which means that *A. utahana* must have dispersed across most of its range over the last 10,000 years. Phylogeographic studies of other non-aerial dispersing spider species have found deep molecular divergence and population monophyly on fairly small geographic scales (i.e. *Nesitcus* species (Hedin, 1997); *Habronattus pugilis* (Masta, 2000); *Apostichus simus* (Bond *et al*., 2001); *Hypochilus* species (Hedin, 2001); *Hypochilus thorelli* (Hedin and Wood, 2002)), consistent with limited dispersal abilities. Our results for *A. aperta*, *A. aleenae*/ *A. spatula*, and *A. longistylus* are similar to these studies in that they show geographic structure and fairly high levels of intra-specific molecular divergence, but other *Agelenopsis* species do not. Complete generic level sampling of each of the study systems mentioned

might reveal a similar pattern with some species having strong geographic structure and others not. However, the two studies that did include most species within a genus found similar patterns of strong geographic structure for most species (Hedin, 1997; 2001).

Unequal mutation rates among species could also account for the differing levels of observed geographic structure. For instance, *A. longistylus* appears to form a very long branch (see Fig. A3). Possibly a higher mutation rate in *A. aperta*, *A. spatula*, *A. aleenae* and *A. longistylus* allows for a stronger signature of geographic structure to be detected within these species.

Higher Level Relationships

Aside from the well-supported species groups discussed above, relationships among *Agelenopsis* species are not completely resolved by the data presented here. However, a few relationships are worth discussing. First, the *A. oregonensis*/ *A. utahana* species group is sister to all other *Agelenopsis* species with high support regardless of gene or optimality criterion used (see Fig. A3, A5-A6). A well-supported sister relationship also occurs between *A. kastoni* and the *A. pensylvanica*/ *A.emertoni*/ *A.potteri* species group. Among the remainder of the *Agelenopsis* species, the positions of *A. longistylus*, *A. naevia* and *A. oklahoma* are fairly ambiguous (i.e. low bootstrap support or posterior probability for any relationships). However, maximum parsimony and bayesian analyses of the 16S and COI genes combined indicates that *A. naevia* is sister to *A. kastoni*, *A. emertoni*, *A. pensylvanica* and *A. potteri*; and that *A.*

oklahoma is sister to *A. aperta*, *A. spatula*, *A. aleenae* and *A. novum* (and possibly *A. longistylus* according to Bayesian analysis; see Fig. A6).

All analyses consistently placed *A. novum* sister to the *A. aleenae*/ *A. spatula* group even though this relationship is never supported by greater than 50% bootstrap probability. Similarly, *A. aperta* is consistently sister to the *A. aleenae*/ *A. spatula*/ *A. novum* group but this relationship only receives greater than 50% bootstrap support in the combined MP analysis and 65% posterior probability in one of the Bayesian analyses (Fig. A6).

Although the relationships noted above among *A. aperta*, *A. aleenae*, *A. spatula* and *A. novum* haplotypes are not well-supported, the male genitalia have similarities that could support monophyly of this group. As we mentioned in the introduction, *A. aleenae* has intermediate genitalia between *A. spatula* and *A. aperta*. The tip of the embolus for *A. aleenae* is spatulate (like *A. spatula*) but also twists (like *A. aperta*) (Fig. A1). Similarly, *A. novum* is intermediate between *A. aleenae* and *A. aperta*. The new species has a twisted embolus like *A. aleenae* and *A. aperta* and the tip of the embolus is broader than *A. aperta's* but does not have the distinct spatulate shape of *A. aleenae*.

Bayesian Posterior Probabilities and Data Partitioning

A number of studies have found that Bayesian posterior probabilities tend to be higher than bootstrap probabilities (Wilcox *et al.* 2002). Most people agree that bootstrap values tend to be conservative estimates of accuracy (Hillis and Bull, 1993)

but the question remains whether bayesian posterior probabilities are also conservative (Wilcox *et al.*, 2002) or if they give unrealistically high confidence measures (e.g. Suzuki *et al.*, 2002; Simmons *et al.*, 2004). When we compare the parsimony bootstrap probabilities for the combined data to the Bayesian posterior probabilities of the partition by gene analysis (Fig. A6) we find that most nodes receive similar levels of support from either measure with posterior probabilities typically only slightly higher.

Interestingly, the Bayesian analysis that used five data partitions (three codon positions for COI, paired versus unpaired for 16S) showed higher posterior probabilities for many nodes than the partition by gene analysis (Fig. A6). The sampling procedure also reached stationarity faster. Possibly, the division of the data into multiple partitions allows for a more realistic model of evolution. The division of data did in fact allow for simpler models of evolution for any one data partition compared to the models of evolution estimated for each gene as a whole (Table A2). By using simpler, possibly more realistic models of evolution, the five data partition Bayesian analysis may better estimate true relationships among haplotypes than the gene partition analysis. However, Suzuki *et al.* (2002) found in a simulation study that Bayesian analyses tend to overestimate the degree of support for a particular clade and that using simpler models of evolution than the simulated data actually fit heightens this problem. Thus, by using simpler models for each data partition, we may actually be artificially inflating the level of support for some clades.

Conclusions

The phylogenetic hypothesis for *Agelenopsis* presented here displays many interesting patterns. First of all, approximately half the species form well-supported monophyletic groups, while half are part of well-supported species groups. The lack of monophyly for some morphologically defined species could mean that genitalia are not good species characters. However, paraphyly or polyphyly of any one species was never well-supported and incomplete sorting may best explain lack of monophyly. Rather than viewing these species groups as problematic we view them as excellent areas for further research on the process of speciation, as they are probably either on the cusp of speciating or have recently speciated. We also feel that these groups present an excellent opportunity to better understand the role of genitalic variation and divergence in the process of speciation. We also found lower levels of sequence divergence and geographic structuring than expected for a non-aerial dispersing spider group, suggesting that dispersal ability of these spiders is better than previously thought. Higher level relationships were not completely resolved by the data presented here but represent an important first step in discerning deeper level evolutionary relationships within *Agelenopsis*.

Acknowledgements

We thank Joseph Spagna for sending leg tissue of *A. oregonensis* individuals from northern California and for sending us *A. potteri* and *A. oregonesis* specimens that had been given to him by R.G. Bennett. Specimens of *A. naevia* and *A. emertoni* from Dallas/Fort Worth, TX were sent to us by Marius Pfiefer. All molecular work was done in the lab of Gary McCracken at the University of Tennessee. Funding for this project came from NSF grants to SER, an NSF graduate fellowship to NAA, and research grants to NAA from the Department of Ecology and Evolutionary Biology at the University of Tennessee. This manuscript was improved by the comments of Florencia Campón.

Literature Cited

Avise, J. C., 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, New York.

Avise, J. C., 2000. Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, MA.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J., Reeb, C., and Saunders, N., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 489-522.

Ayoub, N. A., and Riechert, S. E., in press, Molecular evidence for Pleistocene glacial cycles driving diversification of a North American desert spider, *Agelenopsis aperta*. Mol. Ecol.

Barraclough, T. G., Nee, S., 2001. Phylogenetics and speciation. Trends in Ecology and Evolution. 16, 391-399.

Bond, J. E., Hedin, M. C., Ramirez, M. G., and Opells, B. D., 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Mol. Ecol.. 10, 899-910.

Chamberlin, R. V., and Ivie, W., 1941. North American Agelenidae of the genera *Agelenopsis, Calilena, Ritalena* and *Tortolena.* Ann. Entomol. Soc. Am. 34, 585-628.

Croucher, P. J. P., Oxford, G. S., and Searle, J. B., 2004. Mitochondrial differentiation, introgression and phylogeny of species in the *Tegenaria atrica* group (Araneae: Agelenidae). Biol. J. Linn. Soc. 81, 79-89.

Cunningham, C. W., 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol.. 14, 733-740.

Eberhard, W. G., 1983. Why are genitalia good species characers? In: W. G. Eberhard, Y. D. Lubin, and B. C. Robinson, (Eds.), Proceedings of the ninth international congress of arachnology. Panama, Smithsonian Institution Press, Washington DC, pp. 53-59.

Eberhard, W. G., 1985. Sexual Selection and Animal Genitalia. Harvard University Press, Cambridge, MA.

Eberhard, W. G., Huber, B. A., Rodriguez, R.L., Briceno, R. D., Salas, I., and Rodriguez, V., 1998. One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. Evolution. 52, 415-431.

Farris, J. S., Kallersjo, M. Kluge, A. G., and Bult, C. 1994. Testing significance of incongruence. Cladistics. 10, 315-319.

Folmer, O., Black, M., Hoeh, W., Lutz, R., Vrijenhoek, R., 1994. DNA primers for amplification of mitochondrial cytochrome oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotechnol. 3, 294-299.

Garb, J. E., González, A., Gillespie, R. G., 2004. The black widow genus *Latrodectus* (Araneae: Theridiidae): phylogeny, biogeography, and invasion history. Mol. Phylogenet. Evol. 31, 1127-1142.

Gerring, R. L., 1953. Structure and function of the genitalia in some American agelenid spiders. Smithsonian Miscellaneous Collections. Volume 121, 84 pages. Smithsonian Institution, Washington D.C.

Gillespie, R., 2004. Community assembly through adaptive radiation in Hawaiian spiders. Science. 303, 356-359.

Gutell, R. R., and Fox, G. E., 1988. A compilation of large subunit RNA sequences presented in structural format. Nucleic Acids Res. 16, R175-R313.

Harrison, R. G., 1998. Linking evolutionary pattern and process: The relevance of species concepts for the study of speciation. In: D. J. Howard and S. H. Berlocher, (Eds.) Endless Forms: Species and Speciation. Oxford University Press: New York, pp. 19-31.

Hausmann, A., 1999. Falsification of an entomological rule: polymorphic genitalia in Geometrid moths. Spixiana. 22, 83-90.

Hedin, M. C., 1997. Speciational history in a diverse clade of habitat-specialized spiders (Araneae: Nesticidae: *Nesticus*): Inferences from geographic-based sampling. Evolution. 51, 1929-1945.

Hedin, M. C., 2001. Molecular insights in species phylogeny, biogeography, and morphological stasis in the ancient spider genus *Hypochilus* (Araneae: Hypochilidae). Mol. Phylogenet. Evol. 18: 238-251.

Hedin, M. C., and Maddison, W. P., 2001a. Phylogenetic utility and evidence for multiple copies of elongation Factor-1alpha in the spider genus *Habronattus* (Araneae: Salticidae). Mol. Biol. Evol. 18, 1512-1521.

Hedin, M. C., and Maddison, W. P., 2001b. A combined molecular approach to phylogeny of the jumping spider subfamily Dendryphantinae (Araneae: Salticidae). Mol. Phylogenet. Evol. 18, 386-403.

Hedin, M. C. and Wood, D. A., 2002. Genealogical exclusivity in geographically proximate populations of *Hypochilus thorelli* Marx (Araneae, Hypochilidae) on the Cumberland Plateau of North America. Mol. Ecol.. 11, 1975-1988.

Hillis, D. M., and Bull, J. J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol.. 42, 782-192.

Huber, K. C., Haider, T. S., Manfred, M. W., Huber, B. A., Schweyen, R. J., and Barth, F. G., 1993. DNA sequence data indicates the polyphyly of the family Ctenidae (Araneae). J. Arachnol. 21, 194-201.

Huber, B. A., and González, A. P., 2001. Female genital dimorphism in a spider (Araneae: Pholcidae). J. Zool., Lond. 255, 301-304.

Higgins D., Thompson, J., Gibson, T. Thompson, J. D., Higgins, D. G., Gibson, T. J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673-4680.

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

Huelsenbeck, J., Larget, B., Miller, R., and Ronquist, F., 2002. Potential applications and pitfalls of bayesian inference of phylogeny. Syst. Biol. 51, 673-687.

Johannesen, J., Hennig, A., Dommermuth, B., and Schneider, J. M., 2002. Mitochondrial DNA distributions indicate colony propagation by single matri-

lines in the social spider *Stegodyphus dumicola* (Eresidae). Biol. J. Linn. Soc. 76, 591-600.

Kaston, B. J., 1970. Comparative biology of American black widow spiders. San Diego Soc. Nat. Hist. Trans. 16, 33-82.

Kunze, L., 1959. Die funktionsanatomischen Grundlagen der Kopulation der Zwergkikaden, untersucht en Euscelis plebejus (Fall.) und einigen Typhlocybinen. Dtsch. Entomol. *Z.* 6, 322-387.

Lehtinen, P. T., 1967. Classification of the cribellate spiders and some allied families, with notes on the evolution of the suborder Araneomorpha. Ann. Zool. Fenn. 4, 199-468.

Masta, S., 2000a. Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): Recent vicariance of sky island populations? Evolution. 54, 1699-1711.

Masta, S., 2000b. Mitochondrial sequence evolution in spiders: Intraspecific variation in tRNAs lacking the TgammaC arm. Mol. Biol. Evol.. 17, 1091-1100.

Neigel, , and Avise, J. C., 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Karlin S, Nevo E:, (Eds.) Evolutionary processes and theory. Academic Press, New York. pp. 515-534.

Mound, L. A., Crespi, B. J., Tucker, A., 1998. Polymorphism and kleptoparasitism in thrips (Thysanoptera: Phalaeothripidae) from woody galls on *Casuarina* trees. Aust. J. Entomol. 37, 8-16.

Paison, T., 1997. A biogeographic review of the spider genus *Agelenopsis* (Araneae: Agelenidae). Masters Thesis, University of Tennessee, Knoxville. TN.

Piel, W. H., and Nutt, K. J., 1997. *Kaira* is a likely sister group to *Metepeira*, and *Zygiella* is an araneid (Araneae: Araneidae): Evidence from mitochondrial DNA. J. Arachnol. 25, 262-268.

Piel, W. H., and Nutt, K. J., 2000. One species or several? Discordant patterns of geographic variation between allozymes and mtDNA sequences among spiders in the genus *Metepeira* (Araneae: Araneidae). Mol. Phylogenet. Evol. 15, 414-418.

Platnick, N. I., 2004. The world spider catalog, version 4.5. American Museum of Natural History, online at http://research.amnh.org/entomology/spiders/catalog/index.html

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

Rodriquez, F., Oliver, J. L., Marin, A., and Medina, J. R., 1990. The general stochastic model of nucleotide substitution. J. Theor. Biol. 142: 485-501.

Roth, V. D., 1954. Review of the spider subgenus *Barronopsis* (Arachnida, Agelenidae). American Museum Novitiates. 1678, 1-7.

Sambrook, J. E., Fritsch, F., Maniatis, T., 1987. Molecular cloning: A Laboratory Manual, Cold Spring Harbor, New York.

Shahjahan, R., K. Hughes, R. Leopold, and J. DeVault., 1995. Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. Biotechniques. 19, 333-334.

Simmons, M. P., Pickett, K. M., Miya, M., 2004. How meaningful are Bayesian support values? Mol. Biol. Evol. 21, 188-199.

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

Smith, S. D., Bond, J. E., 2003. An analysis of the secondary structure of the mitochondrial large subunit rRNA gene (16S) in spiders and its implications for phylogenetic reconstruction. J. Arachnol. 31, 44-54.

Suzuki, Y., Glazko, G. V., and Nei, M., 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. Proc. Natl. Acad. Sci. USA, United States. 99, 16138-161433.

Swofford, D. L., 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512-526.

Tan, A. M., Gillespie, R. G., and Oxford, G. S., 1999. Paraphyly of the *Enoplognatha* group (Araneae, Theridiidae) based on DNA sequences. J. Arachnol. 27, 481-488.

Templeton, A. R., 2001. Using phylogeographic analyses of gene trees to test species status and processes. Mol. Ecol.. 10, 779-791.

Ulrich, H., 1988. Das Hypopygium von *Microphor holosericus* (Meigen) (Diptera, Empidoidae). Bonn. Zool. Beitr. 39, 179-219.

Wilcox, T. P., Zwickl, D. J., Heath, T. A., and Hillis, D. M., 2002. Phylogenetic relationships of the dwarf boas and a comparison of Bayesian and bootstrap measures of phylogenetic support. Mol. Phylogenet. Evol. 25, 361-371.

Zehothofer, K., Sturmbauer, C., 1998. Phylogenetic relationships of central European wolf spiders (Araneae: Lycosidae) inferred from 12S ribosomal DNA sequences. Mol. Phylogenet. Evol. 10, 391-398.

Appendix

Table A1. Haplotypes used in this study, with localities. Numbers in parentheses indicate number of individuals from that locality exhibiting that haplotype.

Species	COI haplotypes	Genbank Accession # Top = COI Bottom=16S	locality
A. aleenae Chamberlin and Ivie	aleenae1**	AY770786 AY770755	just outside Meade SP, KS (3)
	aleenae2**	AY77079 AY770752	Hye,TX (5)
	aleenae3	AY770788	Hye, TX (1)
	aleenae4	AY770827	Hye, TX (1)
	aleenae5**	AY770825 AY770751	Hye, TX (3)
	aleenae6	AY770819	Hye, TX (2)
	aleenae7	AY770820	Hye, TX (2)
	aleenae8	AY770821	Hye, TX (2)
	aleenae9	AY770822	Hye, TX (1)
	aleenae10	AY770823	Hye, TX (4)
	aleenae11	AY770824	Hye, TX (1)
	aleenae12**	AY770787 AY770756	Clines Corners, NM (4)
	aleenae13**	AY770788 AY770754	Hobbs, NM (2)
	aleenae14	AY770789	Hobbs, NM (3)
	aleenae15	AY770790	3 mi west of Cimarron, NM (1)
	aleenae16	AY770818	Valmora, NM (2)
A. aperta[1] Gertsch	aperta-G**	AY770778 AY770758	South Western Research Station, Portal, AZ (15), and others[1]
	aperta-WE**	AY770780 AY770762	New Port Beach, CA (6)
	aperta-II**	AY770782 AY770761	Sedalia, CO (8), and others[1]
	aperta-S**	AY770781 AY770760	Sedalia, CO (2)
	aperta-U**	AY770783 AY770759	San Angelo, TX (9)
	aperta-WP**	AY770779 AY770763	near Don Pedro Reservoir, CA (10)
	aperta-T	AY770785	Foss Lake SP, OK (8)

Table A1. Continued.

Species	COI haplotypes	Genbank Accession # locality
		Top = COI
		Bottom=16S	

	aperta-R	AY770784	Hye, TX (25), San Angelo, TX (12)
A. emertoni Chamberlin and Ivie	emertoni1**	AY770805 AY770767	Foss Lake SP, OK (1); Dallas, TX (1)
	emertoni2**	AY770806 AY770765	Loudon, TN (1)
A. kastoni Chamberlin and Ivie	kastoni1	AY770800	Loudon, TN (1)
	kastoni2**	AY770807 AY770766	Loudon, TN (1)
A. longistylus Banks	longistylus1**	AY770814 AY770770	3 mi. west of Cimarron, NM (4)
	longistylus2	AY770828	1.4 mi west of Galisteo, NM (1)
A. naevia Walckenaer	naevia1	AY770801	Loudon, TN (2)
	naevia2**	AY770809 AY770769	Dallas, TX (2)
A. oklahoma Gertsch	oklahoma1**	AY770803 AY770749	Foss Lake SP, OK (1); just outside Meade SP, KS (1); Lake Arrowhead SP, TX (1)
	oklahoma2**	AY770804 AY770750	Trinidad, CO (1)
	oklahoma3	AY770808	Colorado River SP, CO (4)
A. oregonensis Chamberlin and Ivie	oregonensis1**	AY770792 AY770774	8.3 mi west of Del Loma, CA (2)
A. pensylvanica C. Koch	pensylvanica1**	AY770799 AY770768	Foss Lake, OK (1); Knoxville, TN (3); Ellsworth, KS (1); Barber Pond SP, CO (1)
A. potteri Blackwell	potteri1**	AY770793 AY770764	Victoria, BC, Canada (1)
	potteri2	AY770794	Winnipeg, MB, Canada (1)
A. spatula Chamberlin and Ivie	spatula1**	AY770802 AY770753	Foss Lake, OK (1); Caprock Canyons SP, TX (6)

Table A1. Continued.

	COI haplotypes	Genbank Accession # locality
		Top = COI
		Bottom=16S	

A. utahana Chamberlin and Ivie	utahana1**	AY770811 AY770772	White Mountains National Forest, NH (1)
	utahana2**	AY770812 AY770773	Monadnock SP, NH (1)
	utahana3	AY770813	Lincoln, NH (1)
	utahana4**	AY770829 AY770775	Wasatch Mountains SP, UT (1)
	utahana5**	AY770795 AY770771	Wasatch Mountains SP, UT (3)
*A. novum**	novum**	AY770817 AY770757	10 mi south of Balmorhea, TX (2)
Barronopsis texana Gertsch	B. texana1**	AY770796 AY770747	Knoxville, TN (2)
	B. texana2	AY770815	Silver Springs, FL (2)
	B. texana3	AY770816	Hye, TX (1)
*Barronopsis sp**	Barronopsis sp**	AY770810 AY770748	Salt Springs, FL (1)
*Hololena sp**	*Hololena***	AY770797 AY770777	Colorado Springs, CO (1)
*Novolena sp**	*Novolena***	AY770798 AY770776	Yosemite National Park (1)

* Indicates specimens that could not be identified to species.

** Indicates subset of haplotypes used in more extensive phylogenetic analyses. An individual exhibiting this haplotype was also sequenced for 16S.

[1]For complete list of *A. aperta* sample sites and haplotypes see Ayoub and Riechert (2004). Letters at the end of the haplotype name correspond to the haplotype name used in Ayoub and Riechert (2004).

Table A2. Estimates of model parameters for each data partition obtained using the likelihood ratio test carried out in MODELTEST (Posada and Crandall 1998).

	Base Frequencies				Rate Matrix						I	alpha
	A	C	G	T	A< >C	A<>G	A<>T	C<>G	C<>T	G<>T		
COI	0.25	0.10	0.19	0.46	1	8.56	1	1	19.16	1	0.65	2.49
16S	0.40	0.11	0.12	0.43	21.48	14.75	14.06	0.00	166.52	1	0.55	0.46
3rd position	0.28	0.02	0.14	0.56	1	32.8	1	1	82.61	1	NA	3.61
unpaired	0.46	0.12	0.06	0.36	1	2	1	1	10.81	1	NA	0.17
					transition/transversion ratio							
1st position	0.25	0.25	0.25	0.25	3.15							
2nd position	0.13	0.26	0.17	0.44	NA							
paired	0.31	0.14	0.16	0.39	15.40							

Table A3. Levels of intraspecific COI sequence divergence and maximum geographic distance between intraspecific sampling localities.

species	pairwise substitutions per nucleotide (TrN corrected)		maximum geographic distance (km)
	average	maximum	
A. aperta	0.037	0.065	2200
A. aleenae	0.022	0.034	790
A. oklahoma	0.0025	0.0037	1085
A. potteri	0.0098	0.0098	2037
A. pensylvanica	0	0	1895
A. emertoni	0.0077	0.0077	1228
A. kastoni	0.0018	0.0018	0
A. naevia	0.0019	0.0019	1228
A. longistylus	0.089	0.089	152
A. oregonensis	0	0	0
A. utahana	0.014	0.025	3300
A. novum	0	0	0
A. spatula	0	0	214
species groups			
aleenae-spatula	0.023	0.035	527
emertoni-pensylvanica-potteri	0.009	0.012	3467
oregonensis-utahana	0.018	0.029	3893

Figure A1. Genitalia of select *Agelenopsis* species. Shown for each species is the left pedipalp (above) and the female epigynum (below). For A-C we show a close up of the tip of the embolus. A) *A. aperta*, B) *A. aleenae*, C) *A. spatula*, D) *A. pensylvanica*, E) *A. potteri*, and F) *A. emertoni*. Note that the tip of the embolus for *A. aleenae* is intermediate between *A. aperta* and *A. spatula*. In *A. aperta* the tip of the embolus is twisted and in *A. spatula* it is spatulate. In *A. aleenae* it is both twisted and spatulate. Drawings taken from Chamberlin and Ivie (1941).

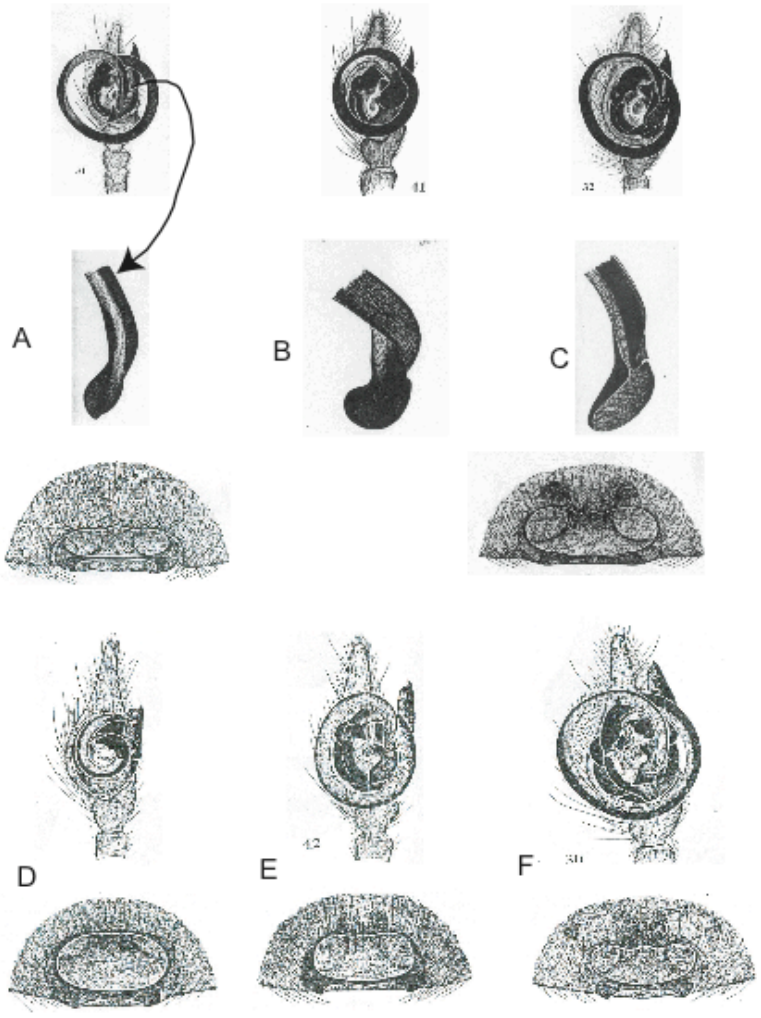

A
B
C
D
E
F

Figure A2. Geographic ranges and sampling localities of the 12 described *Agelenopsis* species sampled in this study and the location of the specimens of *A. novum*. Geographic ranges are based on our collections and those reported in Paison (1997) and Platnick (2004). Species do not necessarily use all available areas within range. Hatched areas indicate where ranges overlap.

[1] *A. utahana* is also found throughout Canada and into Alaska.

[2] *A. aperta's* range extends into northern Mexico.

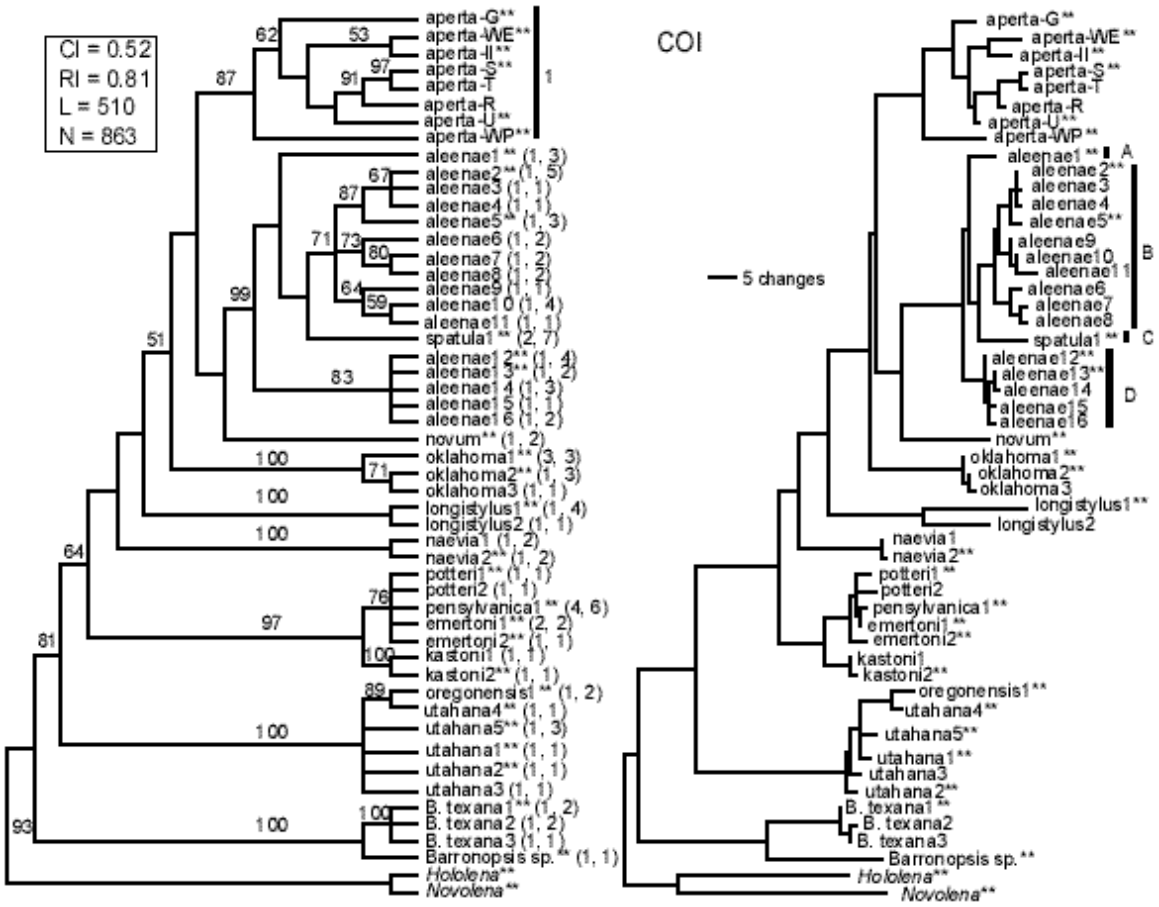


Figure A3. Strict consensus of MP trees (left) and an arbitrarily chosen single MP phylogram (right) for all COI haplotypes observed in this study. Numbers above MP branches represent percent bootstrap support based on 100 replicates. After each haplotype the number of populations followed by the number of individuals exhibiting that haplotype are indicated. A-D represent the four geographically localized gene lineages within the *A. aleenae*/ *A. spatula* clade (see Fig. A7). CI=Consistency index, RI=Retention index, L=length of trees, N=number MP trees.
** Indicates subset of haplotypes used for further analysis and an individual bearing that haplotype was also sequenced for 16S.
[1]For complete list of *A. aperta* populations and haplotypes see Ayoub and Riechert (2004).

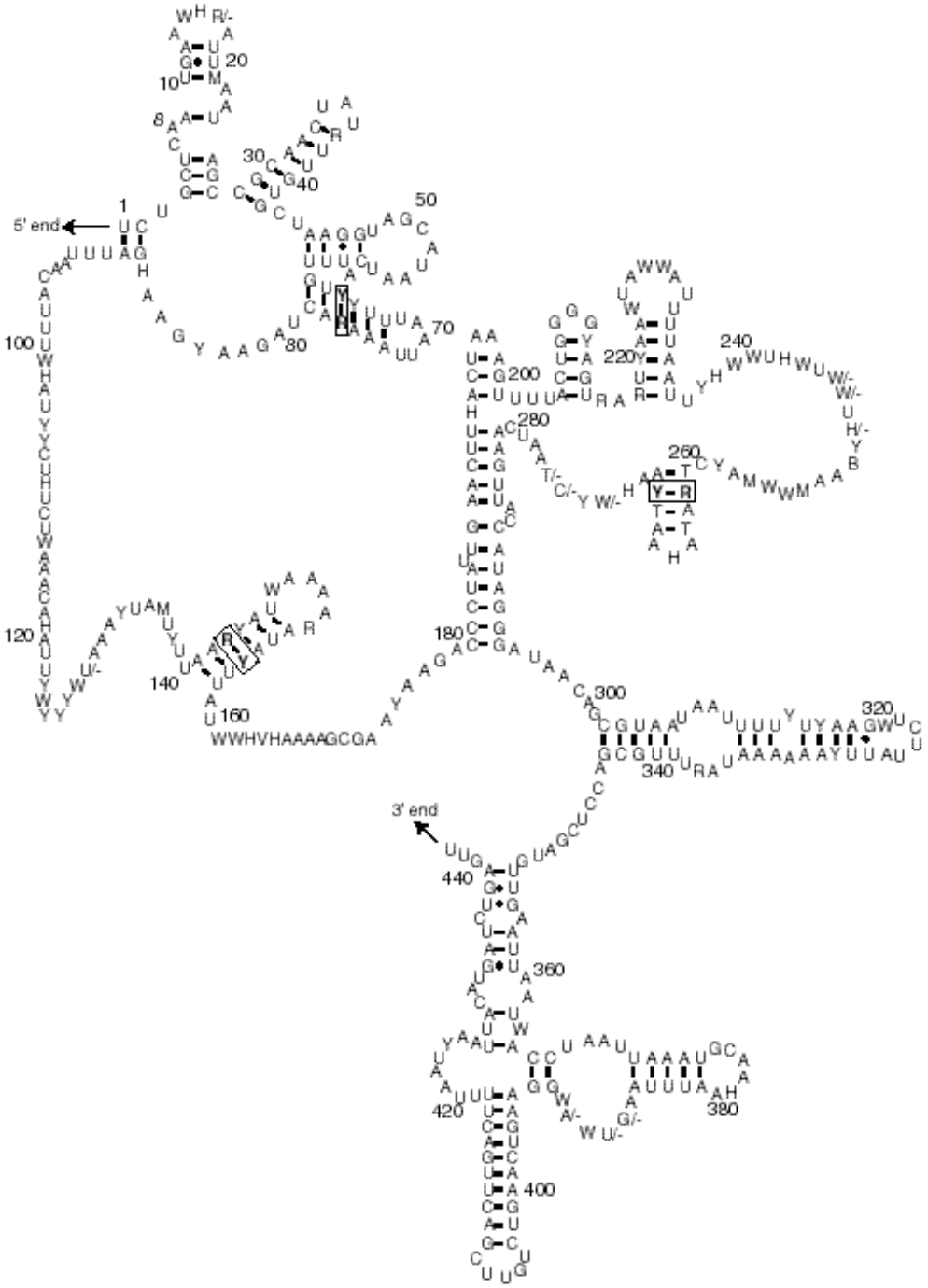


Figure A4. Our proposed secondary structure of agelenid 16S ribosomal RNA. Dashes represent Watson-Crick bonds and circles G-U bonds. Polymorphic sites are indicated by IUPAC symbols. Proposed compensatory mutations are boxed. Indels are depicted as base/-.

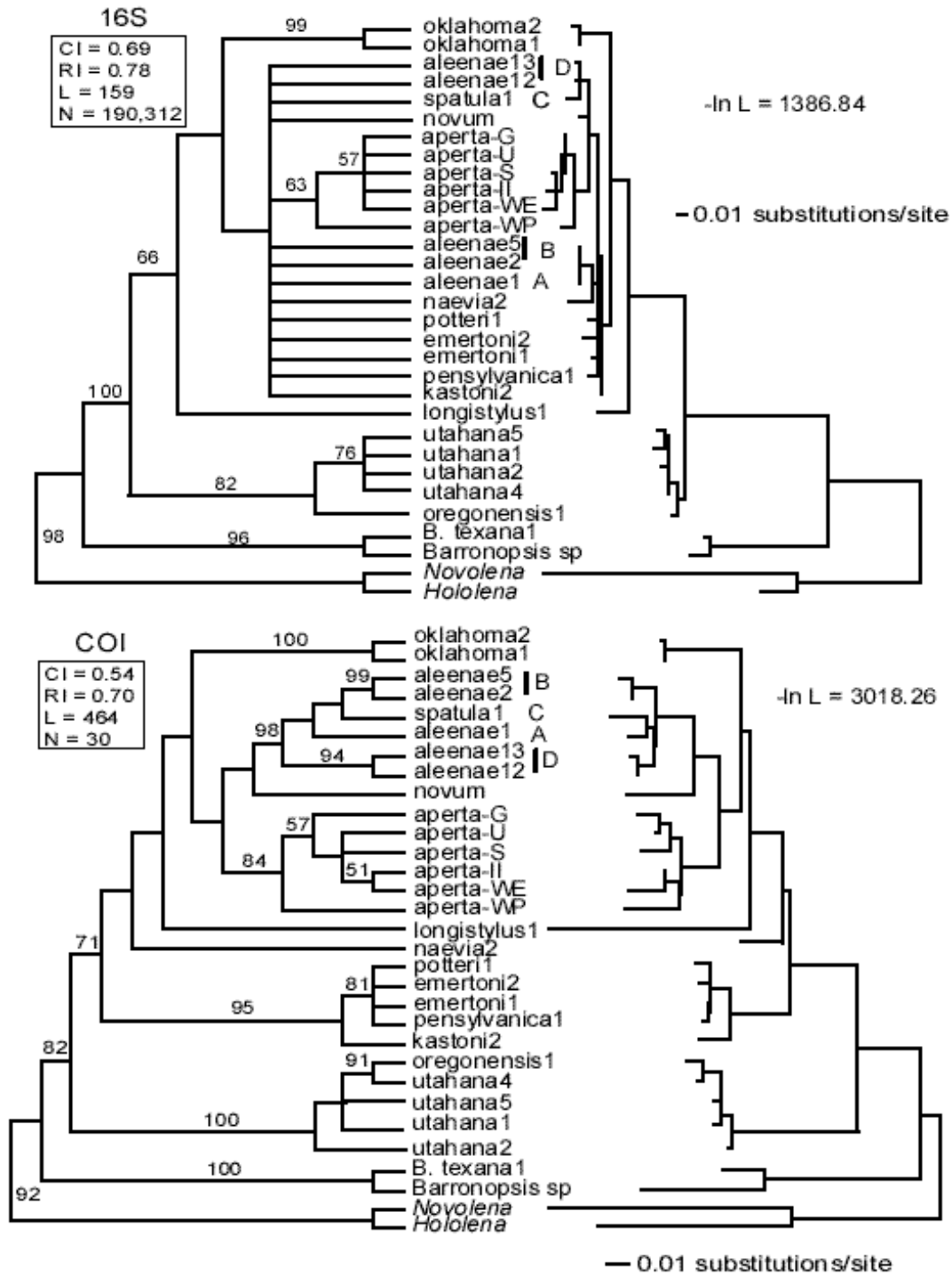


Figure A5. Strict consensus of MP trees (left) and ML trees (right) for16S (top) and COI (bottom). Numbers above MP branches represent percent bootstrap support based on 100 replicates. CI=Consistency index, RI=Retention index, L=length of trees, N=number most parsimonious trees.

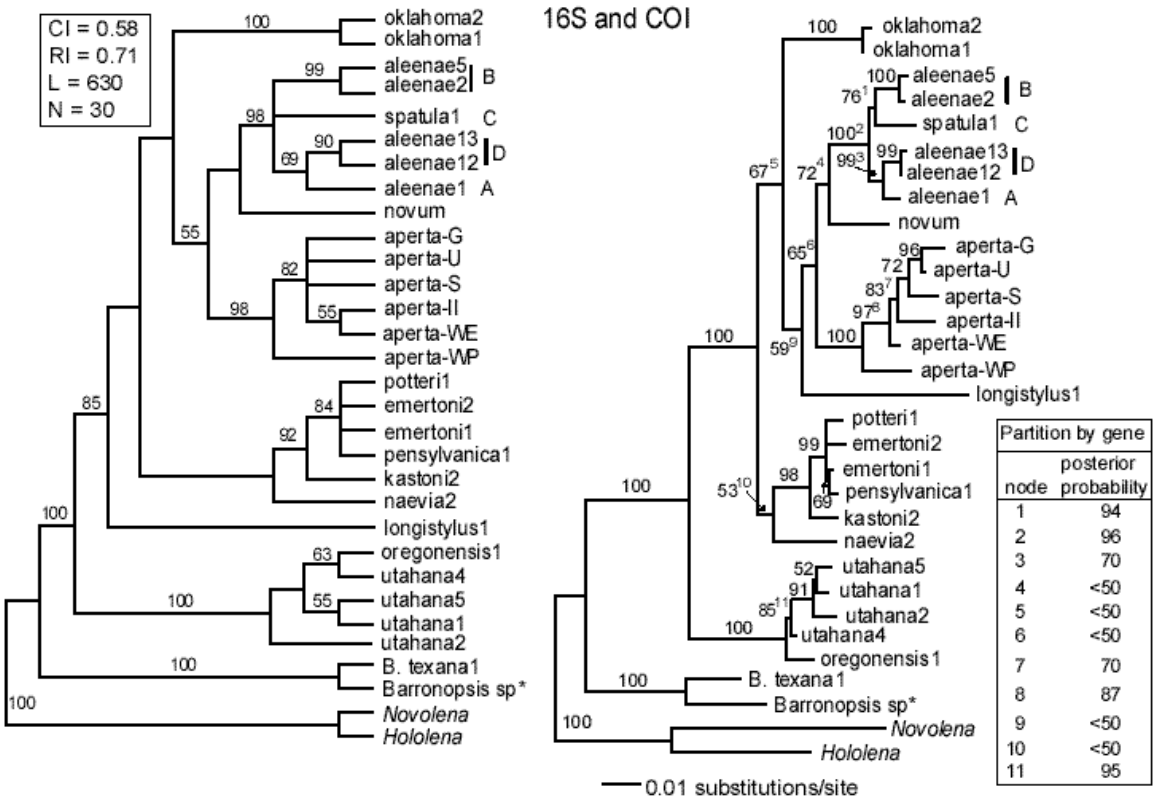


Figure A6. Strict consensus of MP trees (left) for combined 16S and COI analysis and Bayesian tree from five data partition analysis (right). Numbers above MP branches represent percent bootstrap support based on 1000 replicates. Numbers above branches on Bayesian tree represent posterior probability. The posterior probabilities for some nodes differed between the five data partition analysis and the partition by gene analysis. These values are shown in the lower box. Upper box: CI=Consistency index, RI=Retention index, L=length of trees, N=number most parsimonious trees.

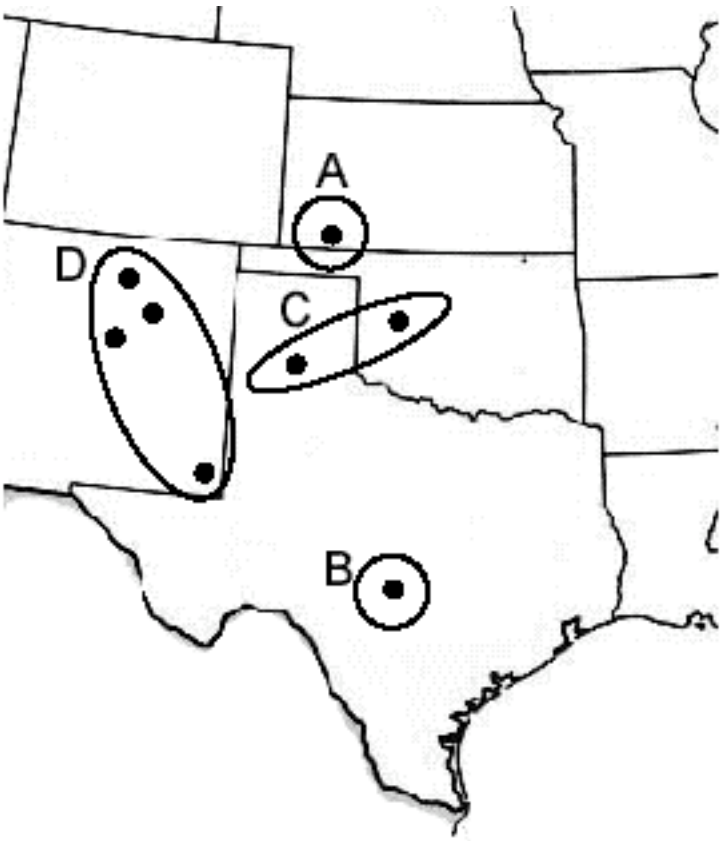


Figure A7. Geographic distribution of lineages within the *A. aleenae - A. spatula* species group. Letters correspond to lettered gene lineages in Figure 3. A= Kansas aleenae, B=Central TX aleenae, C=spatula, D=New Mexico aleenae.

Part III

Evolution of habitat-use in a group of funnel-web spiders, *Agelenopsis*: the potential interaction of history and natural selection in shaping behavioral adaptations.

Abstract

Although historical contingency is often considered an alternative to natural selection as explanations for evolutionary patterns, it is more realistic to view the two processes as interacting. An informative approach to searching for such interactions would be to examine local adaptation within a species in the context of interspecific relationships. Using such an approach, I examined historical relationships among populations of the desert spider, *Agelenopsis aperta,* which exhibits differential behavioral adaptations in riparian and arid habitats. I set this behavior in the context of the phylogeny of the genus and compared *A. aperta's* population structure to *A. aleenae* and *A. spatula*, species that also uses both riparian and arid habitats. Using the mitochondrial gene, cytochrome oxidase I, I found that riparian patches of *A. aperta* distributed throughout the desert are significantly different. This indicates that within *A. aperta* any similarities in behavior would be the result of repeated, independent natural selection rather than gene flow and historical contingency. However, all the *Agelenopsis* species which use arid habitats, *A. aperta, A. aleenae*, *A. spatula* and *Agelenopsis species novum*, formed a monophyletic group, indicating that arid living evolved only once within the genus. The interspecific phylogeny suggests that although arid living evolved once, each species within the arid clade retained the ability to invade riparian habitats. These findings suggest an interaction between history and natural selection in shaping current behavioral characters in *Agelenopsis*.

Introduction

Natural selection, chance, and history are recognized as the three leading factors affecting evolutionary diversity (Gould 1989, Gould and Woodruff 1990, Losos et al. 1998). Although the interaction between chance and natural selection has received extensive attention in the field of population genetics, the potential interaction between history and natural selection has received relatively little attention (Travisano et al. 1995). Accordingly, documented cases of an interaction between selection and history are still rare (but see Taylor and McPhail 2000), possibly because specific traits are often interpreted to be the result of only one process (e.g. Gould and Woodruff 1990, Travisano et al. 1995, Losos et al. 1998). Identifying these interactions will thus fill a gap in our understanding of evolutionary diversification. This study will add further insight into potential interaction effects as it examines the influences of history and previously documented natural selection on the evolutionary diversification of behavioral traits.

The comparative method is used here. It has been widely employed to distinguish the effects of natural selection from historical contingency (Harvey and Pagel 1990, Losos et al. 1998, Pagel 1999) and makes use of a phylogenetic framework to evaluate whether a similar phenotype found in similar environments is contingent upon a single ancestral switch or if the character arose independently multiple times. Multiple independent origins of a trait in similar habitat reflect parallel or convergent evolution and are considered evidence for the importance of natural selection (e.g. Losos et al. 1998).

The comparative method has been used mostly at the interspecific level, but has been applied below the species level as well (e.g. Edwards and Kot 1995, Foster and Cameron 1996, Thorpe 1996, Zamudio 1998, Wiens et al. 1999, Malhotra and Thorpe 2000). Intraspecific comparative studies are relatively rare compared to interspecific studies but are informative about the evolution of adaptive characters because there is less variation among populations within a species than there is between species, and thus fewer confounding factors to evaluate. Another advantage of applying the comparative method in an intraspecific context is that populations may still reside in the habitat where the phenotype of interest originated. The targeted character may also still be under selection and thus the effect of selection can be directly measured (Foster and Cameron 1996).

A potential problem of intraspecific comparative studies, however, is that habitat can be confounded by geography. This can be a problem because populations often form geographically defined phylogenetic groups (Avise 1994). If these geographic groups correspond to different habitat types then it is difficult to determine if phenotypes correlated with habitat are due to current natural selection or history. When geography is confounded by habitat within a species, there is no opportunity to document independent origins because in essence only a single population is found in each habitat type. In order to document multiple independent origins of a trait within a single species, it would thus be necessary for that species to form multiple independent populations in the same habitat type. Such a scenario could exist when two habitat types are found in sympatry (geographic proximity) at many different locations. For instance, fishes found in multiple lakes often use different habitat types within each

lake (eg. benthic and limnetic ecotypes of sticklebacks, Rundle and Nagel 2000; dwarf and normal ecotypes of whitefish, Bernatchez et al. 1999).

By examining the population genetic structure of species that use sympatrically distributed habitats, it is possible to determine if populations of the same habitat are closely related to each other, or if each replicate population is independent. If replicate populations from the same habitat type are closely related then their phenotypic similarities are due to common ancestry and/or gene flow. If instead, replicate populations from the same habitat are not closely related they are evolutionarily independent of each, and natural selection is probably driving the evolution of habitat-associated characters. There is evidence for both patterns. For example in an amphipod and and a galling insect , among others, population structure is strongly correlated with habitat type, suggesting a single switch in habitat-use and associated characters (e.g. Stanhope et al. 1993, Brown et al. 1996, Via 1999, Reusch et al. 2001, Bockelman et al. 2003, Simon et al. 2003). In contrast, in trout and walking sticks there is no correlation with of population structure with habitat type, suggesting that natural selection maintains traits correlated with habitat-use (Docker and Heath 2003, Nosil and Crespi 2004).

Possibly the best way to identify the interaction of history and natural selection is to examine intraspecific variation in the context of other closely related species. For instance, Taylor and McPhail (2000) evaluated population genetic variation of the sympatric species pairs of freshwater sticklebacks, *Gasterosteus aculeatus*, within the context of the marine ancestors. They found that the existence of sympatric species pairs probably resulted from historical double invasions of freshwater lakes, but that

the evolution of similar phenotypes across lakes was due to natural selection acting independently in each lake.

The desert spider, *Agelenopsis aperta* Gertsch, offers an ideal opportunity to evaluate the relative roles of history and natural selection, and their potential interaction, in shaping habitat-associated characters. This spider inhabits arid regions in the southwestern United States and northern Mexico, but also uses riparian habitat that is patchily distributed throughout the arid landscape. Genetically-based behavioral differences between arid and riparian habitats have been described for *A. aperta* (see below). Riparian patches provide replicate populations that are sympatric with arid habitats, permitting examination of the roles of history versus natural selection in the evolution of habitat-associated traits.

Most of the 12 other described *Agelenopsis* (Geibel) species appear to be restricted to more mesic habitats but a few – *A. aleenae*, *A. spatula*, and a potential new species – use arid habitats (see below). In addition, *A. aleenae* and *A. spatula* use both riparian and arid habitats, similar to *A. aperta*. These species are ideal for evaluating the role of history in the evolution of habitat-use. In order to determine the ancestral habitat type for each of these species, their intraspecific phylogenies should be rooted with other *Agelenopsis* species. In addition, the number of times habitat use has changed in the genus as a whole, as well as within *A. aperta*, can be evaluated by considering the interspecific phylogeny.

In this study, I examined intraspecific variation of habitat-use within *A. aperta* in the context of the habitat-use of two other closely related species to evaluate the potential interaction of history and natural selection in shaping the evolution of habitat-

use and associated behavioral characters. Using mitochondrial DNA sequence data, I reconstructed relationships within and among *Agelenopsis* species and mapped the history of habitat-use in *Agelenopsis*. My objectives were to 1) determine the ancestral habitat-type for *A. aperta* and for *Agelenopsis*; 2) determine how many switches in habitat-use have occurred within the genus; 3) evaluate whether riparian *A. aperta* populations are linked by common ancestry or are independent populations; and 4) determine the relative influence of geography versus habitat on the population structure of *A. aperta* and another arid-living species, *A. aleenae*.

Background

Behavioral Adaptations in A. aperta

Agelenopsis aperta is found in a wide range of arid habitats in the southwestern United States and northern Mexico (Figure A1, all figures and tables for this Part are found in the Appendix). Embedded within these deserts are numerous springs, streams, and rivers that support riparian habitat patches. The major source of mortality differs between the two habitat types. Arid populations are limited by prey availability (Riechert 1979, 1993a). In contrast, riparian populations are not food limited but instead experience high mortality from bird predation (Riechert and Hedrick 1990, 1993). *Agelenopsis aperta* exhibits a suite of genetically influenced behavioral traits that are differentially favored in the two environments. Arid-land spiders display high levels of aggression towards competing spiders, show a quick response to prey hitting

their webs, attempt to capture most insects that land on their webs, and quickly return to foraging after exposure to a simulated predator. On the other hand, riparian spiders are less aggressive towards conspecifics, slow to respond to prey cues, feed more selectively, and remain hidden for longer periods of time in response to a simulated predator (Hedrick and Riechert 1989, Riechert and Maynard Smith 1989, Riechert and Hedrick 1990).

Three locations of *A. aperta* populations have been extensively studied (see Riechert 1999 for review): Carrizozo, New Mexico (site 19, see Figure A1)—continuous desert grassland habitat; Hye, Texas (site 1)—continuous riparian habitat; and at the Southwestern Research Station in Portal, Arizona (site 22)—a riparian island within arid habitats. At the Arizona site, *A. aperta* populations are found in extremely dry woodland, extensive stretches of desert scrub and a thin strip of riparian habitat along a stream.

Although spiders in the Arizona riparian habitat are less aggressive than arid-land spiders from New Mexico, they show higher levels of aggression, less selectivity for food, and are bolder after a predator cue than spiders from continuous riparian habitat in Texas (Riechert 1993a). Experimental evidence indicated that the intermediate phenotypes identified in the local riparian population in Arizona result from gene flow from the desert populations countering selection pressure in the riparian habitat (Riechert 1993b). Allozyme analysis of the Arizona riparian local population and five surrounding local arid populations indicated high levels of gene flow between the riparian patch and surrounding arid habitat (Riechert 1987). Drift fence and mating censuses conducted with a marked population in Arizona showed

that gene flow between arid and riparian habitats occurs annually and that it is largely due to unidirectional movement from arid habitats into the riparian strip (Riechert et al. 2001). Reciprocal transplant experiments demonstrated that riparian spiders fail to survive to reproduction in the arid habitats. In contrast, spiders with the aggressive arid phenotype survive in riparian habitat at about 50% the rate of native riparian individuals (Riechert and Hall 2000).

Aggressiveness in *A. aperta* is governed by at least two loci: the tendency to attack (level of aggression) which is sex-linked and dominant, and the tendency to flee (level of fear) which is autosomally inherited and recessive (Riechert and Maynard Smith 1989). F1 hybrid progeny of riparian and arid spiders are extremely aggressive because they have not only a high tendency to attack but also a low tendency to flee. A Mendelian model of F1 riparian and arid hybrids, F2 hybrids, and various backcrosses, shows that hybridization of arid and riparian spiders should lead to some intermediate behavioral types and some extremely fearful or aggressive behavior types (Riechert et al. 2001). This model of the genetic basis for aggression in *A. aperta* is supported by experimental crosses and the fact that spiders from the Arizona riparian habitat, which experience gene flow from arid habitats, exhibit a range of intermediate behavioral types and some extreme behavioral types. From observations of all matings in two separate years within a 3-hectare riparian area in Arizona, Riechert et al. (2001) found that 21% percent of all females were so fearful that they ran from potential mates and 5% attacked every male they encountered in both years (Riechert et al. 2001). Based on laboratory crosses, these "chickens" and "cannibals" appear to be the product of backcrosses among F1 hybrids (Riechert et al. 2001). Thus, there appears to be a

postzygotic barrier to gene flow between arid and riparian adapted populations. Hybrid and backcross progeny of these locally adapted populations often fail to mate at all because of having levels of aggressiveness at either extreme of the continuum.

Habitat-use in other Agelenopsis Species

The genus *Agelenopsis* consists of 13 described species distributed throughout North America (see Platnick 2004 for review of current taxonomy and distributions). Although *A. aperta* is the only *Agelenopsis* species found in the driest regions of North America, a few other species overlap with *A. aperta* in the eastern portion of its range (see Figure A2 in Ayoub et al, in review). These species – *A. aleenae*, *A. spatula*, and an undescribed species, referred to here as "*A. novum*" (see Ayoub et al., in review) - appear to also use arid habitats (areas that receive an average of less than 50cm annual precipitation; Daly 2000). Similar to *A. aperta*, *A. spatula* and *A. aleenae* are also found in riparian habitats within the arid framework (Table A1, Figure A2).

Most of the other known *Agelenopsis* species appear to be restricted to more mesic conditions. I consider an area to be mesic if it receives on average greater than 64cm annual precipitation (Daly 2000), or if it is a riparian habitat if found within areas drier than that. For instance, examination of the species' ranges (see Figure A2 from Ayoub et al., in review) shows that *A. kastoni*, *A. emertoni* and *A. naevia* are all restricted to the eastern United States and are typically found in association with eastern deciduous forest (Paison 1997; current collections - Table A1). In addition, the ranges of *A. potteri* and *A. pensylvanica* indicate an association with mesic areas of the eastern

United States, northwestern United States and southwestern Canada. Although the range of *A. pensylvancia* extends into the more arid regions of Kansas and Colorado, the western populations collected for this study were found in wooded riparian habitats (Table A1). *Agelenopsis oregonensis* has only been recorded from the extremely mesic (greater than 200cm average annual precipitation; Daly 2000) coniferous woodlands of the Cascade Mountains (Paison 1997). Two species, *A. utahana* and *A. actuosa*, found predominantly in northern United States and Canada are also associated with moderate to high levels of precipitation or riparian habitats (Paison 1997; current collections Table A1).

Habitat affinities are harder to identify for the two remaining *Agelenopsis* species, *A. oklahoma* and *A. longistylus*. *Agelenopsis oklahoma* occupies much of the midwestern United States, an area that receives a geographically variable amount of rainfall but much of which receives an average of less than 50cm annual precipitation. From examination of locality records, however, it appears that *A. oklahoma* does not occupy arid habitats that are frequented by *A. aperta*, *A. spatula* and *A. aleenae*. For example, in collections made for this study, *A. oklahoma* was found only in riparian sites (Table A1). This species is often associated with tallgrass prairie or shrub steppe (Muma and Muma 1949, Paison 1997).

It is also difficult to define habitat-use for *A. longistylus*. Most of its range is within arid New Mexico (Figure A2 in Ayoub et al, in review). Although *A. longistylus* webs can be found in areas with very little vegetative cover (Gertsch and Riechert 1976; personal observations), each of the known populations is associated with water (Table A1). I classify *A. longistylus* as a riparian species but further collections and

experimental evidence may indicate that this species is in fact able to use fully arid habitats.

Although phylogenetic relationships among agelenid genera have not been published, the most probable sister genus to *Agelenopsis* is *Barronopsis* Chamberlin and Ivie (J. Spagna, personal communication). *Barronopsis* was once considered a subgenus of *Agelenopsis* (Chamberlin and Ivie 1941) but was elevated to generic status by Lehtinen (1967). Mitochondrial sequences support reciprocal monophyly of the two genera (Ayoub et al., in review). Knowing what habitat the sister genus uses is helpful because it allows determination of the ancestral condition for *Agelenopsis*. Little experimental work has been done with *Barronopsis* but the combined ranges of the described species appear to fall within mesic habitats – southeastern United States, Cuba, and the Bahama Islands (Platnick 2004).

Methods

Collections

Sampling for this study included 12 of the 13 described *Agelenopsis* species, as well as the undescribed species, *A. novum* (see Ayoub et al., in review). Most specimens were collected live in the field but specimens of *A. potteri* and *A. oregonensis* from Oregon, California, and Canada were sent to me preserved in 95% or 100% ethanol (see Table A1 for list of sampling localities). I included multiple individuals of each species from multiple populations when possible.

Because *A. aperta*, *A. aleenae*, and *A. spatula* each use both arid and riparian habitats, these species were sampled more extensively than the other species. *Agelenopsis* individuals were collected from 47 sites spanning most of the previously known ranges of *A. aperta* and *A. spatula*. I collected spiders between June 2000 and August 2003 (Figs. A1-A3, Table A1; some sites yielded other species than *A. aperta*, *A. aleenae*, and *A. spatula*). Initially, sampling focused on the three sites for which A. aperta had been studiend in detail (reviewed in Riechert 1999). At the Arizona site (site 22) described above, collections came from the riparian patch, the adjacent arid woodland, and an arid cactus scrub habitat approximately 10km away from the riparian patch. I employed a geographic sampling scheme analogous to the Arizona site for the continuous arid site in New Mexico (site 19) and the continuous riparian site in central Texas (site 1; this site yielded both *A. aperta* and *A. aleenae*). I then sampled seven more sites, at each of which I collected from a riparian patch, an adjacent arid habitat, and an arbitrarily chosen distant population 4 - 15 km away from the riparian patch. The arbitrarily chosen distant populations were collected from arid habitats except at two sites in which the distant population was collected from riparian habitat, sites 24 and 33. This yielded a total of eight replicate riparian patches and surrounding arid habitats. Seven of these sites yielded *A. aperta* individuals (sites 2,3,13,22,24,28,33), and one site yielded *A. spatula* individuals (site 43). In addition, at site 24, the distant population yielded *A. oklahoma* individuals and was thus excluded from population genetic analyses for *A. aperta*. The different sampling locations within a site will be referred to as "local populations", such that each of the cluster sites included 3 local populations, except site 24, which included two local populations.

At two sites I additionally sampled three local populations evenly distributed along transects (40km at site 1, and 34km at site 4; Table A1). At all other sites samples came from a single local population. This additional geographic sampling allowed for an evaluation of overall patterns of genetic variation within *A. aperta, A. aleenae*, and *A. spatula,* and an assessment of whether molecular genetic differences among sites within *A. aperta* were due to isolation by distance or pure fragmentation.

When sampling in the arid regions of the southwestern United States, I identified riparian habitat by the presence of water (usually a stream or river but in one case a spring-fed lake), and the presence of riparian vegetation, especially cottonwood trees (*Populus fremontii*). Any sites that fell within areas receiving an average of greater than 63cm annual precipitation I considered mesic (see Daly 2000).

For each of the 47 sites within the ranges of *A. aperta* and *A. spatula*, I collected 15-30 individuals per local population irrespective of the age of individuals. Individuals collected as juveniles reached maturity in the laboratory, at which time I made species determination using genitalic characters. I removed 1-8 legs, which I stored at -80°C to be used for DNA extractions, and preserved the remainder of each specimen in 75% ethanol. Voucher specimens for *A. aperta* were deposited at the Colorado Museum of Natural History (Denver, Colorado) or the National Collection of Acari and Arachnids (UNAM; Mexico City, Mexico); voucher specimens for all other species were deposited in the Arachnid Collection at California Academy, San Diego.

DNA Extraction and Sequencing

DNA extractions and sequencing methods were reported in Ayoub and Riechert (2004) and Ayoub et al. (in review). To evaluate overall within and among species variation, I sequenced a 700 base pair (bp) portion of the mitochondrial gene, cytochrome oxidase I (COI) for a total of 576 *Agelenopsis* individuals (474 *A. aperta*, 43 *A. aleenae*, 16 *A. spatula*, and 43 of the other species combined). To gain a better understanding of among species relationships I also sequenced a 500bp portion of the mitochondrial 16S ribosomal RNA gene for a subset of individuals from each species (26 individuals). In all interspecific analyses I also included individuals in the genus *Barronopsis*, as well as two other agelenid genera, *Hololena* and *Novolena*, which were used as outgroups (see Ayoub et al., in review).

For *A. aperta*, *A. aleenae*, and *A. spatula*, I sequenced at least five individuals per local population except in cases where I had too few specimens. I sequenced 10 individuals per local population whenever possible except for sequencing 5 individuals per population from the area containing sites 19-24 (Fig. A1). In my preliminary studies, I failed to find sequence variation within this geographic region for *A. aperta* (Ayoub and Riechert 2004).

Phylogenetic and Network Analyses

I performed maximum likelihood (ML), maximum parsimony (MP) and Bayesian analyses to assess phylogenetic relatedness among *A. aperta* COI haplotypes

(see Ayoub and Riechert 2004) and among all *Agelenopsis* COI and 16S haplotypes (see Ayoub et al, in review).

To test for single origins of habitat-use within *A. aperta* and *A. aleenae*, I used the Shimodaira-Hasegawa test (Shimodaira & Hasegawa 1999) carried out in PAUP*v4.0b10 (Swofford 2002) to test if monophyly of haplotypes found in either riparian or arid habitats was significantly less likely than the ML trees. For each species separately, I searched for ML trees constraining haplotypes from one or the other of the habitat types to be monophyletic. Likelihood searches were heuristic with 10 replications of random stepwise addition sequences with TBR (tree bisection reconnection) branch swapping. For *A. aperta*, I did not constrain the habitats to be reciprocally monophyletic because many haplotypes were found in both habitat types. I did not perform this test on *A. spatula* because I found only one haplotype at the two sampled *A. spatula* sites.

To examine the probable ancestral habitat type for the species that use both arid and riparian habitats I mapped habitat-use onto the MP, ML and Bayesian interspecific trees, using a parsimony criterion in MacClade ver. 4 (Maddison and Maddison 2000). I coded each species as to whether its members were found only in riparian/mesic habitat, only in arid habitat, or in both habitats.

Because intraspecific data often do not conform to the assumption of bifurcating evolutionary lineages that is required by phylogenetic methods (Posada and Crandall 2001), I also created a network of *A. aperta* haplotypes and a network of *A. aleenae* plus *A. spatula* haplotypes. Networks allow ancestral haplotypes to coexist with descendant haplotypes and permit reticulation among haplotypes. I used a statistical

parsimony method (Templeton *et al*. 1992) carried out in TCS v1.13 (Clement *et al*. 2000) and a minimum spanning tree method (Rohlf 1973) carried out in ARLEQUIN ver 2.000 (Schneider *et al*. 2000) to create networks.

Population Genetic Analyses

In addition to determining if habitat type played a role in the phylogenetic history of *A. aperta* populations, I also determined if habitat type was correlated with population genetic structure. I looked at three aspects of population structure. First I examined if riparian populations were more closely related to each other than to surrounding arid populations, or if riparian patches could be considered independent populations. Populations from the same habitat could be more closely related to each other than to the alternate closely located habitat if each habitat shared a common history or if members of populations exchanged genes more readily with members of other populations from the same habitat. Second, given that there is strong selection against arid or riparian adapted individuals in the alternate habitat type (Riechert and Hall 2000), I examined if there was any evidence for restricted gene flow between riparian populations and adjacent arid populations. Finally, I examined whether there was any difference in levels of genetic diversity between riparian populations and surrounding arid populations. Higher genetic diversity in riparian populations might be expected for two reasons. First, riparian populations maintain population densities an order of magnitude greater than arid populations (Riechert, personal observations). Populations with larger effective population sizes tend to maintain higher levels of

genetic diversity. Second, if riparian and arid populations have separate phylogenetic histories, then successful movement of arid-land spiders into riparian habitats could cause mixing of phylogenetically diverged lineages, thus raising genetic diversity in the riparian population.

To determine if *A. aperta* riparian populations were closely related to each other or if they were independent of each other I performed an analysis of molecular variance (AMOVA; Excoffier et al. 1992). I included in the analysis only the seven sampling sites from which I had collected a riparian population and an adjacent arid population. These sites also included a third population 4-15 km away from the adjacent populations, except site 24 in which the distant population had yielded *A. oklahoma*. This third distant population was arid except at site 33, in which the distant population was riparian. By limiting the AMOVA to the sites that included adjacent riparian and arid populations, I could explicitly differentiate the effect of geography versus the effect of habitat on relationships among populations. I carried out two separate AMOVAs. In the first, I grouped local populations according to habitat type. In this case there were two groups, with 8 local populations in the riparian group, and 12 local populations in the arid group. In the second AMOVA, I grouped local populations according to sampling site; there were seven such groups and each contained two or three local populations. For each AMOVA, I divided genetic variation into three hierarchical levels: variation within a local population (F_{ST}), variation among local populations within a group (F_{SC}), and variation among groups (F_{CT}).

I did not perform an AMOVA on *A. aleenae* populations because I did not sample any riparian patches surrounded by arid habitats for *A. aleenae*, and thus habitat

was confounded by geography. Although, I sampled *A. spatula* from a riparian patch and surrounding arid populations, as mentioned above, *A. spatula* showed no sequence variation.

In order to see if any riparian population and adjacent arid population exhibited evidence for restricted gene flow, I calculated pairwise F_{ST} between riparian populations and adjacent arid populations. The significance of F_{ST} values was calculated by 1000 Monte Carlo random permutations of haplotypes among local populations. I evaluated significance at an initial alpha = 0.05 but adjusted this alpha for multiple comparisons using a sequential Bonferroni correction (Sokal and Rohlf 1995). I initially calculated F_{ST} using TrN corrected sequence divergence (Tamura and Nei 1993; see Ayoub and Riechert, 2004). However, these values should reflect historical signatures of restricted gene flow. If restricted gene flow occurred recently then it is unlikely that a sufficient number of mutations would have accumulated to distinguish populations. Thus, I also calculated F_{ST} using conventional methods that simply treat each haplotype as a separate allele without reference to the number of mutations differentiating haplotypes.

To further examine the effects of habitat on population structure in *A. aperta*, I calculated pi (the average number of pairwise substitutions) for local populations within each site containing a riparian patch and surrounding arid habitat. I then tested if riparian populations had higher diversity than the adjacent arid or the distant populations using an analysis of variance (ANOVA) carried out in JMP ver. 5 (SAS Institute, Inc.).

If population genetic structure is not correlated with habitat type, then geography may be a more important determinant of structuring. The AMOVA discussed

above, which grouped local populations according to sampling site, is one way to examine if there is geographic structuring of genetic variation. In addition to this AMOVA, I examined if isolation by distance plays an important role in *A. aperta* or *A. aleenae* population structure. I plotted pairwise F_{ST} values against pairwise geographic distance for each species separately and tested for a significant relationship between the two using the Mantel test (Mantel 1967). A significant positive relationship indicates isolation by distance (Hutchison and Templeton 1999). I also tested if the absolute values of the residuals of the regression of F_{ST} on geographic distance increased with increasing geographic distance using a Mantel test. Under an equilibrium model of isolation by distance these residuals should increase with increasing geographic distance (Hutchison and Templeton 1999). For *A. aperta*, I performed the Mantel tests for all local populations and also within geographically defined clades (see Figure A1). However, I did not test for isolation by distance in the clade restricted to west of the Sierra Nevadas because only two populations from this region had been sampled. All population genetic parameters including AMOVA, F_{ST}, pi and Mantel tests were calculated using ARLEQUIN ver 2.000 (Schneider *et al*. 2000).

Results

At least 535bp of COI sequence data was obtained for each of the 576 *Agelenopsis* individuals sequenced. Ninety different *Agelenopsis* haplotypes (or sequence variants) were identified: 51 in *A. aperta*, 16 in *A. aleenae*, and one in *A. spatula*. These haplotypes were deposited in GenBank (accession numbers published in

Ayoub and Riechert 2004; and Ayoub et al., in review). The distribution of these haplotypes among populations is presented in Table A1.

Phylogenetic and Network Analyses

Phylogenetic and network analyses for *A. aperta* COI haplotypes were described in Ayoub and Riechert (2004). In brief, three major clades were uncovered: 1) west of the Sierra Nevadas, 2) between the Sierra Nevadas and Rocky Mountains, and 3) "eastern" (see Figure A1). Each clade contained haplotypes from riparian and arid habitats and many haplotypes were found in both habitat types (Figure A1). Constraining riparian or arid haplotypes to be monophyletic resulted in significantly less likely trees (arid constrained to be monophyletic: $\partial = 163.3$, $p < 0.001$; riparian constrained to be monophyletic: $\partial = 162.4$, $p < 0.001$).

Phylogenetic analyses of *Agelenopsis* COI plus 16S haplotypes showed *A. aleenae* and *A. spatula* to form a well-supported monophyletic group (100% posterior probability and 98% parsimony bootstrap support), with *A. aleenae* possibly being paraphyletic with respect to *A. spatula* (Ayoub et al., in review). Within this *A. aleenae - A. spatula* species group, phylogenetic and network analyses identified four geographically localized clades (Figure A2). TCS was able to connect COI haplotypes differing by up to 9 mutations with 95% confidence that multiple substitutions per site had not occurred. This resulted in two *A. aleenae* sub-networks. How these two sub-networks and the *A. spatula* haplotype connected was estimated by extending the connection limit in TCS to 15 mutations. While this estimate represents less than 95%

confidence that multiple substitutions had not occurred along those connections, it can be viewed as a "best guess" of how the sub-networks and *A. spatula* connect.

Populations of *A. spatula* from both riparian and arid habitats exhibited a single haplotype. In contrast to *A. spatula* and *A. aperta*, populations of *A. aleenae* from arid habitats formed a cluster in the network. Accordingly, the *A. aleenae* ML tree indicated monophyly of arid haplotypes. Constraining riparian haplotypes to be monophyletic was less likely but not significantly so ($\partial = 3.27$, $p = 0.148$). Note, however, that sampling of *A. aleenae* confounded habitat with geography such that all arid populations fell within New Mexico.

Interspecific MP, ML, and Bayesian analyses of COI alone and COI plus 16S consistently identified *A. aperta*, *A. aleenae*, *A. novum*, and *A. spatula* as forming a monophyletic group (Figure A4). Although this group was not always well-supported by the mitochondrial data (i.e. greater than 50% bootstrap support or 95% posterior probability), there are genitalic characters that support monophyly of this group (see Ayoub et al., in review). Thus, when mapping habitat-use onto the MP, ML, or Bayesian trees, parsimony identified a single origin of arid-living, with *A. aperta, A. aleenae*, and *A. spatula* maintaining an ability to use riparian habitat. Because I only collected *A. novum* from a single site, I do not know if this species can also use riparian habitat in addition to the arid habitat in which it was found.

Population Genetic Analyses

A. aperta - The AMOVA showed that a significant amount of the genetic variation, 64.7%, was distributed among sampling sites (Table A2; $F_{CT} = 0.647$, $p < 0.001$), indicating that geographic location plays a dominant role in structuring genetic variation. In contrast, habitat type did not make a significant contribution to the distribution of genetic variation ($F_{CT} = -0.055$, $p = 0.84$), indicating that genetic variation is randomly distributed between habitat types. In addition to the significant structuring due to sampling site, a significant amount of variation was distributed among local populations within groups. When local populations were grouped according to habitat type, approximately 75% of the variation was distributed among local populations within each habitat type (Table A2; $F_{SC} = 0.71$, $p < 0.001$). This again reflects the importance of geographic locality in structuring genetic variation because most of the local populations within a habitat type were located at different geographic sampling sites. When local populations were grouped according to sampling site, approximately 7% of the genetic variation was due to local populations within each sampling site (Table A2; $F_{SC} = 0.21$, $p < 0.001$). Although this represents a significant amount of variation attributable to the small geographic scale of local populations within each sampling site, it is far less than the amount of variation attributable to the large geographic scale of different samping sites.

In addition to the lack of evidence from the AMOVA for an effect of habitat type on population genetic structure, I found no evidence for restricted gene flow between riparian patches and adjacent arid populations. Adjacent populations shared

many if not all haplotypes and pairwise F_{ST} values were always low (mean = 0.013, range = -0.1 – 0.18) and non-significant (Table A3). Pairwise F_{ST} typically increased when comparing one of the adjacent populations to the third population located between 4 and 15 km away from the adjacent populations. Occasionally these values were significant (Table A3). The same geographic pattern of adjacent populations having a very low F_{ST} and the more distant population having a higher F_{ST} was also apparent for the continuous riparian site (site 1, Table A3).

When comparing nucleotide diversity among each of the three local populations within a cluster site, I found no significant difference in the mean pi of riparian patch, adjacent arid or distant populations ($df = 2$, $F = 0.0252$, $p = 0.98$).

My data indicate that habitat type has no measurable effect on population genetic structure, but that geographic locality plays an important role in shaping genetic diversity. As discussed in Ayoub and Riechert (2004) historical events have probably led to the primary regional phylogenetic differentiation of mitochondrial DNA diversity. However, within-region population structuring is also strong. Each phylogenetically defined region displayed significantly high overall F_{ST} (eastern clade: $F_{ST} = 0.68$, $p < 0.001$; intermontane clade: $F_{ST} = 0.62$, $p < 0.001$; west of Sierra Nevadas clade: $F_{ST} = 0.86$, $p = 0.01$).

Restricted gene flow with isolation by distance may partially explain this significant genetic structuring within regions. There was a significant increase in F_{ST} with geographic distance among all local populations, although only 13% of the variation in F_{ST} could be explained by geographic distance ($r^2 = 0.13$, $p < 0.001$). As mentioned above, however, past fragmentation probably caused regional genetic

structuring. Thus, the increase in F_{ST} with increasing area could be a side effect of this past fragmentation. However, I found the same relationship between F_{ST} and geographic distance within the eastern and intermontane clades as well (eastern clade: $r^2 = 0.093$, $p < 0.001$; intermontane clade: $r^2 = 0.24$, $p < 0.001$). These effects remained significant even when adjacent local populations were treated as a single population (eastern clade: $r^2 = 0.086$, $p < 0.001$; intermontane clade: $r^2 = 0.14$, $p = 0.007$). However, plots of F_{ST} versus geographic distance for each region show a huge scatter in F_{ST} values (Figure A5). In isolation by distance models that are in equilibrium, the scatter of F_{ST} should increase with increasing geographic distance. This is because gene flow should swamp genetic drift at close locations but drift should swamp gene flow at distant locations (Hutchison and Templeton 1999). In the case of mitochondrial DNA variation within *A. aperta*, however, I see an opposite trend with the scatter actually decreasing with increasing geographic distance. In both regions the relationship between degree of scatter and geographic distance was significantly negative (eastern clade: $r^2 = 0.05$, $p = 0.001$; intermontane clade $r^2 = 0.26$, $p < 0.001$). Thus, although F_{ST} generally increases with geographic distance, *A. aperta* mitochondrial variation does not fit an equilibrium model of isolation by distance.

<u>*A. aleenae*</u> - Because habitat was confounded with geography in *A. aleenae* an AMOVA could not be applied to examine the relative amounts of genetic variation distributed between habitat type versus geographic locality. However, the role of geography on population structure in *A. aleenae* could be examined. As mentioned above, 3 geographically localized lineages were observed within *A. aleenae*. In addition, each population of *A. aleenae* exhibits a single haplotype or a group of very

closely related haplotypes (Table A1, Figure A2). The pairwise F_{ST}s among *A. aleenae* populations are significantly positively correlated with geographic distance ($r^2 = 0.56$, $p = 0.023$; I included only populations for which three or more individuals were sequenced). The residuals, however, are not significantly correlated with geographic distance ($r^2 = 0.088$, $p = 0.10$), suggesting that *A. aleenae* does not fit an equilibrium model of isolation by distance. The relationship between geographic distance and F_{ST} appears to be entirely due to the similarities of the local populations within the central Texas site (site 1). F_{ST} is above 0.65 for any populations more distant than 16km. In fact, the relationship between geographic distance and F_{ST} is negative if local populations at site 1 are pooled, although not significantly so ($r^2 = 0.416$, $p=0.18$). Unfortunately, I do not have any sampling points between 16 and 350km, making it impossible to rule out the possibility of isolation by distance contributing to the pattern of geographically localized clades in *A. aleenae*.

Discussion

History of Habitat-use in Agelenopsis

The phylogenetic results presented here indicate that riparian or mesic habitat association is ancestral for the genus *Agelenopsis* (Figure A4). The basal *Agelenopsis* species use riparian or mesic habitats and the probable sister genus, *Barronopsis*, also appears to be restricted to mesic habitats. In addition, there appears to have been a single switch to arid-living within *Agelenopsis*. All the species that use arid habitats, *A.*

aperta, A. aleenae, A. spatula and *A. novum*, consistently form a monophyletic group according to mitochondrial data (Figure A4). Genitalic characters also support monophyly of this group (Ayoub et al., in review).

In order to conclude, however, that only a single switch to arid-living occurred, the other *Agelenopsis* species nust be restricted to more mesic conditions. I feel confident that this is the case for species experimentally shown to be dependent on humid conditions such as *A. naevia* (Jones 1941) and *A. potteri* (Earnshaw 1973), but I feel less confident about other species, especially *A. oklahoma* and *A. longistylus*. These two species inhabit arid regions, areas receiving less than 50cm average annual precipitation, but I only found them associated with water and riparian vegetation. Further sampling and detailed evaluation of habitat as well as experimental measures of physiological tolerance to arid conditions are needed to resolve whether or not these species are restricted to riparian habitats. However, even if *A. oklahoma* and *A. longistylus* are found to use arid habitats, it does not necessarily negate the conclusion of a single origin of arid-living in *Agelenopsis*. Phylogenetic analyses consistently placed *A. oklahoma* sister to the arid species group except when *A. longistylus* was placed sister to the arid species group. In addition, I could not reject monophyly of *A. longistylus, A. aperta, A. aleenae, A. spatul*a and *A. novum* in the parsimony analysis of COI plus 16S, which did not place *A. longistylus* sister to the arid species (constrained MP trees 2 steps longer than unconstrained trees; $t = 0.38$, $p = 0.71$, Kishino-Hasegawa test (Kishino and Hasegawa 1989)).

If there was in fact a single origin of arid-living in *Agelenopsis*, it appears that the ability to use riparian habitats was retained in *A. aperta*, *A. aleenae*, *A. spatula* and

possibly *A. novum*. Southwestern North America was predominantly arid by about four to five million years ago (Axelrod 1985). If the common ancestor of the arid-living *Agelenopses* arose after the region became arid, it is likely that each species spread through arid habitats and reinvaded riparian patches when they were encountered. In the absence of a calibrated molecular clock for *Agelenopsis,* it is difficult to date the common ancestor of the arid species. However, a rough idea of species ages can be formed based on an estimated arthropod mitochondrial mutation rate of 2.3% divergence/million years (Brower 1994). The maximum sequence divergence between any two haplotypes within the arid species group was 6.6% for COI+16S and 8.8% for COI alone (TrN corrected). These levels of sequence divergence translate to a maximum divergence time of sequences between 2.9 and 3.8 million years ago. Considering that gene sequences begin to diverge before divergence of populations or species (Nei & Li 1979; Wilson *et al*. 1985; Edwards & Beerli 2000), divergence among arid species almost certainly occurred well after aridification of the southwest was complete.

Thus, it is likely that the each of the arid species spread through mostly arid habitat and invaded riparian patches as they were encountered. In *A. aperta*, population genetic analyses indicate that sampling sites, which include a riparian population and adjacent arid populations, are significantly differentiated (Table A2), meaning that each riparian patch is likely to be evolutionarily independent of any other riparian patch. Within *A. aperta* at least, then, it appears that there have been multiple, independent invasions of riparian patches.

It is also possible that arid-living arose independently in each of the four arid-land species. For this to be the case, ancestral riparian populations of each species would have had to spread through only riparian habitats. This is plausible because egg sacs of *A. aperta* can float on water (Riechert, personal observation). Today the riparian waterways of the southwest US are not all connected but during Pleistocene glacial maxima riparian areas expanded and could feasibly have been interconnected (Williams et al. 1998). If each of the arid-living species independently evolved the ability to use arid habitats from a riparian-living ancestor, then each of these species should show monophyly of arid populations. The mitochondrial data, however, strongly rejected monophyly of *A. aperta* haploytpes found in arid populations. However, gene flow between the arid and riparian populations could have swamped the historical effect.

Using a strict parsimony criterion, there are fewer changes in habitat-use along the *Agelenopsis* phylogeny if arid-living arose independently four times than if a single switch to arid-living took place followed by multiple switches back to riparian-living. However, it would probably be very difficult for a riparian spider to invade arid habitat – the survivorship of native riparian *A. aperta* transplanted to arid habitat was zero (Riechert and Hall 2001). On the other hand, it is comparatively easy for a native *A. aperta* arid-land spider to invade riparian habitat – only 55% mortality was found in reciprocal transplant experiments (Riechert and Hall 2001). In addition, in drift fence experiments, riparian spiders were never found to leave their native riparian patch, but arid spiders were often found entering a riparian patch (Riechert et al. 2001). These findings indicate that a single switch to arid-living with multiple reinvasions of riparian patches is biologically more plausible then four independent origins of arid-living.

Implications for the Evolution of Habitat-associated Behaviors

The proposed history of habitat-use in *Agelenopsis* has important implications for the evolution of habitat-associated behaviors. Although behavioral data are not presented here, these data have been collected and I predict that new populations of *A. aperta* as well as populations of other *Agelenopsis* species will show similar habitat-associated levels of aggression and fear as have already been documented for the three studied sites of *A. aperta*. If this prediction is supported, then history and natural selection have interacted to shape adaptive behaviors in *Agelenopsis*.

Because each riparian patch population in *A. aperta* is significantly differentiated, the observation of similar behaviors in similar habitats would be strong evidence that natural selection has repeatedly caused the evolution of riparian-associated behaviors. However, because riparian-living is ancestral for the genus, there is probably an underlying historically driven basis for the riparian phenotype. Perhaps arid populations maintain a low frequency of the less aggressive, more fearful behavioral alleles, which are selected for when reinvading riparian patches. Identification of behavioral genes should show the same loci involved in differential behaviors in *A. aperta* populations and probably also in *A. aleenae* and *A. spatula* populations.

Population Genetic Structure

I failed to detect any effects of habitat type on population structure in *A. aperta*. Spiders from neither habitat formed a monophyletic group; adjacent arid and riparian patches were never significantly differentiated; and levels of genetic diversity did not differ among local populations within sampling sites. The same probably holds true for *A. spatula*, but I failed to find sequence variation within *A. spatula*. Although arid populations of *A. aleenae* clustered in the phylogenetic trees and the network, habitat was confounded by geography for *A. aleenae* and I predict a pattern similar to *A. aperta* would be observed if adjacent riparian and arid populations were sampled.

Even though riparian populations of *A. aperta* tend to maintain population densities that are an order of magnitude greater than arid populations (Riechert, personal observations) mitochondrial diversity was not greater in the riparian patch than in the surrounding arid populations within a site. There was also no evidence of restricted gene flow between adjacent riparian and arid patches; local populations within these sites showed similar patterns of differentiation and diversity as the continuous arid and continuous riparian sites (Table A3). Even though there is a fitness cost to hybridization between riparian and arid phenotypes and there is strong selection against either phenotype in the other habitat, there appears to be no restriction on gene flow in any of the riparian patches sampled for this study. These findings do not preclude the possibility that some restriction on gene flow has occurred in the recent past. The close similarity between adjacent sites could be due to incomplete lineage sorting (Neigel and Avise 1986). However, given the similarity of the adjacent

populations in mitochondrial variation, the direct observations of gene flow at site 22 (Riechert et al. 2001), and the lack of differentiation between adjacent populations according to allozyme analysis at site 22 (Riechert 1989), gene flow is probably currently occurring between each riparian patch and its surrounding arid habitat sampled for this study.

In addition to *A. aperta*, a few other organisms that are found in at least two sympatric or adjacent habitats show no evidence for habitat driving population structure. For instance, phylogeographic analysis of the amphipod, *Eogammarus confervicolus*, indicates at least three independent shifts in habitat preference from sandy embankment to adjacent woody debris (i.e. adjacent populations using different habitat types were more closely related than distant populations using the same habitat type; Stanhope et al 1993). Similarly, the walking sticks, *Timema cristinae*, which display different genetically controlled color morphs on two host plants, lack significant neutral genetic differentiation between populations found on adjacent host plants, whereas geographically distant populations are often significantly different (Nosil and Crespi 2004). Sympatric anadromous and freshwater ecotypes of rainbow trout, *Oncorhynchus mykiss*, also show greater structure among geographically separate sampling sites than between ecotypes within a sampling site, suggesting parallel evolution leading to multiple populations of freshwater ecotypes (Docker and Heath 2003). These studies suggest parallel evolution and the importance of current natural selection in maintaining ecotypic variation, but do not take into account relationships with other closely related species. Examining these relationships potentially may show

the importance of historically derived traits in determining current patterns of variation, as is probably the case in *A. aperta*.

In other organisms, however, habitat appears to play an important role in determining phylogeographic or population genetic structure. In addition to multiple switches from sandy embankment to woody debris, the amphipod mentioned above shows a single switch from embankment to algal habitat with algal populations from different sites forming a monophyletic group (Stanhope et al. 1993). The goldenrod ball gall maker, *Eurosta solidaginis*, also appears to show a single switch from galling *Solidago altissima* to galling *S. gigantea*. In both cases range expansions appear to have brought populations on the different habitat/host into sympatry. Other examples of habitat or host being the primary source of population structure are the pea aphids, *Acyrthosiphon pisum*, in Europe (Simon et al. 2003) and North America (Via 1999); sticklebacks, *Gasterosteus aculeatus*, in lake and river habitats in Europe (Reusch et al. 2001); and the grass, *Elymus athericus*, in high and low salt marsh habitats in Europe (Bockelman et al. 2003).

Instead of a signal of habitat-use in the mitochondrial data, there are strong patterns of geographic differentiation in *A. aperta* and *A. aleenae*. For *A. aperta*, at the level of the entire species there are regional geographic groupings that are probably the result of past fragmentation into allopatric glacial refugia with subsequent range expansion into the Rocky Mountains (Ayoub and Riechert 2004). Within each of the regional clades, there is also see a large amount of differentiation among sampling sites.

Differentiation of *A. aperta* populations within regions may be partially explained by isolation by distance. However, the pattern of decreasing scatter of F_{ST}

values with increasing geographic distance indicates that the system is not at an equilibrium state (see Figure A5). This pattern of decreasing scatter can probably be explained by an almost complete lack of gene flow between the most distant locations. Adjacent populations appear to exchange genes readily, but moving just 4-15 km away increases genetic differentiation (Table A3). Although this distance is very short to the human eye, to a ground-dispersing spider this could be a huge distance. Riechert et al. (2001) found the average distance traveled in a day by *A. aperta* males in arid habitat to be 23m and females 8m.

The level of geographic sampling for this study is apparently too coarse to capture the maximum distance over which *A. aperta* female-based gene flow is likely to occur. However, F_{ST} within 15km never exceeded 0.44 whereas just slightly longer distances produced a huge range of F_{ST} values from 0 to 1. This pattern is also apparent from the pairwise F_{ST} values for local populations distributed along transects (sites 1, 4). Although the local populations within a single transect were only 17 to 40km apart, the pairwise F_{ST} values ranged from 0 to 0.86. Such a large scatter of pairwise F_{ST} suggests that very little gene flow can occur even at these short geographic distances.

Because F_{ST} was never below 0.4 after geographic distance between populations reached some apparent critical distance, gene flow appears to be exceedingly rare if not nonexistent between the most distantly located populations. In the intermontane clade this distance was about 450km, and in the eastern clade this distance was about 1400km. In the intervening distances, gene flow probably can occur but apparently the diversifying effect of genetic drift outweighs the homogenizing effect of gene flow. I should mention that the large number of populations with an F_{ST} of zero over a range of

0 to 1300 km in the eastern clade were all fixed for a single haplotype. These populations were at sites 19-25. In Ayoub and Riechert (2004), we concluded that the fixation of this single haplotype over such a large area was most likely due to a post-Pleistocene range expansion. A selective sweep of the mitochondrial genome in this region could also explain the pattern of fixation. Thus, these zero F_{ST} values should not be taken as evidence of current levels of gene flow.

Population structuring of *A. aleenae* also seems to be strongly governed by geography, although the patterns are somewhat different from *A. aperta*. I observed three geographically localized clades of *A. aleenae*, and a geographically localized *A. spatula* haploptype set within the range of *A. aleenae*. As discussed in Ayoub et al. (in review), this pattern is suggestive of four allopatric divergences from a common ancestor with the evolution of the *A. spatula* morphotype in one of these lineages (i.e. parapatric speciation).

Each population of *A. aleenae* exhibits a single haplotype or a group of very closely related haplotypes (Table A1, Figure A2). This is in contrast to population structure of *A. aperta* over the same geographic area. All *A. aperta* populations that fall within the range of *A. aleenae* are part of the eastern *A. aperta* clade and when *A. aperta* populations exhibit multiple haplotypes they are typically not each other's closest relatives (Table A1, Figure A1). The geographic localization of *A. aleenae* lineages suggests that fragmentation may play a stronger role in population structure than isolation by distance. However, the possibility of isolation by distance shaping patterns of genetic diversity in *A. aleenae* cannot be rejected, because there are no samples between geographically localized clades (Templeton 2001). In addition, due to the

stochastic nature of the coalescent process, low dispersal species can form geographically defined genetic groups even when the species is continuously distributed throughout its range (Irwin 2002). However, I sampled sites between some of the geographically defined clades and these sites did not yield *A. aleenae*. If *A. aleenae* is truly not found in these intervening areas, then fragmentation may play a more important role in structuring populations than isoloation by distance. A mechanism that would explain this fragmentation is not immediately clear. In *A. aperta*, geographic clades correlate with the presence of mountain ranges, but in *A. aleenae* no obvious geographic barriers exist that can explain the geographically localized clades. Perhaps *A. aleenae* simply disperses less than *A. aperta*. Or perhaps *A. aperta* actually causes fragmentation of *A. aleenae* populations by out-competing *A. aleenae* in some areas. *Agelenopsis* was collected from a few intermediate sites but these spiders were all *A. aperta*. The only site where I found *A. aperta* and *A. aleenae* co-occurring was in the continuous riparian site (site 1) which has fairly high prey densities, and may accordingly have less intra- and inter-specific competition.

Conclusions

The population genetic and phylogenetic results presented here have important implications for the evolution of habitat-use and predicted associated behavioral characters in *Agelenopsis*. The single origin of arid-living will probably translate into a single origin of high aggression, low fear behaviors. In addition, the finding of strong geographic structuring within *A. aperta*, and that riparian patches are not linked by

common ancestry within *A. aperta* indicates that each riparian patch is independent and any associated behaviors are probably the result of repeated natural selection. Thus, natural selection probably shapes modern behavioral characters. However, behavioral characters associated with riparian habitat are probably dependent on the historical basis of riparian living in the genus. In sum, if habitat-associated behaviors hold for new populations of *A. aperta* and for the other species, history and natural selection have almost certainly interacted to shape the distribution of behavioral characters in *Agelenopsis*.

Literature Cited

Avise, J. C., 1994. Molecular Markers, Natural History and Evolution, Chapman and Hall, New York.

Axelrod, D. I., 1985. Rise of the grassland biome, central North America. The Botanical Review. 51, 163-201.

Ayoub, N.A., and Riechert, S.E, 2004. Molecular evidence for Pleistocene glacial cycles driving diversification in a North American desert spider, *Agelenopsis aperta*. Molecular Ecology. 13, 3453-3465.

Bernatchez, L., Chouinard, A, Lu, G., 1999, Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, Coregonus sp., as a case study. Biological Journal of the Linnean Society. 68, 173-194.

Bockelmann, A. -C., Reusch, T. B. H, Bijlsma, R., and. Bakker, J. P., 2003. Habitat differentiation vs. isolation-by-distance: the genetic population structure of *Elymus athericus* in European salt marshes. Molecular Ecology. 12:505-515.

Brower, A., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconious erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Science USA. 99, 6491-6495.

Brown, J. M., Abrahamson, W. G., Way, P. A., 1996. Mitochondrial DNA phylogeography of host races of the goldenrod ball gallmaker, *Eurosta solidaginis* (Diptera: Tephritidae). Evolution. 50, 777-786.

Chamberlin, R. V., and Ivie, W., 1941. North American Agelenidae of the genera *Agelenopsis, Calilena, Ritalena* and *Tortolena.* Annals of the Entomological Society of America. 34, 585-628.

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

Daly, C., 2000. Spatial Climate Analysis Service, United States Average Annual Precipitation, 1961-1990, *in* National Atlas of the United States, http://nationalatlas.gov.

Docker, M.F., and Heath D. D., 2003. Genetic comparison between sympatric anadromous steelhead and freshwater resident rainbow trout in British Columbia, Canada. Conservation Genetics. 4, 227-231.

Earnshaw, A.P.R. 1973. The ecology, distribution and dispersion of *Agelenopsis utahana* Chamberlin and Ivie, 1933 and *A. potteri* Blackwall, 1846, in the Morgan Arboretum of Macdonald College, P.Q. Ph.D. Dissertation, McGill University, 166pp.

Edwards, S.V. and Kot. M., 1995. Comparative methods at the species level: geographic variation in morphology and group size in grey-crowned babblers (*Pomatostomus temporalis*). Evolution. 46, 1134-1146.

Edwards, S., Beerli, P., 2000. Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution. 54, 1839-1854.

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

Foster, S. and Cameron, S., 1996. Geographic variation in behavior: a phylogenetic framework for comparative studies. *In* Phylogenies and the Comparative Method in Animal Behavior. (ed. E. Martins). pp. 138-165. Oxford University Press, New York.

Gertsch, W. J., and Riechert, S. E., 1976. The spatial and temporal partitioning of a desert spider community, with descriptions of new species. American Museum Novitiates. 2604, 1-25.

Gould, S. J., 1989. Wonderful Life: The Burgess Shale and the Nature of History. Norton, New York.

Gould, S. J., and Woodruff, D. S., 1990. History as a cause of area effects: an illustration from *Cerion* on Great Inagua Bahamas West Indies. Biological Journal of the Linnaean Society. 40, 67-98.

Harvey, P. H., and Pagel, M. D., 1990. The Comparative Method in Evolutionary Biology. Oxford University Press, Oxford.

Hedrick, A. V., and Riechert, S. E., 1989. Genetically-based variation between two spider populations in foraging behavior. Oecologia. 80, 533-539.

Hutchison, D. W., Templeton, A. R., 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution. 53, 1898-1914.

Iriwn, D. E., 2002 Phylogeographic breaks without geographic barriers to gene flow. Evolution. 56, 2383-2394.

Jones, S. E., 1941. Influence of temperature and humidity on the life history of the spider, *Agelena naevia* Walckenaer. Annals of the Entomological Society of America. 34, 557-571.

Kishino, H., and Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution. 29, 170-179.

Lehtinen, P. T., 1967. Classification of the cribellate spiders and some allied families, with notes on the evolution of the suborder Araneomorpha. Ann. Zool. Fenn. 4, 199-468.

Losos, J. B., Jackman, T.R., Larson, A., de Queiroz, K., and Rodriguez-Schettino, L., 1998. Contingency and determinism in replicated adaptive radiations of island lizards. Science. 279, 2115-2118.

Maddison, D. R., and Maddison, W. P., 2000. MacClade, ver4. Sinauer Associates: Sunderland, Massachusetts.

Malhotra, A., and Thorpe, R. S., 2000. The dynamics of natural selection and vicariance in the dominican anole: Patterns of within-island molecular and morphological divergence. Evolution. 54, 245-258.

Mantel, N., 1967. The detection of disease clustering and generalized regression approach. Cancer Research. 27, 209-220.

Muma, M. H., and Muma, K. E., 1949. Studies on a population of prairie spiders. Ecology. 30, 485-503.

Nei, M., Li, W. H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Science USA. 76, 5269-5273.

Neigel, J. E., Avise, J. C., 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. In: Evolutionary processes and theory (eds Karlin S, Nevo E), pp. 515-534. Academic Press, New York.

Nosil, P. and Crespi, B. J., 2004. Does gene flow constrain adaptive divergence or vice versa? A test of using ecomorphology and sexual isolation in *Timema cristinae* walking sticks. Evolution. 58, 102-112.

Pagel, M., 1999. Inferring the historical patterns of biological evolution. Nature. 401, 877-884.

Paison, T. C., 1997. A biogeographic review of the spider genus *Agelenopsis* (Araneae: Agelenidae). MS thesis, University of Tennessee, Knoxville, 123 pages.

Platnick, N. I., 2004. The world spider catalog, version 4.5. American Museum of Natural History, online at http://research.amnh.org/entomology/spiders/catalog/index.html

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

Reusch, T. B. H., Wegner, K. M., Kalbe, M., 2001. Rapid genetic divergence in postglacial populations of threespine stickleback (*Gasterosteus aculeatus*): the role of habitat type, drainage and geographical proximity. Molecular Ecology. 10, 2435-2445.

Riechert, S., 1993a. A test for phylogenetic constraints on behavioral adaptation in a spider system. Behavioral Ecology and Sociobiology. 32, 343-348.

Riechert, S., 1993b. Investigation of potential gene flow limitation of behavior adaptation in an arid ands spider. Behavioral Ecology and Sociobiology. 32, 355-363.

Riechert, S. E. 1999. Using behavioral ecotypes to study evolutionary processes. In: Geographic variation in behavior: Perspectives on evolutionary mechanisms (eds Foster S, Endler J), pp. 3-32. Oxford Univ. Press, New York.

Riechert, S. E., 1987. Between population variation in spider territorial behavior: Hybrid-pure population line comparisons. In M. Huettel ed., Evolutionary Genetics of Invertebrate Behavior. Plenum Press.

Riechert, S., Hedrick, A., 1990. Levels of predation and genetically based anti-predatory behavior in the spider, *Agelenopsis aperta*. Animal Behavior. 40, 679-687.

Riechert, S. A., Hedrick, A., 1993. A test for correlations among fitness-linked behavioral traits in the spider *Agelenopsis aperta*. Animal Behavior. 46, 669-675.

Riechert, S., and Maynard Smith, J., 1989. Genetic analyses of two behavioral traits linked to individual fitness in the desert spider, *Agelenopsis aperta*. Animal Behavior. 37, 624-637.

Riechert, S. E., Hall, R. F., 2000. Local population success in heterogeneous habitats: reciprocal transplant experiments completed on a desert spider. Journal of Evolutionary Biology. 13, 1-10.

Riechert, S. E., Singer, F. D., Jones, T. C., 2001. High gene flow levels lead to gamete wastage in a desert spider system. Genetica, 112-113, 297-319.

Rundle, H. D., Nagel, L. Natural selection and parallel speciation in sympatric sticklebacks. Science. 287, 306-309.

Rohlf, F.J., 1973. Algorithm 76: Hierarchical clustering using the minimum spanning tree. The Computer Journal. 142, 1357-1362.

Schneider, S., Roessli. D., Excoffier, L., 2000. Arlequin ver. 3.000: A software for population genetics data analysis, Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Simon, J. Carre, -C. S., Boutin, M., Prunier-Leterme, N., Sabater-Munoz, B., Latorre, A., Bournoville, R., 2003. Host-based divergence in populations of the pea aphid: insights from nuclear markers and the prevalence of facultative symbionts. Proceedings of the Royal Society of London. 270, 1703-1712.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution, 16, 1114-1116.

Sokal, R. R., Rohlf, F.J., 1995. Biometry, 3rd ed. W.H. Freeman and Company, New York.

Stanhope, M. J., Hartwick, B., Baillie, D., 1993. Molecular phylogeographic evidence for multiple shifts in habitat preference in the diversification of an amphipod species. Molecular Ecology. 2, 99-112.

Swofford, D., 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4, Sinauer Associates, Sunderland, MA.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution. 10, 512-526.

Taylor, E. B., McPhail, J.D., 2000. Historical contingency and ecological determinism interact to prime speciation in sticklebacks, *Gasterosteus*. Proceedings of the Royal Society of London, Series B. 267, 2375-2384.

Templeton, A. R., 2001. Using phylogeographic analyses of gene trees to test species status and processes. Molecular Ecology. 10, 779-791.

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

Thorpe, R. S., 1996. The use of DNA divergence to help determine the correlates of evolution of morphological characters. Evolution. 50, 524-531.

Travisano, M., Mongold, J. A., Bennett, A. F., Lenski, R. E., 1995. Experimental tests of the roles of adaptation, chance, and history in evolution. Science. 267, 87-90.

Wiens,J.J., Reeder T.W., Montes de Oca, A.N., 1999. Molecular phylogenetics and evolution of sexual dichromatism among populations of the yarrow's spiny lizard (*Sceloperus jarrovii*). Evolution. 53, 1884-1897.

Williams, D., Dunkerly, D.,, DeDeckker, P., Kersaw, P., Chappell, M.,1998. Quaternary Environments, Arnold, London.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., *et al.*, 1985. Mitochondrial DNA and two perspectives on evolutionary genetics: Biological Journal of the Linnaean Society. 26, 375–400.

Via, S., 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. Evolution 53, 1446-1457.

Zamudio, K. R., 1998. The evolution of female-biased sexual size dimorphism: a population-level comparative study in horned lizards (*Phrynosoma*). Evolution. 52, 1821-1833.

Appendix

Table A1. Sampling sites of *Agelenopsis*. Numbers in parentheses are number of individuals from that local population exhibiting a particular haplotype.

site #	Town/Park	Habitat type of local population	Latitude, Longitude	Haplotypes exhibited	Elevation (m)
A. aperta Gertsch					
1* **	Hye, TX	riparian	30.27N, 98.34W	R(5)	450
		adjacent riparian	30.28N, 98.55W	R(9), D(2)	450
		distant riparian (northern end of transect)	30.25N, 98.6W	R(5), C(4)	450
		riparian (middle of transect)	30.09N, 98.42W	B(3), R 5), D(1)	450
		riparian (southern end of transect)	29.94N, 98.4W	F(6), D(2), R(1), CC(1)	450
2*	San Angelo, TX	riparian patch	31.53N, 100.54W	V(1), BB(1), R(1), U(7)	560
		adjacent arid	31.53N, 100.54W	R(3), U(5), V(1), P(1)	560
		distant arid	31.47N, 100.51W	U(1), R(8), V(1), A(1)	560
3*	Balmorhea,TX	riparian patch	30.97N, 103.72W	AA(7), G(3)	970
		adjacent arid	30.97N, 103.72W	AA(8), CC(2)	970
		distant arid	30.85N, 103.80W	AA(8), Z(1)	970
4**	Just outside Big Bend National Park, TX	arid (northwestern end of transect)	29.78N, 103.18W	CC(6), AA(1), U(1), P(1)	1235
		arid (middle of transect)	29.65N, 103.08W	P(1), CC(9)	1235
		arid (southeastern end of transect)	29.55N, 102.94W	Q(1), P(6)	1235
5	Amistad Nat'l Rec Area, TX	riparian patch	29.50N, 100.91W	DD(1), D(10)	180
6	Cuatro Cienegas, Coahuila	arid	26.92N, 102.05W	M(9), N(1)	710
7	Lerdo, Durango	arid	25.52N, 103.55W	EE(1), CC(1), GG(8)	1160

Table A1. Continued.

site #	Town/Park	Habitat type	Latitude, Longitude	Haplotypes exhibited	Elevation (m)
8	Plan de Ayupla, Zacatecas	arid	22.62N, 102.78W	O(7), EE(2), GG(1)	1860
9	Tres Marias, Taumalipas	arid	22.47N, 98W	FF(10)	50
10	Brantley Lake SP, NM	arid	32.57N, 103.38W	G(6), K(1), L(1), Y(1)	975
11	Clovis, NM	riparian patch	34.39N, 103.28W	G(10)	1300
12	Ute Lake, NM	arid	35.37N, 103.41W	X(5)	1160
13*	Palo Duro Canyon SP, TX	riparian patch	34.96N, 101.67W	U(10)	630
		adjacent arid	34.96N, 101.67W	U(10)	630
		distant arid	34.98N, 101.70W	U(5), S(5)	670
14	Foss Lake State Park, OK	continuous riparian	35.5N, 99.17W	T(8)	500
15	Colorado City, CO	arid	37.94N, 104.85W	II(7)	1785
16	Avondale, CO	riparian patch	38.22N, 104.39W	II(7), F(1), HH(1)	1390
17	Sedalia, CO	riparian patch	39.43N, 104.95W	II(8), S(2)	1775
18	Nogal Mt, NM	arid	33.48N, 105.8W	AA(2), G(2), I(1)	1800
19*	Carrizozo, NM	arid	33.63N, 105.87W	G(5)	1655
		arid	33.63N, 105.87W	G(5)	1655
		arid	33.61N, 105.85W	G(5)	1655
20	Rio Grande Nature Center, NM	riparian patch	35.08N, 106.65W	G(5)	1510
21	Deming, NM	arid	32.27N, 107.75W	G(5)	1320

Table A1. Continued.

site #	Town/Park	Habitat type	Latitude, Longitude	Haplotypes exhibited	Elevation (m)
22*	South Western Research Station, AZ	riparian patch	31.91N, 109.37W	G(5)	1655
		adjacent arid	31.91N, 109.37W	G(5)	1655
		distant arid	31.8N, 109.21W	G(5)	1655
23	Dead Horse Ranch State Park, AZ	riparian patch	34.6N, 112.1W	G(5)	995
24*	Colorado River State Park, CO	riparian patch	39.05N, 108.55W	G(5)	1400
		adjacent arid	39.05N, 108.55W	G(5)	1400
25	San Diego, CA	arid	32.73N, 117.15W	G(7)	15
26	New Port Beach, CA	arid	33.57N, 117.84W	WF(1), WE(5)	15
27	North LA area, CA	riparian patch	34.07N, 118.47W	WG(7)	15
28*	Big Pine, CA	riparian patch	37.15N, 118.3W	WH(2), WD(3), WC(4)	1400
		adjacent arid	37.15N, 118.3W	WD(2), WC(2), WB(1), WH(5)	1400
		distant arid	37.10N, 118.3W	WC(10)	1220
29	Warm Springs, NV	arid	38.18N, 116.37W	WI(10)	1650
30	Flowell, UT	arid	39.0N, 112.5W	WN(10)	1400
31	Santaquin, UT	arid	39.95N, 111.78W	WN(2), WM(8)	1555
32	exit 62, Hwy 80, UT	arid	40.82N, 112.9W	WI(2), WL(3), WJ(4), WK(1)	1430
33*	Springdale, UT	riparian patch	37.18N, 112.99W	G(3), H(1), WA(6)	1190
		adjacent arid	37.18N, 112.99W	G(3), H(1), WA(4), WO(1)	1190
		distant riparian	37.23N, 112.95W	G(8), WA(2)	1190

Table A1. Continued.

site #	Town/Park	Habitat type	Latitude, Longitude	Haplotypes exhibited	Elevation
34	St. George	arid	37.10N, 113.58W	J(2), G(6)	880
35	near Don Pedro Reservoir, CA	arid	37.82N, 120.3W	WP(10)	440
36	Bells Station, CA	riparian, no adjacent arid	37.04N, 121.31W	WQ(1), WR(1)	15
A. aleenae Chamberlin and Ivie					
1	Hye, TX	adjacent riparian	30.27N, 98.34W	aleenae5(3), aleenae10(2), aleenae11(1)	440
		adjacent riparian	30.27N, 98.34W	aleenae6(2), aleenae7(2), aleenae2(4)	440
		distant riparian	30.27N, 98.19W	aleenae8(2), aleenae5(3), aleenae10(2), aleenae11(1)	440
37	Meade SP, KS	tall grass, wheat, along cow pasture and corn fields	37.17N, 100.45W	aleenae1(3)	760
38	Cimarron Canyon, NM	grassy area by riparian woodland	36.51N, 104.91W	aleenae15(1)	1980
39	Valmora, NM	desert grassland	35.8N, 105.88W	aleenae16(1)	1940
41	Clines Corners, NM (just north)	evergreen trees	35.01N, 105.67W	aleenae12(10)	2170
42	Hobbs, NM	arid	32.700N, 103.133W	aleenae13(5), aleenae14(5)	1100
A. spatula Chamberlin and Ivie					
43*	Caprock Canyons SP, TX	riparian patch	34.4N, 101.1W	spatula1(5)	760
		adjacent arid	34.4N, 101.1W	5 spatula1(5)	760
		distant arid	34.6N, 101.1W	spatula1(5)	760
14	Foss Lake State Park, OK	continuous riparian	35.5N, 99.17W	spatula1(1)	500
A. novum					
3	10 mi south of Balmorhea	distant arid	30.85N, 103.80W	novum(2)	970

Table A1. Continued.

site #	Town/Park	Habitat type	Latitude, Longitude	Haplotypes exhibited	Elevation (m)
A. emertoni Chamberlin and Ivie					
14	Foss Lake State Park, OK	continuous riparian	35.5N, 99.17W	emertoni1(1)	500
45	Dallas/Fort Worth, TX	riparian	32.77N, 96.78W	emertoni1(1)	140
46	Loudon, TN	eastern deciduous forest	35.73N, 84.31W	emertoni2(1)	240
A. kastoni Chamberlin and Ivie					
46	Loudon, TN	eastern deciduous forest	35.73N, 84.31W	kastoni1(1), kastoni2(1)	240
A. longistylus Banks					
38	Cimarron Canyon, NM	grassy area by riparian woodland	36.51N, 104.91W	longistylus1(6)	1980
40	Galisteo, NM (1.4 mi west of)	grass and little bushes by creek	35.4N, 105.9W	longistylus2(1)	1840
A. naevia Walckenaer					
45	Dallas/Fort Worth, TX	riparian	32.77N, 96.78W	naevia2(1)	140
46	Loudon, TN	eastern deciduous forest	35.73N, 84.31W	naevia1(1)	240
A. oklahoma Gertsch					
14	Foss Lake State Park, OK	continuous riparian	35.5N, 99.17W	oklahoma1(1)	500
24	Colorado River State Park, CO	riparian patch	39.05N, 108.55W	oklahoma3(4)	1400
37	Meade SP, KS	along cow pasture and corn fields in lush grass	37.17N, 100.45W	oklahoma1(1)	760
44	Wichita, TX (Lake Arrowhead SP)	mesquite, other trees, pretty lush	33.75N, 98.37W	oklahoma1(1)	280
48	Trinidad, CO by Purgatoire River	grassy area by riparian woodland	37.169N, 104.506W	oklahoma2(1)	1830
49	Monument, CO (1 mile west of)	mixture of arid veg with lusher veg, creeks, pond	39.07N, 104.86W	None sequenced	2120
50	Lon Hagler Reservior, CO	lush grassy area, few trees, near reservoir	40.36N, 105.15W	none sequenced	1550
51	Greeley, CO	juniper bushes galore	40.42N, 104.74W	none sequenced	1420

Table A1. Continued.

site #	Town/Park	Habitat type	Latitude, Longitude	Haplotypes exhibited	Elevation (m)
A. oregonensis Chamberlin and Ivie					
54	Victoria, BC, Canada	mesic[1]	NA[1]	oregonensis3 (1)	NA[1]
A. pensylvanica Koch					
14	Foss Lake State Park, OK	continuous riparian	35.5N, 99.17W	pensylvanica1 (1)	500
47	Knoxville, TN	suburban	35.99N, 83.9W	pensylvanica1 (3)	270
55	Ellsworth, KS	along creek, under deciduous trees	38.76N, 98.22W	pensylvanica1 (1)	490
60	Barber Ponds SP, CO	riparian - grassy, lots of trees	40.17N, 105.01W	pensylvanica1 (1)	1520
A. potteri Blackwell					
54	Victoria, BC, Canada	mesic[1]	NA[1]	potteri1(1)	NA[1]
56	Winnipeg, MB, Canada	mesic[1]	NA[1]	potteri2(1)	NA[1]
A. utahana Chamberlin and Ivie					
52	Wasatch Mtns SP, UT	along creek, deciduous trees	40.55N, 111.49W	oregonensis2 (2), oregonensis4 (2)	1850
57	Lincoln, NH	northern hardwood	44.05N, 71.67W	utahana3(1)	250
58	White Mtns, NH (Jigger Johnson Campground)	northern hardwood	44.2N, 71.3W	utahana1(1)	300
59	Monadnock SP, NH	northern hardwood	42.83N, 72.06W	utahana2(1)	240

[1]These samples were kindly sent to us by Joe Spagna from other collections. Habitat descriptions, coordinates, and elevations were often not included with the specimens.
* At these sites "cluster" sampling was completed, which included a focal population (usually a riparian patch), an adjacent population, and an arbitrarily chosen distant population 4-15 km away from the focal population.
** At these sites three local populations were sampled along a 40 or 32km transect.

Table A2. Analysis of Molecular Variance (AMOVA) table for *A. aperta* grouping populations by habitat type (riparian v. arid) or as cluster site. Local populations used were restricted to ones at riparian patch cluster sites (sites 2,3,13,22,24,28,33).

	% variation	
Variance Partition	Group = Habitat	Group = Cluster
Among Groups	-5.54	64.74*
Among populations within groups	74.59*	7.28*
Within populations	30.95*	27.98*

*p<0.001

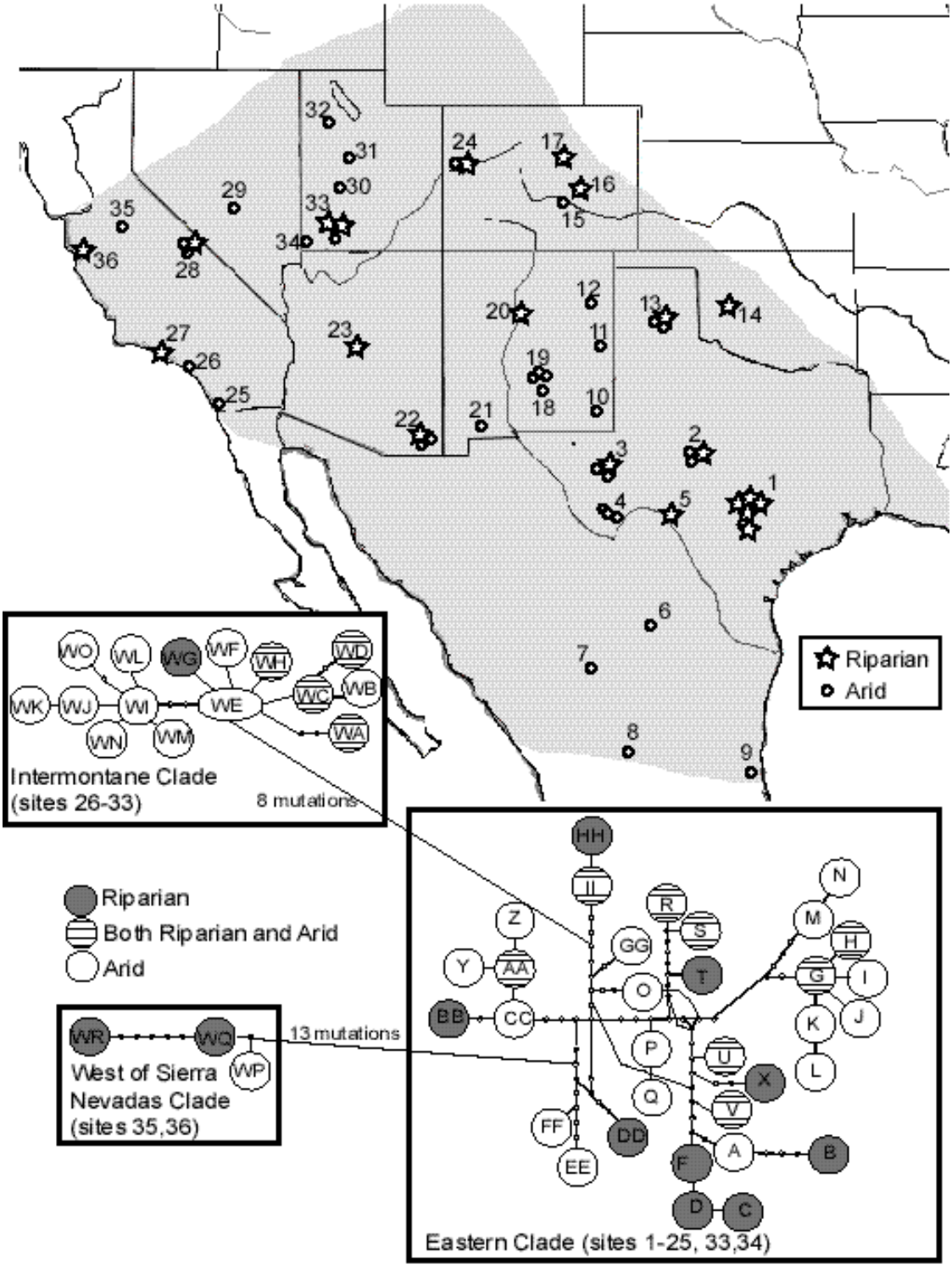
Table A3. Pairwise F_{ST} values for each cluster site. Species is *A. aperta* unless otherwise indicated. F_{ST} was calculated using genetic distance among haplotypes (TrN corrected) and using conventional methods (in parentheses) The first 8 sites are for riparian patches and surrounding local populations. The bottom sites are for clusters within continuous habitats.

Site #	riparian patch - adjacent arid	riparian patch - distant population	adjacent arid - distant population
2	0.021 (-0.02)	**0.45*** (0.41*)	0.19 (0.18*)
3	0.18 (0.07)	0.2 (0.08)	0.06 (-0.003)
13	0 (0)	0.44* (0.44*)	0.44* (0.44*)
22	0 (0)	0 (0)	0 (0)
24	0 (0)	0 (0)	NA[1]
28	-0.009 (0.0075)	0.28* (0.37*)	**0.38* (0.54*)**
33	-0.1 (-0.08)	0.12 (0.25)	0.21 (0.15)
43 (*A. spatula*)	0(0)	0(0)	0(0)
	adjacent populations	adjacent 1 - distant population	adjacent 2 - distant population
1 (continuous riparian - *A. aperta*)	-0.0041 (-0.004)	0.27 (.28)	0.084 (.02)
1 (continuous riparian - *A. aleenae*)	0.26* (0.28*)	0.063 (0.088)	0.016 (0.18*)
19 (continuous arid - *A. aperta*)	0 (0)	0 (0)	0 (0)

*$p<0.05$; those in bold remained significant after Bonferroni correction. Bonferroni correction only considered *A. aperta* populations.
[1]Spiders collected at this local population were all *A. oklahoma*.

Figure A1. Sampling localities and geographically defined clades of *A. aperta*. The area in gray on the map indicates the range of *A. aperta*. However, *A. aperta* does not use areas above 2000m within this range (see Ayoub and Riechert 2004). Numbers correspond to sampling sites in Table A1.

The boxed clades below the map indicate sub-networks identified by TCS; these sub-networks could not be connected by TCS, and connections shown represent a "best guess" (see Ayoub and Riechert 2004). Notice that riparian and arid spiders exhibit haplotypes distributed throughout each clade.


Riparian
Arid
Intermontane Clade
(sites 26-33)
8 mutations
Riparian
Both Riparian and Arid
Arid
13 mutations
West of Sierra
Nevadas Clade
(sites 35,36)
Eastern Clade (sites 1-25, 33,34)

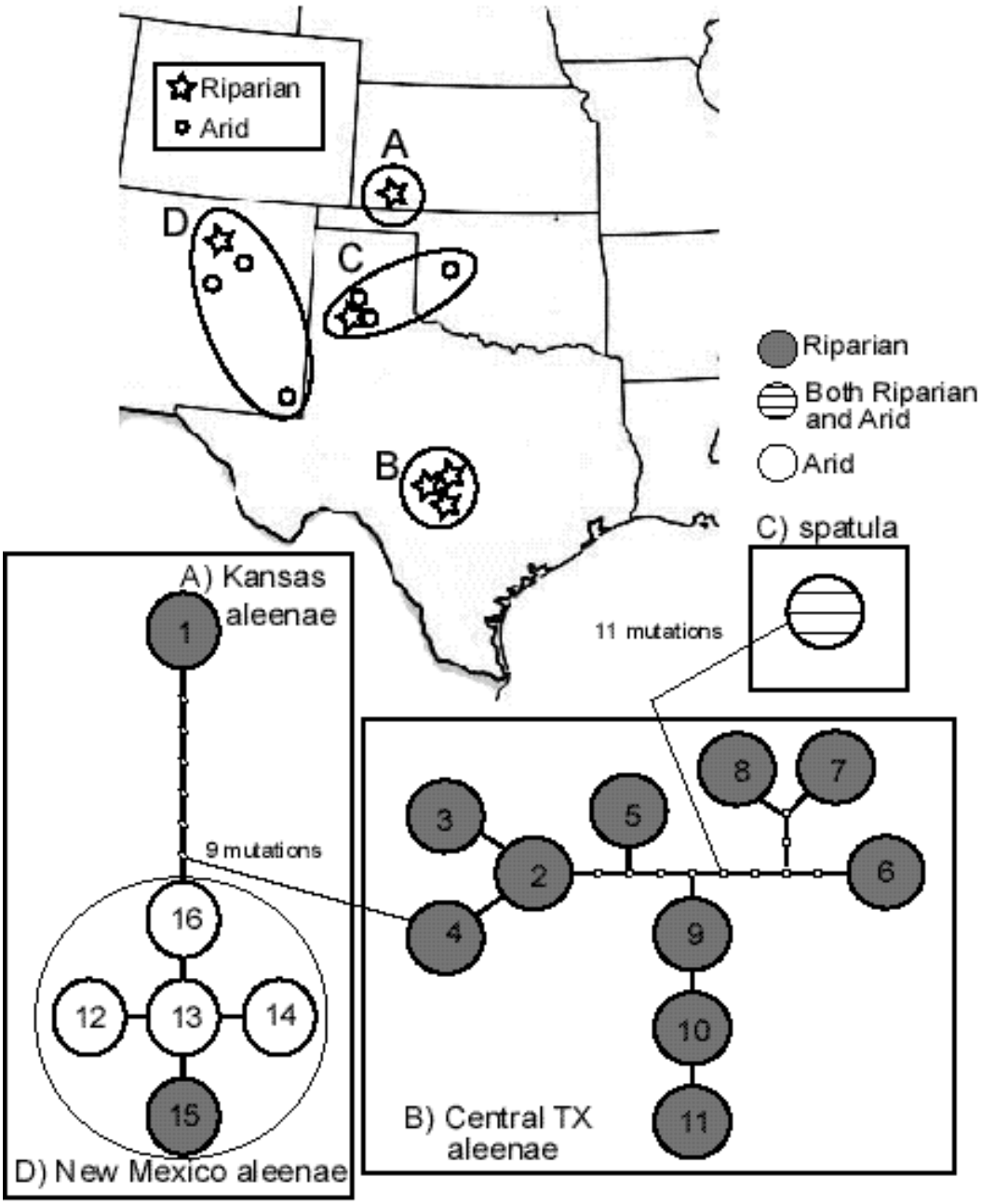


Figure A2. Sampling localities and geographically defined clades of *A. aleenae* and *A. spatula*. Numbers correspond to sampling sites in Table A1. The boxed clades below the map indicate sub-networks identified by TCS; these sub-networks could not be connected by TCS, and connections shown represent a "best guess" (see text). Notice that when habitat is not confounded by geography (*A. spatula*, *A. aleenae* in northern New Mexico), habitat does not display a phylogenetic signal.

Figure A3. Sampling localities of *Agelenopsis* species other than *A. aperta*, *A. aleenae*, or *A. spatula*. Numbers correspond to sampling sites in Table A1.

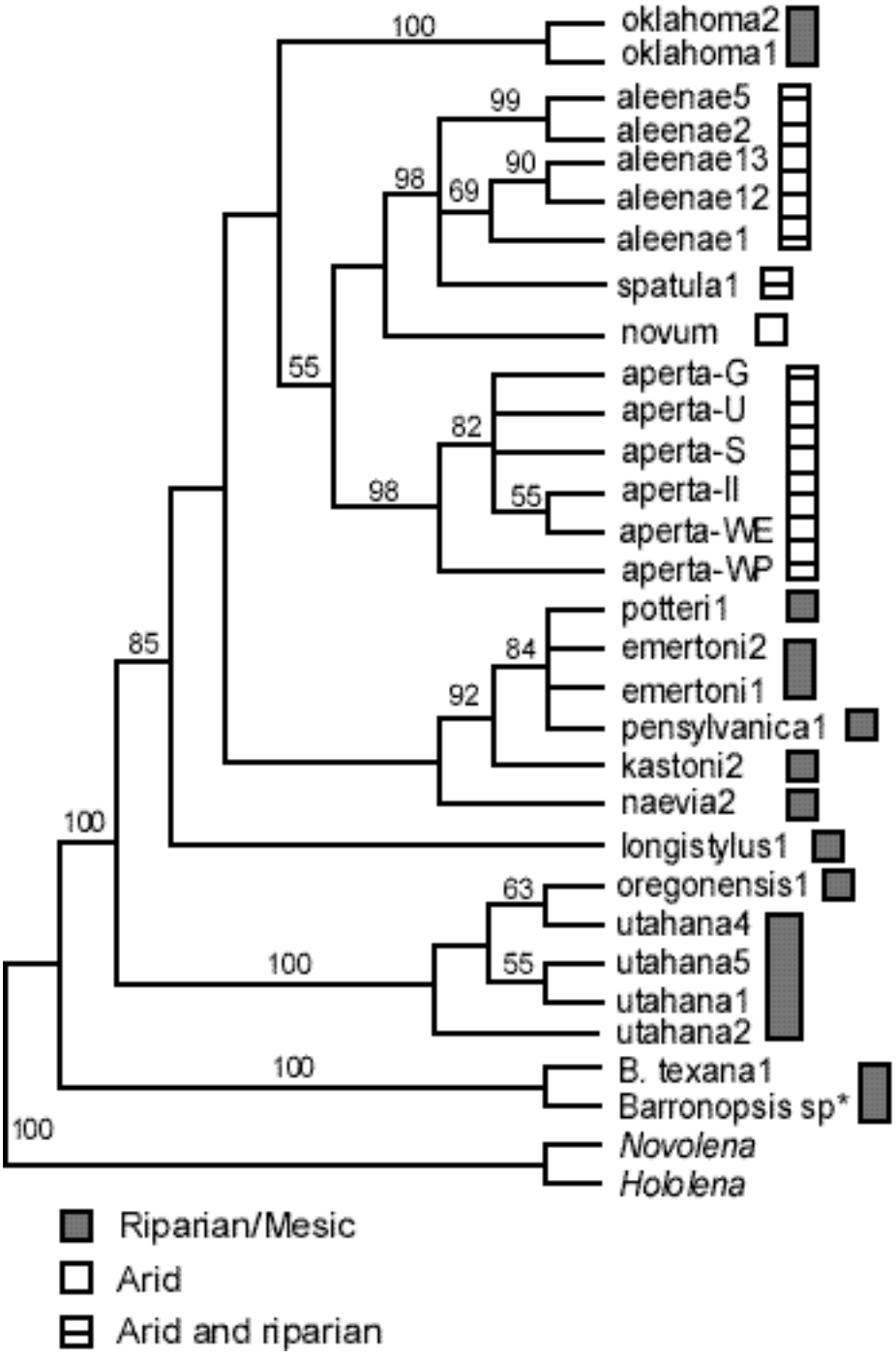


Figure A4. Strict consensus of 30 MP trees of COI plus 16S data combined. Bootstrap probabilities based on 1000 replicates are shown above branches. Note that all species that use arid habitats form a monophyletic group. Habitat-use was coded for each species, not by haplotype.

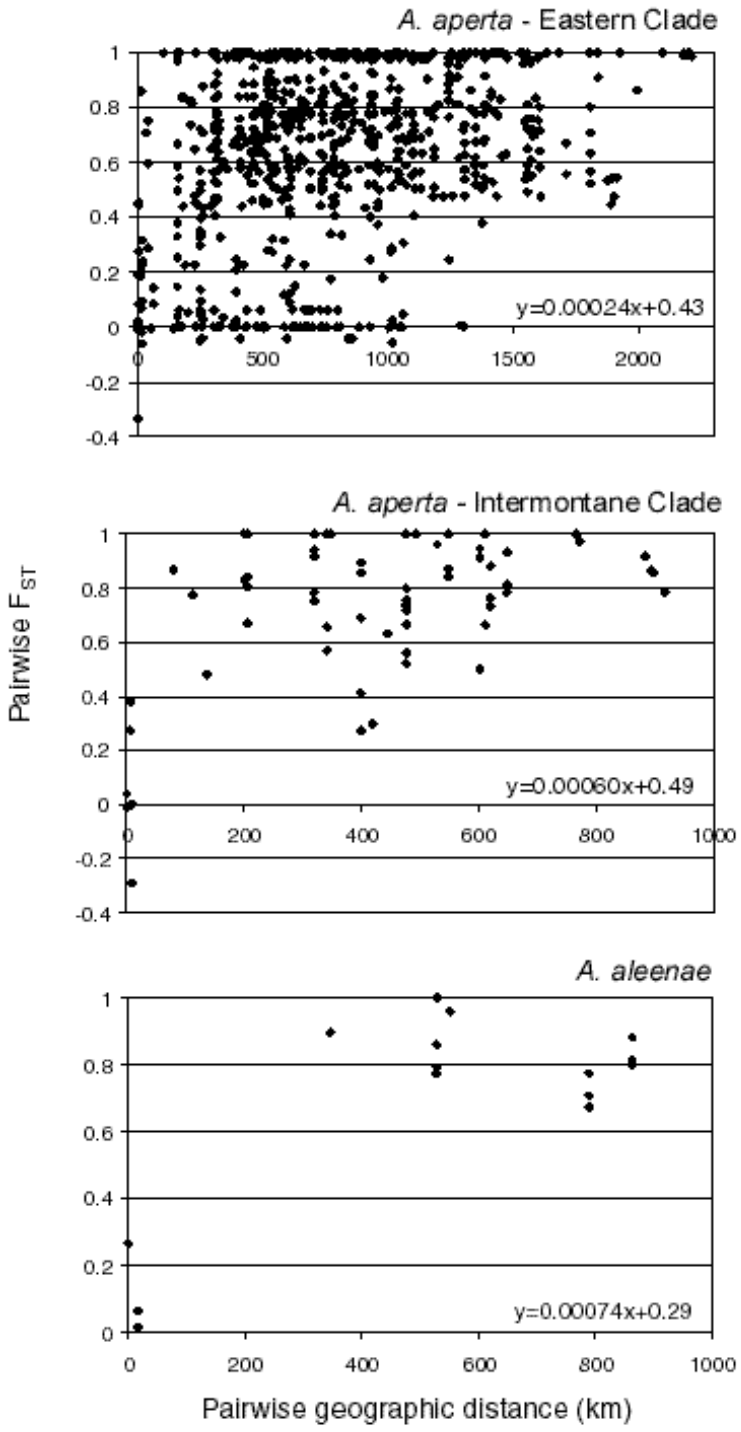


Figure A5. Plots of pairwise geographic distance (km) versus pairwise genetic distance (F_{ST}) for two clades of *A. aperta* and for *A. aleenae*. See Figure A1 for geographic distribution of *A. aperta* clades. A significant positive correlation between geographic and genetic distance was found for each, but see text for caveats.

Concluding Remarks: Summary and Future Work

In this dissertation I use a phylogenetic framework to address the role of history in shaping various evolutionary patterns in the genus *Agelenopsis*, with special attention to the species, *A. aperta*. I reconstruct intra- and inter-specific relationships using mitochondrial DNA (mtDNA) sequence variation. Because evolution is thought to occur at the population level, exploring the interface between population level variation and species level variation allows for insights about evolutionary history that would not be possible when only looking at one or the other. While species level phylogenies are common and have been used extensively to investigate the evolution of some characters, much less attention has been paid to inspecting the evolution of characters at the intraspecific level (but see references in Part IV). In addition, when intraspecific variation is considered, it is often not set in the context of other closely related species. By doing so, I was able to make a number of insights into patterns of speciation and to identify a potential interaction between historical contingency and current natural selection.

Although the mitochondrial genome is effectively a single linked locus and the results of this dissertation thus have limited explanatory power, the general hierarchical framework of considering intraspecific variation in the context of the interspecific phylogeny is, I believe, a powerful way to identify important evolutionary patterns. Once interesting patterns have been identified, various hypotheses regarding the mechanisms contributing to those patterns can then be investigated. After a brief summary of each Part, I describe a number of questions that arose during the course of this dissertation study and offer some ideas about how to address these questions using a hierarchical framework.

Summary

In Part I, I found that patterns and levels of sequence variation were consistent with Pleistocene climate change causing barriers to gene flow in *A. aperta* by restricting populations to lower elevation areas of available habitat. In addition, a low level of sequence variation in populations found in the Rocky Mountains is consistent with a post Pleistocene expansion out of a glacial refugium. While uplift of the Rocky Mountains and Sierra Nevadas could potentially have explained the pattern of geographically defined clades, sequence divergence dated to within the last two million years. This is well within the Pleistocene, and well after final uplift of either mountain chain.

In Part II, I concluded that approximately half of the *Agelenopsis* species examined formed monophyletic groups, while half were part of monophyletic species groups. While these results could be suggestive of poor taxonomy or hybridization, it is also likely that members of these species groups are simply recently diverged and thus the mtDNA maintains ancestral polymorphism common to each of the daughter species. Even though the mtDNA did not resolve all relationships among or within species, the patterns found in the *A. aleenae - A. spatula* group allowed me to reject the hypothesis that *A. aleenae* was simply a hybrid between *A. aperta* and *A. spatula*.

Part III takes advantage of setting intraspecific variation in the context of interspecific phylogeny. I found that populations of *A. aperta* were not differentiated by habitat type, even though habitat-associated behaviors had previously been established

for three study sites of *A. aperta*. Instead, geography played a significant role in structuring neutral genetic variation in *A. aperta* populations, with riparian patches sampled throughout the desert being significantly different from each other. These patterns indicate that any similarities in behavior among riparian patches would be the result of repeated, independent natural selection rather than gene flow and historical contingency within *A. aperta*. However, all the species which use arid habitats, *A. aperta, A. aleenae*, *A. spatula* and *Agelenopsis species novum*, formed a monophyletic group, indicating that arid living evolved only once within *Agelenopsis*. The intraspecific data set within the interspecific phylogeny suggests that although arid living evolved once, each arid species retained the ability to invade riparian habitats. Thus, if new populations of *A. aperta* and other *Agelenopsis* species display similar habitat-associated behavioral characters, selection probably interacts with historically contingent genetic traits to govern behavior in *Agelenopsis*.

Future work

The results of these studies indicate that *Agelenopsis* provides an excellent study system in which to address a number of questions in ecology and evolutionary biology. For most of the questions discussed in the following paragraphs, a hierarchical framework is essential for understanding mechanisms, even if an explicit phylogeny is not needed.

Throughout this dissertation, I referred to geographic ranges of each species. These ranges were used to infer habitat-use and in the case of *A. aperta*, infer that it is

restricted to lower elevation arid regions. However, the mechanisms restricting species ranges are not understood for *Agelenopsis*. This is an important ecological question that applies to any organism. As mentioned in Part I, physiological intolerance or competition could restrain *A. aperta* to lower elevations. In Part IV, I discussed the possibility that competition with *A. aperta* could constrain *A. aleenae's* distribution. Reciprocal transplants of these species could help elucidate the importance of physiological versus competitive effects. Another interesting ecological question that arises from geographic ranges is how do some of the species co-occur over such broad areas and can even be found at the same sampling site. Even the closely related species, *A. pensylvanica*, *A. potteri*, and *A. emertoni* are widely sympatric. Do these species display niche partitioning, or are they functionally equivalent? Perhaps they are not even separate species (see below). If the widely held paradigm in ecology that different species cannot occupy the same niche is true, then how do different *Agelenopsis* species partition habitat? Or perhaps, this paradigm does not hold for *Agelenopsis*.

Another important ecological question that ties into evolution of behavior, is how to define habitat-use. In Part IV, I simplistically split habitat-use into either arid or riparian/mesic types. While these have been shown to be an important dichotomy for *A. aperta* in regard to associated behavioral traits, there may be a finer scale of habitat-use that needs to be examined to understand behavioral and physiological adaptations among *Agelenopsis* species. First of all, a more important indicator of spider habitat than amount of precipitation or presence of water or certain vegetation, is probably prey abundance and predation risk. These two factors are the primary mechanisms underlying differential selection for behavioral phenotypes in *A. aperta* and are

probably important for other species. Secondly, even the cursory view of habitat-use I took in Part IV indicates that species like *A. oklahoma* and *A. longistylus* may not fall into either "arid" or "riparian" habitats. *Agelenopsis oklahoma*, for instance, is found in parts of the United States receiving low levels of precipitation, although this region is not as dry as the southwestern United States. Does *A. oklahoma*, then, use intermediate habitat between arid and riparian/mesic? Precipitation patterns need to be correlated to the life histories of individual species. As a corollary, do *Agelenopsis* species in nature display a dichotomy of behavioral traits or do they in fact fall along a continuum. For instance, does *A. oklahoma* display intermediate levels of aggression and fear between fully arid and fully riparian populations of *A. aperta*? While *A. aperta*, itself, does display some intermediate behaviors in the riparian patch studied, this seems to be a result of gene flow between populations from two selection regimes and does not reflect any intermediacy of the habitat. Can intermediate behaviors be found in intermediate habitats, not as a result of gene flow, or do behaviors always fall into two clear cut levels of aggression and fear? Because behavioral data have already been collected for a number of populations of *A. aperta, A. aleenae*, *A. spatula*, and *A. oklahoma* we should soon be able to address these questions at least in part. However, behavioral analyses should probably also be carried out on a number of populations of fully riparian species, such as *A. naevia*. If *A. oklahoma* shows intermediate levels of aggression or fear between say *A. naevia* and arid populations of *A. aperta*, then the prey availability and predation risk for those intermediate populations should then be measured.

The molecular genetic study of behavioral aggressiveness also needs to be undertaken. For instance, how many loci are involved in determining level of

aggressiveness and where are they found in the genome. Because I have predicted an interaction between history and natural selection governing behavioral adaptations, I would also predict that the same loci are involved in independent riparian populations of *A. aperta* and the other bimodal species, *A. spatula* and *A. aleenae*. One way to get at this question would be to do QTL mapping with F1 and backcross hybrids between riparian and arid populations. However, this type of breeding and mapping is labor intensive and probably could not be done on more than two replicates of riparian and arid populations. Another way to get a coarse scale map of the loci involved would be to employ an approach used by Emelianov et al (2003) that makes use of randomly dispersed, genome wide molecular markers, such as AFLPs (amplified fragment length polymorphism). Essentially, you can use these genome wide markers to search for heterogeneous levels of differentiation among loci between populations. This idea rests on the assumption that gene flow is on going between populations under different selection pressures. In this case, neutral loci would have very low or nonexistent levels of differentiation, while loci under selection (and those loci linked to the loci under selection) should show much higher levels of differentiation. Linkage maps can be made for the AFLP markers using only one or two breeding experiments and then natural populations would be assayed for their levels of differentiation at these various AFLP markers. Using such an approach, a number of replicate adjacent riparian and arid populations could be assayed from *A. aperta, A. spatula* and *A. aleenae*. Same-habitat populations would also be compared with the expectation that these populations would show homogenous levels of differentiation at all AFLP markers. This approach

will provide an answer to the question of whether the same loci are involved in differential behaviors of independent riparian populations from multiple species.

Study of the intraspecific variation of multiple closely related species can also provide an understanding of the patterns and process of speciation. In sexually reproducing animals, the important mechanism of speciation is the evolution of reproductive isolation. Reproductive isolation is thought to often evolve as a byproduct of genetic differentiation in allopatry (Mayr 1963). Thus, we can ask if genetically differentiated populations within a "species" have evolved reproductive isolation. For instance, do geographic clades of *A. aperta* or *A. aleenae* show evidence for reproductive isolation, even though no accompanying genitalic divergence has occurred (for spiders genitalic differences are the hallmark of species). For *A. aleenae*, it is also possible to see if there is a higher level of reproductive isolation between *A. aleenae* populations and *A. spatula* than there is between geographic clades of *A. aleenae*, even though there are similar levels of molecular divergence.

Another speciation problem in *Agelenopsis* is determining the status of the *A. utahana, A. oregonensis* group and the *A. emertoni, A. pensylvanica, A. potteri* group. Are these morphologically defined species in fact reproductively isolated? Or does each group represent a single interbreeding, polymorphic species? Each of these groups shows evidence for paraphyly at the mitochondrial genes used in this dissertation, but to what extent are these results due to hybridization and introgression of the mitochondrial genome? In addition to experimental measures of reproductive isolation, molecular markers from the nuclear genome are also needed to determine current levels of gene flow among the morphologically defined species.

The *Agelenopsis* system also provides the opportunity to examine the evolution of morphological characters, namely genitalic characters, and how differentiation of genitalic characters relates to speciation. While genitalia are considered good species characters for spiders (Eberhard 1985), little empirical work has been done to determine if genitalic characters are in fact discrete, i.e. is there less variation within species than between (but see Eberhard et al. 1998). This is one line of work that can be explored using *Agelenopsis*.

In addition to determining the amount of variation in genitalic characters, it would also be interesting to address the rate of genitalic evolution. Eberhard (1985) proposed the hypothesis that animal genitalia evolve rapidly and divergently in response to female sexual selection. However, phylogenetic studies that included population sampling of two spider genera, *Nesticus* (Nesticidae) and *Hypochilus* (Hypochilidae), found that most speciation events were old (Hedin 1997; 2001), and that populations within species were often highly diverged without any accompanying morphological divergence. The same pattern of high levels of intraspecific genetic divergence without genitalic divergence has been found for other spider species as well (e.g. *Habronattus pugillis* - Masta 2000; *Aptostichus simus* - Bond et al. 2001; *Hypochilus thorelli* - Hedin and Wood 2002). These data suggest that morphological evolution is slow, in opposition to Eberhard's rapid and divergent hypothesis. However, the phylogenetic studies only show that speciation events were old, they do not tell us anything about the time frame of morphological divergence.

Agelenopsis may be a good model system to investigate rates and causes of genitalic divergence because there appears to be a variable rate of evolution. Some

morphologically distinct "species" are quite closely related (i.e. *A. pensylvanica, A. potteri, A. emertoni*), while some morphologically homogenous species have fairly deep intraspecific molecular divergence (i.e. *A. aperta, A. aleenae, A. longistylus*). If sexual selection in fact drives divergence of genitalic characters then it would be interesting to investigate if there is more evidence of sexual selection acting on the closely related species than within species that show high intraspecific genetic divergence. Measuring sexual selection may be difficult to get at though, because the genitalic characters may be confounded with other interspecific differences like pheromones. Another route to take might be to quantitatively measure genitalic shape and size for each species and then estimate the amount of divergence between species of both males and females. Then using the phylogeny of the genus, the rate of divergence for both males and females could be calculated. The expectation under sexual selection by female choice is that rates of male divergence would exceed female divergence. Furthermore, female genitalia should more accurately reflect phylogenetic relationships than male genitalia.

These questions arose from intra- and inter-specific phylogenetic patterns identified in this dissertation. However, the phylogenetic hypotheses presented in this dissertation may not be entirely accurate because only mitochondrial genes were used. Thus, an accurate phylogeny based on multiple independent data sets (such as nuclear loci) is needed to effectively answer evolutionary questions in *Agelenopsis*. However, the various questions presented here show the utility of *Agelenopsis* as a model system to address issues in ecology and evolution.

Literature Cited

Bond, J. E., Hedin, M. C., Ramirez, M. G., and Opells, B. D., 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. Molecular Ecology. 10, 899-910.

Eberhard, W. G., 1985. Sexual Selection and Animal Genitalia. Harvard University Press, Cambridge, MA.

Eberhard, W. G., Huber, B. A., Rodriguez, R.L., Briceno, R. D., Salas, I., and Rodriguez, V., 1998. One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. Evolution. 52, 415-431.

Emelianov, I., Frantisek, M., and Mallet, J., 2004. Genomic evidence for divergence with gene flow in host races of the larch budmoth. Proceedings of the Royal Society of London B, 271, 97-105.

Hedin, M. C., 1997. Speciational history in a diverse clade of habitat-specialized spiders (Araneae: Nesticidae: *Nesticus*): Inferences from geographic-based sampling. Evolution. 51, 1929-1945.

Hedin, M. C., 2001. Molecular insights in species phylogeny, biogeography, and morphological stasis in the ancient spider genus *Hypochilus* (Araneae: Hypochilidae). Molecular Phylogenetics and Evolution. 18: 238-251.

Hedin, M. C. and Wood, D. A., 2002. Genealogical exclusivity in geographically proximate populations of *Hypochilus thorelli* Marx (Araneae, Hypochilidae) on the Cumberland Plateau of North America. Molecular Ecology. 11, 1975-1988.

Masta, S., 2000. Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): Recent vicariance of sky island populations? Evolution. 54, 1699-1711.

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

Vita

Nadia Antoine Ayoub was born in Atlanta, Georgia on June 27, 1976. She attended DeKalb County public schools in Atlanta, GA and graduated from Druid Hills High School in June 1994. She received a Bachelors of Science with Honors in Ecology from University of Georgia, Athens in June 1998. After working as a lab technician in the Veterinary School at the University of Georgia for one year, she entered graduate school at the University of Tennessee, Knoxville in 1999. She received a Doctor of Philosophy degree in Ecology and Evolutionary Biology in December 2004.