Phylogeography, population genetic structure and reproductive biology of Antelope ground squirrels

Joshua Ryan Whorley

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2006

Program Authorized to Offer Degree:
Department of Biology

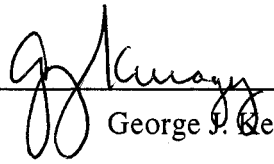
University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Joshua Ryan Whorley

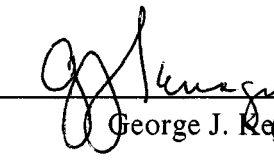
and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the final examining committee have been made.

Chair of the Supervisory Committee:



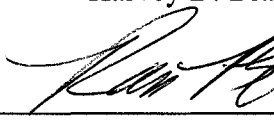
George J. Kenagy

Reading Committee:



George J. Kenagy

Harvey D. Bradshaw



Raymond B. Huey

Date: 17 August 2006

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Joshua R. Wheeley

Date 14 August 2006

University of Washington

Abstract

Phylogeography, population genetic structure and reproductive biology of Antelope ground squirrels

Joshua Ryan Whorley

Chair of the Supervisory Committee:
Professor George J. Kenagy

Zoology

The genetic structure of populations over a wide geographic area should reflect the demographic and evolutionary processes that have shaped a species across its range. Similarly, variation in physiological traits should reflect the degree to which a species can either tolerate, through behavioral or phenotypic plasticity, or adapt to local environments across its range. I examined the distribution and diversity of both neutral genetic variation and reproductive physiology within the Antelope ground squirrel, *Ammospermophilus leucurus*.

In my first and second chapters I examine the population genetic structure of Antelope ground squirrels across the complex of North American deserts from the Great Basin of Oregon to the cape region of the Baja California peninsula. I estimated phylogenies of both the mitochondrial cytochrome-*b* gene and control region gene to

infer population structure and history. The gene trees revealed genetic uniformity in a monophyletic northern clade, and a more diverse, monophyletic southern clade. The distribution and diversity of mutations plus the minimal geographic structure of the northern clade suggest a rapid northward expansion of the population that must have followed a northward desert habitat shift associated with the most recent Quaternary climate warming and glacial retreat. The more variable distribution of pairwise differences and higher haplotype diversity of the southern suggests a longer, more stable history associated with a southern peninsular refugium.

I then examined the diversity and distribution of additional mtDNA haplotypes collected at the phylogeographic break between northern and southern clades, in the Vizcaíno desert region. I also analyzed microsatellite allele distributions to determine whether members of different clades are currently interbreeding or remain reproductively isolated. I recovered no unique haplotypes or hidden diversity but found that most of the total genetic diversity of the previously sampled region (22° Latitude) is contained in the small, 4° Latitude Vizcaíno region. Microsatellite allele distributions suggest that animals from the mid-peninsula represent a single panmictic population, regardless of what mtDNA haplotypes they carry.

In my third chapter, I discuss how the geographic distribution of genetic diversity and reproductive traits within *A.leucurus* compares to an existing gradient of environmental conditions. Gestation and lactation of an individual female require about three months, and individuals breed only once per year. The overall breeding season of populations are relatively short in Oregon and California, but extends over nearly half a year in the southern Baja California peninsula. Litter size decreases significantly from

north to south. A major shift in environmental seasonality, biotic diversity, and ecosystem composition are associated with the longer breeding season and lower litter size at the southern geographic extreme of *A. leucurus*.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Number Page

LIST OF TABLES

Table Number Page

ACKNOWLEDGEMENTS

This dissertation was supported by the Mammalogy Program at the Burke Museum, the Graduate School Fund of the University of Washington, the UW Biology Department's Snyder Award, a Sigma Xi fellowship, and two Grants-in-Aid fellowships from the American Society of Mammalogists. I thank my committee members Ray Huey, Jim Kenagy, Toby Bradshaw and Carl Bergstrom. For advice on the research and manuscript I thank Brian Arbogast, Scott Edwards, Sacha Vignieri, and Xiaoguang Zheng, as well as Mary Kuhner for advice concerning FLUCTUATE. For being incredible friends and colleagues, the EPSMC: Sacha Vignieri, Chris Turner-Himes, Corey Welch, Dou Shuan-Yang, Brian Arboghast, Kathy Hall, and Xiaoguang Zheng. For field-work along the way: Adam Augustyn, Joe Booth, Mayra de La Paz Cuevas, Patrícia Cortez-Calva, Jeff Bradley; Lola Ulvog; The employees at "my other job" at the IMA Climbing Center, Jeremy Bowler and Scott Whitcomb for their motivation and excitement to climb ALWAYS; Kari Stiles for teaching me to climb; Ai Nguyen for her support, love and friendship.

DEDICATION

Since they've been in my mind for many years and are in various ways responsible for my choice to do this: for Fran Berman for teaching me that good writing is explicit, specific and sleek...I tried to not forget that; for Harry Greene and Jim Patton for showing me that if you just sit down and watch wild creatures, they'll tell you what questions need answering; for my advisor, Jim Kenagy, for everything one can imagine. He's the most generous, thoughtful, gentle person I have ever met; for my Mom for support of all kinds whenever I needed it.

Chapter I: Genetic structure of desert ground squirrels over a 20-degree-latitude transect from Oregon through the Baja California peninsula

Introduction

Identifying and interpreting patterns of intraspecific geographic variation provides an ongoing forum for discussion about the relative roles of ecological and evolutionary processes in generating diversity. Darwin (1859) observed that phenotypic variation often coincides with geographic distance. Mayr (1956) suggested that ecogeographic variation has an adaptive basis and may represent a response to clinal variation in environmental conditions across populations. Bergmann's rule (1847), as originally applied intraspecifically to homeotherms, serves as a familiar example of a geographic pattern in adaptation; this "ecogeographic rule" explains larger body size at higher latitude as a response to colder climatic conditions. This suggests that temperature or other environmental characteristics that vary systematically over a geographic gradient could, in principle, generate adaptive variation. The evolutionary response of populations to an environmental gradient may occur either by genotypic adaptation or by evolution of phenotypic plasticity (Via 1994). Mayr (1963) argued and others have demonstrated (Kirkpatrick and Barton 1997) that gene flow should swamp adaptation in such systems. However, other theoretical treatments suggest that selective clines along geographic gradients might maintain adaptive genetic and phenotypic variation (Haldane 1948; Fisher 1950; Kimura & Maruyama 1971) and even promote speciation (Endler 1977).

Although environmental gradients may drive local adaptive evolutionary processes, geological and climatological changes can mold more complex patterns of

genetic diversification over a broader regional scale and a longer time scale (Arbogast & Kenagy 2000; Conroy & Cook 2000; Riddle *et al.* 2000; Matocq 2002). Changes in population size, levels of ancestral polymorphism, and patterns of gene flow can shape the genetic structure of populations and thereby influence their ability to adapt to different environmental settings. As these demographic processes are also the primary modulators of neutral variation, we can better understand population and species-level responses to historical changes of habitat and topography by examining variation at loci mtDNA loci. We use gene geneaologies, which we conceive of as correlated to, but not absolute indicators of population history (Edwards and Beerli, 2000) to infer past levels of population connectivity. This phylogeographic approach (Avise 2000) has been applied to a variety of vertebrate taxa, both individual species (Riddle 1995; Zamudio *et al.* 1997; Ashton & de Queiroz 2001) and regional biotas (Zink 1996; da Silva & Patton 1998; Riddle *et al.* 2000). We also employ new methods that allow estimation of relative times of population separation and divergence within species with greater accuracy, taking into account multiple sources of variance. By tracing demography and historical patterns of gene flow one can begin to assess the potential for adaptive divergence in allopatry (Zheng *et al.* 2003).

To explore the evolution of populations in response to environmental variation, we chose a small mammalian species with an extensive geographic range that includes a lengthy north-south gradient across the four desert ecosystems of western North America (Shreve 1942; MacMahon & Wagner 1985). This latitudinal gradient contains variation in temperature, periodicity and abundance of rainfall, and vegetation types. Thus it represents a good system for investigating the connectivity of populations and

the potential for adaptive differentiation in allopatry. The antelope ground squirrel, *Ammospermophilus leucurus*, is one of the latitudinally most wide ranging mammals in North America, and it is the most widely distributed desert rodent of the family Sciuridae in North America (Hall 1981). It ranges across the complex of North American deserts from the Great Basin Desert in southeastern Oregon, at 43° North Latitude, southward through the Mojave and Sonoran Deserts, and into the Peninsular Desert of Baja California to the Cape Region, at 22° North Latitude.

To assess the potential for local adaptation and to recover details of population history, we obtained samples of *A. leucurus* over a 2200-km transect crossing its geographic range. We examined geographic variation in genetic structure using mitochondrial DNA sequences, both the coding cytochrome-*b* gene and the non-coding control region. We use these data to provide insights into how historical changes in climate and habitat distribution may have influenced genetic structure across the vast environmental gradient that spans the north-south axis of this species' geographic range. We investigated the alternative hypotheses that conspicuous genetic variation would be associated either with a gradient of ecosystem characteristics or with vicariance tied to historical barriers that may have intervened across parts of the transect.

Materials and methods

Sampling

We obtained tissues of 73 specimens from 46 localities in western North America, spanning nearly 20° latitude. We consolidated the specific localities into 13 groups ("areas") based on ranges of about 1° latitude (Table 1.1, Fig. 1.1). Tissue samples are associated with voucher specimens from the following museums: Burke Museum, University of Washington (UWBM); Colección de Mamíferos, Centro de Investigaciones Biológicas (CIB); Museum of Southwestern Biology, University of New Mexico (UNM), and U.S. National Museum, Smithsonian Institution (USNM).

Laboratory techniques

DNA was extracted from liver or kidney tissue using the DNeasy Tissue Kit (QIAGEN, Valencia, CA). We used PCR to amplify 555 base pairs (bp) of the mitochondrial cytochrome-*b* gene for all 73 individuals. The primers used were L14724 (5'-CGAAGCTTGATATGAAAAACCATCGTTG-3'), H15906 (5'-CATTTCCGGTTTACAAGACCAGTGTAAT-3'), and L15162 (5'-GCAAGCTTCTACCATGAGGACAAATATC-3'; Irwin *et al.* 1991), and H15149 (5'-AAACTGCAGCCCCTCAGAATGATATTTGTCCTCA-3'; Kocher *et al.* 1989). We ran standard PCR reactions in a total volume of 25μL containing 10x PCR buffer, 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.0 μM of each primer, 1 unit of *Taq* DNA polymerase (Hoffmann-La Roche Inc., Nutley, NJ), and 1-2 μL genomic DNA. PCR reactions were carried out in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems, Foster City, CA) as follows: 95°C for 2 min; 30 cycles of 95°C for 30 sec,

50°C for 30 sec, and 72°C for 30 sec; and 72°C for 3 min. With each round of PCR reactions we included one negative control to check for contamination in this and the following.

We also amplified 511 bp of the control region for all 73 individuals. The primers used were CTRL-L (5'-CACYWTYAACWCCCAAAGCT-3'; Bidlack & Cook 2001) and H16498 (5'-CCTGAAGTAGGAACCAGATG-3'; Kocher *et al.* 1993). PCR reactions were run in a total volume of 25μL containing 1 x PCR Buffer, 2.0 mM $MgCl_2$, 0.2 mM of each dNTP, 1.0 μM of each primer, 1 unit of *Taq* DNA polymerase (Hoffmann-La Roche, Inc.), and 1μL genomic DNA. The PCR reactions were carried out in a GeneAmp PCR System 9700 thermocycler (Applied Biosystems) as follows: 94°C for 45 sec; 35 cycles of 94°C for 10 sec, 50°C for 15 sec, and 72°C for 30 sec; and 72°C for 3 min.

PCR products were purified, for use in sequencing reactions, with a QIAquick PCR Purification Kit (QIAGEN). Sequencing reactions were carried out in a volume of 10μL with the forward primers using a Big Dye Terminator Cycle Sequencing Ready Reaction Mix (Applied Biosystems). Samples were run on an ABI 377 automated sequencer (Applied Biosystems). To check the quality of obtained sequence data, in one direction, we sequenced some individuals twice and found identical outcomes. Sequences were aligned using SEQUENCHER 4.1 (Gene Codes Corp., Ann Arbor, MI). All sequences were deposited in Genbank (accession numbers XXX-ZZZ).

Phylogenetic analysis

We used the neighbor-joining method in PAUP* 4.0b10 (Swofford 2002) to reconstruct phylogenetic relationships among individuals for both cyt-*b* and control region genes. By applying hierarchical likelihood ratio tests with the program MODELTEST 3.06 (Posada & Crandall 1998), we determined that the HKY model of substitution (Hasegawa *et al.* 1985) plus a gamma distribution (Γ) of rate heterogeneity across all sites provided the best fit to the cyt-*b* data set. The estimated parameters under this model were Γ=0.0928, Ti/Tv=5.1938. For the control region, we found that the HKY model plus invariable sites (I) and a Γ distribution of rate heterogeneity across variable sites provided the best fit. The estimated parameters under this model were Γ=0.8382, I=0.7582, Ti/Tv=6.0700. We used bootstrap analyses with 1000 replicates to evaluate nodal support, as implemented in PAUP* 4.0b10. For both markers we constructed unrooted trees. As a comparison to the neighbor-joining trees, we also constructed trees using maximum likelihood methods in PAUP* 4.0b10, which yielded similar topologies and slightly longer branch lengths.

Molecular diversity

Mean nucleotide diversities within and between clades were calculated with the program ARLEQUIN 2.000 (Schneider *et al.* 2000). We used the substitution model of Tamura and Nei (1993) with gamma corrections of 0.0928 for cyt-*b* and 0.8382 for control region, as estimated with MODELTEST. Mean nucleotide diversity of cyt-*b* data between the two clades was used to estimate divergence while taking into account ancestral polymorphism using the equation for molecular diversity $D_{xy} = D\text{-}0.5x(D_x+D_y)$ (Nei & Li 1979), where D_x and D_y are mean molecular diversities within two different

clades respectively; D is the mean nucleotide diversity between two clades, and D_{xy} is the corrected molecular diversity between two clades.

With ARLEQUIN 2.000 one directly calculates D_x and D_y, but not D. We combined the two clades into one single population (t) and calculated mean nucleotide diversity for this population (D_t) using ARLEQUIN 2.000. Then we used D_x, D_y, and D_t to compute D by the following formula:

$$D_t N_t(N_t-1)/2 = DN_xN_y + D_xN_x(N_x-1)/2 + D_yN_y(N_y-1)/2$$

where N_x, N_y and N_t are sample sizes of populations x, y and t, respectively, and where $N_t = N_x + N_y$.

Historical population dynamics

With the program ARLEQUIN 2.000 (Schneider *et al.* 2000) we calculated Fu's F_s to test for demographic expansion (Fu 1997). This test statistic evaluates the probability of observing a random neutral sample of sequences with a number of alleles similar to or smaller that the observed value. Ramos-Onsins & Rojas (2002) demonstrated that Fu's F_s is better for detecting population growth than other tests, so we used it as an indicator of growth.

We also used ARLEQUIN 2.000 to compute frequencies of pair-wise differences between haplotypes (the mismatch distribution) to evaluate the hypothesis of recent population growth and to estimate growth parameters. Recent growth is expected to generate a unimodal distribution of pairwise differences (Slatkin & Hudson 1991; Schneider & Excoffier 1999). ARLEQUIN uses a non-linear least squares approach to estimate parameters for a stepwise growth model: $\theta_0 = 2uN_0$ (before expansion), $\theta_1 =$

$2uN_1$ (after expansion) and $\tau = 2ut$ (time of expansion). Notice that $u = m_T\mu$ is the mutation rate for the entire DNA sequence under study, where m_T is the number of nucleotides of the sequence, and μ is the mutation rate per nucleotide (Rogers & Harpending 1992). N_0 and N_1 are effective population size of females before and after expansion, respectively.

Approximate confidence intervals for τ, θ_0, and θ_1 are obtained by a parametric bootstrap approach (1000 replicates). The validity of the stepwise expansion model is tested using the same parametric bootstrap approach by a goodness-of-fit statistic P = (number of SSD_{sim} larger or equal to SSD_{obs}) / B. SSD_{sim} is the sum of squared deviation (SSD) between the simulated mismatch distribution and the model expectation, SSD_{obs} is the SSD between the observed mismatch distribution and the model expectation, and B is the number of simulated samples (ARLEQUIN Version 2.000 Manual; Schneider *et al.* 2000).

To produce an additional estimate of population expansion we used the program FLUCTUATE 1.3 (Kuhner *et al.* 1998), based on coalescence, to generate maximum likelihood estimates of a present-day value of $\Theta(\Theta = 2\mu N_F)$ and exponential growth rate (g). Here μ is mutation rate per nucleotide, and N_F is present-day effective population size of females. FLUCTUATE 1.3 uses a Markov chain-Monte Carlo approach with Metropolis-Hastings importance sampling to make estimates of Θ and g by searching through genealogies of highest probability. Transition/transversion ratio for the northern clade was set to 2.6035 and 7.3875 for cyt-*b* and control region data, respectively, and to 14.0429 and 7.4047 for the southern clade. These values were obtained from MODELTEST 3.06. To begin the sampling on genealogies of high probability, we

provided FLUCTUATE 1.3 with initial trees constructed using the programs dnadist and Neighbor within Phylip 3.753c. For each population and each gene we ran 10 iterations of the program. For the cyt-*b* data we ran 10 short chains of 1500-4000 steps each, depending on the information content of the sequences, and two long chains of 15,000 steps each. Sampling increment was set to 20. Using control region data we ran 10 short chains of 1500 steps and two long chains of 10,000 steps. This strategy resulted in low variability among intermediate estimates of Θ and g.

Results

To investigate the genetic structure of *Ammospermophilus leucurus* populations over the entire 2200-km transect, we sampled 73 individuals from 13 major locality areas from Oregon to the southern tip of the Baja California peninsula (Table 1.1, Fig. 1.1). Evidence from both the cytochrome-*b* gene and the control region indicates two distinct clades over the north-south axis of the geographic range (Fig.1.2). The genetic distances between the two clades are 0.022 substitutions per site for cyt-*b* and 0.046 for control region. The southern 25% of the range is occupied by a relatively heterogeneous clade, whereas the northern 75% of the range contains a more homogeneous clade (locality areas in Fig. 1.1b). Northern clade haplotypes do not cluster geographically, except for a relatively well-supported monophyletic sub-tree within northern Baja California (locality areas F, G, H, Fig. 1.2a and b). The southern clade shows less haplotype sharing among locality areas than the northern clade. For cyt-*b* the geographically more restricted southern clade contains 14 haplotypes among 29 individuals, whereas the more geographically expansive northern clade holds only 18 haplotypes among 44 individuals (Fig. 1.2a). Despite the much larger area occupied by the northern clade, only 41% of the individuals had unique haplotypes, whereas 48% of the individuals of the southern clade had unique haplotypes.

Mean nucleotide diversities for the cyt-*b* gene and control region were similar within the northern and the southern clades (Table 1.2). For cyt-*b*, the nucleotide diversity between the two clades was about an order of magnitude greater than that within clades (Table 1.2). The corrected nucleotide diversities are slightly higher than

those shown in the neighbor-joining trees (Fig. 1.2) and closer to those generated by maximum likelihood analysis.

The F_s statistic indicated population expansion in both the northern and southern clades, with $F_s = -12.14$ ($P < 0.001$) and -11.58 ($P = 0.001$), respectively. The observed distribution of pairwise differences does not differ significantly from the simulated and modeled Poisson distributions, indicating recent demographic expansion in both clades (Fig. 1.3). The lack of well-supported branches within each clade (Fig. 1.2), as well as the preponderance of short terminal branches, also suggests recent expansion.

To investigate historical population dynamics, we applied two models of expansion to our control region data. This allowed us to estimate effective female population size and rate or time of expansion. The exponential expansion model indicates rapid increase in population size of both clades (Table 1.3a). The present-day Θ is larger for the southern clade than for the northern clade under this model. This indicates a larger effective population size and consequently more polymorphism. Comparison of the growth parameter *g* indicates demographic expansion of both clades, and the model also estimates a more rapid expansion of the northern clade than the southern clade.

The stepwise model (Table 1.3b) estimated periods of population expansion τ of 9.936 for the southern clade and 6.042 for the northern clade, suggesting that the northern clade expanded relatively recently while the southern clade had a longer, more stable history in the southern Baja California peninsula. The estimated θ_0 before expansion is much smaller than after expansion (θ_1) for both clades, but the estimated θ_1, reflecting current effective population size, is much larger in the south than in the

north. This indicates faster and more recent demographic expansion of the northern clade. This stands out further considering that the northern clade extends over nearly three times the latitudinal distance of the southern clade. Interestingly, the estimated northern θ_0 is larger than the southern θ_0.

Discussion

Ammospermophilus leucurus consists of two distinct mitochondrial clades. Together, they cover the 20°-latitude axis of the species' geographic range, which represents a transect spanning several major desert ecosystems within western North America. The geographically extensive northern clade stretches from Southern Oregon to the northern edge of the Viscaíno Desert in the central Baja California peninsula. The much smaller southern clade occupies only the southern peninsula. The southern clade contains a higher percentage of unique haplotypes and a greater overall haplotype diversity than the northern clade. Our genetic analysis suggests that population size in both clades has expanded recently, and that geographic expansion of the northern clade must have followed a northward desert habitat shift associated with the most recent Quaternary climate warming and habitat shift following glacial retreat. The southern mtDNA clade shows a longer expansion time than the northern clade, and the greater haplotype diversity within the southern clade also suggests a larger effective population size and longer residence time in the southern peninsula. Population demographic models also suggest that the northern clade had a higher rate of population growth than the southern. Estimates of gene-trees and population growth support the hypothesis of rapid northward expansion of the northern clade, probably from a small population in northern Baja California. The occurrence of geographically widespread, similar haplotypes throughout the northern clade bears this out, especially north of the US-Mexican border. The comparative study of the northern and southern population groups will be of interest for exploring possible adaptive divergence in allopatry.

Phylogenetic differentiation and population expansion

Both of our mitochondrial gene trees demonstrate the existence of two distinct clades separated by a relatively high number of mutations. No haplotypes are shared between the two clades, and within-clade divergence is low compared to between-clade divergence. The between-clade nucleotide divergence for cyt-*b* is 0.029 and for control region is 0.043, which accounts for 85% and 69% of the total nucleotide divergence, respectively. By comparison with other mtDNA data sets on rodents, this level of cyt-*b* divergence falls in the middle of the range of within-species comparisons and at the low end of the range for sister species (Bradley & Baker 2001; Spradling *et al.* 2001).

The northern clade is genetically homogeneous, with short terminal branches separating haplotypes that are widespread and closely related. This tree structure also suggests either recent demographic and range expansion or high rates of gene flow relative to mutation. One exception to the general homogeneity of the northern clade is the sub-tree within the northern clade that consists only of haplotypes from the northern Baja California peninsula (areas F, G, H). This section of the northern clade is monophyletic and has relatively longer branches than others and higher bootstrap values separating the branches, especially on the control-region tree (Fig. 1.2). This suggests that the northern clade may have been restricted to northern Baja California for some period before expanding northward to fill its current range. The southern clade contains greater haplotype diversity than the northern, even though its latitudinal extent is only about one third as great. It also has longer branches—relative to the northern clade--separating haplotypes, suggesting that it has been evolving at demographic equilibrium within the southern peninsula for a longer time.

The greatest genetic break in our mtDNA phylogenies occurs at the middle of the Baja California peninsula, near the border between the states of Baja California Norte and Baja California Sur, and falls near the 28° parallel (Fig. 1.2). This result agrees with that of Riddle *et al.* (2000), who initially identified a similar mid-peninsular pattern of phylogeographic structuring in a number of Baja California vertebrates, including *A. leucurus*. In our study of the full latitudinal extent of *A. leucurus* populations, we found the mid-peninsular break to be the only strong genetic break over the entire transect. The structure of our observed gene tree fits a hypothesis of recent demographic and geographic expansion northward from a refugial population of small effective size in the northern peninsula. Based on emerging views of phylogeography (Avise 2000), one can expect low diversity and lack of geographic structure for mitochondrial genes in three situations that may pertain to our system -- (1) recent demographic and geographic expansion from an ancestral population of small size, (2) historically stable population size and geographic range, but with geographically variable levels of gene flow between populations, and (3) a selective sweep, carrying to fixation only a few of the total number of mitochondrial haplotypes in the population (Bertorelle & Slatkin 1995). Unfortunately, we cannot test this last hypothesis against the first two. However, to investigate the first two hypotheses, we subjected our control-region data to treatment by two models of demographic expansion, one that assumed a sudden step-wise expansion (using the program ARLEQUIN, Schneider *et al.* 2000) and another that assumed exponential expansion (using FLUCTUATE, Kuhner *et al.* 1998).

Our first model uses pair-wise differences between sampled DNA sequences (mismatch distribution) to analyze population history (Slatkin & Hudson 1991). Both

theoretical explorations and practical analysis of human DNA sequences have demonstrated that population history leaves detectable impacts on the form of the mismatch distribution, especially when populations are not at demographic equilibrium (DiRienzo & Wilson 1991, Schneider & Excoffier 1999). When Schneider & Excoffier (1999) extended the original method of Rogers & Harpending (1992) to include rate heterogeneity across sites, they found that reasonable estimates of both expansion time τ and initial population size θ_0 are achieved without much bias, whereas estimates of θ_1 suffer from upward bias and highly conservative 95% confidence intervals. Such estimates also reveal that the parameters of older expansions are more precisely recovered than those of more recent expansions. Such difficulties may account for our surprising estimates of a higher θ_0 for the northern clade than for the southern clade. Based on haplotype divergence and diversity, we expected that the southern clade should have the larger initial population size.

Demographic expansion should produce a unimodal distribution of pairwise differences because populations that have recently expanded show fewer coalescent events and have pairwise differences that fall in an intermediate range (Rogers & Harpending 1992). The mismatch distribution for both clades conforms to this shape quite well. Here again, such a distribution can arise by rapid spread of a selectively advantageous mtDNA haplotype (DiRienzo & Wilson 1991; Bertorelle & Slatkin 1995) or by homoplasy due to high mutation rates at some sites (Lundstrom *et al.* 1992). A population bottleneck can also generate a unimodal distribution with elevated upper-tail probabilities (Rogers & Harpending 1992). However, because the divergence time in our system is relatively small (at most several hundred thousand years, based on the

range of independent estimates of mutation rates of cyt-*b*; Arbogast *et al.* 2002), and the sampled area is so large, homoplasy due to sequence saturation is unlikely to have generated the unimodal mismatch distributions. Selection is also unlikely to have confounded our analyses, because intraspecific selection on the mitochondrial genome may not be strong enough to promote complete replacement of unfavored alleles over such a large geographic area. Barring the influence of these confounding factors, our observed mismatch distributions from both clades fit the simulated distributions. Therefore, we accept the hypothesis of demographic expansion of the northern clade of *A. leucurus*.

Because incorporating phylogenetic information into estimates of population growth makes more efficient use of data than analysis only of pairwise differences (Felsenstein 1992), we also used FLUCTUATE to estimate population size and growth rate. This method allows us to account explicitly for homoplasy at rapidly evolving sites. We qualitatively compared directions and magnitudes of change of the various parameters between this method and that of ARLEQUIN. Estimates obtained by FLUCTUATE indicate that the southern clade has a larger present-day effective population size Θ than the northern clade, in agreement with estimates obtained from ARLEQUIN. The ratios of Θ for southern to northern populations are similar, with FLUCTUATE estimating 2.5:1 and ARLEQUIN estimating 4.2:1. The growth rate g indicates more rapid expansion of the northern clade than the southern. These estimate are robust, as multiple iterations of FLUCTUATE generated comparable estimates with sharp likelihood surfaces (data not shown). Estimates of both Θ and g are biased

slightly upward due to cross correlation between the two (Kuhner *et al.* 1998), but correction of this bias does not alter our conclusion.

In the context of western North America's deserts and their history, the northern *A. leucurus* clade appears to have expanded both demographically and geographically, most recently following the last glacial maximum. Our two models of population growth both estimate that the northern clade grew more rapidly and more recently than the southern clade. The contemporary western deserts and their associated biotic communities are of recent origin, and Pleistocene glaciations and climate fluctuations produced southward geographic shifts of the major North American desert habitats (Van Devender *et al.* 1987; Pielou 1991). Assuming that the northern extent of the range of *A. leucurus* has been limited by habitat and climate, as it appears to be today, we conclude that recolonization of the Great Basin desert must have occurred since the last glacial maximum.

Historical biogeography

Our phylogenetic trees together with estimates of population growth indicate rapid expansion of the northern clade of *A. leucurus* from populations restricted to the northern half of the Baja California peninsula. Supported by mtDNA evidence from several vertebrate taxa, Upton & Murphy (1997) and Riddle *et al.* (2000) have offered an explanation of a recurrent pattern of molecular divergence between taxa of the northern and southern Baja California peninsula. They have proposed the formation of a midpeninsular seaway about 1.5 million years ago, which would have isolated formerly continuous populations of these taxa. Under the general scenario of mid-peninsular

vicariant division accounting for the current haplotype distributions, we see two possible explanations for the widespread but genetically homogeneous northern clade. (1) A small panmictic population was isolated immediately north of a mid-peninsular geographic barrier, and some time later that population expanded rapidly northward. (2) The populations of the entire northern peninsula have been structured since the time of a mid-peninsular barrier, whereas populations further to the north were and remain relatively unstructured.

The distributional history of fossil *Ammospermophilus* during the Pliocene and Pleistocene, in relation to climate cycles and associated habitat shifts, suggests that explanation (1) is more likely. *Ammospermophilus* populations were apparently widespread in the Pliocene and early Pleistocene. An extreme southern record for *Ammospermophilus* exists in Baja California Sur from the late Pliocene (Miller 1980), and the extreme northern record is from the Pliocene of south-central Washington (Gustafson 1978). Pliocene fossil *Ammospermophilus* also occur in eastern Oregon's Juntura Basin (Black 1963). This broad fossil distribution contrasts with the pattern seen today and throughout the Pleistocene. No *Ammospermophilus* live in Washington today, and populations of *A. leucurus* barely penetrate into the extreme SE corner of Oregon. Although *A. leucurus* fossils dating to the late Pleistocene occur in the southern Mojave Desert (Goodwin & Reynolds 1989; FAUNMAP Working Group 1994), they are notably absent from the Pleistocene record further north in the Great Basin (Grayson 1993), where, instead, the fossil record is rich with montane rodents. The most likely explanation is that habitat shifts associated with climate change pulled the northern range limits southward throughout the Quaternary, as it did for many other North

American mammals (Pielou 1991). The most recent (Cordilleran) ice sheet reached its southern extreme about 18,000 years ago (Booth 1987). The northern range limit of *A. leucurus* has apparently expanded and contracted, shifting northward and southward, a number of times, during the Pleistocene. Paleobotanical evidence corroborates these kinds of climate-driven habitat shifts; the plant-species contents of packrat (*Neotoma* spp.) middens throughout western deserts indicate that much of the vegetation currently associated with Great Basin populations of *A. leucurus* was not present in the Mojave and Great Basin until at least 14,000 years ago (Van Devender *et al.* 1987).

Adaptive potential along environmental gradients

The north-south axis of *A. leucurus* in western North America represents a system that invites further investigation into intra-specific population genetics and adaptation across an environmental gradient. Theoretical treatments of genetic divergence and speciation (Felsenstein 1981) and laboratory experiments that simulate speciation (Rice & Hostert 1993) suggest that speciation can be readily achieved in allopatry, but that it requires intense, multifarious selection to develop in sympatry. Endler (1973) and May *et al.* (1975) demonstrated that sharp, adaptive, genetic clines can develop and be maintained in the presence of gene flow, but it is unclear whether selective forces in the wild are sufficiently strong and consistent to promote divergence among wild populations (Hoekstra *et al.* 2001; Kingsolver *et al.*2001). Although our current expectations remain that genetic divergence is more likely in allopatry than in sympatry, it is of additional interest to explore the complementary role of phenotypic

adjustment in the response of a species of any environmental gradient that exists over its geographic range.

Although we conclude that the major genetic divergence between two groups of *A. leucurus* populations (mtDNA clades) probably developed with populations in allopatry, associated with an historic geographic barrier, it is theoretically possible that the break was also promoted by selection across an environmental gradient. Alternatively, divergence may have occurred in sympatry across a permeable geographic boundary. We can next ask whether the apparent historic boundary of the mid Baja California peninsula is still being maintained, and, if so, what isolating mechanisms are currently operating. A closer inspection of genetic patterns at the mid-peninsular interface of the two clades will also be relevant to understanding the population biology of a potential zone of hybridization and the processes that lead to discordance between mtDNA and nuclear DNA. We plan to address these questions using additional molecular markers across the mid-peninsular transition.

The exploration of adaptation (Reznick & Travis 1996) is also of interest across the geographic range of *A. leucurus*. Because this system consists of a cline of environmental gradients that follows a linear range of populations, we can examine geographic patterns of genotypic and phenotypic differences and ask whether they are influenced by the cline, and, if so, why. Environmental conditions differ substantially between the southern, middle, and northern parts of the range of *A. leucurus*, and this undoubtedly produces phenotypic differences (Kenagy & Bartholomew 1985; Kenagy *et al.* 2004). We will explore the nature of the environmental changes over the

latitudinal gradient, and then ascertain whether selection has actually produced any genotypic changes.

As an example, the reproductive ecology of *A. leucurus* over its latitudinal range holds a potential for analysis of life-history adaptation across a range of populations. Breeding by *A. leucurus* in the relatively more predictable winter-rainfall environment and floristically simple ecosystem of the Northern Mojave Desert consists of a single, precisely timed annual bout with a relatively large litter (Kenagy & Bartholomew 1985). On the other hand, breeding in the unpredictable summer monsoonal environment and associated floristically diverse ecosystem of the southern Baja California peninsular desert consists of a broad and protracted season of sporadic reproduction with a relatively small litter size (Kenagy *et al.* 2004). The phenotypic and genotypic components of life-history features such as these over the geographic range can be identified through further empirical study. From the review of Reznick & Travis (1996), several patterns are clearly possible: a match of geographic variation in traits with the location of neutral genetic divergence, complex patterns of trait variation at sharp environmental boundaries, or gradual variation of the traits over the course of the environmental gradients. Few previous investigations have successfully addressed the complex interactions of organismal function and genetics that represent geographic adaptation (Powers & Schulte 1998), but such an approach seems to be a reasonable way to investigate the challenging problem of selection and adaptation in natural populations over geographic space (Endler 1986).

Development of our views of geographic evolution and adaptation of populations arose from classic approaches that correlated morphological variation with

differences in habitat. From such correlations it was originally concluded that ecological variation or gradients promoted morphological variation, which may in fact may have been either phenotypic or genotypic. Mitochondrial DNA analysis across geographic space, on the other hand, is capable of generating patterns of substantial variation that are cryptic by comparison to original observations of invariable morphological characters across the same geographic space (Wilcox *et al.* 1997; Baric & Sturmbauer 1999; Bond *et al.* 2001). In the mean time we also have strong inferential examples of adaptive evolutionary divergence across geographic space, where gene flow has been constrained by a vicariant barrier, and ecological diversification may be post-speciational (Peterson *et al.* 1999). The future study of geographic evolution and adaptation appears to be complex. It will benefit from multiple molecular perspectives, from understanding of the geographic facilitation or inhibition of gene flow, from comparisons of morphological and life history data, and from assessments of the roles of phenotypic vs. genotypic responses.

In his classic essay on geography and evolution of pocket gophers (*Thomomys*), Joseph Grinnell (1926) pointed out the enormous geographic variation in body size and coloration of these rodents associated with habitat variation. He concluded that "the evolution of habitats (differentiation areas) must have preceded differentiation of the gopher stocks which came eventually under their impress." It is interesting, by contrast, that Grinnell also assessed geographic variation of antelope ground squirrels, the same species we have studied here. He reported over the continuous range of the species in California that "there are no hindrances to continuous mixing of breed, such as seem essential to subspecific differentiation" (Grinnell & Dixon 1919). Likewise, from the

southern to northern extreme of the state of Nevada, Hall (1946) reported for *A. leucurus* "I cannot detect any geographic variation...in coloration, size of animal, or in proportions of the skull...." These early interpretations of naturalists appear to be valid, and from the molecular perspective that we have now added, we can further conclude that the historical fluctuations north and south in the geographic range of *A. leucurus* have probably been responses to a common general habitat type in which *A. leucurus* lives, and that morphological diversification, as known in other rodents such as pocket gophers, has not been a part of the success of *Ammospermophilus* across the deserts where it is found.

Table 1.1 Locality areas (n=13), specific localities (n=46), and specimens (n=73) of *Ammospermophilus leucurus* in the present study

Locality Area	Locality number	State	Locality: County or local name	Latitude	Longitude	Specimen numbers
A	1	OR	Harney County	42° 17' N	118° 39' W	UWBM 74589
A	2	OR	Harney County	42° 17' N	118° 40' W	UWBM 74591
A	3	OR	Harney County	42° 16' N	118° 40' W	UWBM 74554, UWBM 74561
A	4	OR	Harney County	42° 15' N	118° 40' W	UWBM 74556, UWBM 74559
A	5	OR	Harney County	42° 14' N	118° 39' W	UWBM 74587
B	6	NV	Pershing County	40° 11' N	118° 25' W	UWBM 74597, UWBM 74598, UWBM 74599
B	7	NV	Pershing County	40° 09' N	118° 24' W	UWBM 74594, UWBM 74595, UWBM 74596
B	8	NV	Pershing County	40° 08' N	118° 24' W	UWBM 74593
C	9	CA	Inyo County	37° 12' N	118° 15' W	UWBM 74607, UWBM 74615
C	10	CA	Inyo County	37° 12' N	118° 15' W	UWBM 74624
C	11	CA	Inyo County	37° 11' N	118° 14' W	UWBM 74600, UWBM 74602, UWBM 74605
C	12	CA	Inyo County	37° 11' N	118° 15' W	UWBM 74611
D	13	CA	San Bernardino County	34° 40' N	116° 42' W	UWBM 74626
D	14	CA	San Bernardino County	34° 40' N	116° 42' W	UWBM 74628, UWBM 74631, UWBM 74633
D	15	CA	San Bernardino County	34° 40' N	116° 42' W	UWBM 74636
D	16	CA	San Bernardino County	34° 40' N	116° 42' W	UWBM 74637
D	17	CA	San Bernardino County	34° 36' N	116° 45' W	UWBM 74629
E	18	CA	Imperial County	32° 50' N	114° 51' W	UWBM 74657
E	19	CA	Imperial County	32° 49' N	114° 51' W	UWBM 74645
E	20	CA	Imperial County	32° 49' N	114° 49' W	UWBM 74649, UWBM 74653
E	21	CA	Imperial County	32° 48' N	114° 50' W	UWBM 74639
E	22	CA	Imperial County	32° 48' N	114° 50' W	UWBM 74641
E	23	CA	Imperial County	32° 48' N	114° 47' W	UWBM 74655
F	24	BCN	Valle de La Trinidad	31° 17' N	115° 37' W	UNM 40911, UNM 40913, UNM 40915
F	25	BCN	Valle de La Trinidad	31° 14' N	115° 37' W	UNM 40923
F	26	BCN	San Felipe	31° 10' N	114° 56' W	UNM 40140
G	27	BCN	Rancho Sta. Catarina	29° 43' N	115° 08' W	UNM 42893
H	28	BCN	Mission de San Borja	28° 42' N	113° 55' W	UNM 42833
H	29	BCN	Punta Prieta	28° 40' N	114° 10' W	CIB 2875
H	30	BCN	Rosarito	28° 40' N	113° 59' W	UNM 42767
I	31	BCS	Santa Agueda	27° 16' N	112° 18' W	USNM 531445

Table 1.1 continued

J	32	BCS	Canipole	26° 30' N	111° 39' W	USNM 531429
J	33	BCS	Comondú Viejo	26° 17' N	111° 48' W	USNM 531430
J	34	BCS	Loreto	25° 58' N	111° 22' W	CIB 6099, CIB 6100
K	35	BCS	Puerto Lopez Mateos	25° 14' N	110° 58' W	CIB 6101
K	36	BCS	Ciudad Constitución	24° 59' N	111° 36' W	CIB 6102, CIB 6103, CIB 6104, CIB 6107, CIB 6108, CIB 6109, CIB 6110
L	37	BCS	La Paz	24° 13' N	110° 31' W	CIB 104, CIB 107
L	38	BCS	Santa Rita	24° 13' N	111° 30' W	CIB 6611
L	39	BCS	La Paz	24° 12' N	110° 34' W	CIB 6125
L	40	BCS	La Paz	24° 12' N	110° 34' W	CIB 6141, CIB 6142, CIB 6143
L	41	BCS	La Paz	24° 04' N	110° 37' W	CIB 6146
L	42	BCS	Los Planes	23° 57' N	109° 56' W	CIB 5501, CIB 5502
M	43	BCS	La Burrera	23° 31' N	110° 04' W	USNM 531434
M	44	BCS	El Pulmo	23° 26' N	109° 34' W	USNM 531432, USNM 531433
M	45	BCS	Punta Lobos	23° 26' N	110° 13' W	USNM 531436
M	46	BCS	Migriño	23° 01' N	110° 04' W	CIB 6148, CIB 6149

Table 1.2 Mean nucleotide diversity (substitutions/site ± SD) of the two *A. leucurus* clades

	Mean nucleotide diversity			
Marker	Northern clade (D_x)	Southern clade (D_y)	Between clades (D)	Corrected divergence (D_{xy})
Cytochrome-*b*	0.004572 (±0.002783)	0.004605 (±0.002834)	0.0337	0.0291
Control Region	0.018399 (±0.009561)	0.018831 (±0.009895)	0.0620	0.0434

Table 1.3 Estimates of population expansion. Estimated parameters of population expansion for (a) exponential and (b) stepwise expansion models

(a) Exponential expansion model*

				Exponential expansion model	
Marker	Clade	No. of individuals	No. of haplotypes	$\Theta = 2\mu N_f$	g
Cytochrome-*b*	Northen clade	44	18	0.0350 (±0.0039)	1142.9061 (±110.0933)
	Southern clade	29	14	0.0230 (±0.0038)	1499.9027 (±188.9652)
Control region	Northen clade	44	38	0.9997 (±0.1595)	55.8318 (±4.9328)
	Southern clade	29	25	2.4907 (±0.4834)	29.4929 (±3.1352)

(b) Stepwise expansion model†

		Stepwise expansion model		
Marker	Clade	$\tau=2ut$	$\theta_0=2uN_0$	$\theta_1=2uN_1$
Cytochrome-*b*	Northen clade	2.561 (1.157, 3.256)	0.000 (0.000, 1.464)	5055.000 (32.070, 8710.000)
	Southern clade	2.716 (1.035, 3.685)	0.000 (0.000, 1.722)	4730.000 (26.367, 9902.500)
Control region	Northen clade	6.042 (2.992,16.236)	4.203 (0.000, 13.782)	76.172 (27.156, 6388.672)
	Southern clade	9.936 (6.916,11.857)	0.000 (0.000, 2.979)	321.250 (85.884, 7988.750)

*Population parameters under the exponential model are given as maximum likelihood estimates (±SD).
†Population parameters under the stepwise model are given as estimates (95% confidence limits).

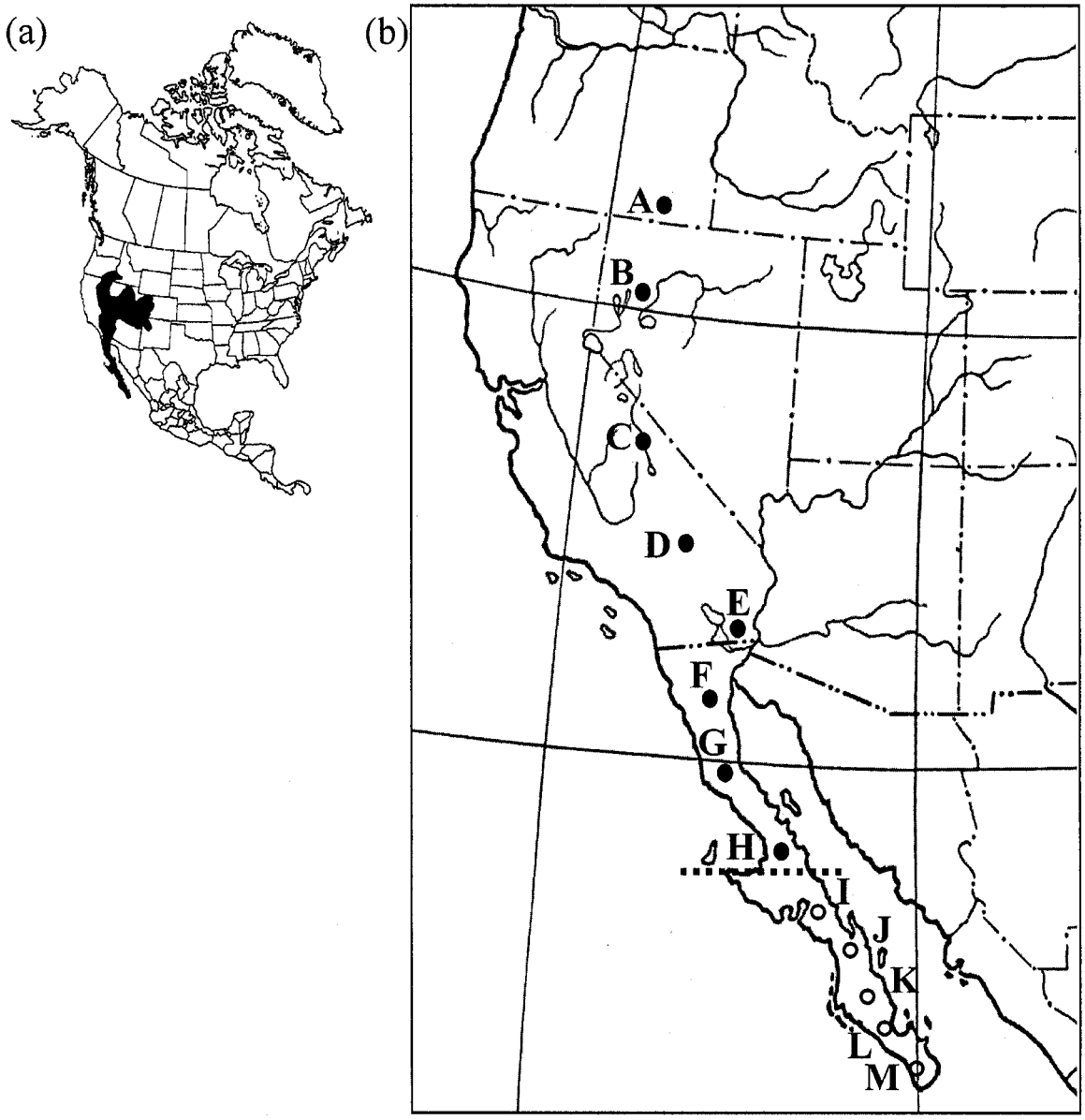


Fig. 1.1 Geographic range of *Ammospermophilus leucurus* and north-south transect of 13 general locality areas. (a) Geographic distribution map modified after Hall (1981). (b) The 13 locality areas (labeled A-M) represent clusters of the 46 specific localities that fall within a range of about one degree latitude (Table 1.1). A geographic break in genetic structure is indicated by the dashed line midway down the Baja California peninsula, separating two mt-DNA clades (black vs. white circles), as demonstrated by data presented in Fig. 2.

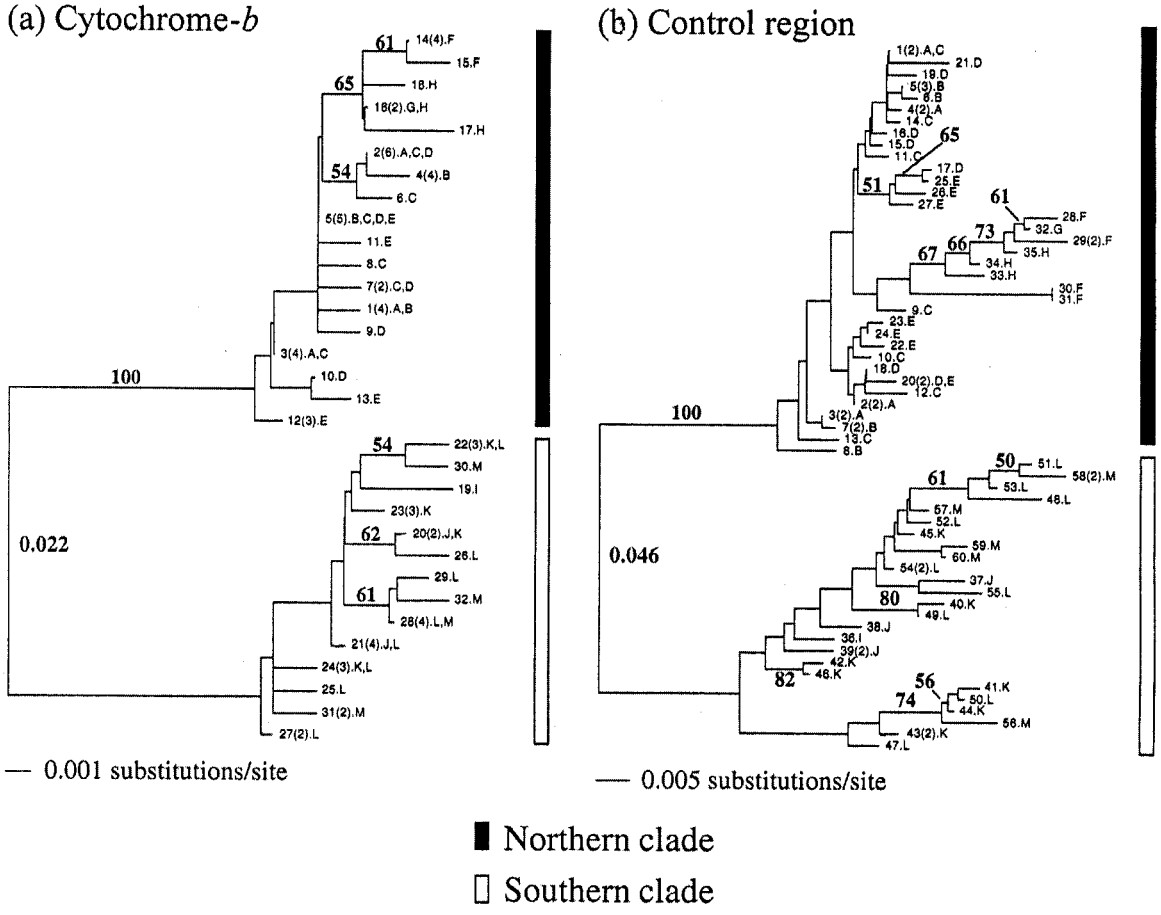


Fig. 1.2 Unrooted neighbor-joining trees for mt-DNA sequences of the cytochrome-*b* gene and control region. (a) Cytochrome-*b* tree based on HKY + Γ model of substitution ($\alpha = 0.0928$, transition/transversion = 5.1938). (b) Control-region tree based on HKY + Γ + I model of substitution ($\alpha = 0.8382$, $p_{inv} = 0.7582$, transition/transversion = 6.0700). Bootstrap support of >50% of 1000 replicates is shown in bold above branches. Branch lengths (substitutions/site) separating the two major clades are shown in bold between the branches. Numbers at each branch tip indicate haplotype number, and numbers in parentheses indicate the number of individuals sharing that haplotype. Letters at branch tips indicate locality areas of each haplotype (Table 1.1).

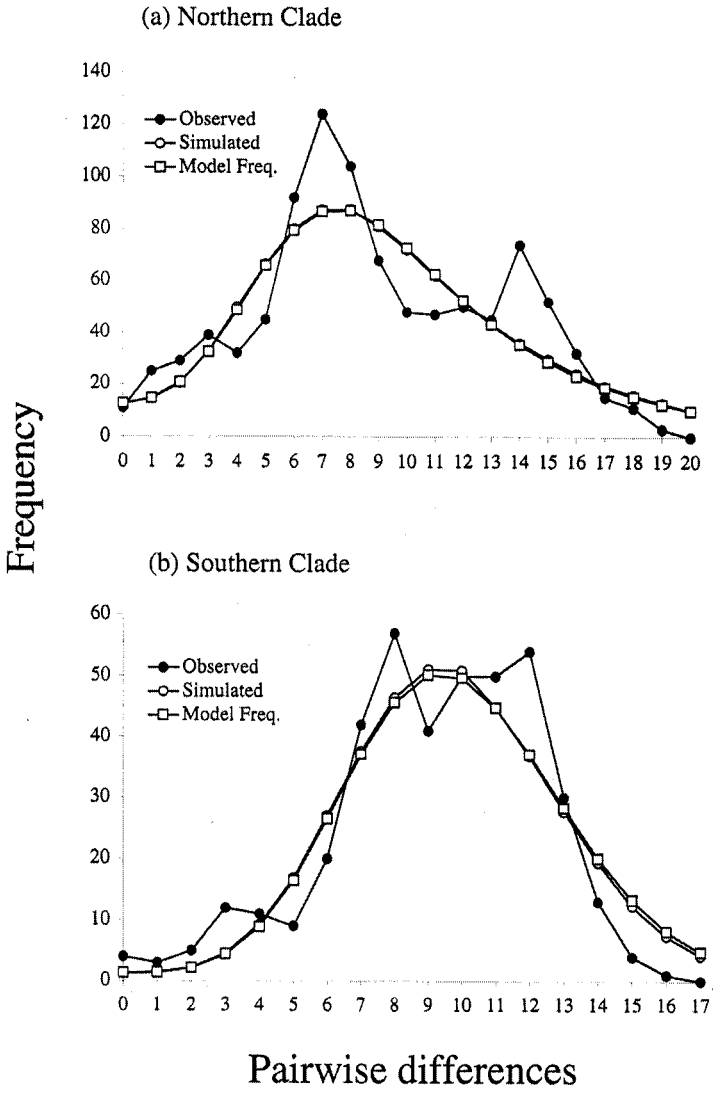


Fig. 1.3 Mismatch distributions (observed, simulated, and stepwise-expansion model) of the two major *Ammospermophilus leucurus* clades based on sequence data for the control region. The curves represent the frequency distribution of pairwise differences for (a) the northern clade, P = 0.76, and (b) the southern clade, P = 0.21. P-values represent the probability that the observed pattern of pairwise differences is different from the simulated pattern based on the estimated model parameters.

Chapter II: Gene flow between two genetically distinct clades of Antelope ground squirrel: Biogeography meets population genetics

Introduction

Population fragmentation has produced within-species divergences and cryptic speciation in multiple taxa across a variety of North American landscapes (Arbogast 1999, Matocq 2002, Whorley *et. al* 2004). The occurrence of geoclimatic change is often cited as the cause of barriers that cause such diversifications (Riddle, 2000). While theoretical perspectives, simulations and laboratory studies demonstrate that population separation can lead to divergence and allopatric speciation (Mayr 1963; Coyne 1992), the effects of historic separation on contemporary gene flow between populations is not well understood. In cases where isolating mechanisms are no longer apparent, we must ask whether populations once isolated by historic barriers remain isolated.

Researchers often use gene genealogies of mitochondrial DNA as proxies for population history and demographic change, and from them infer patterns of connectivity and vicariance (da Silva and Patton 1999, Matocq 2002). When gene trees suggest genetic divergence that is not reflected by phenotype or taxonomy (e.g., Zamudio and Greene 1997), they are sometimes used to reclassify genetically distinct populations as separate species or other taxonomic units. However, when these populations lack current barriers, we can't automatically infer that past separation resulted in speciation. Temporary isolating events, such as changes in sea level (Riddle et al. 2000), habitat shifts due to climate change (Sullivan 2000) or retreat of Pleistocene ice sheets and associated habitat shifts (Arbogast and Kenagy 2001;

Arbogast 1999) may simply leave a transitory genetic imprint of isolation on populations in which gene flow is later resumed. If populations do reconnect prior to speciation, we expect gene flow eventually to erode evidence of past separation (Avise 2000). However if reconnection is recent, within the past 15k years, mtDNA genes alone may not reveal gene flow due to their relatively low mutation rates and lack of recombination. While mtDNA gene trees can reliably identify past population fragmentation (except see Irwin 2002), they cannot address questions concerning current population dynamics. To identify relevant biological species, and to understand how historic isolations shape genetic structure and influence speciation, we must measure current levels of gene flow between historic isolates. To do this we turn to microsatellites, which have the advantage of being diallelic and biparentally inherited. Unlike single-locus, maternally inherited mtDNA, an array of microsatellite loci can be used to identify individual migrants and hybrid individuals to obtain multiple independent measures of gene flow.

The Antelope ground squirrel (*Ammospermophilus leucurus*), which inhabits all four western North American Deserts from the Great Basin in Oregon, south through the Peninsular desert to the tip of Baja California, is an excellent species system for addressing these questions. My initial analysis, based on two mitochondrial markers, of genetic structure over the entire north-south extent of this transect defined a single strong genetic break in the middle of the Baja California peninsula across the Vizcaíno Desert (Whorley *et al.* 2004). This break was identified earlier (Riddle et al. 2000) between two recognized subspecies of *A. leucurus*, and is largely concordant, geographically, with genetic breaks found within white-footed mice (Genus

Peromyscus) two pocket mouse species (*Chaetodipus*), a Kangaroo rat (*Dipodomys*) (all identified in the same study), Le Conte's thrasher and the red-spotted toad (*Bufo punctatus*). These concordant phylogeographic breaks across multiple co-distributed species lead them to hypothesize a common mechanism for population separation. They suggested that an inland seaway had flooded the mid-peninsular lowlands of the region ~ 1.5Mya, effectively bisecting the peninsula and halting gene flow between populations. Whatever the barrier to gene flow was, no obvious barrier now exists and it is not known if contemporary populations of *A. leucurus* just north and south of the suggested historic barrier remain allopatric, or have reconnected and are interbreeding. Additionally, as of 2000, the range boundaries of the recognized subspecies (based largely on morphology) were admittedly poorly defined and deserving of further investigation (Alvarez-Castañeda and Patton 2000). My reconstructed cytochrome-b and control-region gene trees showed little substructure within the two large clades north and south of the break, suggesting that recent population growth and/or migration have produced genetically homogeneous populations to the north and south. A critical remaining question, which I address here, is whether these historically isolated populations have reestablished gene flow across this break.

Materials and Methods:

Sampling—To determine more precisely the geographic location and nature of the previously identified mtDNA break (Whorley et al. 2004) and to analyze contemporary population genetics including potential isolate breaking, I sampled heavily in the Vizcaíno region of the Baja California Peninsula, from 30° - 26° North Latitude. I obtained liver or kidney tissues of 111 specimens from 27 localities within the region. I defined samples collected from within 5 minutes latitude of each other as coming from the same "locality group", which I identified by assigning the name of the nearest town (Table 2.1, Fig. 2.1). Tissue samples are associated with voucher specimens from the following museums: Burke Museum, University of Washington (UWBM); Colección de Mamíferos, Centro de Investigaciones Biológicas (CIB); Museum of Southwestern Biology, University of New Mexico (UNM).

DNA extraction, amplification and sequencing—I extracted and purified genomic DNA from liver or kidney tissues using the Animal Tissues protocol in the DNeasy Tissue Kit from Qiagen (Valencia, CA). I used PCR to amplify 548-554 bp of the mitochondrial control region (CR) for 103 of the sampled individuals using the following primers: CTRL-L (5'-CACYWTYAACWCCCAAAGCT-3'; Bidlack & Cook 2001) and H16498 (5'-CCTGAAGTAGGAACCAGATG-3'; Kocher *et al.* 1993). PCR reactions were run in a total volume of 25μL containing 2.5μL of 1x PCR Buffer, 2.5μL of 2.0 mM $MgCl_2$, 2.0 μL of 10 mM dNTP, 2.5μL of each primer at 10.0 μM, 0.1 μL of 2.5 units/μL *Taq* DNA polymerase (Hoffmann-La Roche, Inc.), and 1μL template DNA. I carried out PCR reactions on an MJ Research Dyad Thermocycler as follows:

94°C for 45 sec; 35 cycles of 94°C for 10 sec, 50°C for 15 sec, and 72°C for 30 sec; and 72°C for 3 min. I purified the resulting PCR products, for use in sequencing reactions, with a QIAquick PCR Purification Kit (QIAGEN).

I carried out sequencing reactions in 10μL volumes with the following reagents: 1.5 μL of purified PCR product, 2.2 μL of Better Buffer (The Gel Company), 0.8 μL each of 1.0μM Forward and Reverse primers, 1.0 μL of ABI Big Dye Terminator Cycle Sequencing Ready Reaction Mix v3.1 (Applied Biosystems) and 4.5 μL of ddH_20. I used a slight modification of the suggested cycle sequencing protocol as follows: Rapid thermal ramp to 96°C and hold for 1 minute; 30 cycles of 94 °C for 10 seconds, 50°C for 10 seconds, 60°C for 30 seconds; a final extension at 60° C for 3 minutes; 1 hold at 4°C for 5 minutes.

I cleaned all samples to remove excess dye terminators using an ethanol precipitation protocol copied from the Swanson lab (Department of Genome Sciences, University of Washington):

1. Add 26 uL of 95% ethanol to each well
2. Mix, and incubate at -20 °C for 15 minutes.
3. Centrifuge for 25 minutes at 3000 x G
4. Centrifuge samples upside down for 1 minute at 350 x G to remove supernatant.
5. Add 125 uL of 70% ethanol to each well.
6. Centrifuge 10 minutes (right side up) at 3000 x G
7. Remove supernatants by centrifuging upside down for 1 minute.

I subsequently sequenced all samples on an ABI/PRISM 3100 Genetic Analyzer (Applied Biosystems) in the University of Washington Biology Department's Comparative Genomics Center. I aligned sequences for phylogenetic analysis using SEQUENCHER 4.1 (Gene Codes Corp., Ann Arbor, MI). All unique haplotype sequences will be deposited in Genbank.

Phylogenetic Analyses—To identify the model of DNA sequence evolution that best explained my data I first used ModelGenerator v0.61 (Keane et al. 2006) to make a maximum likelihood (ML) estimate of the data given each of 56 possible models of evolution. Exploratory analyses suggested some heterogeneity in rate of substitution among sites, so I designated five substitution rate categories, and allowed ModelGenerator to estimate those. Both the Akaike Information Criteria (AIC) and hierarchical Likelihood Ratio Test (hLRT), as implemented in ModelGenerator, used the likelihood estimates of all 56 models of evolution to determine that the HKY model of evolution (Hasegawa *et al.* 1985) including rate heterogeneity across sites and a fraction of invariable sites (HKY+G+I) was the simplest model that adequately fit the data. Their estimates of sequence evolution parameters were identical, with transition/transversion rate ratio (κ) =10.5, Gamma distribution parameter (α) = 0.85, and proportion of invariable sites (I) =0.76. I used these estimates as priors in all subsequent phylogeny reconstructions.

I used both Bayesian and ML methods to reconstruct phylogenetic relationships among my samples. MrBayes v3.1.2 uses Bayesian inference to estimate a phylogeny of highest posterior probability from a sample of DNA sequences. It uses the Markov Chain Monte Carlo (MCMC) method to search through possible tree topologies (with

the option of including informative or non-informative prior estimates of parameters of sequence evolution) and concentrates it's search on tree topologies of high posterior probability. It uses this accumulated collection of tree topologies and parameter estimates to approximate the phylogeny and model of sequence evolution. I ran 2 x 10^6 generations, discarded 1/4 of these as 'burnin' (which allows the markov chain to reach relative stationarity and does not influence the final estimate of the phylogeny), and sampled a topology from the chain every 100 generations thereafter. The final estimate of the phylogeny was made using the final 10,000 trees from this population of sampled trees. I performed several runs, and several markov chains per run, using both informative and non-informative priors to ensure that all runs converged to a similar answer. It has been shown that using a uniform, non-informative prior on branch length can produce incorrect topologies and inflated posterior probabilities of clade support (Yang and Rannala 2005), so I also used two priors on branch length: both a uniform and an exponential with a small mean. I accepted this consistent behavior as evidence that individual runs were not overly influenced by my choice of priors and that results were robust. Phylogenies by Maximum Likelihood (PHYML) uses ML to estimate the phylogeny from a sample of DNA sequences. In its search through tree space, it simultaneously optimizes topologies and branch lengths to arrive at a final estimate of the phylogeny. I made two estimates of the phylogeny using PHYML, assuming five substitution rate categories as above. In one ML search I used the parameter estimates produced by ModelGenerator as fixed parameters. To check that the data were informative and produced consistent results across runs, I ran a second iteration of PHYML in which I allowed it to estimate and optimize κ, α, and I. To assess nodal

support I ran 200 bootstrap replicates on the topologies produced by both the fixed parameter and estimated parameter runs.

Molecular Diversity—For comparison with my original data set, consisting of samples from across the entire 22° Latitude range of the species, I used ARLEQUIN 2.001 to calculate nucleotide diversity (pi), average number of pairwise differences, (theta pi), and estimates of sequence divergence among the samples of my geographically limited, 4° Latitude sampling area of the Vizcaíno region.

Microsatellite Screening—Primers flanking microsatellite loci were chosen from published primer notes and screened for variability and linkage disequilibrium. I screened 43 loci (whose primer sequences I obtained from the literature) that were developed for the following species: 21 from *Sciurus vulgaris* (Hale *et al.* 2001), 9 from *Spermophilus columbianus* (Stevens *et al.* 1997), and 13 from *Spermophilus brunneus* (May *et al.* 1997). I ordered all primers from ABI.

I used a standard amplification protocol under non-stringent PCR conditions in 10uL reaction volumes to check for the presence of variable microsatellites:

1.00 μL of 10x PCR Buffer without $MgCl_2$

0.60 μL of 25mM $MgCl_2$

0.20 μL of 10mM dNTP mix

0.20 μL of 10μM Forward Primer

0.20 μL of 10μM Reverse Primer

0.15 μL of 2.5 units/μL of Jumpstart *Taq* polymerase

6.70 μL of ddH_2O

I then optimized each primer pair by running them at three $MgCl_2$ volumes (0.4, 0.6, and 0.8μL) and four annealing temperatures, separated by ~3 C° and centered on the published optimal annealing temperature. I ran each primer pair under the following thermocycling conditions: One cycle for 2 min at 96° C to denature double stranded DNA, followed by 33 cycles of 30seconds at 94°C, 30 seconds at the annealing temperature, 15 seconds at 72°C to allow *Taq* polymerase to add dNTP's, followed by a final extension step of 4 minutes at 72°C to ensure that all fragments received a terminal Adenine cap.

From this pool I obtained seven loci that appeared to be variable in both populations (Table 2.2). I re-ordered these primers, this time with one of three dyes (HEX, 6-FAM or NED) attached to the 5' end of the forward primer of each pair. I used these labeled primer pairs in another round of amplifications (using thermocyling conditions that were optimal for each locus) to fluorescently tag the alleles at each locus for all individuals. I then diluted the amplified products in a 1:10 dilution with ddH_20, and added 1μL of diluted product to a genotyping cocktail consisting of 13μL formamide, to stabilize products, and 0.2μL GeneScan^{TM} 400HD [ROX] size standard. This final mixture was genotyped using Foundation DataCollection v2.0 on the ABI/PRISM 3100 Genetic Analyzer (Applied Biosystems).

I used GeneMapper v3.7 software to identify microsatellite alleles at each locus and to automatically call alleles. All alleles flagged as suspect were visually inspected and either accepted as accurate calls or rejected. In the latter case, I re-amplified the locus and called the alleles again. If a locus was absent for an individual I coded it as missing data and ignored it in subsequent analysis.

Because my subsequent analyses of genetic structure rely on the assumptions of neutrality and panmixia, I subjected each of these loci to a test of allelic differentiation, linkage disequilibrium and Hardy-Weinberg equilibrium (H-W), as implemented in GenePop (Guindon and Gascuel 2003, Guindon *et al.* 2005). Loci that failed the H-W test were further tested for homozygote excess by the method of Rousset and Raymond (1995) (also implemented in GenePop). An exces of homozygosity across all loci may indicate inbreeding or population sub-structuring, but excess homozygosity in only one or a couple loci probably does not, and may actually give false indications of sub-structure where none exists (Table 2.2).

I tested all loci, both as single populations and as two putative populations, for Hardy-Weinburg equilibrium and linkage disequilibrium, and tested the populations for.

To identify the number of populations in my sample and to characterize population structure I used the program Structure 2.1 (Falush et al. 2003), which is a model based clustering method that assigns multi-locus genotypes to "natural" groupings. I allowed Structure to assign individuals to populations in the microsatellite analysis, regardless of which mtDNA haplotype they carried. The phylogenetic analysis identified two distinct genetic groups with highly divergent haplotypes, which, in conjunction with geographic separation, would suggest population subdivision. However, as many of these individuals were collected in the same or a neighboring locality, with no obvious intervening barrier, it seemed unreasonable to assume that neighbors possessing divergent haplotypes belonged to two reproductively isolated populations. A more reasonable if less simplistic model is to assume that gene flow is occurring between reproductively competent members of historically isolated

populations, and that nuclear loci will reflect this recent mixing. Structure 2.1 provides an admixture model that incorporates this uncertainty. In calculating probable population of origin, it allows for the possibility that each individual inherited some fraction of its genome from the other population in the sample. It also allows for correlation between alleles across loci, which is expected if populations have been isolated long enough for mutation and genetic drift to cause allele distributions to diverge between them. The clear advantage of this model is that it will detect admixture between populations if it exists. Because it accounts for population admixture as well as some correlation among allele frequencies, which is often encountered in real data (Pritchard et al. 2000), I used the admixture model when inferring number of subpopulations.

Structure produces likelihood estimates of the data given the model, allowing the user to infer the number of populations (K) most consistent with the data. The program also produces likelihood estimates of various parameter values associated with population substructure and mixing. I examined the values of Q, the admixture proportion of each individual, Fst and α, the admixture parameter to assess convergence. Large values of α ($>>1$) models individuals as inheriting allele copies from all K populations in equal proportions, while small values ($<<1$) model individuals as inheriting allele copies from primarily one of the K populations.

I performed a series of independent runs of Structure for each value of K between 1 and 3 to assess the performance of the program. I followed the methods of Pritchard et al. (2000) to choose the burn-in period and to assess mixing, and attempted

runs of between 10^6 and 3 x 10^6 iterations or more following a burn-in period of at least 10^6 iterations.

Results:

Model of evolution—Both the hLRT and AIC, as implemented in ModelGenerator, recommended the HKY+I+G model of substitution, with the following parameter values: κ=10.5, α=0.85, I=0.76. I used these parameter estimates as informative priors in the MrBayes analysis, and as fixed values for a PHYML run with fixed parameters.

Phylogenetic structure—Both Bayesian and ML methods of phylogeny reconstruction produced similar topologies with similar branch lengths separating taxa. I present here only the tree produced by PHYML, but include both Bootstrap values and posterior probabilities as measures of nodal support (Fig. 2.1). As in the previous analysis (Whorley et al. 2004), the topology shows only one substantial, well-supported genetic break between two monophyletic clades, which I name southern and northern, near the center of the Baja California peninsula (28° N. Lat.). Though the Vizcaíno region immediately surrounding the previously identified genetic break was weakly sampled before, the intensive sampling and analysis of the present study revealed no new haplotypes of intermediate genetic difference. Within each monophyletic clade internal node support is low; branch lengths separating individual or groups of individuals are short; and numerous unresolved polytomies exist within each clade, producing a star-like phylogeny.

I detected little sub-structure within each clade, which is not unexpected given that samples were obtained from such a geographically small sampling area. In light of this, however, it is surprising that nearly as much haplotype diversity exists within this tiny geographic sample of each clade as exists across the entire geographic range of the

species. The mean nucleotide diversity (D) of the Northern and Southern clades recovered by this analysis accounted for 85 and 88%, respectively, of the nucleotide diversity recovered in the previous analysis of the entire western range of *A. leucurus* (Table 2.3).

All Bayesian analyses, regardless of whether I used informative or non-informative priors, eventually produced topologies and branch lengths similar to each other and to the ML topologies. I used both a uniform prior and an exponential prior with a low mean as suggested by Yang and Rannala (2005).

Both PHYML runs returned the same topology and similar nodal support, suggesting that the parameter estimates generated by ModelGenerator are robust.

Haplotype Distribution and Diversity—The geographic location of the genetic break recovered in the previous analysis is pinpointed here. Haplotypes from both the northern and southern clade appear throughout the Central peninsula (Vizcaíno Desert) region, from 26 – 30° N. Lat., but northern haplotypes are greatly overrepresented. Of the 107 sampled individuals, 75 represent unique haplotypes. More than 83% of these (62 of 75) are northern haplotypes and only 17% (13 of 75) are southern. Similarly, at the four localities where haplotypes from both clades appear, northern haplotypes outnumber southern haplotypes by at least 5 to 1 (Fig 2.2). The Northern clade extends about 2° Latitude further south than was previously detected by the weaker geographic resolution of our previous analysis (Whorley et al. 2004). It covers much of the salty floodplain south of Guerrero Negro and West of the city of Vizcaíno and Hwy 1.Northern and southern clade haplotypes appear to interdigitate along an arc stretching from the southwestern margin of the Vizcaîno desert NE through it, ending on the east

coast in the region of Volcán las tres Vírgenes. Along this border, haplotypes from both clades appear together at the same localities or at adjacent localities (Fig. 2.2).

Population structure: Choice of K—The seven variable microsatellite loci consisted of four di-nucleotide repeats and two, tri-nucleotide repeats (Table 2.2) and provided adequate data to infer population structure. Two loci failed the test of HWE and of homozygote excess. This is not too disturbing because Structure does not appear to have been overly influenced by them. Had it detected significant population structure, then I would perhaps worry that the excess homozygosity was leading to false, or inflated indications of sub-structure. However, as no significant structure was detected (see below), I determined that the distribution of homozygote genotypes was not biasing my results.

I present results based on runs of 3×10^6 iterations or more following a burn-in period of at least 10^6 iterations. I followed the methods of Pritchard et al. (2000) to choose the burn-in period and to assess mixing. The final dataset consists of a tabled summary of 7 – 10 iterations at each K to assess consistency (Table2.4). The probability of the data is highest for K=1 (i.e. the -Ln likelihood is lowest under the assumption of a single population) and point estimates of the likelihood remain stable after an initial burnin of only 200,000 generations in the case of K=1. Point estimates of the likelihood of the data were remarkably similar across runs at K=1. They varied more substantially across runs at K=2 and at K=3, but the variability between runs at different K routinely exceeded the variability encountered among run at the same value of K. This analysis suggests that only one interbreeding population exists across the Vizcaíno region.

For all runs at K>1, values of Fst for each population settled down, but never appeared to converge. Instead, the Fst value for one population shrank to 0.000 while the Fst for other populations assumed some measurable value. For example, at K=2, for 5,000 generations after the burnin the Fst of population 1 remains at 0.000, and population 2 fluctuates, then from generation blah to blah, Fst values switch completely, and population 1 has measurable Fst while population 2 shrinks to zero.

Values of Q (estimates of the ancestry of each individual's genome) for all individuals are evenly divided between source populations, regardless of the value of K that I assume. To illustrate this I include results for a randomly chosen run at K=2 and K=3 for five randomly chosen individuals (Table 2.5). The inference of ancestry at K = 2 is that all individuals inherited ~50% of their genes from one population and ~50% from the other. Similarly, at K = 3, all individuals inherit ~33% of their genome from population 1, ~33% from population 2 and ~33% from population 3.

Discussion:

In my previous analysis of the phylogeography of *A. leucurus*, which included a range-wide sample of individuals covering 20° Latitude, all phylogeny estimates contained a substantial genetic break near the middle of the Baja peninsula and little substructure within the two major clades (Whorley et al. 2004) This analysis extends that work and increases substantially the sample size in the Vizcaíno region, near the geographic location of the genetic break. The increase in sample number and density allowed us to more explicitly quantify the genetic break, to identify areas of haplotype overlap, and to assess population subdivision and gene flow across it.

The phylogenetic results of the previous analysis are reinforced here, with all sampled individuals belonging strongly to one of two monophyletic clades (Fig. 2.1). The lack of geographic structure within each clade across this short Latitudinal band mirrors the lack of structure found previously across the entire range of *A. leucurus*. Similarly, the preponderance of short branch lengths connecting individuals within the two major clades combined with universally low support for internal substructure suggests two things: individuals within each clade share relatively recent common ancestry, and the demographic radiation from that common ancestor was explosive. This is not surprising given the small geographic area sampled, but stands in contrast to measures of haplotype diversity.

Estimates of nucleotide diversity for each clade account for a large percentage of the total diversity found within the species (Table 2.3). Though our phylogeny shows little intra-clade divergence and no geographic structure, most of the species-wide genetic diversity (for both clades) occurs within the narrow 4° Latitude that

encompasses the Vizcaíno region. In light of our previous analysis (Whorley et al. 2004) it makes sense that much of the northern clade's diversity remains present just North of the genetic and geographic break. These new measures of nucleotide diversity support our earlier conclusion that the entire northern group of *A. leucurus* arose recently and explosively from a small founder population isolated just North of 28° Latitude.

The current distribution of mtDNA haplotypes across the Vizcaíno region is complex, with haplotypes from both the northern and southern clade co-occurring at multiple sites. Whatever geographic barrier or dispersal-limiting mechanism was responsible for producing the genetic discontinuity near the mid-peninsula appears to now be absent, as we see haplotypes from both clades spread across most of the intensively sampled region. However, at the four sites where they occur in sympatry, southern clade haplotypes are less numerous than northern clade haplotypes. Because we recovered this same pattern at multiple sampling localities, it is probably not due to chance. This suggests that dispersal across the region between the two clades is biased toward northern clade dispersal southward. The mechanism responsible for this asymmetric gene flow cross the Vizcaino region is unknown, but may be due to asymmetric immigration of northern clade females into the southern clade geographic area.

The pattern of haplotype distribution in the mid-peninsula suggests two models of genetic architecture that the microsatellite data address: Either the mtDNA genealogy identifies two reproductively isolated, monophyletic clades that are syntopic but non-

interbreeding, or the gene genealogy is not an accurate indicator of reproductively isolated populations and the clades have re-connected and resumed gene flow.

The microsatellite data suggest strongly that a single panmictic population exists in the Vizcaíno region of the peninsula. Based on consistent –Ln likelihood estimates and variance in likelihood scores across repeated iterations of Structure, the assumption of one population (K = 1) always yields the lowest likelihood scores. The much larger variation observed between batches of runs at K=2 and at K=3 appears largely due to Structure's inability to consistently assign individuals to one of the two (or three) assumed populations. In other words, individuals are so admixed that functionally, only a single population exists.

Examining values of Q for various individuals at different assumed K also provides insight into the degree of population subdivision. The results from a few individuals (Table 2.5) demonstrate the problem of non-identifiable population of origin: essentially, the allele distributions of all loci, across the entire sample, are undifferentiated and allele frequencies are uncorrelated. When I assume more than one population, every individual becomes as likely to originate from one population as it is from another.

Though populations of *A. leucurus* clearly diverged sometime in the past, either due to habitat shifts or some major geographic barrier (Riddle et al. 2000, Alvarez-Castañeda, pers. comm.), they appear to have reconnected and resumed gene flow. Animals with highly divergent mtDNA haplotypes do not possess similarly divergent microsatellite allele profiles, as would be expected if historic isolation continued to influence gene flow among populations. The contrasting signal of divergence given by

these two genetic markers suggests that recent interest in identifying different species by divergence in mitochondrial genes alone (Bradley and Baker 2001) should be approached with caution. In this case, we would mistakenly call the northern and southern clade distinct species though they are currently interbreeding. I did not recover a stepped or clinal pattern of gene frequency change across our mid-peninsular sample; indeed, I encountered no geographic structure at all. However, this does not mean that it does not exist, perhaps further north and south of the intensively sampled area. Further research into the genetics of the region will demonstrate whether contemporary populations of *A. leucurus* are a reformed single species or a hybrid zone between differently adapted forms (eg Second, *et al.* 2006).

The contrasting signal of differentiation also leads to the more general question of whether populations of peninsular organisms that share the phylogeographic break with *A. leucurus,* also share it's recent population history of reconnection and interbreeding. Quantification of the contemporary population genetics of these syntopically-distributed organisms will provide insight into mechanisms of speciation and the importance of geoclimatically driven barriers in driving speciation across multiple taxa.

Table 2.1 Specific localities (n=43), and specimens (n=107) of *Ammospermophilus leucurus* in the present study

State	Locality	Latitude	Longitude	Specimen number
BCS	Loreto	25° 58' N	111° 22' W	CIB 6099, CIB 6100
BCS	La Purísima	26° 11' N	112° 07' W	CIB 9305, CIB 9306, CIB 9307
BCS	Comondu Viejo	*26° 17' N*	*111° 48' W*	USNM 531430
BCS	Canipole	*26° 22' N*	*111° 44' W*	USNM 531429†
BCS	Ejido Cadejé	26° 22' N	112° 31' W	CIB 7715, CIB 7716
BCS	Punta Abreojos	26° 46' N	113° 38' W	CIB 9302, CIB 9303, CIB 9304
BCS	San Juan de las Pilas	26° 52' N	112° 35' W	CIB 7712, CIB 7714
BCS	Punta Abreojos (NW)	26° 54' N	113° 46' W	CIB 6610
BCS	Punta Abreojos (NW)	26° 58' N	113° 45' W	CIB 9283, CIB 9284
BCS	Punta Abreojos (N)	26° 59' N	113° 28' W	CIB 8518, CIB 8519, CIB 8520, CIB 8521, CIB 8522*
BCS	Pta. Abreojos (NE)	27° 04' N	113° 24' W	CIB 9291
BCS	Bahía Asunción (E)	27° 10' N	114° 09' W	CIB 9292, CIB 9296
BCS	El Rodeo	27° 12' N	113° 13' W	CIB 8961, CIB 8962, CIB 8963, CIB 8964, CIB 8965
BCS	Isla San Marcos	27° 15' N	112° 05' W	CIB 8527
BCS	Santa Agueda	*27° 16' N*	*112° 18' W*	USNM 531445†
BCS	San Ignacio	27° 18' N	112° 54' W	CIB 8506
BCS	San Ignacio (E)	*27° 21' N*	*113° 09' W*	LVT-01716, LVT-01717
BCS	San Ignacio (E)	*27° 22' N*	*113° 10' W*	LVT-01718
BCS	Bahía Asunción (NW)	27° 22' N	114° 28' W	CIB 9288, CIB 9289, CIB 9290
BCS	Santa Rosalía	27° 23' N	112° 22' W	CIB 9849
BCS	Bahía Asunción (NW)	27° 23' N	114° 28' W	CIB 9285, CIB 9286
BCS	Volcan de las Tres Virgenes	27° 30' N	112° 32' W	CIB 8497, CIB 8498, CIB 8499, CIB 8500, CIB 8501, CIB 8502, CIB 8503, CIB 8504, CIB 8505
BCS	San José de Castro	27° 32' N	114° 28' W	CIB 10127, CIB 10128
BCS	San Francisco de la Sierra (SE)	27° 33' N	112° 58' W	CIB 8496
BCS	San Francisco de la Sierra (SE)	27° 34' N	113° 05' W	CIB 8507, CIB 8508, CIB 8509, CIB 8510, CIB 8511, CIB 8512, CIB 8513, CIB 8514, CIB 8515
BCS	San Francisco de la Sierra (SE)	27° 35' N	113° 05' W	CIB 2876
BCS	San Francisco de la Sierra (N)	27° 40' N	113° 01' W	CIB 8489, CIB 8490, CIB 8491, CIB 8492*, CIB 8493, CIB 8495
BCS	Guerrero Negro	27° 59' N	113° 54' W	CIB 9280, CIB 9281
BC	El Arco	28° 01' N	113° 22' W	CIB 8959, CIB 8960
BC	Calmalli	28° 06' N	113° 26' W	CIB 11598, CIB 11599, CIB 11600, CIB 11601*, CIB 11602*

Table 2.1 continued

BC	El Arco (NW)	28° 13' N	113° 31' W	CIB 8958, CIB 11592, CIB 11593, CIB 11596
BC	Cerro Santo Domingo	28° 15' N	114° 05' W	CIB 11587, CIB 11588
BC	El Barril	28° 18' N	112° 52' W	CIB 11578, CIB 11579, CIB 11580, CIB 11583, CIB 11584, CIB 11585*
BC	Guerrero Negro (N)	28° 19' N	113° 49' W	CIB 10120, CIB 10121, CIB 10122
BC	El Barril	28° 20' N	112° 54' W	CIB 6605
BC	San Ignacito	28° 44' N	113° 50' W	CIB 10118
BC	Punta Prieta	*28° 55' N*	*114° 10' W*	LVT-01640 / 1991, LVT-01641, LVT-01642, LVT-01643, LVT-01644, LVT-01645, LVT-01646, LVT-01647, LVT-01648, LVT-01649, LVT-01650, LVT-01651, LVT-01652, LVT-01653
BC	Punta Prieta	28° 56' N	114° 10' W	CIB 8957
BC	San Felipe	30° 56' N	115° 07' W	CIB 10115*
BC	San Felipe	31° 01' N	115° 12' W	CIB 10116
BC	San Felipe	31° 04 'N	115° 00' W	CIB 10117
BC	San Felipe	31° 05' N	114° 58' W	CIB 10108, CIB 10113

Table 2.2 Microsatellite loci used in the analysis of mid-peninsular samples

Locus	Repeat	Primer Sequence (5'-3')	T_a (°C) working	Product size (bp)	# Alleles	HWE
Scv1	$[CA]_{28}$	F:CTCCTCTTCCAAGGGTGACA R:GATGGCCTCTGTTTCTCTGC	48	137-159	8	0.32
Scv31	$[AG]_{29}$	F:CCAAGTTCCAGACCAACCTC R:TCGGGTCTCTAAGGAGATGG	48	197-213	3	0.69
IGS-1	$[CA]_{20}$	F:ATAACAGCACCCTGCTCCAC R:AATCCATCCTCTACCTGTAATGC	57	86-113	16	*0.00
IGS-24d	$[GTT]_8$	F:CCCTCAGATTAAATTGAATTGG R:GCCCTGCATGAAACCTTG	55-52	145-164	7	0.87
IGS-110b	$[TGC]_9$	F:CCATGGAAGCATGTCTGGTG R:TGCTTCCTGATTTCAAAGTTGC	55-52	116-122	3	0.27
GS14	$[TG]_{30}$	F:CAGGTGGGTCCATAGTGTTAC R:TTGTGCCTCAGCACTTCTTTC	48	257-279	12	0.06
GS17	$[TG]_{16}$	F:CAATTCGTGGTGGTTATATC	59	146-176	17	*0.00

Table 2.3 Comparison of molecular diversity.
Comparison of intra-population molecular diversity (substitutions/site ± SD) between a range-wide sample of *A. leucurus* (spanning 20° Latitude) and the mid-peninsular dataset (spanning 4° Latitude) using Control region sequence data.

Sampling Range	Northern clade (D)	Southern clade (D)
20° Latitude	0.018399 (± 0.009561)	0.018831 (± 0.009895)
4° Latitude (this study)	0.015654 (± 0.008164)	0.016577 (± 0.009222)
% Diversity contained in Vizcaíno region	85.1	88.0

Table 2.4 Inferring value of K, the number of populations, for *A. leucurus*.
Inferred number of populations assuming a uniform prior on K ($K \in \{1,2,3\}$), from the sample of 117 individuals across the Vizcaíno region of the Baja California peninsula. The third column estimates the probability of the model, given the data, and suggests the correct number of populations

K	Ln P(X|K)	P(K|X)
1	2310.1	~1.0
2	2413.6	~0.0
3	2504.5	~0.0

Table 2.5 Estimates of Q (the proportion of an individual's genes inherited from each assumed population) for 5 randomly chosen individuals (identified by Specimen #) for one randomly chosen run of K=2 and at K=3.

Assumed # of populations	Specimen #	Q values Population 1	Population 2	Population 3
K=2	LVT-01640/1991	0.456	0.544	NA
	CIB 8521	0.474	0.526	NA
	LVT-01717	0.449	0.551	NA
	CIB 8513	0.479	0.521	NA
	CIB 8503	0.568	0.432	NA
K=3	LVT-01640/1991	0.321	0.372	0.307
	CIB 8521	0.350	0.332	0.318
	LVT-01717	0.347	0.329	0.324
	CIB 8513	0.341	0.332	0.326
	CIB 8503	0.328	0.314	0.358

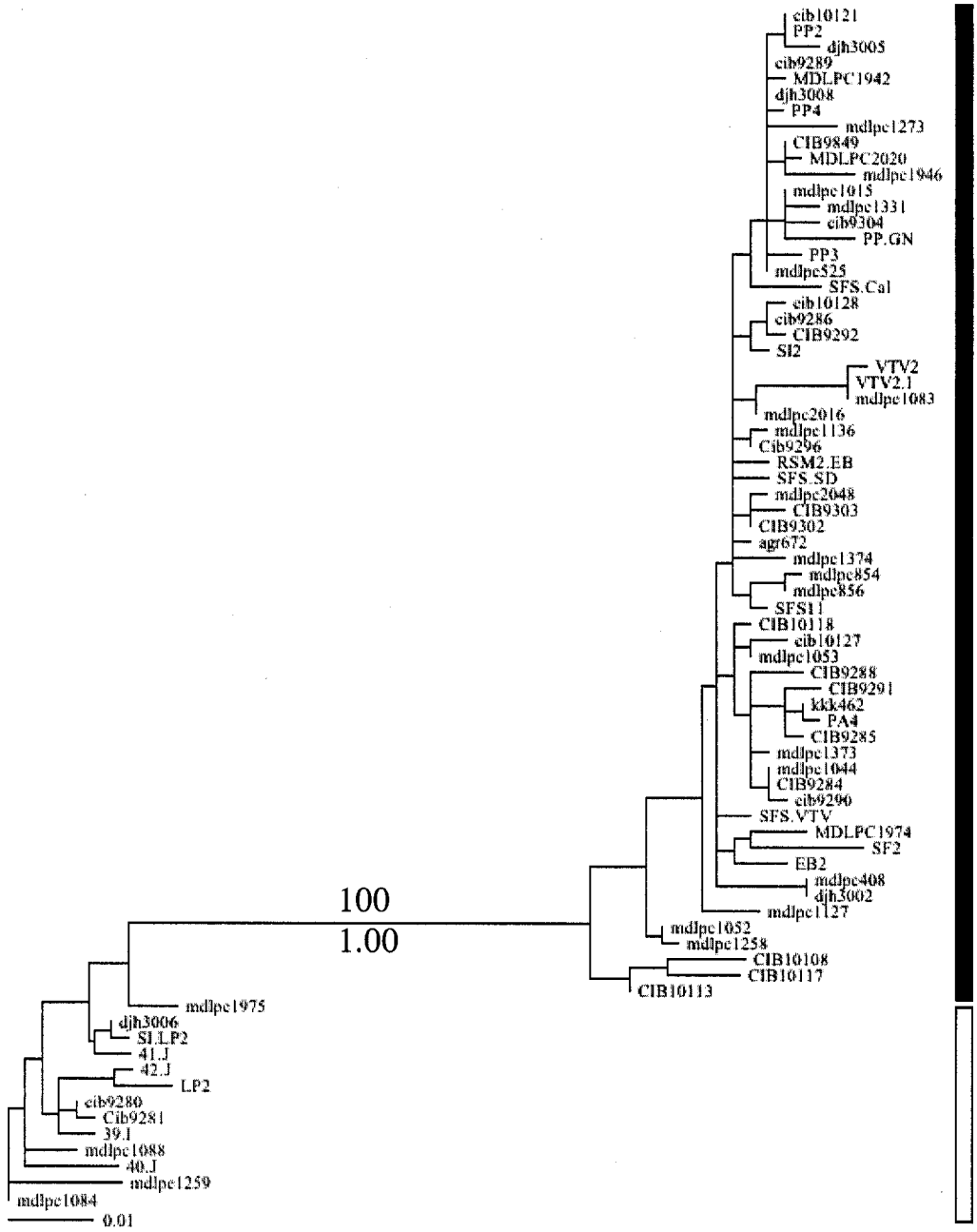


Fig. 2.1 Phylogeny of mid-peninsular samples.
Maximum likelihood tree for mt-DNA sequence of the control region based on HKY + Γ + I model of substitution ($\alpha = 0.85$, $p_{inv} = 0.76$, transition/transversion = 10.5). Bootstrap support of >70% of 1000 replicates is shown above branches; Posterior probabilities of greater than 0.70 are shown below branches.

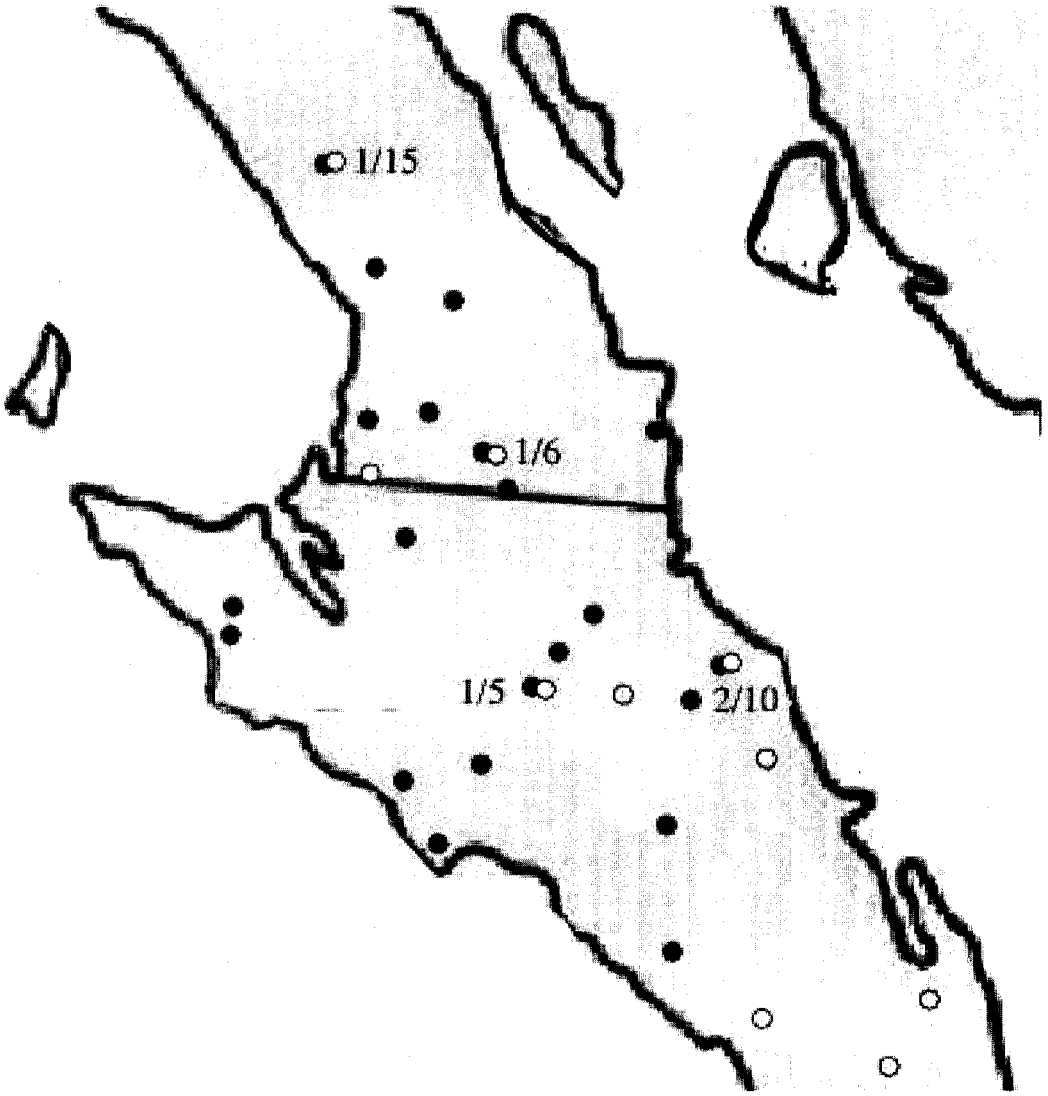


Fig. 2.2 Distribution of haplotypes in the Vizcaíno region of the Baja California peninsula. Black dots mark localities where Northern clade haplotypes were found and white dots mark localities where Southern clade haplotypes were found. Localities where both haplotypes were found are marked with two dots and an associated fraction: the number of southern clade haplotypes (numerator) over the total sample from that locality (denominator).

Chapter III: Variation in reproductive patterns of Antelope ground squirrels, *Ammospermophilus leucurus,* from Oregon to Baja California

Introduction

Biological variation over geographic space and time is produced, both proximately and ultimately, by patterns of environmental variation that include spatial clines and temporal cycles. These kinds of physical environmental variation can shape the expression of variable biological traits that determine the life history of populations over the geographic range of a species. Because of the tight coupling of reproductive traits to fitness and the success of local populations, comparisons of the timing and incidence of reproduction over latitudinal geographic space have been of considerable interest to both ecologists and evolutionary biologists. In fact Darwin (1859) observed that phenotypic variation often coincides with geographic distance, and Mayr (1956) suggested that ecogeographic variation has an adaptive basis and may represent a response to clinal variation in environmental conditions. In terms of contemporary evolutionary ecology, the study of reproduction over a gradient of latitude represents an evaluation of the interplay between flexibility of local phenotypes and the evolution of genotypic adaptation across the range of a species.

The relationships among latitude, length of breeding season, litter (or clutch) size, and lifetime reproductive success, have received wide attention in the literature. Lord (1960) called attention to interspecific patterns of increasing litter size with increasing latitude in North American mammals; he explained this as a response to higher winter mortality at higher latitude. Spencer and Steinhoff (1968) provided an alternative explanation: that the shorter breeding seasons at higher latitudes,

accompanied by reduction in potential longevity, promote natural selection for larger litter size. This follows from Williams' (1966) refinement of Lack's (1954) principle (based on observation of larger clutch size in birds at higher latitude) that the tradeoff between energy allocation to present, vs. future reproductive value leads to reproductive investments that maximize lifetime reproduction. An additional consideration in analyzing reproductive output is the roles of predictability, constancy, and contingency of environmental conditions and resources that influence the timing of breeding seasons (Colwell 1974). We are interested here in the special problem of tradeoffs between adaptation and phenotypic flexibility that come into conflict across large amounts of geographic space, over which gene flow may be reduced or even halted at some point. Relative to geographically distant and disjunct populations, we expect that neighboring populations will show greater similarity in expression of reproductive traits, as well as similarities of genetic structure.

Seasonal breeding results from a process of matching life-history traits and physiological mechanisms to prevailing environmental characteristics. In desert ecosystems, where resources required for reproduction are both sparse and ephemeral, this matching process involves strategies that promote survival under extreme and unpredictable conditions. The threshold for availability of sufficient energy, nutrients, and water to allow breeding is reached only during a restricted season. The relationship between a seasonal pulse of precipitation, the resulting germination of annual plants and proliferation of perennials, and the breeding of desert rodents has been well documented in North American deserts (Beatley 1969; Reichman and van de Graaff 1975; Reynolds and Turkowski 1972). In the Great Basin and Mojave Deserts the precipitation pulse

occurs in winter, associated with Pacific storms, and in the Sonoran Desert it arrives in summer, with Mexican Gulf storms. Within any single desert system, phylogenetic diversity and associated life-history variation of desert rodent species is also associated with differences in the timing and incidence of breeding (Kenagy and Bartholomew 1985). Like other small vertebrate animals, desert rodents show early stages of physiological preparation for reproduction in response to predictive cues such as the seasonal change in day length, or even an associated endogenous rhythm, but the final activation of reproductive function clearly results from a direct response to a favorable supply of food and water (Wingfield and Kenagy 1991).

The enormous north-south extent (2500 km and 21 degrees of latitude) of the geographic range of antelope ground squirrels, *Ammospermophilus leucurus*, presents an opportunity to examine a family of hypotheses and questions that concern proximate local flexibility against a background of genetic relatedness (Whorley et al. 2004). We explore here the application of these questions to patterns of reproduction, first by characterizing reproduction in a population at the northern extreme of the range, in the Alvord Desert of southeastern Oregon. We then evaluate available data on patterns of reproduction, as well as the nature of environmental seasonality, over the full transect from Oregon to the cape region of the Baja California Peninsula. Our study system spans the entire group of North American deserts (Fig. 3.1). *Ammospermophilus leucurus* is the most widely distributed desert-inhabiting sciurid of North America (Hall 1981; Álvarez-Castañeda and Patton 1999). Our new study site in Oregon lies at the northern outpost of the Great Basin Desert, in the intermountain region characterized by shrub-steppe and shadscale plant communities (West and Young 2000). The previously

investigated sites with available comparative reproductive data lie in the northern Mojave Desert (Kenagy and Bartholomew 1985) and at the southern extreme of the Sonoran Desert (Kenagy et al. 2005).

Materials and Methods

Sampling—We collected animals from Harney County, Oregon, in spring of 2004 and 2005, along the Fields-Denio Road, from 0.5 to 14 km south of Fields. We used Sherman live traps and euthanized animals shortly after capture, preserving them on dry ice for laboratory autopsy. Our live-animal procedures were performed in a humane manner, following ASM guidelines and an approved University of Washington animal-use protocol, as well as permits from Oregon Department of Fish and Wildlife.

In 2004 we collected eight males and seven females from 11-13 March to assess the onset of the breeding season. We returned 25 March–3 April to collect pregnant females, capturing 15 individuals. In 2005 we trapped for pregnant females 30 March-4 April, and again 11-19 April, obtaining 16 individuals. As a result of these differences in sampling date between the two years, a bias exists in the potential for detecting differences in dates of breeding for the two years; we thus have pooled the values from both years to represent the general breeding season in the study area.

Autopsy and reproductive analysis.—We dissected and weighed the paired testes, epididymides and seminal vesicles of the eight males collected 11-13 March 2004. We compared the mass of these accessory reproductive tissues to the peak values reported by Kenagy and Bartholomew (1985).

We examined females externally, dissected the reproductive tract, and examined ovaries macroscopically. We determined litter size in two ways: usually by counting embryos or in a few cases by counting recent, discrete uterine scars of sufficiently uniform size and quality to indicate that development had continued to parturition. We

determined the occurrence of lactation from the simultaneous enlargement of nipples and presence of active mammary tissue. To compare the new litter size data from Oregon with two other localities, in California and Baja California (see *Environmental data*, below, for details), we tested for significant differences by ANOVA and corrected for multiple comparisons via Tukey's HSD.

In order to estimate age and to project a date of birth for embryos, we measured the diameter of uterine swellings (embryos implanted within the uterine tract) and, if development had proceeded sufficiently, the crown-rump length and mass of embryos to obtain an average for each pregnancy. We used the growth curve reported by Kenagy and Bartholomew (1985) to estimate the age of embryos and then projected the date of parturition for each litter. We use the values obtained by Pengelley (1966), for gestation (28 days) and lactation (two months) in laboratory-raised females, to describe each component of the breeding season and to estimate its overall timing.

Environmental data.—Records of monthly and annual precipitation were obtained from NOAA's National Climate Data Center (NCDC) website (http://www.ncdc.noaa.gov/oa/ncdc.html). Due to occasional lacking data at each of two nearby weather stations, we used a combination of data from both weather stations, at Fields, Harney County, Oregon, and Denio, Humboldt County, Nevada, (32 km apart on the Fields-Denio Road) to generate a long-term precipitation profile for our study site. Because data for Fields were missing for eight of 24 months during the two years of our observations of breeding seasons, we used only the precipitation data of Denio to characterize those two years, for sake of uniformity. To represent the long-term (normal) precipitation profile we used 52 available years of monthly data from 1952-

2005; only 11 complete years were available for Fields and we substituted the additional 41 years from Denio for the missing values at Fields.

Using the predictability analysis of Colwell (1974) we compared the precipitation data from Oregon (42° 14'N, 118° 38'W) with the localities in California (Kenagy and Bartholomew 1985; 37° 11'N, 118° 14'W) and Baja California Sur (Kenagy *et al.* 2005; 24° 08'N, 110° 34'W). This treatment allowed us to assess the breeding response of *Ammospermophilus leucurus* to local environment at three positions over a large latitudinal gradient.

Results

Normal monthly levels of precipitation range from about 15-25mm throughout autumn, winter, and spring in the Alvord Desert of southeastern Oregon (Table 3.1). This is followed, in July and August, by a consistent summer drought. This pattern of precipitation drives the germination and proliferation of plants that support desert rodent reproduction. During the 2004 and 2005 breeding seasons of antelope ground squirrels, the Alvord Desert received somewhat below-normal precipitation in the first year (74%) and slightly above in the second year (112%). Despite these differences, rainfall sufficed during the critical winter-spring period to promote reproduction by the ground squirrels in both years.

In 2004 we confirmed that males were fertile in the second week of March. Of eight males examined, all had seminal vesicles and epididymides of maximal size, indicating fertility. At this time none of the seven females examined was pregnant. However, ovaries showed enlarged superficial follicles, and vulvas were open or swollen, indicating onset of estrus.

For 2004 and 2005 combined, more than 60% of births (n=31) occurred in the middle two weeks of April (Fig. 3.2). Parturition dates ranged over slightly more than a month, from 27 March to 3 May, with an average date of April 17 (SD10). Projecting back from these parturition dates, most of the mating occurred in March, with an average copulation date of 21 March. Based on a lactation period of eight weeks (two months) for an individual female, the season of lactation for the population spanned just over three months, from the last days of March through the first days of July.

Incidence of breeding by females in the population was high. Of the 37 females captured in two years, 31 (84%) were producing young, consisting of 15 in 2004 and 16 in 2005. Of the six remaining females that were neither pregnant nor lactating when captured, most showed signs of apparent estrus. This means that essentially all females in the population are capable of breeding each year. In conclusion, both the synchrony and incidence of breeding by female *A. leucurus* are high in the Alvord Desert.

Average litter size did not differ between the two years of study (P=0.17, two-sample t-test). For the 31 cases documented, mean litter size was 9.3 (SD 1.8), with a range of 6-14. Partial litter resorptions were occurring in six cases, three each year, with the following reductions: from 12 to 11 (2 cases), from 10 to 8 (2 cases), and from 9 to 8 (2 cases).

Discussion

The Oregon population of *A. leucurus* lies in the Great Basin Desert ecosystem, at the northern extreme of the geographic range of the species. It is the endpoint of a latitudinal cline of two other populations, in California and Baja California, which we compare here. The Oregon population shows a single, relatively brief breeding season in early spring associated with the winter and early spring precipitation typical of the Great Basin. The breeding season further south, in the Mojave Desert, overlaps that of the Great Basin but is shorter (Fig. 3.3). The germination of desert annual plants, vegetative growth of perennials, and further elaboration of the food chain, in response to winter precipitation, support breeding of *A. leucurus* as for other rodents associated with these deserts (Beatley 1969). Summer drought delimits the end of this critical winter-spring period of favorable environmental conditions in the northern populations. In the southern Baja California Peninsula, by contrast, the schedule of breeding is extended markedly (Fig. 3.3). This environment is characterized by a broad and bimodal season of precipitation that is associated with a subtropical monsoonal climate and a more diverse and complex biotic community. We explore here the relationships between ecosystem characteristics and the breeding of *A. leucurus* over its full latitudinal range. The demands for energy, nutrients, and water during reproduction are extremely high for small rodents that bear a large number of young. The energy demands of late lactation can amount to the greatest on a female of any time in her life, because she is supporting a mass of growing young that exceeds her own body mass (Kenagy and Bartholomew 1985; Kenagy et al. 1990).

The relatively shorter intervals of pregnancy and lactation for populations at the northern localities, in the Great Basin and Mojave Deserts (Fig. 3.3), involve a high incidence of pregnancy within the populations. Virtually all females breed, and individuals are synchronized to within only a few weeks of one another. By contrast, females in the Baja California population demonstrate a low incidence of breeding at any time. Over the prolonged period of essentially half of the year, from April through September, females can be found in any stage of reproduction from implantation of embryos to the final stages of weaning young. Thus, the combination of low incidence of breeding and a prolonged reproductive season in Baja California females amounts to a pattern of reduced precision and synchrony.

The antelope ground squirrel is an omnivore and consumes green plant food, seeds, arthropods, and even vertebrate flesh throughout the year (Bradley 1968; Grinnell and Dixon 1919). The ecological flexibility that accompanies omnivory should suit *A. leucurus* for survival over the range of spatial and temporal variability and biotic diversity of food resources that is found across the North American desert ecosystems. However, it is the increase in quality and quantity of food that becomes available following major precipitation pulses that appears to promote successful breeding by these ground squirrels. To understand the nature of the breeding response of *A. leucurus* over its geographic range, we analyzed seasonality and predictability of precipitation at our three study sites (Fig. 3.4). With average total annual precipitation of only 164mm (Oregon), 142mm (California), and 182mm (Baja California Sur), all three areas share values typical of extreme desert aridity. However, the patterns of timing and seasonality differ in important ways among the three sites. The two northern sites show sustained

precipitation in winter, which continues more strongly throughout the spring in Oregon but is truncated earlier in California by summer drought and increased heat. In both of the more northerly environments precipitation is distributed unimodally and generates a single major trophic pulse of winter and springtime biotic production. The temporal distribution of precipitation in Baja California Sur differs sharply from this and is bimodal, showing an initial, high-amplitude pulse in July-September and a secondary, lower pulse in November-January (Fig. 3.4). The heavier pulse of rain in Baja California is associated with late summer monsoonal storms and results in two monthly precipitation totals (August and September) that are more than twice as great as those of any month at the other sites (Fig. 3.4).

To evaluate predictability of precipitation statistically we used the approach of Colwell (1974), which partitions predictability into two additive components: "contingency" (predictability associated with an expected level of precipitation for any particular month) and "constancy" (predictability due to consistency of precipitation across all months). All three sites displayed significant predictability, but the proportions of predictability due to contingency (M/P) and constancy (C/P) differed among them (Table 3.2). Predictability due to constancy influences the timing of precipitation, and thus the potential timing of reproduction, in all three populations. In the two northern populations, where most of the predictability is due to constancy (C/P), *A. leucurus* shows regular and precise timing of its breeding season to match the trophic pulse that is initiated by winter rains. The lower evapotranspiration of water in these cooler northern environments assures that a greater effective moisture supply remains in the seasonal water bank as the reproductive season progresses. The year-to-year

reliability of winter precipitation in this environment appears to account for the routine activation of reproductive function in male *A. leucurus,* which occurs predictably in advance of the actual trophic pulse through the mechanism of an endogenous annual rhythm (Kenagy 1981). In contrast to the pattern of constancy and unimodality of precipitation in the northern parts of the geographic range of *A. leucurus*, our analysis indicates a different situation in the Baja California Peninsula, where predictability is mainly determined by contingency. The bimodality of annual precipitation in the southern Baja California Peninsula contributes to the weaker consistency and thus overall weaker predictability of environmental conditions in this region (Table 3.2). Thus our observations of prolongation of breeding season, weak synchronicity among females, and low incidence of breeding among females (Fig. 3.3; Kenagy et al. 2005) correspond with the divergence and greater variability of the climate of the Baja Peninsula. The lack of a single, major, potentially synchronizing pulse of precipitation in the south simply means that reproductive efforts are more likely to be spread over a greater number of months.

Patterns of precipitation, plant density and structure, and phylogenetic diversity have been characterized for North American deserts by several authors. MacMahon and Wagner (1985) showed that percentage of annual rainfall that occurs in winter decreases from north to south across the major desert ecosystems (Fig. 3.1). Thus, as observed in our comparison (Fig. 3.4), the southern Baja California Peninsula is no longer a "winter rainfall desert." MacMahon and Wagner also demonstrate that percentage plant cover in the Mojave Desert and northern Baja California Peninsula is low, generally less than 15%, but rises to 30-45%, and even higher in some local areas, in the southern

peninsula. They also show, in terms of species isopleths, that Sonoran Desert woody tree diversity increases from 5-10 species in the northern peninsula to 20-30 species in the south. These patterns are indicative of a standing crop or biomass of plant materials in the southern peninsula that far exceeds that of any of the desert systems to the north. Leon de la Luz et al. (1996) describe in detail the great phylogenetic diversity and phenology of the plant communities of the southern Baja California Peninsula. The seasonal patterns of growth and reproduction of this highly diverse flora do not show a unified seasonality, but rather a diverse array of temporal patterns. The basis of this seasonality amounts to diverse strategies of water storage and use, including below-ground and above-ground (e.g., cactus, euphorbs) storage of water. Owing to these temporal patterns and the greater biomass in the southern Baja Peninsula, we conclude that the food and water supply of *A. leucurus* is broadly served by the distinctive characteristics of this ecosystem in a manner that is consistent with the much longer, though less intense breeding season of the ground squirrels in this environment (Fig. 3.3). The more dependable nature of biologically relevant precipitation in the northern range of *A. leucurus*, consisting of the simplicity of unimodality together with Colwell's stronger measures of predictability, might explain the tight coupling of the reproductive cycles of the northern populations to their food and water resource base.

Seasonality in general, as well as the availability and predictability of food and water should influence the number of young produced by a female on each breeding occasion during her life. Comparison of litter size at all three sites reveals significant variation with latitude over the full north-south range of *A. leucurus* (Table 3.2). The number of pups per litter increases with greater latitude; females produce about nine

pups in the highly seasonal, northern environment in Oregon and about six pups in the subtropical desert at the southern end of the range. Average litter size differs statistically among all three sites, though not significantly between California and Baja California Sur. This may reflect our small sample size from Baja California and resulting low power for detecting statistical significance.

Our observation of litter size variation with latitude in *A. leucurus* is consistent with the well-recognized pattern of increase in litter size of mammals (Lord 1960) and in clutch size of birds (Lack 1954) with increasing latitude. The number of litters or clutches produced by a female annually adds to the complexity of litter size or clutch size for each breeding bout. In general these patterns derive from the greater amplitude of environmental seasonality and potentially shorter breeding seasons encountered as one moves away from the equator. Spencer and Steinhoff (1968) have argued, in light of Williams' reasoning (1967), that shorter breeding seasons at higher latitude give rise to reduced longevity and thus that natural selection acts to increase litter/clutch size at each reproductive event. Williams has articulated the evolutionary response as a tradeoff between energy allocation to immediate production of more young versus extension of the investment over a longer lifetime of less young produced per event. In terms of our ecological interpretation of the contrast between the sharper and more intense breeding of northern populations of *A. leucurus* and the less intense breeding in the southern Baja Peninsula, we can offer an additional ecological hypothesis for the latitudinal litter size cline of *A. leucurus*; we suggest that *A. leucurus* in Oregon are able to produce more young in association with the more synchronous and intense pulse of food and water resources that occurs in Oregon. Lord's (1960) initial assessment of

latitudinal litter-size patterns within genera of the squirrel family (Sciuridae) revealed increases in litter size with latitude for "tree squirrels" (*Sciurus*, *Tamiasciurus*, *Glaucomys*) and chipmunks (*Tamias*) but not "ground squirrels" (*Spermophilus*, *Ammospermophilus*), though this lack of confirmation is likely due to lack of sample size and robustness of available data. In the mean time, at the intraspecific level, our data for *A. leucurus* (Table 3.3) and those of Chapman and Lind (1973) both demonstrate significant positive relationships between litter size and latitude in ground squirrels.

Across the observed inter-population gradient in litter size and in timing and incidence of reproduction, antelope ground squirrels also show a significant pattern of genetic variation. Evidence from two neutral, mitochondrial DNA markers (the control region and cytochrome-*b*) has shown that the two northern populations of *A. leucurus* in our study share widespread, similar haplotypes, suggesting that they are closely related and recently expanded from a common ancestor (Whorley et al. 2004). The southern population, on the other hand, is characterized by a diverse and highly divergent set of haplotypes, indicating a more distant relationship to northern populations. In light of this significant genetic distance, concomitant differences in litter size and breeding patterns are perhaps not surprising. Closely related populations should share more traits in common than should more distantly related populations. The biogeographic mechanism responsible for the latitudinal variation of *A. leucurus* populations over the entire north-south range of the species has not been identified.

Table 3.1 Monthly precipitation in the Alvord Desert, Harney County, Oregon, for the "biological year" from September through August. *Monthly average, 1952-2005.

	Monthly precipitation (mm)												Annual total (mm)	Percent of normal
	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug		
2003-04	9.9	0.3	29.5	17.0	0.0	21.8	18.8	6.4	40.1	3.0	3.3	11.4	161.5	74
2004-05	1.3	43.7	9.7	7.1	20.1	19.3	17.5	24.1	85.6	16.8	0.0	0.0	245.1	112
Normal*	12.2	14.1	25.1	20.2	21.5	18.9	23.5	20.2	27.8	20.2	5.4	9.1	218.2	100

Table 3.2 Predictability of precipitation at three localities, analyzed by the method of Colwell (1974) and expressed in components of predictability (P), constancy (C), and contingency (M). The localities represent, respectively, three different North American desert systems (Fig. 3.1): Great Basin, Mojave, and the Baja California Peninsular subdivision of the Sonoran.
(1) Table 3.1.
(2) Kenagy and Bartholomew (1985)
(3) Kenagy et al. (2005)
* $P<0.001$

	Predictability of precipitation				
Locality	P	C	M	C/P	M/P
Oregon (1)	0.319*	0.229	0.090	0.718	0.282
California (2)	0.166*	0.097	0.070	0.581	0.419
Baja California (3)	0.430*	0.208	0.221	0.485	0.515

Table 3.3 Litter size in *Ammospermophilus leucurus.*
Comparison of litter size among the three populations of *Ammospermophilus leucurus* (cf. Fig. 3.3).
(1) Present study.
(2) Kenagy and Bartholomew (1985)
(3) Kenagy et al. (2005)

*Significance, by ANOVA, corrected for multiple comparisons via Tukey's HSD: Oregon vs. California ($P<0.0001$), Oregon vs. Baja California ($P<0.0001$), and California vs. Baja California ($P=0.080$).

Locality	Latitude	Mean*	SD	N	Range
Oregon (1)	42°	9.3	1.8	31	6–14
California (2)	37°	7.4	1.7	43	5–11
Baja California (3)	24°	5.9	1.5	7	4–8

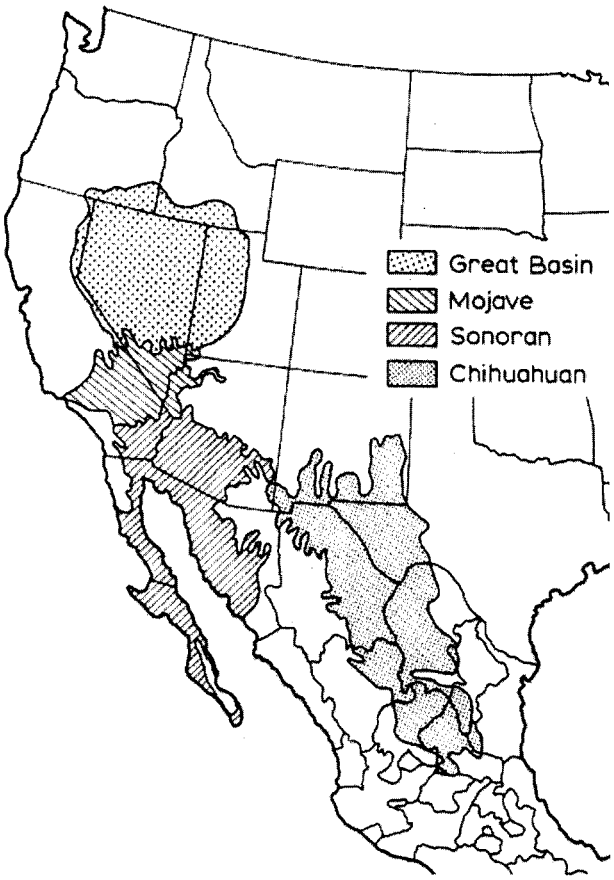


Fig. 3.1 Major North American desert systems (after MacMahon and Wagner 1985). The geographic range of the antelope ground squirrel, *Ammospermophilus leucurus*, extends across three of these deserts, from the northern extreme of the Great Basin to the southern tip of the Baja California Peninsula.

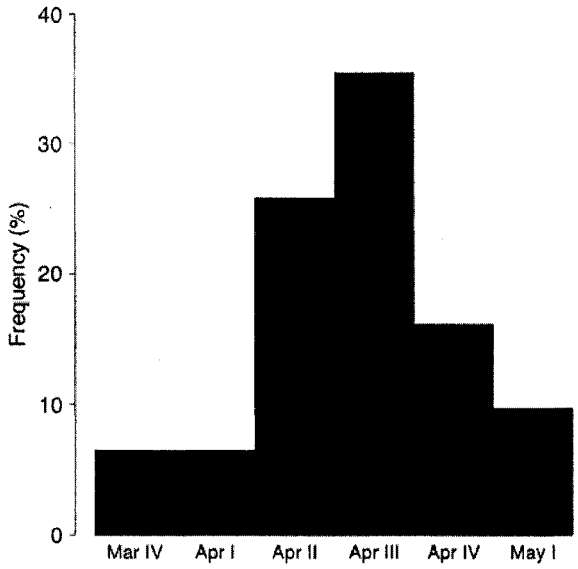


Fig. 3.2 Frequency histogram (by week) of projected parturition dates for 31 female *A. leucurus* over the 2004 and 2005 breeding seasons in the Alvord Desert, Harney County, Oregon. Data for two years are pooled due to differences in sampling dates between the two years.

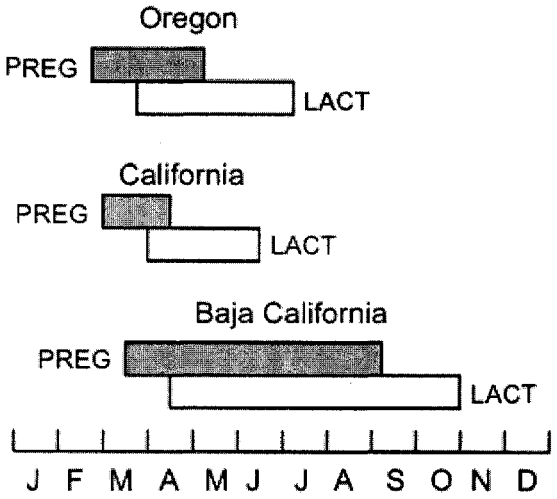


Fig. 3.3 Graphical summary of breeding seasons in *A. leucurus* populations from three major North American desert systems: Great Basin (Oregon; present study), Mojave (California; Kenagy and Bartholomew 1985), and Sonoran-Peninsular (Baja California; Kenagy et al. 2005). Horizontal rectangles represent intervals (by week) of projected pregnancy (PREG) and lactation (LACT) based on field observations of presence of embryos and lactation in females.

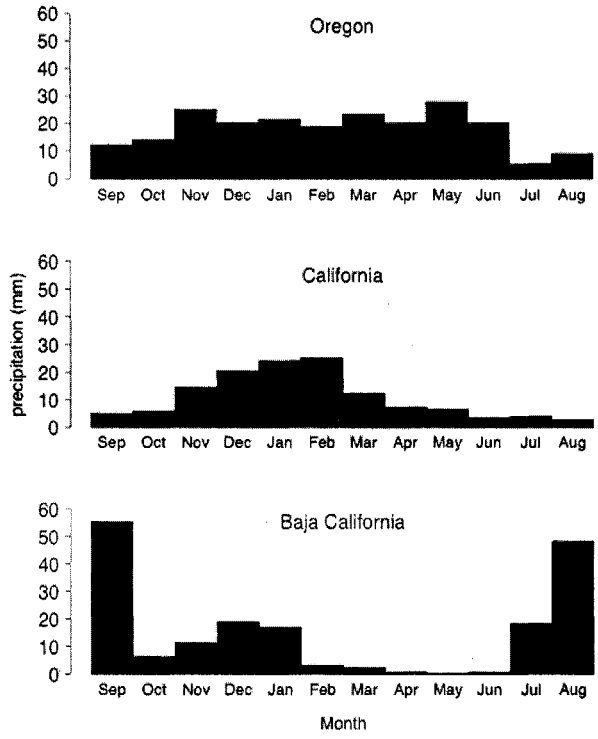


Fig. 3.4 Long-term ("normal") monthly precipitation at three localities spanning the North American desert systems compared here (cf. Fig. 3.3, Table 3.2.)

References

Alderson GW, Gibbs HL, Sealy SG (1999) Parentage and kinship studies in an obligate brood parasite, the brown-headed cowbird (*Molothrus ater*), using microsatellite DNA markers. *Journal of Heredity* **90**, 182-190.

Álvarez-Castañeda ST, Patton. JL (1999) Mamíferos del noroeste de México. Centro de Investigaciones Biológicas del Noroeste, La Paz, Baja California Sur, México.

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819-825.

Arbogast BS (1999) Mitochondrial DNA phylogeography of the New World flying squirrels (Glaucomys): Implications for Pleistocene biogeography. *Journal of Mammalogy*. **80**, 142-155.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002) Estimating divergence times from molecular data on phylogenetic and population genetic time scales. *Annual Review of Ecology and Systematics*, **33**, 707-740.

Ashton KG, de Queiroz A (2001) Molecular systematics of the Western rattlesnake, *Crotalus viridis* (Viperidae), with comments on the utility of the D-loop in phylogenetic studies of snakes. *Molecular Phylogenetics and Evolution*, **21**, 176-189.

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

Baric S, Sturmbauer C (1999) Ecological parallelism and cryptic species in the genus *Ophiothrix* derived from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **11**, 157-162.

Beatley, JD (1969) Dependence of desert rodents on winter annuals and precipitation. *Ecology* **50**:721-724.

Bergmann A (1847) Über die Verhältnesse der Wärmeökonomie der Thiere zu ihrer Grösse. *Göttinger Studien*, **1**, 595-708.

Bertorelle G, Slatkin M (1995) The number of segregating sites in expanding human populations with implications for estimates of demographic parameters. *Molecular Biology and Evolution*, **12**, 887-892.

Bidlack AL, Cook JA (2001) Reduced genetic variation in insular northern flying squirrels (*Glaucomys sabrinus*) along the North Pacific Coast. *Animal Conservation*, **4**, 283-290.

Black CC (1963) A review of the North American Tertiary Sciuridae. *Bulletin of the Museum of Comparative Zoology at Harvard University*, **130**, 109-248.

Bond JE, Hedin MC, Ramirez MG, Opell BD (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.

Booth DB (1987) Timing and processes of deglaciation along the southern margin of the Cordilleran ice sheet. In: *North America and adjacent oceans during the last deglaciation* (eds Ruddiman WF, Wright, HE Jr.), pp. 71-90. Geological Society of America, Boulder, Co.

Bradley WG (1968) Food habits of the antelope ground squirrel in southern Nevada. *Journal of Mammmalogy* **49**,14-21.

Bradley RD, Baker RJ (2001) A test of the genetic species concept: Cytochrome-*b* sequences and mammals. *Journal of Mammalogy*, **82**, 960-973.

Chapman JA, Lind GS (1973) Latitude and litter size of the California ground squirrel, *Spermophilus beecheyi*. *Bulletin of the Southern California Academy of Sciences* **72**,101-105.

Colwell RK (1974) Predictability, constancy, and contingency of periodic phenomena. *Ecology* **55**,1148-1153.

Conroy CJ, Cook JA (2000) Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Rodentia: Muridae). *Molecular Ecology*, **9**, 165-175.

Coyne JA (1992) Genetics and Speciation. *Nature* **355**, 511-515.

da Silva MNF, Patton JL (1998) Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Molecular Ecology*, **7**, 475-486.

Darwin C (1859) *The Origin of Species by Means of Natural Selection*, 6th edition. The Modern Library, New York.

Di Rienzo A, Wilson AC (1991) Branching patterns in the evolutionary tree for human mitochondrial DNA. *Proceedings of the National Academy of Sciences USA*, **88**, 1597-1601.

Edwards SV (1997) Relevance of microevolutionary processes to higher level molecular systematics. In: *Avian Molecular Evolution and Systematics* (ed Mindell DP), pp. 251-278. Academic Press, San Diego, CA.

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

Endler JA (1973) Gene flow and population differentiation. *Science*, **179**, 243-250.

Endler JA (1977) Geographic variation, speciation, and clines. *Monographs in population biology* **10**, Princeton University Press, Princeton, New Jersey.

Endler JA (1986) *Natural Selection in the Wild.* Princeton University Press, Princeton, NJ.

Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genetics* **164**, 1567–1587.

FAUNMAP Working Group (1994) FAUNMAP: a database documenting late quaternary distributions of mammal species in the United States. Illinois State Museum, Springfield, Illinois.

Felsenstein J (1981) Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution*, **35**, 124-138.

Felsenstein J (1992) Estimating effective population size from samples of sequences: inefficiency of pairwise and segregating sites as compared to phylogenetic estimates. *Genetical Research*, **59**, 139-147.

Fisher RA (1950) Gene frequencies in a cline determined by selection and diffusion. *Biometrics*, **6**, 353-361.

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

Garza JC, Dallas J, Duryardi D, Gerasimov S, Croset H, Boursot, P (1997) Social structure of the mound building mouse *Mus spicilegus* revealed by genetic analysis with microsatellites. *Molecular Ecology* **6**, 1009-1017.

Goodwin HT, Reynolds RE (1989) Late Quaternary Sciuridae from low elevations in the Mojave Desert, California. *The Southwestern Naturalist*, **34**, 506-512.

Grayson, DK (1993) The Desert's Past. A Natural History of the Great Basin. Smithsonian Institution Press, Washington.

Grinnell J, Dixon J (1919) Natural history of the ground squirrels of California. *Monthly Bulletin of the State Commission of Horticulture*, **7**, 597-708.

Grinnell J (1926) Geography and evolution in the pocket gopher. University of California Chronicle, **July** 247-262.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.

Guindon S, Lethiec F, Duroux P, Gascuel O (2005) PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* **Jul 1, 33**(Web Server issue), W557-9.

Gustafson EP (1978) The vertebrate faunas of the Pliocene Ringold formation, south-central Washington. *Bulletin of the Museum of Natural History University of Oregon*, **23**, 1-62.

Haldane JBS(1948) The theory of a cline. *Journal of Genetics*. **48**, 277-284.

Hale ML, Bevan R, Wolff K (2001) New polymorphic microsatellite markers for the red squirrel (*Sciurus vulgaris*) and their applicability to the grey squirrel (*S. carolinensis*). *Molecular Ecology Notes*, **1**, 47-49.

Hall ER (1946) The mammals of Nevada. University of California Press, Berkeley, California.

Hall ER (1981) The mammals of North America. Second ed. John Wiley and Sons, NewYork.

Hasegawa M, Kishino K, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160-174.

Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hoang A, Hill CE, Gibert P, Beerli P (2001) Strength and tempo of directional selection in the wild. *Proceedings of the National Academy of Sciences of the USA*, **98**, 9157-9160.

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

Irwin DM, Kocher TD, Wilson AC (1991) Evolution of the cytochrome *b* gene of mammals. *Journal of Molecular Evolution*, **32**, 128-144.

Keane, TM Creevey CJ, Pentony, MM, Naughton TJ, JO McInerney (2006) Assessment of methods for amino acid matrix selection and their use on empirical data

shows that ad hoc assumptions for choice of matrix are not justified, *BMC Evolutionary Biology*, **6**, 29.

Kenagy GJ (1981) Endogenous annual rhythm of reproductive function in the non-hibernating desert ground squirrel, *Ammosphermophilus leucurus*. Journal of Comparative Physiology. **142**,251-258.

Kenagy GJ, Bartholomew GA (1985) Seasonal reproductive patterns in five coexisting California desert rodent species. *Ecological Monographs,* **55**, 371-397.

Kenagy GJ, Masman D, Sharbaugh SM, Nagy KA (1990) Energy expenditure during lactation in relation to litter size in free-living golden-mantled ground squirrels. *Journal of Animal Ecology* **59**,73-88.

Kenagy, GJ, Whorley JR, Cortés-Calva P, Álvarez-Castañeda ST (2005) Timing of reproduction in antelope ground squirrels, *Ammospermophilus leucurus*, near La Paz, Baja California Sur. Pp. 269-274 In: *Contribuciones mastozoológicas en homenaje a Bernardo Villa.* (V. Sánchez-Cordero and R. A. Medellín, eds.) Instituto de Biología, UNAM; Instituto de Ecología, UNAM; CONABIO. México.

Kimura M, Maruyama T (1971) Patterns of neutral polymorphism in a geographically structured population. *Genetical Research.* 18, 125-131.

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P (2001) The strength of phenotypic selection in natural populations. *American Naturalist*, **157**, 245-261.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *American Naturalist*, **150**, 1-23.

Kocher, TD, WK Thomas, A Meyer, SV Edwards, S Paabo, FX Villablanca and AC Wilson (1989) Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences USA*, **86**, 6196-6200.

Kocher TD, Conroy JA, McKaye KR, Stauffer JR (1993) Similar morphologies of cichlid fish in Lakes Tanganyika and Malawi are due to convergence. *Molecular Phylogenetics and Evolution*, **2**, 158-165.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics*, **149**, 429-434.

Lack, D (1954) The natural regulation of animal numbers. Clarendon Press, Oxford, U.K.

Leon de la Luz, JL, R Corio Benet, M Cruz Estrada. (1996) Fenología floral de una comunidad árido-tropical de Baja California Sur, México. *Acta Botánica Mexicana* **35**,45-64.

Lord, RD (1960) Litter size and latitude in North American mammals. *American Midland Naturalist* **64**,488-499.

Lundstrom R, Ţavare S, Ward RH, (1992) Modelling evolution of the human mitochondrial genome. *Mathematical Biosciences*, **122**, 319-335.

MacMahon JA, Wagner FH (1985) The Mojave, Sonoran, and Chihuahuan deserts of North America. In: *Hot Deserts and Arid Shrublands* (eds Evenari M, et al.), pp 105-202. Elsevier Science Publishers, B. V., Amsterdam.

Mantel, N (1967) The detection of disease clustering and a generalized regression approach. *CancerResearch* **27**, 209-220.

Matocq DM (2002) Phylogeographical structure and regional history of the dusky-footed woodrat, *Neotoma fuscipes*. *Molecular Ecology*, **11**, 229-242.

May B, Gavin TA, Sherman PW, Korves TM (1991) Characterization of microsatellite loci in the Idaho ground squirrel *Spermophilus brunneus brunneus*. *Molecular Ecology*, **6**, 399-400.

May RM, Endler JA, McMurtrie RE (1975) Gene frequency clines in the presence of selection opposed by gene flow. *American Naturalist*, **109**, 659-676.

Mayr E (1956) Geographic character gradients and climatic adaptation. *Evolution*, **10**, 105-108.

Mayr E (1963) Animal Species and Evolution. Harvard University Press. Cambridge, Massachusetts

Miller WE (1980) The late Pliocene Las Tunas fauna from southernmost Baja California, México. *Journal of Paleontology*, **54**, 762-805.

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

Paetkau D, Calvert W, Sterling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**, 347-354.

Pengelley ET (1966) Differential developmental patterns and their adaptive value in various species of the genus *Citellus*. *Growth* **30**, 137-142.

Petersen AT, Soberón J, Sánchez-Cordero V (1999) Conservation of ecological niches in evolutionary time. *Science*, **285**, 1265-1267.

Pielou EC (1991) *After the Ice Age: the Return of Life to Glaciated North America.* University of Chicago Press, Chicago.

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

Powers DA and Schulte PM (1998) Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: A multidisciplinary approach to address the million-year saga of a small fish. *Journal Of Experimental Zoology*, **282**, 71–94.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Queller DC, Strassman JE, Hughes CR (1993) Microsatellites and Kinship. *Trends in Ecology and Evolution*, **8**, 285-288.

Ramos-Onsins SE, Rojas J (2002) Statistical properties of new neutrality tests against population growth. *Molecular Biology and Evolution*, **19**, 2092-2100.

Raymond M, Rousset F (1995) An exact test for population differentiation. Evolution, 49 : 1280-1283.

Reichman OJ, van de Graaf KM (1975) Association between ingestion of green vegetation and desert rodent reproduction. *Journal of Mammalogy* **56**, 503-506.

Reynolds HG, Turkowski F (1972) Reproductive variations in the round-tailed ground squirrel as related to winter rainfall. *Journal of Mammalogy* **53**, 893-898.

Reznick D,Travis J (1996) The empirical study of adaptation in natural populations. In: *Adaptation* (eds Rose MR, Lauder GV), pp 243-289. Academic Press, San Diego, CA.

Rice WR, Hostert EE (1993) Laboratory experiments on speciation: What have we learned in 40 years? *Evolution*, **47**, 1637-1653.

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

Riddle BR, Hafner DJ, Alexander LF, Jaeger JR (2000) Cryptic vicariance in the historical assembly of a Baja California peninsular desert biota. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 14438-14443.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution,* **9**, 552-569.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079-1089.

Schneider S, Roessli D, Excoffier L (2000*) ARLEQUIN, Version 2.000: a Software for Population Genetic Data Analysis.* Genetics and Biometry Laboratory, Department of Anthropology and Ecology, University of Geneva, Switzerland.

Second J, Faivre B, Bensch S (2006) Spreading introgression in the wake of a moving contact zone. *Molecular Ecology*, **15**, 2463-2475.

Shreve F (1942) The desert vegetation of North America. *Botanical Review*, **8**, 195-246.

Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555-562.

Spradling TA, Hafner MS, Demastes JW (2001) Differences in rate of cytochrome-*b* evolution among species of rodents. *Journal of Mammalogy*, **82**, 65–80.

Spencer AW, Steinhoff. HW (1968) An explanation of geographic variation in litter size. *Journal of Mammalogy* **49**, 281-286.

Steven S, Coffin J, Strobeck C (1997) Microsatellite loci in Columbian ground squirrels *Spermophilus columbianus*. Molecular Ecology, **6**, 493-495.

Sullivan J, Arellano E, Rogers DS (2000) Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations *American Naturalist* **155**, 755–768.

Swofford, DL (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. *Molecular Biology and Evolution*, **10**, 512-526.

Upton DE, Murphy RW (1997) Phylogeny of the side-blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. *Molecular Phylogenetics and Evolution*, **8**, 104-113.

Van Devender TR, Thompson RS, Betancourt JL (1987) Vegetation history of the desert of southwestern North America; The nature and timing if the Late Wisconsin-Holocene transition In: *North America and adjacent oceans during the last deglaciation* , v. K-3 (ed. Ruddiman WF, Wright HE), pp. 323-352 The Geological Society of America, Boulder, Colorado.

Via S (1994) The evolution of phenotypic plasticity: what do we really know? In: *Ecological Genetics* (ed Real LA), pp.35-57. Princeton University Press, Princeton, NJ.

West NE, Young JA (2000) Intermountain valleys and lower mountain slopes. Pp. 255-284 in North American terrestrial vegetation, 2nd ed. (M. G. Barbour and W. D. Billings, eds.). Cambridge University Press, Cambridge.

Whorley JR, Alvarez-Castaneda ST, Kenagy GJ (2004) Genetic structure of desert ground squirrels over a 20-degree-latitude transect from Oregon through the Baja California Peninsula. *Molecular Ecology* **13**, 2709-2720.

Wilcox TP, Hugg L, Zeh JA, Zeh DW (1997) Mitochondrila DNA sequencing reveals extreme genetic differentiation in a cryptic species complex of neotropical pseudoscorpions. *Molecular Phylogenetics and Evolution*, **7**, 208-216.

Williams G C (1966) Natural selection, the costs of reproduction, and a refinement of Lack's principle. American Naturalist, **100**, 687-690.

Wingfield JC, Kenagy GJ (1991) Natural regulation of reproductive cycles. Pp. 181-241 in *Vertebrate endocrinology: Fundamentals and biomedical implications.*

Volume 4, Part B (P. K. T. Pang and M. P. Schreibman, eds.). Academic Press, New York.

Wright S (1951) The genetical structure of populations. *Annual Eugenics*, **15**, 323-354.

Yang Z, Rannala B (2005) Branch length prior influences Bayesian posterior probability of phylogeny *Systematic Biology*, **54**, 455-470.

Zamudio KR Greene HW (1997) Phylogeography of the bushmaster (*Lachesis muta:* Viperidae): Implications for neotropical biogeography, systematics, and conservation. *Biol. J. Linn. Soc.***62:** 421–442.

Zamudio KR, Jones KB, Ward RH (1997) Molecular systematics of short-horned lizards: Biogeography and taxonomy of a widespread species complex. *Systematic Biology*, **46,** 284-305.

Zheng X, Arbogast BS, Kenagy GJ (2003) Historical demography and genetic structure of sister species: deermice (*Peromyscus*) in the North American temperate rain forest. *Molecular Ecology*, **12**, 711-724.

Zink RM (1996) Comparative phylogeography in North American birds. *Evolution*, **50**, 308-330.

Vita

Joshua R. Whorley was born in Inglewood, CA. He lived in Colorado as a young boy, but grew up mostly in Southern California. He earned his Bachelors of Arts in Integrative Biology from the University of California at Berkeley in 1998. He earned his PhD from the University of Washington at Seattle in 2006. He lives in Seattle now and hopes to stay. He wants to climb the world and teach high school biology.