COALESCENT ANALYSIS OF FIFTEEN NUCLEAR LOCI REVEALS PLEISTOCENE SPECIATION AND LOW GENETIC DIVERSITY IN THE MOJAVE FRINGE-TOED LIZARD, *UMA SCOPARIA*

By

Andrew David Gottscho

A Thesis

Presented to

The Faculty of Humboldt State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Arts

Biological Sciences

March, 2010

COALESCENT ANALYSIS OF FIFTEEN NUCLEAR LOCI REVEALS PLEISTOCENE SPECIATION AND LOW GENETIC DIVERSITY IN THE MOJAVE FRINGE-TOED LIZARD, *UMA SCOPARIA*

By

Andrew David Gottscho

We certify that we have read this study and that it conforms to acceptable standards of scholarly presentation and is fully acceptable, in scope and quality, as a thesis for the degree of Master of Arts.

Approved by the Master's Thesis Committee:

_____

Dr. W. Bryan Jennings, Co-Major Professor Date

_____

Dr. Sharyn B. Marks, Co-Major Professor Date

_____

Dr. Edward C. Metz, Committee Member Date

_____

Dr. Jeffrey W. White, Committee Member Date

_____

Dr. Michael R. Mesler, Graduate Coordinator Date

_____

Jená Burges, Vice Provost Date

ABSTRACT

COALESCENT ANALYSIS OF FIFTEEN NUCLEAR LOCI REVEALS PLEISTOCENE SPECIATION AND LOW GENETIC DIVERSITY IN THE MOJAVE FRINGE-TOED LIZARD, *UMA SCOPARIA*

Andrew D. Gottscho

Analyzing DNA sequence data from multiple unlinked nuclear loci in a coalescent Isolation-with-Migration (IM) model is a statistically powerful method for estimating population divergence times, effective population sizes, and gene flow. This approach was used to reconstruct the evolutionary history of the Mojave fringe-toed Lizard, *Uma scoparia*, which is restricted to windblown sand habitats in the Mojave and Colorado Deserts of southern California and western Arizona, a region that is thought to have undergone dramatic climatic change during and since the Pleistocene epoch. To shed light on the origin of this species, I analyzed 15 nuclear loci (621,694 total bp) representing twenty localities of *U. scoparia* and four localities from its sister species to the south, *U. notata*. I found a latitudinal gradient in heterozygous SNPs and indels, low nucleotide diversity in *U. scoparia* (π = 0.148%, SD = 0.167%), particularly relative to *U. notata* (π = 0.469%, SD = 0.366%), and reciprocal monophyly in 3/15 gene trees. Using the IM model, I estimated with 95% confidence that *U. scoparia* and *U. notata* speciated in the Pleistocene epoch (~1 – 1.4 mya, 95% CI ~0.7 mya – 2.1 mya) without significant gene flow ($2Nm < 1$), an estimate that is robust to violations of the no-recombination assumption. I also found that *U. notata* has 2-5 times the effective population size of *U.*

scoparia. These findings suggest that *U. scoparia* originated in the Pleistocene epoch and was confined to a Colorado Desert refuge during glacial maxima; northern populations represent a recent range expansion.

ACKNOWLEDGEMENTS

I would like to thank the U.S. Army at Fort Irwin National Training Center, the Joshua Tree National Park Association, Community Foundation, Needles Field Office of the Bureau of Land Management (BLM), Humboldt State University, and Judith and Alistair McCrone for funding this research. Brad Hollingsworth of the San Diego Natural History Museum kindly provided a tissue sample that I used to create my genomic library, and Robert Murphy of the Royal Ontario Museum donated the *U. inornata* sample used in this study. Cameron Rognan, Jeff Jarvis, Collin Grant, Jennifer Taylor, Bonnie Riddell, Mitch Mulks, Kendra Gietzen, Sharyn Marks, Bryan Jennings and Barry Sinervo helped collect tissue samples. Bryan Jennings, Sean Reilly, Mitch Hart, Anthony Baker, and Maria Harper provided assistance with laboratory work. Jim Andre and Tasha La Doux at the Granite Mountains Reserve and Bill Presch at the Zzyzx Desert Studies Center provided us with a place to sleep and eat while collecting tissue samples, and taught me much about desert natural history. Adam Leache provided helpful tips for data analysis, and Robert Reynolds assisted with my geologic interpretations. Bryan Jennings and Sharyn Marks advised me throughout the duration of this project, and Ed Metz and Jeff White helped revise my thesis. I would like to thank the California Department of Fish and Game, Arizona Department of Game and Fish, BLM, and National Park Service for providing me with permits to collect DNA samples.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Recently separated allopatric populations or species are of great interest to evolutionary biologists who are interested in the mechanisms of speciation, but these sibling taxa also pose several challenging questions. How has the physical and climatic landscape influenced population divergence and gene flow? At what point during the continuous speciation process should a geographically disjunct population be considered an incipient species or a full species? The answers to these questions have important implications not only for speciation theory, but also for biodiversity conservation in the face of anthropogenic impacts like habitat fragmentation and global climate change.

To address these questions, many phylogeneticists and phylogeographers have relied on mitochondrial DNA (mtDNA) or microsatellite loci to estimate population divergence times or gene flow (Avise 1998; Brito and Edwards 2009). For a single locus, mtDNA harbors a disproportionately large amount of information regarding demographic history due to its high mutation rate (relative to autosomal nuclear DNA) and its small effective population size (Jennings and Edwards 2005; Moore 1995; Avise 1998). However, a single gene tree does not necessarily reflect the true population history, which can be confounded by deep coalescence (i.e. incomplete lineage sorting), branch length heterogeneity, and gene flow (Brito and Edwards 2009; Edwards 2009). In other words, ancestral populations often have many polymorphisms, and the sorting of these polymorphisms into descendent populations is a stochastic process; the topology of any one gene tree does not necessarily reflect the topology of the true species tree. Recent

studies have demonstrated a random nature to the coalescent process, which leads to high heterogeneity among gene trees (Edwards 2009). Because gene divergence usually occurs before population divergence, mtDNA studies may overestimate population divergence due to ancestral polymorphisms, yet the extent of overestimation is impossible to measure from one locus (Jennings and Edwards 2005; Edwards and Beerli 2000; Beerli and Edwards 2002). Because it is maternally inherited, mtDNA cannot be used to assess male gene flow; this is significant because existing studies suggest that males may disperse more than females (Doughty *et al*. 1994).

Microsatellites, although useful for identifying individual organisms, nevertheless suffer from high mutation rates (and thus high homoplasy), and contain insufficient information for their gene trees to be reconstructed. Additionally, they are afflicted with ascertainment bias, defined as a distortion of the true frequency of a phenomenon due to non-random sampling (Jennings and Edwards 2005). Recent bottleneck events in one or both populations can increase estimates of genetic divergence, even with substantial gene flow, and vicariance events involving large subpopulations on either side of an impermeable barrier could incorrectly be interpreted as a sign of high gene flow (Hey *et al.* 2004; Zhivotovsky 2001). Most microsatellite studies have made heavy use of summary statistics like F_{ST}. However, a given F_{ST} value can be explained by assuming short divergence times and little gene flow or long divergence times and strong gene flow (Nielsen and Wakeley 2001).

Fortunately, a coalescent-based Isolation with Migration (IM) model that uses a Markov Chain Monte Carlo (MCMC) approach is now available to analyze multiple

unlinked nuclear loci to distinguish between ancestral lineage sorting and gene flow as explanations for genetic divergence (Wakeley 1996; Nielsen and Wakeley 2001; Hey *et al.* 2004). This model assumes panmictic populations (random mating), no intra-locus recombination, and no natural selection. By analyzing multiple non-coding regions of the nuclear genome, this method has several key advantages over mtDNA and microsatellite methods (Brito and Edwards 2009). Because anonymous loci represent random segments of nuclear DNA, they represent variation across the entire genome better than any one marker does (Lee and Edwards 2008). Unlike microsatellites, anonymous loci do not suffer from ascertainment bias. Since non-coding loci comprise approximately 90% of the genome, they can be harvested in virtually unlimited amounts. Most non-coding DNA presumably isn't under natural selection, and therefore meets coalescent assumptions of neutrality. With multiple loci, one can estimate confidence intervals for demographic parameters, including population divergence times, effective population sizes, and gene flow (Hey *et al.* 2004). Finally, unlike mitochondrial DNA, which is maternally inherited, nuclear DNA can be used to estimate gene flow due to male-biased dispersal (Brito and Edwards 2009).

Fringe-toed lizards (genus *Uma*, family Phrynosomatidae) are restricted to sand dunes, hummocks, flats, and other fine windblown (aeolian) sand habitats in the deserts of North America (Stebbins 1944; Norris 1958). *Uma* possess a number of morphological and behavioral adaptations for life in this extremely hot, arid, and exposed landscape (Figure 1; Stebbins 1944; Carothers 1986; Luke 1986). Two disjunct eastern species, *U. exsul* and *U. paraphygas*, occur in the Chihuahuan desert of central Mexico, separated

from their western kin by the Sierra Madre Occidental. *Uma* to the west of the Sierra Madre Occidental diverged into a northern form, *U. scoparia*, and its sister southern form, the *U. notata* species complex. This split corresponds precisely with a long yet narrow chain of mountains that lies parallel to the San Andreas Fault in California's Riverside and Imperial Counties. In some areas, such as around Joshua Tree National Park, the two species approach within thirty miles of each other (Figure 2). There are two easily recognized morphological differences between the *U. notata* species complex and *U. scoparia* (Figure 3).

The *U. notata* species complex occurs to the south of these mountains, in the lower Colorado River Valley, and is composed of three closely allied yet allopatric subspecies: the Colorado Desert fringe-toed lizard, *U. n. notata* in the lower Colorado Desert of California and Baja Norte, Mexico; *U. n. inornata* in the Coachella Valley of Riverside County, California; and *U. n. rufopunctata* in southwest Arizona and Sonora, Mexico, east of the Colorado River (Trepanier and Murphy 2001; Norris 1958). Together, these subspecies occupy the lower Colorado and Sonoran Desert life zones of North America.

The focal species of this study, the Mojave Fringe-toed Lizard, *Uma scoparia*, ranges approximately from the Colorado River at Parker, Arizona, north to southern Death Valley, California, and west to Barstow and Twenty-nine Palms. Although most of this area is considered the Mojave Desert, the species' common name is a misnomer, as the southern part of its range near the Colorado River is widely considered to be part of the Colorado Desert biome, or at least a transition zone between the two desert biomes (Brown 1994). Within the Mojave Desert, *U. scoparia* occupies low, hot, arid valleys in

the creosote bush shrub desert, below the Joshua tree/pinyon pine/juniper zone (personal observation).

The taxonomy within the *Uma notata* species complex has been dynamic over the past sixty years (Heifetz 1941; Norris 1958; Carpenter 1963; Mayhew 1964; Adest 1977; Trepanier and Murphy 2001). Norris (1958) pointed out that because all the western forms of *Uma* are allopatric, the taxonomic problem is of the most difficult sort, and species designations may be subjective. In particular, *U. n. inornata*, listed as endangered (Chen *et al.* 2006), is widely granted full species status (Stebbins 2003), whereas *U. n. rufopunctata* has only recently been elevated to a full species (Trepanier and Murphy 2001). Because I am examining the deepest divide in the western forms, between *U. scoparia* and *U. notata*, and the morphological, genetic, and geological evidence concur that all subspecies of *U. notata* are more closely related to each other than any are to *U. scoparia* (Norris 1958; Trepanier and Murphy 2001), I will treat the *U. notata* species complex as one species in my model (see Materials and Methods for more discussion on this point). Herein, I will refer to the entire *U. notata* species complex as such, or simply as *U. notata*. I will refer to the *U. notata* and *U. scoparia* clade as "western *Uma*" to distinguish them from the Chihuahuan species.

There is an extensive literature on the effects of the Pleistocene epoch on shaping genetic diversity and promoting speciation in terrestrial animals (Avise *et al.* 1998; Hewitt 1996, 2000, 2004; Klicka and Zink 1997; Edwards and Beerli 2000). Recent evidence supports the idea that desert-adapted floral and faunal communities in western North America were fragmented into isolated refugia during Pleistocene glacial maxima.

Particularly, studies on the composition of fossilized packrat (*Neotoma*) middens have shown that much of southwestern North America, particularly the valleys of the Mojave Desert, contained open coniferous woodland that is currently restricted to higher elevations or latitudes (Bentancourt *et al*. 1990a, 1990b; Thompson and Anderson 2000). Although glaciers did not intrude into the range of *Uma*, the climate was considerably more mesic, and numerous permanent lakes and streams graced the region, most notably Lake Manix in the vicinity of Barstow, CA and Lake Manly in Death Valley (Enzel *et al*. 2003). In parts of the Mojave Desert, fossils of giant sloths, horses, and other large herbivorous mammals have been found; these animals could not have survived in the modern desert climate (Norris 1958).

However, *Neotoma* midden studies also demonstrated that much of the lower Colorado Desert retained a desert flora and fauna during glacial maxima (Cole 1986). They were southerly enough, and low elevation enough to retain a climate suitable for desert-adapted species, including the creosote bush and the Joshua tree. These studies have also suggested that a smaller northern refuge was also present in Death Valley, which also has a low-elevation, hyper-arid climate. Recent paleoclimatic niche modeling also supports the idea that both the lower Colorado Desert and Death Valley sheltered isolated desert communities during the Pleistocene epoch (Jezkova *et al*. 2009).

A strong connection existed between Pleistocene climatic cycles and sand dune formation. There are three essential requirements for sand dune formation: 1) a source of sand, typically supplied by rivers or oceans, 2) exposure to wind, and 3) a place for the sand to settle out, often at the base of a mountain range or another obstacle that causes a

reduction in wind velocity. During mesic Pleistocene glacial maxima, the Mojave and Colorado Deserts had more permanent lakes and streams that transported sand to the inter-mountain basins. During more arid cycles of the Pleistocene and Holocene, the water dried up, exposing the sand to the wind, creating dune systems (Lancaster and Tehakerian 2003; Enzel *et al.* 2003).

Given the dynamic nature of eolian systems, and the specialization of *U. scoparia* for this habitat, the historical biogeography of this species has generated much interest. Norris (1958) found clinal morphological variation between the northwestern, intermediate, and southern populations of *U. scoparia*. He presented considerable evidence suggesting that *U. notata* and *U. scoparia* diverged in the recent geological past, yet the two forms have been separated long enough for fixed differences to have developed in several morphological traits (Figure 3). He also found that northern *U. scoparia* had surprisingly minor levels of morphological variation, despite the fact that they occur in isolated metapopulations. Based on this evidence and his knowledge of the climate and geology of the region, Norris hypothesized that *U. scoparia* originated during the early Pleistocene in the Colorado Desert, was displaced from the northern part of its range during glacial maxima, and then colonized the Mojave Desert by migrating northward along the edges of stream courses.

Murphy *et al*. (2006) constructed a mitochondrial gene tree for *U. scoparia*, and found that the most basal split is between the Amargosa River in the north, and all other populations in the south. A population at Red Pass Dune (located in the Mojave River drainage just upstream from the Amargosa River) possesses mitochondrial haplotypes

from both the Amargosa River population and the Mojave River population. The authors postulated that the Amargosa River population was likely isolated from the other populations 500,000 years ago, and that Death Valley was a Pleistocene refuge for this species. Based on these mtDNA data, Murphy *et al*. defined two distinct population segments (as defined in the Endangered Species Act) within *U. scoparia*: the Amargosa River population and the Red Pass Dune population, located at the northern end of the Mojave River drainage. These findings of distinct northern lineages are similar to those of other habitat associates of *Uma*, including the desert pocket mouse (*Chaetodipus penicillatus*; Jezkova *et al.* 2009), the sidewinder (*Crotalus cerastes*), and the speckled rattlesnake (*C. mitchelli*; Douglas *et al.* 2006).

Here, I reconstruct the evolutionary history of *U. scoparia* by analyzing fifteen nuclear loci (620,000 bp total) in a coalescent IM model in an attempt to distinguish between the alternate hypotheses of Norris (1958) and Murphy *et al.* (2006). I aim to 1) estimate the population divergence time (speciation time) between *U. scoparia* and *U. notata*, 2) estimate gene flow since speciation between *U. scoparia* and *U. notata*, 3) determine ancestral and modern effective population sizes, 4) determine the location of any Pleistocene refugia within western *Uma* by categorizing heterozygous single nucleotide polymorphisms (SNPs) and heterozygous indels by population, and 5) examine the gene trees for any evidence of a divergent Amargosa River population.

MATERIALS AND METHODS

Field Collections

DNA samples were collected throughout the range of *Uma scoparia* between March and June 2008; in May 2009, I collected samples from three representative populations of *U. notata* in the Colorado Desert of California and southwestern Arizona (Table 1; Figure 2). I actively searched for lizards in dune habitat during the day, usually in the mornings and late afternoons, when I found them to be most active. After stalking them carefully, my field assistants and I captured most of the lizards with a fishing pole noose. A few of the lizards were captured by hand by tracking the lizard to the burial site. This method was most effective in the early morning following a spell of windy weather that cleared the dunes of old tracks. Tail tips (usually less than 1 cm) were removed at an inter-vertebral break point and preserved in 100% ethanol for DNA analysis. Date and time of capture, snout-vent length, sex, and general weather conditions were noted. A GPS unit was used to record latitude, longitude, and elevation. Sand surface temperatures in the sun and shade and lizard body temperatures were recorded using an infrared temperature gun. Finally, each lizard was photographed and released at the same location within minutes of capture.

Genomic Data Acquisition

For reasons stated earlier, multiple anonymous nuclear loci are the molecular markers of choice. I assume that these loci recombine amongst themselves freely, since

they are from random portions of the nuclear genome (Edwards and Beerli 2000; Arbogast *et al*. 2002). To design primers for these loci, a small-insert genomic library was created. Genomic DNA was extracted from liver tissue of an *Uma scoparia* using standard phenol-chloroform methods (Maniatis *et al.* 1982). Pressured nitrogen was used to nebulize the DNA into fragments between 1 and 4 kb in length. This was accomplished by placing a nebulization vessel on ice, filling it with approximately 6.8 µg DNA and 750 µL shearing buffer, and running the gas at 14 PSI for approximately 30 seconds. After nebulization, the liquid was transferred to a 1.5 mL microcentrifuge tube. Eighty µL of 3 molar sodium acetate, 4 µL of 20 mg/mL glycogen, and 700 µL 100% isopropyl alcohol were added and mixed, and placed in a -80° C freezer for 15 minutes. After that, I centrifuged the mixture at 12,000 RPM for 15 minutes at 4° C. The pellet was washed with cold 80% ethanol, and the centrifugation step repeated. The fluid was decanted and the tube spun dry until a DNA pellet was observed. I re-suspended the pellet with 55 µL of distilled water, and then I ran the re-suspended mixture on a 1% agarose gel with a ladder to confirm that the nebulization step produced DNA fragments of the appropriate size (1 – 4 kb).

The nebulized DNA was then blunt-end repaired by adding blunting buffer, BSA, dNTPs, T4 polymerase, and Klenow polymerase. This mixture was incubated for 30 minutes at room temperature, followed by an incubation at 70° C for 20 minutes. To dephosphorylate the blunt-ended mixture, I added dephosphorylation buffer and calf intestinal phosphatase (CIP), and incubated the mixture at 37° C for 60 minutes. I then centrifuged the mixture to create a pellet, which was subsequently washed with ethanol,

dried, and resuspended with water. The blunt-ended fragments of DNA were then ligated into a blunt-end TOPO vector before being transformed into chemically-competent *Escherichia coli* cells (Invitrogen, Carlsbad, CA). Clones were then picked at random with a toothpick, plasmids were purified using a mini-prep kit (Qiagen, Valencia, CA), and sequenced using vector primers from both ends. I then designed polymerase chain reaction (PCR) primers nested within each lizard-insert sequence. The inserts in eight clones were fully sequenced. A Genbank BLAST search was conducted to determine whether any of the sequences were homologous with coding DNA.

Out of the eight primer pairs tested, only five successfully amplified their target loci in PCR reactions. Recent empirical data suggest that as the number of loci in an IM analysis increases from five to fifteen, the 95% confidence intervals for parameter estimates narrow rapidly; beyond fifteen loci, confidence in parameter estimates are not much improved (Lee and Edwards 2008; Jennings and Edwards 2005). Therefore, I attempted other ways to acquire primers for nuclear loci in order to reach the optimum number, fifteen. First, I ordered primers for four introns that have been used successfully in studies of other phrynosomatid lizards, bringing my total to nine loci (Leache 2009; Leache and McGuire 2006). However, I still needed six more loci.

A recent study on turtles demonstrated that anonymous loci primers can be used on increasingly distantly related taxa with varying degrees of success (Thomson *et al*. 2008). Generally speaking, most anonymous loci primers tend to be fairly specific to the taxa on which they were developed, at least compared to introns; this is because the primer sequences are usually located in non-coding regions of the genome which are free

to mutate without the constraints of natural selection. However, if enough loci from one organism are tested on a distantly related organism, a fraction of them are likely to successfully amplify the target sequence. Therefore, I decided to test 38 anonymous loci primers that successfully amplified PCR products for several fence lizard (*Sceloporus*) species, which are also in the family Phrynosomatidae (Rosenblum *et al*. 2007). I used a gradient thermocycler to test each primer pair at 12 different annealing temperatures between 50-60° C, in some cases up to 67° C. Six loci successfully amplified their targets in both *U. notata* and *U. scoparia*, bringing my total to fifteen loci. The primer sequences and annealing temperatures are given in Table 2.

The following PCR settings were used to directly sequence the fifteen nuclear loci: 1) 94° for 5 minutes, 2) 94° for 20 seconds, 3) annealing temperature (Table 2) for 20 seconds, 4) 72° for 1 minute, 5) go to step 2 thirty-five times, 6) 72° for 3 minutes. The success of the PCR reactions was verified by running the products out on a 2% agarose gel, staining it with ethidium bromide, and imaging it with UV light. DNA sequencing was outsourced to the University of Washington High-Throughput Genomics Center.

Designation of Populations

In order to categorize population level genetic diversity, as well as to conduct IM analyses, it is necessary to make *a priori* population designations (Table 1). Fortunately, in the case of *Uma*, these populations are easily defined by the landscape topography, at least on a regional scale – mountains are natural barriers that define the course of sand

dune movement due to wind action (Muhs *et al.* 2003). I group all subspecies of *U. notata* as one population; none of these subspecies are separated by mountain ranges, and gene flow was likely common before humans fragmented dune ecosystems and dammed the Colorado River in the past century. Although this may violate the assumption of panmictic populations in the IM model (Hey *et al.* 2004), recent empirical data suggest that the model is robust to violations of this assumption (Strasburg and Rieseberg 2009). Furthermore, I am principally concerned with the speciation of *U. scoparia*, not the intra-specific history of the *U. notata* species complex.

Within *U. scoparia*, I define three populations, stratified by latitude. The southern Colorado River (Colorado) population includes the Amboy, Cadiz, Dale, Pinto, Palen, Rice, and Bouse localities. All of these populations are either in the Colorado River drainage, or in a sand transport corridor that directly connects to the Colorado River (Muhs *et al.* 2003). This region is generally considered an ecological transition zone from true Colorado Desert, to Mojave Desert (Brown 1994). To the north of the Colorado population, separated by a low divide, is the Mojave River (Mojave) population, which includes the following localities: Lenwood Road, Coyote Dry Lake, Bitter Springs, Red Pass Dune, West Red Pass, The Whale, Cronese Lakes, Afton Canyon, Rasor Road, and Kelso Dunes. Lizard metapopulations at all of these locales were almost certainly interconnected by the Mojave River during the late Pleistocene and early Holocene (Enzel *et al*. 2003). Finally, the Amargosa River population, as defined by Murphy *et al*. (2006), consists of Ibex Dunes, Dumont Dunes, and Coyote Holes. This population represents the northern range limit for *Uma*.

Sequence Processing and Data Characteristics

I first used FinchTV v1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com) to browse through all sequences and dispose of those with moderate to high levels of background noise, which would interfere with the detection of SNPs in directly sequenced PCR products. I then used CodonCode Aligner v3.5.2 (CodonCode Inc.; Dedham, MA, USA; http://www.codoncode.com/aligner) to align and convert the remaining high-quality .abi files into contigs. One of the main advantages of CodonCode aligner is that it allows for the automatic resolution of heterozygous indels, as well as base calling for heterozygous SNPs (double peaks in the chromatogram). Heterozygous SNPs and indels are important because they are a reliable proxy for measuring genetic diversity, especially with large sample sizes (Abdulla *et al*. 2009; Novembre *et al*. 2008; Wakeley *et al.* 2000). I trimmed the ends of the contigs that had a high degree of mutations due to background noise in the chromatograms, then I used the "Change Bases" command to automatically call bases as ambiguities when the second peak was more than 50% of the height of the first peak. All segregating sites were manually verified to ensure that bases calls were accurate, and I made manual edits as necessary, which is critical to prevent excess noise from biasing the dataset.

I counted all heterozygous SNPs and indels (insertions or deletions), and categorized them by population and by locus. This enabled me to calculate the total SNPs per population. This raw number is not meaningful by itself, however, due to my uneven sample sizes among loci and populations. Therefore, to correct for the uneven sampling, I

divided the total SNPs by the total number of base pairs analyzed per population. This gave me a frequency of SNPs per population (Total SNPs/Total Base Pairs = SNP Frequency).

Determination of Haplotype Sequences and Recombination

The program PHASE v2.1 was used to infer haplotypes (Stephens *et al.* 2001; Stephens and Scheet 2005). This program reconstructs haplotype sequences of heterozygous individuals mainly based on other homozygous sequences at the same locus. However, PHASE is not infallible, so it calculates a probability value for each inferred allele. Generally, most authors only use inferred haplotype sequences with 90% probability or greater; however, this may bias against unique SNPs in the dataset (Becquet and Przeworski 2009). To illustrate this point, let us consider a directly sequenced PCR product that has one unique heterozygous SNP. In this case, PHASE (or the user) can easily determine the haplotype sequence. However, if the sequence has two SNPs, one of which is unique to that sequence, then there are two possible ways of inferring the haplotypes, each with a 50% probability. Therefore PHASE will incorrectly estimate the phase of unique SNPs 50% of the time. To include all inferred alleles with 50% probability or higher in an IM analysis will inevitably create four gamete types (essentially a false recombination event), which will bias the estimates of demographic parameters. If one deals with this by excluding all inferred alleles with less than 90% confidence, the number of recombination events is drastically reduced, but a new bias emerges against unique SNPs, which are a biologically real component of the data. This

is particularly important if one population in the analysis has an unusually high number of unique SNPs, as might be expected in a Pleistocene refuge. By excluding recombination, one is biasing the data by removing unique, real polymorphisms; and by including recombinant alleles, one is violating the assumptions of the IM model.

To deal with this paradox, I decided to run my analysis both ways to see how robust the results are to these violations and biases. First I used PHASE to reconstruct all alleles with 50% confidence or greater, thereby preserving a complete dataset with all unique SNPs. I conducted the first IM analysis including all of these data. For the second analysis, I first used the 4-gamete test in DNAsp 5.10 to detect recombination events in the complete dataset (Librado and Rozas 2009; Hudson and Kaplan 1985). For loci that did not pass the test, I ran PHASE again, this time only preserving alleles with 90% confidence or greater. This removed most, but not all, of the detected recombination events. The final few recombination events were dealt with by either truncating the locus to the shortest non-recombining block or removing one or two "problem" alleles.

The number of alleles per population per locus is given in Table 3. I used the program DNAsp 5.10 (Librado and Rozas 2009) to calculate nucleotide diversity (π), the number of segregating sites (S), and the number of fixed differences (F_{sn}) between *U. scoparia* and *U. notata*, as well as between the Amargosa and Mojave populations within *U. scoparia* (F_{am}). Tajima's *D* test for neutrality was calculated to ensure that the loci are not under natural selection and therefore meet the molecular clock assumption of the IM model (Hey *et al.* 2004).

Mutation Models and Gene Tree Reconstruction

I used FindModel, a web implementation of ModelTest, to find the best mutation model (Posada and Crandall 1998). To reconstruct my gene trees I used the maximum-likelihood phylogenetic software Garli v1.0, which uses a rapid stochastic algorithm to simultaneously determine the topology, branch lengths and substitution model parameters that maximize the log-likelihood (Zwickl 2006; Lewis 1998). For each locus, I conducted six independent runs of Garli, and I checked that they all came up with the same best tree. I used the midpoint rooting method with the software FigTree v 1.3.1.

Coalescent Demographic Analyses

The program IM was used to estimate effective population sizes, population divergence time, and gene flow between *U. scoparia* and *U. notata* (Hey and Nielsen 2004; Nielsen and Wakeley 2001). The IM program assumes that the loci are unlinked and free from natural selection and recombination, and that the populations are panmictic. Six population parameters scaled by mutation rate (μ) were estimated: the effective population sizes for the ancestral population (θ_A, $\theta = 4N\mu$) and for the two daughter populations (θ_1 and θ_2); the divergence time between the two descendent populations (t); and migration rates between daughter populations (m_1, m_2), in the coalescent. Since the coalescent goes backward in time, m_1 is the rate at which genes enter population 1 from population 2 as time goes forward (Hey *et al.* 2004). I applied the HKY mutation model to my data. In IM, several preliminary runs were made until prior boundaries were

optimized (Reilly 2009). The final simulations were carried out with a geometric heating scheme, six chains, 5 million steps, and a burn-in of 2 million steps. Effective sample size (ESS) values were monitored to ensure proper mixing of the Markov chain. The following priors were used: -q1 12 -q2 12 -qA 12 -t 5 -m1 5 -m2 5 -b 2000000 -L 5000000 -fg -n6 -g1 0.8 -g2 0.9.

To convert model parameter estimates from the IM program into demographic quantities (population size, migration rate, and divergence time), I started with a mutation rate of 2.2×10^{-9} substitutions/site/year for mammalian nuclear DNA (Kumar and Subramanian 2002). This mutation rate was multiplied by the number of base pairs in each locus to calculate the per-locus mutation rate μ (substitutions/year). The geometric mean of all four per-locus mutation rates was taken to calculate the divergence time by dividing the time parameter by the mutation rate ($t = t/\mu$)**.** To convert the effective population size parameter (θ) to population size in terms of number of breeding adults (N), I used the equation $N = \theta/(4G\mu)$, where G is the number of years until sexual maturity. To estimate the population migration rate 2Nm (the effective rate at which genes come into a population per generation), I used the equation $2Nm = \theta * m/2$. A 2Nm value less than one indicates that gene flow is not significant (Wright 1931).

RESULTS

Analysis of Heterozygosity

I examined 273,480 bp of directly sequenced nuclear DNA, before inferring the haplotypes, to categorize heterozygous SNPs and indels by population (Table 4; Figure 4). I excluded locus Uma7r from the analysis because it had an atypically large number of heterozygous SNPS (179), with few homozygotes, in violation of what one would expect from a locus in Hardy-Weinberg equilibrium. All other loci had between 0 and 59 total SNPs. Within the remaining fourteen loci, *U. notata* had the highest heterozygous SNP frequency of any population (0.33% of 60,836 bp were SNPs). Within *U. scoparia*, the southern Colorado population had the highest SNP frequency (0.09% of 81,578 bp). The northern Mojave and Amargosa populations had similarly low frequencies of SNPs (0.05% of 73,369 bp and 0.07% of 57,697 bp respectively). Using SNP frequencies as a proxy for genetic diversity, *U. notata* is about three times more diverse than southern *U. scoparia*, and *U. notata* is approximately five times more diverse than either northern (Mojave & Amargosa) *U. scoparia* population. The occurrence of heterozygous indels, which were rare in this study, is consistent with the pattern of heterozygous SNPs. *Uma notata* had five heterozygous indels, the Colorado population of *U. scoparia* had three, and the Mojave and Amargosa populations had zero.

Data Characteristics

After using PHASE and SeqPhase to infer all haplotypes with 50% probability or higher, I had 621,694 bp of data, the characteristics of which are summarized in Table 5. The total length of all sites is 7,496 bp, and the mean locus length is 500 bp (SD = 150 bp). I found a total of 171 segregating sites in all 15 loci, 14 (8.2%) of which are fixed differences between *U. notata* and *U. scoparia*, which translates to a 0.19% sequence divergence between the two species (14 fixed differences / 7496 total bp). There were zero fixed differences between the Amargosa and Mojave populations (0% sequence divergence). The average number of segregating sites per locus was 11.4 (SD = 4.75), or 2.69% of the total base pairs on average (SD = 1.89%). The mean nucleotide diversity (π) of *U. notata* (0.00469, SD = 0.00366) was approximately three times higher than that of *U. scoparia* (0.00148, SD = 0.00167). None of the results of Tajima's *D* test were significant, with the exception of locus Uma08, indicating that the majority of my loci are not under natural selection and thus meet the assumptions of the IM model.

For the no-recombination analysis, a total of 35 recombination events were detected; most of these were caused by PHASE error, because by subsequently excluding all alleles with less than 90% confidence, I removed all but 8 of these recombination events. The final recombination was dealt with by excluding 12 allele sequences and truncating two loci to the shortest non-recombining block. I also excluded locus Uma07, because recombination was so common in this locus.

Characteristics of Gene Trees

The fifteen midpoint-rooted gene trees are given in Figures 5-19. Three loci (RAG-1, Sun08, and Uma06) show reciprocal monophyly between the two species. Four of the loci (Sun07, Sun10, Sun18, and Uma05) had monophyletic *U. scoparia* nested within *U. notata*, so that *U. notata* is paraphyletic with respect to *U. scoparia*. In seven of the loci (BDNF, PNN, R35, Sun12, Sun28, Uma03, and Uma08), the outgroup was composed entirely of *U. notata*, while the ingroup had a mixture of *U. notata* and *U. scoparia* alleles, so that neither species was monophyletic. The fourteen gene trees mentioned so far all have an outgroup of pure *U. notata*, and *U. scoparia* is always either reciprocally monophyletic or nested within the *U. notata* genealogy. The sole exception to this pattern occurs in Locus Uma07 (Figure 18), which has an unusually large number of SNPs; in this case *U. notata* is nested in several locations within the *U. scoparia* lineage.

One of the most striking aspects of the gene trees is the lack of population structure within *U. scoparia*; most of the gene trees have a bush-like topology. The most extreme example is in locus BDNF (Figure 5), in which all 56 *U. scoparia* alleles representing the entire species range are 100% identical. By contrast, *U. notata* had much more structure even within a single locality, especially at the Algodones Dunes.

Coalescent Analyses

I conducted two separate IM analyses, the first including all of my data with recombination, and the second a pared-down dataset with no recombination that biases against unique SNPs. All parameters were estimated well, because all the posterior probability distributions were shaped as single, symmetrical, narrow peaks. The raw parameter estimates and the converted demographic parameters are shown in Table 6. For the first analysis, including all data with recombination, the geometric mean of mutation rates (μ) was calculated to be 1.037×10^{-6} mutations/year. The effective population size for *U. notata* (576,529 individuals; 95% CI 428,694 – 855,761 individuals; Figure 20) was approximately five times larger than that of *U. scoparia* (123,190 individuals; 95% CI 90,338 – 162,610 individuals; Figure 21), while the ancestral population size (106,764 individuals; 95% CI 21,335 – 215,167 individuals; Figure 22) was slightly smaller than that of *U. scoparia*. The speciation time between *U. notata* and *U. scoparia* was estimated to be 1.04 million years ago (mya) (95% CI 0.71 – 1.53 mya; Figure 23). The population migration rate distributions from *U. scoparia* to *U. notata* ($2N_1m_1 = 0.0897$; Figure 24) and from *U. notata* to *U. scoparia* ($2N_2m_2 = 0.003833$; Figure 25) indicated negligible amounts of gene flow.

For the second analysis, excluding all recombination, the geometric mean of mutation rates (μ) was calculated to be 1.009×10^{-6} mutations/year. The effective population size for *U. notata* (200,979 individuals; 95% CI 137,785 – 291,994 individuals; Figure 26) was approximately three times larger than that of *U. scoparia*

(56,878 individuals; 95% CI 41,717 – 84,697 individuals; Figure 27), while the ancestral population size (1,263 individuals; 95% CI 1,263 – 324,866 individuals; Figure 28) was much smaller than any other estimated population size. The speciation time between *U. notata* and *U. scoparia* was estimated to be 1.41 mya (95% CI 0.74 - 2.08 mya; Figure 29). Again, gene flow was negligible in both directions ($2N_1m_1 = 0.096$; $2N_2m_2 = 0.007$; Figures 30 and 31).

DISCUSSION

I examined 621,694 base pairs of DNA (after PHASE), representing fifteen independent nuclear loci, in a coalescent IM model to reconstruct the evolutionary history of the Mojave Fringe-toed Lizard, *Uma scoparia*, a sand dune specialist native to the Mojave and Colorado deserts of California and Arizona. This study not only demonstrates that *U. scoparia* has low genetic diversity compared to its sister species, *U. notata*, whether measured by SNPs, indels, nucleotide diversity or effective population size, but it also is consistent with the hypothesis that *U. scoparia* arose via founder effect speciation in the Pleistocene epoch, a finding that is robust to violations of the no-recombination assumption. I interpret these results to mean that western *Uma* were confined to the Colorado Desert during Pleistocene glacial maxima, in support of Norris' (1958) hypothesis. It is also consistent with the hypothesis of Avise *et al.* (1998) that the Pleistocene played an important role in creating many modern sister species in vertebrates. In addition, this study demonstrates the feasibility of using multiple anonymous nuclear loci in studies of recently diverged species that still share some alleles. The six anonymous nuclear loci developed for *S. undulatus* (Rosenblum *et al.* 2007) that also amplify PCR products for *Uma*, across the most basal split in Phrynosomatidae (Wiens *et al.* 2010), should be of use to researchers interested in the evolution of lizards in this family.

The first striking pattern observed in the dataset, before using PHASE to infer haplotypes, is the asymmetrical distribution of heterozygous SNP and indels between the

northern and southern populations of western *Uma*. In particular, I found that *U. notata* has approximately three times more heterozygous SNPs than the southern Colorado population of *U. scoparia,* which in turn has almost double the SNPs found in the Mojave River population. The SNP frequency in the Amargosa River population is intermediate between that of the Mojave River and Colorado River populations. Heterozygous indels, rare in this study, were only found in lizards in the Colorado Desert or the transition zone between the Colorado and Mojave Deserts.

This finding is especially significant because my sampling was biased towards *U. scoparia*; I only sampled *U. notata* within a subset of its full range, and therefore I likely underestimated the true genetic diversity of *U. notata*. For the most part, these raw data represent a latitudinal gradient in heterozygosity, although the SNP diversity in the northernmost Amargosa population is intermediate between the Colorado and Mojave River populations. A clear latitudinal gradient of genetic diversity has also recently been demonstrated in human populations in eastern Asia using genomic data; more than 90% of northeast Asian haplotypes were found in southeastern Asia (Abdulla *et al.* 2009). Here, the accepted explanation is that southeastern Asia was a major geographic source of northeastern populations. This pattern is essentially the same as that seen in western *Uma*.

My analysis of nucleotide diversity (π) revealed similar results. *Uma notata* (0.469%) was approximately three times more diverse than *U. scoparia* (0.148%), which is the same ratio observed in heterozygous SNP diversity between the two species. How does this nucleotide diversity compare to the genomes of other vertebrates? Genomic data

is currently only available for a few non-model organisms, but one can make several general comparisons. The mean nucleotide diversity for nuclear DNA of *U. scoparia* is about 1.5 times higher than that of humans (0.088%) and southern elephant seals (0.09%), which are generally regarded to have low genetic diversity (Zhao *et al.* 2000; Slade *et al.* 1998). By contrast, *U. scoparia* is approximately seven times less diverse than the black salamander, *Aneides flavipunctatus* (1.1%; Reilly 2009) and about ten times less diverse than the red-backed fairy wren, *Malurus melanochephalus* (1.6%; Lee and Edwards 2008).

Between *U. scoparia* and *U. notata*, I found that the level of fixed differences in the nuclear DNA (14 out of 7,496 bp; 0.19% sequence divergence) is low. This is not surprising considering that the two species are morphologically similar and ecologically identical to one another (Stebbins 1944; Norris 1958). The fixed differences between *U. scoparia* and *U. notata* occur across the continuous mountain ranges associated with the San Andreas Fault (personal observation). By contrast, the Amargosa River population does not have a single fixed difference from the Mojave River population (0% sequence divergence), which is not unexpected given that the Mojave River overflows into the Amargosa River when its current terminus at Silver Lake reaches capacity, and no mountains exist that might have impeded the movement of sand dunes and lizards between these drainages in historical times.

The patterns in the gene trees reflect the youth of *U. scoparia*. No geographic structuring within *U. scoparia* is evident, particularly between the Mojave and Amargosa populations, which is expected given that they have 0% sequence divergence. In fact,

most of the *U. scoparia* alleles resemble a "bush" in each gene tree. By contrast, there is more divergence within the Algodones Dunes, within a few hundred square meters, than there is in the entire Mojave Desert, suggesting founder effect speciation for *U. scoparia*. This is strikingly similar to a study that shows that some people within southern Africa are more genetically divergent from one another than Europeans are from Asians (Schuster *et al.* 2010). The widely accepted interpretation of this pattern is that modern humans originated in southern Africa and a small subset of this ancestral population migrated north to colonize the Sahara Desert, Europe, and Asia. Applying the same logic to my data, the Colorado Desert near the delta of the Colorado River was likely the center of origin for western *Uma.*

While the gene trees don't support splitting *U. scoparia* into distinct populations, there is evidence that *U. scoparia* and *U. notata* are truly divergent: three of the fifteen gene trees show reciprocal monophyly between *U. scoparia* and *U. notata*, and seven of the fifteen trees show *U. scoparia* as a monophyletic group. *U. notata* is paraphyletic with respect to *U. scoparia* in 11 of 15 trees, yet the reverse pattern is not observed, except in Locus Uma07, which has an unusually high number of SNPs and probably can't be trusted as a reliable marker of demographic history. Again, this strongly suggests that *U. scoparia* originated as a northward expansion of *U. notata*. These data are consistent with coalescent theory, which predicts that recently diverged species will share alleles at some loci and yet be reciprocally monophyletic at other loci, if effective population sizes are large enough, due to the stochastic nature of lineage sorting (Brito and Edwards 2009).

Analyses of both datasets suggest a mid-Pleistocene speciation time (~1 mya for the full dataset and ~1.4 mya excluding recombination) with the 95% confidence intervals falling entirely within the Pleistocene epoch (~0.7 mya – 2.1 mya). All gene flow values, and their 95% confidence values, were not significant ($2Nm < 1$). Therefore, my finding of allopatric Pleistocene speciation is robust to violations of the recombination assumption, and the 0.3 mya bias is not large relative to the width of the confidence intervals.

Using the full dataset with recombination, the effective population size of *U. notata* is about five times larger than that of *U. scoparia*, which is consistent with the other data characteristics (SNP frequencies, indels, nucleotide diversity) suggesting that *U. notata* is more genetically diverse. Excluding recombination, the effective population sizes are smaller: *U. notata* had only double the effective population size of *U. scoparia*. This is expected because this dataset biases against unique SNPs in alleles that already have SNPs, which were primarily found in *U. notata*.

Evolutionary History of *Uma scoparia*

The geologic events leading to the speciation of *U. scoparia* and *U. notata* began well before the actual speciation event itself. The Colorado River was the main sand source for the Gran Desierto, a vast sea of dunes where the ancestor to western *Uma* probably adapted to arenicolous life. The Salton Trough, the northern end of the Gulf of California where the Colorado River currently empties into the Gulf, began to open 6 mya as the Baja peninsula continued to drift north with the Pacific Plate along the strike-

slip San Andreas Fault system (McKibben 2008; Elders *et al.* 1972). Shortly thereafter (5 mya), the Colorado River delta was formed as sediment from the river filled the basin (McKibben 2008). This delta supplied the sand for the Gran Desierto, the most extensive dune field in North America.

The late Pliocene and Pleistocene epochs were periods of great orogeny in western North America; in particular, the Sierra Nevada Mountains, the Transverse Ranges, and the Peninsular Ranges were rapidly uplifted, creating a rain shadow to the east (Thorne 1986). The uplift of the Transverse Ranges, particularly the San Bernardino Mountains, in turn created the Mojave River (Meisling and Weldon 1989; Enzel *et al.* 2003; Lancaster and Tchakerian 2003). The first cold glacial maxima of the Pleistocene likely displaced western *Uma* to the southern Colorado Desert (Norris 1958). The actual speciation event between *U. scoparia* and *U. notata* may have been caused by dispersal north along the Colorado River, during which time *U. scoparia* clearly went through a founder effect as its effective population size was greatly reduced. As the climate became more hot and arid, *U. scoparia* continued to expand north following stream courses and sand transport corridors from the Colorado River to the newly created Mojave River, losing genetic diversity along the way (Figure 32; Norris 1958; Muhs *et al.* 2003).

The Mojave River, which supplies much of the sand presently occupied by *U. scoparia*, flows northeast from the San Bernardino Mountains and currently terminates in Silver Lake, near Baker, California (Figure 2; Enzel *et al*. 2003). The river was created over two million years ago when the Transverse Ranges were uplifted; it took over a million years to fill its upstream basins with sediment before creating Lake Manix.

Approximately 25,000-15,000 years ago, Lake Manix overflowed its boundaries and incised Afton Canyon, subsequently creating the Cronese Lakes, Silver Lake, Soda Lake, and the Kelso Dunes (Enzel *et al.* 2003; Lancaster and Tchakerian 2003). During peak periods of flow, the Mojave River drained from the north end of Silver Lake into the Amargosa River (Enzel *et al*. 2003). Today that prehistoric drainage is still visible, and takes the form of a dry wash (personal observation). Further evidence that the Mojave River is geologically young is provided by geological dating of sand dunes; for example, the Kelso Dunes at the terminus of the Mojave River drainage are no more than 35 kya (Lancaster and Tehakerian 2003).

Since this study revealed zero genetic divergence between the geographically isolated Mojave and Amargosa populations at all fifteen nuclear loci, how did *U. scoparia* disperse between these dune systems? In considering this question, it is important to recall that sand dunes are highly dynamic and mobile entities. Norris (1958) speculated that "dune deposits were probably often more continuous then than they are today. Reworking and transport of old dunes have destroyed what continuity existed and isolated the sand deposits and the biota restricted to them." Miller and Stebbins (1964) wrote that the distribution of *U. scoparia* "doubtless has become greatly fragmented with the breaking up and sorting of the lake shore and river bank sand deposits, which must have been more continuous in the more humid past than now. Since the species is so completely restricted to aeolian sand, one can picture some of the populations carried about on their sand 'islands' as the deposits have moved over the centuries." These moving dunes are known as bachran dunes (Norris 1958), and even large dunes may

move dozens of meters per year. Norris also noted that blow-ups and sand shadows have allowed migration across low divides between river drainages, including the divide between the Mojave and Amargosa Rivers, which at the time of his writing was covered with sand and occupied with *Uma,* although this doesn't appear to be the case today (personal observation).

The patterns in the nuclear DNA are not entirely consistent with the patterns in mtDNA in *U. scoparia*. There are some similarities – for example, Murphy *et al*. (2006) noted that *U. scoparia* from the Cadiz Dunes in the Colorado River population have higher mtDNA haplotype diversity than do lizards from the Ibex Dunes in the Amargosa River population. However, there are important distinctions as well; Trepanier and Murphy (2001) found reciprocal monophyly between *U. notata* and *U. scoparia*, but I only found this pattern in 3/15 gene trees. Murphy *et al.* (2006) found that the basal split in the mtDNA gene tree of *U. scoparia* corresponds with a split between the Amargosa River population and all other populations, yet I did not observe an Amargosa River clade in any of the gene trees, nor did I find a single fixed difference between the Amargosa and Mojave River populations.

What could account for these discrepancies between the patterns of mitochondrial and nuclear DNA? This same question has been debated with many other organisms, particularly humans, who are also thought to have gone through a recent population bottleneck or rapid population expansion (Hey 1997; Hey and Harris 199; Fay and Wu 1999). Compared to mtDNA, non-coding nuclear loci have low mutation rates and four times the effective population size, which means that nuclear gene trees will nearly

always be less resolved than the corresponding mtDNA gene tree; the polymorphism patterns of loci with different population sizes may be out of phase with each another if the period of population size fluctuation is short and the fluctuation has occurred recently (Hey and Harris 1999; Fay and Wu 1999). Also, because gene divergence always predates population divergence, it is easy to over-estimate divergence times based on a single mitochondrial gene tree. Finally, mtDNA only represents part of the population history due to its maternal inheritance. If males tend to disperse to new localities while females remain near their hatch site, significant gene flow could occur among nuclear genes but not mitochondrial genes. This may be the case in iguanid lizards, including *Uma*, where most species have highly territorial males that drive out younger conspecifics of the same sex (Doughty *et al.* 2006). For all these reasons, this multi-locus coalescent study emphasizes the necessity of incorporating nuclear DNA in phylogeographic analyses.

The findings of this study are a reminder that the desert is a rapidly changing, geologically young landscape that was much different during glacial maxima of the Pleistocene. Plate tectonics and climatic cycles exert powerful influences on the creation and structure of mountain ranges and sand dunes. During repeated glacial maxima, heavy precipitation eroded sand from mountains and deposited it into stream and lake beds. Dunes were smaller in extent and had more stabilizing vegetation, and were more continuously distributed along river and lake shorelines. Soil-producing processes dominated. During more arid periods, vegetation died off and lakes and rivers became playas and washes. Thus more sand was exposed to the wind, creating expansive dune fields (e.g., Kelso Dunes). During these arid periods, dunes grew in size, yet were

fragmented from each other. This explains the interesting patchy distribution of *Uma scoparia*. To the untrained eye, the Mojave Desert seems ancient, and its sand dunes seem like isolated islands in a sea of mountains, dry lake beds, and alluvial fans, but this is an illusion – on a geologic time scale, the desert is relatively young, and dunes are highly mobile.

Comparisons with other Mojave Desert Phylogeography Studies

It is interesting to compare the results of this study with recent phylogeography studies of other co-distributed organisms in the Mojave Desert, such as the sidewinder and speckled rattlesnakes (*Crotalus cerastes* and *C. mitchelli/stephensi*; Douglas *et al.* 2006), the desert pocket mouse (*Chaetodipus penicillatus*; Jezkova *et al.* 2009), the night lizard (*Xantusia vigilis*; Leavitt *et al.* 2007) and brittlebush (*Encelia farinosa*; Fehlberg and Ranker 2009). It appears that the Colorado Desert was the only major Pleistocene refuge for the night lizard and brittlebush. On the other hand, there is evidence of a distinct Amargosa/Death Valley lineage in the desert pocket mouse, the sidewinder and the speckled rattlesnake. This discrepancy can likely be explained by the fact that pocket mice and rattlesnakes (even the sand-adapted sidewinder) are not as restricted to sand dunes as are fringe-toed lizards, and they could have more easily dispersed over alluvial fans or low mountain ranges, unlike *Uma*. The geographic distribution of the pocket mice and rattlesnakes also extends further north than that of *Uma,* for example into northern Death Valley, southern Nevada, and the Owens Valley. Fringe-toed lizards do not occur in the upper Amargosa River dunes, or the Death Valley dunes, which should be perfectly

suitable habitat, since these locations are occupied by most of the desert habitat associates of *Uma*, including squamates like the zebra-tailed lizard (*Callisaurus draconoides*), the desert iguana (*Dipsosaurus dorsalis*), and the sidewinder (*C. cerastes*), as well as plants including creosote (*Larrea tridentata*) and burroweed (*Ambrosia dumosa*). Therefore *U. scoparia* would probably thrive in these dunes if they could disperse there. Norris (1958) rightly pointed out that all these unoccupied dunes are separated from the present range of *U. scoparia* by unsuitable habitat – large expanses of rocky terrain, salt flats, or alluvial fans.

Future Directions

Although this study has revealed much about the evolutionary history of western *Uma*, there is much more to learn. In particular, the relationships within the *U. notata* species complex are in need of further investigation, as well as the divergence between eastern and western *Uma* (that is, across the Sierra Madre Occidental).

This study has demonstrated the utility of using multiple unlinked nuclear loci to address questions regarding speciation, phylogeography and population genetics. However, as genomics technology continues to improve, researchers will continue to gather ever-larger datasets, and intra-locus recombination will become an increasingly important issue. I believe that the future of the field lies in new sequencing methods that allow researchers to sample millions of base pairs from non-model organisms in a single day, allowing for mass-screening of SNPS. With this much data, each SNP site can be treated as an independent locus and the no-recombination assumption can be avoided

altogether (e.g., Novembre *et al.* 2008). It will be interesting to see if the pattern observed in this study, the latitudinal gradient in heterozygosity, holds out with genomic sample sizes. As DNA sequencing costs continue to decrease, and coalescent computing techniques become more advanced, multi-locus techniques will continue to revolutionize evolutionary biology and biogeography and help us better understand the process of speciation.

LITERATURE CITED

Abdulla, M. A., I. Ahmed, A. Assawamakin, J. Bhak, S. K. Brahmachari, G. C. Calacal, A. Chaurasia, C.-H. Chen, J. Chen, Y.-T. Chen, J. Chu, E. M. C. Cutiongco-de la Paz, M. C. A. De Ungria, F. C. Delfin, J. Edo, S. Fuchareon, H. Ghang, T. Gojobori, J. Han, S.-F. Ho, B. P. Hoh, W. Huang, H. Inoko, P. Jha, T. A. Jinam, L. Jin, J. Jung, D. Kangwanpong, J. Kampuansai, G. C. Kennedy, P. Khurana, H.-L. Kim, K. Kim, S. Kim, W.-Y. Kim, K. Kimm, R. Kimura, T. Koike, S. Kulawonganunchai, V. Kumar, P. S. Lai, J.-Y. Lee, S. Lee, E. T. Liu, P. P. Majumder, K. K. Mandapati, S. Marzuki, W. Mitchell, M. Mukerji, K. Naritomi, C. Ngamphiw, N. Niikawa, N. Nishida, B. Oh, S. Oh, J. Ohashi, A. Oka, R. Ong, C. D. Padilla, P. Palittapongarnpim, H. B. Perdigon, M. E. Phipps, E. Png, Y. Sakaki, J. M. Salvador, Y. Sandraling, V. Scaria, M. Seielstad, M. R. Sidek, A. Sinha, M. Srikummool, H. Sudoyo, S. Sugano, H. Suryadi, Y. Suzuki, K. A. Tabbada, A. Tan, K. Tokunaga, S. Tongsima, L. P. Villamor, E. Wang, Y. Wang, H. Wang, J.-Y. Wu, H. Xiao, S. Xu, J. O. Yang, Y. Y. Shugart, H.-S. Yoo, W. Yuan, G. Zhao, B. A. Zilfalil, H. P. A. S. Consortium, and C. Indian Genome Variation. 2009. Mapping Human Genetic Diversity in Asia. Science (Washington D C) 326:1541-1545.

Adest, G. A. 1977. Genetic relationships in the genus *Uma* (Iguanidae). Copeia:47-52.

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. Pp. 707-740 *in* D. J. Futuyma, ed. Annual Review of Ecology and Systematics. Volume 33.

Avise, J. C. 1998. The history and purview of phylogeography: A personal reflection. Molecular Ecology 7:371-379.

Avise, J. C., D. Walker, and G. C. Johns. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. Proceedings of the Royal Society of London Series B Biological Sciences 265:1707-1712.

Becquet, C., and M. Przeworski. 2009. Learning about modes of speciation by computational approaches. Evolution 63:2547-2562.

Beerli, P., and S. V. Edwards. 2002. When did Neanderthals and modern humans diverge? Evolutionary Anthropology 11:60-63.

Bentancourt, J., T. Van Devender, and P. Martin. 1990a. Introduction. Pp. 2-11 *in* J. Bentancourt, T. Van Devender, and P. Martin, eds. Packrat middens: the last 40,000 years of biotic change. University of Arizona Press, Tucson.

Bentancourt, J., T. Van Devender, and P. Martin. 1990b. Synthesis and Prospectus. Pp. 435-447 *in* J. Bentancourt, T. Van Devender, and P. Martin, eds. Packrat middens: the last 40,000 years of biotic change. University of Arizona Press, Tucson.

Brito, P. H., and S. V. Edwards. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica (Dordrecht) 135:439-455.

Brown, D. E. 1994. Biotic Communities: Southwestern United States and Northwestern Mexico. University of Utah Press, Salt Lake City, UT.

Carothers, J. H. 1986. An experimental confirmation of morphological adaptation: toe fringes in the sand-dwelling lizard *Uma scoparia.* Evolution 40:871-874.

Carpenter, C. C. 1963. Patterns of Behavior in Three Forms of the Fringe-Toed Lizards (*Uma*-Iguanidae). Copeia 1963:406-412.

Chen, X., C. W. Barrows, and B.-L. Li. 2006. Is the Coachella Valley fringe-toed lizard (*Uma inornata*) on the edge of extinction at Thousand Palms Preserve in California? Southwestern Naturalist 51:28-34.

Cole, K. L. 1986. The lower Colorado River valley, USA: a Pleistocene desert. Quaternary Research (Orlando) 25:392-400.

Doughty, P., B. Sinervo, and G. M. Burghardt. 1994. Sex-biased dispersal in a polygynous lizard, *Uta stansburiana*. Animal Behaviour 47:227-229.

Douglas, M. E., M. R. Douglas, G. W. Schuett, and L. W. Porras. 2006. Evolution of rattlesnakes (Viperidae; *Crotalus*) in the warm deserts of western North America shaped by Neogene vicariance and Quaternary climate change. Molecular Ecology 15:3353-3374.

Edwards, S. V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63:1-19.

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

Elders, W. A., R. W. Rex, T. Meidav, P. T. Robinson, and S. Biehler. 1972. Crustal Spreading in Southern California. Science 178:15-24.

Enzel, Y., S. Wells, and N. Lancaster. 2003. Late Pleistocene lakes along the Mojave River, southeast California. Pp. 61-78 *in* Y. Enzel, S. Wells, and N. Lancaster, eds. Paleoenvironments and paleohydrology of the Mojave and southern Great Basin deserts.

Fay, J. C., and C.-I. Wu. 1999. A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. Molecular Biology and Evolution 16:1003-1005.

Fehlberg, S. D., and T. A. Ranker. 2009. Evolutionary history and phylogeography of *Encelia farinosa* (Asteraceae) from the Sonoran, Mojave, and Peninsular Deserts. Molecular Phylogenetics and Evolution 50:326-335.

Heifetz, W. 1941. A Review of the Lizards of the Genus *Uma*. Copeia 1941:99-111.

Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. Nature (London) 405:907-913.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnean Society 58:247-276.

Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. Philosophical Transactions of the Royal Society of London B Biological Sciences 359:183-195.

Hey, J. 1997. Mitochondrial and nuclear genes present conflicting portraits of human origins. Molecular Biology and Evolution 14:166-172.

Hey, J., and E. Harris. 1999. Population bottlenecks and patterns of human polymorphism. Molecular Biology and Evolution 16:1423-1426.

Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167:747-760.

Hey, J., Y.-J. Won, A. Sivasundar, R. Nielsen, and J. A. Markert. 2004. Using nuclear haplotypes with microsatellites to study gene flow between recently separated Cichlid species. Molecular Ecology 13:909-919.

Hudson, R. R., and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147-164.

Jennings, W. B., and S. V. Edwards. 2005. Speciational history of Australian grass finches (*Poephila*) inferred from thirty gene trees. Evolution 59:2033-2047.

Jezkova, T., J. R. Jaeger, Z. L. Marshall, and B. R. Riddle. 2009. Pleistocene impacts on the phylogeography of the desert pocket mouse (*Chaetodipus penicillatus*). Journal of Mammalogy 90:306-320.

Klicka, J., and R. M. Zink. 1997. The importance of recent ice ages in speciation: A failed paradigm. Science (Washington D C) 277:1666-1669.

Kumar, S., and S. Subramanian. 2002. Mutation rates in mammalian genomes. Proceedings of the National Academy of Sciences of the United States of America 99:803-808.

Lancaster, N., and V. Tehakerian. 2003. Late Quaternary eolian dynamics, Mojave Desert, California. Pp. 231-249 *in* Y. Enzel, S. Wells, and N. Lancaster, eds.

Leaché, A. D. 2009. Species tree discordance traces to phylogeographic clade boundaries in North American fence lizards (*Sceloporus*). Systematic Biology 58:547-559.

Leaché, A. D., and J. A. McGuire. 2006. Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: Evidence for a misleading mitochondrial gene tree. Molecular Phylogenetics and Evolution 39:628-644.

Leavitt, D. H., R. L. Bezy, K. A. Crandall, and J. W. Sites, Jr. 2007. Multi-locus DNA sequence data reveal a history of deep cryptic vicariance and habitat-driven convergence in the desert night lizard *Xantusia vigilis* species complex (Squamata : Xantusiidae). Molecular Ecology 16:4455-4481.

Lee, J. Y., and S. V. Edwards. 2008. Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). Evolution 62:3117-3134.

Lewis, P. O. 1998. A genetic algorithm for maximum-likelihood phylogeny inference using nucleotide sequence data. Molecular Biology and Evolution 15:277-283.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics (Oxford) 25:1451-1452.

Luke, C. 1986. Convergent evolution of lizard toe fringes. Biological Journal of the Linnean Society 27:1-16.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual.

Mayhew, W. 1964. Taxonomic Status of California Populations of the Lizard Genus *Uma*. Herpetologica 20:170-183.

McKibben, M. A. 2008. Introduction to the Salton Sea Rift. Pp. 91-95 *in* R. E. Reynolds, ed. Trough to Trough: The Colorado River and the Salton Sea. California State University, Desert Studies Consortium and LSA Associates, Inc., Zzyzx, CA.

Meisling, K.E., and R.J. Weldon. 1989. Late Cenozoic tectonics of the northwestern San Bernardino Mountains, southern California. Geological Society of America Bulletin 101:106-128.

Miller, A., and R. Stebbins. 1964. The Lives of Desert Animals in Joshua Tree National Monument. University of California Press, Berkeley, California.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. Evolution 49:718-726.

Muhs, D. R., R. L. Reynolds, J. Been, and G. Skipp. 2003. Eolian sand transport pathways in the southwestern United States: importance of the Colorado River and local sources. Quaternary International 104:3-18.

Murphy, R. W., T. L. Trepanier, and D. J. Morafka. 2006. Conservation genetics, evolution and distinct population segments of the Mojave fringe-toed lizard, *Uma scoparia*. Journal of Arid Environments 67:226-247.

Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. Genetics 158:885-896.

Norris, K. S. 1958. The evolution and systematics of the iguanid genus *Uma* and its relation to the evolution of other North American desert reptiles. American Museum of Natural History Bulletin 114:247-326.

Norris, R., and R. Webb. 1990. Geology of California: Second Edition. John Wiley & Sons, New York.

Novembre, J., T. Johnson, K. Bryc, Z. Kutalik, A. R. Boyko, A. Auton, A. Indap, K. S. King, S. Bergmann, M. R. Nelson, M. Stephens, and C. D. Bustamante. 2008. Genes mirror geography within Europe. Nature (London) 456:98.

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

Reilly, S. 2009. Estimating gene flow between black salamander (*Aneides flavipunctatus*) populations: a multi-locus coalescent approach. Biological Sciences. Humboldt State University, Arcata, California.

Rosenblum, E. B., N. M. Belfiore, and C. Moritz. 2007. Anonymous nuclear markers for the eastern fence lizard, *Sceloporus undulatus*. Molecular Ecology Notes 7:113-116.

Schuster, S. C., W. Miller, A. Ratan, L. P. Tomsho, B. Giardine, L. R. Kasson, R. S. Harris, D. C. Petersen, F. Zhao, J. Qi, C. Alkan, J. M. Kidd, Y. Sun, D. I. Drautz, P. Bouffard, D. M. Muzny, J. G. Reid, L. V. Nazareth, Q. Wang, R. Burhans, C. Riemer, N. E. Wittekindt, P. Moorjani, E. A. Tindall, C. G. Danko, W. S. Teo, A. M. Buboltz, Z. Zhang, Q. Ma, A. Oosthuysen, A. W. Steenkamp, H. Oostuisen, P. Venter, J. Gajewski, Y. Zhang, B. F. Pugh, K. D. Makova, A. Nekrutenko, E. R. Mardis, N. Patterson, T. H. Pringle, F. Chiaromonte, J. C. Mullikin, E. E. Eichler, R. C. Hardison, R. A. Gibbs, T. T. Harkins, and V. M. Hayes. 2010. Complete Khoisan and Bantu genomes from southern Africa. Nature 463:943-947.

Slade, R. W., C. Moritz, A. R. Hoelzel, and H. R. Burton. 1998. Molecular population genetics of the southern elephant seal *Mirounga leonina*. Genetics 149:1945-1957.

Stebbins, R. 2003. Western Reptiles and Amphibians. Peterson Field Guide Series, New York.

Stebbins, R. C. 1944. Some Aspects of the Ecology of the Iguanid Genus *Uma*. Ecological Monographs 14:311-332.

Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. American Journal of Human Genetics 76:449-462.

Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68:978-989.

Strasburg, J., and L. Rieseberg. 2010. How Robust Are "Isolation with Migration" Analyses to Violations of the IM Model? A Simulation Study. Mol Biol Evol 27:297-310.

Thompson, R. S., and K. H. Anderson. 2000. Biomes of western North America at 18,000, 6000 and 0 14C yr BP reconstructed from pollen and packrat midden data. Journal of Biogeography 27:555-584.

Thomson, R. C., A. M. Shedlock, S. V. Edwards, and H. B. Shaffer. 2008. Developing markers for multilocus phylogenetics in non-model organisms: A test case with turtles. Molecular Phylogenetics and Evolution 49:514-525.

Thorne, R. F. 1986. A historical sketch of the vegetation of the Mojave and Colorado USA Deserts of the America Southwest. Annals of the Missouri Botanical Garden 73:642-651.

Trepanier, T. L., and R. W. Murphy. 2001. The Coachella Valley fringe-toed lizard (*Uma inornata):* Genetic diversity and phylogenetic relationships of an endangered species. Molecular Phylogenetics and Evolution 18:327-334.

Wakeley, J. 1996. Distinguishing migration from isolation using the variance of pairwise differences. Theoretical Population Biology 49:369-386.

Wakeley, J., R. Nielsen, K. Ardlie, S. N. Liu-Cordero, and E. S. Lander. 2000. The discovery of single nucleotide polymorphisms and inferences of human demographic history. American Journal of Human Genetics 67:210.

Wiens, J., C. Kuczynski, S. Arif, and T. Reeder. 2010. Phylogenetic relationships of phrynosomatid lizards based on nuclear and mitochondrial data, and a revisited phylogeny for *Sceloporus*. Pp. 150-161. Molecular Phylogenetics and Evolution.

Wright, S. 1931. Evolution in Mendelian populations. Pp. 97-159. Genetics.

Zhao, Z., L. Jin, Y.-X. Fu, M. Ramsay, T. Jenkins, E. Leskinen, P. Pamilo, M. Trexler, L. Patthy, L. B. Jorde, S. Ramos-Onsins, N. Yu, and W.-H. Li. 2000. Worldwide DNA sequence variation in a 10-kilobase noncoding region on human chromosome 22. Proceedings of the National Academy of Sciences of the United States of America 97:11354-11358.

Zhivotovsky, L. A. 2001. Estimating divergence time with the use of microsatellite genetic distances: Impacts of population growth and gene flow. Molecular Biology and Evolution 18:700-709.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Thesis, University of Texas at Austin.

Tables and Figures

Table 1. Sampling localities with GPS Coordinates in WGS-84 decimal degrees. *Sample courtesy of R. Murphy at the Royal Ontario Museum, Toronto, Canada.

Locality	**Species**	**Population Designation**	**Latitude**	**Longitude**
Ibex Dunes	*U. scoparia*	Amargosa	35°41.607	116°21.958
Dumont Dunes	*U. scoparia*	Amargosa	35°40.107	116°14.505
Coyote Holes	*U. scoparia*	Amargosa	35°38.650	115°57.264
Bitter Springs	*U. scoparia*	Mojave	35°13.904	116°26.119
Red Pass Dune	*U. scoparia*	Mojave	35°14.953	116°21.807
West Red Pass	*U. scoparia*	Mojave	35°15.460	116°22.597
The Whale	*U. scoparia*	Mojave	35°13.201	116°28.158
Coyote Dry Lake	*U. scoparia*	Mojave	35°06.136	116°45.423
Cronese Lakes	*U. scoparia*	Mojave	35°08.625	116°17.561
Afton Canyon	*U. scoparia*	Mojave	35°02.604	116°25.110
Rasor Road	*U. scoparia*	Mojave	35°04.968	116°09.506
Kelso Dunes	*U. scoparia*	Mojave	34°53.484	115°42.881
Lenwood Road	*U. scoparia*	Mojave	34°53.438	117°07.440
Amboy Crater	*U. scoparia*	Colorado	34°33.855	115°48.019
Cadiz Dunes	*U. scoparia*	Colorado	34°23.592	115°25.561
Dale Dry Lake	*U. scoparia*	Colorado	34°06.012	115°37.601
Pinto Basin	*U. scoparia*	Colorado	33°54.758	115°49.278
Rice Valley	*U. scoparia*	Colorado	34°03.016	114°51.815
Bouse Wash, AZ	*U. scoparia*	Colorado	34°04.454	114°16.039
Palen Dry Lake	*U. scoparia*	Colorado	33°48.139	115°11.094
Thousand Palms*	*U. notata inornata*	Notata	n/a	n/a
Borrego Springs	*U. n. notata*	Notata	33°14.846	116°17.744
Algodones	*U. n. notata*	Notata	32°59.528	115°07.934
Mohawk Dunes	*U. n. rufopunctata*	Notata	32°41.843	113°48.571

Table 2. Polymerase Chain Reaction (PCR) primer sequences, annealing temperatures, and citations. A = annealing temperature in degrees Celcius.

Locus	Forward Primer	Reverse Primer	A	Citation
BDNF	GAC CAT CCT TTT CCT (G/T)AC TAT GGT TAT TTC ATA CTT	CTA TCT TCC CCT TTT AAT GGT CAG TGT ACA AAC	61	Leache 2009
RAG-1	CAA AGT (A/G)A GAT CAC TTG AGA AGC	ACT TG(C/T) AGC TTG AGT TCT CTT AG(A/G) CG	55	Leache 2009
PNN	ACA GGT AAT CAG CAC AAT GAY GTA GA	TCT YYT GCC TGA YCG ACT ACT YTC TGA	57	Leache 2009
R35	GAC TGT GGA YGA YCT GAT CAG TGT GGT GCC	GCC AAA ATG AGS GAG AAR CGC TTC TGA GC	66.5	Leache 2009
Sun07	TTT CTG TCA CGA TGA AAA TTG TAA ACT A	TAA ACA CAA TGC TCA CAT TAG GAA AAA T	61	Rosenblum et al. 2007
Sun08	CTC TTG AAG TTC ACA GGG TTT TCT TAG	TAG CCT AGC TTC CTT ACA GTT TGA TAC	63	Rosenblum et al. 2007
Sun10	CAG AAA GTA AAT CCA CTG TAG CTA GGA	CTA ATA ATG GCA TAG CAA GGA GTG TAG	61	Rosenblum et al. 2007
Sun12	TAC AGA GTC TCC TCT TGA CTG GAT ATT	TTG GTA CAC TAA CTC AAG CAA ACC T	54	Rosenblum et al. 2007
Sun18	ATG ACA GAA GTT GTG GTT CAA CAG TAT	AGT GAG ATA GAA GTG GCT TTC TGA TTA C	61	Rosenblum et al. 2007
Sun28	AAT CTT ATT TCT GCA GTT GAT GTA CTT T	ATA AAT GCA ATG CCA CAA ATA TAA TAA G	57.5	Rosenblum et al. 2007
Uma03	CGC ATG AGA ATT CTG TGT TA	TGC TAA TGC TGA TGA AAA TG	59	This study
Uma05	TTG TTC CCA TAG CTG AAA CT	TTT GGT AAT GAA TCC CAC TC	50	This study
Uma06	GCT CTT ACC CTC TGT TTG AA	AGG TGA CAG ATG GAG CTA AA	50	This study
Uma07	ACA CGA TCT GAG GAG ATG AC	CTT CTC ACC CTC TTG AC AAA	57	This study
Uma08	AGG TGT TTT GAA CTG CAA CT	ACT TTC CCA GCA CAT AAA AA	59	This study

Table 3. Population sampling: number of alleles per population per locus.

	U. notata	Colorado	Mojave	Amargosa	Total Alleles
BDNF	20	20	20	16	**76**
RAG-1	20	22	22	10	**74**
PNN	12	20	16	18	**66**
R35	18	22	20	18	**78**
Sun07	20	22	18	16	**76**
Sun08	2	14	2	6	**24**
Sun10	20	20	18	16	**74**
Sun12	18	22	16	18	**74**
Sun18	10	0	12	4	**26**
Sun28	20	22	18	14	**74**
Uma03	20	46	34	28	**128**
Uma05	18	14	16	12	**60**
Uma06	20	46	68	30	**164**
Uma07	18	52	54	32	**156**
Uma08	28	52	54	38	**172**
Total Alleles	**264**	**394**	**388**	**276**	

Table 4. Analysis of heterozygosity by population, excluding Locus Uma07. The Colorado River, Mojave River, and Amargosa River population all belong to *U. scoparia*. N= number of sequences (before PHASE), I = number of heterozygous indels in analysis, SNPs = total heterozygous single nucleotide polymorphisms (aka ambiguities), L = number of base pairs, before removal of gaps.

LOCUS	*U. notata*			Colorado River			Mojave River			Amargosa River				
	N	I	SNPs	N	I	SNPs	N	I	SNPs	N	I	SNPs	L	Total SNPS per locus
BDNF	10	0	11	10	0	0	10	0	0	8	0	0	628	11
RAG-1	10	0	4	11	0	8	11	0	2	5	0	1	705	15
PNN	6	0	8	10	0	6	8	0	2	9	0	3	600	19
R35	9	0	25	11	0	6	10	0	6	9	0	0	577	37
Sun07	10	0	8	11	0	2	9	0	0	8	0	0	546	10
Sun08	1	0	0	7	0	1	1	0	1	3	0	1	619	3
Sun10	10	5	27	10	0	9	9	0	4	8	0	9	651	49
Sun12	9	0	21	11	0	0	8	0	2	9	0	4	419	27
Sun18	5	0	0	0	0	0	6	0	0	2	0	0	379	0
Sun28	10	0	23	11	3	18	9	0	0	7	0	2	571	43
Uma03	10	0	5	23	0	2	17	0	0	14	0	0	523	7
Uma05	9	0	25	7	0	5	8	0	6	6	0	1	298	37
Uma06	10	0	22	23	0	14	34	0	4	15	0	16	146	56
Uma08	14	0	21	26	0	3	27	0	5	19	0	2	384	31
TOTAL	123	5	**200**	171	3	**74**	167	0	**32**	122	0	**39**	**7046**	345

	Total SNPs	Total BP	SNPS/BP	Frequency
U. notata	200	60836	0.00329	0.33%
Colorado River	74	81578	0.00091	0.09%
Mojave River	35	73369	0.00048	0.05%
Amargosa River	36	57697	0.00062	0.07%

Table 5. Data Characteristics. L = length after PHASE, after trimming some gaps, but before trimming recombination; bolded numbers have been trimmed. S = segregating sites, %S = percentage of total base pairs that are segregating, D = results of Tajima's D test, π = nucleotide diversity, R_m = minimum number of recombination events. F_{ns} = # fixed differences between *U. notata* and *U. scoparia*, F_{am} = # fixed differences between Amargosa and Mojave Populations.
* Because this dataset includes PHASED haplotypes with probabilities of 50% or greater, most of this recombination is likely an artifact of incorrect haplotype estimation.

Locus	**L**	**S**	**%S**	**D**	**Stat sig?**	**π**	**π *U. notata***	**π *U. scoparia***	**Mut. Model**	**R_m***	**F_{ns}**	**F_{am}**
BDNF	628	4	0.64%	0.45971	no	0.00159	0.00173	0	HKY+G	0	0	0
RAG-1	**703**	10	1.42%	-0.7639	no	0.00208	0.00057	0.00088	GTR	0	2	0
PNN	600	13	2.17%	-1.72647	no	0.00181	0.00343	0.00121	HKY	0	0	0
R35	577	13	2.25%	-0.42659	no	0.00387	0.00565	0.00155	HKY	5	0	0
Sun07	546	11	2.01%	-1.32457	no	0.00211	0.00212	0.00013	TrN	2	2	0
Sun08	**613**	7	1.14%	-1.15004	no	0.00193	0	0.00069	TrN + G	0	5	0
Sun10	**643**	22	3.42%	-1.16492	no	0.00455	0.00665	0.00157	TrN + G	4	1	0
Sun12	419	9	2.15%	-1.17765	no	0.00192	0.00564	0.00047	TrN + G	3	0	0
Sun18	379	7	1.85%	0.92327	no	0.0063	0.00375	0.00062	HKY	0	2	0
Sun28	**567**	14	2.47%	-0.31408	no	0.00484	0.00586	0.00154	TrN + G	5	0	0
Uma03	**522**	6	1.15%	-0.82323	no	0.00131	0.00241	0.00007	F81	0	0	0
Uma05	298	11	3.69%	-0.08262	no	0.00768	0.00774	0.00252	GTR + G	4	0	0
Uma06	146	12	8.22%	-0.35149	no	0.01244	0.01514	0.006	GTR + G	4	2	0
Uma07	**471**	13	2.76%	-0.09488	no	0.00473	0.00487	0.0043	HKY+G	6	0	0
Uma08	384	19	4.95%	-2.02247	< 0.05	0.00216	0.00482	0.00064	HKY+G	2	0	0
Total	**7496**	**171**	-	-	-	-	-	-	-	**35**	**14**	**0**
Mean	**500**	**11.4**	2.69%	-	-	**0.00395**	**0.00469**	**0.00148**	-	**2.3**	**0.9**	**0**
Std Dev	**150**	**4.75**	1.89%	-	-	**0.00303**	**0.00366**	**0.00167**	-	**2.2**	**1.4**	**0**

Table 6. Results of the IM analysis.

A) With Recombination

	θ_{notata}	$\theta_{scoparia}$	$\theta_{ancestor}$	t	m_1	m_2
High Point	4.7839	1.0222	0.8859	1.0775	0.0375	0.0075
95% low	3.5572	0.7496	0.1772	0.7325	0.0075	0.0025
95% high	7.1009	1.3493	1.7854	1.5875	0.1475	0.1275

B) Without Recombination

	θ_{notata}	$\theta_{scoparia}$	$\theta_{ancestor}$	t	m_1	m_2
High Point	1.6226	0.4592	0.0102	1.4225	0.1175	0.0325
95% low	1.1124	0.3368	0.0102	0.7475	0.0375	0.0075
95% high	2.3574	0.6838	2.6228	2.0975	0.3875	0.2575

C) With Recombination - Converted Demographic Parameters

	N_{notata}	$N_{scoparia}$	$N_{ancestor}$	t (years)	$2N_1m_1$	$2N_2m_2$
High Point	576,529	123,190	106,764	1,038,835	0.0897	0.003833
95% low	428,694	90,338	21,355	706,215	0.01794	0.001278
95% high	855,761	162,610	215,167	1,530,534	0.35281	0.065165

D) Without Recombination - Converted Demographic Parameters

	N_{notata}	$N_{scoparia}$	$N_{ancestor}$	t (years)	$2N_1m_1$	$2N_2m_2$
High Point	200,979	56,878	1,263	1,409,556	0.09533	0.007462
95% low	137,785	41,717	1,263	740,698	0.02928	0.001799
95% high	291,994	84,697	324,866	2,078,414	0.30252	0.061749

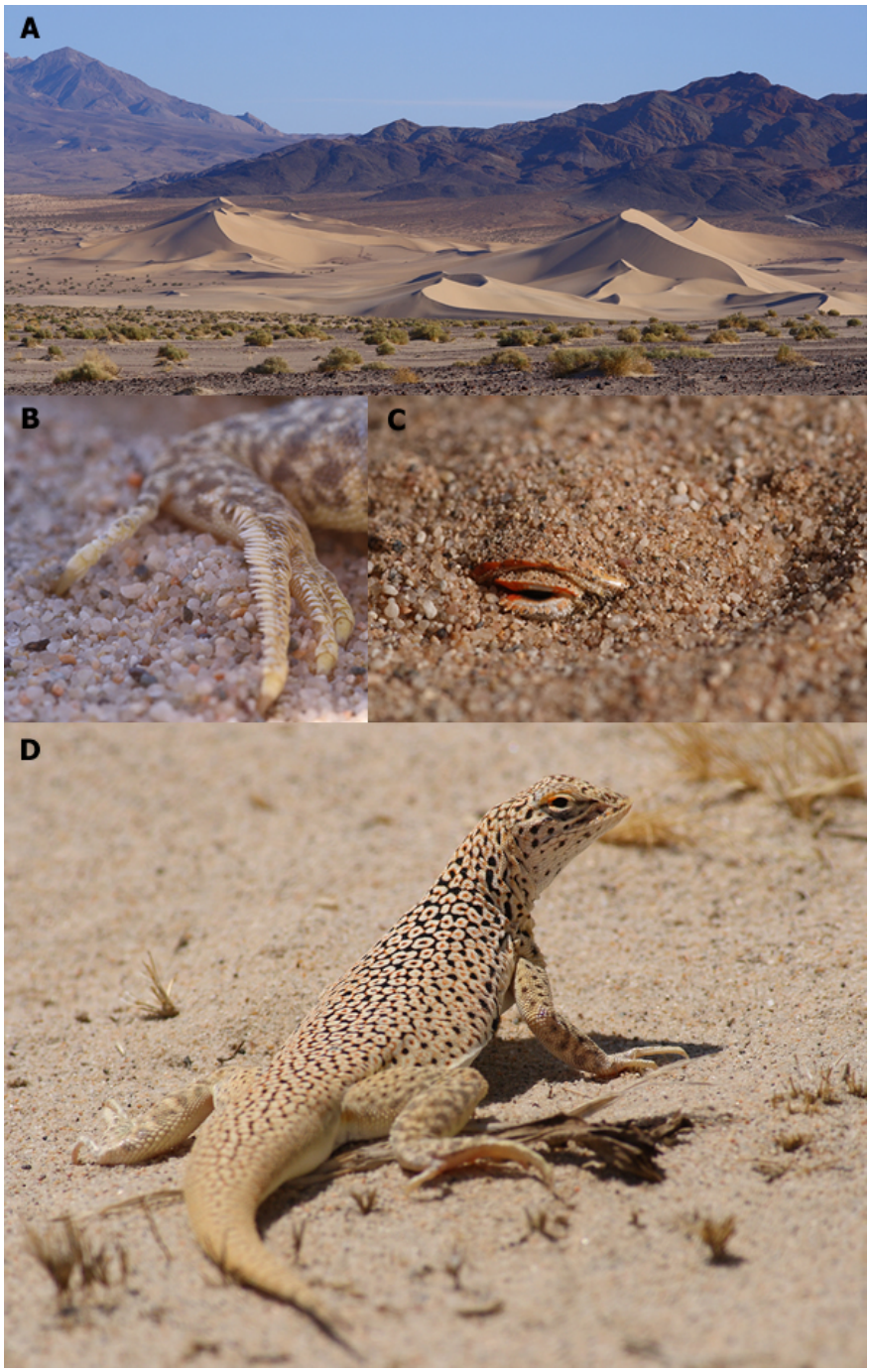

Figure 1. Habitat and adaptations of fringe-toed lizards. A) Ibex Dunes, Death Valley National Park, the northernmost locality where this genus occurs. B) Fringed toes increase traction and locomotion efficiency on loose sand. C) A shovel-shaped snout facilities burial, and D) an ocellated pattern increases crypsis in an exposed habitat. Photographs courtesy of Cameron Rognan.

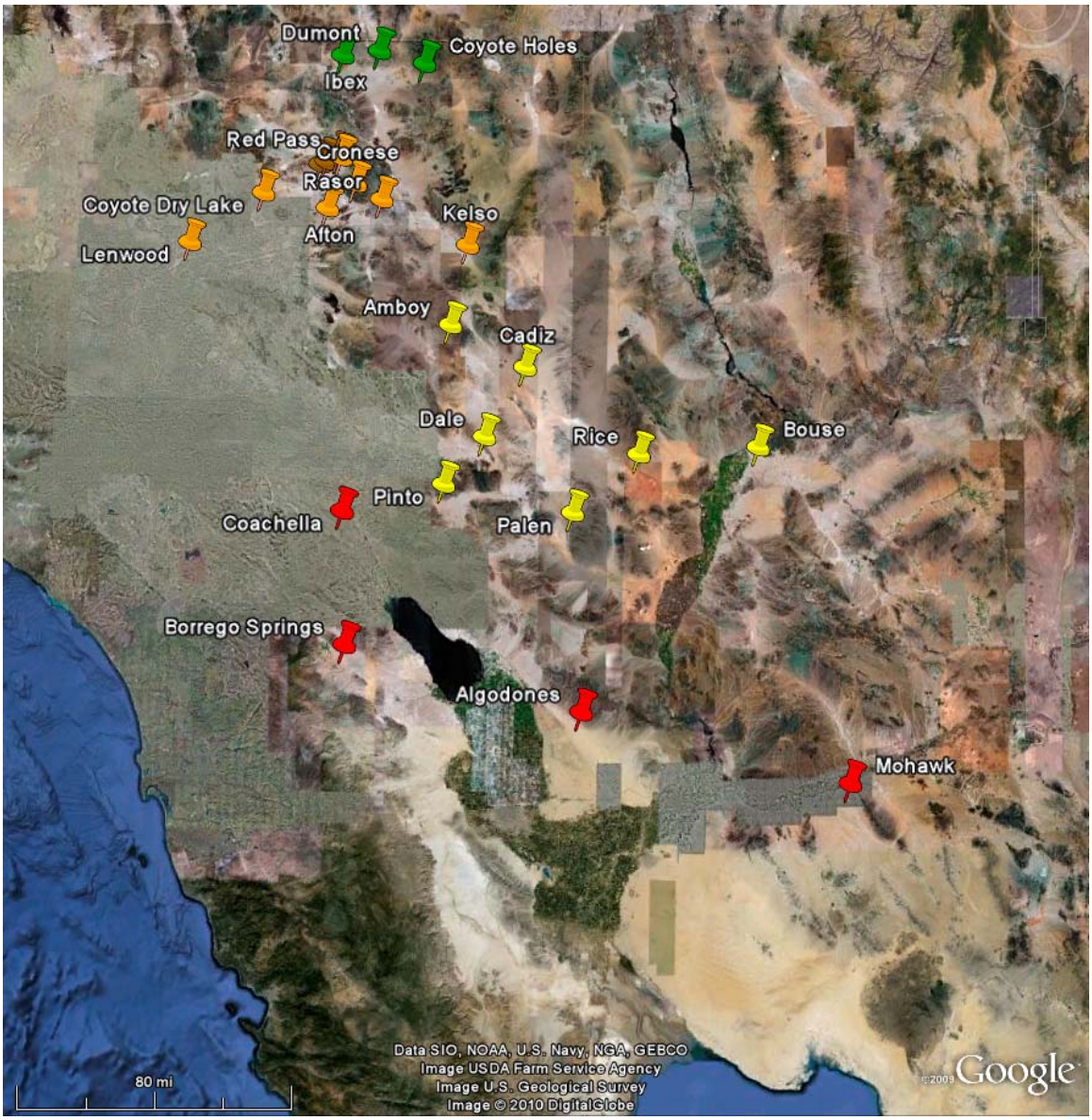

Figure 2. Sampling Localities. Red = *Uma notata* species complex, Yellow = Colorado River *U. scoparia*, Orange = Mojave River *U. scoparia*, Green = Amargosa River *U. scoparia*.

Figure 3. Fixed and unfixed morphological differences between the *Uma notata* species complex and *U. scoparia*. *Uma notata* (left) has ocelli that coalesce to form lines over the shoulders, and it has streaked markings on the throat. The shoulder ocelli of *U. scoparia* (right) do not coalesce into lines, and it usually has crescent-shaped markings on the throat. Photographs by Cameron Rognan and the author.

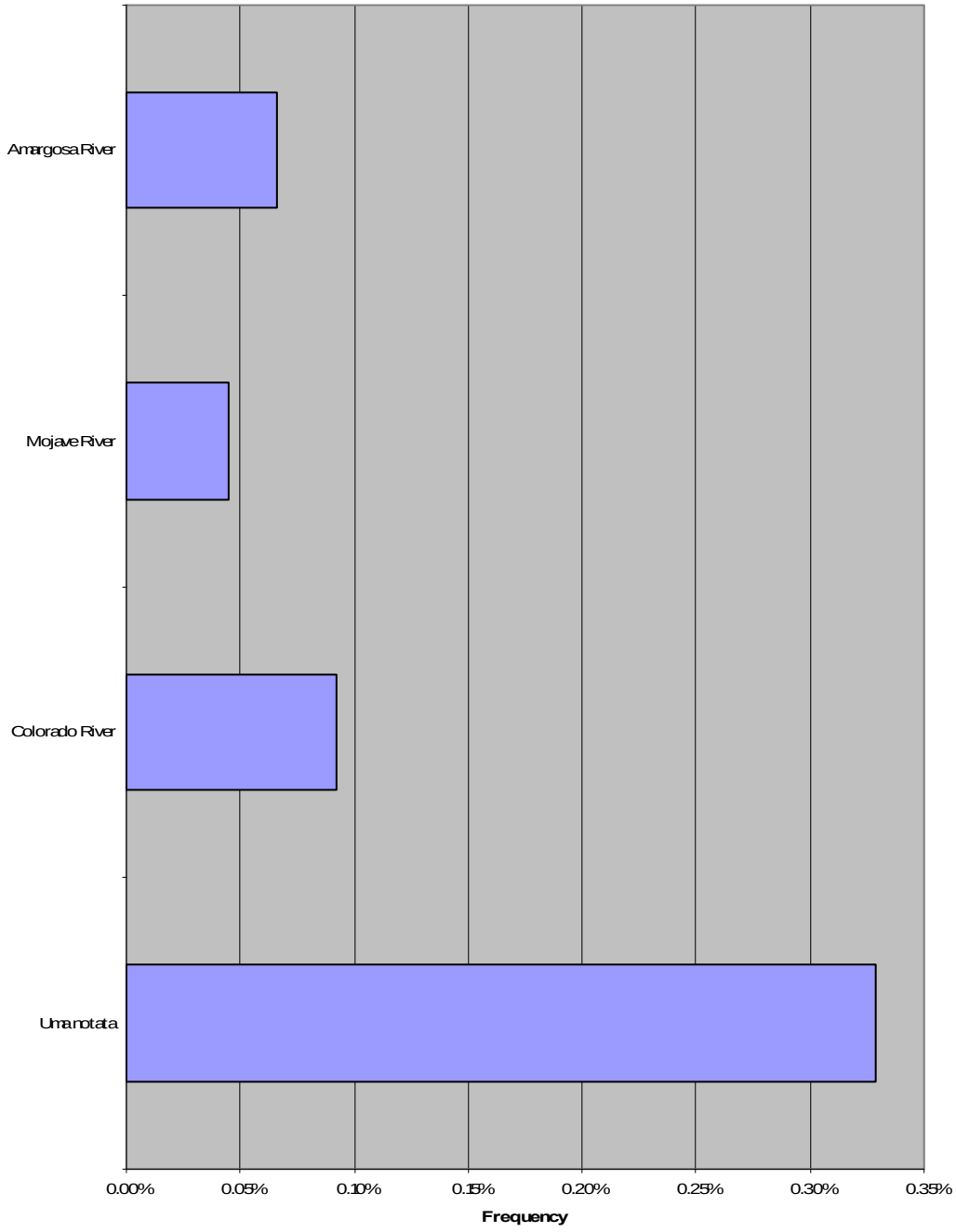

Figure 4. Frequency of heterozygous SNPs by population, excluding Locus Uma07 (see also Table 4).

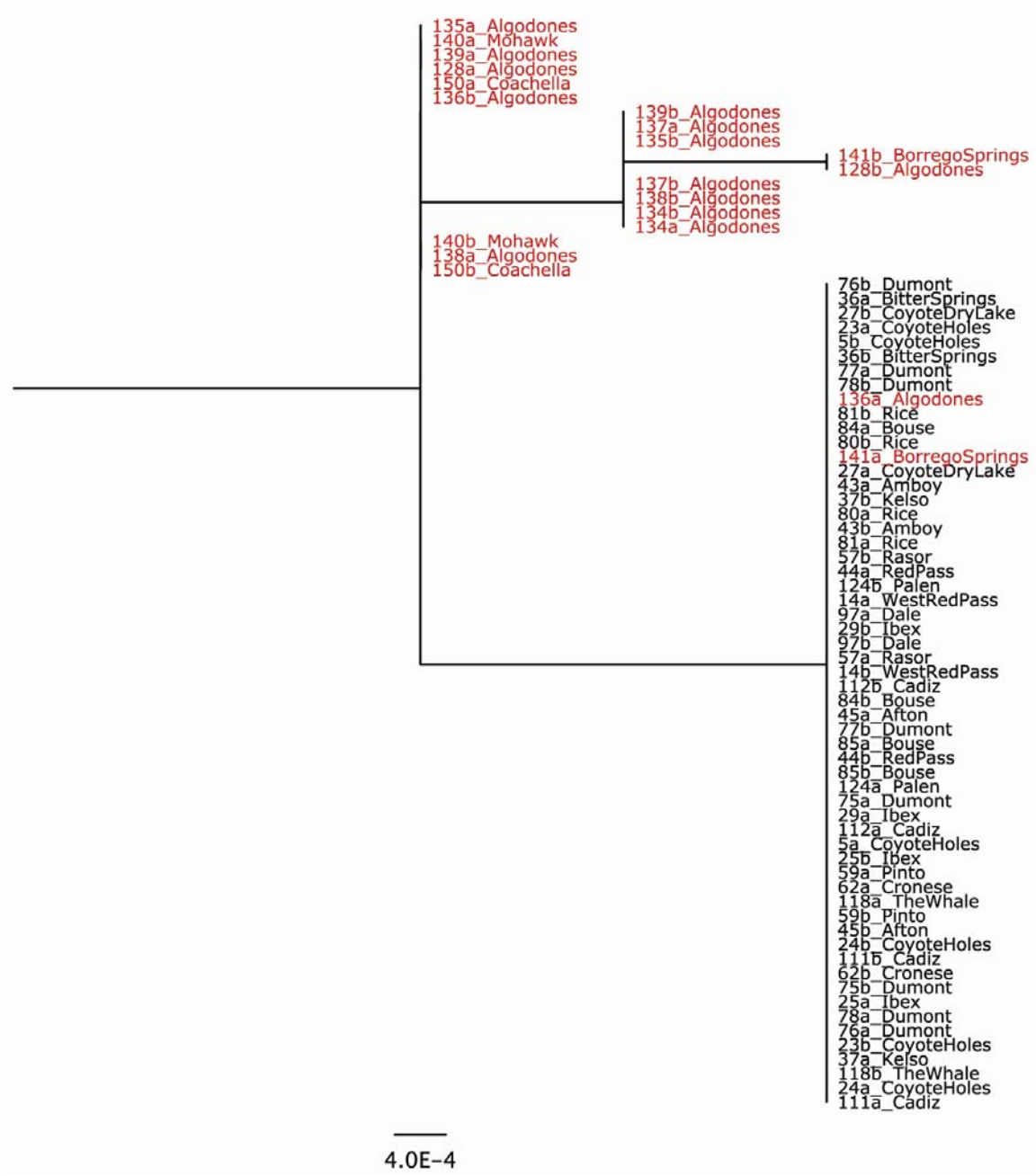

Figure 5. Maximum likelihood gene tree for locus BDNF, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

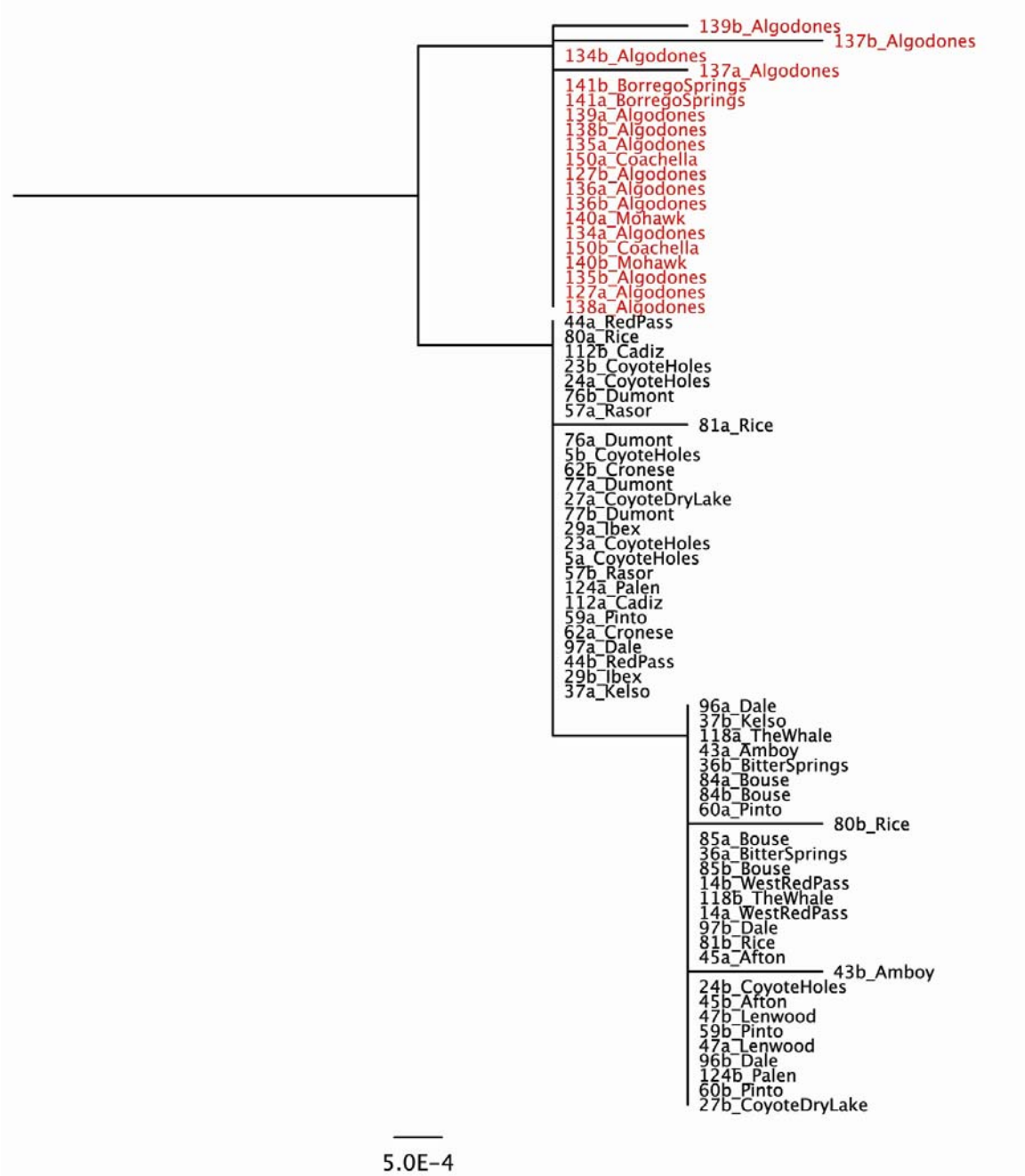

Figure 6. Maximum likelihood gene tree for locus RAG-1, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

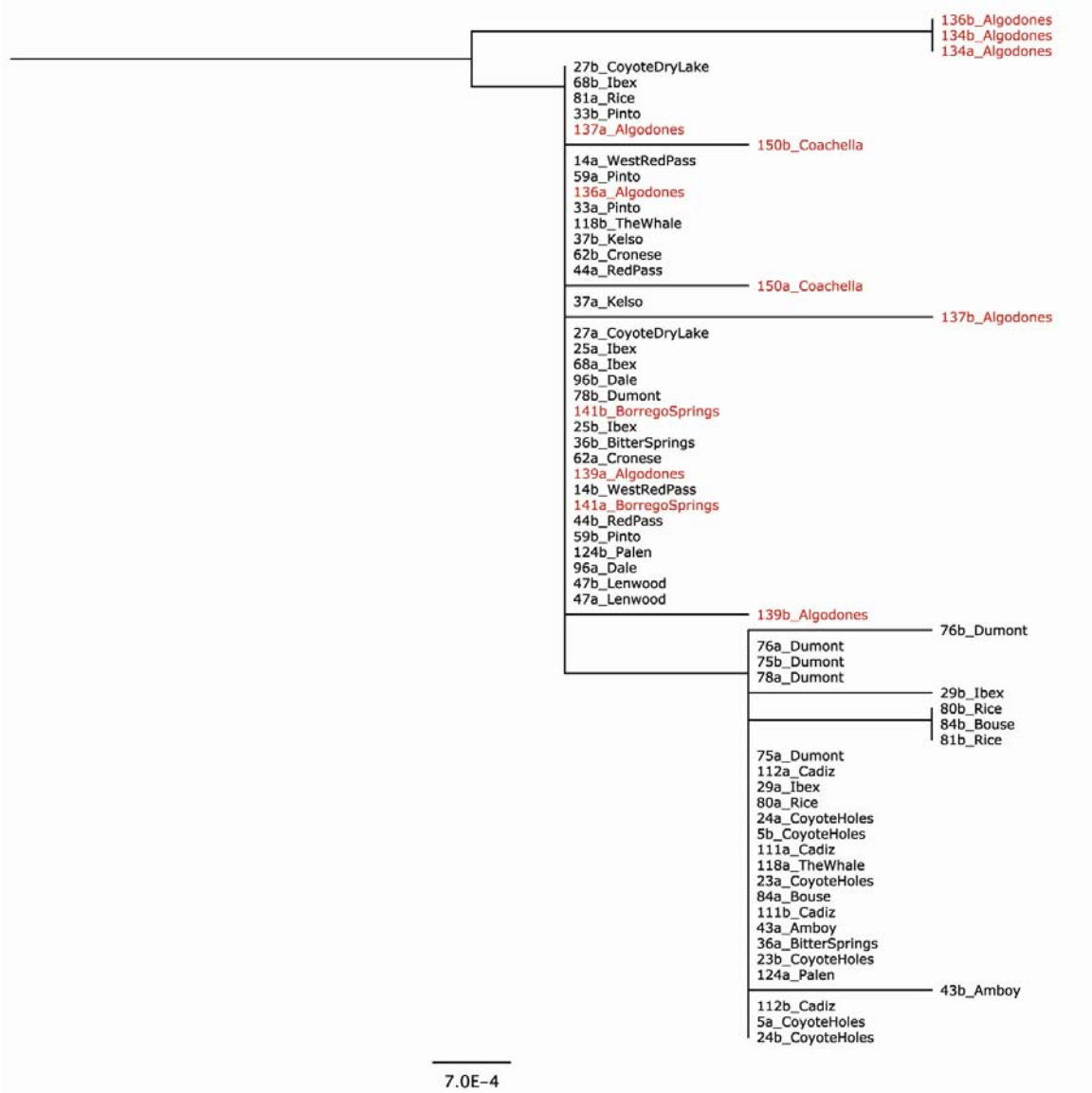

Figure 7. Maximum likelihood gene tree for locus PNN, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

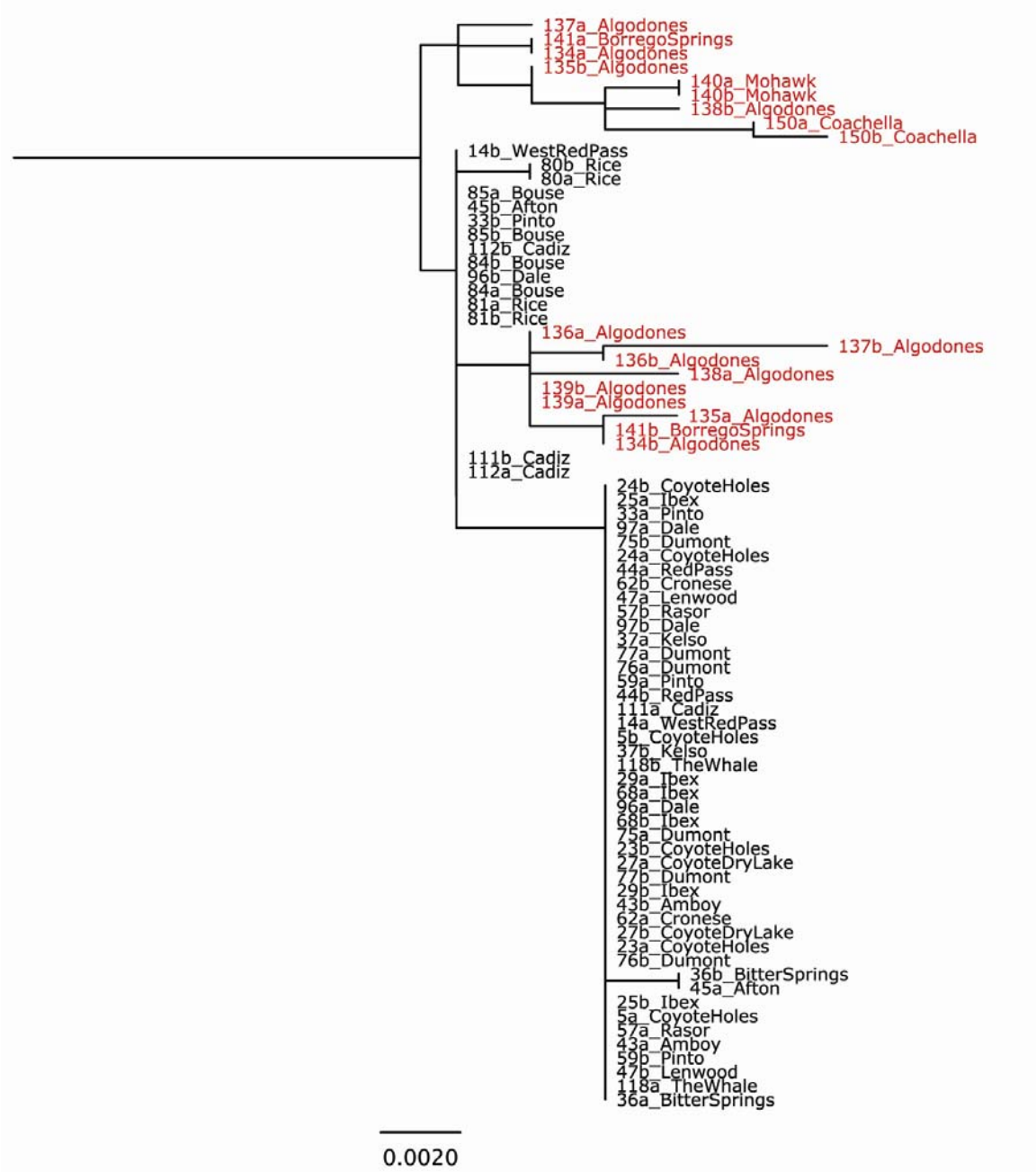

Figure 8. Maximum likelihood gene tree for locus R35, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

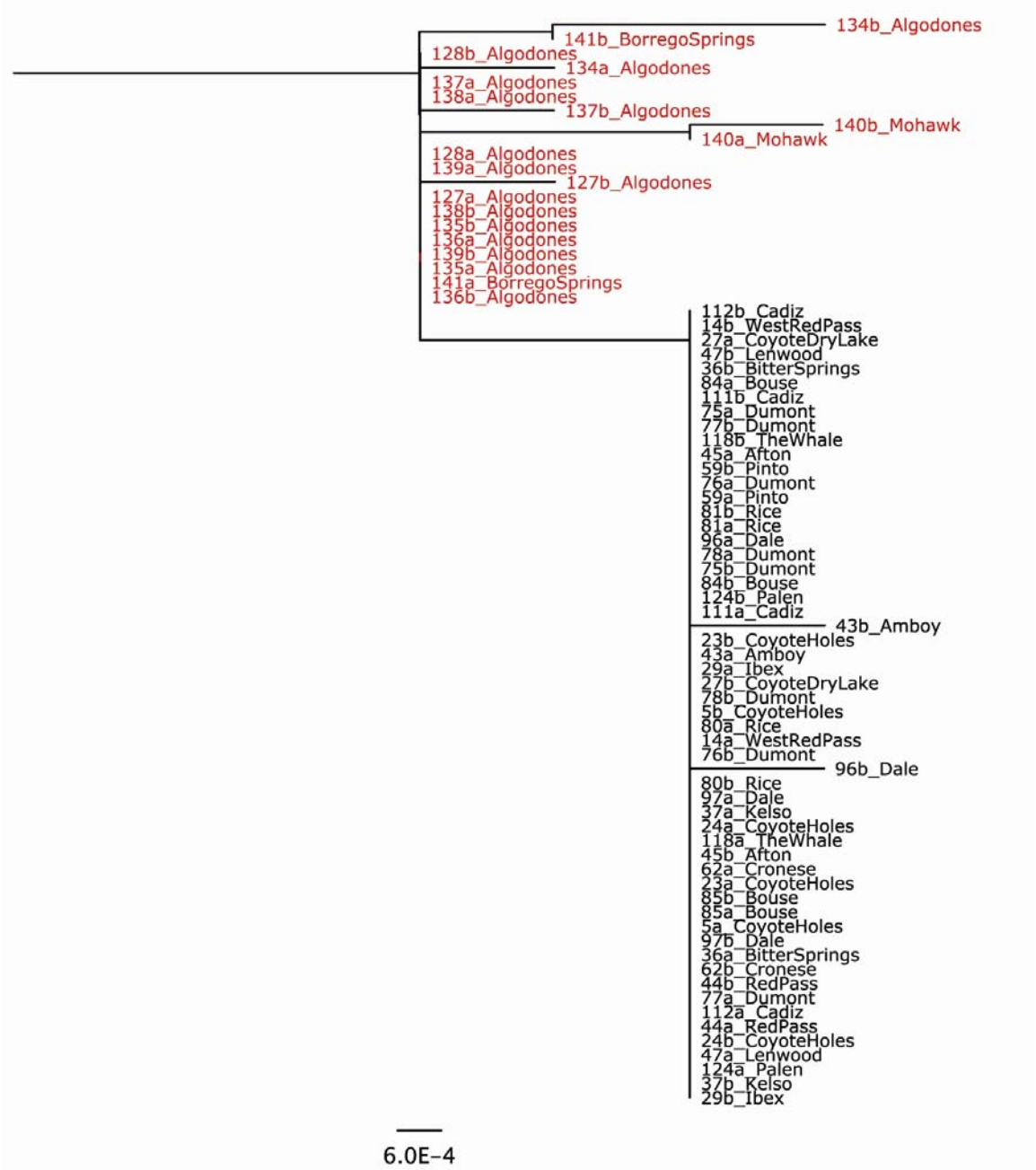

Figure 9. Maximum likelihood gene tree for lócus Sun07, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

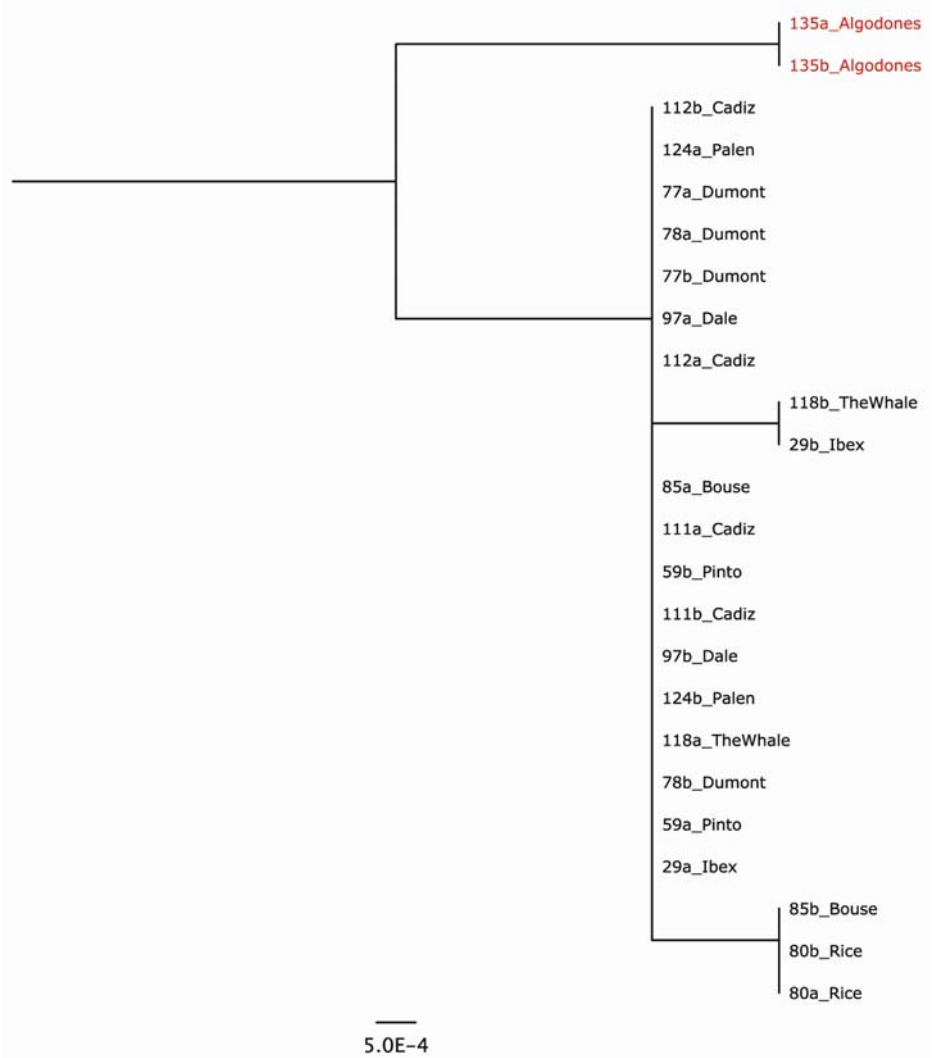

Figure 10. Maximum likelihood gene tree for locus Sun08, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

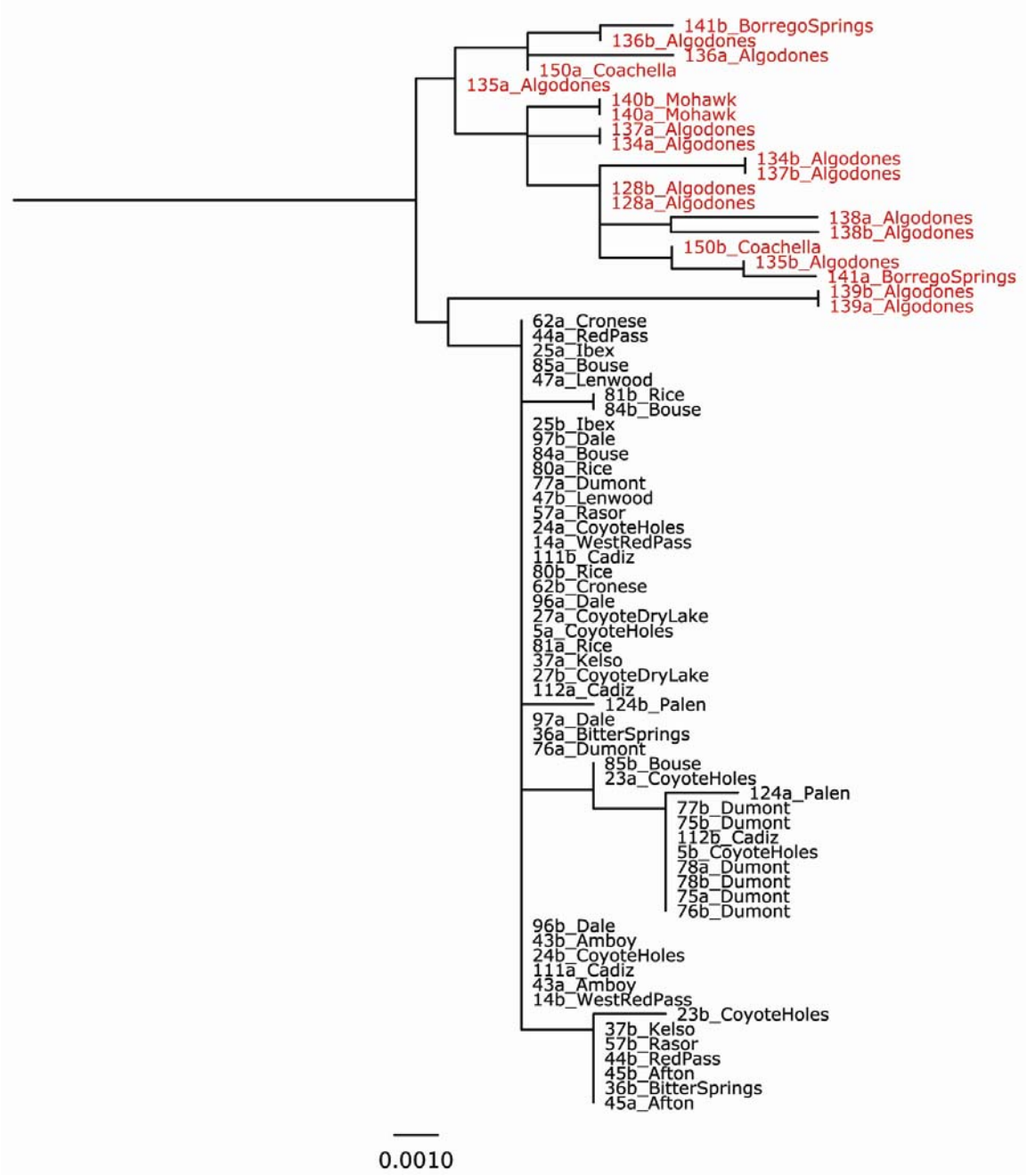

Figure 11. Maximum likelihood gene tree for locus Sun10, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

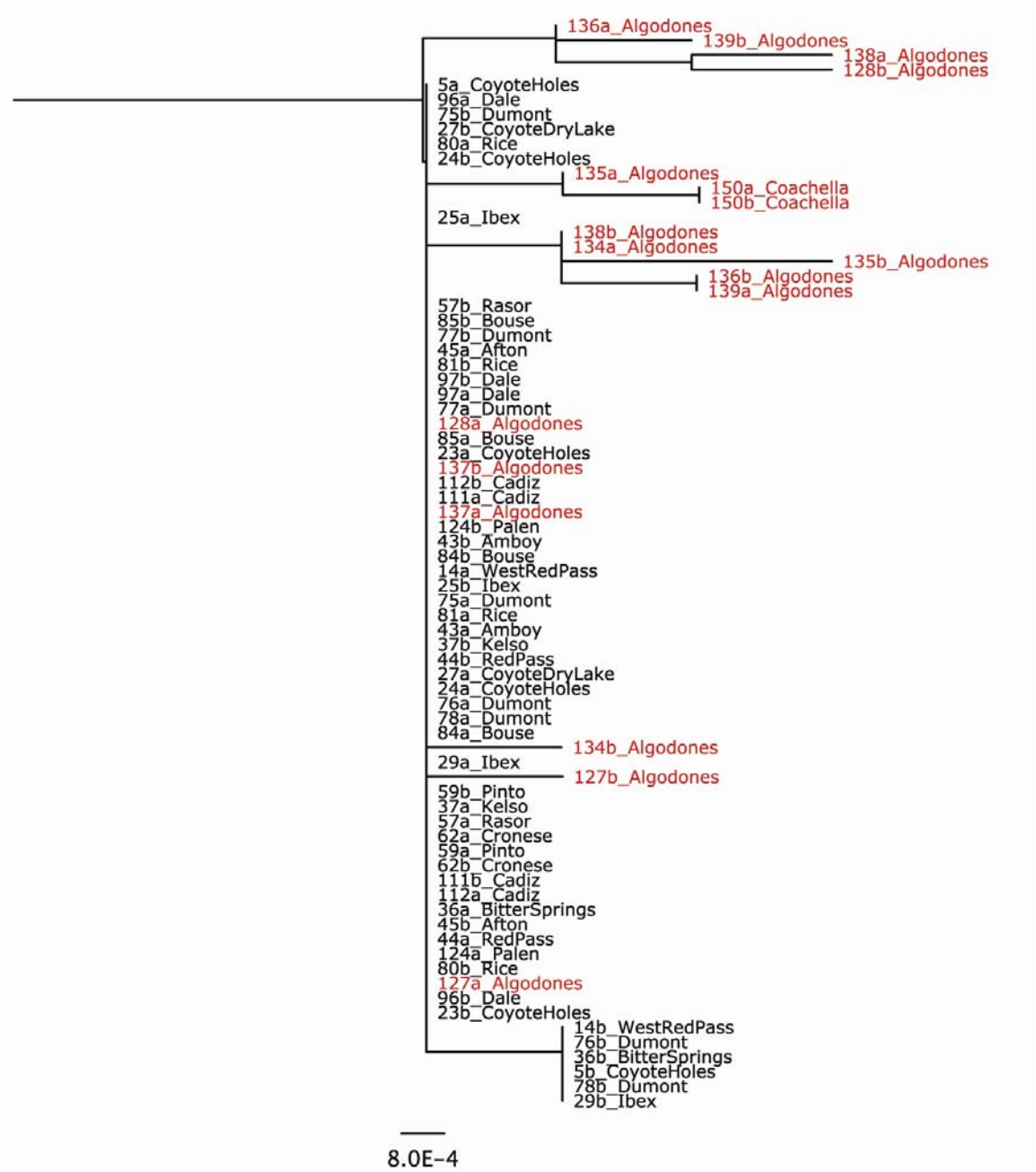

Figure 12. Maximum likelihood gene tree for locus Sun12, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

Figure 13. Maximum likelihood gene tree for locus Sun18, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

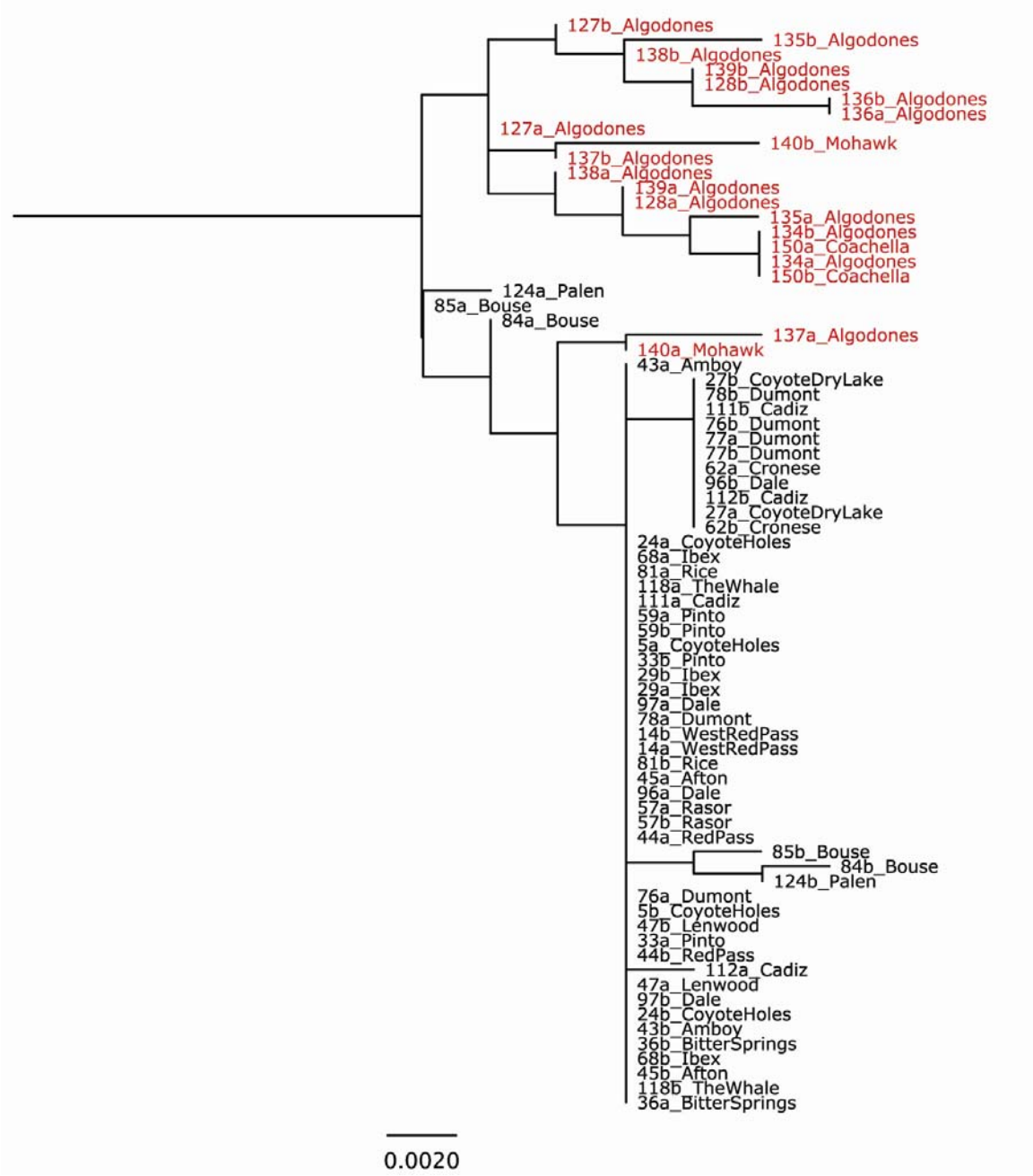

Figure 14. Maximum likelihood gene tree for locus Sun28, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

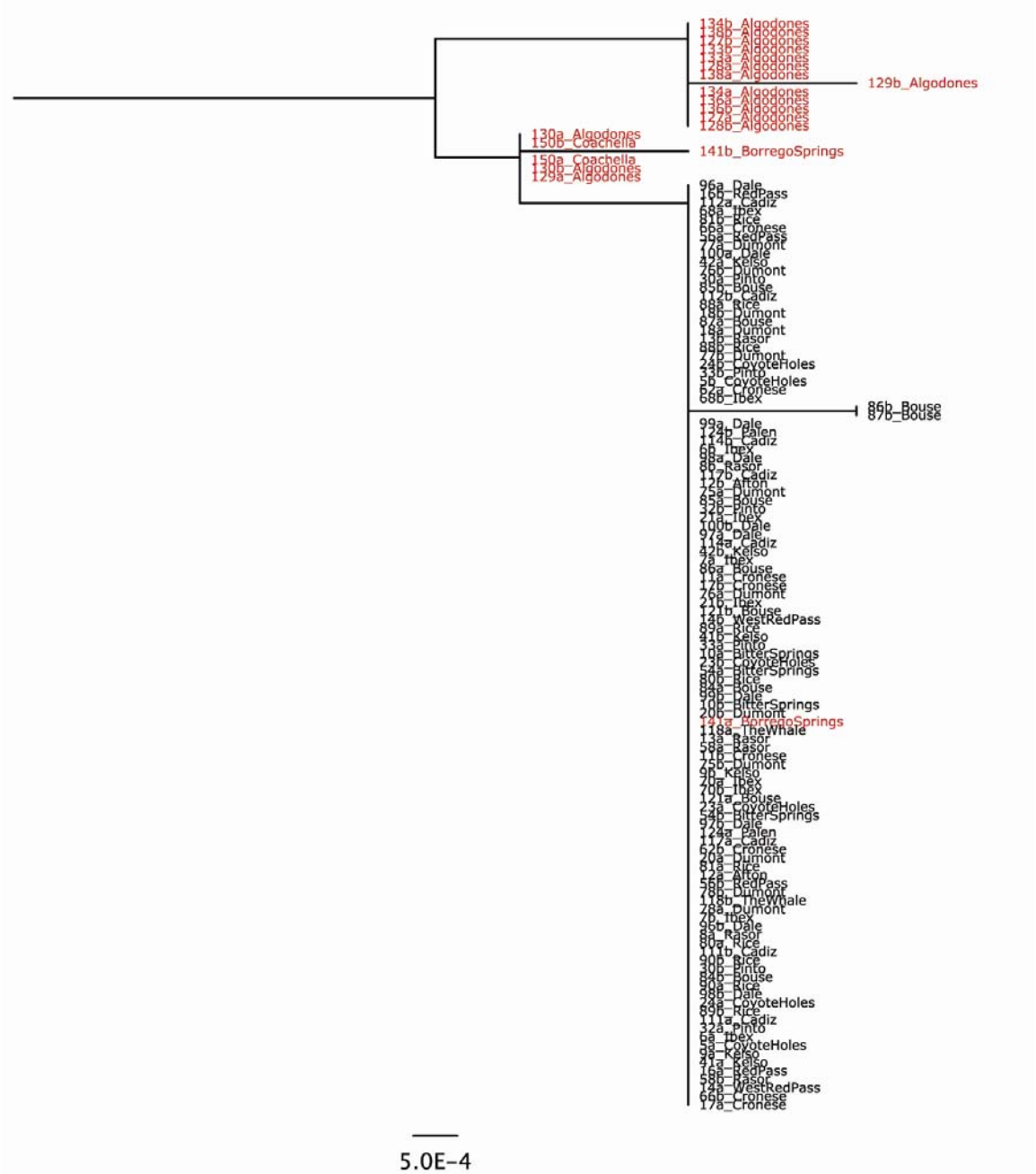

Figure 15. Maximum likelihood gene tree for locus Uma03, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

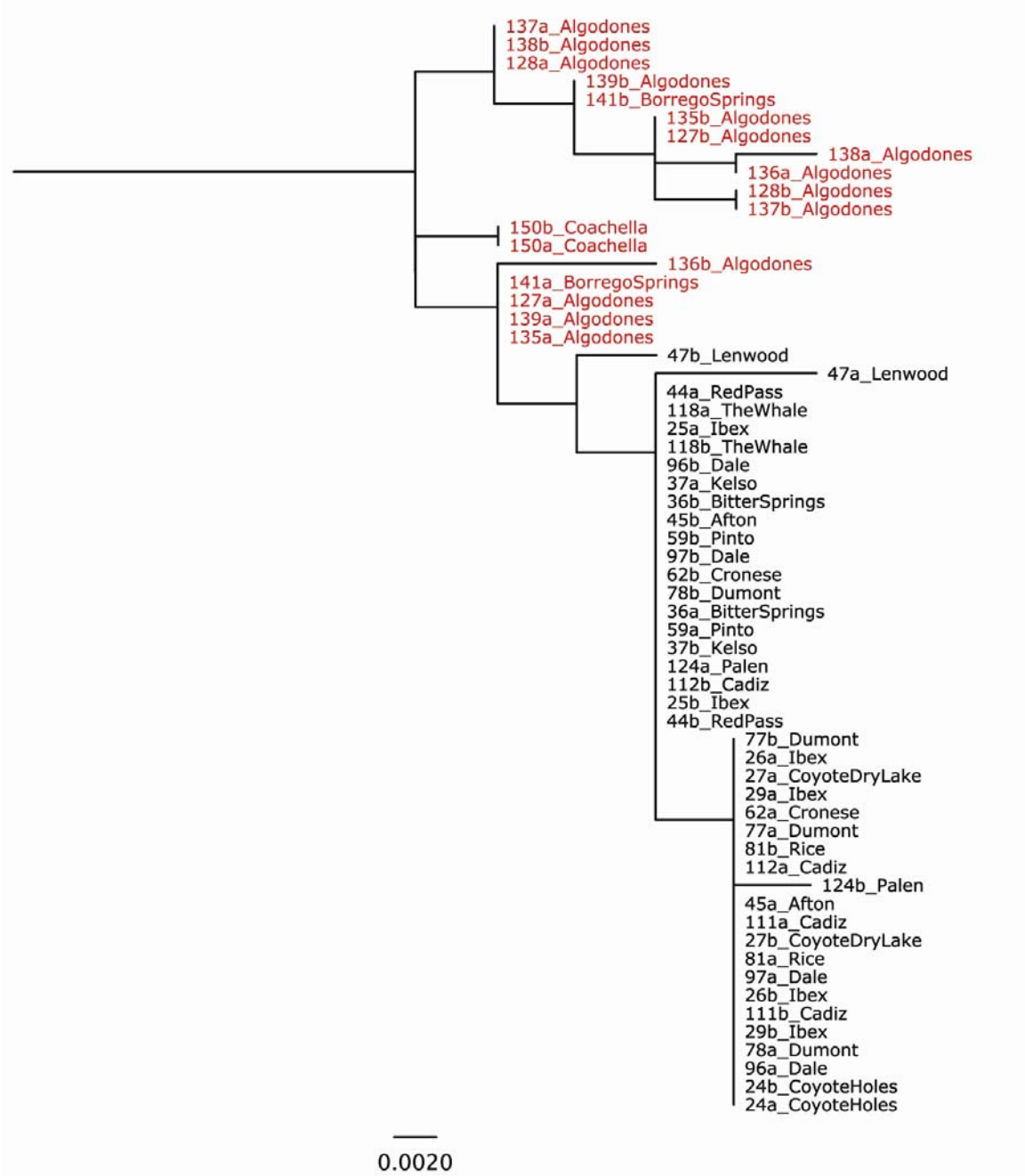

Figure 16. Maximum likelihood gene tree for locus Uma05, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

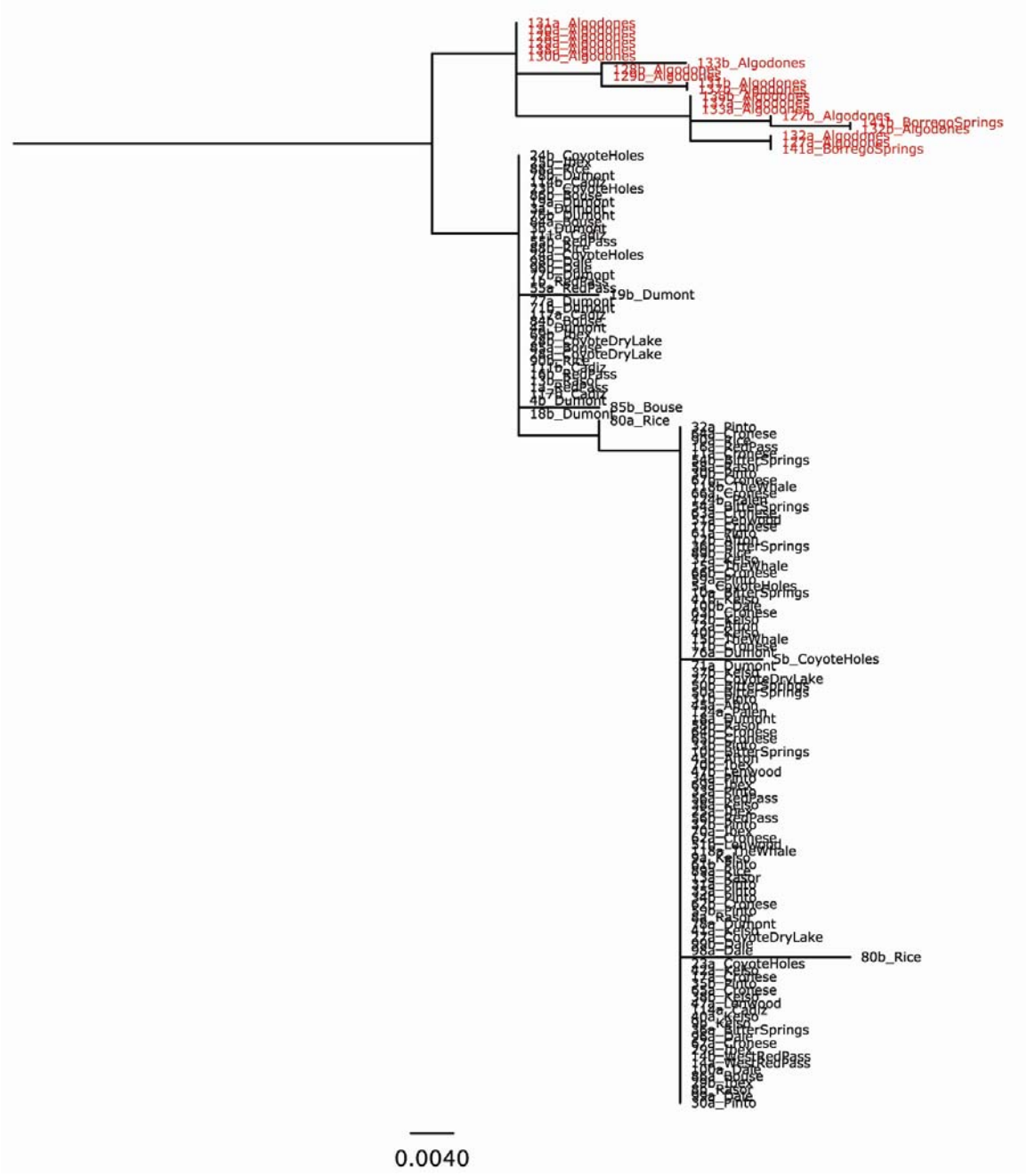

Figure 17. Maximum likelihood gene tree for locus Uma06, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

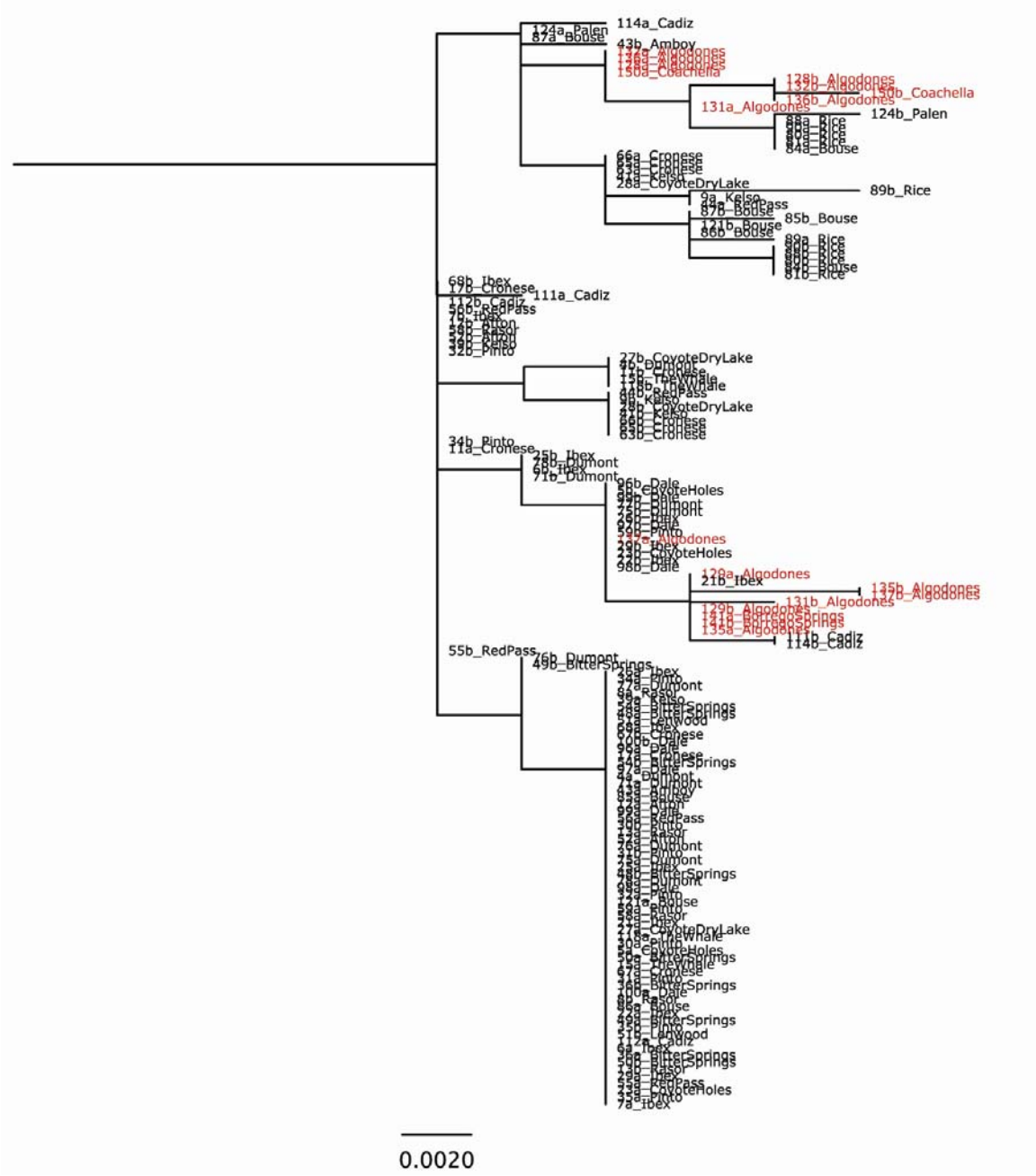

Figure 18. Maximum likelihood gene tree for locus Uma07, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

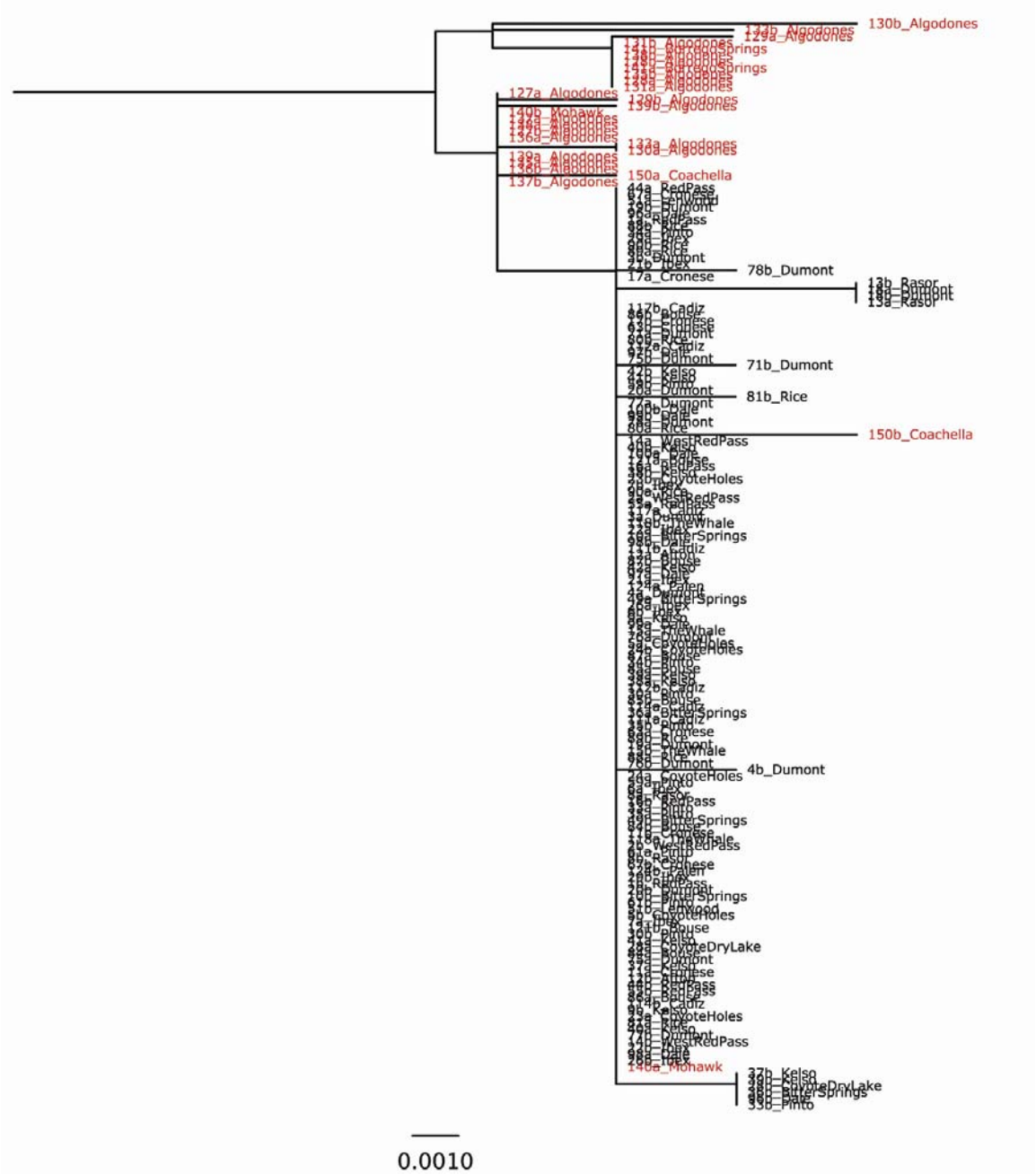

Figure 19. Maximum likelihood gene tree for locus Uma08, generated using Garli v. 1.0. *Uma notata* are shown in red, *U. scoparia* in black. The scale bar is in units of substitutions/site.

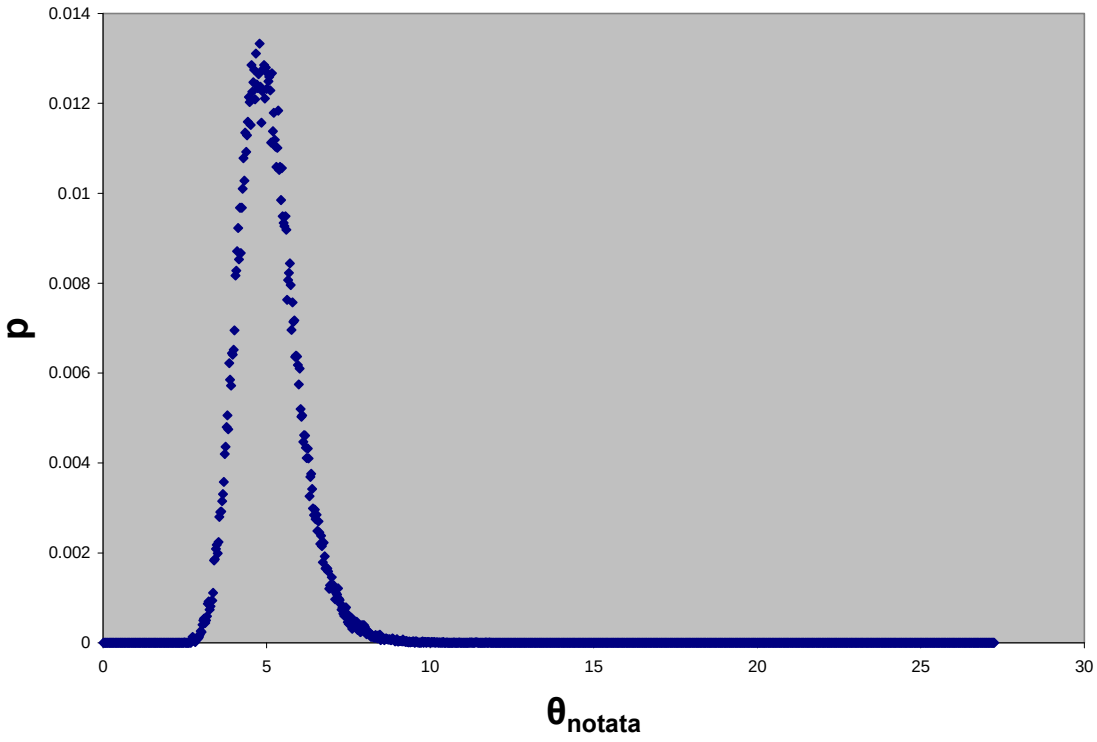

Figure 20. Effective population size for *U. notata* scaled by mutation rate (μ). All data included.

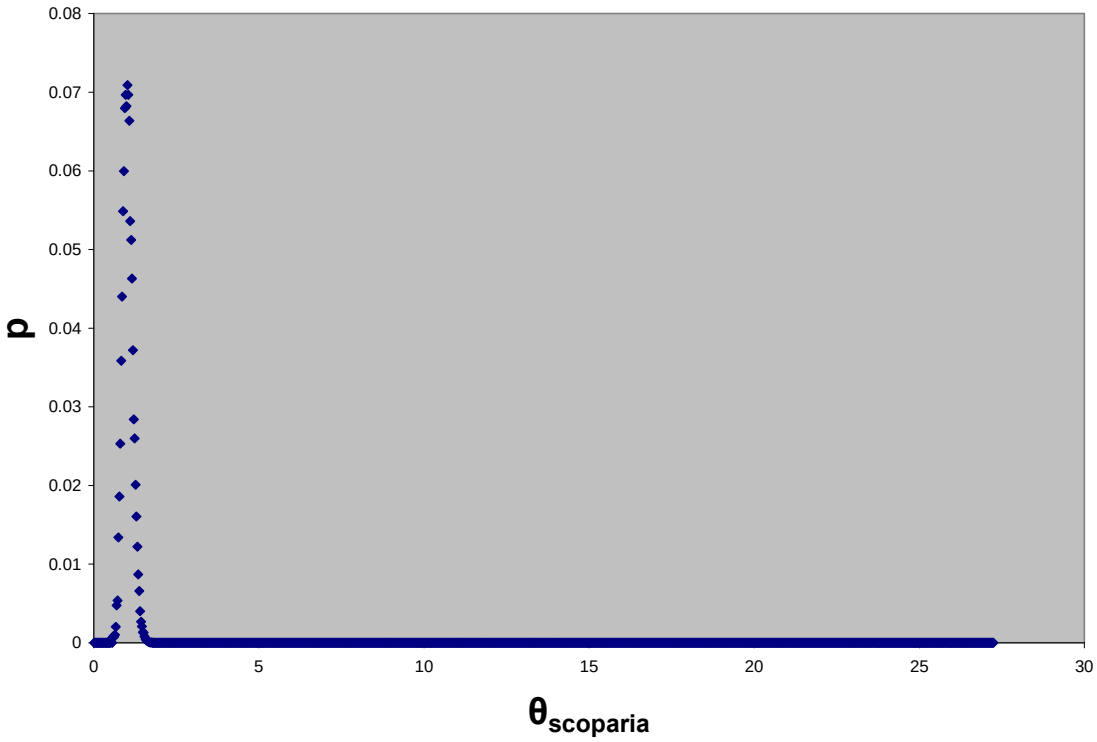

Figure 21. Effective population size for *U. scoparia* scaled by mutation rate (μ). All data included.

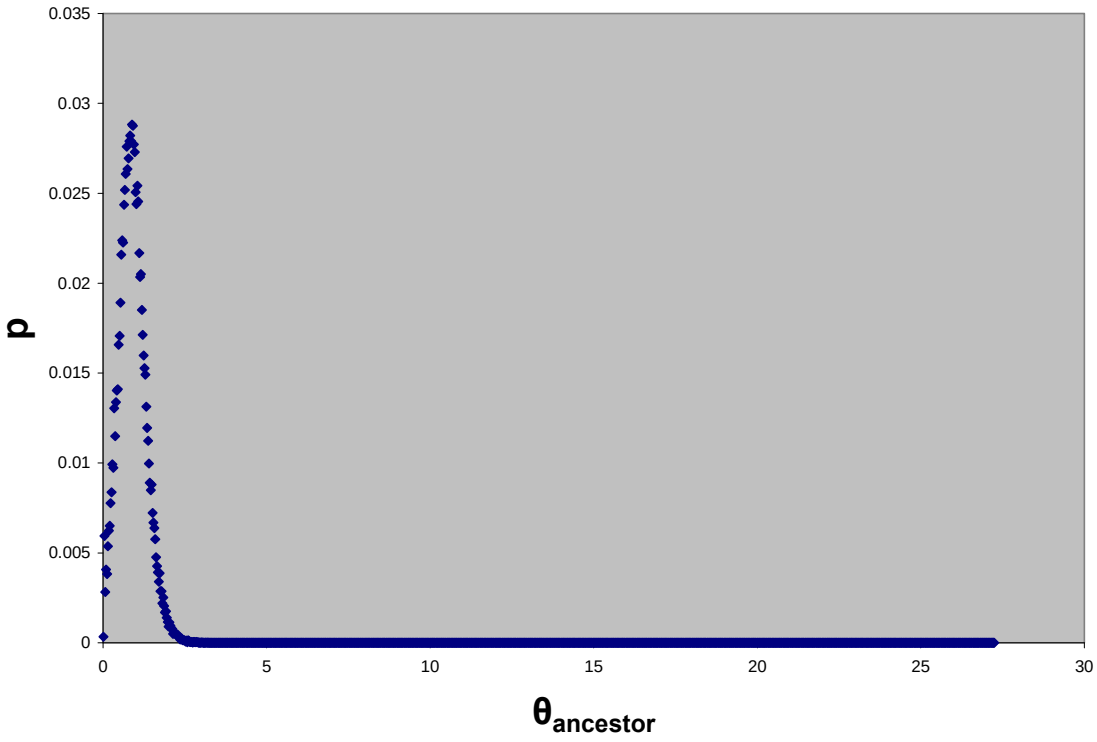

Figure 22. Effective population size for the ancestral population scaled by mutation rate (μ). All data included.

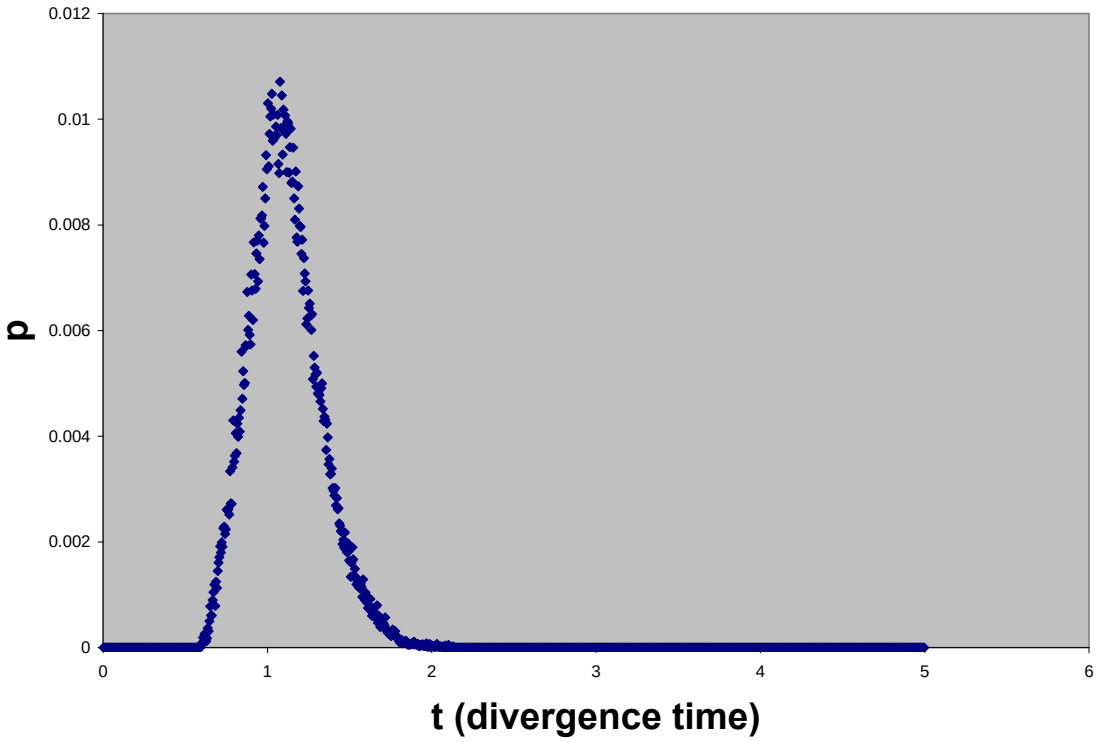

Figure 23. Population divergence (speciation) time between *U. notata* and *U. scoparia*, scaled by mutation rate (μ). All data included.

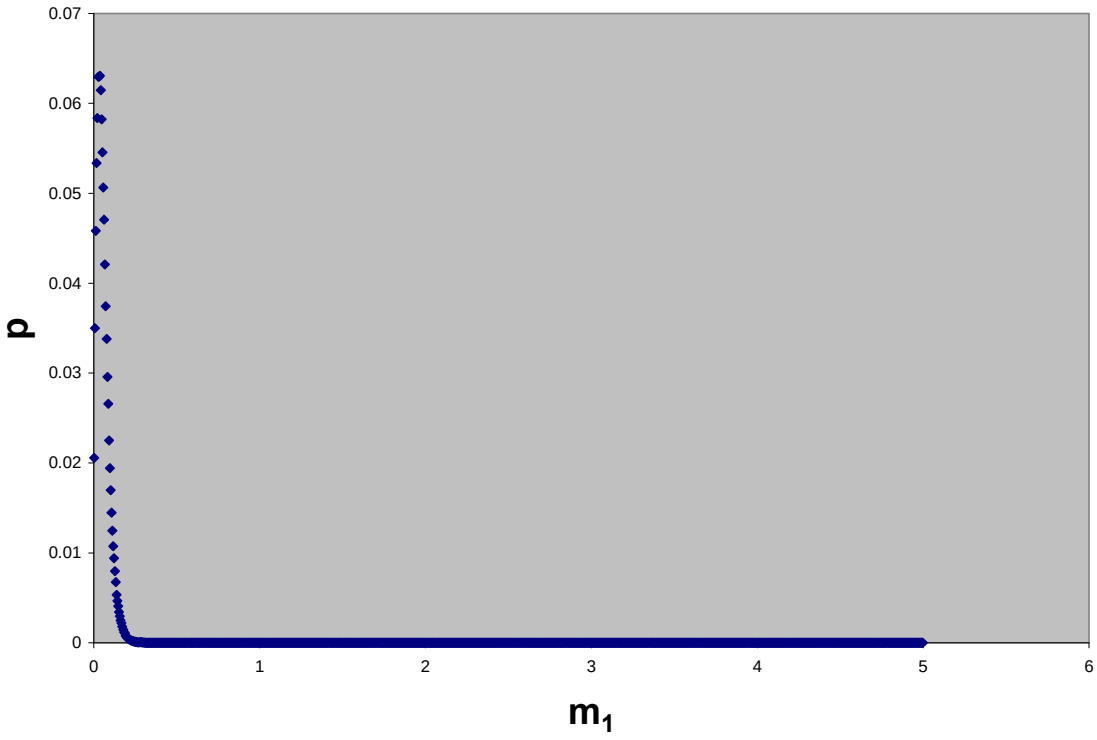

Figure 24. Migration rate (gene flow) into *U. notata* from *U. scoparia*, scaled by mutation rate (μ). All data included.

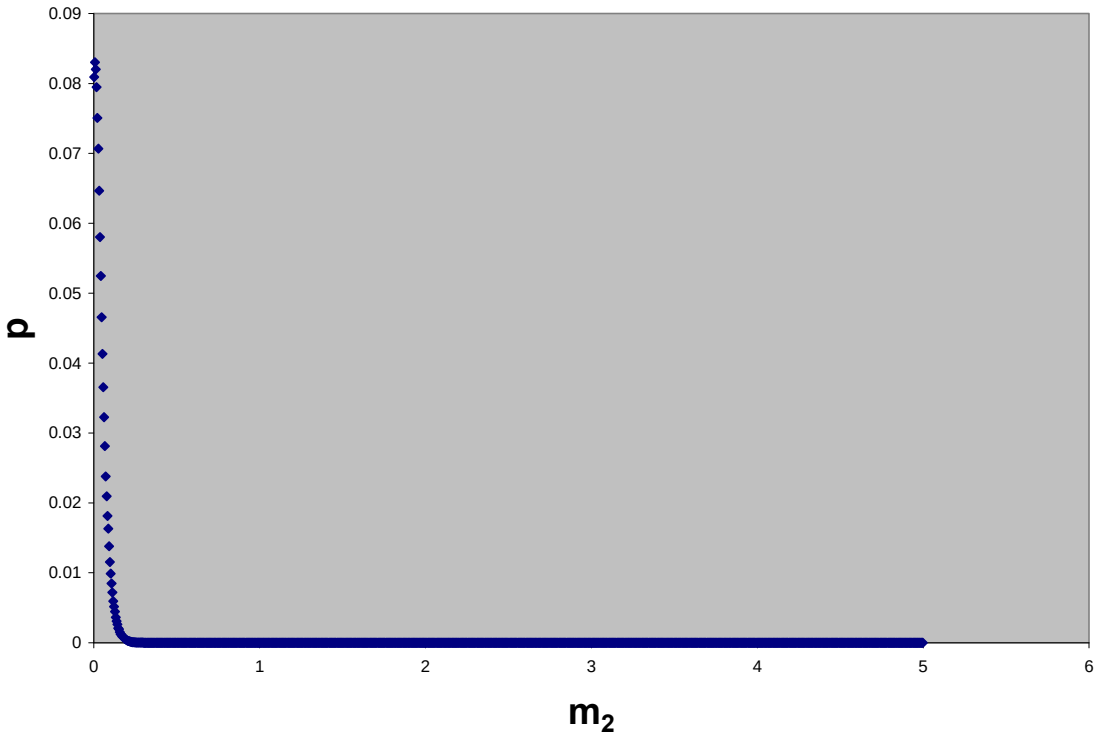

Figure 25. Migration rate (gene flow) into *U. scoparia* from *U. notata*, scaled by mutation rate (μ). All data included.

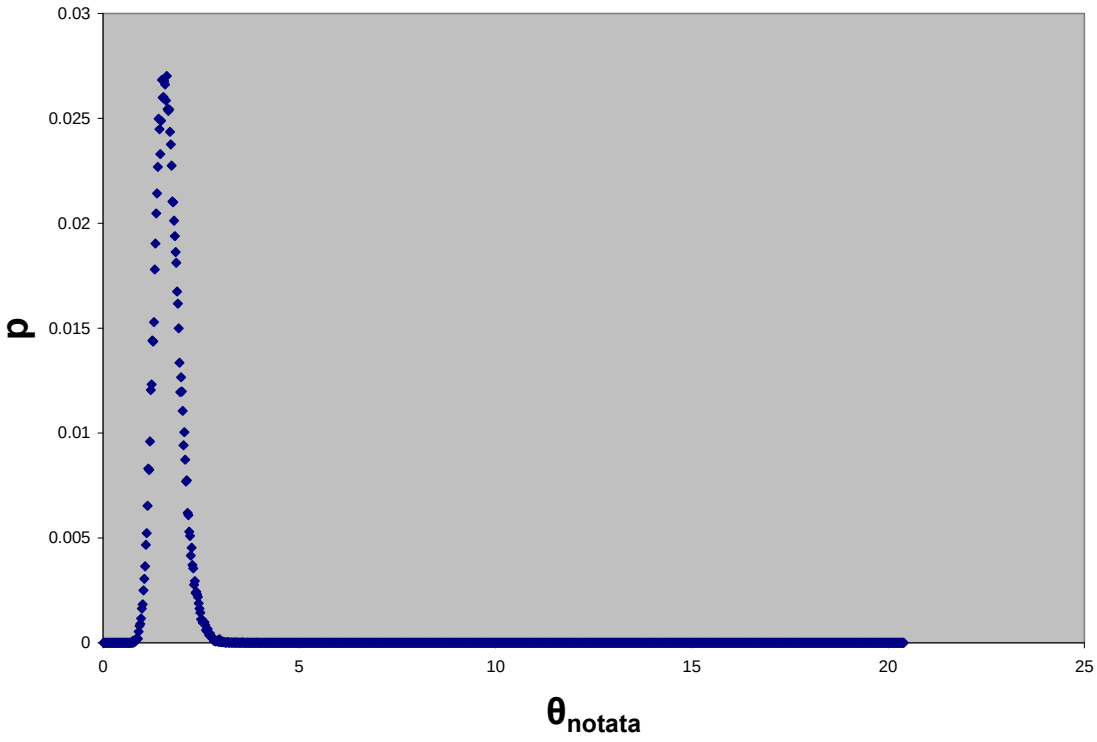

Figure 26. Effective population size for *U. notata* scaled by mutation rate (μ). All recombination excluded.

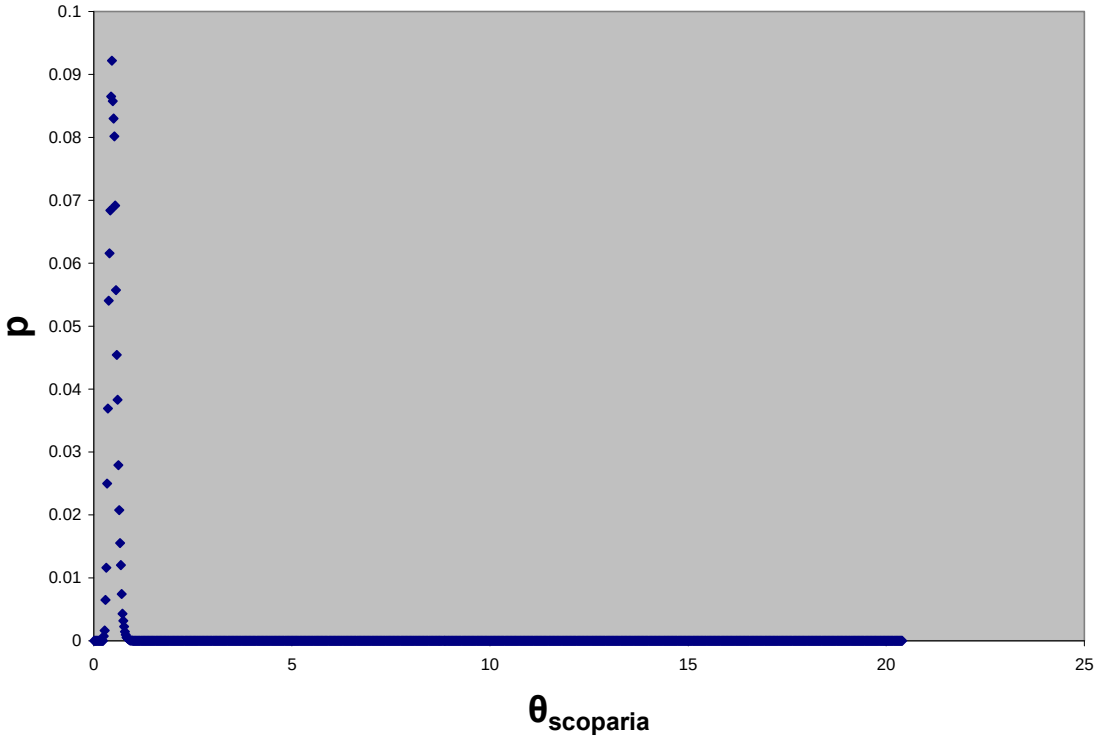

Figure 27. Effective population size for *U. scoparia* scaled by mutation rate (μ). All recombination excluded.

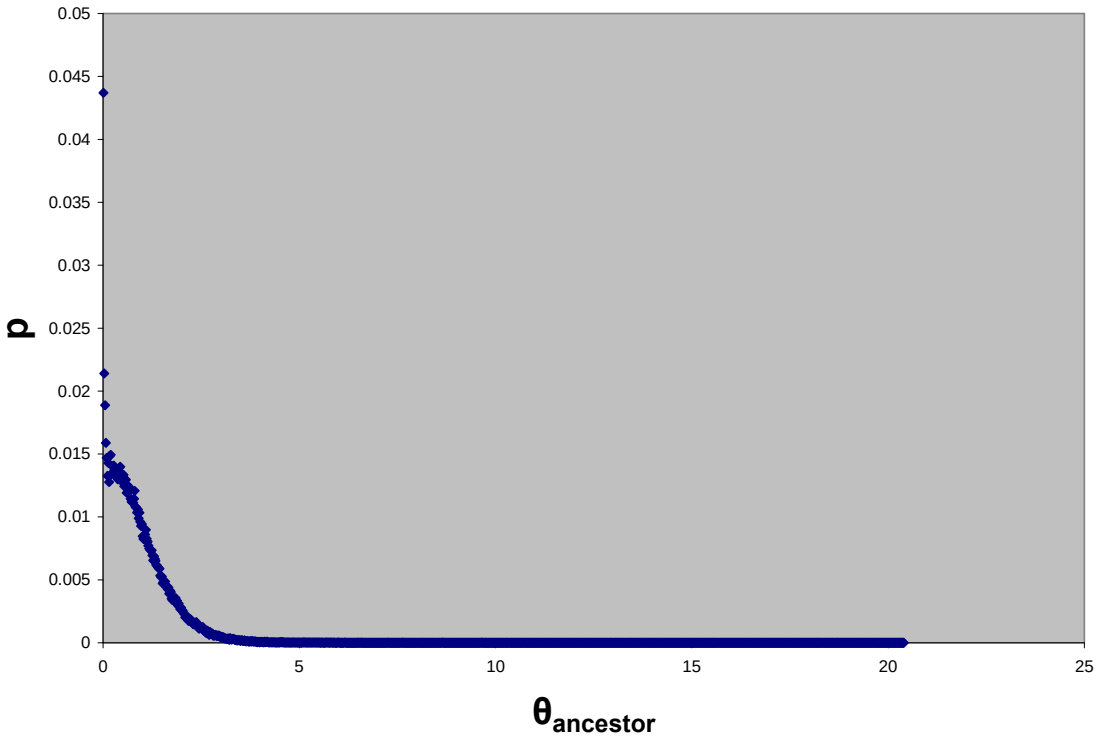

Figure 28. Effective population size for the ancestral population scaled by mutation rate (μ). All recombination excluded.

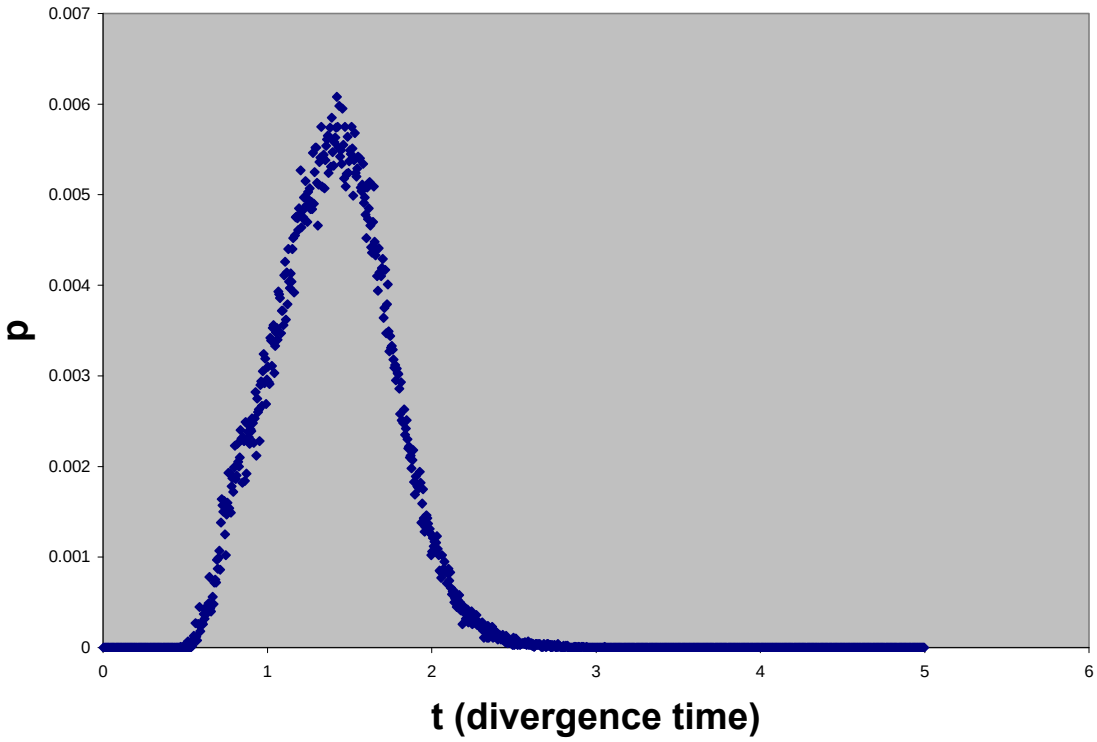

Figure 29. Population divergence (speciation) time between *U. notata* and *U. scoparia*, scaled by mutation rate (μ). All recombination excluded.

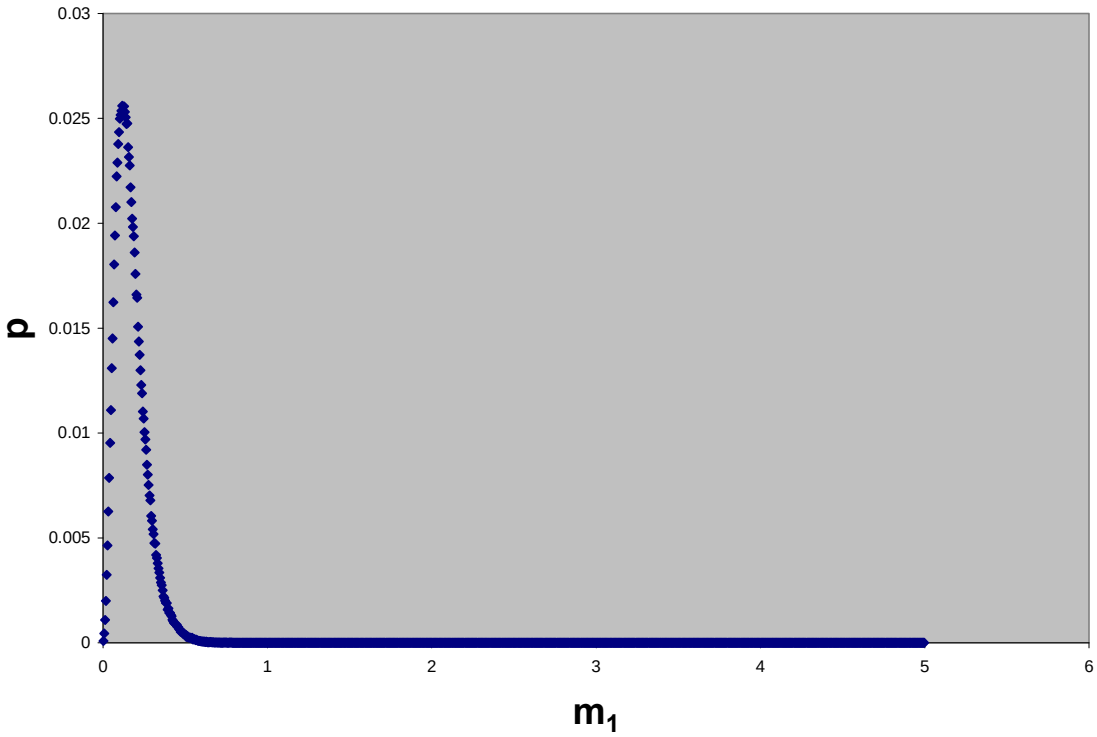

Figure 30. Migration rate (gene flow) into U. notata from U. scoparia, scaled by mutation rate (μ). All recombination excluded.

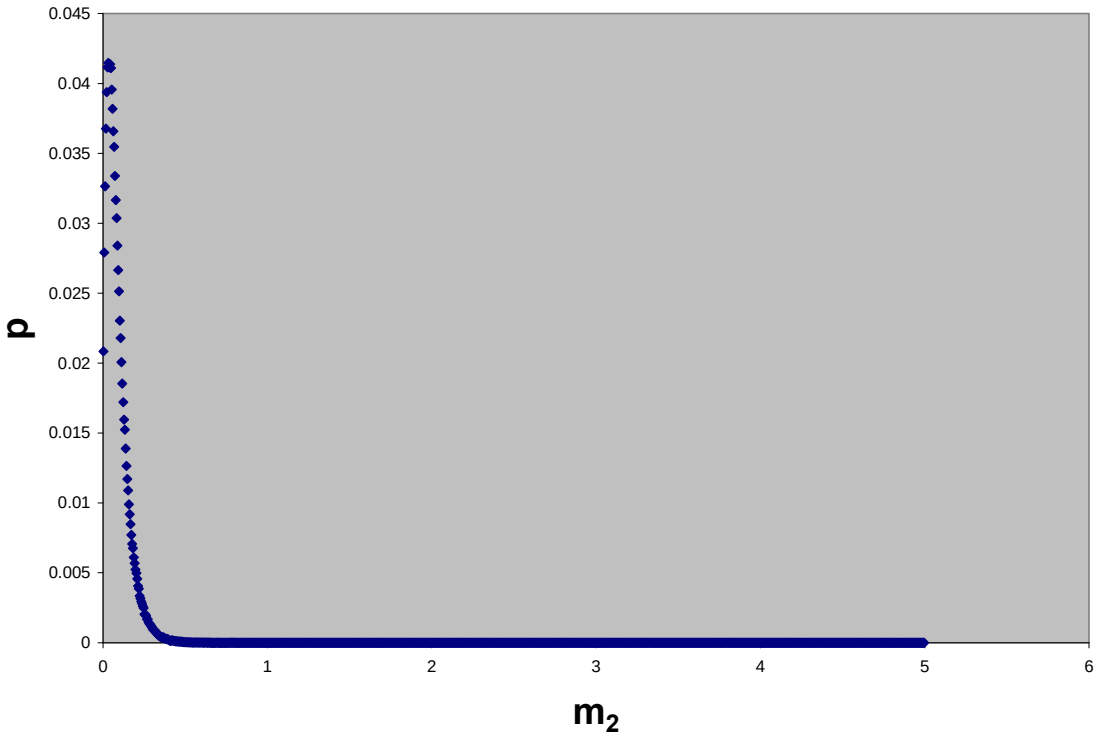

Figure 31. Migration rate (gene flow) into *U. scoparia* from *U. notata*, scaled by mutation rate (μ). All recombination excluded.

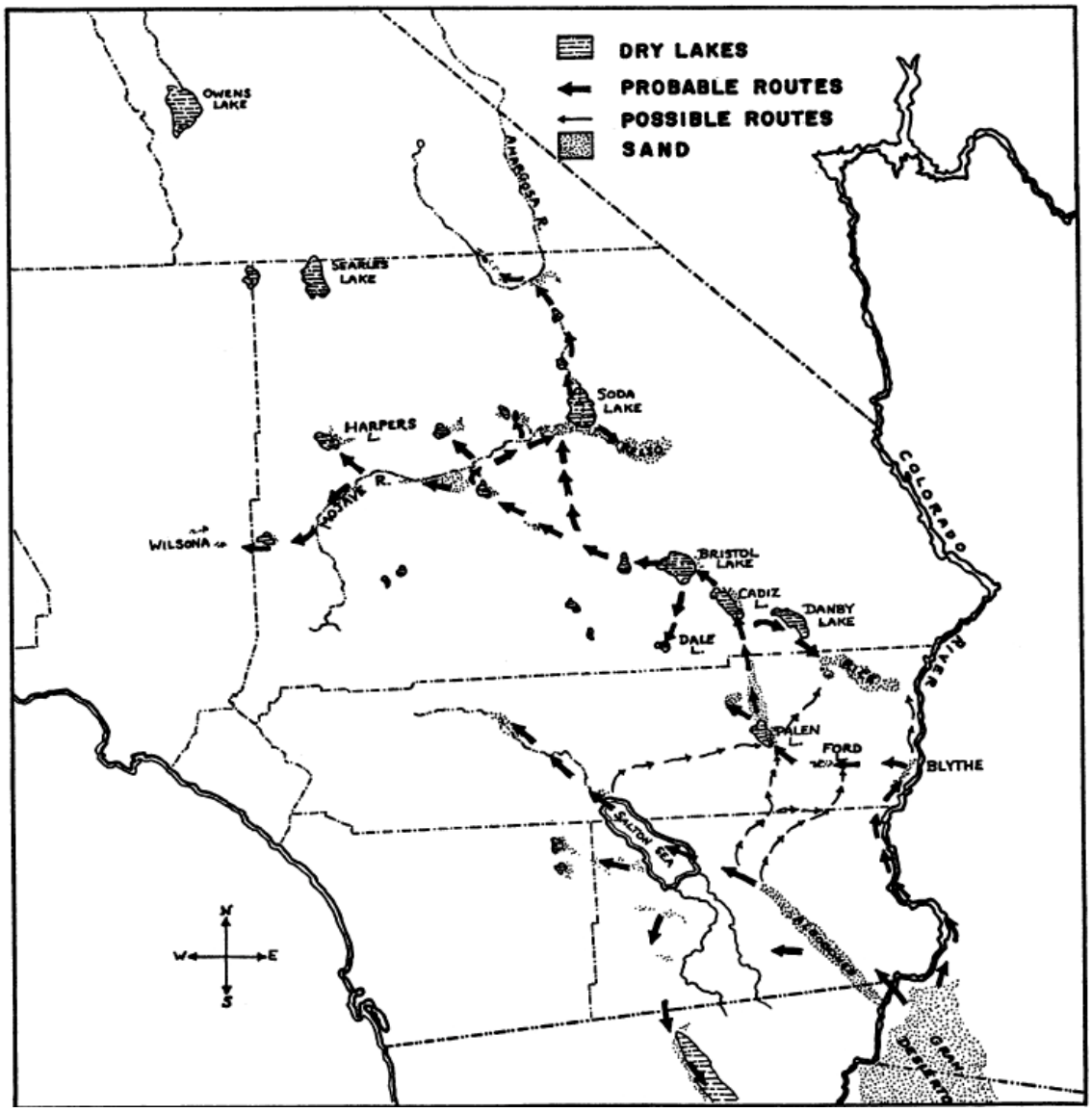

Figure 32. Map of possible migration routes followed by *U. scoparia, U. n. inornata*, *U. n. notata,* and *U. n. rufopunctata* during the glacial cycles of the Pleistocene. Adapted from Norris (1958).