Speciation pattern and process in the trapdoor spider genus *Aptostichus* (Araneae: Mygalomorphae: Euctenizidae)

By

LACIE NEWTON

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Entomology

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Jason Bond, Chair

Lynn Kimsey

Phil Ward

Committee in Charge

2022

Abstract

Spiders placed in the infraorder Mygalomorphae (tarantulas, trapdoor spiders and their kin) are generally recognized as an ancient cosmopolitan lineage that has persisted for over 250 million years. Mygalomorph life history traits that include limited dispersal abilities, habitat specialization, and site fidelity altogether make them ideal organisms for studying speciation pattern and process, phylogeography, and adaptation. Evolutionary studies of mygalomorphs at both shallow and deeper phylogenetic levels have been limited prior to the advent of next generation sequencing approaches, with the majority of such studies relying on morphological characters or limited targeted locus approaches for phylogenetic reconstruction. Thus, it is imperative to implement larger genomic-scale datasets for confident reconstruction of relationships. My dissertation focuses on species delimitation in two trapdoor spider groups, *Antrodiaetus unicolor* complex and *Aptostichus icenoglei* sister species complex, and evaluation of interspecific relationships within the genus *Aptostichus*. To address species boundaries in the *A. unicolor* species complex, I implemented genomic-scale data (i.e., restriction-site associated DNA sequencing, RADseq) in conjunction with morphological, behavioral, and ecological data to evaluate cohesion species identity (Chapter I). Similarly, assessing species boundaries in the *Aptostichus icenoglei* sibling species complex involved a target capture approach for subgenomic data (i.e., ultraconserved elements, UCEs) and ecological data to evaluate genetic and ecological exchangeability, as per the cohesion species-based delimitation approach from a previous study (Chapter II). Lastly, to resolve interspecific relationships within *Aptostichus*, I used a sequence capture method (i.e., UCEs) to generate molecular data in conjunction with morphological data for a combined-evidence phylogeny (Chapter III). These chapters all have the same overarching

theme: to understand the forces underlying divergence between units of diversity and how those units are defined.

CHAPTER I

Integrative species delimitation reveals cryptic diversity in the southern Appalachian *Antrodiaetus unicolor* (Araneae: Antrodiaetidae) species complex

Introduction

Cataloguing and describing species is a crucial first step towards understanding Earth's biodiversity. Species are the foundation of biological questions including evolutionary processes, ecological systems, physiological mechanisms, and are key to formulating conservation efforts. However, defining what species are (Freudenstein et al., 2016; Hey, 2001; Wake, 2006) and the importance/order of certain processes by which new species arise continue to be contentious topics (Butlin et al., 2008). A considerable number of different species concepts have emerged in the scientific literature (e.g., 24 listed in Mayden, 1997), many of which can be incompatible in the sense that they determine a different limit for species (i.e., conclude a different number of species; De Queiroz, 2007). Because these concepts are based on different biological properties (e.g., niche divergence, morphological differences, genetic divergence), certain concepts define varying species limits at particular points in the speciation process (i.e., the gray zone in Figure 1 of De Queiroz, 2007). Although it is common in the literature to discretize modes of speciation, for instance the spatial aspect of divergence ranging from sympatric to allopatric or amount of gene flow (Butlin et al., 2008), speciation is typically considered a multi-level process that occurs across varying temporal and spatial scales, ultimately leading to complete reproductive isolation (Abbott et al., 2013). Therefore, given that a characteristic may be variable in its role in speciation, a species concept that utilizes one data type may mislead estimates of species diversity (e.g., Battey & Klicka, 2017; Bond et al., 2001; Bond & Stockman, 2008; Starrett et al., 2018; Weisrock & Larson, 2006).

Specifically, evaluating species limits is a difficult task in taxa with strong geographic structuring and prone to cryptic diversity (i.e., morphologically indistinguishable species that exhibit extensive molecular divergence; Bickford et al., 2007). Morphological homogeneity may be the result of recent divergence (i.e., insufficient time has elapsed for morphological traits to evolve) or niche conservatism (i.e., geographically isolated species remain morphologically conserved due to high similarity in their ecological niches; Wiens & Graham, 2005). Cryptic species cannot be distinguished by morphology alone; thus, additional data types, such as molecular and ecological data, are necessary for robust species delimitation (Bickford et al., 2007; Stockman & Bond, 2007). For example, recent studies using molecular data in addition to other data types have unveiled cryptic species across many animal taxa, including birds (e.g., Battey & Klicka, 2017; Garg et al., 2016), arthropods (Bond et al., 2001; Bond & Sierwald, 2002; Derkarabetian et al., 2011; Derkarabetian & Hedin, 2014; Garrick et al., 2018; Starrett & Hedin, 2007; Y. Zhang & Li, 2014), amphibians (e.g., Chan et al., 2017; Moritz et al., 2016; Ortega-Andrade et al., 2015; Ramírez-Reyes et al., 2017; Reynolds et al., 2012; Weisrock & Larson, 2006), cnidarians (e.g., Holland et al., 2004), and annelids (e.g., Barroso et al., 2010). These studies highlight that species delimitation utilizing only morphology potentially underestimates species diversity in taxa with relative morphological homogeneity.

Spiders in the infraorder Mygalomorphae (tarantulas, trapdoor spiders, and kin) (Bond et al., 2012; Opatova et al., 2019) present a perplexing situation for robust species delimitation when compared to the more diverse Araneomorphae ('true spiders'), as well as many other taxa. Araneomorph spider species delimitation relies predominately on morphological criteria (e.g., distinctive differences in male genitalia, body size, etc.; Dupérré and Tapia, 2015; Richardson, 2016; also see Bond et al., in review for summary) whereas mygalomorph spiders are relatively

morphologically homogeneous (Bond & Stockman, 2008; Hamilton et al., 2014; Harvey et al., 2018; Hendrixson & Bond, 2005a, 2005b; Huey et al., 2019; Opatova & Arnedo, 2014b) and subject to significant population genetic structuring at microgeographical scales (Bond et al., 2001; Hedin et al., 2015; Starrett et al., 2018). Mygalomorph life history traits that include limited dispersal abilities (with few exceptions; see Coyle, 1983, 1985), relatively long life spans (10-30 years), habitat specialization, and site fidelity altogether make them ideal organisms for studying population divergence and phylogeography (Hamilton et al., 2014; Hedin et al., 2013; Hendrixson et al., 2013; Hendrixson & Bond, 2005a, 2005b; Opatova & Arnedo, 2014a; Satler et al., 2011; Starrett & Hedin, 2007; Stockman & Bond, 2007). Although morphology tends to under-split mygalomorph species, DNA barcoding and related approaches such as GMYC tend to over-split species owing to their population structuring at very fine geographical scales; it has been long recognized (Bond et al., 2001) that mygalomorph populations are highly structured over relatively short distances. Hamilton et al. (2014) posited that GMYC greatly overestimated species diversity in the tarantula genus, *Aphonopelma* (recognizing 114 species), whereas more integrative type approaches recovered fewer species (34). In addition to single locus approaches, multispecies coalescent methods using many loci are more inclined to detect population structure rather than speciation events (Sukumaran & Knowles, 2017). As a result, such attributes of mygalomorph morphological similarity and population structure typify a system requiring more integrative approaches when delimiting species; that is, it is important to evaluate molecular, geographical, ecological, and morphological lines of evidence as opposed to relying on only one data type, analysis, or conceptual approach (e.g., Bond & Stockman, 2008; Hamilton et al., 2014; Hedin et al., 2015). A system prone to such high levels of population subdivision is fertile ground for understanding a classical model of allopatric speciation in which populations diverge

geographically, with subsequent potential for adaptation to local ecological conditions in the absence of significant gene flow (Barraclough, 2019).

The focus of our study is the *Antrodiaetus unicolor* mygalomorph species complex. Like many related mygalomorph spider species, they have relatively long life spans (> 10 years), high site and microhabitat fidelity (e.g., mesic forests and stream banks), and limited dispersal capabilities (Coyle, 1971; Hendrixson & Bond, 2005a, 2005b). Sympatry within the species complex occurs where *Antrodiaetus unicolor* (Hentz, 1842) and *A. microunicolor* (Hendrixson & Bond, 2005a) are co-distributed across part of their ranges in the eastern United States (Figure 1; Hendrixson & Bond 2005a, 2005b). *Antrodiaetus microunicolor* is found only in the southwestern region of the Appalachian Mountains, whereas *A. unicolor* is centered in the central and southern regions of the Appalachian Mountains with peripheral populations extending as far west as the Ozarks in Arkansas and east near the Atlantic Coast, as far south as the Gulf Coast, and as far north to Pennsylvania. In the mesic forests of the Appalachians these spiders are primarily found along creek banks or steep ravines, whereas in the peripheral populations they are isolated in small ravines in humid hardwood forests (Coyle, 1971; Hendrixson & Bond, 2005a). These spiders build subterranean burrows covered by a unique collapsible collar door where they sit and wait for prey; as a result, they rarely depart from their burrows unless disturbed, or to seek a mate in the case of mature males (Coyle, 1971; Hedin et al., 2019).

Antrodiaetus spiders show both morphological stasis and variation across their distribution (Coyle, 1971). Individuals from localities separated by hundreds of kilometers may be morphologically indistinguishable, yet spiders at the same location may exhibit significant disparities in size and coloration (Coyle, 1971; Hendrixson & Bond, 2005a). Due to difficulty

interpreting this intraspecific variation, Coyle (1971) was conservative when revising the genus by maintaining all populations of *A. unicolor* as one species. Later, Hendrixson and Bond (2005a) used morphological and behavioral data to distinguish two forms different in size, coloration, and setal patterns from the Coweeta Long Term Ecological Research site in southwestern North Carolina and described *A. microunicolor* as a new species. A subsequent molecular analysis using two genetic markers (mtDNA gene cytochrome oxidase I and nuclear ribosomal RNA gene 28S) to evaluate species boundaries between these two forms showed that *A. unicolor* is paraphyletic with respect to *A. microunicolor* (Hendrixson & Bond, 2005b), which they attributed to improper taxonomy after exploring other potential interpretations (e.g., inadequate phylogenetic information, interspecific hybridization, incomplete lineage sorting, unrecognized paralogy). Because *A. microunicolor* is morphologically and behaviorally distinct, its nesting within *A. unicolor* suggests potential for multiple cryptic species among the paraphyletic assemblage of sister lineages; that is, the hierarchical level of genetic divergence coupled with apparent reproductive isolation (e.g., differences in male dispersal) may be reflective of underlying changes in phenotype among the other lineages, to include reproductive isolation, that is not readily apparent using standard taxonomic data.

With the advent of next-generation sequencing methods, it is now feasible to generate genomic-scale data for non-model organisms (Baird et al., 2008; Faircloth et al., 2012; Lemmon et al., 2012). These much larger, more comprehensive data sets provide a framework for more rigorous tests of species boundaries in systems where previously only a few targeted loci were available. RADseq (restriction-site associated DNA sequencing) approaches are one of the most widely used techniques for generating a reduced representation of the nuclear genome with extensively sampled homologous loci (Andrews et al., 2016; Baird et al., 2008). Various

RADseq techniques have been utilized for addressing population genetic and phylogeographic studies (Andrews et al., 2016; Baird et al., 2008; Emerson et al., 2010), reconstructing phylogenetic relationships among both closely (e.g., Eaton et al., 2016) and distantly (e.g., Leaché et al., 2015) related species, and evaluating species boundaries and speciation processes (e.g., Battey & Klicka, 2017; De Jesús-Bonilla et al., 2019; Delgado□Machuca et al., 2019).

We generated RADseq data using 3RAD, a three-enzyme protocol that reduces DNA chimaera and adapter-dimer formation (Bayona-Vásquez et al., 2019), to investigate species boundaries and phylogenetic relationships within the *A. unicolor* species complex. Specifically, we employed the cohesion species concept (CSC; Templeton, 1989), which defines a species as a set of populations that derives from a single evolutionary lineage and meet the criteria of being genetically exchangeable and/or ecologically interchangeable. The CSC can be applied across all types of taxa, takes into account multiple important biological properties for evaluating potential adaptive divergence (thus coupling speciation process and species delimitation), and provides a methodological framework for species hypothesis testing (e.g., Bond & Stockman, 2008). This integrative approach is particularly insightful when evaluating species limits in morphologically cryptic taxa with high population genetic structuring (e.g., Bond & Stockman, 2008; Hendrixson et al., 2013; Stockman & Bond, 2007). For example, Bond & Stockman (2008) uncovered extensive molecular divergence in addition to potential adaptive (ecological) divergence (i.e., coastal dune habitat versus inland) in the mygalomorph *Aptostichus atomarius* species complex, leading to the recognition of three additional cohesion species. In a follow up study, Garrison et al. (2020) recently evaluated gene family conservation across the same sister species complex and found gene families associated with venom production, metabolism, and sensory systems under positive selection in dune endemic lineages; thus, adaptive divergence (habitat/ecological

divergence) appears to be reflected at a genomic scale. To evaluate cohesion species identity within the *A. unicolor* species complex, we assessed the amount of genetic population structure using clustering analyses to identify the number of evolutionary lineages (i.e., are these lineages genetically exchangeable?), and constructed niche-based distribution or ecological niche models (ENMs) for each lineage and then compared to evaluate niche overlap/similarity (i.e., are these lineages ecologically interchangeable?).

Methods

Genomic Library Prep

We sampled 157 individuals from 103 localities throughout the geographic distribution of the *Antrodiaetus unicolor* species complex (Supplementary Table S1; Figure 1). We extracted genomic DNA from leg tissue for each individual using the Qiagen DNeasy Blood and Tissue kit (Qiagen, Valencia, CA) following the manufacturer's protocol. DNA was quantified using a Qubit 3.0 Fluorometer (Life Technologies, Inc., Grand Island) and checked for quality using an agarose gel.

RADseq libraries were prepared following the Adapterama III protocol (Bayona-Vásquez et al., 2019), which is a modified version of ddRAD (Peterson et al., 2012) that uses three enzymes for digesting genomic DNA (3RAD). This protocol alleviates the need for large quantities of DNA since the third enzyme prevents adapter dimers and DNA chimaeras from forming during the reaction (Bayona-Vásquez et al., 2019; Graham et al., 2015; Hoffberg et al., 2016). Based on previous studies that tested different enzyme combinations (Bayona-Vásquez et al., 2019; Burns et al., 2017), we chose EcoRI-HF, MspI, and ClaI as our cohort of enzymes for genomic DNA digestion.

For each sample, 100 ng of genomic DNA were digested for 2 hr at 37°C in an 18 µl solution consisting of 1X Cutsmart buffer, 20U EcoRI-HF, 20U MspI, 10U ClaI, 0.28 µM each of forward and reverse adapters. Immediately following digestion, the reaction was brought to 24 µl total consisting 0.25x Ligase buffer, 100U T4 DNA Ligase, 0.75 µM rATP, and incubated at 22°C for 20 min and 37°C for 10 min for 3 cycles and a final 20 min enzyme kill step at 80°C. Libraries were pooled and cleaned with 1.5X volume Sera-Mag Speedbeads (Fisher Scientific, Pittsburgh, PA) and quantified using the Qubit.

PCR was used to attach the iTru5 primer. Six replicate 50 µl reactions consisting of 10 µl of pooled DNA, 1X Kapa HiFi Fidelity Buffer (Kapa Biosystems, Massachusetts, EUA), 0.3 µM dNTP mix, 0.5 µM iTru5_8N primer, and 1U KAPA HiFi Hotstart DNA polymerase with conditions at 95°C for 2 min of initial denaturation, followed by 98°C for 20 sec, 61°C for 30 sec, and a 5.5 min 72°C extension. The pools were combined and then 2X volume bead-cleaned. Twelve iTru7 primers were added using PCR with three reactions consisting of 10 µl of pooled DNA, 1X Kapa HiFi Fidelity Buffer, 0.3 µM dNTP mix, 0.5 µM P5 primer, 0.125 µM of four different iTru7 primers, and 1U KAPA HiFi Hotstart DNA polymerase with conditions at 95°C for 2 min; 6 cycles at 98°C for 20 sec, 61°C for 15 sec, and 72°C for 30 sec; and a final extension at 72°C for 5 min. The products were pooled and 2X volume bead-cleaned and quantified.

Size selection was performed using Pippin Prep (Sage Science, Beverly, MA) to capture 570 bp +/- ~18% (470-670 bp range) fragments using a 1.5% agarose cassette. The last enrichment PCR for the final size-selected pool was run with 10 µl of pooled DNA, 1X Kapa HiFi Fidelity Buffer, 0.3 µM dNTP mix, 0.5 µM each of P5 and P7 flowcell binding primers, and 1U KAPA HiFi Hotstart DNA polymerase with conditions at 95°C for 2 min; 9 cycles at 98°C for 20 sec, 61°C for 15 sec, 72°C for 45 sec; and lastly an extension at 72°C for 5 min. Final

libraries were sent to the Georgia Genome and Bioinformatics Core or UC Davis Genome Center for paired-end 150 bp mid-output sequencing with the Illumina NextSeq 550 platform.

Sequence Analysis

Short read processing was conducted using ipyrad v.0.7.1 (Eaton, 2014) on the Hopper HPC Cluster at Auburn University. The workflow for ipyrad involves the following steps: 1) demultiplexing raw reads; 2) quality filtering reads; 3) de novo clustering of data within samples using VSEARCH (Rognes et al., 2016) with clusters then aligned using MUSCLE (Edgar, 2004); 4) jointly estimating heterozygosity and error rate; 5) estimating consensus allele sequences from clustered reads; 6) de novo clustering of data and alignment across samples; 7) filtering/formatting output files for downstream analyses. The majority of parameters were set to default values (see params file in Dryad repository for more details). However, since locus occupancy (i.e., the amount of samples that have a particular locus) is known to have a substantial effect on the total number of loci retained in the data set (Crotti et al., 2019; Eaton et al., 2016; Huang & Knowles, 2016), branched assemblies were created with varying amounts of locus occupancy percentages (i.e., altered the minimum sample per locus parameter; see Table 1) in addition to a differing number of individuals present in the matrix (i.e., all individuals and only *A. unicolor*B individuals). The data matrix including all individuals to be used for downstream analyses was chosen after comparing the efficacy of each built from the branched assemblies (see Figure 3 & Results).

Phylogenetic Analysis

We estimated the phylogeny for several datasets (see Table 1 for dataset descriptions; Supplemental Figures S1-3) within a maximum likelihood framework implemented in RAxML v8.2.4 (Stamatakis, 2014). We used the GTR+G model of sequence evolution. For each analysis,

1000 bootstrap replicates were calculated using the rapid bootstrapping option implemented in RAxML. The phylogenies constructed from these datasets were viewed in Figtree v1.4.1 with midpoint rooting. To verify that our rooting was not misplaced, we also sampled the root in our species tree analysis employed in SNAPP (see below) and recovered the same rooting with full support (posterior probability value = 1). The phylogenies were then compared to look for congruence among clades and bootstrap values, with the final tree having the largest amount of taxon coverage and high support values used for downstream analyses (All_M20 dataset with all 157 individuals and a minimum of 20% locus occupancy; Figure 2).

Bayesian inference of the All_M20 data set was conducted in ExaBayes v1.4.1 (Aberer et al., 2014). Two independent runs of $6x10^7$ generations with four coupled chains each were run simultaneously starting with a random starting tree. Standard deviation of split frequencies was monitored (< 0.05), and the first 25% were discarded as burn-in. An extended majority rule consensus tree was generated using the program consense within ExaBayes (Supplemental Figure S4; Aberer et al., 2014).

We generated a coalescent based tree from a set of gene trees using the software ASTRAL-III (Supplemental Figure S5; Zhang et al., 2018). The MAGNET v0.1.5 pipeline (Bagley, 2019) was used to estimate a gene tree for each RAD locus within the All_M20 data set using RAxML v.8.2.4, with each RAxML run implementing the GTR+G model. These gene trees were used as input for ASTRAL-III. We inferred a species tree (Supplemental Figure S6) using SNAPP (Bryant et al., 2012) implemented in Beast v2.5.2 (Bouckaert et al., 2019). Due to computational constraints, six representatives spaced out across their geographic range were chosen from each of the five lineages as input for SNAPP. Three independent runs were conducted with all default parameters except for a chain length of two million generations.

LogCombiner v.2.5.2 and TreeAnnotator v2.5.2, both implemented in Beast v2.5.2 (Bouckaert et al., 2019), were used to combine posterior distributions of trees from each run and construct the Maximum Clade Credibility Tree with ten percent burn-in, respectively. Convergence was assessed by monitoring stationarity of the MCMC chains and ESS values with Tracer v1.7.1 (Rambaut et al., 2018).

Cohesion Species Delimitation

We employed the Bond and Stockman (2008) approach by utilizing two criteria to define cohesion species: genetic exchangeability and/or ecological interchangeability, which means that two or more populations exchanging genes and having similar ecological niche attributes are considered a cohesion species. To identify the number of evolutionary lineages, we used the RAxML topology (Figure 2) to establish a baseline evolutionary framework.

Genetic Exchangeability

Following the Bond and Stockman (2008) flowchart for delimiting cohesion species, we first established a "basal lineage" starting point, resulting in five monophyletic lineages (see Figure 2 for highlighted lineage designations). Lineage name designations were based on clade labels from Hendrixson and Bond (2005b): *A. microunicolor* and four lineages within *A. unicolor* (A, B1, B2, and B3). We assessed the potential for gene flow (i.e., genetic exchangeability) between sister lineages by evaluating the distributions of these lineages (Figure 1).

To help elucidate the potential for gene flow, several clustering analyses were performed. STRUCTURE v 2.3.4 (Pritchard et al., 2000) was run with several data matrices of unlinked SNPs for 1,000,000 generations and 100,000 burn-in using the admixture model featuring K values ranging from 1-10, with five replicate runs for each K value. The package pophelper

(Francis, 2017) in R v.3.5.3 (R Core Team, 2019) was used to calculate the optimal K value via the Evanno method of calculating **Δ**K (Evanno, Regnaut, and Goudet 2005) as well as view the output (Figure 4a). In addition to STRUCTURE, we used the R package adegenet (Jombart, 2008; Jombart & Ahmed, 2011) to perform a principal components analysis (dudi.pca function) and visualize the clustering of each lineage (Figure 4b). Also, an unsupervised machine learning approach, specifically a Variational Autoencoder (VAE) derived from Bayesian probability theory, was implemented to provide another clustering algorithm (Figures 4c & 4d; for details see Derkarabetian et al., 2019). A VAE, a class of neural networks, takes input data with a high dimensionality and compresses it through multiple encoding layers into two-dimensional latent variables, with subsequent reconstruction of the data by un-compressing the latent variables through multiple decoding layers (Derkarabetian et al., 2019). The latent variables, which are given as a normal distribution with a mean (μ) and standard deviation (σ), provide a two-dimensional representation of our unlinked SNP data, with relatively clear visualization for assessing cluster uncertainty due to the standard deviation around samples/clusters (e.g., see Figure 4d; Derkarabetian et al., 2019). This approach, which relies on inherent structure within the data for clustering individuals, alleviates the need for a priori hypotheses of species number, level of divergence, and any population parameters.

Genetic exchangeability was rejected in cases of allopatric lineages with distinguishable clusters in clustering analyses. However, in cases of sympatric or parapatric lineages with uncertainty of whether a barrier to gene flow is present, we tentatively accept genetic exchangeability unless the clustering analyses indicate separate clusters. Because cohesion species can be rejected by either not being genetically exchangeable **or** ecologically interchangeable, these lineages were also tested for ecological interchangeability.

Ecological Interchangeability

We evaluated a proxy for ecological interchangeability by measuring the overlap between ENMs for the sister lineages being compared (generally following the approach of Stockman & Bond, 2007) as well as taking into account previous morphological and behavioral criteria for distinguishing *A. microunicolor* (Hendrixson & Bond, 2005a). Current climate data from 1970-2000 for 19 bioclimatic variables at 30 arc-second resolution were downloaded for tiles 12 and 13 from WorldClim v.2 (http://worldclim.org/version2; Fick and Hijmans, 2017). Climate data from the tiles were then merged into layers, cropped to the area of interest, and converted to a raster stack using the packages 'raster' (Hijmans, 2015) and 'rgdal' (Bivand et al., 2019) in R. The software ENMTools v1.4.4 (Warren et al., 2010) was used to estimate the amount of correlation between the 19 bioclimatic variables, with significant correlation defined as $r > 0.9$ following Jezkova et al. (2011). Six variables were removed and the 13 retained variables (BIO 2-9,12,13,15,16,18) were used for generating ENMs. Maxent v.3.4.1 (Phillips et al., 2006) was used to estimate ENMs with default settings for the lineages. The receiver-operating characteristic (ROC) plot's area under the curve (AUC) was used as a measure of model prediction accuracy, with values > 0.9 indicating optimal model performance as opposed to values < 0.7 indicating poor model performance (Swets, 1988). Occurrence records were based on specimens collected for this study and specimens from prior publications (Hendrixson & Bond, 2005a, 2005b).

To statistically compare the ENMs of each lineage, we conducted analyses of niche overlap, niche identity, and niche similarity (i.e., background) tests in ENMTools (Warren et al., 2008, 2010). Niche overlap for each sister lineage comparison was quantified using Schoener's *D* (Schoener, 1968), which ranges from 0 (no overlap between ENMs) to 1 (complete overlap of

ENMs). To assess the significance of *D*, we employed both niche identity and niche similarity tests. The niche identity test compares whether the niche overlap values differ significantly from the null distribution, with a significant difference suggesting that individual occurrences in an area is not random. A hundred pseudoreplicates were used to construct a null distribution of niche overlap compared to the observed overlap value using a one-tailed test (Supplemental Figure S7). Warren et al. (2008) highlighted that the niche identity test may be too strict and results in the null hypothesis often being rejected as it relies strictly on occurrence points. Thus, we conducted the niche similarity test to compare niche overlap between lineages relative to the niche spaces available to those lineages (i.e., habitats within the estimated geographic range). Here, a distribution of niche overlap estimates for occurrence points of one lineage and randomly selected points from the background distribution of the other lineage is generated. Niche overlap that is significantly more similar or different relative to the background than expected by chance indicate niche conservatism or niche divergence, respectively (i.e., not a function of geography). We estimated the background regions for each lineage based on circular buffers around occurrence points in ArcMap v10.7 (ESRI). Occurrence points of one lineage were tested against random points from the background region of the other lineage and vice versa. A hundred pseudoreplicates were used to construct a null distribution of niche similarity compared to the observed overlap value using a two-tailed test (Supplemental Figures S8-11). In order to reject ecological interchangeability with the niche similarity test, both of the background tests for a pair of lineages should have a *D* value that is more different (niche divergence) than expected by chance. Since some of our comparisons involved parapatric/sympatric lineages and results may be affected by the defined background region, we conducted sensitivity tests by defining

alternative background regions based on buffer circles from occurrence points of differing radii (25 km, 50 km, 75 km, and 100 km) in ArcMap v10.7.

Results

3RAD Data

For the 157 specimens sampled throughout the distribution of the *A. unicolor* species complex, 234,116,852 raw reads were obtained ranging from 365,011 to 4,698,473 and averaging 1,491,190 reads per sample (Table 1). Table 1 outlines the number of loci and SNPs generated in each dataset with different minimum locus occupancy percentage (min_sample_locus) parameter values. Additional clustering thresholds (0.85, 0.88, 0.90) were tested as well, but there were minimal differences in the number of retained loci/SNPs across these thresholds; therefore, we chose the threshold value of 0.85 in ipyrad for generating our data matrices. Lower values for minimum locus occupancy resulted in more loci and SNPs retained for both the full dataset (All_samples) and the subset with *A. unicolor*B samples only (A.unicolorB_only).

Phylogenetic Analysis

A summary tree depicting support values for the five designated lineages across all phylogenetic analyses is shown in Figure 3 (for individual trees see Supplemental Figures S1-6). Maximum likelihood phylogenies produced from the All_M10, All_M20, All_M30, and All_M40 datasets all support three main clades: *A. microunicolor*, *A. unicolor*A, and *A. unicolor*B (Figure 3). However, low support values (< 70) for nodes in the All_M30 and All_M40 are more frequent when compared to All_M10 and All_M20 (Figure 2 and Supplemental Figures S1-3; also see Discussion). Both All_M10 and All_M20 recovered a

similar topology except for several sister relationships at the tips but with the majority of nodes having higher support values (> 70). The data matrix All_M20 was preferred over All_M10 for downstream analyses due to All_M20 having a higher taxon coverage than All_M10 while still maintaining well-supported relationships.

The Bayesian inference phylogeny produced from the All_M20 dataset was overall well-supported with posterior probabilities ≥ 0.95 and very closely resembled the maximum likelihood All_M20 phylogeny with the exception of a few sister relationships within *A. unicolor*B (Supplemental Figure S4). Specifically, the clade comprising MY2542, MY2541, MY2806, MY2807, ANTR74, and ANTR73 is sister to the rest of *A. unicolor*B in the Bayesian analysis, whereas the clade is sister to all other samples within *A. unicolor*B3 in the maximum likelihood phylogeny. The node denoting the split between this clade and its sister group has low support for both All_M20 RAxML and Bayesian phylogenies; however, the All_M10 RAxML tree, which has the same sister group relationship as the RAxML All_M20, has a moderate bootstrap support value of 73 (Supplemental Figure S1). Therefore, the All_M20 maximum likelihood phylogeny was preferred over the Bayesian tree for downstream analyses.

The quartet-based species tree estimated in ASTRAL-III (Supplemental Figure S5) comprised 868,916,301 induced quartet trees from the gene trees, which represented 46.19% of all quartets present in the species tree. The low normalized quartet score (0.46) indicates a very high level of gene tree discordance. This species tree yielded a different topology from both the Bayesian and maximum likelihood trees (Figure 2; Supplemental Figures S4 & S5). Two of the main clades *A. microunicolor* and *A. unicolor*A remain monophyletic; however, there is significant uncertainty in the relationships between individuals within *A. unicolor*B, which reflects a high level of gene tree discordance (see Discussion). Additionally, the species tree

inferred in SNAPP yielded similar results: the split between *A. microunicolor* and *A. unicolor*A is highly supported (PP =0.98) while relationships within *A. unicolor*B have low support (PP values of 0.82 for B1 sister to B2 and 0.69 for B1 + B2 sister to B3; Supplemental Figure S6).

Genetic Exchangeability - Clustering Analyses

Three different genetic clustering analyses were used to evaluate the potential for gene flow (see Table 2 for summary). The STRUCTURE analysis with All_M20 resulted in an optimal K value of eight (**Δ**K=8; Figure 4a), with all individuals within the monophyletic lineages *A. microunicolor* and *A. unicolor*A forming separate clusters (i.e., all individuals forming a cluster correspond to the monophyletic lineages seen in Figure 2); however, *A. unicolor*B formed six clusters that did not correspond to monophyletic groups. STRUCTURE analyses with UniB-only matrices produce two to three clusters with all data sets recovering clusters that do not correspond to the monophyletic lineages with the exception of UniB_M30 data set (Figure 5). UniB_M30 does recover clusters consistent with the three monophyletic lineages within *A. unicolor*B, but some individuals have mixed cluster assignments (Figure 5). In contrast, both PCA and VAE clustering analyses (Figure 4b-d) for All-M20 indicate three distinct clusters consistent with monophyletic lineages: *A. microunicolor*, *A. unicolor*A, and *A. unicolor*B. The VAE analysis with UniB_M30 shows some separation between *A. unicolor*B lineages; however, there is still some overlap between them when considering standard deviation around samples (Figure 5). The lineages within *A. unicolor*B were not recovered as discrete clusters in any of the three clustering analyses.

Ecological Interchangeability - Niche Identity/Similarity Analyses

Table 3 summarizes the ecological interchangeability analyses. Although *A. microunicolor* has morphological and behavioral differences indicating adaptive divergence from *A. unicolor*A, the other lineage comparisons do not possess any obvious morphological or known behavioral differences (Hendrixson & Bond, 2005a, 2005b). Niche identity was rejected for all four sister lineage comparisons. However, for the niche similarity test three of the comparisons were not significantly more different or similar than background regardless of the buffer size threshold: 1) *A. microunicolor* and *A. unicolor*A; 2) *A. unicolor*B1 and *A. unicolor*B2; 3) *A. unicolor*B1 + *A. unicolor*B2 and *A. unicolor*B3. The fourth comparison between *A. microunicolor* + *A. unicolor*A and *A. unicolor*B had conflicting results with three of the four background regions (50km, 75km, and 100km) while the 25km background region indicated no niche divergence or conservatism. Comparing *A. microunicolor* + *A. unicolor*A occurrences to the background regions of *A. unicolor*B suggests niche divergence (i.e., significantly different from the null distribution), whereas comparing the occurrences of *A. unicolor*B to the background regions of *A. microunicolor* + *A. unicolor*A indicates niche conservatism (i.e., significantly more similar than the null distribution).

Discussion

3RAD - Analytical considerations/caveats

First, when assembling 3RAD data it is essential to consider the effects various parameters will have on downstream analyses, especially the locus occupancy. Studies have shown that the amount of missing data can greatly affect resulting phylogenetic inferences (e.g., Crotti et al., 2019; Eaton et al., 2016; Huang & Knowles, 2016). This result has been attributed to larger data sets comprising more phylogenetic data/signal and informative sites being excluded when a greater percentage of taxon coverage is required for a locus to be retained in the

final data matrix; therefore, quickly mutating sites become disproportionately omitted with increasing taxon coverage and exclude potentially variable and informative characters (Crotti et al., 2019).

We examined the effect of increasing the minimum sample per locus parameter for data sets containing all individuals as well as those with *A. unicolor*B samples only and found that increasing the number of samples required greatly lowered the number of SNPs retained in the final matrix (summarized in Table 1). Our phylogenetic inferences using these data sets (Figure 2; Supplementary Figures S1, S2, & S3) reflect findings of previous studies (e.g., Crotti et al., 2019) showing disproportionate loss of informative SNPs resulting in phylogenies with ambiguous or unresolved evolutionary relationships. All_M30 (30% locus occupancy) and All_M40 (40% locus occupancy) matrices contain many nodes with low support (< 50) and polytomies present; however, All_M20 and All_M10 have very similar topologies and include a majority of highly-supported (≥ 95) nodes.

Not only did the amount of missing data affect our phylogenetic inferences, but it also influenced our STRUCTURE and VAE analyses (Figure 5). We explored the effects of missing data on these clustering analyses with our *A. unicolor*B only data sets (UniB_M30, UniB_M40, UniB_M50, UniB_M60, and UniB_M70). For STRUCTURE, the number of clusters decreased from **Δ**K=3 for UniB_M30 to **Δ**K=2 for UniB_M40, UniB_M50 and UniB_M60 while UniB_M70 had a peculiar increase back to **Δ**K=3. Although UniB_M30's clusters reflect population structure (albeit with some admixture) corresponding to the three lineages within *A. unicolor*B, the data sets with **Δ**K=2 cluster B2+B3 together with varying amounts of admixture. The three clusters inferred from the UniB_M70 dataset do not necessarily reflect the structure found in UniB_M30 because several individuals are not clustering with individuals from their

lineage. This is most likely due to the greatly decreased number of informative sites to accurately detect any structure in the data.

VAE clustering across these data sets reflected the same general trend as STRUCTURE whereby increasing taxon coverage greatly decreases the amount of structure detectable in the data (Figure 5). UniB_M30 clusters had the greatest amount of structure detected for each lineage, though some overlap between them is still present. UniB_M70 has virtually no noticeable structure in the data, which most likely results from the small number and mostly uniform loci retained in the data matrix. VAE, which uses the structure inherent in the data to train the model, is likely affected by the deficiency of informative sites.

Speciation and phylogeography

Like most mygalomorphs, the *Antrodiaetus unicolor* species complex has certain life history traits (e.g., long generation times and limited dispersal capabilities; Coyle, 1971) that drastically influence population structure due to increased vulnerability of losing genetic variation by genetic drift (Laporte & Charlesworth, 2002; Whitlock & Barton, 1997). This in conjunction with extensive molecular divergence would most likely indicate that the populations have been isolated from gene flow for an extended period of time, which would increase speciation potential (Barraclough, 2019). While that seems to be the case when comparing A. *microunicolor*, *A.unicolor*A, and *A.unicolor*B, lineages within *A.unicolor*B remain ambiguous. Specifically, the STRUCTURE analyses for both All_M20 and several UniB datasets recover conflicting clusters (Figures 4a & 5; also see 3RAD section above). These clusters generated for all datasets were incongruent with monophyletic groups, which might be reflective of the use of unlinked SNPs for STRUCTURE versus using all variation for phylogenetic inference. While VAE for UniB_M30 provides some evidence of structure separating these lineages, there is still

slight overlap between them (Figure 5a). Derkarabetian et al. (2019) used STRUCTURE and VAE for species delimitation, which revealed clear VAE and STRUCTURE clusters agreeing with multiple lines of evidence (i.e., COI clades, DAPC, morphology) for species-level divergence within the harvestman *Metanonychus*, a group that is also known to have high population structure, conservative somatic morphology, with allopatry between species in the focal species group. Our study, however, includes a taxon with sympatry/parapatry across multiple lineages in addition to being morphologically homogeneous, with *A. microunicolor* being the exception. Given the more recent divergence in the *Antrodiaetus unicolor* species complex (~11.5 mya; Hendrixson & Bond, 2007) compared to *Metanonychus* (~25 mya; Derkarabetian et al., 2019), the underlying genetic patterns indicate speciation may not necessarily be complete and/or as clearly reflected in the VAE clusters.

Although the mixed cluster assignments of individuals suggest admixture between populations is occurring, it could also indicate incomplete lineage sorting (i.e., ancestral polymorphism) and/or recent range expansion (Avise, 2009). In addition, the species tree generated in ASTRAL (Supplemental Figure S5) showed considerable uncertainty for most of the relationships within *A. unicolor*B indicating an appreciable amount of gene tree discordance. There are several potential reasons for gene tree/species tree discordance (incomplete lineage sorting, reticulation, gene duplication, or horizontal gene transfer; see Maddison, 1997) but among these, incomplete lineage sorting (ILS) is the most likely explanation due to the prevalence of ILS in taxa with long generation times, large effective population sizes, and/or low mutation rates (Degnan & Rosenberg, 2009). Additionally, mtDNA data in previous analyses hinted at the possibility of a recent range expansion (Hendrixson, unpublished data; see speciation scenarios below).

All spiders in this species complex appear to have similar habitat requirements (Coyle, 1971; Hendrixson & Bond, 2005a). However, morphology, behavior, and large-scale ecological (i.e., climatic) data offered some plausible insight into potential adaptive divergence. Morphological, geographical, and molecular data clearly demarcate *A. microunicolor* and *A. unicolor*A as distinct lineages, yet the niche similarity tests between these two lineages reveals no significantly divergent or conserved niches. Two of the background areas (75km and 100km) approach significant niche divergence when comparing *A. microunicolor* occurrence points to the background area of *A. unicolor*A but not vice versa. It is possible that niche divergence important for speciation in this complex occurs in other ecological variables not tested in our study; however, our clustering analyses show minimal gene flow between these lineages and suggests a speciation model in which reproductive isolation accumulated in allopatry without much ecological differentiation (i.e., niche divergence may not have been a large driver in the speciation process). With subsequent sympatry between *A. microunicolor* and *A. unicolor*B these lineages potentially maintained reproductive isolation through premating barriers, specifically character displacement with disparities in size and breeding periods (e.g., Bond and Sierwald, 2002).

The other lineages, largely sympatric, appear morphologically and behaviorally similar, thus we use ENM ecological data. *A. microunicolor* + *A. unicolor*A and *A. unicolor*B niche similarity analyses yielded conflicting results. The comparison of *A. microunicolor* + *A. unicolor*A occurrences to the background area of *A. unicolor*B would suggest niche divergence (i.e., significantly more different compared to the null distribution) whereas the comparison of occurrences of *A. unicolor*B to the background area of *A. microunicolor* + *A. unicolor*A suggests niche conservatism (i.e., significantly more similar compared to the null distribution). Because

these results are inconclusive, we cannot rule out either niche conservatism or niche divergence. For niche divergence, one possible explanation is that *A. unicolor*B is much more widespread than both *A. microunicolor* and *A. unicolor*A and, therefore, has more potential for heterogeneous environmental conditions in the habitat available to them (e.g., more tolerance for lower elevation compared to *A. microunicolor* and *A. unicolor*A). Additionally, niche conservatism indicates *A. unicolor*B prefers similar environmental conditions to *A. microunicolor* and *A. unicolor*A (e.g, relatively mesic regions) within its broad niche background.

The niche similarity analyses involving lineages within *A. unicolor*B yielded no obvious results for either niche divergence or niche conservatism. The niche similarity analyses for *A. unicolor*B1 and *A. unicolor*B2 show that their niches are no more different or similar than expected for any of the background areas tested, although the 100km background area hinted at niche conservatism (not significant). The Little Tennessee River may be a potential geographic barrier between these two lineages facilitating reproductive isolation, but not enough time has occurred for the accumulation of neutral divergence indicative of allopatric speciation. Numerous taxa in the southern Appalachians exhibit similar phylogeographical patterns, identifying either the Asheville or Little Tennessee River Basins as potential barriers to gene flow: *Hypochilus* araneomorph spiders (Keith & Hedin, 2012), harvestmen *Sabacon cavicolens* (Hedin & McCormack, 2017) and *Fumontana deprehendor* (Thomas & Hedin, 2008), *Dasycerus carolinensis* beetles (Caterino & Langton-Myers, 2019), *Narceus* millipedes (Walker et al., 2009), and plethodontid salamanders (Kozak & Wiens, 2010; Weisrock & Larson, 2006). Consequently, these lineages could prospectively be considered distinct species after sufficient time and divergence in allopatry. Additionally, the niche similarity analyses for *A. unicolor*B1 +

*A. unicolor*B2 to *A. unicolor*B3, while not consistently significant for all background areas, does show signs of niche divergence. For the 100km and 75km background areas *A. unicolor*B3 occurrences were significantly different from the background of *A. unicolor*B1 and *A. unicolor*B2 but not quite significantly different for the reciprocal comparison. Also, the 50km background area hints at niche divergence for both comparisons. One explanation is that the ancestor of *A. unicolor*B lineages had undergone range expansion where *A. unicolor*B3 later became geographically isolated from *A. unicolor*B1 and *A. unicolor*B2 for a significant period of time generating reproductive isolation, with subsequent secondary contact potentially driving niche divergence and, therefore, reinforcing reproductive isolation in sympatry.

One caveat for using large-scale ecological data like these is the potential for not having the resolution needed for detecting microhabitat differences, which are often important for divergent selection and driving speciation (Massatti & Knowles, 2014; Soberón, 2007). Mygalomorph spiders like *Antrodiaetus* are known for habitat specialization at small scales, such as preferring north-facing slopes, shaded ravines, and particular soil types (Hedin et al., 2015; Starrett et al., 2018). Even with the best resolution available ENMs cannot necessarily detect these microhabitat similarities/differences; thus, it is possible that we are overlooking significant microhabitat differences that potentially drive niche divergence, with speciation following. Another potential issue could be how we defined the background areas. Considering that these spiders have low dispersal capabilities, it may be that the regions with higher amounts of distance incorrectly show niche divergence due to incorporating potentially uninhabitable environments and, therefore, misinterpreting possible speciation mechanisms (Warren et al., 2008).

Species limits of the *A. unicolor* species complex

Species delimitation of the *Antrodiaetus unicolor* species complex has long been challenging due to both morphological stasis and variation across their distribution (Coyle, 1971; Hendrixson & Bond, 2005a). As already discussed, *A. microunicolor* is unambiguously recognized as a species based on morphological, behavioral, and genetic divergence from *A. unicolor*. However, the now paraphyletic assemblage of *A. unicolor* lineages lacks distinctive features that could be used to distinguish them. Because no nomenclatural changes have been made to elevate these lineages to species status, the species-level diversity remains underestimated in this complex.

This study employed genome-wide SNPs for several clustering analyses and niche-based distribution modeling to evaluate genetic and ecological exchangeability, which further elucidated potential species boundaries. Clustering analyses with all individuals present (All_M20) coincide with the three species hypothesis (*A. microunicolor*, *A. unicolor*A, and *A. unicolor*B). Alternatively, the niche identity analyses and STRUCTURE/VAE analyses for *A. unicolor*B individuals only (UniB_M30) support the five species hypothesis (additional splitting of *A. unicolor*B into B1, B2, and B3). Although it is possible that these lineages were once allopatric and diverged before recent secondary contact in the Southern Appalachians, the tentative interpretation of potential incipient speciation events within *A. unicolor*B and less conservative niche identity test results remain too ambiguous for confident species delineation within the clade. Our integrative approach for delimiting species in the *A. unicolor* complex utilizing previous morphological and behavioral differences (Hendrixson & Bond 2005a) as well as our implementation of a substantial amount of molecular characters with ecological niche modeling provided ample evidence of an additional cohesion species within the complex for a total of three (*A. microunicolor*, *A. unicolor*A, and *A. unicolor*B), which potentially warrants

distinguishing these two genetically different lineages as separate species. In that case, one lineage would be designated as the true name-bearing *Antrodiaetus unicolor* and one with a new name. Considering that Hendrixson and Bond (2005a) designated a neotype for *Antrodiaetus unicolor* from Desoto State Park in northeastern Alabama and specimens included in this study from the surrounding areas cluster together in the lineage *A. unicolor*B, this lineage should be considered the true name-bearing *Antrodiaetus unicolor.* To our knowledge, none of the existing available names (junior synonyms of *A. unicolor* sensu lato) would be attributed to any of the *A. unicolor*A individuals based on geography. As such a new name will need to be proposed for *A. unicolor*A. Formal taxonomic description of *A. unicolor*A will soon be published elsewhere.

Overall, our study demonstrates the efficacy of genomic-scale data for recognizing cryptic species, suggesting that species delimitation with one data type, whether one mitochondrial gene or morphology, underestimates species diversity in taxa with low vagility and relative morphological stasis. Specifically, our clustering analyses highlight the importance gene flow constraints likely play in influencing cryptic diversity in *A. unicolor*, with speciation likely occurring due to a vicariant event isolating populations from gene flow for an extended period of time and generating reproductive isolation without much ecological differentiation (i.e., allopatric speciation). Subsequent sympatry between *A. microunicolor* and *A. unicolor*B potentially maintained reproductive isolation through premating barriers, specifically character displacement with differences in size and breeding periods (e.g., Bond & Sierwald, 2002). Incorporating multiple lines of evidence (i.e., morphological, behavioral, geographical, and ecological diversity) in addition to genomic-scale data underscores the robustness of integrative species delimitation approaches across all organismal diversity despite differences in biological or ecological characteristics. We were able to resolve species-level paraphyly within the *A.*

unicolor complex by delimiting an additional putative species despite morphological homogeneity. In addition, our study highlights another instance of cryptic speciation in the southern Appalachians, with phylogeographical patterns contributing to our understanding of the processes generating biodiversity in this rich, geologically and environmentally complex region.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Aberer, A. J., Kobert, K., & Stamatakis, A. (2014). ExaBayes: Massively parallel Bayesian tree inference for the whole-genome era. *Molecular Biology and Evolution*, *31*(10), 2553–2556. https://doi.org/10.1093/molbev/msu236

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nrg.2015.28

Avise, J. C. (2009). Phylogeography: Retrospect and prospect. *Journal of Biogeography*, *36*(1), 3–15. https://doi.org/10.1111/j.1365-2699.2008.02032.x

Bagley, J. C. (2019). *MAGNET v0.1.5* [GitHub Package]. http://github.com/justincbagley/MAGNET

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping

using sequenced RAD markers. *PLoS ONE*, *3*(10), e3376. https://doi.org/10.1371/journal.pone.0003376

Barraclough, T. (2019). *The Evolutionary Biology of Species*. Oxford University Press.

Barroso, R., Klautau, M., Solé-Cava, A. M., & Paiva, P. C. (2010). Eurythoe complanata (Polychaeta: Amphinomidae), the 'cosmopolitan' fireworm, consists of at least three cryptic species. *Marine Biology*, *157*(1), 69–80. https://doi.org/10.1007/s00227-009-1296-9

Battey, C. J., & Klicka, J. (2017). Cryptic speciation and gene flow in a migratory songbird species complex: Insights from the red-eyed vireo (*Vireo olivaceus*). *Molecular Phylogenetics and Evolution*, *113*, 67–75. https://doi.org/10.1016/j.ympev.2017.05.006

Bayona-Vásquez, N. J., Glenn, T. C., Kieran, T. J., Pierson, T. W., Hoffberg, S. L., Scott, P. A., Bentley, K. E., Finger, J. W., Louha, S., Troendle, N., Díaz-Jaimes, P., Mauricio, R., & Faircloth, B. C. (2019). Adapterama III: Quadruple-indexed, double/triple-enzyme RADseq libraries (2RAD/3RAD). *BioRxiv*. https://doi.org/10.1101/205799

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K. L., Meier, R., Winker, K., Ingram, K. K., & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in Ecology & Evolution*, *22*(3), 148–155. https://doi.org/10.1016/j.tree.2006.11.004

Bivand, R., Keitt, T., & Rowlingson, B. (2019). *rgdal: Bindings for the "geospatial" data abstraction library* [R package version 1.4-4]. . https://CRAN.R-project.org/package=rgdal

Bond, J. E., Hedin, M. C., Ramirez, M. G., & Opell, B. D. (2001). Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune

endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology*, *10*(4), 899–910. https://doi.org/10.1046/j.1365-294X.2001.01233.x

Bond, J. E., Hendrixson, B. E., Hamilton, C. A., & Hedin, M. (2012). A reconsideration of the classification of the spider Infraorder Mygalomorphae (Arachnida: Araneae) based on three nuclear genes and morphology. *PLoS ONE*, *7*(6), e38753. https://doi.org/10.1371/journal.pone.0038753

Bond, J. E., & Sierwald, P. (2002). Cryptic speciation in the *Anadenobolus excisus* millipede species complex on the island of Jamaica. *Evolution*, *56*(6), 1123–1135.

Bond, J. E., & Stockman, A. K. (2008). An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology*, *57*(4), 628–646. https://doi.org/10.1080/10635150802302443

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology*, *15*(4), e1006650. https://doi.org/10.1371/journal.pcbi.1006650

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring Species Trees Directly from Biallelic Genetic Markers: Bypassing Gene Trees in a Full Coalescent Analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Burns, M., Starrett, J., Derkarabetian, S., Richart, C. H., Cabrero, A., & Hedin, M. (2017). Comparative performance of double-digest RAD sequencing across divergent arachnid lineages. *Molecular Ecology Resources*, *17*(3), 418–430. https://doi.org/10.1111/1755-0998.12575

Butlin, R. K., Galindo, J., & Grahame, J. W. (2008). Sympatric, parapatric or allopatric: The most important way to classify speciation? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1506), 2997–3007. https://doi.org/10.1098/rstb.2008.0076

Caterino, M. S., & Langton-Myers, S. S. (2019). Intraspecific diversity and phylogeography in Southern Appalachian *Dasycerus carolinensis* (Coleoptera: Staphylinidae: Dasycerinae). *Insect Systematics and Diversity*, *3*(6). https://doi.org/10.1093/isd/ixz022

Chan, K. O., Alexander, A. M., Grismer, L. L., Su, Y.-C., Grismer, J. L., Quah, E. S. H., & Brown, R. M. (2017). Species delimitation with gene flow: A methodological comparison and population genomics approach to elucidate cryptic species boundaries in Malaysian Torrent Frogs. *Molecular Ecology*, *26*(20), 5435–5450. https://doi.org/10.1111/mec.14296

Coyle, F. A. (1971). Systematics and natural history of the mygalomorph spider genus *Antrodiaetus* and related genera (Araneae: Antrodiaetidae). *Bull. Mus. Comp. Zool*, *141*, 269–402.

Coyle, F. A. (1983). Aerial dispersal by Mygalomorph spiderlings (Araneae, Mygalomorphae). *Journal of Arachnology*, *11*(2), 283–286.

Coyle, F. A. (1985). Ballooning behavior of *Ummidia* spiderlings. (Araneae: Ctenizidae). *Journal of Arachnology*, *13*, 137–138.

Crotti, M., Barratt, C. D., Loader, S. P., Gower, D. J., & Streicher, J. W. (2019). Causes and analytical impacts of missing data in RADseq phylogenetics: Insights from an African frog (*Afrixalus*). *Zoologica Scripta*, *48*(2), 157–167. https://doi.org/10.1111/zsc.12335

De Jesús-Bonilla, V. S., Meza-Lázaro, R. N., & Zaldívar-Riverón, A. (2019). 3RAD-based systematics of the transitional Nearctic-Neotropical lubber grasshopper genus *Taeniopoda* (Orthoptera: Romaleidae). *Molecular Phylogenetics and Evolution*, *137*, 64–75. https://doi.org/10.1016/j.ympev.2019.04.019

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology & Evolution*, *24*(6), 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Delgado□Machuca, N., Meza□Lázaro, R. N., Romero□Nápoles, J., Sarmiento□Monroy, C. E., Amarillo□Suárez, Á. R., Bayona□Vásquez, N. J., & Alejandro, Z. (2019). Genetic structure, species limits and evolution of the parasitoid wasp genus *Stenocorse* (Braconidae: Doryctinae) based on nuclear 3RAD and mitochondrial data. *Systematic Entomology*. https://onlinelibrary.wiley.com/doi/abs/10.1111/syen.12373

Derkarabetian, S., Castillo, S., Koo, P. K., Ovchinnikov, S., & Hedin, M. (2019). A demonstration of unsupervised machine learning in species delimitation. *Molecular Phylogenetics and Evolution*, *139*, 106562. https://doi.org/10.1016/j.ympev.2019.106562

Derkarabetian, S., & Hedin, M. (2014). Integrative taxonomy and species delimitation in harvestmen: A revision of the Western North American genus *Sclerobunus* (Opiliones:

Laniatores: Travunioidea). *PLoS ONE*, *9*(8), e104982. https://doi.org/10.1371/journal.pone.0104982

Derkarabetian, S., Ledford, J., & Hedin, M. (2011). Genetic diversification without obvious genitalic morphological divergence in harvestmen (Opiliones, Laniatores, *Sclerobunus robustus*) from montane sky islands of western North America. *Molecular Phylogenetics and Evolution*, *61*(3), 844–853. https://doi.org/10.1016/j.ympev.2011.08.004

Dupérré, N., & Tapia, E. (2015). Discovery of the first telemid spider (Araneae, Telemidae) from South America, and the first member of the family bearing a stridulatory organ. *Zootaxa*, *4020*(1), 191. https://doi.org/10.11646/zootaxa.4020.1.9

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*(13), 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2016). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, syw092. https://doi.org/10.1093/sysbio/syw092

Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/nar/gkh340

Emerson, K. J., Merz, C. R., Catchen, J. M., Hohenlohe, P. A., Cresko, W. A., Bradshaw, W. E., & Holzapfel, C. M. (2010). Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences*, *107*(37), 16196–16200. https://doi.org/10.1073/pnas.1006538107

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*(5), 717–726. https://doi.org/10.1093/sysbio/sys004

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas: new climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Francis, R. M. (2017). pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32. https://doi.org/10.1111/1755-0998.12509

Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2016). Biodiversity and the species concept—Lineages are not enough. *Systematic Biology*, syw098. https://doi.org/10.1093/sysbio/syw098

Garg, K. M., Tizard, R., Ng, N. S. R., Cros, E., Dejtaradol, A., Chattopadhyay, B., Pwint, N., Päckert, M., & Rheindt, F. E. (2016). Genome-wide data help identify an avian species-level lineage that is morphologically and vocally cryptic. *Molecular Phylogenetics and Evolution*, *102*, 97–103. https://doi.org/10.1016/j.ympev.2016.05.028

Garrick, R. C., Newton, K. E., & Worthington, R. J. (2018). Cryptic diversity in the southern Appalachian Mountains: Genetic data reveal that the red centipede, *Scolopocryptops*

sexspinosus, is a species complex. *Journal of Insect Conservation*, *22*(5–6), 799–805. https://doi.org/10.1007/s10841-018-0107-3

Garrison, N. L., Brewer, M. S., & Bond, J. E. (in press). Shifting evolutionary sands: Transcriptome characterization of the *Aptostichus atomarius* species complex. *BMC Evolutionary Biology*, 1–12. https://doi.org/10.1186/s12862-020-01606-7

Graham, C. F., Glenn, T. C., McArthur, A. G., Boreham, D. R., Kieran, T., Lance, S., Manzon, R. G., Martino, J. A., Pierson, T., Rogers, S. M., Wilson, J. Y., & Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, *15*(6), 1304–1315. https://doi.org/10.1111/1755-0998.12404

Hamilton, C. A., Hendrixson, B. E., Brewer, M. S., & Bond, J. E. (2014). An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: A case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution*, *71*, 79–93. https://doi.org/10.1016/j.ympev.2013.11.007

Harvey, M. S., Hillyer, M. J., Main, B. Y., Moulds, T. A., Raven, R. J., Rix, M. G., Vink, C. J., & Huey, J. A. (2018). Phylogenetic relationships of the Australasian open-holed trapdoor spiders (Araneae: Mygalomorphae: Nemesiidae: Anaminae): multi-locus molecular analyses resolve the generic classification of a highly diverse fauna. *Zoological Journal of the Linnean Society*, *184*(2), 407–452. https://doi.org/10.1093/zoolinnean/zlx111

Hedin, M., Carlson, D., & Coyle, F. A. (2015). Sky island diversification meets the multispecies coalescent—Divergence in the spruce-fir moss spider (*Microhexura montivaga,*

Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. *Molecular Ecology*, *24*(13), 3467–3484. https://doi.org/10.1111/mec.13248

Hedin, M., Derkarabetian, S., Alfaro, A., Ramírez, M. J., & Bond, J. E. (2019). Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci. *PeerJ*, *7*, e6864. https://doi.org/10.7717/peerj.6864

Hedin, M., & McCormack, M. (2017). Biogeographical evidence for common vicariance and rare dispersal in a southern Appalachian harvestman (Sabaconidae, *Sabacon cavicolens*). *Journal of Biogeography*, *44*(7), 1665–1678. https://doi.org/10.1111/jbi.12973

Hedin, M., Starrett, J., & Hayashi, C. (2013). Crossing the uncrossable: Novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. *Molecular Ecology*, *22*(2), 508–526. https://doi.org/10.1111/mec.12130

Hendrixson, B. E., & Bond, J. E. (2005a). Two sympatric species of *Antrodiaetus* from southwestern North Carolina (Araneae, Mygalomorphae, Antrodiaetidae). *Zootaxa*, *872*(1), 1. https://doi.org/10.11646/zootaxa.872.1.1

Hendrixson, B. E., & Bond, J. E. (2005b). Testing species boundaries in the *Antrodiaetus unicolor* complex (Araneae: Mygalomorphae: Antrodiaetidae): "Paraphyly" and cryptic diversity. *Molecular Phylogenetics and Evolution*, *36*(2), 405–416. https://doi.org/10.1016/j.ympev.2005.01.021

Hendrixson, B. E., & Bond, J. E. (2007). Molecular phylogeny and biogeography of an ancient Holarctic lineage of mygalomorph spiders (Araneae: Antrodiaetidae: *Antrodiaetus*).

Molecular Phylogenetics and Evolution, *42*(3), 738–755. https://doi.org/10.1016/j.ympev.2006.09.010

Hendrixson, B. E., DeRussy, B. M., Hamilton, C. A., & Bond, J. E. (2013). An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, *Aphonopelma*). *Molecular Phylogenetics and Evolution*, *66*(1), 327–340. https://doi.org/10.1016/j.ympev.2012.10.004

Hey, J. (2001). The mind of the species problem. *Trends in Ecology & Evolution*, *16*(7), 326–329.

Hijmans, R. J. (2015). raster: Geographic data analysis and modeling [R package version 2.5-2]. https://cran.r-project.org/package=raster

Hoffberg, S. L., Kieran, T. J., Catchen, J. M., Devault, A., Faircloth, B. C., Mauricio, R., & Glenn, T. C. (2016). RADcap: Sequence capture of dual-digest RADseq libraries with identifiable duplicates and reduced missing data. *Molecular Ecology Resources*, *16*(5), 1264–1278. https://doi.org/10.1111/1755-0998.12566

Holland, B. S., Dawson, M. N., Crow, G. L., & Hofmann, D. K. (2004). Global phylogeography of *Cassiopea* (Scyphozoa: Rhizostomeae): molecular evidence for cryptic species and multiple invasions of the Hawaiian Islands. *Marine Biology*, *145*(6), 1119–1128. https://doi.org/10.1007/s00227-004-1409-4

Huang, H., & Knowles, L. L. (2016). Unforeseen consequences of excluding missing data from next-generation sequences: Simulation study of RAD sequences. *Systematic Biology*, *65*(3), 357–365. https://doi.org/10.1093/sysbio/syu046

Huey, J. A., Hillyer, M. J., & Harvey, M. S. (2019). Phylogenetic relationships and biogeographic history of the Australian trapdoor spider genus *Conothele* (Araneae:

Mygalomorphae: Halonoproctidae): diversification into arid habitats in an otherwise tropical radiation. *Invertebrate Systematics*. https://doi.org/10.1071/IS18078

Jezkova, T., Olah-Hemmings, V., & Riddle, B. R. (2011). Niche shifting in response to warming climate after the last glacial maximum: Inference from genetic data and niche assessments in the chisel-toothed kangaroo rat (*Dipodomys microps*). *Global Change Biology*, *17*(11), 3486–3502. https://doi.org/10.1111/j.1365-2486.2011.02508.x

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., & Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*, *27*(21), 3070–3071. https://doi.org/10.1093/bioinformatics/btr521

Keith, R., & Hedin, M. (2012). Extreme mitochondrial population subdivision in southern Appalachian paleoendemic spiders (Araneae: Hypochilidae: *Hypochilus*), with implications for species delimitation. *Journal of Arachnology*, *40*(2), 167–181.

Kozak, K. H., & Wiens, J. J. (2010). Niche conservatism drives elevational diversity patterns in Appalachian salamanders. *The American Naturalist*, *176*(1), 40–54. https://doi.org/10.1086/653031

Laporte, V., & Charlesworth, B. (2002). Effective population size and population subdivision in demographically structured populations. *Genetics*, *162*, 501–519.

Leaché, A. D., Chavez, A. S., Jones, L. N., Grummer, J. A., Gottscho, A. D., & Linkem, C. W. (2015). Phylogenomics of Phrynosomatid lizards: Conflicting signals from sequence capture versus restriction site associated DNA sequencing. *Genome Biology and Evolution*, *7*(3), 706–719. https://doi.org/10.1093/gbe/evv026

Lemmon, A. R., Emme, S. A., & Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, *61*(5), 727–744. https://doi.org/10.1093/sysbio/sys049

Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology*, *46*(3), 523–536.

Massatti, R., & Knowles, L. L. (2014). Microhabitat differences impact phylogeographic concordance of codistributed species: Genomic evidence in montane sedges (*Carex*) from the Rocky Mountains. *Evolution*, *68*(10), 2833–2846. https://doi.org/10.1111/evo.12491

Moritz, C., Fujita, M. K., Rosauer, D., Agudo, R., Bourke, G., Doughty, P., Palmer, R., Pepper, M., Potter, S., Pratt, R., Scott, M., Tonione, M., & Donnellan, S. (2016). Multilocus phylogeography reveals nested endemism in a gecko across the monsoonal tropics of Australia. *Molecular Ecology*, *25*(6), 1354–1366. https://doi.org/10.1111/mec.13511

Opatova, V., & Arnedo, M. A. (2014a). From Gondwana to Europe: Inferring the origins of Mediterranean *Macrothele* spiders (Araneae: Hexathelidae) and the limits of the family Hexathelidae. *Invertebrate Systematics*, *28*(4), 361. https://doi.org/10.1071/IS14004

Opatova, V., & Arnedo, M. A. (2014b). Spiders on a hot volcanic roof: Colonisation pathways and phylogeography of the Canary Islands endemic trapdoor spider *Titanidiops canariensis* (Araneae, Idiopidae). *PLoS ONE*, *9*(12), e115078. https://doi.org/10.1371/journal.pone.0115078

Opatova, V., Hamilton, C. A., Hedin, M., De Oca, L. M., Král, J., & Bond, J. E. (2019). Phylogenetic systematics and evolution of the spider infraorder Mygalomorphae using genomic scale data. *Systematic Biology*, *syz064*. https://doi.org/10.1093/sysbio/syz064

Ortega-Andrade, H. M., Rojas-Soto, O. R., Valencia, J. H., Espinosa de los Monteros, A., Morrone, J. J., Ron, S. R., & Cannatella, D. C. (2015). Insights from integrative

systematics reveal cryptic diversity in *Pristimantis* frogs (Anura: Craugastoridae) from the Upper Amazon Basin. *PLOS ONE*, *10*(11), e0143392. https://doi.org/10.1371/journal.pone.0143392

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5), e37135. https://doi.org/10.1371/journal.pone.0037135

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*(3–4), 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 949–959.

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

R. L. Mayden. (1997). A hierarchy of species concepts: The denouement in the saga of the species problem. In *Species: The units of biodiversity* (pp. 381–424). Chapman & Hall.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Ramírez-Reyes, T., Piñero, D., Flores-Villela, O., & Vázquez-Domínguez, E. (2017). Molecular systematics, species delimitation and diversification patterns of the *Phyllodactylus lanei* complex (Gekkota: Phyllodactylidae) in Mexico. *Molecular Phylogenetics and Evolution*, *115*, 82–94. https://doi.org/10.1016/j.ympev.2017.07.008

Reynolds, R. G., Niemiller, M. L., & Fitzpatrick, B. M. (2012). Genetic analysis of an endemic archipelagic lizard reveals sympatric cryptic lineages and taxonomic discordance. *Conservation Genetics*, *13*(4), 953–963. https://doi.org/10.1007/s10592-012-0344-z

Richardson, B. J. (2016). New genera, new species and redescriptions of Australian jumping spiders (Araneae: Salticidae). *Zootaxa*, *4114*(5), 501. https://doi.org/10.11646/zootaxa.4114.5.1

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Satler, J. D., Starrett, J., Hayashi, C. Y., & Hedin, M. (2011). Inferring species trees from gene trees in a radiation of California trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus*). *PLoS ONE*, *6*(9), e25355. https://doi.org/10.1371/journal.pone.0025355

Schoener, T. W. (1968). The *Anolis* lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, *49*(4), 704–726. https://doi.org/10.2307/1935534

Soberón, J. (2007). Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, *10*(12), 1115–1123. https://doi.org/10.1111/j.1461-0248.2007.01107.x

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Starrett, J., Hayashi, C. Y., Derkarabetian, S., & Hedin, M. (2018). Cryptic elevational zonation in trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus janus* complex) from the California southern Sierra Nevada. *Molecular Phylogenetics and Evolution*, *118*, 403–413. https://doi.org/10.1016/j.ympev.2017.09.003

Starrett, J., & Hedin, M. (2007). Multilocus genealogies reveal multiple cryptic species and biogeographical complexity in the California turret spider *Antrodiaetus riversi* (Mygalomorphae, Antrodiaetidae): Cryptic diversification in Californian *A. riversi*. *Molecular Ecology*, *16*(3), 583–604. https://doi.org/10.1111/j.1365-294X.2006.03164.x

Stockman, A. K., & Bond, J. E. (2007). Delimiting cohesion species: Extreme population structuring and the role of ecological interchangeability. *Molecular Ecology*, *16*(16), 3374–3392. https://doi.org/10.1111/j.1365-294X.2007.03389.x

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Swets, J. (1988). Measuring the accuracy of diagnostic systems. *Science*, *240*(4857), 1285–1293. https://doi.org/10.1126/science.3287615

Templeton, A. R. (1989). The meaning of species and speciation: A genetic perspective. In *Speciation and its Consequences* (pp. 3–27). Sinauer.

Thomas, S. M., & Hedin, M. (2008). Multigenic phylogeographic divergence in the paleoendemic southern Appalachian opilionid *Fumontana deprehendor* Shear (Opiliones, Laniatores, Triaenonychidae). *Molecular Phylogenetics and Evolution*, *46*(2), 645–658. https://doi.org/10.1016/j.ympev.2007.10.013

Wake, D. B. (2006). Problems with species: Patterns and processes of species formation in salamanders. *Annals of the Missouri Botanical Garden*, *93*(1), 8–23.

Walker, M. J., Stockman, A. K., Marek, P. E., & Bond, J. E. (2009). Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus *Narceus*:

Evidence from population genetic, phylogeographic, and paleoclimatic data. *BMC Evolutionary Biology*, *9*(1), 25. https://doi.org/10.1186/1471-2148-9-25

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, *62*(11), 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x

Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*. https://doi.org/10.1111/j.1600-0587.2009.06142.x

Weisrock, D. W., & Larson, A. (2006). Testing hypotheses of speciation in the *Plethodon jordani* species complex with allozymes and mitochondrial DNA sequences: Diagnosing plethodontid salamander species. *Biological Journal of the Linnean Society*, *89*(1), 25–51. https://doi.org/10.1111/j.1095-8312.2006.00655.x

Whitlock, M. C., & Barton, N. H. (1997). The effective size of a subdivided population. *Genetics*, *146*, 427–441.

Wiens, J. J., & Graham, C. H. (2005). Niche conservatism: Integrating evolution, ecology, and conservation biology. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 519–539. https://doi.org/10.1146/annurev.ecolsys.36.102803.095431

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*(S6). https://doi.org/10.1186/s12859-018-2129-y

Zhang, Y., & Li, S. (2014). A spider species complex revealed high cryptic diversity in South China caves. *Molecular Phylogenetics and Evolution*, *79*, 353–358. https://doi.org/10.1016/j.ympev.2014.05.017

Tables

Table 1. Summary of the data obtained in matrices with an 85% clustering threshold. Matrix names refer to the individuals included (All or UnicolorB only), M meaning the minimum amount of samples needed for a locus to be retained in the dataset, and the number is the locus occupancy percentage (percentage is calculated by the minimum sample per locus divided by individuals included). Minimum sample per locus parameter refers to the minimum number of samples with a particular locus that is needed to be included in the final data set. SNPs refers to the total number of SNPs in the data set, with Var referring to the number of variable sites and Pis referring to the number of parsimony-informative sites.

Matrix Individuals	Matrix Names	Minimum sample per locus parameter	Total # of loci after filtering	Mean # of loci per sample	SNPs - Var	SNPs - Pis	Mean # of SNPs (Pis) per locus
All individuals (157 individuals)	All_M10	16	9561	1633	260312	105737	11.06
	All_M20	33	1719	576	56130	24348	14.16
	All_M30	49	574	285	18615	8081	14.08
	All_M40	66	290	186	9042	3752	12.94
	All_M50	82	189	84	5620	2287	12.1
	All_M60	99	145	68	4306	1754	12.1
	All_M70	115	100	50	3035	1248	12.48
	All_M80	132	56	27	1777	716	12.79
*A.unicolor*B-only (120 individuals)	UniB_M10	12	15651	3207	372129	154859	9.89
	UniB_M20	24	5384	1844	164350	71112	12.19

UniB_M30	36	2408	1133	78778	34803	14.45
UniB_M40	48	1277	750	41952	18561	14.53
UniB_M50	60	781	530	25070	10897	13.95
UniB_M60	72	512	385	15877	6834	13.35
UniB_M70	84	325	265	9732	4108	12.64
UniB_M80	96	202	174	6108	2548	12.61
UniB_M90	108	41	38	1221	505	12.32

Table 2. Summary of the genetic exchangeability (GE) criterion for cohesion species delimitation between populations of the *A. unicolor* species complex.

Lineage Comparison	Geographic Barrier	STRUCTURE	PCA	VAE	Conclusions
*A. unicolor*A to *A. microunicolor*	Allopatric	Separate clusters	Separate clusters	Separate clusters	Reject GE
*A. unicolor*A+*A. microunicolor* to *A. unicolor*B	None/parapatric	Separate clusters	Separate clusters	Separate clusters	Reject GE
*A. unicolor*B1 to *A. unicolor*B2	None/parapatric	admixture	Cluster overlap	Cluster overlap	Fail to reject GE
*A. unicolor*B1+*A. unicolor*B2 to *A. unicolor*B3	None/parapatric	admixture	Cluster overlap	Cluster overlap	Fail to reject GE

Table 3. Summary of the ecological interchangeability (EI) criterion for cohesion species delimitation between populations of the *A. unicolor* species complex. We evaluated a proxy for ecological interchangeability by measuring the overlap between ENMs for the sister lineages comparisons as well as taking into account previous morphological and behavioral criteria for distinguishing *A. microunicolor.* N_a and N_b values are the number of occurrence records for the first and second lineages used in a comparison, respectively. Niche identity refers to the "niche equivalency test", and niche similarity refers to the "niche similarity (i.e., background) test" proposed by Warren et al. (2008) (see Methods for details). [a] ENMs are no more similar or

different than expected by chance; [b] ENMs are more similar than expected by chance for one comparison, yet ENMs are more different than expected by chance for the reciprocal comparison; see Discussion.

Lineage Comparison	Morphology/ Behavior Differences	N_a, N_b	Niche Overlap Value, Niche Identity, Niche Similarity	Conclusions
*A. unicolor*A to *A. microunicolor*	Size, setal characters, non-overlapping mating season	6, 20	0.2964, $P < 0.05$, $P > 0.05$[a]	Reject EI
*A. unicolor*A+*A. microunicolor* to *A. unicolor*B	No obvious differences	26, 82	0.5899, $P < 0.05$, $P < 0.05$[b]	Fail to reject EI
*A. unicolor*B1 to *A. unicolor*B2	No obvious differences	29, 37	0.2041, $P < 0.05$, $P > 0.05$[a]	Fail to reject EI
*A. unicolor*B1+*A. unicolor*B2 to *A. unicolor*B3	No obvious differences	66, 16	0.7566, $P < 0.05$, $P > 0.05$[a]	Fail to reject EI

Figures

Figure 1. Geographic distributions of each *A. unicolor* lineage as defined in the Methods. See the legend denoting color and symbol for each lineage.

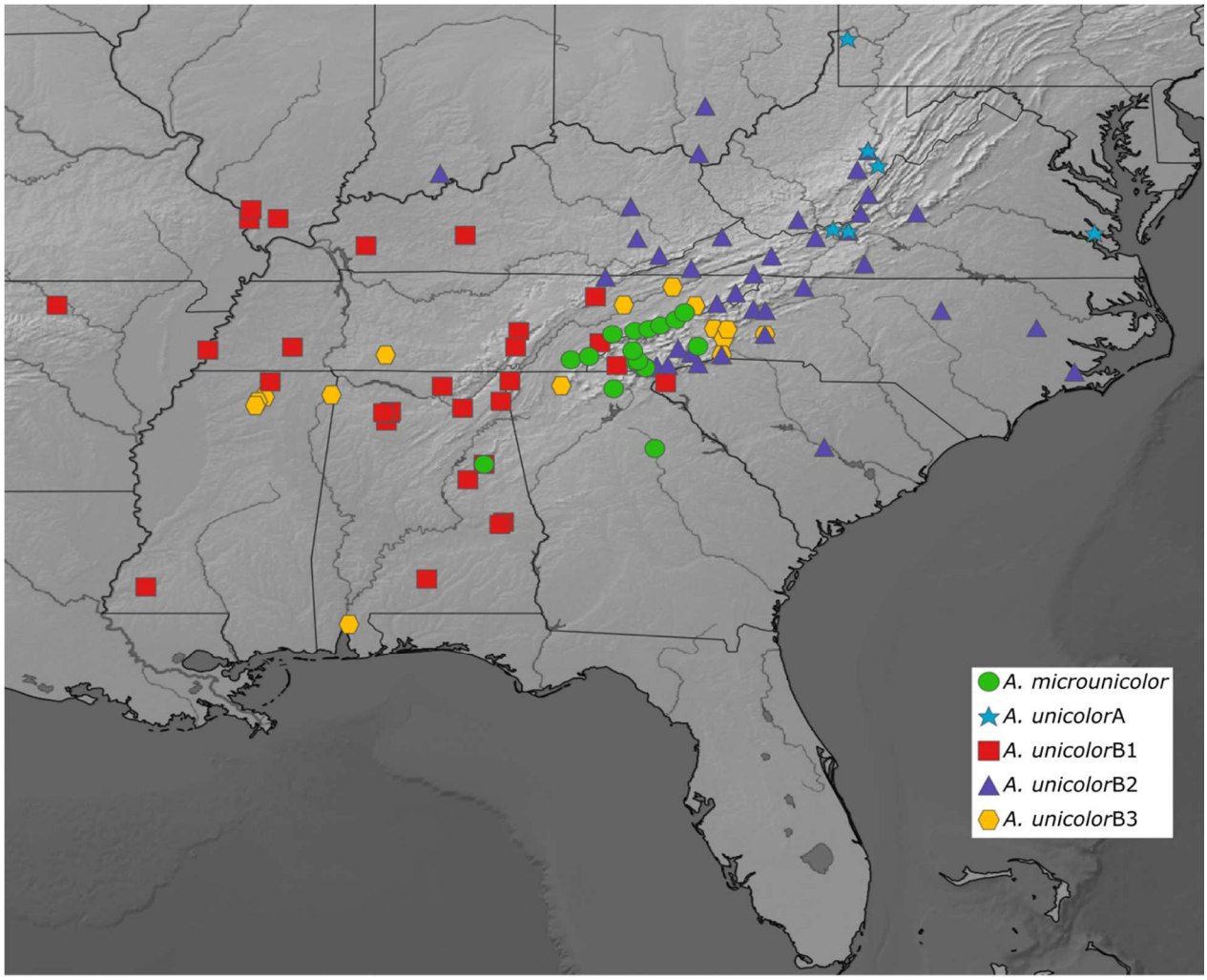


Figure 2. All_M20 RAxML phylogeny with midpoint rooting. Although trees typically depict highly supported nodes, we denote nodes with support < 90 with black dots since there are far fewer nodes with lower support. Each lineage has been labeled and highlighted with color scheme from Figure 1. Bottom left: male (top spider) and female (bottom spider) *Antrodiaetus unicolor* specimens.

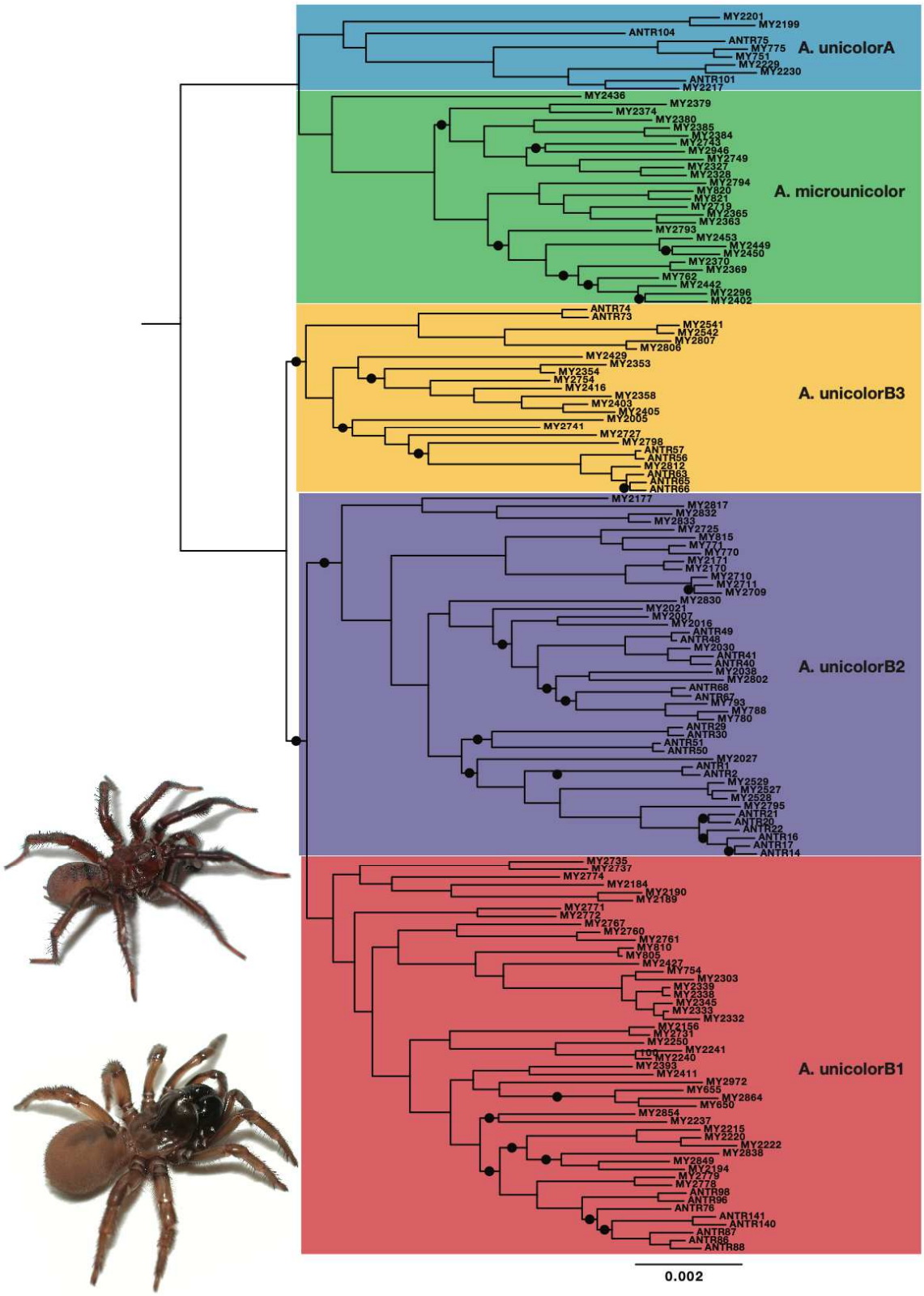

A. unicolorA
A. microunicolor
A. unicolorB3
A. unicolorB2
A. unicolorB1
0.002

Figure 3. Species tree from SNAPP analyses for the designated lineages. Boxes on nodes represent support values attained in each approach (left to right): RAxML bootstrap support across four datasets (All_M10, All_M20, All_M30, and All_M40), ExaBayes Bayesian posterior probabilities, ASTRAL support values, SNAPP Bayesian posterior probabilities. Color coding of boxes corresponds to discrete ranges of support values as seen in the bottom left corner.

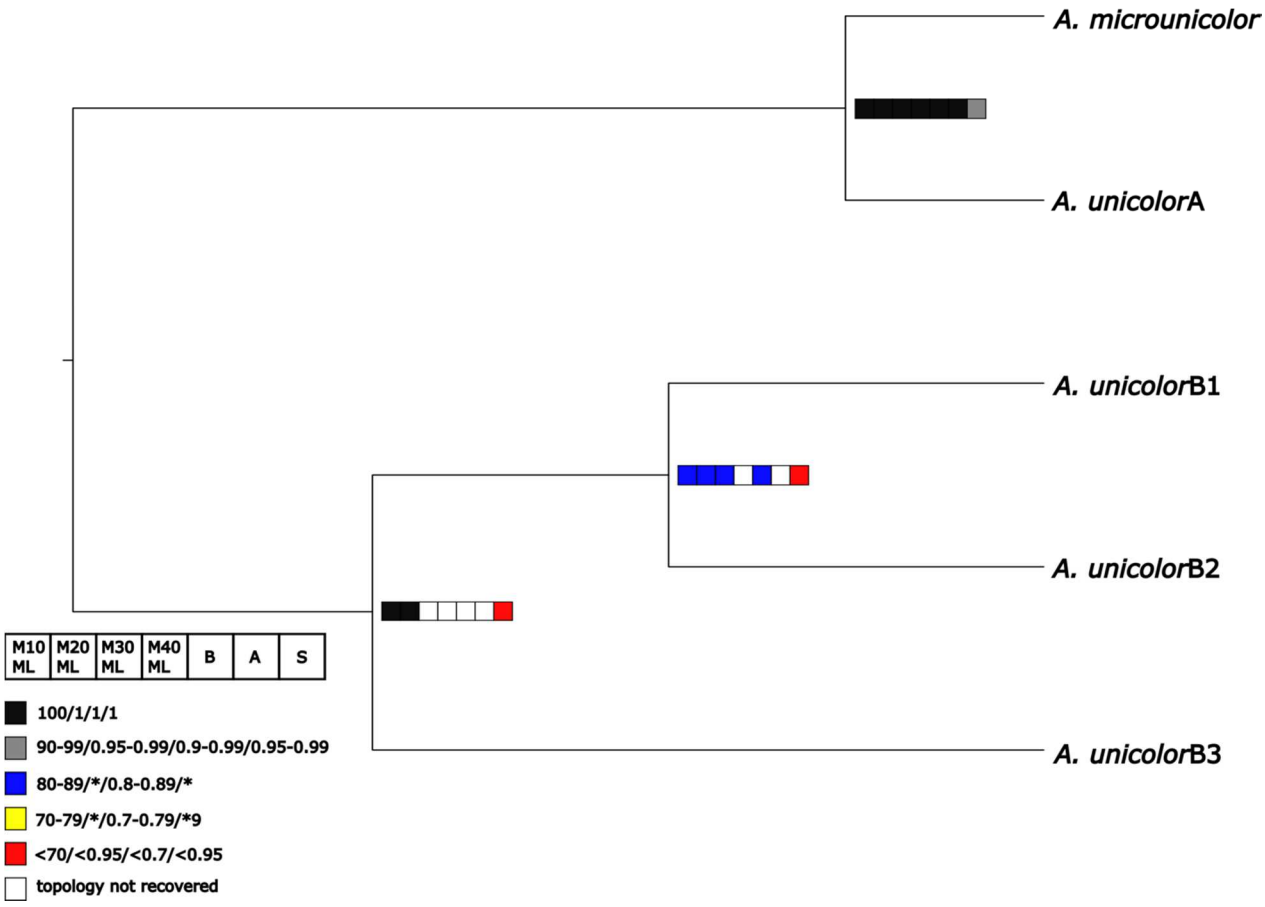


Figure 4. Cluster analysis plots of the All_M20 dataset with the same color scheme for each lineage as previous figures unless otherwise stated. a) STRUCTURE plot with an optimal K value of 8 (**Δ**K=8), showing a large amount of admixture. *A. microunicolor* and *A. unicolor*A retain their color from previous figures since each cluster retains all lineage individuals while *A.*

*unicolor*B does not due to additional clusters not corresponding to our designated *A. unicolor*B (B1, B2, and B3) lineages (see text labels for cluster coloration). b) PCA plot, showing significant cluster overlap of *A. unicolor*B lineages. c) VAE plot, displaying no clear distinctions among *A. unicolor*B lineages and d) VAE results with encoded mean (μ) and standard deviation (σ) for each sample.

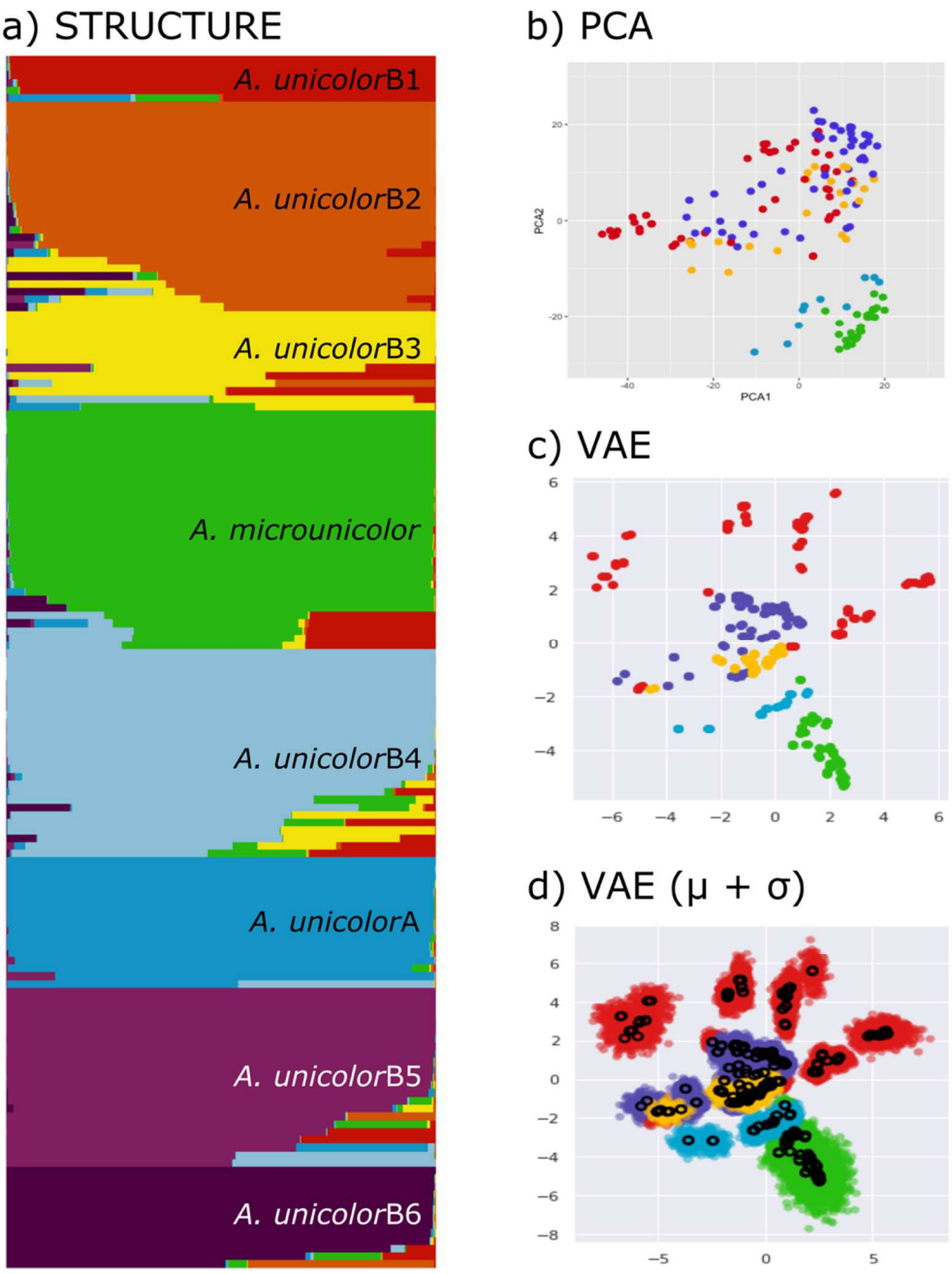
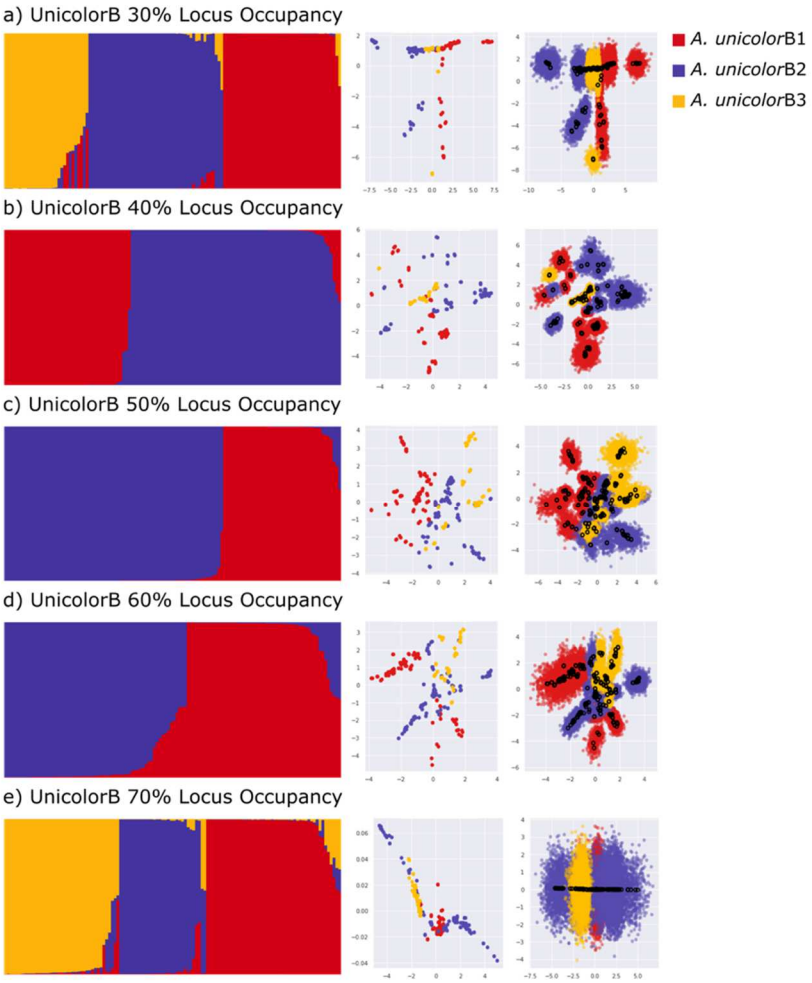


Figure 5. STRUCTURE and VAE analyses using the datasets that include only *A. unicolor*B individuals with varying locus occupancies (UniB_M30, UniB_M40, UniB_M50, UniB_M60, and UniB_M70; see Table 1 for dataset details). The color scheme for each *A. unicolor*B lineage is the same as previous figures (see legend for coloration) except for the 40%, 50%, and 60% locus occupancy STRUCTURE plots (**Δ**K=2): *A. unicolor*B2 and *A. unicolor*B3 individuals cluster together as one group and is represented by *A. unicolor*B2's coloration. a) 30% locus occupancy. b) 40% locus occupancy. c) 50% locus occupancy. d) 60% locus occupancy. e) 70% locus occupancy.



CHAPTER II

Integrative species delimitation and phylogeography of California endemic trapdoor spiders within the *Aptostichus icenoglei* sibling species complex (Araneae: Mygalomorphae: Euctenizidae)

Introduction

Delimiting species is a fundamental first step in establishing a foundation upon which all biological inquiries (e.g., assessing physiological pathways or extinction risk) are evaluated. Yet, the conceptual definition of what constitutes a species and the relative importance of the various processes by which new species arise continue to be debatable and contentious topics of discussion (de Queiroz, 2007; Hey, 2001; Wells et al., 2021). Species concepts typically emphasize disparate intrinsic biological properties (e.g., morphological differences, niche divergence, genetic divergence) that are differentially important with respect to species recognition and/or speciation process. Depending on the taxon and the particular point in the speciation process, various concepts may be incompatible or delimit species in different ways; that is, for the same set of populations, one concept may recognize multiple distinct species whereas another may lump them together (de Queiroz, 2007).

Assessing species limits is particularly difficult in taxa with limited dispersal capabilities when reduced gene flow leads to high levels of population structuring. Taxa with high levels of genetic divergence and no gene flow can sometimes lead to speciation in the absence of notable morphological differentiation, essentially obscuring otherwise cryptic species boundaries. Specifically, non-vagile taxa are closely tied to the landscape, such that as geological and topographical changes occur over time, populations become geographically isolated with

severely limited opportunity for gene flow (e.g., Bond et al., 2001; Derkarabetian et al., 2021; Starrett & Hedin, 2007; Weisrock & Larson, 2006). As these populations remain spatially isolated over relatively long periods of time and accumulate random mutations, genetic divergence builds through genetic drift and/or natural selection for adaptive alleles in population(s) that inhabit newly available niche space. When spatial isolation is coupled with occupation of new niche space, one would expect each population to not only exhibit genetic divergence but also morphological, behavioral, and/or physiological differences (Freudenstein et al., 2016). However, if genetically diverged populations remain stationary in niche space (i.e., niche conservatism; Wiens & Graham, 2005), then it would be plausible for morphological stasis to occur in the absence of differing selective pressures, resulting in genetic lineages that are morphologically indistinguishable (Bond et al., 2001; Cerca et al., 2021; Mas-Peinado et al., 2018). In that case, it is likely that species diversity will be underestimated because traditional approaches that primarily apply morphological distinctiveness are what is typically applied in species delimitation (Bond et al., 2021). Thus, implementing a species concept focusing on one biological property/data type could potentially misrepresent the actual number of species present if that property was not important in the speciation process (Abbott et al., 2013; de Queiroz, 2007).

The species concept applied in a given system has severe implications for downstream delimitation decisions and thus must be explicitly stated in any species delimitation study. As already alluded to above, a species concept that focuses strictly on morphological differentiation has the potential to overlook cryptic species that may otherwise be genetically diverged to the point that genomic incompatibilities preclude gene flow (Barroso et al., 2010; Battey & Klicka, 2017; Holland et al., 2004; Weisrock & Larson, 2006). Alternatively, molecular approaches to

species delimitation have been shown to overestimate species diversity when local population structuring is viewed as 'species divergence'. Specifically, single-locus approaches such as DNA barcoding along with GMYC, as well as multiple-locus approaches (e.g., multispecies coalescent methods) that assume panmixia are prone to identifying population structure as opposed to speciation events (Hamilton et al., 2014; Hedin et al., 2015; Sukumaran & Knowles, 2017). In such systems localized divergence of neutral alleles may be inconsequential when populations come into secondary contact, so any species delimitation approach that relies primarily on genetic differentiation (e.g., general lineage species concept) will vastly overestimate species diversity when applying these methods. Consequently, a species concept that incorporates multiple biological properties as an integrative species delimitation approach that weighs evidence from multiple independent sources is likely to be more successful at identifying true evolutionary species diversity.

A largely overlooked species concept, the Cohesion Species Concept (CSC; Templeton, 1989), has arguably already solved the problems of too little versus too much gene flow and, unlike the unified species concept (de Queiroz, 2007), provides the hypothetical and conceptual foundation for framing integrative species delimitation. The CSC posits that a cohesion species must constitute an independently evolving evolutionary lineage with populations that are genetically exchangeable and/or ecologically interchangeable (Templeton, 1989). This concept can be applied to essentially all taxa, integrates multiple biological properties that are potentially important in the speciation process, and provides a methodological framework in which species hypotheses can be tested (Barraclough, 2019; Bond & Stockman, 2008; Templeton, 1989; Wells et al., 2021). Thus, it is particularly useful when evaluating species boundaries in morphologically homogenous taxa prone to cryptic diversity in conjunction with a high amount

of population structuring at small spatial scales (Bond & Stockman, 2008; Hendrixson et al., 2013, 2015; Newton et al., 2020). Our primary focus in this paper will be the explicit implementation of the CSC to a species delimitation problem in a previously characterized lineage of trapdoor spiders in the genus *Aptostichus*. The *Aptostichus icenoglei* sibling species complex contains three species that can be easily differentiated using approaches with standard taxonomic characters, yet two of these species are geographically widespread and known to be genetically distinct (Bond, 2012). The questions we pose are related first to genetic exchangeability – do these populations constitute distinct genetic lineages (i.e., no evidence of gene flow between lineages), and if so, are they ecologically interchangeable or not? If these genetically distinct lineages are ecologically interchangeable, then the unavoidably subjective question arises of how heavily one weights genetic divergence versus ecological/adaptive divergence, or the lack thereof (discussed below).

Mygalomorphae, the spider infraorder containing tarantulas and trapdoor spiders, are notorious for being morphologically static relative to the other, more diverse spider groups placed in the sister infraorder Araneomorphae (Opatova et al., 2019). In addition, having relatively long lifespans along with limited dispersal capabilities make mygalomorphs more vulnerable to significant population genetic structuring at small spatial scales (Bond & Stockman, 2008; Cooper et al., 2011; Hamilton et al., 2014; Harvey et al., 2018; Starrett & Hedin, 2007), thus underscoring the interplay of genetic versus ecological interchangeability when evaluating divergence at the species/population interface in these highly structured taxa. Considering that species delimitation in spiders has primarily relied on morphological differentiation (see Bond et al., 2021 for summary of spider species decriptions from 2008-2018), delimiting mygalomorph spider species using these morphological criteria has led to the

perception that species diversity has been vastly underestimated. In contrast, strong population genetic structuring often leads to an overestimation of mygalomorph species diversity when only using molecular species delimitation methods (e.g., DNA barcoding, GYMC, BPP; Hamilton et al., 2014; Hedin et al., 2015; Hendrixson et al., 2013). As a result, only using morphology or molecules as the data type for species delimitation criteria likely leads to an incomplete evaluation of species boundaries; therefore, integrating multiple independent data types is vital for a robust species delimitation framework.

The focus of our study is the *Aptostichus icenoglei* sibling species complex, which comprises three species that are all allopatric (*A. icenoglei*, *A. isabella,* and *A. barackobamai*; Bond, 2012). Like other *Aptostichus* species, they construct thin wafer trapdoors from silk and the surrounding substrate; they are geographically widespread throughout three regions in the California Floristic Province (CAFP), a known biodiversity hotspot (Bond, 2012; Myers et al., 2000). The CAFP has a complex geological, climatic, and topographical history which has highly influenced the speciation pattern and process of many plants (e.g., Anacker et al., 2011; Baldwin et al., 2011; Cole et al., 2011; Eckert et al., 2008; Grivet et al., 2006; Kraft et al., 2010; Liston et al., 2007; Rundel, 2011) and animals (Alexander & Burns, 2006; Chatzimanolis & Caterino, 2007; Leaché et al., 2009; Oliver & Shapiro, 2007; Pardikes et al., 2017; Rios & Álvarez-Castañeda, 2010; Sgariglia & Burns, 2003; Spinks & Shaffer, 2005; Vandergast et al., 2006), especially non-vagile taxa such as salamanders (Jockusch et al., 2020; Martínez-Solano et al., 2007; Wake, 1997), harvestmen (Emata & Hedin, 2016), scorpions (Bryson et al., 2016), and mygalomorph spiders (Bond, 2012; Bond & Stockman, 2008; Hedin et al., 2013; Leavitt et al., 2015; Satler et al., 2011). Dispersal-limited taxa have proven to be particularly useful in broadening our understanding of the historical biogeography of the CAFP (Emata & Hedin,

2016; Hedin et al., 2013; Martínez-Solano et al., 2007). Evolutionary divergence, influenced by barriers to dispersal either because of biotic (e.g., competition or predation) or abiotic factors (e.g., geologic, geographic, or environmental factors), can be "remarkably fine-grained spatially and can extend to shallow timescales" for these low-dispersal taxa (Hedin et al., 2013). The combination of dispersal-limited taxa generally being relatively morphologically homogenous yet having significant genetic divergence suggests the primary mode of divergence would be influenced by vicariance events, such as geological activity creating barriers to gene flow, as opposed to adaptive divergence (e.g., niche divergence through competition). Thus, evidence for biogeographical barriers remains intact in these systems for longer time periods and can potentially reveal multiple barriers to dispersal (i.e., both long-term and short-term barriers; Hedin et al., 2013; Martínez-Solano et al., 2007), so one can expect patterns seen in genetic variation of low-dispersal organisms to closely reflect the geological history of the region in which they are distributed.

Aptostichus barackobamai and *A. icenoglei* are relatively widespread and exhibit evidence of cryptic diversity (i.e., morphologically similar yet found in a variety of habitats across a sizable geographic range) found in other mygalomorph groups (e.g., Hamilton et al., 2014; Hendrixson et al., 2013; Hendrixson & Bond, 2005; Starrett et al., 2018; Starrett & Hedin, 2007) as well as other *Aptostichus* species (Bond et al., 2001; Bond & Stockman, 2008). *Aptostichus isabella* on the other hand, is only known from one specimen collected from the type locality near Lake Isabella in the southern Sierran foothills. *Aptostichus icenoglei* is widely distributed throughout the Transverse Ranges from the Los Angeles Basin to the Santa Ana/San Jacinto Mountains as well as the mountains and hills surrounding San Diego (Bond, 2012). The primary habitat types for *A. icenoglei* include coastal chaparral forest and coastal range open

woodland shrub and coniferous forest (Bond, 2012). *Aptostichus barackobamai* is also relatively widespread in primarily mixed redwood and coniferous forests in the northern Coastal Ranges as well as along the northern rim of the Central Valley, with one population in the Sutter Buttes (Bond, 2012). Altogether, these likely represent a diversity of habitat types spread across a number of different California ecoregions. Mitochondrial data from Bond (2012) indicate considerable population genetic structuring, especially in *A. icenoglei*, which is likely influenced greatly by the typical mygalomorph life history traits already discussed in detail above. This, in conjunction with notable molecular divergence as well as a diversity of habitats suggest that *A. barackobamai* and *A. icenoglei* populations, respectively, have been isolated from gene flow for an extended period of time, which would increase speciation potential (i.e., either likely comprises more than one species; Barraclough, 2019).

The primary objective of this study is to use genomic-scale data, specifically ultraconserved elements (UCEs; Faircloth et al., 2012), to evaluate phylogenetic relationships, species boundaries, and historical biogeography within the *A. icenoglei* complex. We explicitly tested species hypotheses within this assemblage by implementing a CSC-based approach. We first evaluated genetic exchangeability using clustering analyses to assess the potential for gene flow and then assessed ecological interchangeability of genetic lineages with a niche-based distribution modeling approach. Additionally, biogeographic analysis was used to investigate the likelihood of dispersal versus vicariant events that may have potentially influenced speciation pattern and process across the CAFP's complex topographic and geologic landscape.

Methods

Taxon Sampling

We sampled 62 individuals overall for the three species within the complex using both specimens from Bond (2012) and new records (Figure 6; see Table 4 for locality information). *Aptostichus barackobamai* was collected across its geographic range in northern California for a total of 21 samples, and *A. icenoglei* was collected throughout its range in southern California for a total of 40 samples. Only one specimen of *A. isabella* was included in this study due to collecting constraints (i.e., only one individual of this species has ever been collected and a burrow has not yet been found containing this species; Bond 2012).

Sequence Capture

Data for ultraconserved elements was produced following the methods described in Faircloth et al. (2012) with subsequent modifications in Starrett et al. (2017), Hedin et al. (2018), and Kulkarni et al. (2020). We extracted genomic DNA from leg tissue for *A. barackobamai* and *A. icenoglei* individuals using the Blood and Tissue DNeasy kit (Qiagen) following the manufacturer's protocol. The lone *A. isabella* individual, a museum voucher kept in 80% ethanol and stored at room temperature, had DNA extracted from leg tissue following the 'MMYT protocol' from Tin et al. (2014) with modifications in Derkarabetian et al. (2019). DNA quantification and quality check were performed using Qubit 3.0 Fluorometer (Life Technologies) and agarose gel, respectively. Excluding *A. isabella*, 250 ng of DNA was sonicated into fragments ranging from 200-1000 bp using an ultrasonicator (Covaris E220). UCE libraries were generated with the KAPA Hyperprep kit (Roche) with universal adapters and iTru5/7 barcodes (Glenn et al., 2019; BadDNA@UGA) with slight modifications on a few steps for *A. isabella* (for details see Newton et al., in prep). Libraries were hybridized at 60°C for 24 hours to the Spider probeset (Kulkarni et al., 2020) following the version 4 chemistry protocol (Arbor Biosciences). Hybridization enriched library pools were sequenced with 150 bp paired-

end reads on the HiSeq4K at the University of California Davis DNA Technologies Core. Additional individuals were sent to Rapid Genomics (Florida) for library prep and sequencing. Sequence processing and analyses were performed on the Farm Community HPC at the University of California, Davis. Reads were filtered and trimmed using Illumiprocessor (Faircloth, 2013) and Trimmomatic (Bolger et al., 2014) in the Phyluce 1.7.1 pipeline (Faircloth, 2015). De novo assemblies with the cleaned paired-end and single-end reads were performed using SPAdes v. 3.14.1 with the isolate option (Prjibelski et al., 2020). Scaffolds were matched with 65% identity and 65% coverage to the modified probe list from Maddison et al. (2020), which is a blend of the Arachnid (Faircloth, 2017; Starrett et al., 2017) and Spider (Kulkarni et al., 2020) probesets. MAFFT was used to align individual locus datasets, and alignments with locus occupancy minimums of 50%, 75%, and 90% were obtained. Alignment masking was performed with TrimAl v.1.2 (Capella-Gutierrez et al., 2009) using default settings.

SNP datasets were generated for *A. icenoglei* only with Phyluce from the 50%, 75%, and 90% minimum occupancy loci. Reads were mapped against corresponding scaffolds with BWA, implemented in Phyluce, and phased alignments were generated for each minimum locus occupancy set. Phased alignments were screened for SNPs, with five sets of single random SNP per locus generated to test for SNP set sensitivity.

Phylogenetic & Biogeographic Analyses

Phylogenies were estimated for 3 different data sets (50, 75, and 90 percent locus occupancy; Figures 7-9) with a maximum likelihood inference using IQ-TREE v2.1.2 (Minh et al., 2020). Model selection was performed by ModelFinder (Kalyaanamoorthy et al., 2017), which is implemented in IQ-TREE, and support values were inferred from 1000 replicates of ultrafast bootstrapping (Hoang et al., 2018). Our phylogenies were visualized in FigTree v1.4.1

with midpoint rooting (large scale analyses of the entire genus support this rooting so outgroups were omitted to maximize locus recovery; Newton et al., in prep) and compared to assess congruence among clades. We also conducted two coalescent based analyses for the 75p and 90p data sets. Gene trees for each locus were constructed using RAxML v8.0.12 for each data set and used to generate a coalescent based tree with ASTRAL-III (Figures 10 and 11; Zhang et al., 2018). Multispecies coalescent (MSC) bootstrapping was run with ASTRAL v.5.7.4 and 100 pseudoreplicates (Figures 12 and 13; Simmons et al., 2019). For downstream analyses, we decided to employ the ML phylogeny based on the largest amount of taxon coverage and with the most robust support values (i.e., the phylogeny with 90 percent minimum locus occupancy).

Biogeographic analyses were generated using Reconstruct Ancestral State in Phylogenies (RASP; Yu et al., 2015) with dispersal constraints (i.e., dispersal multipliers set to 0.01 for adjacent areas and 0.0001for non-adjacent areas) to account for their limited dispersal capacity and using our 90p consensus tree from IQ-TREE. Model testing was conducted using the R package BioGeoBEARS (Matzke, 2014), implemented in RASP, and the best fit model was chosen based on the weighted AICc scores (Figure 14). The distribution range of this complex was divided into seven regions: (A) lower San Diego county; (B) upper San Diego county/Santa Ana Mtns; (C) Transverse Ranges (San Gabriel & San Bernardino Mtns); (D) southern Sierras; (E) northern rim of Central Valley; (F) Sutter Buttes; and (G) Northern Coast Ranges.

Cohesion Species Delimitation

To assess species boundaries within *A. barackobamai* and *A. icenoglei* we employed the methodological framework for delimiting cohesion species from Bond & Stockman (2008) that evaluates two cohesion mechanisms: genetic exchangeability and ecological interchangeability; thus, a cohesion species will comprise populations that exchange genes, or have a clear potential

to exchange genes, **and** occupy similar niche space. We used our 90p topology as the baseline evolutionary framework for establishing the 'basal starting point' to identify potential separately evolving lineages (for details see flowchart in Bond & Stockman, 2008) within *A. icenoglei* and *A. barackobamai*. Due to the paraphyletic grade within *A. barackobamai*, we designated all of the individuals as part of one evolving lineage and was not tested further for genetic and ecological exchangeability. For *A. icenoglei*, we also used our topology from the MSC tree resampling (Figure 13) as additional guidance for establishing lineage designations, which resulted in 3 lineages: North, Central, and South (see Figure 7). We evaluated the distributions of these lineages as well as performed morphological and genetic clustering analyses to assess the potential for gene flow. Genetic exchangeability was rejected if any allopatric lineage forms an apparently separate clustering pattern from other lineages, or if any parapatric lineage has a separate clustering pattern *and* an obvious barrier to gene flow.

For morphological data, we quantified 25 continuous character measurements for 30 males (10 males from each lineage; Table 5). All measurements were recorded in millimeters and were quantified with a Leica M165C stereomicroscope using the Leica Application Suite software and a digital camera. Measurements were transformed to log normal values, and a principal component analysis was conducted using ggplot2 in R (R core team), following Hamilton et al. (2016). We conducted two genetic clustering analyses. Variational AutoEncoder (VAE), an unsupervised machine learning approach derived from Bayesian probability theory, was used to visualize clustering of these lineages (Figure 15; for details see Derkarabetian et al., 2019). This class of neural networks takes large-scale SNP data as input and compresses this high dimensional data through several encoding layers into two-dimensional latent variables, which is subsequently reconstructed by uncompressing the latent variables through several

decoding layers (Derkarabetian et al., 2019). SNP datasets were converted to one-hot encoding and used as input for VAE analyses. Three replicates per random SNP set were conducted for each dataset (15 total replicates per dataset), and the replicate for each dataset with the least amount of loss during decoding was used for visualization (Figure 15). CLADES, a supervised machine learning approach, was used to further test species hypotheses (Pei et al., 2018). A 90% minimum locus occupancy data set with all *A. icenoglei* individuals was the input data for CLADES. The delimitation analysis was performed using a training model of genetic characteristics of species generated from a short-range endemic arachnid genus (*Metanonychus*) that has similar natural history characteristics to mygalomorph spiders (Metano_CLADES model from Derkarabetian et al., 2022).

Niche-based species distribution modeling (SDMs) with measures of SDM overlaps for each lineage were used as a proxy for ecological interchangeability, with ecological interchangeability rejected if both the niche equivalency and niche similarity tests are more different than expected by chance (i.e., niche divergence). Current climate data for 19 bioclimatic variables averaged from 1970-2000 were downloaded from WorldClim version 2.1 at 30 arc-second spatial resolution (https://www.worldclim.org/data/worldclim21.html; Fick & Hijmans, 2017). Climate layers were cropped to encompass the geographic area of interest and converted to a raster stack using R packages raster (Hijmans, 2015) and rgdal (Bivand et al., 2019). Highly correlated variables with a Pearson correlation coefficient > 0.80, estimated using the R package ENMTools (Warren et al., 2021), were removed. The remaining bioclimatic variables (see Table 6) were used in conjunction with occurrence records from the current study as well as records from Bond (2012) that could confidently be assigned to a lineage to generate SDMs, with duplicate records deleted prior to SDM construction. The R package ENMeval

(Kass et al., 2021) was used to estimate the SDM for each lineage by implementing MaxEnt (Phillips & Dudík, 2008), which is a machine learning program that uses maximum entropy algorithm. Multiple points within a 30 arc second grid cell were removed (i.e., only retaining one record per grid cell) by ENMeval during the modeling step to reduce potential for spatial autocorrelation. To limit the likelihood of overfitting while also accounting for goodness of fit, multiple feature classes and regularization multipliers were chosen to generate a total of 30 models (see Tables 7-10 for model parameters and stats). Model selection was based on AICc, with the best model having a delta AICc of zero and was subsequently used in downstream analyses (Figure 16).

Statistical comparisons of SDMs for each sister lineage comparison were conducted with niche overlap, niche equivalency, and niche similarity tests in ENMTools (Warren et al., 2008, 2010). We used the Schoener's *D* statistic (Schoener, 1968) to calculate the niche overlap for each lineage comparison, which ranges from 0 (no overlap) to 1 (complete overlap). We carried out two tests, niche equivalency and niche similarity, to evaluate the significance of niche overlap with a randomization procedure (Warren et al., 2008, 2010). The niche equivalency test, a one-tailed test, assesses whether the two niches being compared are identical or not. If the observed niche overlap value is significantly lower than the null distribution of randomized *D* values, then the niches are not identical (i.e., not equivalent; Figure 17). Considering the limitation of relying only on occurrence records for the niche equivalency test (Warren et al., 2008), we also employed the niche similarity test, a two-tailed test, to assess whether niche overlap between lineages relative to the niche spaces available to those lineages is more similar or different than expected by chance (niche conservatism or niche divergence, respectively; Figures 18 and 19). We estimated 3 potential background regions for each lineage in ArcGIS Pro

v2.8 (ESRI): 1) minimum area polygons based on occurrence points; 2) minimum area polygons based on SDM raster grid cells with a habitat suitability score threshold greater than 0.5 (i.e., a polygon was generated around every grid cell with a habitat suitability score greater than 0.5); and 3) minimum area polygons based on SDM raster grid cells with a habitat suitability score threshold greater than 0.75 (see Figures 20-22 for reference).

Results

UCE Stats

The UCE data are summarized in Table 11. Across all individuals, contigs that matched to the probes had a mean length of 1010 bp, with an average of 545 contigs over 1kb per individual. After aligning, filtering, and trimming these UCE contigs we had three data matrices with varying minimum locus occupancy percentages: 50p containing 1336 loci with 1101054 total bp, 75p containing 835 loci with 692091 total bp, and 90p containing 242 loci with 204134 total bp. For each *A. icenoglei* SNP data set there were 1120 SNPs, 668 SNPs, and 195 SNPS for the 50p, 75p, and 90p respectively.

Phylogenetic & Biogeographic Analyses

All estimated phylogenies fully supported (i.e., 100 for IQTREE or 1 for ASTRAL analyses) species level divergence among the three previously delineated morphological species within this sibling complex (see Figures 7-13). Also, all *A. icenoglei* lineage divergences (North, Central, and South) were highly, if not fully, supported across the majority of phylogenetic trees (i.e., all IQTREE and ASTRAL analyses); however, despite recovering North and South lineages as monophyletic and highly supported (i.e., > 90) in both the 75p and 90p MSC bootstrapping

analyses, there was uncertainty in placement of several Central lineage individuals with both analyses (Figures 12 and 13).

RASP analysis (Figure 14) revealed an unresolved ancestral range for the ancestor of the complex, with only 0.1968 probability of a relatively widespread ancestor along the southern Sierras and in the Transverse Ranges that then dispersed to the north with a vicariance event, splitting the ancestor of *A. barackobamai* from the ancestor of *A. isabella*+*A. icenoglei*. Within *A. barackobamai*, there were dispersal events northeastward along the northern Coast Ranges to the northern rim of the Central Valley/Sierra Nevada, and vicariance events splitting the northern Coast Ranges populations from northeastern populations as well as the Sutter Buttes population from the northern Sierra Nevada population. The split between the ancestor of *A. isabella* and the ancestor of *A. icenoglei*, with a 0.25 probability, was potentially the result of dispersal further south and a vicariance event. For the *A. icenoglei* lineages, the most likely scenario involved dispersal to the south towards the Peninsular Ranges with vicariance splitting the North from Central+South lineages and subsequent vicariance splitting Central from South.

Cohesion Species Delimitation

Table 12 summarizes results for each genetic exchangeability analysis of the nominal species, *Aptostichus icenoglei*. Geographic distribution assessments for each of our lineage comparisons were considered parapatric. One comparison had no obvious barrier to gene flow (i.e., Central and South lineages), whereas the other comparison of North and Central+South lineages had an unlikely chance of gene flow occurring due to the LA Basin acting as a likely geographic barrier (see Figure 6). Three clustering analyses, one with morphological data and two with molecular data, were also used to inform the possibility of gene flow. PCA analysis of the quantitative morphological measurements reveal no distinct clustering for any of the lineages

(Figure 15). Similarly, the CLADES analysis with the Metano_CLADES training model identified one species. In contrast, the VAE analysis for 50p indicate three very distinct clusters corresponding to each lineage for both the mean and standard deviation (Figure 23); however, although VAE analyses for 75p and 90p show separation between the lineages for the mean, there is a small amount of overlap for the standard deviation between Central and South lineages (Figure 15).

Table 12 also summarizes results for each ecological interchangeability analysis. Niche equivalency was rejected for both of the lineage comparisons, indicating that their niches are not identical (Figure 17). Niche similarity test results were different depending on the background region selected. Central occurrence points compared to the South background, determined by a minimum bounding polygon connecting its occurrence points, were not significantly different; however, the reciprocal comparison of South occurrence points to Central background was significantly more similar than expected (i.e., niche conservatism; Figure 18). Central occurrences compared to the South background, determined by minimum bounding polygons based on raster grid cells with either habitat suitability scores > 0.5 or > 0.75, and vice versa indicated niche conservatism (i.e., more similar than expected compared to the null distribution; Figure 18). When comparing the Central+South occurrence records to the minimum bounding polygon connecting occurrence points defining the background region of North, the results show no significant difference; yet, the reciprocal comparison is significantly more similar than expected (Figure 19). All comparisons of North versus Central+South and vice versa suggest niche conservatism when background regions are defined by minimum bounding polygons based on raster grid cells with either habitat suitability scores > 0.5 or > 0.75 (Figure 19).

Discussion

Species delimitation of the *A. icenoglei* assemblage previously identified three species based on the traditional morphological distinctiveness criterion; however, there was molecular (i.e., mitochondrial) evidence that the two geographically widespread species (*A. icenoglei* and *A. barackobamai*) could include additional cryptic diversity (Bond, 2012). Our study used a cohesion species-based delimitation approach from Bond & Stockman (2008) to expand on the evaluation of species boundaries within the complex. Once our evolutionary lineages were delineated, which was based on the topology and high support values (i.e., > 0.95) from both the 90p IQ-TREE and the 90p MSC bootstrapping (Figures 7 and 13), we recognized three distinct lineages with *A. icenoglei*: North, Central, and South. In contrast, the paraphyletic grade within *A. barackobamai* leads us to retain the current species boundaries at this time (i.e., *A. barackobamai* populations comprise a single species owing to their apparent genetic exchangeability). Although it is possible that sampling more populations is warranted, particularly where a modest-sized gap exists between the Coast Range populations and northern Central Valley rim/Sierra populations (Figure 6), intensive sampling efforts in parts of the Mendocino National Forest area did not yield any additional populations. As such, we have limited our targeted assessment of genetic and ecological exchangeability to *A. iceonoglei* lineages. Specifically, we utilized morphological measurements, genomic-scale SNP data, and niche-based distribution modeling to further evaluate and test cohesion species boundaries within *A. icenoglei*; these analyses produced conflicting results inferring one to three species. The unsupervised machine learning (VAE) analysis with the 50p dataset and niche equivalency tests coincide with the three species hypothesis (i.e., North, Central, and South lineages are all separate cohesion species). The two species hypothesis (i.e., North and Central+South lineages are cohesion species) is supported by a likely geographic barrier to gene flow (i.e., LA Basin)

and VAE analyses with both 75p and 90p datasets. In contrast, the morphological data, supervised machine learning (CLADES) analysis, and niche similarity tests support what is essentially the null hypothesis that *A. icenoglei* lineages all comprise one single species. When taking into account all lines of evidence, limitations of datasets and analyses, and mygalomorph life history characteristics we tentatively choose to retain the current species delimitation of *A. icenoglei* as one cohesion species (discussed further below).

Speciation & phylogeography

Spiders in the *A. icenoglei* complex, like most mygalomorphs, have limited dispersal capabilities and relatively long generation times (Bond, 2012; Harvey et al., 2018; Hedin et al., 2013; Hendrixson et al., 2013), which contributes to their tendency to have population structure at relatively small spatial scales. The molecular data show genetic divergence across the *A. icenoglei* complex and within *A. icenoglei* populations, thus populations have likely been isolated from gene flow for a long period of time, inferring the increased potential for speciation (Barraclough, 2019). The three nominal species (*A. barackobamai*, *A. isabella*, and *A. icenoglei*) are distributed across different areas of the CAFP and have been delimited based on distinct morphological differences in secondary mating structures (clasper morphology; Bond 2012), indicating that gene flow has not occurred between them for a relatively long period of time. Within *A. icenoglei* lineages, our VAE analyses with 75p and 90p datasets along with the LA Basin acting as a geographic barrier, support the lack of gene flow between the North and Central+South lineages over an extended time period.

All species within the complex, with potentially the exception of *A. isabella*, seem to have similar microhabitat requirements (e.g., north-facing shady slopes) despite their occurrence in different ecoregions, similar to other mygalomorph taxa in the CAFP (e.g., Hedin et al., 2013;

Hedin & Carlson, 2011; Leavitt et al., 2015). Our niche similarity tests show evidence of niche conservation within *A. icenoglei* lineages, with the caveat that these analyses yielded different results for minimum bounding polygons versus raster grid cell thresholds. Many studies have used the niche similarity test to evaluate overlap in niche space (e.g., Hendrixson et al., 2013; McCormack et al., 2009; Newton et al., 2020; Starrett et al., 2018; Warren et al., 2008), yet very few are explicit about the background region they chose to incorporate into the analysis (McCormack et al., 2009; Newton et al., 2020; Starrett et al., 2018). In addition, as far as we are aware, no other study other than Newton et al. (2020) has explicitly tested multiple background regions to evaluate the impact background region choice has on the analysis. Our background region choices for the current study were chosen based on previous studies (minimum bounding polygon; McCormack et al., 2009; Starrett et al., 2018) and polygons reflecting raster grid cells with habitat suitability score thresholds (> 0.5 and > 0.75) that better reflect the suitable habitat space available (i.e., not including large gaps of uninhabitable areas that are included in the minimum bounding polygon). The minimum bounding polygon yielded conflicting results for both North versus Central+South comparisons and Central versus South comparisons, which is most likely attributed to the aforementioned uninhabitable areas included that potentially obscured signal of niche conservatism. Although we attempted to incorporate a more biologically realistic background region, it is possible that our habitat suitability thresholds slightly inflated the inference of niche conservatism and thus should be tested more extensively in a future study.

Niche conservatism, in conjunction with restricted gene flow, suggests that speciation scenarios in which vicariant events separate populations with subsequent reproductive isolation through genetic drift, as opposed to ecological differentiation, may apply across the complex.

This pattern is also supported by our biogeographic analysis, but caution should be used when interpreting these results considering our ultrametric tree was not dated and, thus, cannot pinpoint with certainty the geological/climatic events that potentially influenced these splits. First, a vicariant event (after range expansion; Figure 14) is inferred for the split of the ancestor of *A. barackobamai* and ancestor of *A. isabella+A.icenoglei*. This phylogeographic break potentially coincides with uplift of the Transverse Ranges approximately 5 mya (Norris & Webb, 1990), which likely cut off the potential for gene flow, and has been hypothesized for other taxa in the CAFP (Alexander & Burns, 2006; Calsbeek et al., 2003; Feldman & Spicer, 2006; Reilly et al., 2015; Rissler et al., 2006). Second, the split between *A. isabella* and *A. icenoglei* possibly occurred due to vicariance. This split could be attributed to the Tehachapi Mountains acting as a barrier to dispersal, which has also been inferred for other taxa (Calsbeek et al., 2003; Chatzimanolis & Caterino, 2007; Rissler et al., 2006). Third, vicariance was inferred for the split between the North lineage and Central+South lineages, which could be associated with periodic inundations of the LA Basin (Jacobs et al., 2004) that might have resulted in habitat fragmentation, also hypothesized for the mahogany Jerusalem cricket (*Stenopelmatus* 'mahogani'; Vandergast et al., 2006) and a stream-dwelling frog (*Pseudacris cadaverina*; Phillipsen & Metcalf, 2009).

Species Limits within *A. icenoglei* and Taxonomic Implications

Although our integrative approach took into account multiple independent lines of evidence, our conflicting results circle back to the unavoidably subjective question of how much weight should be given to genetic divergence versus morphological/ecological divergence (or lack thereof) when delimiting species with extreme population structuring. Should we elevate genetically diverged lineages to species status despite the lack of observed

morphological/ecological differences? One could argue that identifying and describing cryptic diversity can be important not only for more accurate biodiversity measures but also conservation management plans. For example, Fennessy et al. (2016) delimited four species of giraffe based on a genetic isolation criterion and placed special emphasis on conservation management of the northern giraffe *Giraffa camelopardalis* and its four recognized subspecies due to the severity of population declines when compared to other *Giraffa* species. In our case, the North lineage has been severely threatened by fires over the last 20 years compared to Central+South lineages (Newton et al., unpublished data). Specifically, more than half of the North lineage population occurrence records fall within a fire perimeter that occurred within the years 2000-2020, compared to approximately twenty percent for Central+South occurrence records falling within a fire perimeter. Failing to recognize the obvious genetic diversity in the North lineage could result in its loss due to lack of a management plan targeting their distribution in the Transverse Ranges or trying to manage all of *A. icenoglei* as one species could also result in not having adequate recognition and consequently protection for the North lineage.

Alternatively, one could argue that there is no practical value of recognizing genetically diverged lineages as separate species considering the lack of any visible diagnostic character/difference in ecological role (Freudenstein et al., 2016). Specifically, Freudenstein et al., (2016) argued that a unique ecological role "with phenotypic difference as critical" is imperative when recognizing distinct species units, and that "one needs enough evidence (of lineage or role) to build a persuasive case for a particular real-world instance". However, even this argument is rife with subjectivity; for example, how much phenotypic difference is enough to distinguish lineages as separate species? Also, it has been established that the speciation process is a continuum in which certain biological properties can be affected at different points

along that continuum (Abbott et al., 2013; de Queiroz, 2007). Thus, it is feasible for geographically separated populations to accumulate enough genetic divergence for reproductive isolation despite still having morphological and ecological stasis. However, if one were to view species only in the context of a small snapshot in time (i.e., time-limited view of species; Freudenstein et al., 2016) and assume reproductive isolation based only on genetic divergence, then that raises the question of what happens if/when secondary contact occurs with a sibling sister 'species' or lineage (i.e., time-extended view of species; Freudenstein et al., 2016). One of two options are possible if secondary contact occurs: 1) morphological and/or ecological differences could emerge to maintain reproductive isolation (reinforcement), or 2) hybridization occurs and genetic divergence between populations is eliminated via the effects of gene flow. Freudenstein et al., (2016) argues that viewing species "as historically connected populations with unique role" combines both the temporal and phenotypic natures of species and alleviates the ambiguity of whether or not genetically diverged yet morphologically/ecologically homogenous lineages will remain diverged in the future. Thus, the most conservative taxonomic approach would be to require rejection of *both* genetic and ecological interchangeability for identifying separate cohesion species.

Previous studies spanning different animal taxa that have utilized CSC-based delimitation approaches have highlighted the importance of evidence for adaptive divergence when delimiting species (e.g., Bond & Stockman, 2008; Leaché et al., 2009; Newton et al., 2020; Rengifo-Correa et al., 2021). For mygalomorphs, Bond and Stockman (2008), the study upon which our CSC framework is based, delimited *Aptostichus miwok* and *A. stephencolberti* within the *A. atomarius* species complex based on mitochondrial data plus evidence of adaptive divergence (i.e., coastal dune habitats and lighter abdominal coloration). In a follow-up study,

Garrison et al. (2020) found evidence of chemosensory-associated gene families under selection in dune endemics compared to their inland sister lineages, further elucidating patterns of ecological differentiation between coastal and inland sister species. Another example within mygalomorphs includes Newton et al. (2020) who initially identified five genetically distinct lineages within the *Antrodiaetus unicolor* species complex; however, genetic and ecological exchangeability assessments led to the delimitation of three species, not five, based on molecular, behavioral, and morphological data. In a similar study, Leaché et al., (2009) identified five phylogeographic groups within the coast horned lizard *Phyrnosoma coronatum* species complex based on molecular data, yet an assessment of climatic niche models and morphometrics of cranial horn shapes led to the delimitation of three species based on multiple operational criteria. Lastly, another example involves the difficult taxonomic status of kissing bugs within the *Triatoma phyllosoma* species group, where species limits have been hard to establish given occurrences of hybridization and cryptic diversity (Rengifo-Correa et al., 2021). Despite relatively low genetic divergence and the potential for hybridization, species within the *T. phyllosoma* complex can be distinguished based on morphological characters (i.e., head phenotype) and are all considered separate cohesion species.

Our analytical results separately inferred one to three species within *A. icenoglei* depending on the dataset and analysis used, but the final decision, arguably subjective, comes down to acknowledging mygalomorph life history characteristics and limitations for each data type/analysis (discussed further below). The three species hypothesis was dismissed due to: 1) the less conservative niche equivalency test (Warren et al., 2008), 2) the possibility that the 50 percent locus occupancy SNP dataset spuriously inflated cluster separation between Central + South, and 3) no obvious barrier to gene flow between Central and South lineages. The two

species hypothesis is not substantiated based on morphological and ecological similarity between lineages, yet it is supported by rejecting genetic exchangeability as inferred by the VAE cluster separation with higher/more conservative locus occupancy percentage datasets (75p and 90p) and a probable barrier to gene flow, the LA Basin, between North and Central+South. Although the LA Basin is likely impeding gene flow due to urban development and habitat fragmentation, the small likelihood of a potential corridor of habitat suitable for dispersal along the northern Basin rim/southeastern San Bernardino mountains cannot be completely dismissed (e.g., Figure 1 in Vandergast et al., 2006). The one species hypothesis is supported by morphological and ecological data as well an implementation of a supervised machine learning analysis on the 90p SNP dataset. However, the CLADES training model used in our study is potentially not appropriate for mygalomorphs, and the prevalence of morphological homogeneity (e.g., Bond & Stockman, 2008; Harvey et al., 2018; Hedin et al., 2013; Leavitt et al., 2015; Newton et al., 2020) and ecological similarity (e.g., Cooper et al., 2011; Hedin & Carlson, 2011; Rix et al., 2020) among mygalomorphs could obscure actual evolutionary diversity. The flowchart in Bond & Stockman (2008) suggests that rejecting genetic exchangeability for parapatric lineages, but not rejecting ecological interchangeability, can still potentially indicate separate cohesion species if niche conservatism is occurring. However, this view has to be balanced with acknowledging that sparse, if any, evidence for adaptive divergence could indicate that reproductive isolation is not complete (i.e., ecological divergence is usually correlated with reproductive isolation; Freudenstein et al., 2016; Rissler & Apodaca, 2007), especially for parapatric lineages that still have the potential for gene flow in the future. Considering the lack of congruence across data types and analyses, we are taking the most conservative approach by retaining species

boundaries within *A. icenoglei* until more data, both ecological and whole genomes, can be collected for a more robust evaluation of ecological interchangeability.

Limitations of Analyses & Future Prospects

We believe that the supervised machine learning analysis has limitations due to potential shortcomings with the training data set devised using unrelated taxa. Although we see great value in attempting to establish a training dataset integrating biologically/ecologically relevant characteristics, it is difficult to assess how applicable this dataset can be to other dispersal-limited taxa, especially across different taxonomic orders and biogeographical regions (Derkarabetian et al., 2022). First, the taxon *Metanonychus*, on which the training dataset was established, diverged approximately 25 mya, whereas the *A. icenoglei* sibling species complex likely diverged much later, which could artificially conflate deeper divergences with a predetermined 'species cutoff' value, even if shallower species divergences are observed. Second, *Metanonychus* is found throughout the Pacific Northwest (Derkarabetian et al., 2019) whereas the *A. icenoglei* complex is found throughout the California Floristic Province, a biodiversity hotspot characterized by the intense complexity of geological, climatic, and topographic changes (Myers 2000). One could argue that the overall complexity of the CAFP might influence the speciation process of low dispersal taxa in a different manner from how topographic changes in the Pacific Northwest would to the point that the genetic signatures may manifest differently. Specifically, as there are more topographical changes (both in number and intensity) the more chances there are for speciation through vicariance when compared to fewer/less drastic topography shifts (Badgley et al., 2017).

Our VAE analysis with the lower locus occupancy dataset (50p) showed obvious separation between all three of the *A.icenoglei* lineages, whereas our higher locus occupancy

datasets (75p and 90p) retained only enough signal to maintain the North lineage as a separate cluster but not for Central or South lineages (Figure 15). VAE relies on the inherent structure present in the data (Derkarabetian et al., 2019), and previous studies have shown that VAE analyses have been heavily influenced by the filtering parameters for the SNP datasets (Martin et al., 2021; Newton et al., 2020). Specifically, if a lower threshold for locus occupancy is allowed in a dataset the more likely it is to 'over-split', whereas more stringent filtering (i.e., a high threshold for locus occupancy) can remove potentially important signal and 'under-split' the amount of diversity. Because our higher locus occupancy datasets retained the same clustering patterns, we are confident that they accurately reflect genetic divergence, and that the 50p dataset separation pattern for Central and South is an artifact of the filtering choice. Thus, it is important to be mindful of the potential filtering strategies for these SNP datasets, and best practices if utilizing VAE as a species delimitation method would be to use multiple filtering strategies to identify possible data artifacts versus actual structure.

There are a few caveats for using niche-based distribution modeling approaches as a proxy for evaluating ecological interchangeability. First, it has to be acknowledged that large-scale ecological data, which are based on a very small time frame of thirty years (i.e., 1970-2000), used for building the SDMs potentially lacks the resolution needed for detecting very small-scale habitat differences which may be important for detecting adaptive divergence (Massatti & Knowles, 2014; Newton et al., 2020; Starrett et al., 2018). Second, as discussed above, background region choice can heavily impact the results of niche similarity tests, thus incorporating multiple regions with biologically relevant constraints may provide a more rigorous application. Third, considering that our proxy of ecological interchangeability was only based on the abiotic factors contributing to niche space (i.e., bioclimatic variables and occurrence

records to build an SDM), one could argue that there were other potential avenues of ecological divergence that could have been included in this study for a more robust evaluation of ecological interchangeability. There are potential biotic factors (e.g., competition with other taxa, difference in prey items across microhabitats, or non-overlapping breeding periods) that could distinguish lineages from one another. For example, other studies delimiting mygalomorph species have included behavioral traits when applicable (e.g., non-overlapping breeding periods; Hendrixson et al., 2015; Hendrixson & Bond, 2005; Prentice, 1997). Unfortunately, the lack of available natural history data for many fossorial mygalomorphs (Bond, 2012; Hedin et al., 2013; Starrett et al., 2017) have limited using this type of data in species delimitation decisions.

Given these limitations, there are many potential avenues in which researchers can begin to bridge these gaps in knowledge. First, generating more datasets comprising low dispersal taxa with varying divergence times and across other biogeographical regions that can be used to train models for supervised machine learning methods such as CLADES, will likely aid the robustness of this approach (Derkarabetian et al., 2022). Second, accumulating more natural history data for mygalomorphs will not only provide valuable general ecological information but may also be used as additional evidence in species delimitation. For example, installing pitfall traps in areas where occurrence records of each species/lineage of interest is well-known to collect specimens can be informative for both breeding period times and gut content analysis to identify prey items that are being ingested (i.e., can inform potential for ecological divergence). Third, the advent of assembled and annotated genomes for non-model taxa, specifically in *Aptostichus* (e.g., Bond et al., in prep), will likely pave the way towards utilizing these data not only for reconstructing evolutionary relationships but also identifying genes that contribute to potential adaptive divergence across the landscape (Johnson, 2018).

Overall, our study emphasized the efficacy of implementing a cohesion species-based delimitation approach across all taxa, but especially for assessing the potential of cryptic diversity. Using genome-wide UCEs in conjunction with morphological and ecological data to evaluate genetic and ecological exchangeability provided multiple independent lines of evidence that covered multiple biological properties potentially important in the speciation process. Specifically, this integrative approach underscored how different data types or approaches alone could either over- or under-split diversity estimates, yet taking them all into consideration led to a more robust species delimitation hypothesis within the *A. icenoglei* complex. In addition, our biogeographic analysis revealed that vicariance likely played a dominant role in the speciation process across this complex, highlighting the impact of the complex geological, climatic, and topographical changes across the CAFP on the speciation process.

References:

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Alexander, M. P., & Burns, K. J. (2006). Intraspecific phylogeography and divergence in the white-headed woodpecker. *The Condor*, *108*, 489–508.

Anacker, B. L., Whittall, J. B., Goldberg, E. E., & Harrison, S. P. (2011). Origins and consequences of serpentine endemism in the California flora. *Evolution*, *65*(2), 365–376. https://doi.org/10.1111/j.1558-5646.2010.01114.x

Badgley, C., Smiley, T. M., Terry, R., Davis, E. B., DeSantis, L. R. G., Fox, D. L., Hopkins, S. S. B., Jezkova, T., Matocq, M. D., Matzke, N., McGuire, J. L., Mulch, A., Riddle, B. R., Roth, V. L., Samuels, J. X., Strömberg, C. A. E., & Yanites, B. J. (2017). Biodiversity and topographic complexity: Modern and geohistorical perspectives. *Trends in Ecology & Evolution*, *32*(3), 211–226. https://doi.org/10.1016/j.tree.2016.12.010

Baldwin, B. G., Kalisz, S., & Armbruster, W. S. (2011). Phylogenetic perspectives on diversification, biogeography, and floral evolution of *Collinsia* and *Tonella* (Plantaginaceae). *American Journal of Botany*, *98*(4), 731–753. https://doi.org/10.3732/ajb.1000346

Barraclough, T. (2019). *The Evolutionary Biology of Species*. Oxford University Press.

Barroso, R., Klautau, M., Solé-Cava, A. M., & Paiva, P. C. (2010). *Eurythoe complanata* (Polychaeta: Amphinomidae), the 'cosmopolitan' fireworm, consists of at least three cryptic species. *Marine Biology*, *157*(1), 69–80. https://doi.org/10.1007/s00227-009-1296-9

Battey, C. J., & Klicka, J. (2017). Cryptic speciation and gene flow in a migratory songbird Species Complex: Insights from the Red-Eyed Vireo (Vireo olivaceus). *Molecular Phylogenetics and Evolution*, *113*, 67–75. https://doi.org/10.1016/j.ympev.2017.05.006

Bivand, R., Keitt, T., & Rowlingson, B. (2019). *rgdal: Bindings for the "geospatial" data abstraction library* [R package version 1.4-4]. . https://CRAN.R-project.org/package=rgdal

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bond, J. E.. (2012). Phylogenetic treatment and taxonomic revision of the trapdoor spider genus *Aptostichus* Simon (Araneae, Mygalomorphae, Euctenizidae). *ZooKeys*, *252*, 1–209. https://doi.org/10.3897/zookeys.252.3588

Bond, J.E., Hedin, M. C., Ramirez, M. G., & Opell, B. D. (2001). Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology*, *10*(4), 899–910. https://doi.org/10.1046/j.1365-294X.2001.01233.x

Bond, J. E., Godwin, R. L., Colby, J. D., Newton, L. G., Zahnle, X. J., Agnarsson, I., Hamilton, C. A., & Kuntner, M. (2021). Improving Taxonomic Practices and Enhancing Its Extensibility—An Example from Araneology. *Diversity*, *14*(1), 5. https://doi.org/10.3390/d14010005

Bond, J. E., & Stockman, A. K. (2008). An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology*, *57*(4), 628–646. https://doi.org/10.1080/10635150802302443

Bryson, R. W., Savary, W. E., Zellmer, A. J., Bury, R. B., & McCormack, J. E. (2016). Genomic data reveal ancient microendemism in forest scorpions across the California Floristic Province. *Molecular Ecology*, *25*(15), 3731–3751. https://doi.org/10.1111/mec.13707

Calsbeek, R., Thompson, J. N., & Richardson, J. E. (2003). Patterns of molecular evolution and diversification in a biodiversity hotspot: The California Floristic Province. *Molecular Ecology*, *12*(4), 1021–1029. https://doi.org/10.1046/j.1365-294X.2003.01794.x

Capella-Gutierrez, S., Silla-Martinez, J. M., & Gabaldon, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Cerca, J., Rivera-Colón, A. G., Ferreira, M. S., Ravinet, M., Nowak, M. D., Catchen, J. M., & Struck, T. H. (2021). Incomplete lineage sorting and ancient admixture, and speciation without morphological change in ghost-worm cryptic species. *PeerJ*, *9*, e10896. https://doi.org/10.7717/peerj.10896

Chatzimanolis, S., & Caterino, M. S. (2007). Limited phylogeographic structure in the flightless ground beetle,*Calathus ruficollis*, in southern California. *Diversity and Distributions*, *13*(5), 498–509. https://doi.org/10.1111/j.1472-4642.2007.00352.x

Cole, K. L., Ironside, K., Eischeid, J., Garfin, G., Duffy, P. B., & Toney, C. (2011). Past and ongoing shifts in Joshua tree distribution support future modeled range contraction. *Ecological Applications*, *21*(1), 137–149. https://doi.org/10.1890/09-1800.1

Cooper, S. J. B., Harvey, M. S., Saint, K. M., & Main, B. Y. (2011). Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: Evidence for long-term refugia within refugia. *Molecular Ecology*, *20*(15), 3219–3236. https://doi.org/10.1111/j.1365-294X.2011.05160.x

de Queiroz, K. (2007). Species Concepts and Species Delimitation. *Systematic Biology*, *56*(6), 879–886. https://doi.org/10.1080/10635150701701083

Derkarabetian, S., Baker, C. M., & Giribet, G. (2021). Complex patterns of Gondwanan biogeography revealed in a dispersal□limited arachnid. *Journal of Biogeography*, jbi.14080. https://doi.org/10.1111/jbi.14080

Derkarabetian, S., Castillo, S., Koo, P. K., Ovchinnikov, S., & Hedin, M. (2019). A demonstration of unsupervised machine learning in species delimitation. *Molecular Phylogenetics and Evolution*, *139*, 106562. https://doi.org/10.1016/j.ympev.2019.106562

Derkarabetian, S., Starrett, J., & Hedin, M. (2022). Using natural history to guide supervised machine learning for cryptic species delimitation with genetic data. *Frontiers in Zoology*, *19*(1), 8. https://doi.org/10.1186/s12983-022-00453-0

Eckert, A. J., Tearse, B. R., & Hall, B. D. (2008). A phylogeographical analysis of the range disjunction for foxtail pine (*Pinus balfouriana*, Pinaceae): The role of Pleistocene glaciation. *Molecular Ecology*, *17*(8), 1983–1997. https://doi.org/10.1111/j.1365-294X.2008.03722.x

Emata, K. N., & Hedin, M. (2016). From the mountains to the coast and back again: Ancient biogeography in a radiation of short-range endemic harvestmen from California. *Molecular Phylogenetics and Evolution*, *98*, 233–243. https://doi.org/10.1016/j.ympev.2016.02.002

Faircloth, B. C. (2013). *Illumiprocessor: A trimmomatic wrapper for parallel adapter and quality trimming*. https://github.com/faircloth-lab/illumiprocessor

Faircloth, B. C. (2015). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788.

Faircloth, B. C. (2017). Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods in Ecology and Evolution*, *8*(9), 1103–1112. https://doi.org/10.1111/2041-210X.12754

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning

multiple evolutionary timescales. *Systematic Biology*, *61*(5), 717–726. https://doi.org/10.1093/sysbio/sys004

Feldman, C. R., & Spicer, G. S. (2006). Comparative phylogeography of woodland reptiles in California: Repeated patterns of cladogenesis and population expansion. *Molecular Ecology*, *15*(8), 2201–2222. https://doi.org/10.1111/j.1365-294X.2006.02930.x

Fennessy, J., Bidon, T., Reuss, F., Kumar, V., Elkan, P., Nilsson, M. A., Vamberger, M., Fritz, U., & Janke, A. (2016). Multi-locus Analyses Reveal Four Giraffe Species Instead of One. *Current Biology*, *26*(18), 2543–2549. https://doi.org/10.1016/j.cub.2016.07.036

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas: new climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2016). Biodiversity and the species concept—Lineages are not enough. *Systematic Biology*, syw098. https://doi.org/10.1093/sysbio/syw098

Garrison, N. L., Brewer, M. S., & Bond, J. E. (2020). Shifting evolutionary sands: Transcriptome characterization of the Aptostichus atomarius species complex. *BMC Evolutionary Biology*, 1–12. https://doi.org/10.1186/s12862-020-01606-7

Glenn, T. C., Pierson, T. W., Bayona-Vásquez, N. J., Kieran, T. J., Hoffberg, S. L., Thomas IV, J. C., Lefever, D. E., Finger, J. W., Gao, B., Bian, X., Louha, S., Kolli, R. T., Bentley, K. E., Rushmore, J., Wong, K., Shaw, T. I., Rothrock Jr, M. J., McKee, A. M., Guo, T. L., … Faircloth, B. C. (2019). Adapterama II: Universal amplicon sequencing on Illumina platforms (TaggiMatrix). *PeerJ*, *7*, e7786. https://doi.org/10.7717/peerj.7786

Grivet, D., Deguilloux, M.-F., Petit, R. J., & Sork, V. L. (2006). Contrasting patterns of historical colonization in white oaks (*Quercus* spp.) in California and Europe: *Molecular Ecology*, *15*(13), 4085–4093. https://doi.org/10.1111/j.1365-294X.2006.03083.x

Hamilton, C. A., Hendrixson, B. E., & Bond, J. E. (2016). Taxonomic revision of the tarantula genus *Aphonopelma* Pocock, 1901 (Araneae, Mygalomorphae, Theraphosidae) within the United States. *ZooKeys*, *560*, 1–340. https://doi.org/10.3897/zookeys.560.6264

Hamilton, C. A., Hendrixson, B. E., Brewer, M. S., & Bond, J. E. (2014). An evaluation of sampling effects on multiple DNA barcoding methods leads to an integrative approach for delimiting species: A case study of the North American tarantula genus *Aphonopelma* (Araneae, Mygalomorphae, Theraphosidae). *Molecular Phylogenetics and Evolution*, *71*, 79–93. https://doi.org/10.1016/j.ympev.2013.11.007

Harvey, M. S., Hillyer, M. J., Main, B. Y., Moulds, T. A., Raven, R. J., Rix, M. G., Vink, C. J., & Huey, J. A. (2018). Phylogenetic relationships of the Australasian open-holed trapdoor spiders (Araneae: Mygalomorphae: Nemesiidae: Anaminae): multi-locus molecular analyses resolve the generic classification of a highly diverse fauna. *Zoological Journal of the Linnean Society*, *184*(2), 407–452. https://doi.org/10.1093/zoolinnean/zlx111

Hedin, M., & Carlson, D. (2011). A new trapdoor spider species from the southern Coast Ranges of California (Mygalomorphae, Antrodiaetidae, *Aliatypus coylei*, sp. Nov,), including consideration of mitochondrial phylogeographic structuring. *Zootaxa*, *2963*(1), 55. https://doi.org/10.11646/zootaxa.2963.1.3

Hedin, M., Carlson, D., & Coyle, F. A. (2015). Sky island diversification meets the multispecies coalescent—Divergence in the spruce-fir moss spider (*Microhexura montivaga,* Araneae,

Mygalomorphae) on the highest peaks of southern Appalachia. *Molecular Ecology*, *24*(13), 3467–3484. https://doi.org/10.1111/mec.13248

Hedin, M., Derkarabetian, S., Ramírez, M. J., Vink, C., & Bond, J. E. (2018). Phylogenomic reclassification of the world's most venomous spiders (Mygalomorphae, Atracinae), with implications for venom evolution. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-19946-2

Hedin, M., Starrett, J., & Hayashi, C. (2013). Crossing the uncrossable: Novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California. *Molecular Ecology*, *22*(2), 508–526. https://doi.org/10.1111/mec.12130

Hendrixson, B. E., & Bond, J. E. (2005). Testing species boundaries in the *Antrodiaetus unicolor* complex (Araneae: Mygalomorphae: Antrodiaetidae): "Paraphyly" and cryptic diversity. *Molecular Phylogenetics and Evolution*, *36*(2), 405–416. https://doi.org/10.1016/j.ympev.2005.01.021

Hendrixson, B. E., DeRussy, B. M., Hamilton, C. A., & Bond, J. E. (2013). An exploration of species boundaries in turret-building tarantulas of the Mojave Desert (Araneae, Mygalomorphae, Theraphosidae, *Aphonopelma*). *Molecular Phylogenetics and Evolution*, *66*(1), 327–340. https://doi.org/10.1016/j.ympev.2012.10.004

Hendrixson, B. E., Guice, A. V., & Bond, J. E. (2015). Integrative species delimitation and conservation of tarantulas (Araneae, Mygalomorphae, Theraphosidae) from a North American biodiversity hotspot. *Insect Conservation and Diversity*, *8*(2), 120–131. https://doi.org/10.1111/icad.12089

Hey, J. (2001). The mind of the species problem. *Trends in Ecology & Evolution*, *16*(7), 326–329.

Hijmans, R. J. (2015). *raster: Geographic data analysis and modeling* [R package version 2.5-2]. https://cran.r-project.org/package=raster

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Holland, B. S., Dawson, M. N., Crow, G. L., & Hofmann, D. K. (2004). Global phylogeography of *Cassiopea* (Scyphozoa: Rhizostomeae): Molecular evidence for cryptic species and multiple invasions of the Hawaiian Islands. *Marine Biology*, *145*(6), 1119–1128. https://doi.org/10.1007/s00227-004-1409-4

Jacobs, D. K., Haney, T. A., & Louie, K. D. (2004). Genes, diversity, and geologic process on the pacific coast. *Annual Review of Earth and Planetary Sciences*, *32*(1), 601–652. https://doi.org/10.1146/annurev.earth.32.092203.122436

Jockusch, E. L., Hansen, R. W., Fisher, R. N., & Wake, D. B. (2020). Slender salamanders (genus *Batrachoseps*) reveal Southern California to be a center for the diversification, persistence, and introduction of salamander lineages. *PeerJ*, *8*, e9599. https://doi.org/10.7717/peerj.9599

Johnson, B. R. (2018). Taxonomically Restricted Genes Are Fundamental to Biology and Evolution. *Frontiers in Genetics*, *9*, 407. https://doi.org/10.3389/fgene.2018.00407

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Kass, J. M., Muscarella, R., Galante, P. J., Bohl, C. L., Pinilla□Buitrago, G. E., Boria, R. A., Soley□Guardia, M., & Anderson, R. P. (2021). ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. *Methods in Ecology and Evolution*, *12*(9), 1602–1608. https://doi.org/10.1111/2041-210X.13628

Kraft, N. J. B., Baldwin, B. G., & Ackerly, D. D. (2010). Range size, taxon age and hotspots of neoendemism in the California flora: California plant neoendemism. *Diversity and Distributions*, *16*(3), 403–413. https://doi.org/10.1111/j.1472-4642.2010.00640.x

Kulkarni, S., Wood, H., Lloyd, M., & Hormiga, G. (2020). Spider□specific probe set for ultraconserved elements offers new perspectives on the evolutionary history of spiders (Arachnida, Araneae). *Molecular Ecology Resources*, *20*(1), 185–203. https://doi.org/10.1111/1755-0998.13099

Leaché, A. D., Koo, M. S., Spencer, C. L., Papenfuss, T. J., Fisher, R. N., & McGuire, J. A. (2009). Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proceedings of the National Academy of Sciences*, *106*(30), 12418–12423. https://doi.org/10.1073/pnas.0906380106

Leavitt, D. H., Starrett, J., Westphal, M. F., & Hedin, M. (2015). Multilocus sequence data reveal dozens of putative cryptic species in a radiation of endemic Californian mygalomorph spiders (Araneae, Mygalomorphae, Nemesiidae). *Molecular Phylogenetics and Evolution*, *91*, 56–67. https://doi.org/10.1016/j.ympev.2015.05.016

Liston, A., Parker-Defeniks, M., Syring, J. V., Willyard, A., & Cronn, R. (2007). Interspecific phylogenetic analysis enhances intraspecific phylogeographical inference: A case study in Pinus lambertiana. *Molecular Ecology*, *16*(18), 3926–3937. https://doi.org/10.1111/j.1365-294X.2007.03461.x

Maddison, W. P., Beattie, I., Marathe, K., Ng, P. Y. C., Kanesharatnam, N., Benjamin, S. P., & Kunte, K. (2020). A phylogenetic and taxonomic review of baviine jumping spiders (Araneae, Salticidae, Baviini). *ZooKeys*, *1004*, 27–97. https://doi.org/10.3897/zookeys.1004.57526

Maddison, W. P., Evans, S. C., Hamilton, C. A., Bond, J. E., Lemmon, A. R., & Lemmon, E. M. (2017). A genome-wide phylogeny of jumping spiders (Araneae, Salticidae), using anchored hybrid enrichment. *ZooKeys*, *695*, 89–101. https://doi.org/10.3897/zookeys.695.13852

Martin, B. T., Chafin, T. K., Douglas, M. R., Placyk, J. S., Birkhead, R. D., Phillips, C. A., & Douglas, M. E. (2021). The choices we make and the impacts they have: Machine learning and species delimitation in North American box turtles (*Terrapene* spp.). *Molecular Ecology Resources*, *21*(8), 2801–2817. https://doi.org/10.1111/1755-0998.13350

Martínez-Solano, I., Jockusch, E. L., & Wake, D. B. (2007). Extreme population subdivision throughout a continuous range: Phylogeography of *Batrachoseps attenuatus* (Caudata: Plethodontidae) in western North America. *Molecular Ecology*, *16*(20), 4335–4355. https://doi.org/10.1111/j.1365-294X.2007.03527.x

Mas-Peinado, P., Buckley, D., Ruiz, J. L., & García-París, M. (2018). Recurrent diversification patterns and taxonomic complexity in morphologically conservative ancient lineages of *Pimelia* (Coleoptera: Tenebrionidae). *Systematic Entomology*, *43*(3), 522–548. https://doi.org/10.1111/syen.12291

Massatti, R., & Knowles, L. L. (2014). Microhabitat differences impact phylogeographic concordance of codistributed species: Genomic evidence in montane sedges (*Carex*) from the Rocky Mountains. *Evolution*, *68*(10), 2833–2846. https://doi.org/10.1111/evo.12491

Matzke, N. J. (2014). Model Selection in Historical Biogeography Reveals that Founder-Event Speciation Is a Crucial Process in Island Clades. *Systematic Biology*, *63*(6), 951–970. https://doi.org/10.1093/sysbio/syu056

McCormack, J. E., Zellmer, A. J., & Knowles, L. L. (2009). Does niche divergence accompany allopatric divergence in *Aphelocoma* jays as predicted under ecological speciation?: Insights from tests with niche models. *Evolution*. https://doi.org/10.1111/j.1558-5646.2009.00900.x

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403*(6772), 853–858. https://doi.org/10.1038/35002501

Newton, L. G., Starrett, J., Hendrixson, B. E., Derkarabetian, S., & Bond, J. E. (2020). Integrative species delimitation reveals cryptic diversity in the southern Appalachian *Antrodiaetus unicolor* (Araneae: Antrodiaetidae) species complex. *Molecular Ecology*, *29*(12), 2269–2287. https://doi.org/10.1111/mec.15483

Norris, R. M., & Webb, R. W. (1990). *Geology of California* (Second Edition). John Wiley.

Oliver, J. C., & Shapiro, A. M. (2007). Genetic isolation and cryptic variation within the *Lycaena xanthoides* species group (Lepidoptera: Lycaenidae). *Molecular Ecology*, *16*(20), 4308–4320. https://doi.org/10.1111/j.1365-294X.2007.03494.x

Opatova, V., Hamilton, C. A., Hedin, M., De Oca, L. M., Král, J., & Bond, J. E. (2019). Phylogenetic systematics and evolution of the spider infraorder Mygalomorphae using genomic scale data. *Systematic Biology*, *syz064*. https://doi.org/10.1093/sysbio/syz064

Pardikes, N. A., Harrison, J. G., Shapiro, A. M., & Forister, M. L. (2017). Synchronous population dynamics in California butterflies explained by climatic forcing. *Royal Society Open Science*, *4*(7), 170190. https://doi.org/10.1098/rsos.170190

Pei, J., Chu, C., Li, X., Lu, B., & Wu, Y. (2018). CLADES: A classification-based machine learning method for species delimitation from population genetic data. *Molecular Ecology Resources*, *18*(5), 1144–1156. https://doi.org/10.1111/1755-0998.12887

Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*(2), 161–175. https://doi.org/10.1111/j.0906-7590.2008.5203.x

Phillipsen, I. C., & Metcalf, A. E. (2009). Phylogeography of a stream-dwelling frog (*Pseudacris cadaverina*) in southern California. *Molecular Phylogenetics and Evolution*, *53*(1), 152–170. https://doi.org/10.1016/j.ympev.2009.05.021

Prentice, T. R. (1997). Theraphosidae of the Mojave Desert West and North of the Colorado River (Araneae, Mygalomorphae, Theraphosidae). *The Journal of Arachnology*, *25*, 137–176.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, *70*(1). https://doi.org/10.1002/cpbi.102

Reilly, S. B., Corl, A., & Wake, D. B. (2015). An integrative approach to phylogeography: Investigating the effects of ancient seaways, climate, and historical geology on multi-locus phylogeographic boundaries of the Arboreal Salamander (Aneides lugubris). *BMC Evolutionary Biology*, *15*(1), 241. https://doi.org/10.1186/s12862-015-0524-9

Rengifo-Correa, L., Juan Luis Téllez-Rendón, J. L. x, Esteban, L., Huerta, H., & Morrone, J. J. (2021). The *Triatoma phyllosoma* species group (Hemiptera: Reduviidae: Triatominae), vectors of Chagas disease: Diagnoses and a key to the species. *Zootaxa*, *5023*(3), 335–365. https://doi.org/10.11646/zootaxa.5023.3.2

Richardson, B. J. (2016). New genera, new species and redescriptions of Australian jumping spiders (Araneae: Salticidae). *Zootaxa*, *4114*(5), 501. https://doi.org/10.11646/zootaxa.4114.5.1

Rios, E., & Álvarez-Castañeda, S. T. (2010). Phylogeography and systematics of the San Diego pocket mouse (*Chaetodipus fallax*). *Journal of Mammalogy*, *91*(2), 293–301. https://doi.org/10.1644/09-MAMM-A-135

Rissler, L. J., & Apodaca, J. J. (2007). Adding More Ecology into Species Delimitation: Ecological Niche Models and Phylogeography Help Define Cryptic Species in the Black Salamander (Aneides flavipunctatus). *Systematic Biology*, *56*(6), 924–942. https://doi.org/10.1080/10635150701703063

Rissler, L. J., Hijmans, R. J., Graham, C. H., Moritz, C., & Wake, D. B. (2006). Phylogeographic lineages and species comparisons in conservation analyses: A case study of California herpetofauna. *The American Naturalist*, *167*(5), 12.

Rix, M. G., Wilson, J. D., & Harvey, M. S. (2020). First phylogenetic assessment and taxonomic synopsis of the open-holed trapdoor spider genus Namea (Mygalomorphae: Anamidae): a highly diverse mygalomorph lineage from Australia's tropical eastern rainforests. *Invertebrate Systematics*. https://doi.org/10.1071/IS20004

Rundel, P. W. (2011). The Diversity and Biogeography of the Alpine Flora of the Sierra Nevada, California. *Madroño*, *58*(3), 153–184. https://doi.org/10.3120/0024-9637-58.3.153

Satler, J. D., Starrett, J., Hayashi, C. Y., & Hedin, M. (2011). Inferring species trees from gene trees in a radiation of California trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus*). *PLoS ONE*, *6*(9), e25355. https://doi.org/10.1371/journal.pone.0025355

Schoener, T. W. (1968). The *Anolis* lizards of Bimini: Resource partitioning in a complex fauna. *Ecology*, *49*(4), 704–726. https://doi.org/10.2307/1935534

Sgariglia, E. A., & Burns, K. J. (2003). Phylogeography of the California thrasher (*Toxostoma redivivum*) based on nested-clade analysis of mitochondrial DNA variation. *The Auk*, *120*, 16.

Simmons, M. P., Sloan, D. B., Springer, M. S., & Gatesy, J. (2019). Gene-wise resampling outperforms site-wise resampling in phylogenetic coalescence analyses. *Molecular Phylogenetics and Evolution*, *131*, 80–92. https://doi.org/10.1016/j.ympev.2018.10.001

Spinks, P. Q., & Shaffer, H. B. (2005). Range-wide molecular analysis of the western pond turtle (*Emys marmorata*): Cryptic variation, isolation by distance, and their conservation

implications. *Molecular Ecology*, *14*(7), 2047–2064. https://doi.org/10.1111/j.1365-294X.2005.02564.x

Starrett, J., Derkarabetian, S., Hedin, M., Bryson, R. W., McCormack, J. E., & Faircloth, B. C. (2017). High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. *Molecular Ecology Resources*, *17*(4), 812–823. https://doi.org/10.1111/1755-0998.12621

Starrett, J., Hayashi, C. Y., Derkarabetian, S., & Hedin, M. (2018). Cryptic elevational zonation in trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus janus* complex) from the California southern Sierra Nevada. *Molecular Phylogenetics and Evolution*, *118*, 403–413. https://doi.org/10.1016/j.ympev.2017.09.003

Starrett, J., & Hedin, M. (2007). Multilocus genealogies reveal multiple cryptic species and biogeographical complexity in the California turret spider *Antrodiaetus riversi* (Mygalomorphae, Antrodiaetidae): Cryptic diversification in Californian *A. riversi*. *Molecular Ecology*, *16*(3), 583–604. https://doi.org/10.1111/j.1365-294X.2006.03164.x

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences*, *114*(7), 1607–1612. https://doi.org/10.1073/pnas.1607921114

Templeton, A. R. (1989). The meaning of species and speciation: A genetic perspective. In *Speciation and its Consequences* (pp. 3–27). Sinauer.

Tin, M. M.-Y., Economo, E. P., & Mikheyev, A. S. (2014). Sequencing Degraded DNA from Non-Destructively Sampled Museum Specimens for RAD-Tagging and Low-Coverage Shotgun Phylogenetics. *PLoS ONE*, *9*(5), e96793. https://doi.org/10.1371/journal.pone.0096793

Vandergast, A. G., Bohonak, A. J., Weissman, D. B., & Fisher, R. N. (2006). Understanding the genetic effects of recent habitat fragmentation in the context of evolutionary history: Phylogeography and landscape genetics of a southern California endemic Jerusalem cricket (Orthoptera: Stenopelmatidae: *Stenopelmatus*). *Molecular Ecology*, *16*(5), 977–992. https://doi.org/10.1111/j.1365-294X.2006.03216.x

Wake, D. B. (1997). Incipient species formation in salamanders of the *Ensatina* complex. *Proceedings of the National Academy of Sciences*, *94*(15), 7761–7767. https://doi.org/10.1073/pnas.94.15.7761

Warren, D. L., Glor, R. E., & Turelli, M. (2008). Environmental niche equivalency versus conservatism: Quantitative approaches to niche evolution. *Evolution*, *62*(11), 2868–2883. https://doi.org/10.1111/j.1558-5646.2008.00482.x

Warren, D. L., Glor, R. E., & Turelli, M. (2010). ENMTools: A toolbox for comparative studies of environmental niche models. *Ecography*. https://doi.org/10.1111/j.1600-0587.2009.06142.x

Warren, D. L., Matzke, N. J., Cardillo, M., Baumgartner, J. B., Beaumont, L. J., Turelli, M., Glor, R. E., Huron, N. A., Simões, M., Iglesias, T. L., Piquet, J. C., & Dinnage, R. (2021). ENMTools 1.0: An R package for comparative ecological biogeography. *Ecography*, ecog.05485. https://doi.org/10.1111/ecog.05485

Weisrock, D. W., & Larson, A. (2006). Testing hypotheses of speciation in the *Plethodon jordani* species complex with allozymes and mitochondrial DNA sequences: Diagnosing plethodontid salamander species. *Biological Journal of the Linnean Society*, *89*(1), 25–51. https://doi.org/10.1111/j.1095-8312.2006.00655.x

Wells, T., Carruthers, T., Muñoz-Rodríguez, P., Sumadijaya, A., Wood, J. R. I., & Scotland, R. W. (2021). Species as a Heuristic: Reconciling Theory and Practice. *Systematic Biology*, syab087. https://doi.org/10.1093/sysbio/syab087

Wiens, J. J., & Graham, C. H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 519–539. https://doi.org/10.1146/annurev.ecolsys.36.102803.095431

Yu, Y., Harris, A. J., Blair, C., & He, X. (2015). RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Molecular Phylogenetics and Evolution*, *87*, 46–49. https://doi.org/10.1016/j.ympev.2015.03.008

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*(S6). https://doi.org/10.1186/s12859-018-2129-y

Tables:

Table 4. Locality information for all samples used in this study.

Specimen ID	Specific Epithet	Locality Details	Latitude	Longitude	Country	State	County	Collection Date	Collectors
BME102225	*Aptostichus barackobamai*	CA-Hwy 253, E of Boonville	39.0547	-123.245	United States	California	Sonoma	1/16/2021	L. Newton, E. Joachim
BME102234	*Aptostichus barackobamai*	CA-Hwy 36, S of Platina	40.3177	-122.7822	United States	California	Shasta	2/3/2021	L. Newton

BME10237	*Aptostichus barackobamai*	Trinity Mtn Rd	40.6851	-122.6395	United States	California	Shasta	2/4/2021	L. Newton
BME10238	*Aptostichus barackobamai*	S Shore Dr, Whiskeytown Lake	40.6081	-122.5586	United States	California	Shasta	2/4/2021	L. Newton
BME10239	*Aptostichus barackobamai*	Middle Creek Rd	40.5955	-122.4528	United States	California	Shasta	2/4/2021	L. Newton
BME10240	*Aptostichus barackobamai*	Middle Creek Rd	40.5955	-122.4528	United States	California	Shasta	2/4/2021	L. Newton
BME10241	*Aptostichus barackobamai*	Middle Creek Rd	40.5955	-122.4528	United States	California	Shasta	2/4/2021	L. Newton
BME10242	*Aptostichus barackobamai*	S Cow Creek Rd	40.54	-122.1131	United States	California	Shasta	2/24/2021	L. Newton
BME10305	*Aptostichus barackobamai*	CA-Hwy 128, S of Boonville	38.8295	-123.044	United States	California	Sonoma	3/15/2021	L. Newton, E. Joachim
MY1098	*Aptostichus barackobamai*	Hwy 3, near Douglas City	40.64972	-122.94109	United States	California	Trinity	03/15/2006	M Hedin, J Starrett, S Thomas, R Keith, S Derkarabetian,

								M McCormack, D Marxsen
MY3025	*Aptostichus barackobamai*	Orr Springs Rd, Ackerman Creek-1st stream crossing W of Hwy 101	39.1807	-123.23307	United States	California	Mendocino	03/14/2005 J. Bond, A. Stockman, D. Beamer
MY3026	*Aptostichus barackobamai*	Orr Springs Rd, 2.4 mi W of bridge	39.19242	-123.26595	United States	California	Mendocino	03/14/2005 J. Bond, A. Stockman, D. Beamer
MY3027	*Aptostichus barackobamai*	Orr Springs Rd, W of Hwy 101, 8.3 mi W of 1st bridge	39.22932	-123.34265	United States	California	Mendocino	03/14/2005 J. Bond, A. Stockman, D. Beamer
MY3038	*Aptostichus barackobamai*	Orr Springs Rd, 24.1 mi W of Acker	39.26095	-123.54085	United States	California	Mendocino	03/14/2005 J. Bond, A. Stockman, D. Beamer

		man Creek							
MY3158	*Aptostichus barackobamai*	County Rd 201, 4.2 mi N of JCT w/ Hwy 175 (201 is turnoff for Hoplan d Field STN, but do not turn onto Univers ity Dr.)	39.028 32	-123.13 034	United States	Califor nia	Mendo cino	05/18/2 005	A.K. Stockm an
MY3173	*Aptostichus barackobamai*	Platina Rd, 2.9 mi NE of Hwy 36	40.365 66	-122.85 8	United States	Califor nia	Shasta	05/22/2 005	A. Stockm an
MY3175	*Aptostichus barackobamai*	Lower Springs Rd, 0.5 mi S of Hwy 299	40.580 76	-122.45 019	United States	Califor nia	Shasta	05/23/2 005	A. Stockm an
MY3621	*Aptostichus barackobamai*	Hwy32, S of Deer Creek, 3.3mi SW Potato	40.459 12	-121.78 287	United States	Califor nia	Teham a	12/19/2 007	M Hedin, J Starrett, D Leavitt

		Patch CG							
MY3622	*Aptostichus barackobamai*	Hwy32, S of Deer Creek, 3.3mi SW Potato Patch CG	40.15993	-121.5704	United States	California	Tehama	12/19/2007	M Hedin, J Starrett, D Leavitt
MY3803	*Aptostichus barackobamai*	Cotton wood	40.316774	-122.34998	United States	California	Tehama	10/15/2009	C.S. Will
MY729	*Aptostichus barackobamai*	Sutter Buttes, Dean Place	39.223066	-121.78128	United States	California	Sutter	04/04/2003	Hedin, Paquin, Starrett
BME102752	*Aptostichus icenoglei* (Central)	Cleveland Forest Rd, Cleveland NF	33.5289	-117.3885	United States	California	Riverside	5/13/2021	L. Newton, J. Starrett
BME102753	*Aptostichus icenoglei* (Central)	Cleveland Forest Rd, Cleveland NF	33.5289	-117.3885	United States	California	Riverside	5/13/2021	L. Newton, J. Starrett
MY2465	*Aptostichus icenoglei* (Central)	North of Fallbrook on DeLuz Road	33.41095	-117.28984	United States	California	San Diego	01/30/2004	J. Bond & M.Hedin

MY2467	*Aptostichus icenoglei* (Central)	North of Fallbrook on DeLuz Road	33.41095	-117.28984	United States	California	San Diego	01/30/2004	J. Bond & M.Hedin
MY2480	*Aptostichus icenoglei* (Central)	Bautista Canyon, along Bautista Canyon Road	33.70998	-116.87756	United States	California	Riverside	02/01/2004	J. Bond
MY2492	*Aptostichus icenoglei* (Central)	Ortega HWY H74, ~1.7 miles North Orange Co/Riverside Co line	33.61276	-117.43462	United States	California	Riverside	02/02/2004	J. Bond
MY2505	*Aptostichus icenoglei* (Central)	Winchester, just east of Icenogle residence, end of Grand Ave	33.71568	-117.09365	United States	California	Riverside	01/29/2004	J. Bond
MY2512	*Aptostichus icenoglei* (Central)	Winchester, just east of Icenogl	33.71568	-117.09365	United States	California	Riverside	01/29/2004	J. Bond

		e residence, end of Grand Ave							
MY2523	*Aptostichus icenoglei* (Central)	Winchester, just east of Icenogle residence, end of Grand Ave	33.71568	-117.09365	United States	California	Riverside	01/29/2004	J. Bond
MY2597	*Aptostichus icenoglei* (Central)	Winchester, Leona Rd ~1.0 m South of intersection with Patton Ave	33.67712	-117.11578	United States	California	Riverside	03/13/2004	J. Bond, W. Icenogle, et al
MY2668	*Aptostichus icenoglei* (Central)	Cleveland National Forest, along H74	33.629719	-117.42525	United States	California	Riverside	03/18/2004	J. Bond, C. Spruill, D. Beamer
MY2669	*Aptostichus icenoglei* (Central)	Cleveland National	33.629719	-117.42525	United States	California	Riverside	03/18/2004	J. Bond, C. Spruill,

		Forest, along H74							D. Beamer
MY3776	*Aptostichus icenoglei* (Central)	Winchester, Grand Ave ~.6 mile east intersection of Grand and Matthews	33.714781	-117.11009	United States	California	Riverside	05/17/2009	J. Bond
MY3777	*Aptostichus icenoglei* (Central)	Winchester, Grand Ave ~.6 mile east intersection of Grand and Matthews	33.714781	-117.11009	United States	California	Riverside	05/17/2009	J. Bond
MY718	*Aptostichus icenoglei* (Central)	De Luz Murrieta Road	33.49565	-117.24338	United States	California	Riverside	01/11/2003	M. Hedin
BME102526	*Aptostichus icenoglei* (North)	Angeles NF, Angeles Forest Hwy (N3) near jct w/ Big Tujunga	34.2907	-118.1706	United States	California	Los Angeles	5/11/2021	L. Newton, J. Bond, J. Starrett, L. Chamberland

		Canyon Rd								
BME102534	*Aptostichus icenoglei* (North)	Cleghorn Rd, San Bernardino NF, Silverwood Lake Rec Area, managed by CA State Parks	34.2853	-117.3728	United States	California	San Bernardino	5/12/2021	L. Newton, J. Bond, J. Starrett, L. Chamberland
BME102535	*Aptostichus icenoglei* (North)	Cleghorn Rd, San Bernardino NF, Silverwood Lake Rec Area, managed by CA State Parks	34.2853	-117.3728	United States	California	San Bernardino	5/12/2021	L. Newton, J. Bond, J. Starrett, L. Chamberland
BME102536	*Aptostichus icenoglei* (North)	Cleghorn Rd, San Bernardino NF, Silverwood Lake	34.2853	-117.3728	United States	California	San Bernardino	5/12/2021	L. Newton, J. Bond, J. Starrett, L.

		Rec Area, managed by CA State Parks							Chamberland
BME102537	*Aptostichus icenoglei* (North)	CA Hwy 138, San Bernardino NF	34.2588	-117.2944	United States	California	San Bernardino	5/12/2021	L. Newton, J. Bond, J. Starrett, L. Chamberland
BME102748	*Aptostichus icenoglei* (North)	CA Hwy 38	34.0767	-117.0568	United States	California	San Bernardino	5/12/2021	L. Newton, J. Bond, J. Starrett, L. Chamberland
MY2600	*Aptostichus icenoglei* (North)	Puente Hills, intersection of Azusa & Tomich Rd	33.98161	-117.93351	United States	California	Los Angeles	03/14/2004	J. Bond, C. Spruill, D. Beamer
MY3759	*Aptostichus icenoglei* (North)	Mt. Baldy Rd, ~0.2 km N of jct w/ N Mountain Ave	34.1773	-117.6767	United States	California	Los Angeles	02/15/2009	M Hedin, J Satler, J Starrett, C Richart

MY3763	*Aptostichus icenoglei* (North)	Lytle Creek Rd, near Scotland	34.244	-117.4952	United States	California	San Bernardino	02/15/2009	M Hedin, J Satler, J Starrett, C Richart
BME102828	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102829	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102830	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102831	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas,

									B. Gibson
BME102832	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102833	*Aptostichus icenoglei* (South)	CA Hwy 76, Morettis Junction	33.2015	-116.7118	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102837	*Aptostichus icenoglei* (South)	CA Hwy 76, Cleveland NF, across from picnic area	33.2533	-116.7922	United States	California	San Diego	6/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME102842	*Aptostichus icenoglei* (South)	Lyons Valley Rd	32.7526	-116.6715	United States	California	San Diego	6/23/2021	L. Newton, J. Starrett, R. Ruedas
BME102844	*Aptostichus icenoglei* (South)	Lyons Valley Rd	32.7526	-116.6715	United States	California	San Diego	6/23/2021	L. Newton, J. Starrett,

									R. Ruedas
BME102845	*Aptostichus icenoglei* (South)	Otay Mtn Ecological Reserve, ~1 mile up trail	32.6365	-116.884	United States	California	San Diego	6/23/2021	L. Newton, J. Starrett, R. Ruedas, M. Hedin
BME102847	*Aptostichus icenoglei* (South)	El Monte Rd	32.8838	-116.8214	United States	California	San Diego	6/24/2021	L. Newton, J. Starrett, M. Hedin
BME102851	*Aptostichus icenoglei* (South)	Torrey Pines State Reserve Extension, Mar Scenic Trail	32.9459	-117.2543	United States	California	San Diego	6/24/2021	L. Newton, J. Starrett
MY305	*Aptostichus icenoglei* (South)	El Monte Park Rd.	32.88395	-116.82145	United States	California	San Diego	01/19/2002	M.C. Hedin
MY306	*Aptostichus icenoglei* (South)	El Monte Park Rd.	32.88395	-116.82145	United States	California	San Diego	01/19/2002	M.C. Hedin
MY3635	*Aptostichus icenoglei* (South)	E. Lakeside between El	32.88369	-116.82239	United States	California	San Diego	02/23/2008	M.C. Hedin

		Monte Park & entrance to El Capitan Res, El Monte Rd								
MY719	*Aptostichus icenoglei* (South)	Nate Harrison Grade Road, 2.1 mi E jnct w/HWY 76	33.32697	-116.96523	United States	California	San Diego	01/11/2003	M. Hedin
MY3824	*Aptostichus isabella*	Erskine Creek Rd., 3.5 mi. E of int. w/ Lake Isabella Blvd., E of Bodfish	35.5689	-118.4383	United States	California	Kern	10/08/2010	J. Satler

Table 5. Morphological measurements of each *A. icenoglei* lineage.

Voucher	Clade/Gr	Carapace	Carapace	Posterior	Sternum	Sternum	Labium	Labium	Femur I	Patella I	Tibia I	Metatars	Tarsus I	Femur	Patella	Tibia IV	Metatars	Tarsus	Tibia I	Tibia I	Tibia I	Patella I	Metatars	Metatars	Palpal	Palpal	Bulb L
AP 886	South	4.784	3.925	0.863	2.657	2.213	0.355	0.661	4.793	2.405	3.372	3.12	2.114	4.811	2.072	4.2	4.668	2.588	3	2	6	1	2	0	1.873	0.727	1.005

AP 825 Central	AP 989 Central	AP 927 South	AP 878 South	AP 914 South	AP 019 South	AP 941 South	AP 1069 South	AP 1092 South	AP 912a South	AP 912b South
5.33	4.717	4.577	5.868	4.801	4.947	5.197	5.836	5.603	4.929	5.811
4.35	3.883	3.84	4.526	3.911	3.932	4.298	4.741	4.69	3.925	4.749
0.88	0.883	0.812	0.901	0.831	0.888	0.853	0.977	1.009	0.874	1.026
3.028	2.745	2.51	3.01	2.622	2.633	2.859	3.26	3.142	2.816	3.167
2.272	1.988	2.05	2.367	2.106	2.196	2.229	2.55	2.567	2.241	2.493
0.391	0.338	0.413	0.481	0.351	0.33	0.362	0.443	0.358	0.373	0.417
0.723	0.717	0.641	0.78	0.648	0.679	0.568	0.755	0.795	0.621	0.834
5.231	5.095	4.258	5.146	4.693	4.675	4.948	5.812	5.363	4.825	5.393
2.655	2.413	2.123	2.656	2.245	2.23	2.442	2.905	2.627	2.308	2.663
3.992	3.893	3.224	3.682	3.248	3.378	3.429	4.141	3.968	3.556	3.946
3.767	3.655	3.234	3.526	3.261	3.335	3.279	4.067	3.809	3.277	3.707
2.683	2.594	2.263	2.857	2.346	2.339	2.463	2.961	2.796	2.49	2.843
5.316	5.136	4.562	5.468	4.852	4.88	5.195	5.658	5.788	5.154	5.619
2.302	2.24	2.027	2.533	2.07	2.045	2.35	2.603	2.425	2.196	2.525
4.491	4.496	3.981	4.406	4.249	4.204	4.499	5.129	4.729	4.535	4.814
5.068	5.131	4.421	4.839	4.484	4.627	4.769	5.619	5.137	4.613	5.288
2.647	2.596	2.369	2.733	2.523	2.329	2.397	2.78	2.602	2.731	3.106
3	3	2	3	2	3	3	3	2	3	3
2	2	0	2	3	2	2	2	0	2	2
7	7	8	5	6	7	7	7	8	8	7
1	1	0	1	0	1	1	1	1	1	1
2	1	1	2	0	2	1	1	1	2	2
0	0	0	0	1	0	0	0	0	0	0
2.129	2.137	1.894	2.16	1.974	1.995	2.031	2.492	2.306	2.041	2.225
0.727	0.682	0.679	0.765	0.678	0.686	0.766	0.786	0.794	0.706	0.764
0.802	1.024	0.876	1.086	0.858	0.884	1.038	1.168	1.163	1.025	1.057

AP 32 North	AP 1171 North	AP 1215 North	AP 899 Central	AP 1144 Central	AP 875 Central	AP 1134 Central	AP 963 Central	AP 979 Central	AP 1113 Central	AP 1231 Central
5.43	5.016	6.029	5.323	5.88	4.846	4.541	5.768	5.88	5.983	4.778
4.303	4.081	4.809	4.46	4.935	3.787	3.811	4.621	4.715	4.866	3.918
0.933	0.856	0.964	0.905	1.047	0.854	0.772	0.97	1.032	1.007	0.85
2.888	2.917	3.273	3.117	3.359	2.71	2.627	3.224	3.257	3.14	2.874
2.334	2.141	2.561	2.342	2.516	2.006	2.08	2.405	2.608	2.39	2.233
0.321	0.327	0.399	0.386	0.303	0.298	0.232	0.381	0.408	0.341	0.272
0.762	0.851	0.787	0.707	0.783	0.553	0.633	0.757	0.753	0.754	0.72
5.379	4.99	5.479	5.29	5.543	4.657	4.68	5.591	5.712	5.448	4.978
2.65	2.38	2.504	2.684	2.743	2.222	2.292	2.695	2.883	2.747	2.409
3.968	3.993	4.203	4.085	4.298	3.61	3.604	4.168	4.507	4.142	3.696
3.899	3.479	4.07	3.916	4.171	3.556	3.474	4.114	4.163	3.589	3.48
2.617	2.57	2.787	2.759	2.813	2.261	2.535	2.82	2.969	2.586	2.616
5.484	4.626	6.028	5.532	5.764	4.741	4.752	5.667	5.605	5.678	5.241
2.323	2.059	2.652	2.403	2.454	2.052	1.876	2.375	2.557	2.377	2.108
4.715	4.502	5.078	4.774	4.99	4.267	4.141	4.664	4.776	4.871	4.478
5.245	5.227	5.757	5.091	5.46	4.562	4.56	5.616	5.739	5.415	5.053
3.036	2.684	2.994	2.604	3.035	2.557	2.538	2.694	3.024	3.015	2.782
2	2	1	2	2	3	3	2	3	3	2
2	2	1	1	2	2	2	2	2	2	2
8	6	9	8	8	8	7	7	9	7	7
1	0	1	0	2	0	0	1	2	1	0
1	1	0	1	2	1	1	3	2	1	1
0	0	0	0	0	0	0	0	0	0	0
2.216	2.143	2.366	2.275	2.335	2.037	1.976	2.205	2.35	2.224	2.073
0.682	0.687	0.771	0.743	0.779	0.649	0.645	0.741	0.749	0.762	0.662
1.035	0.967	1.173	1.067	1.079	0.951	0.912	1.087	1.102	1.07	0.945

AP 21a North	AP 25 North	AP 1168 North	AP 1242 North	AP 021b North	AP 49 North	AP 1173 North
5.744	6.082	5.361	5.275	4.157	5.935	4.951
4.614	4.768	4.294	4.34	3.492	4.794	4.205
0.968	0.986	0.929	0.965	0.812	0.92	0.956
3.165	3.277	2.957	3.006	2.468	3.37	2.984
2.532	2.622	2.258	1.791	1.783	2.367	2.107
0.368	0.403	0.373	0.245	0.335	0.353	0.394
0.813	0.847	0.708	0.734	0.576	0.68	0.731
5.017	5.809	4.841	5.254	4.279	5.651	5.309
2.596	2.32	2.53	2.513	2.252	2.717	2.41
3.688	4.642	3.688	4.003	3.439	4.37	4.036
3.685	4.051	3.424	3.743	3.207	3.925	3.769
2.549	2.983	2.487	2.735	2.238	2.844	2.669
5.418	5.9	5.019	5.489	4.142	5.622	5.462
2.448	2.115	2.038	2.377	1.878	2.441	2.274
4.547	5.109	4.351	4.696	3.805	5.111	4.609
4.784	5.758	4.799	5.217	4.277	5.732	5.378
2.79	3.052	2.326	2.642	2.386	2.902	2.906
3	3	2	2	2	3	2
1	3	2	1	0	2	0
7	6	8	8	6	8	6
1	2	0	1	0	1	0
1	1	1	1	2	1	1
0	0	0	0	0	0	0
2.128	2.345	2.041	1.969	1.792	2.351	2.287
0.745	0.727	0.672	0.718	0.613	0.776	0.726
1.073	1.124	1.022	0.893	0.848	0.992	1.059

Table 6. Bioclimatic variables used and their percent contribution to each species distribution model.

Bioclimatic Variables	% Contribution			
	North	**Central+South**	**Central**	**South**
bio2 - mean diurnal range (mean of monthly (max temp - min temp))	16.6163	0.8843	2.2109	0
bio4 - temperature seasonality (standard deviation * 100)	0	25.2361	15.3477	45.2929

bio8 - mean temperature of wettest quarter	0	1.7744	0.2544	0.0687
bio9 - mean temperature of driest quarter	6.4207	0.8392	0.3835	0.1458
bio11 - mean temperature of coldest quarter	0	13.4691	13.5075	18.3141
bio14 - precipitation of driest month	2.5498	24.776	50.2496	23.4304
bio15 - precipitation seasonality (coefficient of variation)	16.6163	13.2709	3.4537	4.2049
bio17 - precipitation of driest quarter	0.2437	0.2556	0	0.2486
bio18 - precipitation of warmest quarter	0	4.1435	2.0492	7.0691
bio19 - precipitation of coldest quarter	57.5531	15.3509	12.5434	1.2254

Table 7. MaxEnt model parameters and stats for the North lineage.

tune.args	auc.train	auc.diff.avg	auc.diff.sd	auc.val.avg	auc.val.sd	AICc	delta.AICc
rm.1_fc.L	0.9629851	0.0477946	0.3923257	0.9498648	0.4455652	549.88471	12.083926
rm.1.5_fc.L	0.9628444	0.0477676	0.4120915	0.9495444	0.4622118	551.22761	13.426829
rm.2_fc.L	0.9626296	0.0477817	0.4310545	0.9489740	0.4784984	548.42725	10.626462
rm.2.5_fc.L	0.9627148	0.0474170	0.4414480	0.9484629	0.4865563	546.15603	8.3552488

rm.3_fc.L	0.9626888 9	0.0473946	0.4543244	0.9477296 3	0.4974188 8	547.36302 2	9.5622351 8
rm.1_fc.L Q	0.9678111 1	0.0597511 4	0.6828543 5	0.9401074 1	0.7241791 4	554.74619 7	16.945410 2
rm.1.5_fc. LQ	0.9652814 8	0.0544054 1	0.5477880 6	0.9431592 6	0.5932112 2	550.94303 4	13.142246 6
rm.2_fc.L Q	0.9642425 9	0.0524532 8	0.5286328 6	0.94455	0.5728645 7	550.28026 5	12.479477 6
rm.2.5_fc. LQ	0.9641240 7	0.0506226 5	0.5101353 8	0.9455870 4	0.5530151 4	551.46850 2	13.667715 4
rm.3_fc.L Q	0.9640629 6	0.0492687 3	0.4954859 1	0.9459722 2	0.5370412 1	549.02467 4	11.223887 2
rm.1_fc.H	0.9762555 6	0.0331965 1	0.2604949 2	0.9638555 6	0.2967624	800.70288 4	262.90209 7
rm.1.5_fc. H	0.9748055 6	0.0336431 6	0.2389182 1	0.9655203 7	0.2828658 3	610.70956 8	72.908780 7
rm.2_fc.H	0.9734666 7	0.0350408 8	0.2385859 7	0.9648444 4	0.2866404 7	599.25091 9	61.450131 6
rm.2.5_fc. H	0.9721296 3	0.0362104 7	0.2425729	0.9640407 4	0.2930054 1	619.37925 5	81.578468 3
rm.3_fc.H	0.9711148 1	0.0364767 8	0.2397286 6	0.9642666 7	0.2917815 9	575.82525	38.024462 8
rm.1_fc.L QH	0.9762777 8	0.0606836 2	0.8874413 8	0.9357481 5	0.9113453 8	625.57934 4	87.778556 8
rm.1.5_fc. LQH	0.9735222 2	0.0609995	0.8428971	0.9373666 7	0.8715909 1	544.28720 4	6.4864169 7
rm.2_fc.L QH	0.9729074 1	0.0593517 8	0.7805355 4	0.9400185 2	0.8120470 6	546.73812 4	8.9373369 5
rm.2.5_fc. LQH	0.9721703 7	0.0589411 7	0.7560956 9	0.9404555 6	0.7891223 3	549.27703 4	11.476247 2

rm.3_fc.LQH	0.9715296	0.0580315	0.7187361	0.9410963	0.7532752	551.98781	14.187031
rm.1_fc.LQHP	0.9786814	0.0454836	0.5465515	0.9496592	0.5709002	589.07075	51.269964
rm.1.5_fc.LQHP	0.9772185	0.0424527	0.4511517	0.9537666	0.4809244	555.46300	17.662218
rm.2_fc.LQHP	0.9764740	0.0408289	0.3895596	0.9559703	0.4235067	542.57431	4.7735297
rm.2.5_fc.LQHP	0.9752592	0.0400275	0.3495182	0.9574166	0.3878705	546.65298	8.8521978
rm.3_fc.LQHP	0.9737481	0.0388396	0.3172238	0.9582388	0.3579212	537.80078	0
rm.1_fc.LQHPT	0.9824759	0.0439695	0.5252246	0.9512129	0.5470276	592.62825	54.827471
rm.1.5_fc.LQHPT	0.9786611	0.0407871	0.4378645	0.9549351	0.4656726	557.37636	19.575573
rm.2_fc.LQHPT	0.9770981	0.0388585	0.3807904	0.9576537	0.4126648	538.63203	0.8312473
rm.2.5_fc.LQHPT	0.9760185	0.0381433	0.3444891	0.9585277	0.3796265	543.62495	5.8241643
rm.3_fc.LQHPT	0.9748074	0.0374591	0.3164356	0.9591055	0.3542870	543.02911	5.2283244

Table 8. MaxEnt model parameters and stats for the Central+South lineage.

tune.args	auc.train	auc.diff.avg	auc.diff.sd	auc.val.avg	auc.val.sd	AICc	delta.AICc
rm.1_fc.L	0.9291128	0.0541256	0.3771339	0.9258397	0.5998118	1646.1697	59.499740
rm.1.5_fc.L	0.9285615	0.0550577	0.3858884	0.9252717	0.6117928	1647.5168	60.846818

rm.2_fc.L	0.9279897 4	0.0561398 2	0.3990187 5	0.9244269 2	0.6271908 3	1651.5676 2	64.897606 5
rm.2.5_fc. L	0.9273615 4	0.0572990 2	0.4141694	0.9235769 2	0.6452634 9	1653.3367 4	66.666721 1
rm.3_fc.L	0.9264076 9	0.0583372 2	0.4229683 7	0.9231410 3	0.6574612 1	1655.2238 3	68.553810 1
rm.1_fc.L Q	0.9486346 2	0.0368368 9	0.2541046 4	0.9449602 6	0.405763	1607.0571 4	20.387120 2
rm.1.5_fc. LQ	0.9466153 8	0.0402281 6	0.2709620 6	0.9421730 8	0.4398240 5	1613.0727	26.402687 6
rm.2_fc.L Q	0.9438166 7	0.0428591 4	0.2782847 3	0.9397435 9	0.4620282 3	1620.8646 2	34.194601 8
rm.2.5_fc. LQ	0.9422371 8	0.0441936 7	0.2866835	0.9382397 4	0.4762874 5	1622.6654 3	35.995411 4
rm.3_fc.L Q	0.940775	0.0456288 5	0.2987932 4	0.9370205 1	0.4939026	1623.7974 1	37.127396 3
rm.1_fc.H	0.9655865 4	0.0362674 2	0.2356075 8	0.9593435 9	0.3895928 1	1586.6700 2	0
rm.1.5_fc. H	0.9619762 8	0.0389595 8	0.2467847 2	0.9560717 9	0.4153076 7	1594.9404 1	8.2703909 9
rm.2_fc.H	0.9591557 7	0.0412849 8	0.2523317	0.9543076 9	0.4358438 6	1603.3535 9	16.683575 4
rm.2.5_fc. H	0.9574378 2	0.0427319 8	0.2667521 3	0.9521673 1	0.4545210 8	1616.1943 4	29.524320 3
rm.3_fc.H	0.9552397 4	0.0437811	0.2764279 4	0.9510775 6	0.4684025 7	1615.5369 9	28.866974 3
rm.1_fc.L QH	0.96825	0.0330674 2	0.2355062 4	0.9603115 4	0.3648360 8	1610.0417 6	23.371745 6
rm.1.5_fc. LQH	0.9604525 6	0.0363532	0.2251430 1	0.9549076 9	0.3840634 4	1599.5168 9	12.846869 7

rm.2_fc.LQH	0.9558064	0.0388500	0.2379061	0.9508769	0.4099038	1599.9395	13.269554
	1	6	7	2	9	7	7
rm.2.5_fc.LQH	0.9527564	0.0405723	0.2498123	0.9478910	0.4288992	1610.7104	24.040437
	1		5	3	1	5	3
rm.3_fc.LQH	0.9494141	0.0413396	0.2617715	0.9450307	0.4418661	1610.9090	24.239033
		6	9	7	9	5	8
rm.1_fc.LQHP	0.9693769	0.0324988	0.2516660	0.9629833	0.3730060	1615.3478	28.677859
	2	3	5	3	8	7	5
rm.1.5_fc.LQHP	0.9651846	0.0366311	0.2479068	0.9581756	0.3992946	1603.7779	17.107914
	2	2		4	1	3	3
rm.2_fc.LQHP	0.9588461	0.0410485	0.2463940	0.9525871	0.4309694	1602.9478	16.277878
	5	4	2	8	7	9	6
rm.2.5_fc.LQHP	0.9535288	0.0427562	0.2504532	0.9494974	0.4449998	1601.7671	15.097168
	5	6	2	4	3	8	1
rm.3_fc.LQHP	0.9516865	0.0436959	0.2586165	0.9477807	0.4579731	1613.2856	26.615659
	4	2	1	7	8	8	9
rm.1_fc.LQHPT	0.9761230	0.0306473	0.2686300	0.9657025	0.3660955	1640.7173	54.047316
	8	2	2	6	6	3	4
rm.1.5_fc.LQHPT	0.9702564	0.0366132	0.2629582	0.9603807	0.4046657	1604.1553	17.485377
	1	5	4	7	3	9	7
rm.2_fc.LQHPT	0.9623435	0.0415952	0.2632919	0.9543371	0.4423781	1608.3208	21.650851
	9	4	2	8	4	7	6
rm.2.5_fc.LQHPT	0.9560769	0.0437276	0.2613493	0.9507974	0.4570601	1611.8483	25.178283
	2	6	7	4	8		4
rm.3_fc.LQHPT	0.9534448	0.0444206	0.2634580	0.9486993	0.4649338	1608.3079	21.637894
	7	9	6	6	5	1	1

Table 9. MaxEnt model parameters and stats for the Central lineage.

tune.args	auc.train	auc.diff.avg	auc.diff.sd	auc.val.avg	auc.val.sd	AICc	delta.AICc

rm.1_fc.L	0.9439291 7	0.0543644 4	0.3882350 8	0.9370763 9	0.4955913 6	756.96212 9	43.702322 6
rm.1.5_fc. L	0.9437152 8	0.0525440 5	0.3556072 3	0.9383152 8	0.4647441 3	751.77725 7	38.517451 5
rm.2_fc.L	0.9444625	0.0511104 8	0.3318123 9	0.9394736 1	0.4416010 4	746.19612 5	32.936318 6
rm.2.5_fc. L	0.9451375	0.0499348 8	0.3145186 5	0.9402958 3	0.4241114 2	746.6486	33.388793 9
rm.3_fc.L	0.9457041 7	0.0489957 9	0.2955073 9	0.9412111 1	0.4061378 8	747.15569 1	33.895885
rm.1_fc.LQ	0.9588458 3	0.0481094 4	0.3859591 1	0.9479069 4	0.4645682	735.64498 5	22.385179 2
rm.1.5_fc. LQ	0.9567013 9	0.0457511 9	0.2677988 5	0.9494402 8	0.3700486 3	733.63702 7	20.377221 3
rm.2_fc.LQ	0.9546430 6	0.0452102 4	0.2108407 5	0.9490791 7	0.3319668 5	738.30160 3	25.041797 4
rm.2.5_fc. LQ	0.9528236 1	0.0458309 5	0.2029459 2	0.9477541 7	0.3296426 7	738.64352 6	25.383719 7
rm.3_fc.LQ	0.9517069 4	0.0459457 9	0.1969850 5	0.9468847 2	0.3273901 1	741.62379 4	28.363988 5
rm.1_fc.H	0.9769305 6	0.0303296 4	0.2989176 4	0.9673361 1	0.3385893 3	734.36601 8	21.106211 9
rm.1.5_fc. H	0.9744097 2	0.0293813 1	0.2280561 6	0.9682569 4	0.2784530 4	731.18947 1	17.929664 7
rm.2_fc.H	0.9724333 3	0.0289951 6	0.1735387 3	0.9681652 8	0.2365247 8	748.01323 8	34.753432 2
rm.2.5_fc. H	0.9708597 2	0.0295567 5	0.1484598 5	0.9673944 4	0.2226564 3	730.53907 8	17.279271 7
rm.3_fc.H	0.9690541 7	0.0306956	0.1311568 8	0.9658319 4	0.2175822 3	720.77335	7.5135436 6

rm.1_fc.LQH	0.9762097 2	0.0323095 2	0.4032461 6	0.9652097 2	0.4356389 3	718.51106 2	5.2512558 4
rm.1.5_fc.LQH	0.9714263 9	0.0341375 4	0.3493803 3	0.9635430 6	0.3941583 4	726.32595 2	13.066146 3
rm.2_fc.LQH	0.9693263 9	0.0347119	0.3101559 9	0.9629402 8	0.3624690 5	713.25980 6	0
rm.2.5_fc.LQH	0.9677305 6	0.0354073	0.2768168 4	0.9619541 7	0.3373925 1	718.84512 3	5.5853170 6
rm.3_fc.LQH	0.9660069 4	0.0368369 4	0.2435658 5	0.9601847 2	0.3169424 6	724.53439 2	11.274586
rm.1_fc.LQHP	0.9777430 6	0.0320955 6	0.4207485	0.9661930 6	0.4498349 6	715.39198	2.1321742 7
rm.1.5_fc.LQHP	0.9739736 1	0.0330015 1	0.3455787 2	0.9650458 3	0.3866296 2	723.33166 6	10.071860 4
rm.2_fc.LQHP	0.9703708 3	0.0329958 7	0.2435157 2	0.9647513 9	0.3037382 2	719.03614 3	5.7763369 5
rm.2.5_fc.LQHP	0.9679847 2	0.0350177 4	0.1910431 8	0.9620041 7	0.2737460 8	723.04647 5	9.7866694
rm.3_fc.LQHP	0.9639013 9	0.0380290 1	0.1616459 9	0.9591361 1	0.2667763 2	726.77287 5	13.513069 2
rm.1_fc.LQHPT	0.9846736 1	0.0285225 8	0.3658320 1	0.9722625	0.3906634 6	910.53346	197.27365 4
rm.1.5_fc.LQHPT	0.9780625	0.0339515 5	0.3533200 5	0.9658625	0.3936638 9	760.19398 6	46.934179 7
rm.2_fc.LQHPT	0.9717625	0.0340260 3	0.2561526 3	0.9649125	0.3146430 2	729.53473 6	16.274930 5
rm.2.5_fc.LQHPT	0.9694597 2	0.0352384 5	0.2086975 5	0.9629986 1	0.2845233 1	728.29766	15.037853 6
rm.3_fc.LQHPT	0.9657319 4	0.0381873	0.1721447 5	0.9597652 8	0.2717780 2	728.41608 2	15.156275 6

Table 10. MaxEnt model parameters and stats for the South lineage.

tune.args	auc.train	auc.diff.avg	auc.diff.sd	auc.val.avg	auc.val.sd	AICc	delta.AICc
rm.1_fc.L	0.9422583	0.0515193	0.2567954	0.9377345	0.4087404	876.99636	64.994147
rm.1.5_fc.L	0.9400654	0.0530493	0.2669897	0.9359988	0.4226837	877.07630	65.074092
rm.2_fc.L	0.9375559	0.0557154	0.2892786	0.9334869	0.4499205	880.86540	68.863190
rm.2.5_fc.L	0.9345607	0.0584057	0.3138422	0.9311916	0.4777822	881.80561	69.803404
rm.3_fc.L	0.9332369	0.0603526	0.3329055	0.9296559	0.5003034	880.76773	68.765517
rm.1_fc.LQ	0.9586607	0.0485257	0.2901524	0.9523011	0.4165409	848.07968	36.077466
rm.1.5_fc.LQ	0.9551297	0.0507955	0.2891957	0.9495916	0.4255187	856.39616	44.393945
rm.2_fc.LQ	0.9534869	0.0510214	0.2840440	0.9485416	0.4239162	851.85821	39.855998
rm.2.5_fc.LQ	0.9510416	0.0521827	0.2826356	0.9471869	0.4282558	857.95942	45.957215
rm.3_fc.LQ	0.9483321	0.0545451	0.2919219	0.9446702	0.4450242	860.86996	48.867753
rm.1_fc.H	0.9879904	0.0147384	0.1026645	0.9834095	0.1353453	839.21460	27.212395
rm.1.5_fc.H	0.9828059	0.0219336	0.1450593	0.9771559	0.2017487	825.66722	13.665005
rm.2_fc.H	0.9745154	0.0359667	0.2282897	0.9664654	0.3280060	834.41132	22.409111
rm.2.5_fc.	0.9658107	0.0460097	0.2492369	0.9617928	0.3840071	835.59170	23.589487

H	1	3	2	6	9	2	5
rm.3_fc.H	0.9616821 4	0.0513671 3	0.2676814 5	0.9589404 8	0.4192812 7	832.12381 8	20.121603 8
rm.1_fc.LQH	0.9863297 6	0.0179804 3	0.1308907 3	0.9805702 4	0.1697064 5	1042.7149 4	230.71272 9
rm.1.5_fc.LQH	0.9800940 5	0.0238850 8	0.1407316 6	0.9766821 4	0.2047539 2	819.44679 8	7.4445845 1
rm.2_fc.LQH	0.9744535 7	0.0332006 4	0.1880013 2	0.9696464 3	0.2802693 3	821.98997	9.9877559
rm.2.5_fc.LQH	0.9689440 5	0.0407816 2	0.2229754 6	0.9644738 1	0.3381648 8	826.06625 4	14.06404
rm.3_fc.LQH	0.9632583 3	0.0488021 5	0.2686720 6	0.9587083 3	0.4083274 5	825.61650 7	13.614293 5
rm.1_fc.LQHP	0.9875142 9	0.0141643 4	0.1024044	0.9839690 5	0.1331038 4	870.10981 3	58.107598 8
rm.1.5_fc.LQHP	0.9844297 6	0.0187702 1	0.1129537 8	0.9798154 8	0.1631254 7	837.98907 6	25.986861 5
rm.2_fc.LQHP	0.9787178 6	0.0258718 9	0.1415112 2	0.975125	0.2163621 4	812.00221 4	0
rm.2.5_fc.LQHP	0.9728988 1	0.0339680 6	0.1755675 8	0.9699607 1	0.2754224 1	822.48489 6	10.482682 5
rm.3_fc.LQHP	0.9695583 3	0.0374696 9	0.1903508 2	0.9675226 2	0.3009816 3	817.52450 4	5.5222903 4
rm.1_fc.LQHPT	0.9894809 5	0.0151662 6	0.1177814 9	0.9834809 5	0.1467849 5	813.67299 4	1.6707797 4
rm.1.5_fc.LQHPT	0.9863440 5	0.0182235 8	0.1188491 5	0.9807988 1	0.1622668	826.00961 8	14.007404 1
rm.2_fc.LQHPT	0.9812642 9	0.0247746 8	0.1348616 2	0.9756154 8	0.2066947 2	830.45254 1	18.450326 7

rm.2.5_fc.LQHPT	0.97460595	0.0320628	0.16397314	0.97108929	0.25980884	817.959958	5.95774382
rm.3_fc.LQHPT	0.97089167	0.03650499	0.18686754	0.96787738	0.29544553	821.706309	9.70409497

Table 11. Stats for UCE data.

Species	**Sample ID**	**Total Raw Reads**	**Total Cleaned Reads**	**Total Contigs**	**Total Base Pairs**	**Mean Contig Length**	**95 CI length**	**Min Contig Length**	**Max Contig Length**	**Median Length**	**Contigs > 1kb**
Aptostichus barackobamai	BME102225	4420078	7830491	1332	1384323	1039.3	11.7	229	7453	1018	708
Aptostichus barackobamai	BME102234	3812115	6628564	1303	1380869	1059.8	12.4	230	6012	1034	716
Aptostichus barackobamai	BME102237	3330201	5663593	1344	1517872	1129.4	14.1	230	7288	1074.5	827
Aptostichus barackobamai	BME102238	3707668	6442682	1302	1267103	973.2	10.4	155	5966	959.5	576
Aptostichus barackobamai	BME102239	3510383	6056168	1306	1262668	966.8	10.4	218	6109	962	571

591	733	540	296	744	665	759
963	1029	989	860.5	1137.5	1042	1171.5
6847	4637	8145	2887	2835	3456	3269
185	241	231	230	230	234	179
10.8	11.6	14.4	9.5	9.3	9.4	10.4
987.2	1055.8	1002.5	856.9	1102.3	1023.5	1155.9
1304052	1429568	1122821	843236	1216927	1225139	1241439
1321	1354	1120	984	1104	1197	1074
4722596	5867580	8457382	7602546	4686229	7912442	6578386
2715448	3388498	4821353	4304076	2415859	4268144	3350390
BME102240	BME102241	BME102242	BME102305	MY1098	MY3025	MY3026
Aptostichus barackobamai	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*

926	924	897	989	989	1128	1089
1352	1304	1344	1470	1509.5	1512	1500
6358	5002	3869	7753	6611	4614	4498
242	210	204	219	240	230	231
14.5	12.4	13.3	15.3	15.5	12.8	12.9
1351	1287.9	1329.2	1460	1487.3	1486	1478.6
1554991	1506863	1487339	1680454	1707380	1893179	1827553
1151	1170	1119	1151	1148	1274	1236
6744182	8233007	8520364	8180139	8027550	8196361	8629470
3451363	4215949	4358520	4188958	4111218	4191434	4413976
MY3027	MY3038	MY3158	MY3173	MY3175	MY3621	MY3622
Aptostichus barackobamai	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*	*Aptostichus barackobamai*

950	1073	2	678	941	588	328	558
1278	1580	290	1115	1186.5	1022	863	1000
5061	7389	1077	7522	6235	2537	4290	2344
239	229	229	230	230	235	233	231
11.7	15.2	5.4	14.5	12.2	9.2	8.8	8.6
1275	1548	321.6	1126.6	1211.7	1001.4	873.5	982.8
1538978	1857610	112875	1186312	1543663	1101546	1037676	1094827
1207	1200	351	1053	1274	1100	1188	1114
9700833	5535308	410418	5908597	4822054	7327284	3855530	5422034
4948423	2838864	255356	3018802	2465496	3745511	1992694	2776732
MY3803	MY729	MY3824	BME102535	BME102536	BME102537	BME102748	BME102752
Aptostichus barackobamai	*Aptostichus barackobamai*	*Aptostichus isabella*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*

573	247	262	295	156	105	347	347	196
986	831	867	858	761	732	881	888	788
3599	3267	3042	3852	5540	3176	3204	2278	5750
229	191	227	193	154	191	229	230	238
9.1	8.7	9.7	10.3	9	7.1	8.8	9.5	8.8
970	831.5	859.7	865.7	770	737	885.7	890	789.7
1159099	991958	810743	909845	779982	899874	1024709	906935	818933
1195	1193	943	1051	1013	1221	1157	1019	1037
4417498	4487134	4665418	4113040	5155788	3408090	4279630	4411190	3812588
2776732	2316700	2417689	2128137	2671417	1759619	2212181	2299688	1975434
BME102753	BME102828	BME102829	BME102830	BME102831	BME102832	BME102833	BME102837	BME102842
Aptostichus icenoglei	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*

237	346	396	205	376	370	574	508	582
782	902.5	894	795.5	886.5	888	1032	967	1015.5
4227	3826	3031	7274	1929	2792	8634	4132	3993
232	229	237	232	178	168	230	225	173
9.7	10.4	9.4	11.2	10.2	10.3	15.1	12	10.7
799.6	908	908.7	818.4	895.9	891	1049.9	983.3	1011.3
952279	882606	1067725	841342	886950	902588	1126520	1085531	1110392
1191	972	1175	1028	990	1013	1073	1104	1098
3796994	3777170	4004664	3811742	9232542	9008121	9587452	8243968	8499404
1964343	1958630	2073595	1966455	10149464	9805774	5172058	4440918	4614117
BME102844	BME102845	BME102847	BME102851	MY2465	MY2467	MY2480	MY2492	MY2505
Aptostichus icenoglei	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*

349	594	654	743	433	731	328	467	264
865	1045	1101	1100	938.5	1126.5	858	951.5	829
2798	3665	3758	3905	2885	5716	2228	4356	3849
185	226	236	180	218	230	205	185	232
10.2	12.6	11.9	11	10.1	13.7	9.1	10.3	9
878.9	1066.4	1094.7	1096.5	938.3	1126.9	854.2	950	827.9
874528	1113337	1189932	1298298	964622	1250851	901193	1025976	1024072
995	1044	1087	1184	1028	1110	1055	1080	1237
7461057	8315839	9193740	8991862	7854572	8670518	7482832	8162959	4303471
4058619	4530508	4530508	4885983	4259227	4744563	4060054	4437521	2311741
MY2512	MY2523	MY2597	MY2600	MY2668	MY2669	MY305	MY306	MY3635
Aptostichus icenoglei	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*	*Aptostichus icenoglei*

Aptostichus icenoglei	MY3759	4785470	9033324	1031	814409
Aptostichus icenoglei	MY3763	4658003	8736611	1099	926173
Aptostichus icenoglei	MY3776	5701942	10709261	1031	829842
Aptostichus icenoglei	MY3777	5578622	10517313	1012	828068
Aptostichus icenoglei	MY718	5235415	9565653	1020	1054146
Aptostichus icenoglei	MY719	3890591	7047193	1061	1087100

789.9	8.4	119	1627	790	241
842.7	8.7	183	2113	849	341
804.9	8.4	220	1708	815	264
818.2	8.8	230	2579	824	272
1033.5	13.1	187	6305	1013	529
1024.6	11.1	186	2717	1026	572

Table 12. Summary of *Aptostichus icenoglei* cohesion species delimitation assessment. N_a and N_b values are the number of occurrence records for the first and second lineages used in a comparison, respectively.

Lineage Comparison	*Genetic Exchangeability*				
	Geographical barrier	**PCA (morphology)**	**VAE (molecules)**	**CLADES (molecules)**	**Conclusion**
Central to South	Parapatric, no obvious barrier	significant overlap	Small overlap of clusters	1 species	Fail to reject GE

North to Central + South	Parapatric, potential barrier (LA Basin)	significant overlap	Separate clusters	1 species	Reject GE
	Ecological Interchangeability				
	N_a, N_b	**Niche overlap value**	**Niche equivalency test**	**Niche similarity test**	**Conclusion**
Central to South	42, 55	0.4595	$p < 0.05$	$p < 0.05$, niche conservatism	Fail to reject EI
North to Central + South	29, 97	0.3873	$p < 0.05$	$p < 0.05$, niche conservatism	Fail to reject EI

Figures:

Figure 6. Geographic distributions of each lineage within the *Aptostichus icenoglei* sibling species complex. See legend denoting color for each lineage.

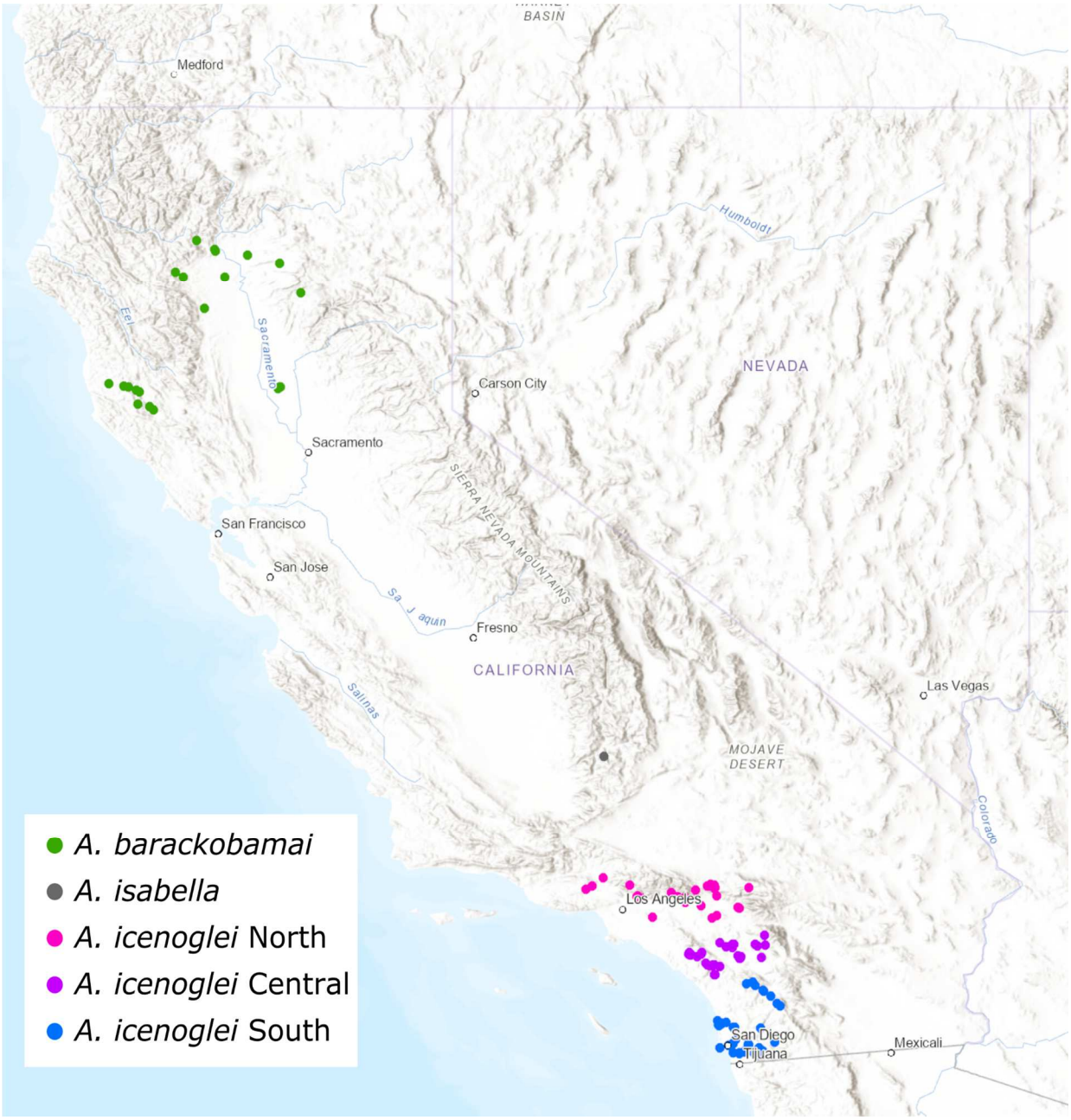


Figure 7. IQTree consensus tree for 90p data set. Bootstrap values with support < 90 are denoted by black nodes.

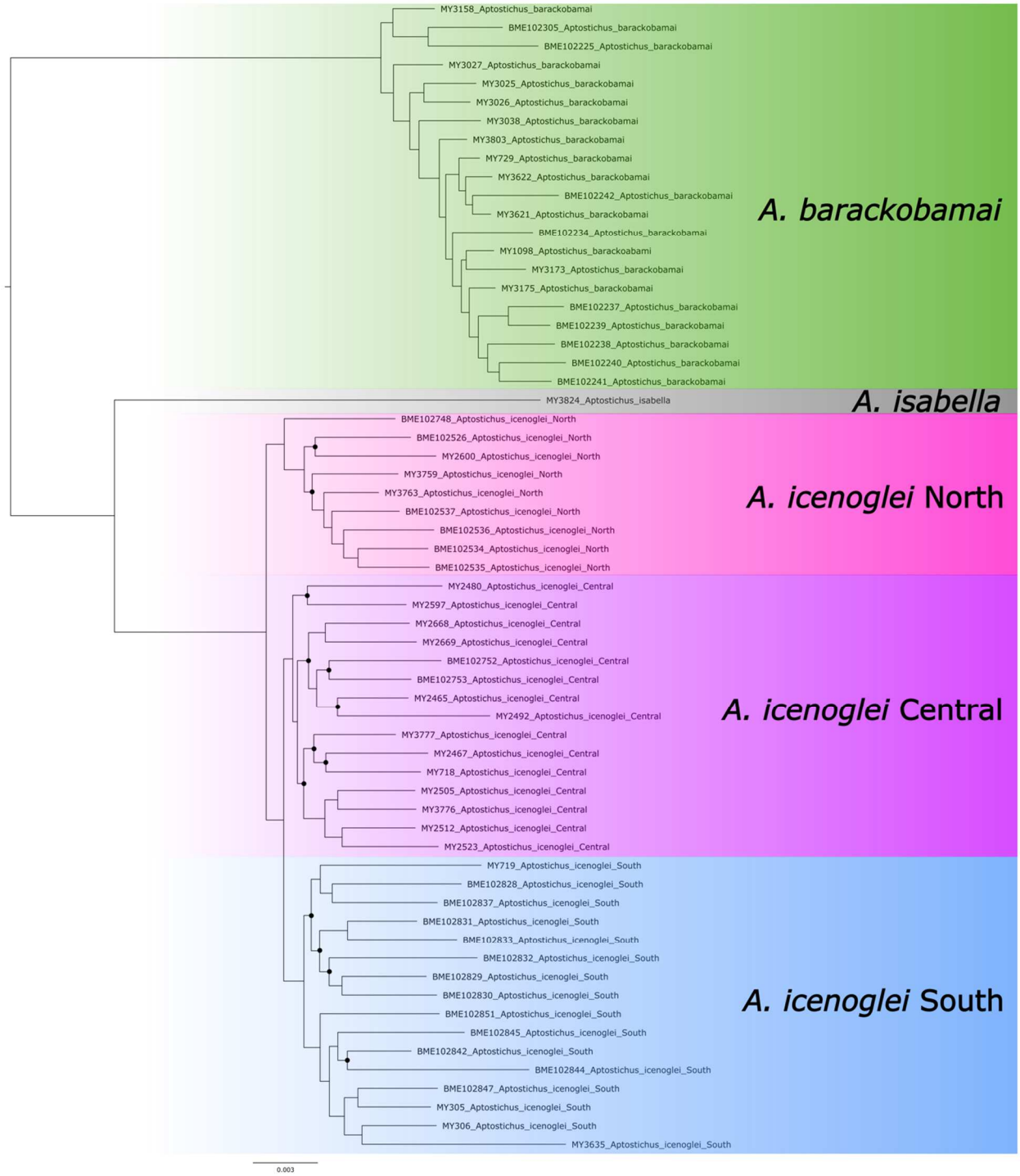


Figure 8. Raw results of the 50p IQ-TREE phylogeny.

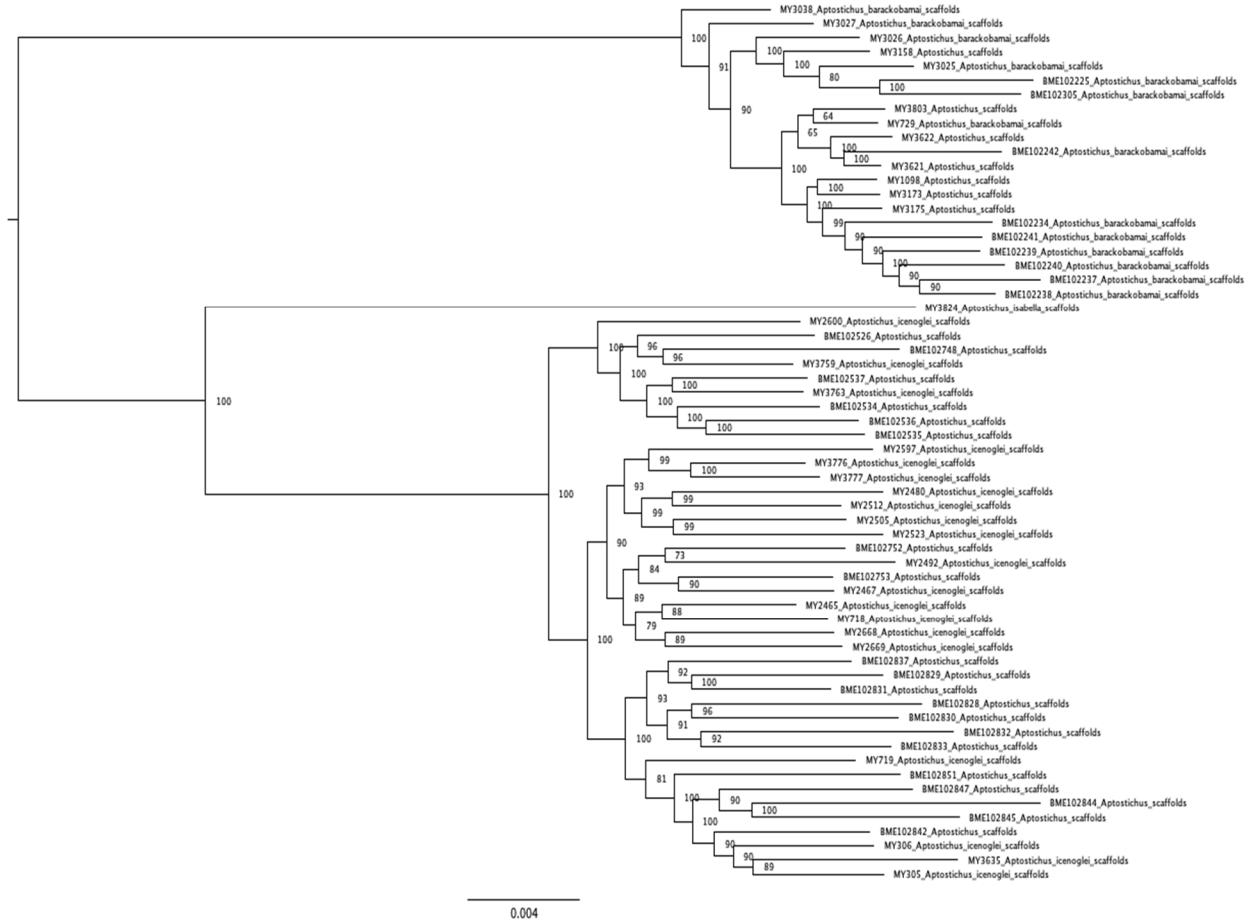


Figure 9. Raw results of the 75p IQ-TREE phylogeny.

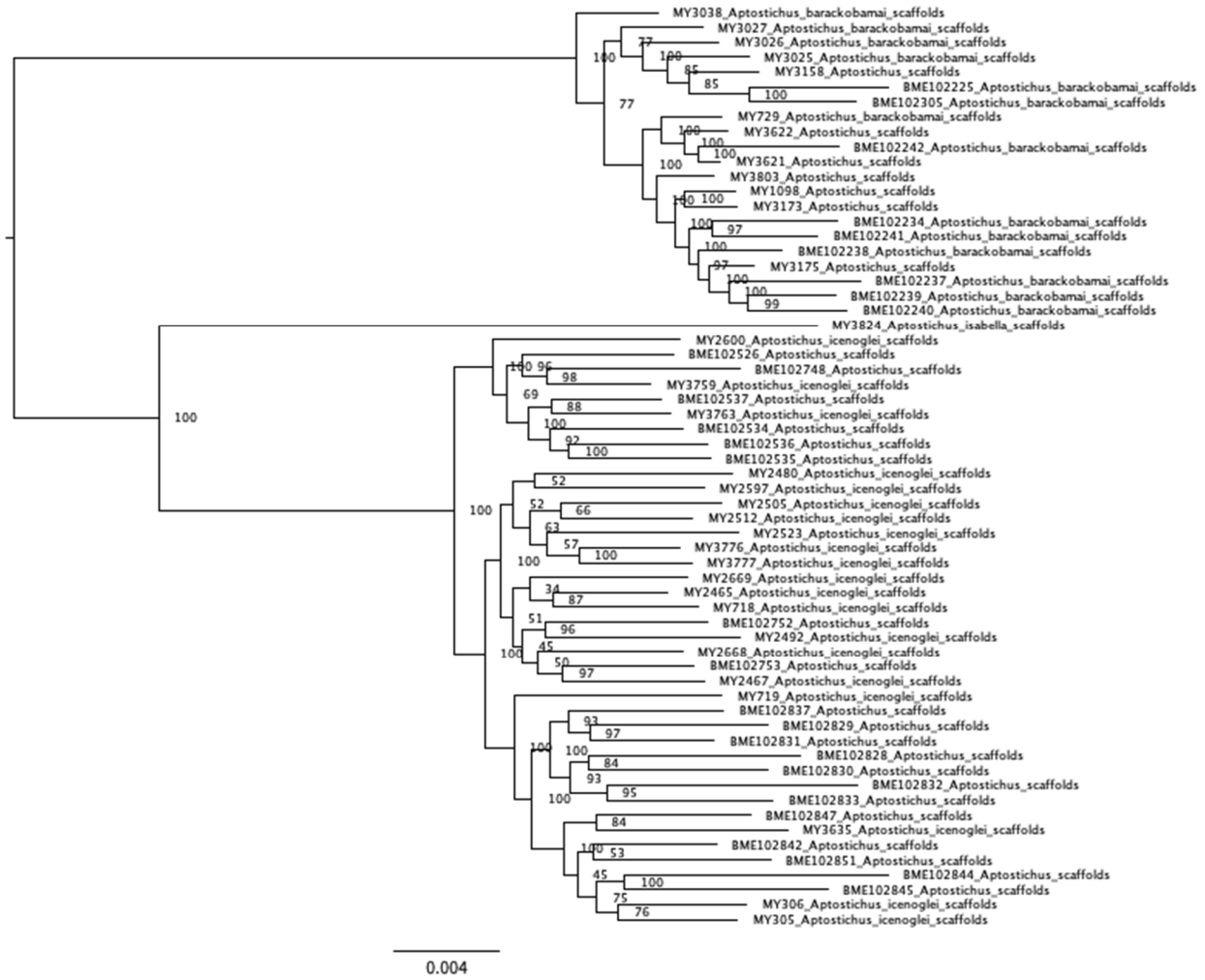


Figure 10. Raw results of the 75p ASTRAL phylogeny.

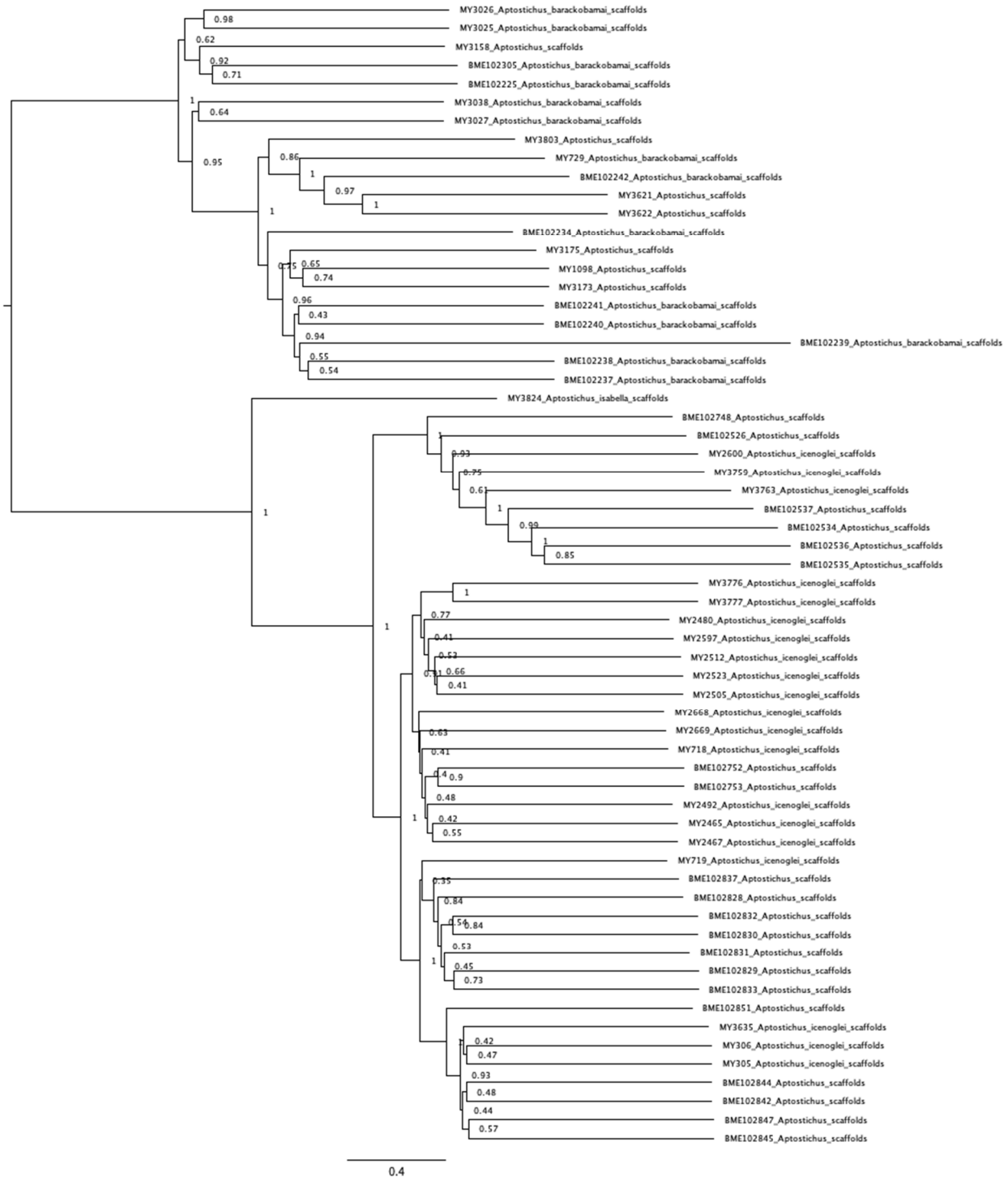

MY3026_Aptostichus_barackobamai_scaffolds
MY3025_Aptostichus_barackobamai_scaffolds
MY3158_Aptostichus_scaffolds
BME102305_Aptostichus_barackobamai_scaffolds
BME102225_Aptostichus_barackobamai_scaffolds
MY3038_Aptostichus_barackobamai_scaffolds
MY3027_Aptostichus_barackobamai_scaffolds
MY3803_Aptostichus_scaffolds
MY729_Aptostichus_barackobamai_scaffolds
BME102242_Aptostichus_barackobamai_scaffolds
MY3621_Aptostichus_scaffolds
MY3622_Aptostichus_scaffolds
BME102234_Aptostichus_barackobamai_scaffolds
MY3175_Aptostichus_scaffolds
MY1098_Aptostichus_scaffolds
MY3173_Aptostichus_scaffolds
BME102241_Aptostichus_barackobamai_scaffolds
BME102240_Aptostichus_barackobamai_scaffolds
BME102239_Aptostichus_barackobamai_scaffolds
BME102238_Aptostichus_barackobamai_scaffolds
BME102237_Aptostichus_barackobamai_scaffolds
MY3824_Aptostichus_isabella_scaffolds
BME102748_Aptostichus_scaffolds
BME102526_Aptostichus_scaffolds
MY2600_Aptostichus_icenoglei_scaffolds
MY3759_Aptostichus_icenoglei_scaffolds
MY3763_Aptostichus_icenoglei_scaffolds
BME102537_Aptostichus_scaffolds
BME102534_Aptostichus_scaffolds
BME102536_Aptostichus_scaffolds
BME102535_Aptostichus_scaffolds
MY3776_Aptostichus_icenoglei_scaffolds
MY3777_Aptostichus_icenoglei_scaffolds
MY2480_Aptostichus_icenoglei_scaffolds
MY2597_Aptostichus_icenoglei_scaffolds
MY2512_Aptostichus_icenoglei_scaffolds
MY2523_Aptostichus_icenoglei_scaffolds
MY2505_Aptostichus_icenoglei_scaffolds
MY2668_Aptostichus_icenoglei_scaffolds
MY2669_Aptostichus_icenoglei_scaffolds
MY718_Aptostichus_icenoglei_scaffolds
BME102752_Aptostichus_scaffolds
BME102753_Aptostichus_scaffolds
MY2492_Aptostichus_icenoglei_scaffolds
MY2465_Aptostichus_icenoglei_scaffolds
MY2467_Aptostichus_icenoglei_scaffolds
MY719_Aptostichus_icenoglei_scaffolds
BME102837_Aptostichus_scaffolds
BME102828_Aptostichus_scaffolds
BME102832_Aptostichus_scaffolds
BME102830_Aptostichus_scaffolds
BME102831_Aptostichus_scaffolds
BME102829_Aptostichus_scaffolds
BME102833_Aptostichus_scaffolds
BME102851_Aptostichus_scaffolds
MY3635_Aptostichus_icenoglei_scaffolds
MY306_Aptostichus_icenoglei_scaffolds
MY305_Aptostichus_icenoglei_scaffolds
BME102844_Aptostichus_scaffolds
BME102842_Aptostichus_scaffolds
BME102847_Aptostichus_scaffolds
BME102845_Aptostichus_scaffolds
0.4

Figure 11. Raw results of the 90p ASTRAL phylogeny.

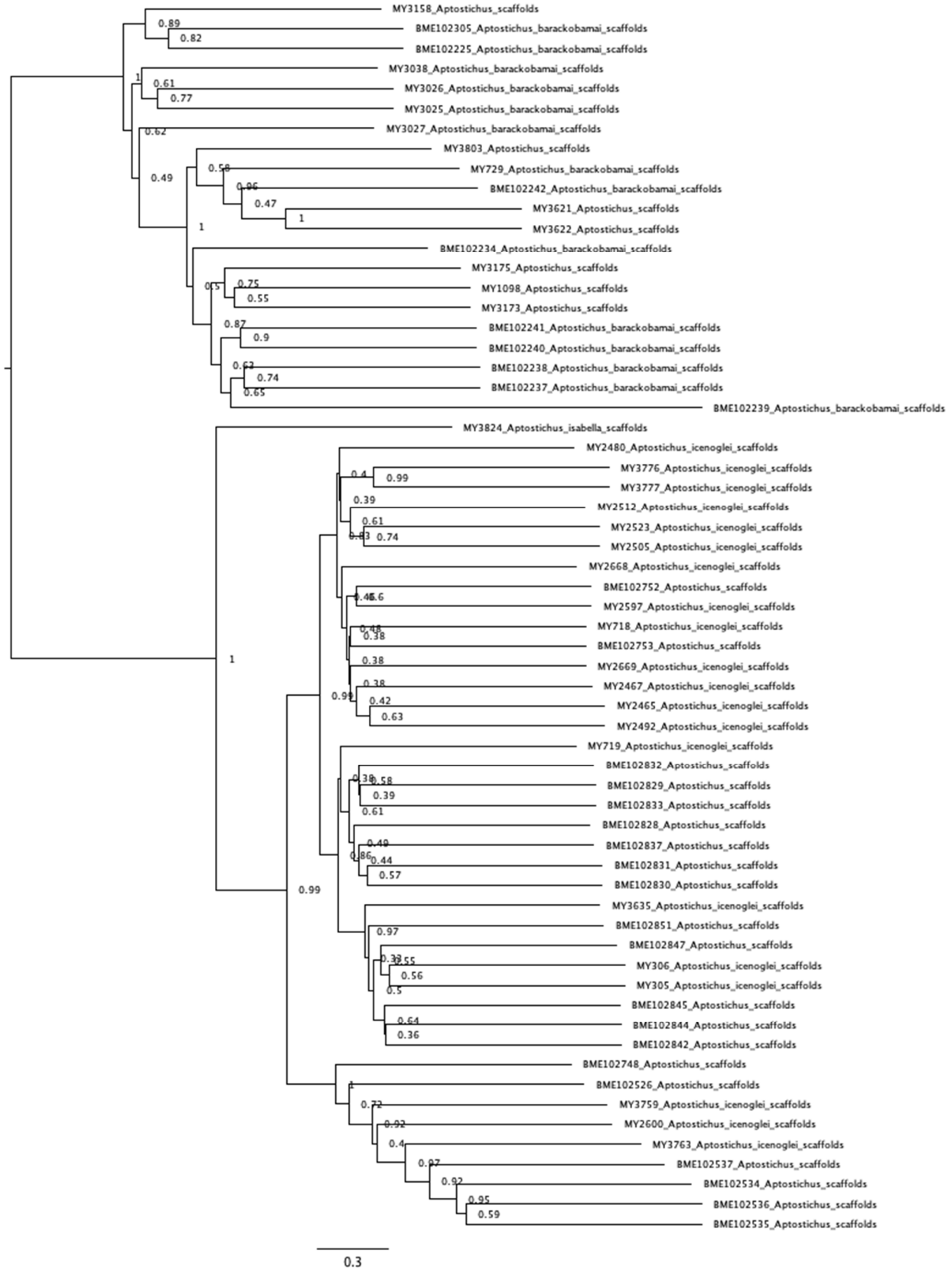


Figure 12. Raw results of the 75p MSC resampling phylogeny.

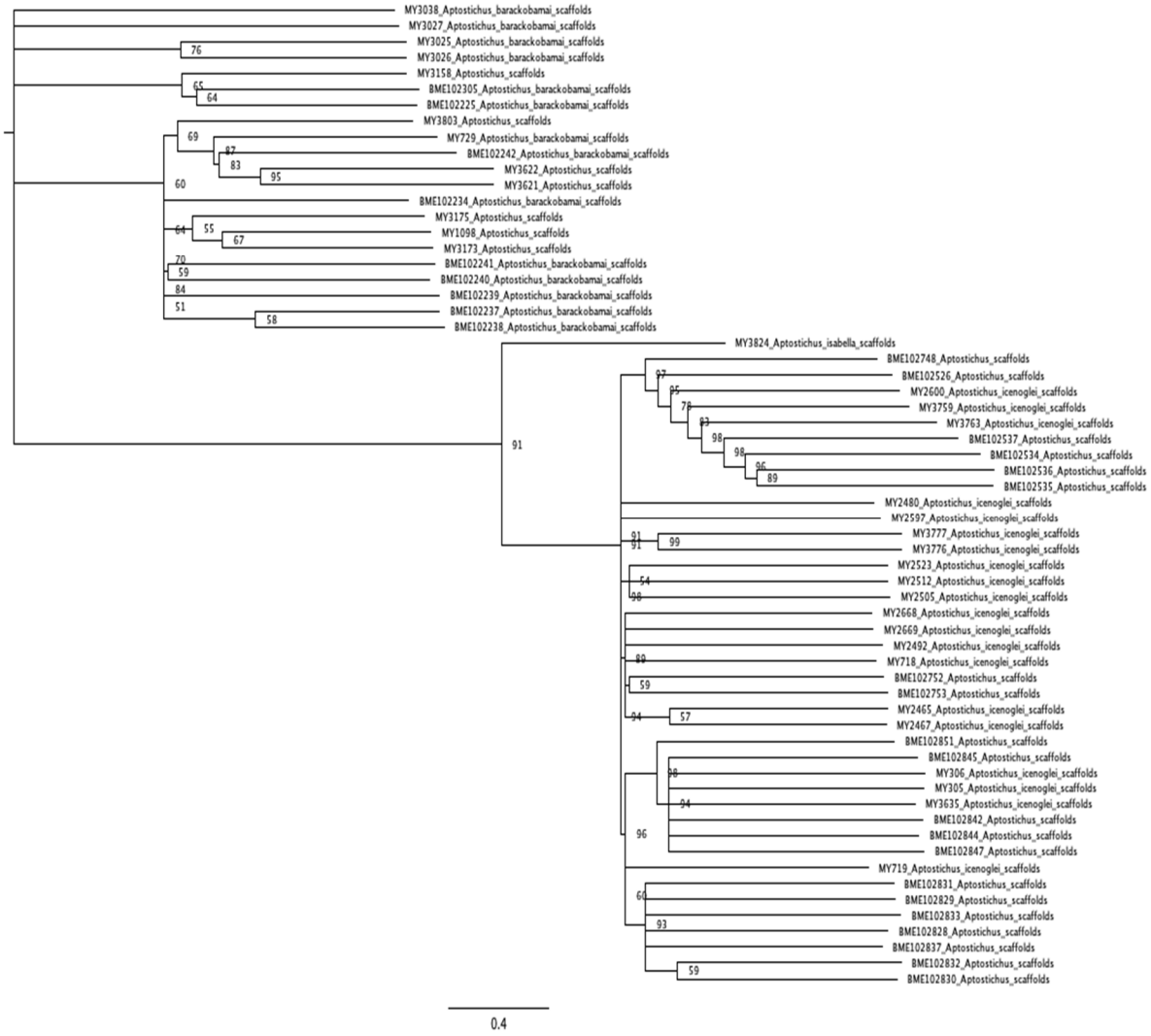


Figure 13. Raw results of the 90p MSC resampling phylogeny.

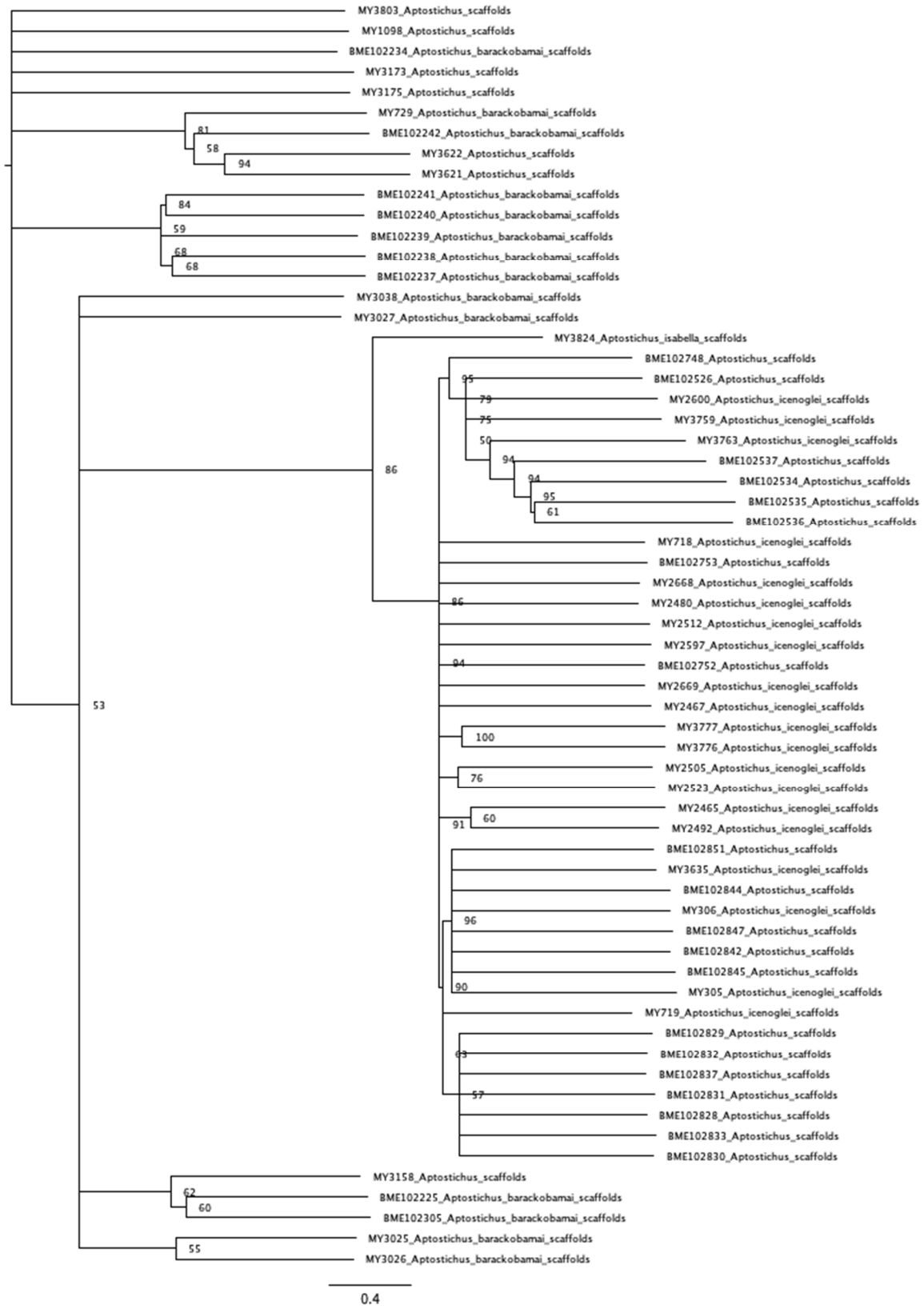

MY3803_Aptostichus_scaffolds
MY1098_Aptostichus_scaffolds
BME102234_Aptostichus_barackobamai_scaffolds
MY3173_Aptostichus_scaffolds
MY3175_Aptostichus_scaffolds
MY729_Aptostichus_barackobamai_scaffolds
81
BME102242_Aptostichus_barackobamai_scaffolds
58
MY3622_Aptostichus_scaffolds
94
MY3621_Aptostichus_scaffolds
BME102241_Aptostichus_barackobamai_scaffolds
84
BME102240_Aptostichus_barackobamai_scaffolds
59
BME102239_Aptostichus_barackobamai_scaffolds
68
BME102238_Aptostichus_barackobamai_scaffolds
68
BME102237_Aptostichus_barackobamai_scaffolds
MY3038_Aptostichus_barackobamai_scaffolds
MY3027_Aptostichus_barackobamai_scaffolds
MY3824_Aptostichus_isabella_scaffolds
BME102748_Aptostichus_scaffolds
95
BME102526_Aptostichus_scaffolds
79
MY2600_Aptostichus_icenoglei_scaffolds
75
MY3759_Aptostichus_icenoglei_scaffolds
50
MY3763_Aptostichus_icenoglei_scaffolds
94
BME102537_Aptostichus_scaffolds
94
BME102534_Aptostichus_scaffolds
95
BME102535_Aptostichus_scaffolds
61
BME102536_Aptostichus_scaffolds
86
MY718_Aptostichus_icenoglei_scaffolds
BME102753_Aptostichus_scaffolds
MY2668_Aptostichus_icenoglei_scaffolds
86
MY2480_Aptostichus_icenoglei_scaffolds
MY2512_Aptostichus_icenoglei_scaffolds
MY2597_Aptostichus_icenoglei_scaffolds
94
BME102752_Aptostichus_scaffolds
MY2669_Aptostichus_icenoglei_scaffolds
53
MY2467_Aptostichus_icenoglei_scaffolds
MY3777_Aptostichus_icenoglei_scaffolds
100
MY3776_Aptostichus_icenoglei_scaffolds
MY2505_Aptostichus_icenoglei_scaffolds
76
MY2523_Aptostichus_icenoglei_scaffolds
MY2465_Aptostichus_icenoglei_scaffolds
91
60
MY2492_Aptostichus_icenoglei_scaffolds
BME102851_Aptostichus_scaffolds
MY3635_Aptostichus_icenoglei_scaffolds
BME102844_Aptostichus_scaffolds
MY306_Aptostichus_icenoglei_scaffolds
96
BME102847_Aptostichus_scaffolds
BME102842_Aptostichus_scaffolds
BME102845_Aptostichus_scaffolds
90
MY305_Aptostichus_icenoglei_scaffolds
MY719_Aptostichus_icenoglei_scaffolds
BME102829_Aptostichus_scaffolds
63
BME102832_Aptostichus_scaffolds
BME102837_Aptostichus_scaffolds
57
BME102831_Aptostichus_scaffolds
BME102828_Aptostichus_scaffolds
BME102833_Aptostichus_scaffolds
BME102830_Aptostichus_scaffolds
MY3158_Aptostichus_scaffolds
62
BME102225_Aptostichus_barackobamai_scaffolds
60
BME102305_Aptostichus_barackobamai_scaffolds
MY3025_Aptostichus_barackobamai_scaffolds
55
MY3026_Aptostichus_barackobamai_scaffolds
0.4

Figure 14. Ancestral area distributions estimation was inferred with DIVALIKE+j analysis implemented in RASP. Terminals are color coded according to the region of their sampling location depicted in the map (bottom left). Inferred ancestral distributions coloration corresponds to the assigned geographic regions or combination of regions (i.e., AD, BC, BD, and CD) as depicted in the legend (top left). Biogeographic events are marked on the nodes as follows: Di = dispersal; V = vicariance.

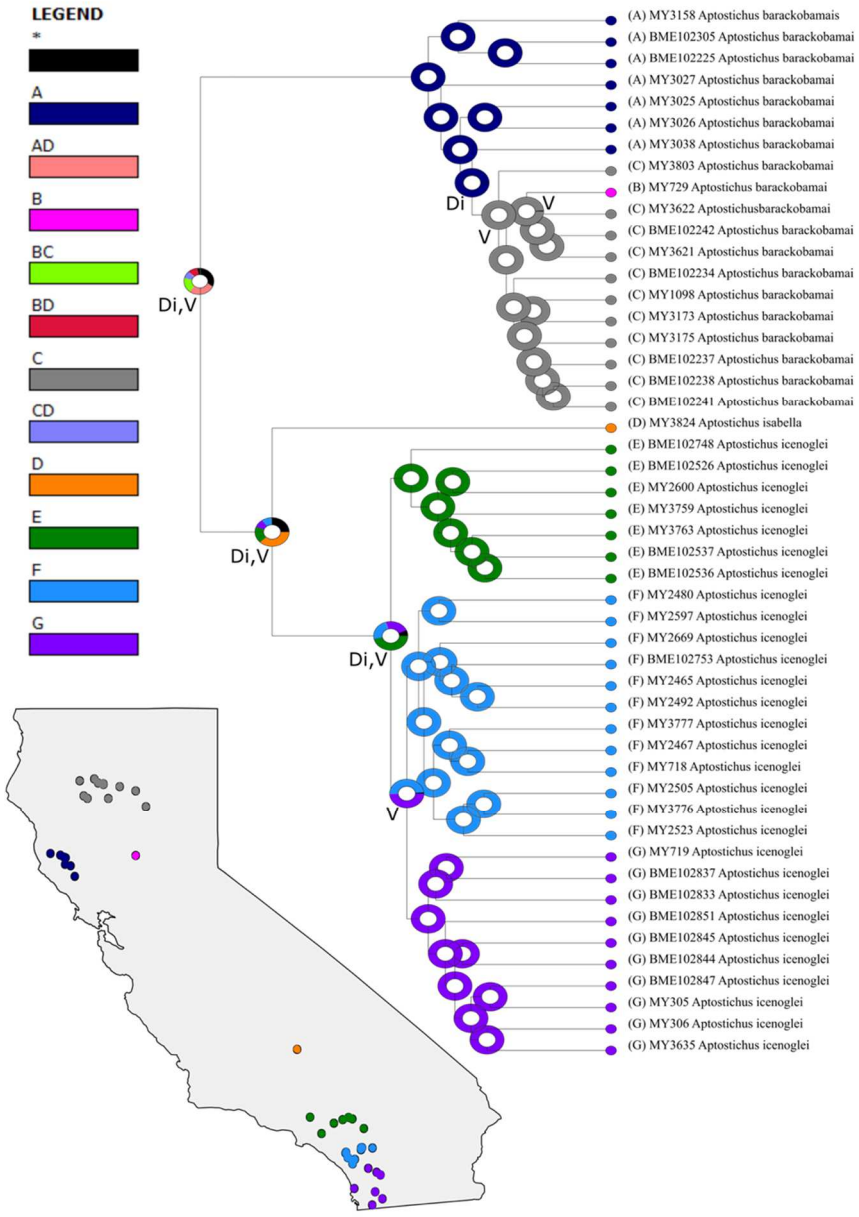


Figure 15. Clustering analysis plots of *A. icenoglei* lineages for both molecular and morphological data sets. Same color scheme for each lineage as previous figures. (A) VAE plot constructed from the 75p SNP data set. (B) VAE plot constructed from the 90p SNP data set. (C) PCA plot, with PC1 and PC2, constructed from morphological measurements. (D) PCA plot, with PC2 and PC3, constructed from morphological measurements.

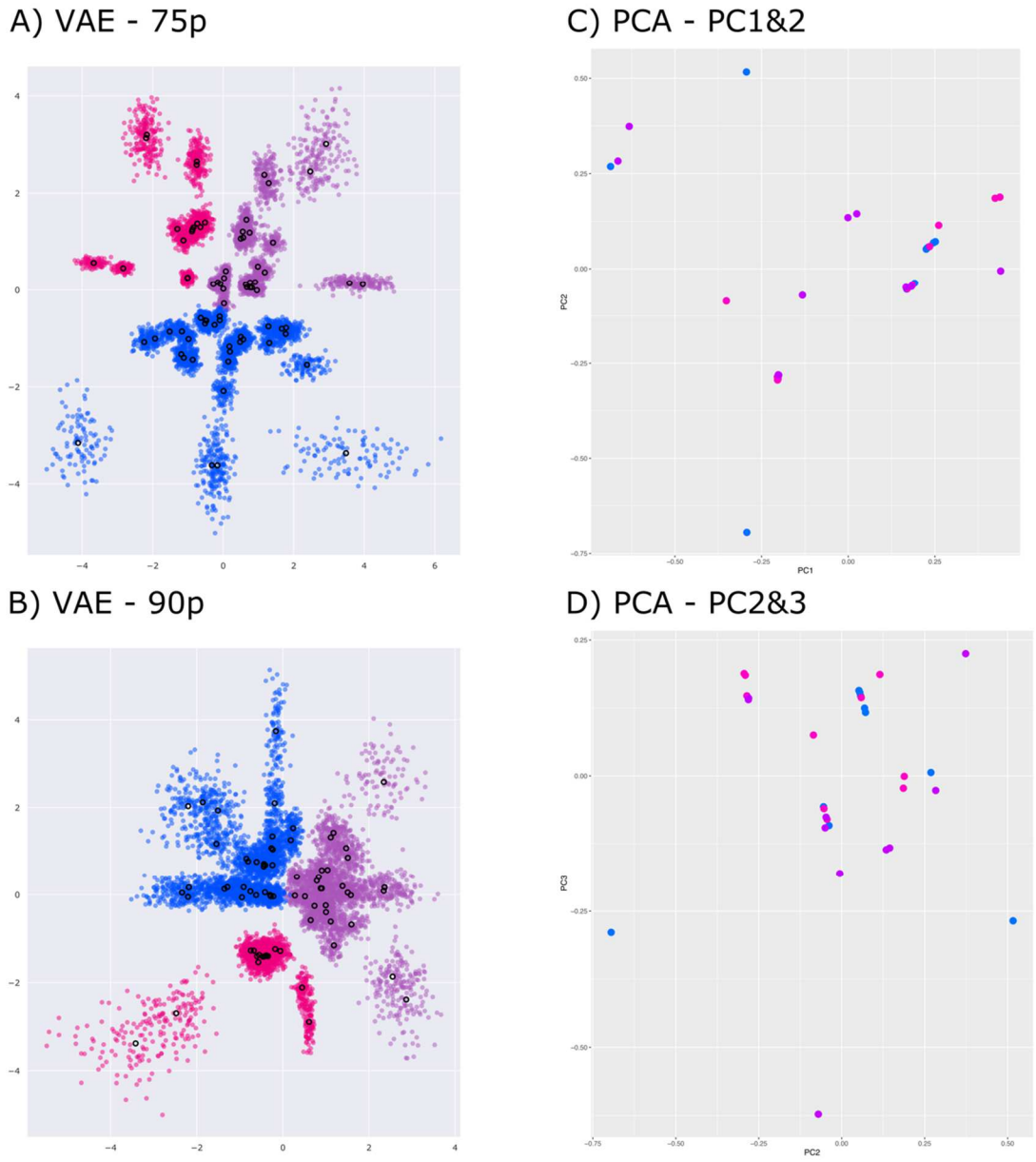


Figure 16. Species distribution models for each *Aptostichus icenoglei* lineage. A) SDM for *A. icenoglei* North lineage. B) SDM for *A. icenoglei* Central lineage. C) SDM for *A. icenoglei* South lineage.

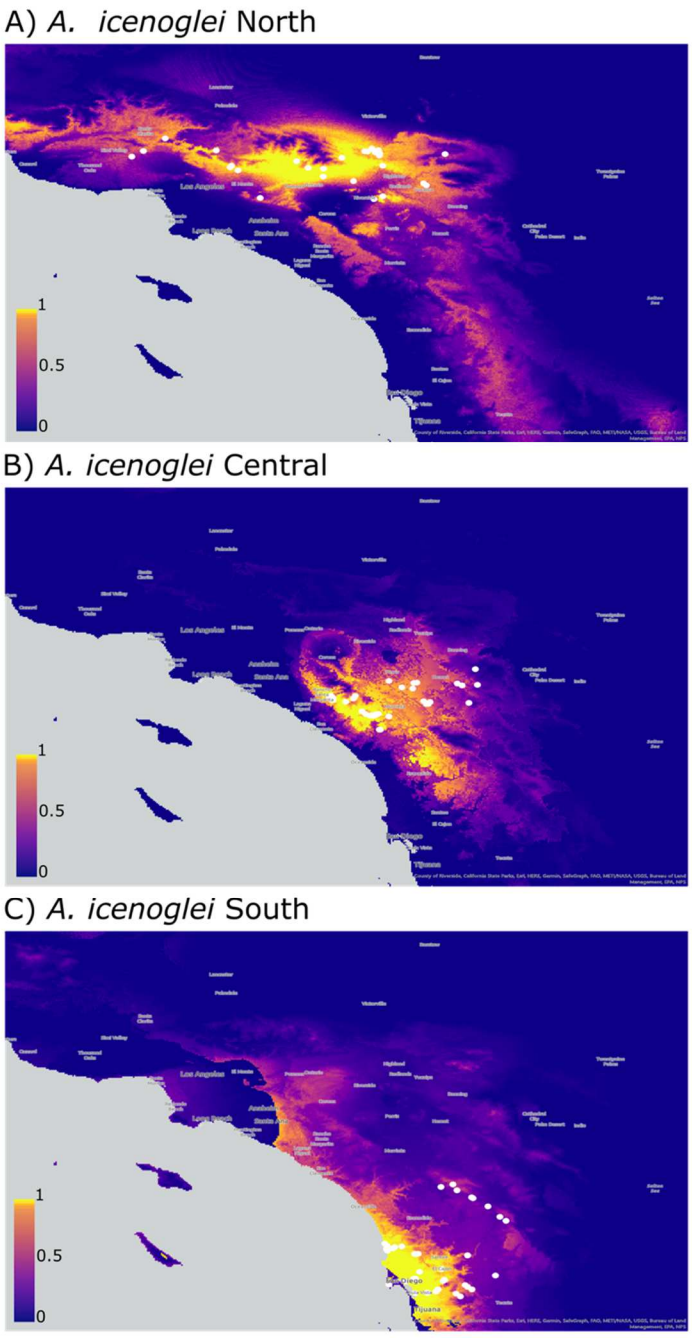


Figure 17. Niche equivalency test results. Top = Central versus South. Bottom = South+Central versus North.

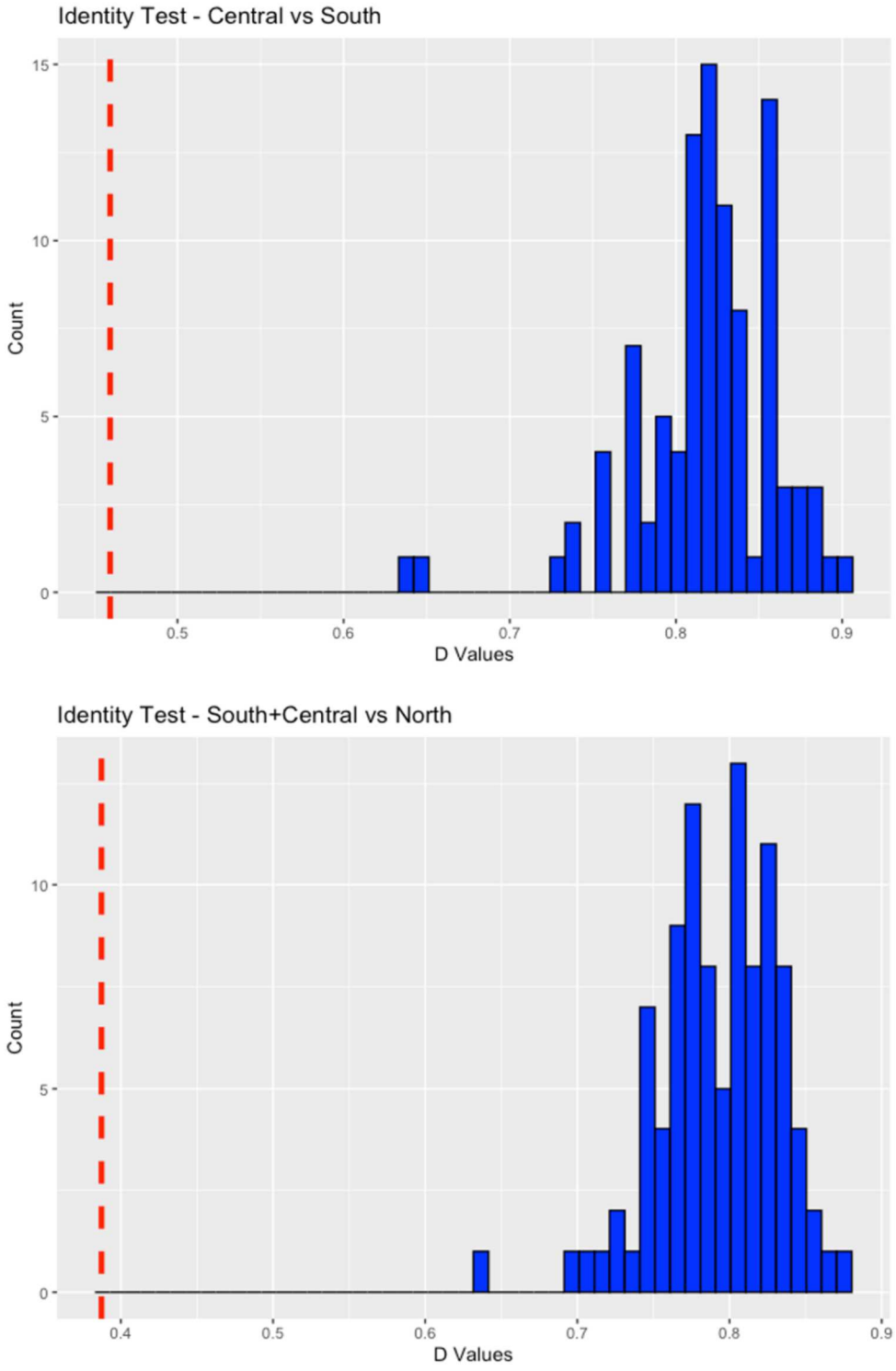


Figure 18. Niche similarity test results for North versus Central+South. A) background region of minimum bounding polygon. B) background region with raster polygons where only grid cells

with habitat suitability scores > 0.75 were retained. C) background region with raster polygons where only grid cells with habitat suitability scores > 0.5 were retained.

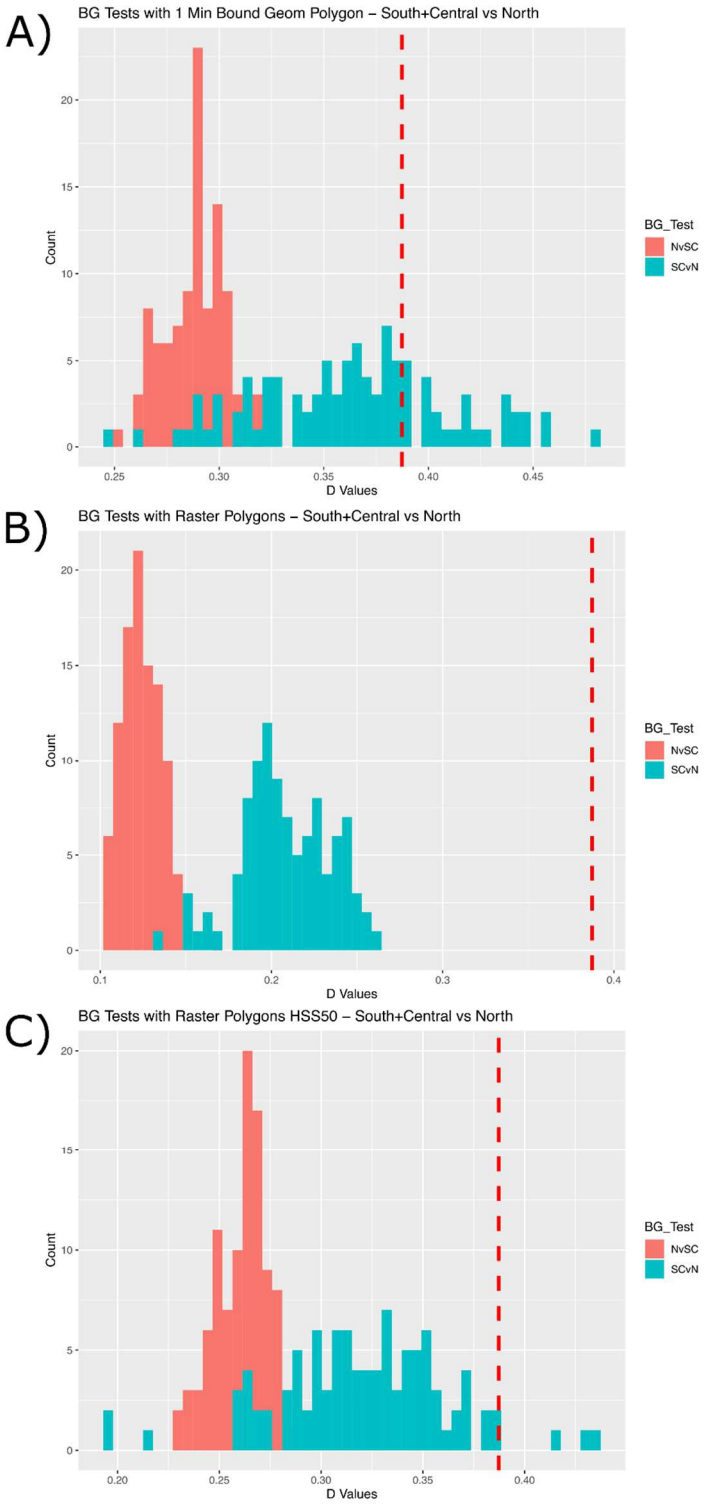


Figure 19. Niche similarity test results for Central versus South. A) background region of minimum bounding polygon. B) background region with raster polygons where only grid cells with habitat suitability scores > 0.75 were retained. C) background region with raster polygons where only grid cells with habitat suitability scores > 0.5 were retained.

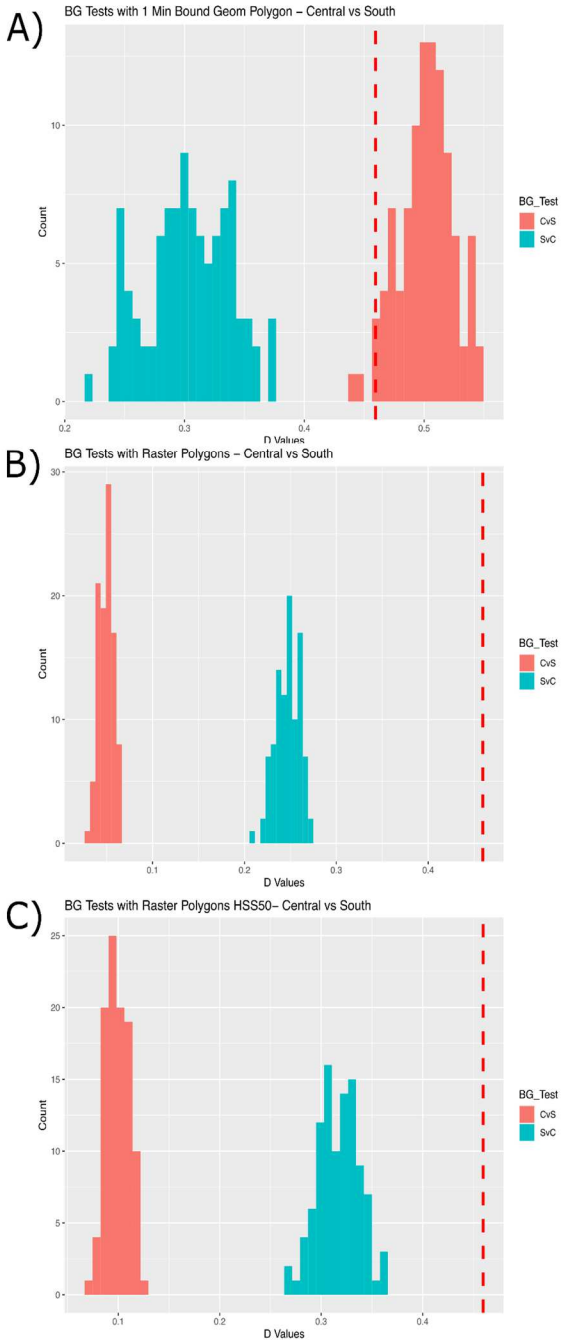
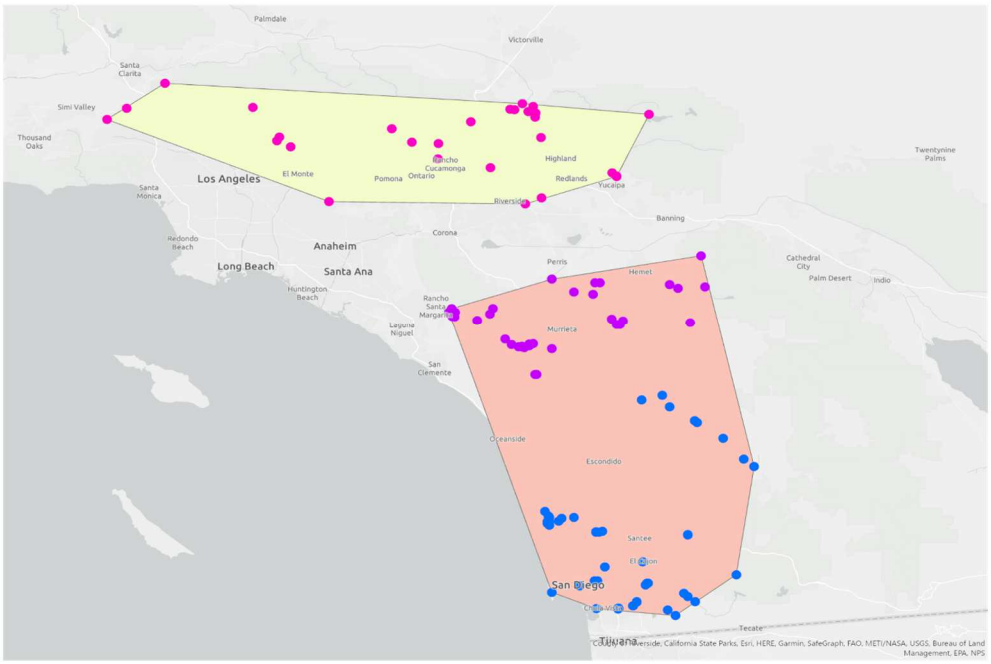


Figure 20. Minimum Bound Geometry polygon background regions for niche similarity tests. A) polygons for the North lineage and Central+South lineage. B) polygons for the Central lineage and South lineage.

A)



B)

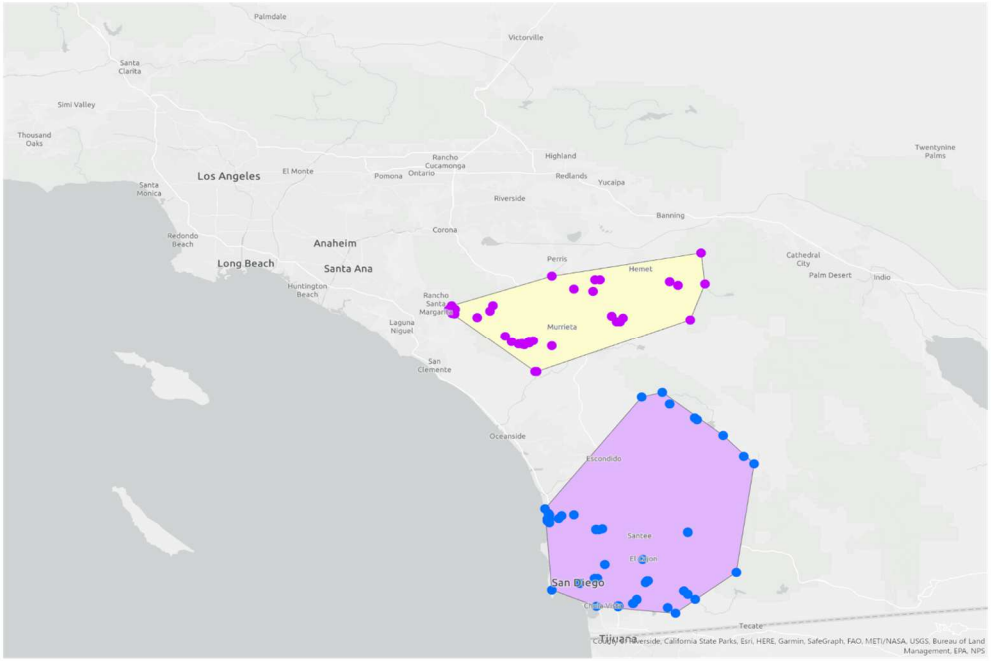


Figure 21. Raster Polygons with HSS > 0.5 and > 0.75 background regions for North and Central+South niche similarity tests. A) North lineage polygons with HSS > 0.75. B) North lineage polygons with HSS > 0.5. C) Central+South lineage polygons with HSS > 0.75. D) Central+South lineage polygons with HSS > 0.5.

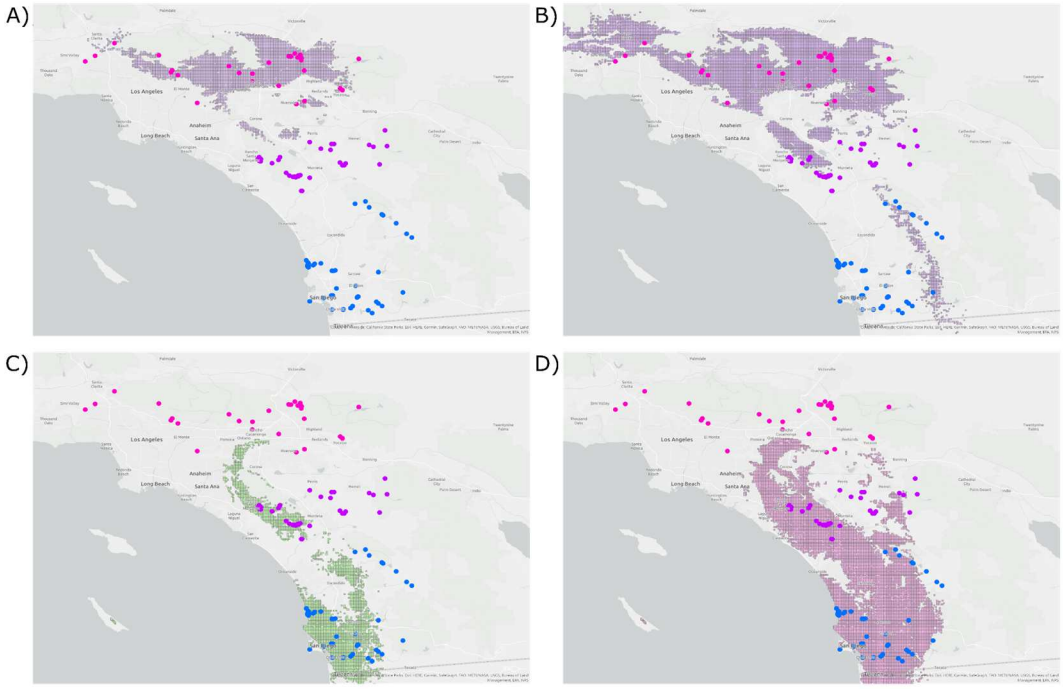


Figure 22. Raster Polygons with HSS > 0.5 and > 0.75 background regions for Central and South niche similarity tests. A) Central lineage polygons with HSS > 0.75. B) Central lineage polygons with HSS > 0.5. C) South lineage polygons with HSS > 0.75. D) South lineage polygons with HSS > 0.5.

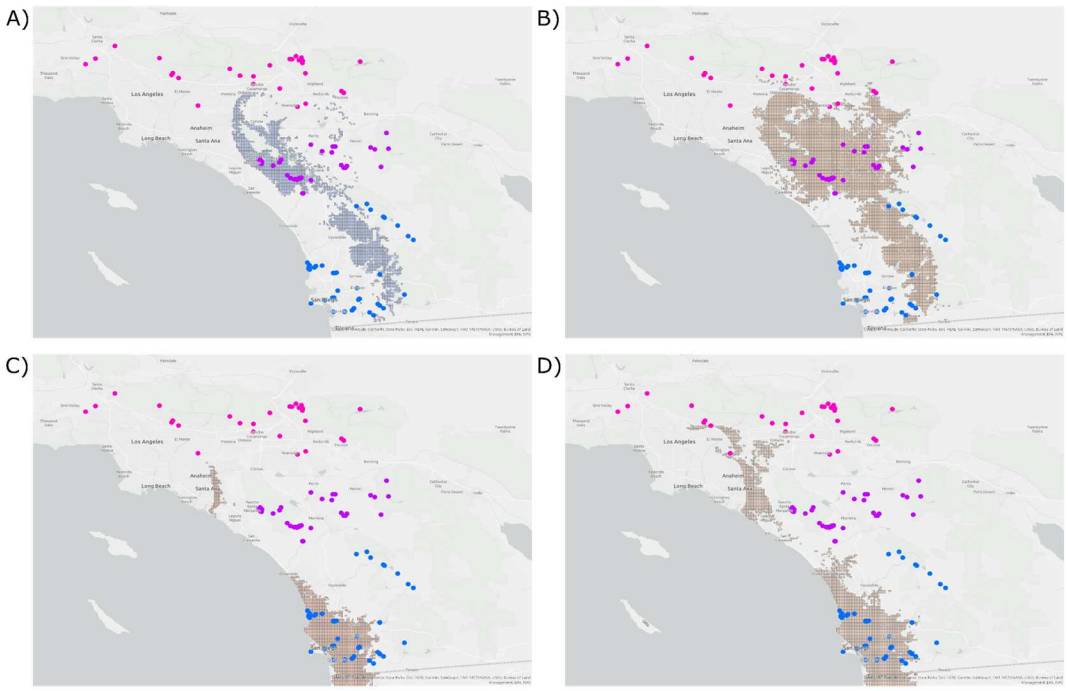
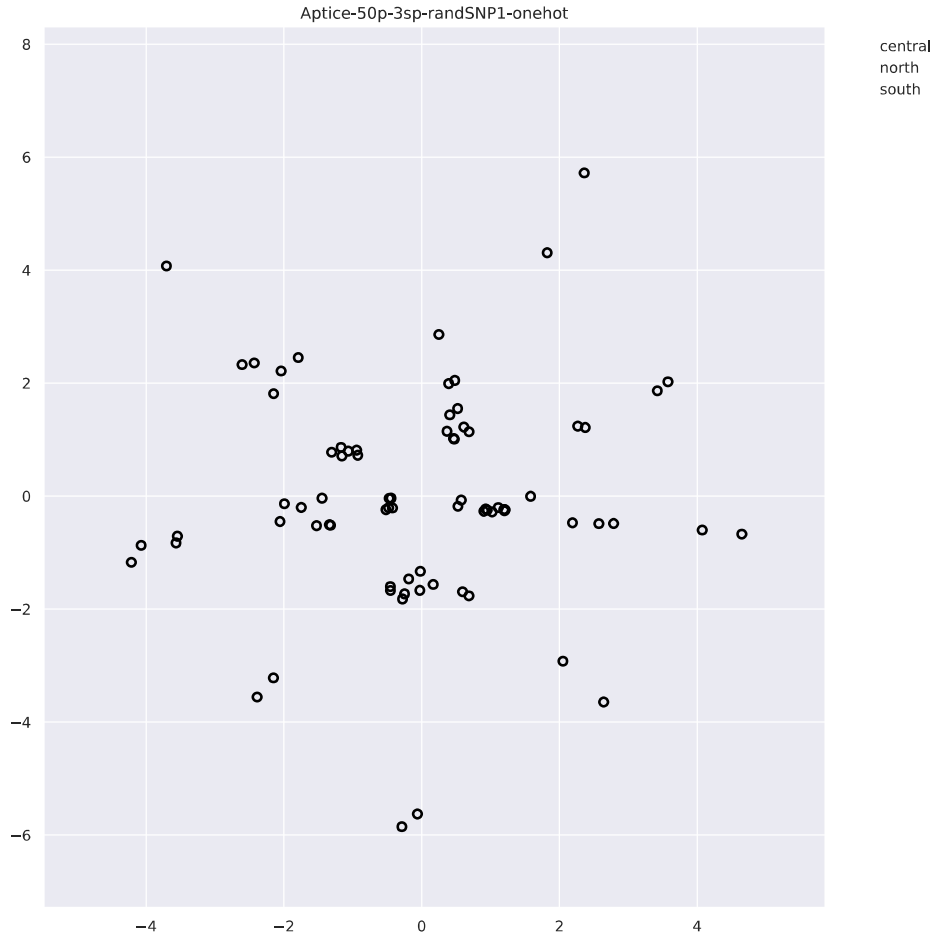


Figure 23. 50p VAE analysis.



CHAPTER III

Combined-evidence phylogeny, habitat evolution, and biogeography of the trapdoor spider genus *Aptostichus* (Araneae: Mygalomorphae: Euctenizidae)

Introduction:

Spiders placed in the infraorder Mygalomorphae (tarantulas, trapdoor spiders and their kin) are generally recognized as an ancient cosmopolitan lineage that has persisted for over 250 million years (Opatova et al., 2019). In addition to their fossorial lifestyle, they are relatively unique as a consequence of their extended life spans, which are known to exceed 20 years or more (Mason et al., 2018), extremely limited vagility leading to localized population divergence and genetic structuring (Bond et al., 2001; Candia-Ramírez & Francke, 2020; Cooper et al., 2011; Hamilton et al., 2011; Hedin et al., 2015; Hedin & Carlson, 2011; Rix et al., 2020), and morphological stasis (Bond et al., 2001; Bond & Stockman, 2008; Godwin et al., 2018; Hendrixson et al., 2015; Leavitt et al., 2015; Newton et al., 2020; Starrett & Hedin, 2007). These features resulting from limited dispersal capability make them attractive candidates for studying speciation pattern and process, biogeography, and adaptation and character evolution, yet also present a set of perplexing challenges. Evolutionary studies of mygalomorphs at both shallow and deeper phylogenetic levels have been limited prior to the advent of next generation sequencing approaches, with the majority of such studies relying on morphological characters or limited targeted locus approaches for phylogenetic reconstruction. A number of recent studies (e.g., Bond et al., 2020; Hedin et al., 2019; Montes de Oca et al., 2022; Opatova et al., 2019) that have employed genomic data show significant departure in relationships when compared to these older studies, both morphological (e.g., Bond & Opell, 2002; Goloboff, 1993; Raven, 1985) and targeted locus (Bond et al., 2012; Hedin & Bond, 2006), consequently requiring considerable

reorganization of classification schemes that span the infraorder, families, and relationships within genera. It has become clear that larger genomic-scale datasets are required to confidently reconstruct relationships among these enigmatic taxa.

One family that greatly owes its taxonomic rank to the implementation of genomic data is Euctenizidae. Raven (1985) previously recognized euctenizids as a subfamily within Cyrtaucheniidae, Euctenizinae. Raven's (1985) analysis was based on morphology, yet a subsequent morphological study by Bond & Opell (2002) suggested that Cyrtaucheniidae was probably polyphyletic with euctenizines forming an independent monophyletic group, but their status with respect to cyrtaucheniids remained unresolved. A subsequent study by Bond et al. (2012) finally resulted in the elevation of euctenizines to family level (based on total evidence, targeted molecular and morphological data) in light of a number of previous molecular studies that had already inferred monophyly (Bond & Hedin, 2006; Hedin & Bond, 2006). Recent subgenomic analyses using anchored hybrid enrichment have confirmed the family's monophyly and also established well-supported relationships among the genera along with its placement as a sister group to the largely southern hemisphere family Idiopidae (Bond et al., 2020; Opatova et al., 2019). Euctenizidae currently comprises eight genera and is subdivided into two subfamilies: Apomastinae and Euctenizinae (Bond et al., 2020). Euctenizinae is currently composed of five genera, which include *Eucteniza* Ausserer 1875, *Promyrmekiaphila* Simon 1891, *Neoapachella* Bond & Opell 2002, *Cryptocteniza* Bond et al. 2020, and *Entychides* Bond & Opell 2002. Euctenizines are distributed across the American Southwest in Texas, New Mexico, Arizona, and California with a number of genera whose distributions extend into Mexico (Bond et al., 2020). Apomastines comprise two genera, *Apomastus* Bond & Opell 2002 and *Aptostichus* Simon 1891, both of which are found in California, with some *Aptostichus* species in Nevada, Arizona, and

Mexico (Bond, 2012; Bond et al., 2020). Over the last 10-15 years these genera have undergone major taxonomic revision, to include *Myrmekiaphila* Atkinson 1886 (a southeastern United States endemic; Bond & Platnick, 2007), *Eucteniza* (Bond & Godwin, 2013), *Promyrmekiaphila* (Stockman & Bond, 2008), *Apomastus* (Bond, 2004), and *Aptostichus* (Bond, 2012). Species-level diversity within Euctenizidae (73 nominal species) is relatively high compared to other mygalomorph families found in North America (with the exception of *Ummidia*; Godwin & Bond, 2021). The number of species within each genus varies widely (Bond et al., 2020), with two monotypic genera (*Cryptocteniza* and *Neoapachella*), two genera with twelve and fourteen species (*Myrmekiaphila* and *Eucteniza*, respectively), and the genus *Aptostichus* with forty species (one species in Mexico, likely misplaced) being the most diverse euctenizine genus. Bond et al. (2020) identified a significant diversification rate shift across the lineage suggesting a large increase in speciation events, thus underscoring a hypothesized pattern of an adaptive radiation in *Aptostichus* across the CAFP.

Given its apparent higher rate of diversification, distribution that largely spans the California Floristic Province, and numerous disparate ecoregions that include chaparral, desert (both Mojave and Colorado), alpine, and coastal dunes, *Aptostichus* presents an interesting system for studying evolutionary diversification in trapdoor spiders. The primary objective of this study is to reconstruct the relationships among all *Aptostichus* species thereby providing a framework for a number of questions related to the evolution of spider adaptation in varying environments. We build upon a previous study using morphological data (Bond, 2012) by adding genomic-scale data (i.e., ultraconserved elements; UCEs) for the majority of species and reconstructing evolutionary relationships using a combined evidence approach. This phylogeny will then allow us to test hypotheses about evolution of arid environments.

Genus Aptostichus

Aptostichus was originally described by Simon (1891), with *A. atomarius* as the type species. Over the next century only three other species were described in the genus: *A. stanfordianus* Smith 1908, *A. simus* Chamberlin 1917, and *A. hesperus* Chamberlin 1919. However, extensive efforts have been made over the last twenty years to describe the large amount of diversity, both morphological and genetic, present in *Aptostichus*. A large taxonomic revision of the genus by Bond (2012) led to the description of 33 new species based primarily on morphological data (e.g., male secondary sexual characters). Subsequently, Valdez-Mondragón & Cortez-Roldán (2016) described a new *Aptostichus* species from Oaxaca, Mexico based on morphology; however, this species is clearly misplaced and should be attributed to a different family/genus (in prep). In addition to morphological diversity, preliminary mitochondrial data (Bond, 2012) for well-sampled species hint at the possibility of cryptic diversity for several relatively widespread species. This phenomenon has long been recognized in *Aptostichus* prior to the 2012 revision, first in *A. simus* (Bond et al., 2001) and then within the *A. atomarius* species complex where Bond & Stockman (2008) described three new cohesion species based on molecular and ecological data: *A. stephencolberti* Bond & Stockman 2008, *A. angelinajolieae* Bond & Stockman 2008, and *A. miwok* Bond & Stockman 2008. These studies indicate that *Aptostichus* species numbers may be much higher than what can be gleaned from morphology alone, and additional studies are needed to resolve species limits (e.g., Newton et al., in prep).

The majority of *Aptostichus* species are known from a well-known biodiversity hotspot, the California Floristic Province (Figures 24-26), yet three additional taxa are known from Nevada (*A. pennjillettei*; Figure 24) and Arizona (*A. chiricahua* and *A. edwardabbeyi*; Figures 25 & 26, respectively). Across species there are a mixture of distributional patterns that range from

narrowly distributed (known from a single locality) to more widespread. For example, *A. icenoglei* is distributed throughout the Transverse and Peninsular Ranges in southern California, whereas the comparatively rare species, *A. anzaborrego*, is only known from the type locality in Anza-Borrego Desert State Park. Additionally, these distributions span disparate ecoregions across the CAFP. One can find *Aptostichus* species in coastal dunes, arid environments such as deserts (both Mojave and Colorado deserts), dry steppe, and chaparral, as well as mesic environments (i.e., mixed forest, mixed redwood, and Sierran alpine habitat; Figure 27). These spiders inhabit some of the most arid environments on the planet in the Mojave Desert including Death Valley (e.g., *A. elisabethae*, *A. sarlacc*, *A. derhamguilianii*, and *A. fornax*). In all of these habitats, like most euctenizids, *Aptostichus* species construct subterranean burrows that are covered by a thin, wafer trapdoor made of silk and the surrounding substrate, often making their doors quite cryptic, where they can safely forage for prey at the burrow entrance (Figure 27). Their burrows are lined with heavy and distinctive white silk, with burrow depths varying from relatively shallow to quite deep, and, like a number of other euctenizids (e.g., *Promyrmekiaphila*), store prey remains inside their burrows.

The first phylogeny evaluating interspecific relationships within *Aptostichus* was constructed in Bond (2012) using 72 morphological characters, both qualitative and quantitative, and established four monophyletic species groups: Atomarius, Hesperus, Simus, and Sierra. Nevertheless, this phylogeny was considered preliminary due to putative homoplasy and overreliance of genitalic features and characters only scored for one sex. Specifically, some characters are associated with habitat (e.g., psammophilic conditions influencing lighter coloration) and approximately one-third of the taxa are only known from one sex, thus some character systems were not assessed for all taxa. As seen in Bond (2012), relationships within the

Atomarius and Hesperus species groups are uncertain or unresolved. A follow up genomic study that included enough taxa to evaluate some interspecific relationships revealed that the Hesperus species group is nested within the Atomarius species group, with the exception of *A. hedinorum*, which is sister to the Atomarius+Hesperus assemblage (Bond et al., 2020). However, this study had limited taxon sampling (i.e., 15 species represented out of 40), thus many interspecific relationships within *Aptostichus* still remain unresolved.

Our Approach

We aim to build upon previous work by acquiring subgenomic UCE data from at least one representative of each species for a complete molecular dataset to use in conjunction with the previous morphological dataset from Bond (2012). The rarity and/or presumed extinction of a number of *Aptostichus* species, which makes collecting them nearly impossible, have previously hindered attempts to construct a comprehensive phylogeny using molecular data. Recent advancements in historical museum DNA extraction protocols (e.g., Tin et al., 2014) have made it possible for researchers to extract viable DNA from these traditionally morphological vouchers. Thus, when fresh material could not be sampled, we attempted to extract DNA and recover UCEs from ethanol-preserved museum specimens ranging from 11-63 years old (see Table 13 for details). Although some of these extractions did not yield usable UCEs (*A. pennjillettei*, *A. nateevansi*, *A. chiricahua*, *A. killerdana*, *A. chemehuevi*, *A. sarlacc*, *A. derhamguilianii*, *A. anzaborrego*, *A. satleri*, *A. lucerne*, *A. bonoi*, and *A. dorothealangeae*), we were successful in adding ten additional species to our UCE data matrix (*A. elisabethae*, *A. fornax*, *A. fisheri*, *A. cajalco*, *A. sierra*, *A. huntington*, *A. chavezi*, *A. serrano*, *A. isabella*, and *A. muiri*). Given the incomplete taxon sampling in our UCE dataset we employ a total evidence approach that combines molecular and morphological data for phylogenetic inference.

Methods:

Taxon Sampling

We sampled four outgroup species from related euctenizid genera (*Myrmekiaphila*, *Promyrmekiaphila*, *Entychides*, and *Apomastus*) and 57 *Aptostichus* taxa (including all but one misplaced species for a total of 40; Table 13); we did not sample the Mexico species *A. sabinae*. Specimens included herein were either collected previously (Bond, 2012; Bond & Stockman, 2008) or recently, using standard field techniques. Morphological character scorings are from Bond (2012), which included 40 *Aptostichus* species plus euctenizid outgroups scored for 72 qualitative characters, a subset of which were originally quantitative measurements (21 total) that were discretized.

Sequence Capture

Ultraconserved element (UCE) data were generated following the methods described in Faircloth et al. (2012) with subsequent modifications in Starrett et al. (2017), Hedin et al. (2019), and Kulkarni et al., (2020). If possible, we used tissue samples stored in -80°C (i.e., 'fresh' material). In the cases where we had no freshly collected tissue, we attempted to extract DNA from ethanol-preserved museum specimens (typically stored in 70-80% ethanol at room temperature). We extracted genomic DNA from leg tissue for 'fresh' material individuals using the Blood and Tissue DNeasy kit (Qiagen) following the manufacturer's protocol. To extract DNA from museum voucher specimens we followed the 'MMYT protocol' from Tin et al. (2014) with modifications in Derkarabetian et al. (2019). To mitigate the potential for contamination, we implemented sterilization techniques with MMYT extractions: frequently

sterilizing all laboratory surfaces (i.e., lab bench and fume hood where extractions took place) and tools used, using filtered tips for pipettes, and wearing appropriate PPE. DNA quantification was performed using Qubit 3.0 Fluorometer (Life Technologies), with quality check assessed with an agarose gel.

For freshly collected specimens, 250 ng of DNA was sonicated into fragments ranging from 200-1000 bp using an ultrasonicator (Covaris E220), whereas museum DNA was not sonicated due to its high amount of fragmentation as seen with the agarose gel (i.e., highly degraded). UCE libraries were generated with the KAPA Hyperprep kit (Roche) using universal adapters and iTru5/7 barcodes (Glenn et al., 2019; BadDNA@UGA) with minor modifications for museum material extractions. Specifically, all bead cleanup ratios were increased to capture smaller DNA fragments and adapter ligation time was increased to 60 minutes (see Derkarabetian et al., 2019 for more details). Libraries were hybridized at 60°C for 24 hours to the Spider probeset (Kulkarni et al., 2020) following the version 4 chemistry protocol (Arbor Biosciences). Hybridization enriched library pools were sequenced with 150 bp paired-end reads on the HiSeq4K at the University of California Davis DNA Technologies Core. Some individuals were sent to Rapid Genomics (Florida) for library prep and sequencing (Table 13).

Sequence processing and analyses were performed on the Farm Community High Performance Computing Cluster at the University of California, Davis. Reads were first filtered using Illumiprocessor (Faircloth, 2013) and then trimmed using Trimmomatic (Bolger et al., 2014) in the Phyluce 1.7.1 pipeline (Faircloth, 2015). Cleaned paired-end and single-end reads were assembled de novo with SPAdes v. 3.14.1 with the isolate option (Prjibelski et al., 2020). Because many of the museum individuals have fewer reads and thus fewer UCE loci compared to the fresh material individuals, we attempted to bolster some of these museum specimens by

combining contigs from conspecific individuals. We imported contig files into Geneious v10.0.5 where we re-assembled them using default parameters. Then we retained all 'unique' contigs (i.e., reads not assembled) as well as overlapping contigs with a 90% threshold consensus to generate a combined contig file for each species (i.e., AP321 and AP333 combined for *A. elisabethae* and AP418 and AP578 combined for *A. fornax*; Table 13). Scaffolds were matched with 65% identity and 65% coverage to the modified probe list from Maddison et al., (2020), which is a combination of the Arachnid (Faircloth, 2017; Starrett et al., 2017) and Spider (Kulkarni et al., 2020) probesets. MAFFT was used to align individual locus datasets, and alignments with locus occupancy minimums of 50% and 75% were assembled. Alignment masking was performed with TrimAl v.1.2 (Capella-Gutierrez et al., 2009) using default settings.

Phylogenetic Analyses

Our phylogenetic reconstructions were generated with maximum likelihood (ML) inference. For partitioning schemes, we designated the morphological data as one partition and each UCE locus as a partition. Our ML analyses were executed in IQTREE v2.1.2 (Minh et al., 2020), with the MK model selected for the morphological data partition and model selection performed for each UCE locus using ModelFinder (Kalyaanamoorthy et al., 2017) with branch lengths linked between partitions. Support values were inferred from 1000 replicates of ultrafast bootstrapping (i.e., UFBoot; Hoang et al., 2018). The flag -bnni, which is recommended when severe model violations are present, was necessary for some of the larger constrained analyses as an additional step to further improve UFBoot trees. We explored 16 combinations with our datasets for ML phylogenetic inference (see Table 14 for analysis names): 1) morphological data only (MA), 2) only UCE data including fresh and museum individuals (UA_50p and UA_75p), 3) only UCE data excluding museum individuals (UN_50p and UN_75p), 4) combined

morphological and UCE data including fresh and museum individuals (UA_MA_50p and UA_MA_75p), 5) combined morphological and UCE data excluding museum individuals (UN_MA_75p), 6) combined UCE data including fresh and museum individuals and only morphological data with corresponding UCE data (UA_ME_50p and UA_ME_75p) and 7) combined UCE data including only fresh samples and only morphological data with corresponding UCE data (UN_ME_75p). Museum harvested samples had high amounts of missing data, with some varying in topological placement depending on different datasets. Thus, topological constraints, which were based primarily on previously identified species groups (with the exceptions of *A. hedinorum* and *A. cabrillo*; see Bond et al., 2020 and unconstrained analyses), were implemented to mitigate discordance among phylogenies. Because analyses with large proportions of missing data produce spuriously long branch lengths, we computed the average branch lengths for tips and internal nodes in adjacent clades to reassign the values for each of these long branches (see Table 15 for details). This is similar to the "stolen branch length" approach (Darriba et al., 2016), however simplified because we did not scale for partition length, which were relatively homogenous across clades.

Evolution of Habitat Type

To investigate the evolution of habitat type across *Aptostichus*, we employed an ancestral state reconstruction analysis using the R package corHMM (Beaulieu et al., 2013) on an ultrametric scaled tree. Habitat scoring for each species was based on the scoring previously used by Bond (2012): (0) mixed forest and coastal range; (1) chaparral; (2) alpine meadow; (3) desert; (4) coastal dune; (5) mixed redwood; and (6) dry steppe. All nominal *Aptostichus* species previously scored in Bond (2012) were reevaluated to verify habitat scoring accuracy, and four newly sampled *Aptostichus* individuals (i.e., MY4535, MY4536, BME101842, and *A. n sp.*

madera) were scored based on their locality information. If a species inhabits multiple habitat types, then they were scored as polymorphic. Our preferred tree topology was converted to a relative-rate scaled ultrametric tree using the R package ape ('chronopl'; Paradis & Schliep, 2019) with lambda=0.1 and node.states=marginal, thus providing a way to understand evolutionary changes over relative 'time' given the branches are scaled to evolutionary rates as opposed to a dated phylogeny (Bond et al., 2020). As per Bond et al. (2020), character state optimizations utilized models with either all rates being equal (ER) or all rates differ (ARD), with the preferred model chosen by comparing AICc values computed by corHMM.

Results:

UCE Data

Table 16 summarizes the UCE data. All individuals, including both recently collected material ("new collections", NC) and museum collections harvested material (MUS), averaged 1909144.5 cleaned reads, with a mean number of 1001.2 contigs and average contig length of 817 bp. After steps to align, filter, and trim these contigs, we established four final datasets with 50% and 75% minimum locus occupancy percentages for both NC and MUS material (UA) and NC material only (i.e., excluding museum specimens; UN). For UA, the 50p data matrix included a total number of 1333 loci and 898644 bp, with an average of 896.5 loci retained per individual; whereas, the 75p data matrix comprised 461 loci and 337300 bp, with an average of 350.2 loci retained per individual. For UN, the 50p dataset comprised 1337 loci, with 1070.1 loci retained per individual on average, and 1037549 bp. The UN 75p dataset contained 853 loci and 652747 bp, with a mean of 719.9 loci retained for each individual.

Comparison of NC versus MUS UCE data is summarized in Table 17. Overall, DNA extracted from recently collected tissues averaged higher clean read counts, number of contigs, and number of loci for both 50p and 75p UA data matrices when compared to the ethanol-preserved museum specimens. NC material averaged 2,440,254.9 cleaned reads and 1169.3 contigs, whereas MUS material averaged 356668.2 cleaned reads and 509.9 contigs, which is about 7X fewer reads and 2.3X fewer contigs than NC material. MUS material averaged 473.9 and 195.8 loci for the UA 50p and 75p datasets, respectively; in contrast, NC samples retained an average of 1041.1 loci for the 50p data matrix and 403 loci for the 75p data matrix, which is a little over 2X more loci for both datasets compared to MUS tissues.

Phylogenetic Analyses

Our taxon sampling for phylogenetic analyses included all nominal *Aptostichus* species except *A. sabinae* because the specimen described as *A. sabinae* has two tarsal claws, which does not fit the generic diagnosis of *Aptostichus* or the family diagnosis of Euctenizidae. Thus, its taxonomic status will be reassessed (in prep). Table 14 and Figures 28-43 summarize the results from the phylogenetic analyses and their log likelihood values (see analysis names in Table 4 for reference – typically in the format U[A = all taxa or N = no MUS specimens + 50p/75p = locus occupancy percentage] + M[A = all taxa scored for morphology or E = only taxa scored for morphology and UCEs] + Con [with constraints if applicable]). Our preferred phylogeny, the UA75p_MA_Con, had many well supported nodes (bootstrap values > 70), although some nodes towards the tips had weak support (Figure 29). All tree topologies primarily correspond to the overall grouping of the four previously established monophyletic species groups (Sierra, Simus, Hesperus, and Atomarius) by Bond (2012); however, the Atomarius species group is now paraphyletic with respect to the Hesperus group, all now forming one clade (hereafter referred to

as Hesperus+Atomarius). The Hesperus group, with the exception of *A. hedinorum*, remains largely monophyletic with its species forming a clade within the larger Hesperus+Atomarius clade. The UN phylogenies, both 50p and 75p, recover the species group topology (i.e., Sierra, Simus, and Hesperus+Atomarius; Supplementary Figures 33 & 34) with all nodes having high support (i.e., all bootstrap values = 100). Conversely, adding UCE MUS data and morphological data result in varying topologies within species groups across analyses.

Sierra Clade

Bond (2012) recovered four species within the Sierra species group: *A. sierra*, *A. huntington*, *A. dorothealangeae*, and *A. chavezi*. Recent collections yielded four individuals (i.e., MY4535, MY4536, BME101842, *A. n sp. madera*) that grouped with the other Sierra clade individuals in all analyses, with the exception of UN_MA (see Figure 38) as well as *A. sierra* and *A. huntington* in unconstrained analyses. One clade, MY4536+ BME101842+*A. n sp. madera*, was consistent across all analyses; however, its sister relationship was different depending on the locus occupancy percentage and in combination with morphological data. UN50p/75p and UN_ME recovered the same topology with high support (bootstrap values = 100; Figures 34, 35, and 43). Adding MUS UCE data for *A. sierra*, *A. huntington*, and *A. chavezi* resulted in varying topologies depending on locus occupancy percentage and whether a constraint was applied. UA50p and UA50p_ME recovered the same topology, with *A. sierra* and. *A. huntington* grouping with *Apomastus* and MY4535 sister to *A. chavezi*+MY4536+ BME101842+*A. n sp. madera* (Figures 30 & 41). UA75p recovered *A. sierra* sister to *Apomastus* and *A. huntington* sister to the rest of Sierra group species (Figure 32), which contrasts slightly with UA75p_ME placing *A. huntington*+*A. chavezi* as sister to the rest of Sierra group species (Figure 42). The constrained analyses recovered very similar topologies with the exception of *A.*

chavezi either recovered as sister to MY4535+MY4536+ BME101842+*A. n sp. madera* in UA75p_Con (Figure 33) or sister to MY4536+ BME101842+*A. n sp. madera* in UA50p_Con (Figure 31).

Adding morphological data resulted in differing topologies, with all analyses except for one recovering a monophyletic Sierra group. UN_MA recovered a paraphyletic Sierra group with respect to the Simus group. Specifically, *A. sierra*+*A. huntington*+*A. dorothealangeae*+*A. chavezi* was sister to the Simus clade, which was sister to MY4535+MY4536+ BME101842+*A. n sp. madera* (Figure 38). Our preferred tree recovered a monophyletic Sierra group with *A. sierra*+*A. huntington*+*A. dorothealangeae*+*A. chavezi* sister to MY4535+MY4536+ BME101842+*A. n sp. madera* (Figure 29), which slightly contrasts with UA75p_MA in which *A. sierra* was recovered as sister to *Apomastus* (Figure 37). UA50p_MA recovers several different relationships from our preferred tree, with *A. sierra* and *A. huntington* grouping with *Apomastus* and MY4535 being sister to *A. dorothealangeae*+*A. chavezi*+MY4536+ BME101842+*A. n sp. madera* (Figure 36). UA50p_MA_Con resulted in *A. sierra* sister to *A. huntington*+all other Sierra group species (Figure 39).

Simus Clade

Bond (2012) recovered eight species within the Simus species group: *A. simus*, *A. elisabethae*, *A. fisheri*, *A. fornax*, *A. bonoi*, *A. lucerne*, *A. satleri*, and *A. cajalco*. All phylogenetic analyses recovered a monophyletic Simus species group regardless of data type. However, sister relationships within the Simus clade differed depending on UCE locus occupancy percentage and inclusion of morphological data. Analyses with only freshly collected specimen UCE data (i.e., UN50p/75p and UN_ME; Figures 34, 35, & 43) included only *A. simus*, which was sister to the Sierra clade. Analyses with both freshly collected and museum

specimen UCE data recovered similar topologies, the only exception being the placement of *A. fisheri* as sister to *A. simus*+*A. fornax*+*A. cajalco* (i.e., UA50p, UA50p_Con, and UA50p_ME; Figures 30, 31, & 41) versus *A. cajalco* (i.e., UA75p, UA75p_Con, and UA75p_ME; Figures 32, 33, & 42).

Adding morphological data in conjunction with UCE data for combined evidence analyses resulted in varying topologies, with *A. fisheri*+*A. bonoi* being the only consistent sister relationship across all analyses. Our preferred tree topology recovered *A. elisabethae*+*A. satleri*, but the placement of *A. satleri* was different for most of the combined evidence analyses: *A. satleri*+*A. fornax* (UA50p_MA), *A. satleri*+*A. cajalco*+*A. fornax* (UA50p_MA_Con), *A. satleri*+*A. cajalco* (UA75p_MA and UN_MA), and *A. satleri*+*A. simus*+*A. fornax*+*A. elisabethae*+*A. lucerne*+*A. fisheri*+*A.bonoi* (UN_MA_Con). In addition, *A. elisabethae* was recovered as sister to all other Simus group species in UA50p_MA, UA50p_MA_Con, and UA75p_MA, but *A. elisabethae* was sister to *A. fornax* in UN_MA and UN_MA_Con. Our preferred tree and UA75p_MA recovered *A. lucerne* as the sister taxon of A. *fornax*, but the remaining analyses (i.e., UA50p_MA, UA50p_MA_Con, UN_MA, and UN_MA_Con) recovered *A. lucerne*+*A. fisheri*+*A. bonoi*, which contrasts with *A. cajalco* being the sister taxon of *A. fisheri*+*A. bonoi* in the preferred tree. Lastly, the arrangement of *A. simus* shifted across the majority of analyses: 1) sister to *A. lucerne*+*A. fornax*+*A. cajalco*+*A. fisheri*+*A. bonoi* (preferred tree), 2) sister to *A. lucerne*+*A. fornax*+*A. cajalco*+*A. satleri*+*A. fisheri*+*A. bonoi* (UA75p_MA), 3) sister to *A. cajalco*+*A. satleri*+*A. fornax* (UA50p_MA and UA50p_MA_Con), and 4) sister to *A. fornax*+*A. elisabethae*+*A. lucerne*+*A. fisheri*+*A. bonoi* (UN_MA and UN_MA_Con).

Hesperus+Atomarius Clade

Bond (2012) recovered thirteen species within the monophyletic Hesperus species group (*A. hesperus*, *A. hedinorum*, *A. sinnombre*, *A. mikeradtkei*, *A. edwardabbeyi*, *A. aguacaliente*, *A. cahuilla*, *A. killerdana*, *A. chemehuevi*. *A. serrano*, *A. anzaborrego*, *A. sarlacc*, and *A. derhamguilianii*) and fifteen species within the monophyletic Atomarius species group (*A. atomarius*, *A. dantrippi*, *A. stephencolberti*, *A. barackobamai*, *A. icenoglei*, *A. isabella*, *A. muiri*, *A. chiricahua*, *A. angelinajolieae*, *A. stanfordianus*, *A. miwok*, *A. pennjillettei*, *A. nateevansi*, *A. asmodaeus*, and *A. cabrillo*). In contrast, all of our phylogenetic analyses did not recover these two species groups as monophyletic. As stated above, the Atomarius group was rendered paraphyletic with respect to the Hesperus group, thus forming the Hesperus+Atomarius clade. Although the Hesperus group was largely monophyletic within the larger clade, *A. hedinorum* was sister to all other species in the Hesperus+Atomarius clade; additionally, *A. cabrillo*, which was previously an Atomarius species group individual, was nesting within other Hesperus group species. Consistent through all of our analyses were the sister relationships between *A. stephencolberti* and *A. atomarius*+*A. dantrippi* as well as *A. cahuilla* and *A. hesperus*. Analyses with only NC UCE data (i.e., UN50p/75p and UN_ME) recovered the same topology with high support (bootstrap values > 95; Figures 34, 35, & 43).

Similar to the other clades, sister relationships varied when including MUS UCE and morphological data. Adding MUS UCE data for *A. isabella* and *A. serrano* resulted in very similar topologies (UA50p/75p and UA_ME), with the exception of *A. isabella* in the constrained analyses (i.e., UA50p_Con and UA75p_Con). Specifically, the constrained analyses placed *A. isabella* as sister to *A. icenoglei* as opposed to the rest of the Atomarius group (excluding *A. barackobamai*). The inclusion of morphological data recovered several different relationships depending on the locus occupancy percentage and constraint application. Similar to

above, *A. isabella* was recovered as sister to *A. icenoglei* in our preferred tree topology and UA50p_MA_Con and sister to the rest of the Atomarius group (excluding *A. barackobamai*) for UA75p_MA and UA50p_MA. Both unconstrained and constrained UN_MA analyses recovered *A. isabella* sister to *A. barackobamai*. The sister relationship *A. derhamguilianii+A. sarlacc* was recovered for all UA_MA analyses except for UA50p_MA_Con, which placed *A. sarlacc* sister to *A. serrano* and *A. derhamguilianii* sister to *A. chemehuevi+A. aguacaliente+A. sarlacc+A. serrano*. Our preferred tree, UA50p_MA, and UN_MA50p/75p recovered *A. killerdana+A. mikeradtkei*, but *A. killerdana* forms different sister group relationships for UA75p_MA (*A. aguacaliente+A. chemehuevi+A. serrano*) and UA50p_MA_Con (*A. cahuilla+A. hesperus+A. edwardabbeyi+A. anzaborrego+A. derhamguilianii+A. chemehuevi+A. aguacaliente+A. sarlacc+A. serrano*).

The sister taxon of *A. chiricahua* was *A. stephencolberti+A. dantrippi+A. atomarius+A. angelinajolieae+A. stanfordianus+A. miwok+A. pennjillettei+A. nateevansi+A. asomdaeus* in our preferred tree, UA75p_MA, and UN_MA, with a similar placement recovered in UN_MA_Con except for adding *A. muiri* to the sister group; however, UA50p_MA_Con placed *A. chiricahua* sister to *A. miwok+A. pennjillettei+A. nateevansi+A. asmodaeus* and UA50p_MA sister to *A. nateevansi*. All analyses recovered *A. pennjillettei+A. nateevansi* sister to *A. asmodaeus*, with the exception of UA50p_MA placing *A. pennjillettei* as sister to all other Atomarius group species (excluding *A. barackobamai*, *A. icenoglei*, and *A. isabella*). Our preferred tree and UA50p_MA recover *A. anzaborrego* sister to *A. serrano*; in contrast, *A. anzaborrego* sister relationships differed in several ways across other analyses: 1) sister to all other Hesperus groups species (excluding *A. sinnombre*) in UA75p_MA and UN_MA_Con, 2) sister to *A. sinnombre* in UN_MA, and 3) sister to *A. derhamguilianii+A. chemehuevi+A. aguacaliente+A. sarlacc+A.*

serrano in UA50p_MA_Con. Lastly, the position of *A. chemehuevi* varied across several analyses: 1) sister to *A. derhamguilianii+A. sarlacc+A. anzaborrego+A. serrano* in the preferred tree, 2) sister to *A. aguacaliente* in UA50p_MA, UA50p_MA_Con, and UN_MA, 3) sister to all other Hesperus group species (excluding A. sinnombre, A. anzaborrego, and A. cabrillo) in UN_MA_Con, and 4) sister to *A. serrano* in UA75p_MA.

Evolution of Habitat Type

Figure 44 depicts the ancestral state reconstruction estimates of habitat type across *Aptostichus*. Based on AICc values for equal rates (ER) model and all rates different (ARD) model (ER = 154.7165; ARD = 927.8662), the equal rates model (ER) generated the favored reconstruction. The uncertainty of habitat type in the deeper nodes (i.e., habitat types are practically equivocal) make it difficult to assess whether or not the *Aptostichus* ancestor lived in multiple environments, with the desert being one of them. The Simus species group ancestor was primarily in desert environments, with one shift to alpine habitat (*A. satleri*) and one to coastal dunes (*A. simus*). The Hesperus+Atomarius species group most recent common ancestor predominantly lived in arid environments (chaparral and desert), with a subsequent shift towards deserts within the Hesperus group and a shift towards chaparral within the Atomarius group. Within the Atomarius group, *A. pennjillettei* independently shifted from chaparral back to the desert and two species (*A. stephencolberti* and *A. miwok*) each transitioned to coastal dune habitats independently. Overall, there were three independent derivations of coastal dune occupancy (i.e., the three coastal dunes species each invaded coastal dune areas separately) and at least two independent derivations of desert inhabitance.

Discussion:

The North American trapdoor spider genus *Aptostichus* is diverse, both in terms of species and ecological breadth, yet its interspecific relationships have received minimal attention using more modern approaches. Specifically, a 2012 (Bond) preliminary morphology-based phylogeny left many relationships unresolved, thereby limiting downstream comparative evolutionary studies. To fully resolve species level relationships, we attempted to sample molecular data (UCEs) from rare and/or presumed extinct *Aptostichus* species which were then included in UCE datasets. Contrasting Bond, (2012), our preferred tree topology recovered three clades: Sierra, Simus, and Hesperus+Atomarius. The addition of museum harvested samples helped elucidate a few interspecific relationships within *Aptostichus*; nevertheless, the inclusion of morphological data in conjunction with UCE data, both NC and MUS, resulted in conflicting phylogenetic relationships within the three clades. Consequently, this brings forward several questions: 1) generally how robust is the MUS UCE data, 2) can MUS UCE data confidently be employed for reconstructing phylogeny across *Aptostichus*, and 3) similarly, how confident can we be in the results of the combined-evidence analyses?

The Highs and Lows of Museomics

Museomics refers to the process of obtaining DNA sequences from historical museum specimens (i.e., combination of 'museum' and 'genomics'; Raxworthy & Smith, 2021). There are a wide range of applications for historical museum specimens, most of which have historically served as morphological vouchers. However, a recent number of studies across a diversity of taxa have employed historical museum specimens as a potential resource for subgenomic data; these recently developed protocols ostensibly allow one to use degraded DNA samples for a number of downstream applications (e.g., Buenaventura, 2021; Derkarabetian et al., 2019; Grewe et al., 2021; Hosner et al., 2016; O'Connell et al., 2022; Ruane & Austin, 2017;

Wood et al., 2018). For example, Grewe et al. (2021) sequenced several genetic markers of the *Xerces* blue butterfly to infer the butterfly was in fact a distinct species, with the caveat that their species delimitation method only included the COI barcoding gene, that had been driven to extinction. Numerous studies have now highlighted the efficacy of museum harvested specimens for phylogenetic inference across multiple taxonomic levels: ordinal (Derkarabetian et al., 2019; Hosner et al., 2016), suborder (Ruane & Austin, 2017), superfamily (Wood et al., 2018), family (Buenaventura, 2021), and species-level population genomics (O'Connell et al., 2022).

Our study successfully sampled 10 ethanol-preserved *Aptostichus* museum specimens spanning material collected 11-63 years ago. Although they averaged about half the number of loci compared to NC samples (Table 3), the majority of these taxa were largely consistent for most species group topological placements, if not retaining exactly the same sister group relationships, across analyses. However, the placements of *A. sierra* and *A. huntington* varied, thus we ultimately added a constraint to retain the well-established Sierra group as monophyletic, which then resulted in the same placement across analyses (i.e., *A. sierra* sister to *A. huntington*+all other Sierra group individuals). The displacement of these two taxa in unconstrained analyses is likely the result of very high levels of missing data (i.e., gap/ambiguity percentage > 90%), which is a natural consequence of using degraded DNA samples in target capture approaches. Moreover, the amount of allelic dropout/missing data in this study for all museum specimens was considerably higher than newly collected specimens (museum specimens > 60% compared to newly collected specimens < 50% for UCE datasets; data not shown) despite the relative topological stability across analyses. Similarly, O'Connell et al. (2022) found that higher levels of missing data introduced discordance between historical museum and freshly collected specimens in PC space, but did *not* negatively affect the

topological placement of museum specimens (i.e., they clustered with corresponding species replicates in the phylogeny). Thus, we are confident that using historical museum samples helped to inform our understanding of interspecific relationships, and that the applicability of using historical museum specimens for phylogenomics across diverse taxa is supported.

Although we were able to include 10 historical museum specimens in our UCE datasets, we did have a number of samples for which we were unable to recover sufficient numbers of loci and/or had wildly spurious phylogenetic placements across preliminary analyses (data not shown). Generally, recoverable DNA concentrations vary across parts of the specimen tissues sampled, preservation techniques, tissue types, and method for obtaining genomic-scale datasets (e.g., RADseq or target capture; O'Connell et al., 2022; Raxworthy & Smith, 2021). Thus, various factors likely contributed to the small amount of viable DNA left from which to recover UCEs from these suboptimal extractions: field collection technique (i.e., pitfall trap versus burrow extraction), preservation differences (e.g., 70% versus 80% ethanol), potential contamination, date of collection, and more (Card et al., 2020; Nakahama, 2020; Raxworthy & Smith, 2021). Each museum specimen has a unique collection history, unfortunately not always documented, leading to high variability in DNA quality even for similarly aged individuals (Raxworthy & Smith, 2021). For example, our specimens for *A. isabella* and *A. satleri* were collected 12 years ago within the same month via pitfall traps, yet *A. isabella* yielded twice the amount of extracted DNA and retained enough informative UCE loci for relatively confident phylogenetic placement, whereas *A. satleri* had lower quantity/quality DNA and a spurious phylogenetic placement.

Although the various factors potentially affecting the efficacy of historical museum DNA are not necessarily mutually exclusive, we think that one reason some of our extractions were

unsuccessful was because many *Aptostichus* species are collected via pitfall traps and consequently some samples may have been exposed to undesirable conditions over a relatively long period of time leading to various states of DNA degradation/contamination. For example, when comparing a similarly aged sample taken from a burrow and promptly placed in 80% ethanol (AP1263 = *A. muiri*) versus a sample taken from a pitfall trap where it sat for an unknown amount of time before placed in ethanol (AP901 = *A. bonoi*), the burrow extraction sample *A. muiri* yielded 8X more DNA and was consistently found as sister to a freshly collected *A. muiri* specimen. Out of the 51 attempted DNA extractions from older museum material, 17 individuals were explicitly documented as pitfall trap collections, with only four of those yielding successful extractions. However, a large number of specimens had no documentation for how the specimen was collected; therefore, eight of the successful extractions with no labeled field collection technique could have potentially been from pitfall traps.

Another reason for failure to recover UCEs could be attributed to the age of specimens. For example, Card et al. (2020) evaluated UCE capture success across three different studies and found that the proportion of target UCEs sequenced decreased by 0.34% per year on average for dried bird tissue (McCormack et al., 2016), pinned insects (Blaimer et al., 2016), and ethanol-preserved arachnids (Derkarabetian et al., 2019). We did not assess UCE recovery success rates for all museum individuals, but our evaluation of DNA quantity across museum samples did not find a clear-cut trend of decreased quantity over time; in fact, DNA quantity was quite variable through time, even those collected within the same year. For example, three individuals collected in 1972 had 0.817 (AP578 = *A. fornax*), 1.77 (AP84 = *A. dorothealangeae*), and 3.2 ng (AP409 = *A. muiri*) of DNA per µl, and two individuals collected in 2000 yielded 6.39 and 1.35 ng DNA per µl for AP1080 (*A. mikeradtkei*) and AP901 (*A. bonoi*), respectively. In addition, the number

of loci recovered for museum specimens in our data matrices did not directly correspond to age (i.e., no trend was seen for decreased loci with increased age; see Table 16). For example, our second oldest museum specimen AP400 (*A. sierra*) collected in 1959 recovered 57 loci in the UA75p data matrix whereas our sample AP562 (*A. chavezi*) collected in 1983 recovered 54 loci in the UA75p data matrix. Consequently, it seems that age alone does not account for the varying success of museum specimens, but rather potentially a combination of unknown factors. Overall, historical museum specimens have varying histories that can affect the efficacy of DNA extraction/UCE recovery, thus best practices are currently being developed that can alleviate some of these issues in the future (Derkarabetian et al., 2019; Nakahama, 2020; Raxworthy & Smith, 2021).

Combined-Evidence Phylogeny Comparisons

The morphology alone data set produced results that differ quite markedly from the phylogenies that include molecular data. For example, the morphology-only tree recovered the outgroup *Apomastus* nested with *Aptostichus* species and a paraphyletic Sierra species group (Figure 28), whereas molecular analyses placed *Apomastus* as sister to all *Aptostichus* and recovered a monophyletic Sierra species group (Figure 29). Using morphological data alone to inform our understanding of relationships between *Aptostichus* species may be difficult to justify given the prevalence of relative morphological homogeneity across mygalomorphs in general (e.g., Hendrixson & Bond, 2005; Leavitt et al., 2015; Starrett & Hedin, 2007), putative homoplasy of some characters (e.g., carapace and abdominal coloration), and the incomplete nature of our data matrix given that approximately one-third of the taxa are only known from a single sex (usually males). Such factors in combination are likely affecting topological uncertainty in the morphology-only tree (Figure 28). Specifically, the morphological characters

scored included some that potentially reflect convergent evolution across the group. For example, lighter carapace and abdominal coloration tend to be prevalent in desert and coastal dune species and thus are likely a product of the environment the species inhabit. Consequently, these species are more likely to group together based on that character alone. Anecdotally, this can be seen readily in the Simus group in which all of the desert species and the coastal dune endemic *A. simus* forms a sister group (Figure 28), whereas with UCE data that is not the case (i.e., darker colored *A. cajalco* sister to *A. fornax* and *A. elisabethae* sister to all other Simus group taxa; Figure 29). Additionally, the morphology-only tree placed the outgroup, *Apomastus*, as sister to the Sierra+Simus clade (Figure 28), which is clearly refuted by the molecular data and in all previous studies (i.e., UN50&75p; Bond et al., 2020, 2012; Bond & Opell, 2002).

For our combined evidence analyses, although a couple of taxa with only morphological data were consistently placed (i.e., *A. dorothealangeae* and *A. bonoi*), the remaining taxa had at least one conflicting sister group relationship (i.e., *A. lucerne*, *A. satleri*, *A. killerdana*, *A. chemehuevi*, *A. derhamguilianii*, *A. sarlacc*, *A. anzaborrego*, *A. chiricahua*, *A. pennjillettei*, and *A. nateevansi*). Among these conflicting topologies, all but two taxa (*A. killerdana* and *A. nateevansi*) are only known from male specimens. Hence, the topological instability of these morphology-only taxa may likely be attributed at least somewhat to the imbalance of characters scored for male-only taxa versus taxa known from both sexes. Overall, the discordance between morphological and molecular data is likely due to the issues listed above, namely homoplasy and missing data.

These analyses beg the question, is it worthwhile to combine data when specimens for molecular analysis are unavailable? More precisely, do we think it is better to have a molecular phylogeny with incomplete taxon sampling or a tree with combined molecular and

morphological data that potentially introduces problematic sister group hypotheses? We exhausted attempts to extract viable DNA from rare/presumed extinct museum taxa (e.g., sampled the breadth of available vouchers and implemented sterile conditions during DNA extractions), but not all extractions were successful. Consequently, this leaves us in an analogous situation to tip dating with fossils where the phylogenetic position is potentially unresolved due to a variety of factors (e.g., overall sparse fossil record, morphological character scoring for fossils are incomplete), which can lead to a high level of uncertainty in divergence times (Parham et al., 2012). As outlined above, we have to acknowledge that morphological data alone introduces issues leading to uncertain sister group relationships; however, the integration of UCE data for over half of *Aptostichus* taxa does abrogate, at least to some extent, confounding issues introduced with morphological data alone. As such, we believe that it is better to take a total evidence approach that combines all the specimens and data available to formulate a complete as possible phylogenetic hypothesis, which is analogous to the implementation of the fossilized birth-death process when combining fossil data and molecular sequences of extant taxa to simultaneously reconstruct the phylogeny and compute divergence times (Heath et al., 2014).

Habitat Type Evolution

The distributions of *Aptostichus* species span disparate ecoregions across the CAFP, making them ideal candidates for assessing adaptation evolution associated with the invasion of habitat types. Using our preferred tree topology, the ancestral state reconstruction estimates for habitat type indicate at least two independent shifts to deserts and three independent shifts to chaparral (Figure 44). This pattern of repeated invasion of arid habitats has been seen in other trapdoor spider families (e.g., Huey et al., 2019; Rix et al., 2017, 2021). For example, the large Australian genus *Aname* has radiated across the Australian arid zone multiple times, with some

arid-adapted ancestral taxa recolonizing mesic habitat and their descendants subsequently returning to arid zones (Rix et al., 2021). Cloudsley-Thompson (1983) reviewed several adaptations observed in desert dwelling spiders: 1) larger size for a smaller surface-to-volume ratio, 2) deeper burrows for a stable microclimate, 3) brooding females plug burrow entrances during dry season, 4) aestivation of young spiderlings in plugged burrows, 5) varying burrow structures to prevent flooding (e.g., turrets, plugs), 6) cryptic coloration, 7) longer lifespan of females, 8) burrow rim modifications to increase foraging area (e.g., 'twig-lining'), and 9) breeding and dispersal synced to the rainy season. Additionally, a study by Mason et al. (2013) comparing standard metabolic rate (SMR) and evaporative water loss (EWL) of mesic- and arid-adapted trapdoor spiders found no reduction in SMR for the arid-adapted spiders and only moderately lower EWL in arid-adapted species compared to the mesic-adapted species. This suggests that behavioral modifications, especially burrow structures, and to a lesser extent physiological/morphological modifications, have a profound effect on the adaptability of these spiders to desert and overall arid (e.g., chaparral) environments.

Increasing aridification and wildfire (both in abundance and intensity) because of factors such as climate change and fire suppression tactics (Keeley & Syphard, 2019; Pausas & Keeley, 2021) is a growing concern for trapdoor spiders, especially throughout the CAFP (Newton et al., in prep) and Australia (Main, 1995; Mason et al., 2019). However, the behavior of plugging burrows (i.e., forming a soil plug to seal the burrow; Cloudsley-Thompson, 1983) with spiderlings aestivating during the dry season may contribute to their ability to survive in drought and fire prone landscapes (Cloudsley-Thompson, 1983). For example, this behavior has recently been observed in a related arid-adapted, chaparral dwelling euctenizid genus *Promyrmekiaphila* approximately nine months post-fire (L. Newton, personal observation). Additionally, a number

of arid-adapted species in the trapdoor spider genus *Aliatypus*, also distributed across the CAFP, seal their burrows during the summer dry season and have been observed in post-fire landscapes (Coyle & Icenogle, 1994). Overall, burrow plugging and the adaptations listed above may contribute to the long-term survival of the arid-adapted trapdoor spider populations in increasingly harsh conditions. Because a large number of *Aptostichus* desert species are still only known from male specimens collected via pitfall traps, it will be imperative to obtain female specimens as vouchers and document their natural history data to provide a more well-rounded investigation of desert adaptations in *Aptostichus*. Specifically, we first need to evaluate whether or not females are performing the above-mentioned behavioral adaptations, and then use vouchers for genomic/transcriptomic data to evaluate the potential for genes under selection compared to non-desert species (i.e., physiological modifications).

Our findings also show that each of the coastal dune species (i.e., *A. simus*, *A. stephencolberti*, and *A. miwok*) independently migrated to coastal dune habitats. This contrasts with the results of Bond (2012) that supported only two shifts to coast dune environments (i.e., once in *A. simus* and once for the ancestor of *A. stephencolberti* and *A. miwok*). A preliminary study by Garrison et al. (2020) evaluated potential evidence of positive selection in two of the coastal endemic lineages (*A. stephencolberti* and *A. miwok*) compared to their inland sister species. The inferred positive selection for genes related to venom production, metabolism, and sensory systems in coastal endemics. This suggests that several adaptations are potentially needed to colonize coastal dune habitats: 1) morphological shifts to reflect psammophilic lifestyle, 2) adjustment of venom peptides for a different assemblage of prey items (i.e., divergent from inland habitats), 3) behavioral modifications (e.g., chemical communication) due to an altered signaling landscape (i.e., substrate and vegetation differences), and 4) adjustment of

metabolic responses to deal with environmental changes like temperature extremes, increased salinity, and decrease in soil moisture (Garrison et al., 2020). The newly annotated genome of *A. stephencolberti* (Bond et al., in prep) will likely be instrumental in further narrowing down the underlying genetic changes in coastal dune endemics potentially facilitating adaptation to this unstable environment.

Lastly, our findings reveal at least five independent derivations of invading alpine habitats in the Sierra Nevada. The behavioral and/or physiological adaptations would likely diverge from the arid-adapted taxa in most aspects. For example, the winters can be harsh (i.e., cold and long) and thus would be plausible for alpine populations to be inactive for most of the winter season, contrasting the behavioral adaptation of arid-adapted taxa remaining dormant during the summer dry season. As such, cold-adapted spiders will likely disperse earlier than earlier species, in the late summer or early fall to avoid intolerably cold conditions; a similar pattern has been observed in populations of the trapdoor spider *Aliatypus janus* occurring in the Sierra Nevada (Coyle & Icenogle, 1994). Interestingly, a number of *Aliatypus janus* burrows have been observed to add plant material attached around the burrow entrance rim, which may extend their foraging area like in arid-adapted spiders (Coyle & Icenogle, 1994). Because all alpine *Aptostichus* species are rare and difficult to collect, there has been no formal assessment of the potential adaptation mechanisms for alpine habitat so far. Similar to the desert species, future collecting schemes aimed towards collecting females and natural history data in alpine habitats will be imperative to develop a better understanding of the underlying forces driving this habitat type shift.

Conclusions and Future Questions

We implemented a comprehensive approach using both molecular and morphological data to reconstruct the phylogeny of *Aptostichus*, a group with a number of rare species and relative morphological homogeneity. In addition, our study underscored the utility of including historical museum specimens in sequence capture phylogenomics despite some caveats. Our re-evaluation of habitat type evolution revealed an additional independent derivation of coastal dune inhabitance compared to Bond (2012) for a total of three, and at least two independent derivations of desert inhabitance, which potentially lines up with the previous inference of three independent shifts to desert in Bond (2012). Although we are confident that our preferred tree is more robust than Bond (2012) given the inclusion of UCE data for a number of taxa, we cannot rule out that the interspecific relationships within our combined evidence phylogeny might still be confounded by morphological characters potentially associated with habitat. Thus, future collecting schemes with an emphasis on collecting vouchers *and* natural history data for enigmatic and rare species will achieve more rigorous tests of character adaptation in desert, and other environments. For coastal dune adaptations, preliminary data from transcriptomes indicate evidence of positive selection for chemosensory-associated gene families in coastal dune endemics in relation to their inland sister species (Garrison et al., 2020). In addition, the newly assembled and annotated genome of *A. stephencolberti* (Bond et al., in prep) will be the backbone of future studies investigating the potential convergence of coastal dune adaptations across these coastal dune endemic species.

References:

Atkinson, G. F. (1886). Descriptions of some new trapdoor spiders; their nests and food habits. *Entomologica Americana* **2**: 109-117, 128-137.

Ausserer, A. (1875). Zweiter Beitrag zur Kenntniss der Arachniden-Familie der Territelariae Thorell (Mygalidae Autor). *Verhandlungen der Kaiserlich-Königlichen Zoologisch-Botanischen Gesellschaft in Wien* **25**: 125-206, pl. 5-7.

Beaulieu, J. M., O'Meara, B. C., & Donoghue, M. J. (2013). Identifying Hidden Rate Changes in the Evolution of a Binary Morphological Character: The Evolution of Plant Habit in Campanulid Angiosperms. *Systematic Biology*, *62*(5), 725–737. https://doi.org/10.1093/sysbio/syt034

Blaimer, B. B., Lloyd, M. W., Guillory, W. X., & Brady, S. G. (2016). Sequence Capture and Phylogenetic Utility of Genomic Ultraconserved Elements Obtained from Pinned Insect Specimens. *PLOS ONE*, *11*(8), e0161531. https://doi.org/10.1371/journal.pone.0161531

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bond, J. E. (2004). Systematics of the Californian euctenizine spider genus *Apomastus* (Araneae: Mygalomorphae : Cyrtaucheniidae): the relationship between molecular and morphological taxonomy. *Invertebrate Systematics*, *18*(4), 361. https://doi.org/10.1071/IS04008

Bond, J. E. (2012). Phylogenetic treatment and taxonomic revision of the trapdoor spider genus *Aptostichus* Simon (Araneae, Mygalomorphae, Euctenizidae). *ZooKeys*, *252*, 1–209. https://doi.org/10.3897/zookeys.252.3588

Bond, J. E. & Godwin, R. (2013). Taxonomic revision of the Trapdoor spider genus *Eucteniza* Ausserer (Araneae, Mygalomorphae, Euctenizidae). *ZooKeys*, *356*, 31–67. https://doi.org/10.3897/zookeys.356.6227

Bond, J. E., Hamilton, C. A., Godwin, R. L., Ledford, J. M., & Starrett, J. (2020). Phylogeny, Evolution, and Biogeography of the North American Trapdoor Spider Family Euctenizidae (Araneae: Mygalomorphae) and the Discovery of a New 'Endangered Living Fossil' Along California's Central Coast. *Insect Systematics and Diversity*, *4*(5). https://doi.org/10.1093/isd/ixaa010

Bond, J. E., & Hedin, M. (2006). A total evidence assessment of the phylogeny of North American euctenizine trapdoor spiders (Araneae, Mygalomorphae, Cyrtaucheniidae) using Bayesian inference. *Molecular Phylogenetics and Evolution*, *41*(1), 70–85. https://doi.org/10.1016/j.ympev.2006.04.026

Bond, J. E., Hedin, M. C., Ramirez, M. G., & Opell, B. D. (2001). Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology*, *10*(4), 899–910. https://doi.org/10.1046/j.1365-294X.2001.01233.x

Bond, J. E., Hendrixson, B. E., Hamilton, C. A., & Hedin, M. (2012). A reconsideration of the classification of the spider Infraorder Mygalomorphae (Arachnida: Araneae) based on three nuclear genes and morphology. *PLoS ONE*, *7*(6), e38753. https://doi.org/10.1371/journal.pone.0038753

Bond, J. E., & Opell, B. D. (2002). Phylogeny and taxonomy of the genera of south-western North American Euctenizinae trapdoor spiders and their relatives (Araneae: Mygalomorphae, Cyrtaucheniidae). *Zoological Journal of the Linnean Society*, *136*(3), 487–534. https://doi.org/10.1046/j.1096-3642.2002.00035.x

Bond, J. E., & Platnick, N. I. (2007). A Taxonomic Review of the Trapdoor Spider Genus *Myrmekiaphila* (Araneae, Mygalomorphae, Cyrtaucheniidae). *American Museum*

Novitates, *3596*(1), 1–30. https://doi.org/10.1206/0003-0082(2007)3596[1:ATROTT]2.0.CO;2

Bond, J. E., & Stockman, A. K. (2008). An integrative method for delimiting cohesion species: Finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology*, *57*(4), 628–646. https://doi.org/10.1080/10635150802302443

Buenaventura, E. (2021). Museomics and phylogenomics with protein-encoding ultraconserved elements illuminate the evolution of life history and phallic morphology of flesh flies (Diptera: Sarcophagidae). *BMC Ecology and Evolution*, *21*(1), 70. https://doi.org/10.1186/s12862-021-01797-7

Candia-Ramírez, D. T., & Francke, O. F. (2020). Another stripe on the tiger makes no difference? Unexpected diversity in the widespread tiger tarantula Davus pentaloris (Araneae: Theraphosidae: Theraphosinae). *Zoological Journal of the Linnean Society*. https://doi.org/10.1093/zoolinnean/zlaa107

Capella-Gutierrez, S., Silla-Martinez, J. M., & Gabaldon, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Card, D., Shapiro, B., Giribet, G., Moritz, C., & Edwards, S. V. (2020). Museum Genomics. *Annual Review of Genetics*, *55*, 633–659.

Chamberlin, R. V. (1917). New spiders of the family Avicylariidae. *Bulletin of the Museum of Comparative Zoology* **61**: 25-75.

Chamberlin, R. V. (1919). New Californian spiders. *Journal of Entomology and Zoology* **12**: 1-17.

Cloudsley-Thompson, J. L. (1983). Desert adaptations in spiders. *Journal of Arid Environments*, *6*(4), 307–317. https://doi.org/10.1016/S0140-1963(18)31410-1

Cooper, S. J. B., Harvey, M. S., Saint, K. M., & Main, B. Y. (2011). Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: Evidence for long-term refugia within refugia. *Molecular Ecology*, *20*(15), 3219–3236. https://doi.org/10.1111/j.1365-294X.2011.05160.x

Coyle, F. A., & Icenogle, W. (1994). Natural History of the Californian Trapdoor Spider Genus *Aliatypus* (Araneae, Antrodiaetidae). *The Journal of Arachnology*, *22*(3), 225–255.

Darriba, D., Weiß, M., & Stamatakis, A. (2016). Prediction of missing sequences and branch lengths in phylogenomic data. *Bioinformatics*, *32*(9), 1331–1337. https://doi.org/10.1093/bioinformatics/btv768

Derkarabetian, S., Benavides, L. R., & Giribet, G. (2019). Sequence capture phylogenomics of historical ethanol□preserved museum specimens: Unlocking the rest of the vault. *Molecular Ecology Resources*, *19*(6), 1531–1544. https://doi.org/10.1111/1755-0998.13072

Faircloth, B. C. (2013). *Illumiprocessor: A trimmomatic wrapper for parallel adapter and quality trimming*. https://github.com/faircloth-lab/illumiprocessor

Faircloth, B. C. (2015). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics*, *32*, 786–788.

Faircloth, B. C. (2017). Identifying conserved genomic elements and designing universal bait sets to enrich them. *Methods in Ecology and Evolution*, *8*(9), 1103–1112. https://doi.org/10.1111/2041-210X.12754

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*(5), 717–726. https://doi.org/10.1093/sysbio/sys004

Garrison, N. L., Brewer, M. S., & Bond, J. E. (2020). Shifting evolutionary sands: Transcriptome characterization of the *Aptostichus atomarius* species complex. *BMC Evolutionary Biology*, 1–12. https://doi.org/10.1186/s12862-020-01606-7

Glenn, T. C., Pierson, T. W., Bayona-Vásquez, N. J., Kieran, T. J., Hoffberg, S. L., Thomas IV, J. C., Lefever, D. E., Finger, J. W., Gao, B., Bian, X., Louha, S., Kolli, R. T., Bentley, K. E., Rushmore, J., Wong, K., Shaw, T. I., Rothrock Jr, M. J., McKee, A. M., Guo, T. L., … Faircloth, B. C. (2019). Adapterama II: Universal amplicon sequencing on Illumina platforms (TaggiMatrix). *PeerJ*, *7*, e7786. https://doi.org/10.7717/peerj.7786

Godwin, R. L., & Bond, J. E. (2021). Taxonomic revision of the New World members of the trapdoor spider genus *Ummidia* Thorell (Araneae, Mygalomorphae, Halonoproctidae). *ZooKeys*, *1027*, 1–165. https://doi.org/10.3897/zookeys.1027.54888

Godwin, R. L., Opatova, V., Garrison, N. L., Hamilton, C. A., & Bond, J. E. (2018). Phylogeny of a cosmopolitan family of morphologically conserved trapdoor spiders (Mygalomorphae, Ctenizidae) using Anchored Hybrid Enrichment, with a description of the family, Halonoproctidae Pocock 1901. *Molecular Phylogenetics and Evolution*, *126*, 303–313. https://doi.org/10.1016/j.ympev.2018.04.008

Goloboff, P. A. (1993). A reanalysis of Mygalomorph spider families (Araneae). *American Museum Novitates*, *3056*, 1–32.

Grewe, F., Kronforst, M. R., Pierce, N. E., & Moreau, C. S. (2021). Museum genomics reveals the Xerces blue butterfly (*Glaucopsyche xerces*) was a distinct species driven to extinction. *Biology Letters*, *17*(7), 20210123. https://doi.org/10.1098/rsbl.2021.0123

Hamilton, C. A., Formanowicz, D. R., & Bond, J. E. (2011). Species Delimitation and Phylogeography of *Aphonopelma hentzi* (Araneae, Mygalomorphae, Theraphosidae): Cryptic Diversity in North American Tarantulas. *PLoS ONE*, *6*(10), e26207. https://doi.org/10.1371/journal.pone.0026207

Heath, T. A., Huelsenbeck, J. P., & Stadler, T. (2014). The fossilized birth–death process for coherent calibration of divergence-time estimates. *Proceedings of the National Academy of Sciences*, *111*(29). https://doi.org/10.1073/pnas.1319091111

Hedin, M., & Bond, J. E. (2006). Molecular phylogenetics of the spider infraorder Mygalomorphae using nuclear rRNA genes (18S and 28S): Conflict and agreement with the current system of classification. *Molecular Phylogenetics and Evolution*, *41*(2), 454–471. https://doi.org/10.1016/j.ympev.2006.05.017

Hedin, M., & Carlson, D. (2011). A new trapdoor spider species from the southern Coast Ranges of California (Mygalomorphae, Antrodiaetidae, *Aliatypus coylei*, sp. Nov,), including consideration of mitochondrial phylogeographic structuring. *Zootaxa*, *2963*(1), 55. https://doi.org/10.11646/zootaxa.2963.1.3

Hedin, M., Carlson, D., & Coyle, F. A. (2015). Sky island diversification meets the multispecies coalescent—Divergence in the spruce-fir moss spider (*Microhexura montivaga,* Araneae, Mygalomorphae) on the highest peaks of southern Appalachia. *Molecular Ecology*, *24*(13), 3467–3484. https://doi.org/10.1111/mec.13248

Hedin, M., Derkarabetian, S., Alfaro, A., Ramírez, M. J., & Bond, J. E. (2019). Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci. *PeerJ*, *7*, e6864. https://doi.org/10.7717/peerj.6864

Hendrixson, B. E., & Bond, J. E. (2005). Testing species boundaries in the *Antrodiaetus unicolor* complex (Araneae: Mygalomorphae: Antrodiaetidae): "Paraphyly" and cryptic diversity. *Molecular Phylogenetics and Evolution*, *36*(2), 405–416. https://doi.org/10.1016/j.ympev.2005.01.021

Hendrixson, B. E., Guice, A. V., & Bond, J. E. (2015). Integrative species delimitation and conservation of tarantulas (Araneae, Mygalomorphae, Theraphosidae) from a North American biodiversity hotspot. *Insect Conservation and Diversity*, *8*(2), 120–131. https://doi.org/10.1111/icad.12089

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Hosner, P. A., Faircloth, B. C., Glenn, T. C., Braun, E. L., & Kimball, R. T. (2016). Avoiding Missing Data Biases in Phylogenomic Inference: An Empirical Study in the Landfowl (Aves: Galliformes). *Molecular Biology and Evolution*, *33*(4), 1110–1125. https://doi.org/10.1093/molbev/msv347

Huey, J. A., Hillyer, M. J., & Harvey, M. S. (2019). Phylogenetic relationships and biogeographic history of the Australian trapdoor spider genus *Conothele* (Araneae: Mygalomorphae: Halonoproctidae): diversification into arid habitats in an otherwise tropical radiation. *Invertebrate Systematics*. https://doi.org/10.1071/IS18078

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Keeley, J. E., & Syphard, A. D. (2019). Twenty-first century California, USA, wildfires: Fuel-dominated vs. wind-dominated fires. *Fire Ecology*, *15*(1), 24, s42408-019-0041–0. https://doi.org/10.1186/s42408-019-0041-0

Kulkarni, S., Wood, H., Lloyd, M., & Hormiga, G. (2020). Spider□specific probe set for ultraconserved elements offers new perspectives on the evolutionary history of spiders (Arachnida, Araneae). *Molecular Ecology Resources*, *20*(1), 185–203. https://doi.org/10.1111/1755-0998.13099

Leavitt, D. H., Starrett, J., Westphal, M. F., & Hedin, M. (2015). Multilocus sequence data reveal dozens of putative cryptic species in a radiation of endemic Californian mygalomorph spiders (Araneae, Mygalomorphae, Nemesiidae). *Molecular Phylogenetics and Evolution*, *91*, 56–67. https://doi.org/10.1016/j.ympev.2015.05.016

Maddison, W. P., Beattie, I., Marathe, K., Ng, P. Y. C., Kanesharatnam, N., Benjamin, S. P., & Kunte, K. (2020). A phylogenetic and taxonomic review of baviine jumping spiders (Araneae, Salticidae, Baviini). *ZooKeys*, *1004*, 27–97. https://doi.org/10.3897/zookeys.1004.57526

Main, B. Y. (1995). Survival of trapdoor spider during & after fire. *CALMScience Supplement*, *4*, 207–216.

Mason, L., Bateman, P. W., Miller, B. P., & Wardell-Johnson, G. W. (2019). Ashes to ashes: Intense fires extinguish populations of urban short-range endemics. *Austral Ecology*, *44*(3), 514–522. https://doi.org/10.1111/aec.12685

Mason, L. D., Tomlinson, S., Withers, P. C., & Main, B. Y. (2013). Thermal and hygric physiology of Australian burrowing mygalomorph spiders (*Aganippe* spp.). *Journal of Comparative Physiology B*, *183*(1), 71–82. https://doi.org/10.1007/s00360-012-0681-8

Mason, L. D., Wardell-Johnson, G., & Main, B. Y. (2018). The longest-lived spider: Mygalomorphs dig deep, and persevere. *Pacific Conservation Biology*, *24*(2), 203. https://doi.org/10.1071/PC18015

McCormack, J. E., Tsai, W. L. E., & Faircloth, B. C. (2016). Sequence capture of ultraconserved elements from bird museum specimens. *Molecular Ecology Resources*, *16*(5), 1189–1203. https://doi.org/10.1111/1755-0998.12466

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Montes de Oca, L., Indicatti, R. P., Opatova, V., Almeida, M., Pérez-Miles, F., & Bond, J. E. (2022). Phylogenomic analysis, reclassification, and evolution of South American nemesioid burrowing mygalomorph spiders. *Molecular Phylogenetics and Evolution*, *168*, 107377. https://doi.org/10.1016/j.ympev.2021.107377

Nakahama, N. (2020). Museum specimens: An overlooked and valuable material for conservation genetics. *Ecological Research*, *36*(1), 1–11. https://doi.org/10.1111/1440-1703.12181

Newton, L. G., Starrett, J., Hendrixson, B. E., Derkarabetian, S., & Bond, J. E. (2020). Integrative species delimitation reveals cryptic diversity in the southern Appalachian

Antrodiaetus unicolor (Araneae: Antrodiaetidae) species complex. *Molecular Ecology*, *29*(12), 2269–2287. https://doi.org/10.1111/mec.15483

O'Connell, K. A., Mulder, K. P., Wynn, A., Queiroz, K., & Bell, R. C. (2022). Genomic library preparation and hybridization capture of formalin□fixed tissues and allozyme supernatant for population genomics and considerations for combining capture□ and RADseq□based single nucleotide polymorphism data sets. *Molecular Ecology Resources*, *22*(2), 487–502. https://doi.org/10.1111/1755-0998.13481

Opatova, V., Hamilton, C. A., Hedin, M., De Oca, L. M., Král, J., & Bond, J. E. (2019). Phylogenetic systematics and evolution of the spider infraorder Mygalomorphae using genomic scale data. *Systematic Biology*, *syz064*. https://doi.org/10.1093/sysbio/syz064

Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Parham, J. F., Donoghue, P. C. J., Bell, C. J., Calway, T. D., Head, J. J., Holroyd, P. A., Inoue, J. G., Irmis, R. B., Joyce, W. G., Ksepka, D. T., Patané, J. S. L., Smith, N. D., Tarver, J. E., van Tuinen, M., Yang, Z., Angielczyk, K. D., Greenwood, J. M., Hipsley, C. A., Jacobs, L., … Benton, M. J. (2012). Best Practices for Justifying Fossil Calibrations. *Systematic Biology*, *61*(2), 346–359. https://doi.org/10.1093/sysbio/syr107

Pausas, J. G., & Keeley, J. E. (2021). Wildfires and global change. *Frontiers in Ecology and the Environment*, *19*(7), 387–395. https://doi.org/10.1002/fee.2359

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., & Korobeynikov, A. (2020). Using SPAdes De Novo Assembler. *Current Protocols in Bioinformatics*, *70*(1). https://doi.org/10.1002/cpbi.102

Raven, R. J. (1985). The spider infraorder Mygalomorphae (Araneae): Cladistics and systematics. *Bulletin of the American Museum of Natural History*, *182*, 1–180.

Raxworthy, C. J., & Smith, B. T. (2021). Mining museums for historical DNA: Advances and challenges in museomics. *Trends in Ecology & Evolution*, S0169534721002147. https://doi.org/10.1016/j.tree.2021.07.009

Rix, M. G., Cooper, S. J. B., Meusemann, K., Klopfstein, S., Harrison, S. E., Harvey, M. S., & Austin, A. D. (2017). Post-Eocene climate change across continental Australia and the diversification of Australasian spiny trapdoor spiders (Idiopidae: Arbanitinae). *Molecular Phylogenetics and Evolution*, *109*, 302–320. https://doi.org/10.1016/j.ympev.2017.01.008

Rix, M. G., Wilson, J. D., & Harvey, M. S. (2020). First phylogenetic assessment and taxonomic synopsis of the open-holed trapdoor spider genus *Namea* (Mygalomorphae: Anamidae): a highly diverse mygalomorph lineage from Australia's tropical eastern rainforests. *Invertebrate Systematics*. https://doi.org/10.1071/IS20004

Rix, M. G., Wilson, J. D., Huey, J. A., Hillyer, M. J., Gruber, K., & Harvey, M. S. (2021). Diversification of the mygalomorph spider genus *Aname* (Araneae: Anamidae) across the Australian arid zone: Tracing the evolution and biogeography of a continent-wide radiation. *Molecular Phylogenetics and Evolution*, *160*, 107127. https://doi.org/10.1016/j.ympev.2021.107127

Ruane, S., & Austin, C. C. (2017). Phylogenomics using formalin□fixed and 100+ year□old intractable natural history specimens. *Molecular Ecology Resources*, *17*(5), 1003–1008. https://doi.org/10.1111/1755-0998.12655

Simon, E. (1891). Liste des espéces de la famille des Avicullariidae qui habitent le Mexique et l'Amérique du Nord. *Actes de La Société Linnéenne de Bordeaux*, *44*, 307–326.

Smith, C. P. (1908). A preliminary study of the Araneae Theraphosidae of California. *Annals of the Entomological Society of America* **1**: 207-236.

Starrett, J., Derkarabetian, S., Hedin, M., Bryson, R. W., McCormack, J. E., & Faircloth, B. C. (2017). High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. *Molecular Ecology Resources*, *17*(4), 812–823. https://doi.org/10.1111/1755-0998.12621

Starrett, J., & Hedin, M. (2007). Multilocus genealogies reveal multiple cryptic species and biogeographical complexity in the California turret spider *Antrodiaetus riversi* (Mygalomorphae, Antrodiaetidae): Cryptic diversification in Californian *A. riversi*. *Molecular Ecology*, *16*(3), 583–604. https://doi.org/10.1111/j.1365-294X.2006.03164.x

Stockman, A., & Bond, J. E. (2008). A taxonomic review of the trapdoor spider genus *Promyrmekiaphila* Schenkel (Araneae, Mygalomorphae, Cyrtaucheniidae, Euctenizinae). *Zootaxa*, *1823*(1), 25. https://doi.org/10.11646/zootaxa.1823.1.2

Tin, M. M.-Y., Economo, E. P., & Mikheyev, A. S. (2014). Sequencing Degraded DNA from Non-Destructively Sampled Museum Specimens for RAD-Tagging and Low-Coverage Shotgun Phylogenetics. *PLoS ONE*, *9*(5), e96793. https://doi.org/10.1371/journal.pone.0096793

Valdez-Mondragón, A., & Cortez-Roldán, M. R. (2016). On the trapdoor spiders of Mexico: Description of the first new species of the spider genus *Aptostichus* from Mexico and the description of the female of *Eucteniza zapatista* (Araneae, Mygalomorphae, Euctenizidae). *ZooKeys*, *641*, 81–102. https://doi.org/10.3897/zookeys.641.10521

Wood, H. M., González, V. L., Lloyd, M., Coddington, J., & Scharff, N. (2018). Next-generation museum genomics: Phylogenetic relationships among palpimanoid spiders using

sequence capture techniques (Araneae: Palpimanoidea). *Molecular Phylogenetics and Evolution*, *127*, 907–918. https://doi.org/10.1016/j.ympev.2018.06.038

Tables:

Table 13. Locality information for all taxa sampled in this study.

Specimen ID	Specific Epithet	Museum/Fresh Material	Sequencing Location	Locality Details	Latitude	Longitude	Country	State	County	Collection Date	Collectors
AP400	*Aptostichus sierra*	Museum	UC Davis DNA Technologies Core	Shaver Lake	37.1129	-119.3095	USA	California	Fresno	09/12/1959	W.J. Gertsch and V. Roth
AP408	*Aptostichus huntingtion*	Museum	UC Davis DNA Technologies Core	Billy Creek at Huntington Lake	37.2379	-119.2295	USA	California	Fresno	08/21/1984	J. Halstead
MY4535	*Aptostichus* sp.	Fresh	Rapid Genomics	Horse Creek CG	36.3887	-118.9548	USA	California	Tulare	3/14/2011	A Schönhofer
MY4536	*Aptostichus* sp.	Fresh	Rapid Genomics	Horse Creek CG	36.3887	-118.9548	USA	California	Tulare	3/14/2011	A Schönhofer
BME101842	*Aptostichus* sp.	Fresh	Rapid Genomics	Hwy 245 N of Woodlake	36.53616	-119.12818	USA	California	Tulare	08/10/2007	J. Starrett, J. Dejas, D. Leavitt, S. Thomas

MY3261	*Aptostichus* n. sp. madera	Fresh	Rapid Genomics	County Rd 222, 2.9 mi S of CR 200	37.18369	-119.50723	USA	California	Madera	06/06/2005	A. Stockman
AP562	*Aptostichus chavezi*	Museum	UC Davis DNA Technologies Core	Ash Mountain, Kaweah Power Station #3 (40 miles Northeat Visalia)	36.488	-118.837	USA	California	Tulare	03/03/1983	D.J. Burdick
AP321	*Aptostichus elisabethae*	Museum	UC Davis DNA Technologies Core	Pisgah Crater	34.7465	-116.3755	USA	California	San Bernardino	02/25/1961	Norris & Heath
AP333	*Aptostichus elisabethae*	Museum	UC Davis DNA Technologies Core	Pisgah Crater	34.7465	-116.3755	USA	California	San Bernardino	02/11/1961	Norris & Heath
BME101032	*Aptostichus simus*	Fresh	UC Davis DNA Technologies Core	San Buenaventura St. Bch.	34.26914	-119.27978	USA	California	Ventura	12/29/2019	JE Bond
BME101348	*Aptostichus simus*	Fresh	UC Davis DNA Technologies Core	Malibu, Broad Beach	34.03385	-118.85172	USA	California	Los Angeles		M. Ramirez

AP1238	*Aptostichus cajalco*	Museum	UC Davis DNA Technologies Core	1.8 miles (road) west of Lake Matthews, just south of Cajalco Rd	33.8256	-117.4957	USA	California	Riverside	09/06/1999	W.R. Icengole
AP418	*Aptostichus fornax*	Museum	UC Davis DNA Technologies Core	Panamint Valley	36.09167	-117.25917	USA	California	Inyo	02/17/1972	Derham Giuliani
AP578	*Aptostichus fornax*	Museum	UC Davis DNA Technologies Core	Panamint Valley	36.09167	-117.25917	USA	California	Inyo	09/27/1974	J. Pinto & A. Cohen
AP954	*Aptostichus fisheri*	Museum	UC Davis DNA Technologies Core	east Dove Spring Canyon	35.43813	-118.04607	USA	California	Kern	10/01/2003	USGS-BRD San Diego Sta.
MY3826	*Aptostichus satleri*	Museum	UC Davis DNA Technologies Core	Keyesville Recreation Area, off hwy 155, ~1.6km SW of Lake Isabella Dam	35.6358	-118.4953	USA	California	Kern	10/8-29/2010	J. Satler

MY3828	*Aptostichus satleri*	Museum	UC Davis DNA Technologies Core	Keyesville Recreation Area, off hwy 155, ~1.6km SW of Lake Isabella Dam	35.6358	-118.4953	USA	California	Kern	10/8-29/2010	J. Satler
MY645	*Aptostichus hedinorum*	Fresh	UC Davis DNA Technologies Core	ABSP, ~0.5 miles NE of Hayden Springs	32.71118	-116.11602	USA	California	San Diego	10/19/2002	M. Hedin
UCD25	*Aptostichus hedinorum*	Fresh	UC Davis DNA Technologies Core	Anza Borrego			USA	California	San Diego	11/3/2018	J. Bond & M. Hedin
MY2496	*Aptostichus hesperus*	Fresh	UC Davis DNA Technologies Core	Winchester, just east of Icenogle residence, end of Grand Ave	33.71568	-117.09365	USA	California	Riverside	01/29/2004	J. Bond
MY2521	*Aptostichus cahuilla*	Fresh	UC Davis DNA Technologies Core	just south of Winchester on Leona Rd, ~1.0	33.67712	-117.11578	USA	California	Riverside	02/01/2004	J. Bond

				miles south of intersection with Patton Avenue							
BME102848	*Aptostichus mikeradtkei*	Fresh	Rapid Genomics	Ran Ranch West Trail @ Los Penasquitos Canyon Preserve	32.9429	-117.1737	USA	California	San Diego	06/24/2021	L. Newton, J. Starrett
MY2279	*Aptostichus edwardabbeyi*	Fresh	UC Davis DNA Technologies Core	Santa Rita Mtns., Madera Canyon	31.725	-110.8794	USA	Arizona	Santa Cruz	05/06/1997	D. Maddison
MY3761	*Aptostichus sinnombre*	Fresh	UC Davis DNA Technologies Core	Anza Borrego Desert State Park, Indian Gorge, ~0.25 mi W mouth Torote Canyon	32.86923	-116.2374	USA	California	San Diego	02/20/2009	M.C. Hedin
AP712	*Aptostichus serrano*	Museum	UC Davis DNA Techno	Joshua Tree National Park turnout	34.08667	-115.47778	USA	California	San Bernardino	01/17/1997	J. Bond & W. Icenogle

			logies Core								
MY2515	*Aptostichus aguacaliente*	Fresh	UC Davis DNA Technologies Core	Windy Point Area, Snow Creek Rd. exit off of HWY111	33.91099	-116.67579	USA	California	Riverside	02/05/2004	J. Bond
MY2487	*Aptostichus atomarius*	Fresh	UC Davis DNA Technologies Core	Ortega HWY H74, ~1.7 miles North Orange Co/Riverside Co line	33.61276	-117.43462	USA	California	Riverside	02/02/2004	J. Bond
BME102834	*Aptostichus atomarius*	Fresh	Rapid Genomics	CA-Hwy 76, Cleveland NF across from San Luis Rey picnic area	33.2533	-116.7922	USA	California	San Diego	06/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
BME101054	*Aptostichus stephencolberti*	Fresh	UC Davis DNA Technologies Core	Moss Landing State Beach	36.81462	-121.7905	USA	California	Monterey	10/01/2019	J. Bond

MY3070	*Aptostichus stephencolberti*	Fresh	UC Davis DNA Technologies Core	Marina Dunes Natural Preserve	36.6905	-121.8105	USA	California	Monterey	03/17/2005	J. Bond, D. Beamer, A. Stockman
MY302	*Aptostichus miwok*	Fresh	UC Davis DNA Technologies Core	Clam Beach Co. Park, just S Crannell Rd exit on Clam Beach Dr, near Little River	41.01333	-124.10923	USA	California	Humboldt	01/13/2002	J. Bond & M. Hedin
MY3817	*Aptostichus dantrippi*	Fresh	UC Davis DNA Technologies Core	Breckenridge Rd, 33km E int w Comanche Dr, NE Edison	35.4843	-118.6477	USA	California	Kern	3/28/2011	J Satler, S Derkarabetian, C Richart, P van Niekerk
BME101021	*Aptostichus dantrippi*	Fresh	UC Davis DNA Technologies Core	1.5 miles N intersection Caliente Bodfish Rd & Caliente Crk Rd.	35.32234	-118.59161	USA	California	Kern	12/20/2020	JE Bond & J Starrett

MY632	*Aptostichus angelinajolieae*	Fresh	UC Davis DNA Technologies Core	Monterey, ~100 yards N of intersection of Vieja & Valenzuela Rds	36.57597	-121.89967	USA	California	Monterey	07/27/2002	M. Hedin, P. Paquin, J. Starrett
UCD30	*Aptostichus angelinajolieae*	Fresh	UC Davis DNA Technologies Core	Hastings Natural History Reservation			USA	California	Monterey	1/20/2019	R. Godwin
MY705	*Aptostichus stanfordianus*	Fresh	UC Davis DNA Technologies Core	Mount Madonna County Park, along Mt Madonna road, 0.45 rd miles from junction Pole Line Rd	37.01167	-121.72083	USA	California	Santa Cruz	12/15/1999	J. Bond & M. van der Merwe
BMEA 101040	*Aptostichus asmodaeus*	Fresh	UC Davis DNA Technologies Core	Mt. Diablo	37.87631	-121.96114	USA	California	Contra Costa	09/23/2020	J. Bond

BMEA 101047	*Aptostichus asmodaeus*	Fresh	UC Davis DNA Technologies Core	Mt. Diablo	37.8653	-121.93124	USA	California	Contra Costa	09/23/2020	J. Bond
MY2600	*Aptostichus icenoglei*	Fresh	UC Davis DNA Technologies Core	Puente Hills, intersection of Azusa & Tomich Rd	33.98161	-117.93351	USA	California	Los Angeles	03/14/2004	J. Bond, C. Spruill, D. Beamer
BME10 2752	*Aptostichus icenoglei*	Fresh	Rapid Genomics	Cleveland Forest Rd., Cleveland National Forest	33.5289	-117.3885	USA	California	Riverside	05/13/2021	L. Newton, J. Starrett
BME10 2828	*Aptostichus icenoglei*	Fresh	Rapid Genomics	CA-Hwy 76, Morettis Junction	33.2015	-116.7118	USA	California	San Diego	06/22/2021	L. Newton, J. Starrett, R. Ruedas, B. Gibson
MY3824	*Aptostichus isabella*	Museum	UC Davis DNA Technologies Core	Erskine Creek Rd., 5.6km E or intersection with Lake Isabella Blvd.,	35.5689	-118.4383	USA	California	Kern	10/8-29/2010	J. Satler

				E of Bodfish							
BME102237	*Aptostichus barackobamai*	Fresh	UC Davis DNA Technologies Core	Trinity Mountain Rd	40.6851	-122.6395	USA	California	Shasta	02/04/2021	L. Newton
BME102241	*Aptostichus barackobamai*	Fresh	UC Davis DNA Technologies Core	Middle Creek Rd.	40.5955	-122.4528	USA	California	Shasta	02/04/2021	L. Newton
MY3801	*Aptostichus cabrillo*	Fresh	UC Davis DNA Technologies Core	Cabrillo National Monument, Point Loma	32.7101	-117.2523	USA	California	San Diego	07/13/2009	J. Satler
BME102850	*Aptostichus cabrillo*	Fresh	Rapid Genomics	Torrey Pines State Reserve Extension, Mar Scenic Trail	32.9459	-117.2543	USA	California	San Diego	06/24/2021	L. Newton, J. Starrett
AP1263	*Aptostichus muiri*	Museum	UC Davis DNA Technologies Core	Yosemite National Park, west facing slope of Valley, off of "4	37.72261	-119.59438	USA	California	Mariposa	05/10/1997	J. Bond

				Mile" trailhead							
BME101845	*Aptostichus muiri*	Fresh	Rapid Genomics	Auberry Rd., 0.4 mi S of N Fork Rd.	37.21371	-119.50503	USA	California	Madera	06/25/2012	J. Starrett

Table 14. IQ-TREE Analyses summary.

Analysis Name	**UCE Data Matrix**	**Morphology Data Matrix**	**Constraint Applied**	**Log Likelihood Value**
UA50p	50% minimum locus occupancy dataset with both fresh and musueum taxa	N/A	No	-4815810.896
UA50p_Con	50% minimum locus occupancy dataset with both fresh and musueum taxa	N/A	Yes	-4910971.596
UA75p	75% minimum locus occupancy dataset with both fresh and musueum taxa	N/A	No	-1806459.477
UA75p_Con	75% minimum locus occupancy dataset with both fresh and musueum taxa	N/A	Yes	-1843708.312
UN50p	50% minimum locus occupancy dataset with only fresh taxa	N/A	No	-5280041.221
UN75p	75% minimum locus occupancy dataset with only fresh taxa	N/A	No	-3332343.322
MA	N/A	all *Aptostichus* species and 4 outgroup taxa	No	-1222.828

UA50p_MA	50% minimum locus occupancy dataset with both fresh and musueum taxa	all taxa except 4 recently sampled individuals	No	-4814744.471
UA75p_MA	75% minimum locus occupancy dataset with both fresh and musueum taxa	all taxa except 4 recently sampled individuals	No	-1807979.483
UN_MA	75% minimum locus occupancy dataset with only fresh taxa	all taxa except 4 recently sampled individuals	No	-3333706.854
UA50p_MA_Con	50% minimum locus occupancy dataset with both fresh and musueum taxa	all taxa except 4 recently sampled individuals	Yes	-4829291.659
UA75p_MA_Con	75% minimum locus occupancy dataset with both fresh and musueum taxa	all taxa except 4 recently sampled individuals	Yes	-1808772.536
UN_MA_Con	75% minimum locus occupancy dataset with only fresh taxa	all taxa except 4 recently sampled individuals	Yes	-3335075.922
UA50p_ME	50% minimum locus occupancy dataset with both fresh and musueum taxa	only taxa with corresponding UCE data	No	-4814714.101
UA75p_ME	75% minimum locus occupancy dataset with both fresh and musueum taxa	only taxa with corresponding UCE data	No	-1807810.107
UN_ME	75% minimum locus occupancy dataset with only fresh taxa	only taxa with corresponding UCE data	No	-3333294.754

Table 15. Calculations for averaging branch lengths of MUS individuals.

sierra clade internal nodes	**simus group tips**
0.0113	0.0043
0.0149	0.0054
0.0215	**0.00485**

0.0159	
sierra clade tips	**simus group internal nodes**
0.0056	0.0245
0.0063	0.016
0.0077	0.061
0.0149	0.0288
0.008625	**0.032575**
hesperus clade tips	**hesperus clade sister group**
0.0021	0.0131
0.0017	0.0014
0.0029	0.0092
0.0064	**0.0079**
0.0066	
0.0094	**hesperus clade next level**
0.0099	0.0015
0.0095	0.0018
0.0065	0.0024
0.006111111111	0.0016
	0.001825
Ice clade tips	**atomarius clade tips**
0.0044	0.0061
0.0026	0.0041
0.0064	0.0033
0.0049	0.0046
0.004575	0.0038

	0.0021
	0.004

Table 16. UCE Data for both MUS and NC samples.

Specimen ID	**Specific Epithet**	**Total Cleaned**	**Total Contigs**	**Total Base**	**Mean Contig**	**95 CI length**	**Min Contig**	**Max Contig**	**Median Length**	**Contigs > 1kb**	**# Loci - UA50p**	**# Loci - UA75p**	**# Loci - UN 50p**	**# Loci - UN 75p**
AP321	AP562	MY3261	BME101842	MY4536	MY4535	AP408	AP400							
Aptostichus	*Aptostichus*	*Aptostichus* n. sp.	*Aptostichus* sp.	*Aptostichus* sp.	*Aptostichus* sp.	*Aptostichus*	*Aptostichus sierra*							
140128	191920	3643970	1891630	1761212	1079946	97666	33096							
234	179	1178	1066	1050	368	34	138							
65500	50063	1637484	1327625	775662	233211	9273	40217							
279.9	279.7	1390.1	1245.4	738.7	633.7	272.7	291.4							
8	4.7	12.8	13.5	7.8	15.9	5.2	4.9							
229	229	233	173	179	230	230	229							
2010	862	5557	3177	4147	3878	339	561							
259	264	1421	1269	742	607.5	270	273.5							
1	0	996	786	90	27	0	0							
201	140	1054	948	929	284	27	113							
86	54	404	391	379	153	11	57							
N/A	N/A	1085	979	956	305	N/A	N/A							
N/A	N/A	728	680	657	227	N/A	N/A							

UCD25	MY645	AP954	AP578	AP418	AP1238 B	AP1238 A	BME101 348	BME101 032	AP333
Aptostichus	*Aptostichus*	*Aptostichus fisheri*	*Aptostichus fornax*	*Aptostichus fornax*	*Aptostichus*	*Aptostichus*	*Aptostichus simus*	*Aptostichus simus*	*Aptostichus*
5167328	3608492	50594	761916	158818	489314	750544	2562058	4167416	84892
1349	1233	577	769	19	1119	1138	1329	1347	74
1507421	1505580	191153	236537	4623	478272	466856	1245754	1517101	20241
1117.4	1221.1	331.3	307.6	243.3	427.4	410.2	937.4	1126.3	273.5
11	12.7	8.1	2.6	13.7	7	3.8	10.4	15.5	6.4
230	233	229	101	80	229	211	229	229	80
4524	5429	3738	934	291	2808	1303	4867	9660	492
1157	1233	289	293	255	382	395	941	1122	264
937	896	7	0	0	29	5	542	886	0
1144	1115	463	625	625	914	939	1183	1159	201
420	421	204	261	261	371	370	434	433	86
1175	1146	N/A	N/A	N/A	N/A	N/A	1211	1194	N/A
771	757	N/A	N/A	N/A	N/A	N/A	791	786	N/A

BME101 054	BME102 834	MY2487	MY2515	AP712	MY3761	MY2279	BME102 848	MY2521	MY2496
Aptostich us	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*	*Aptostich us*
888856	979192	2643118	2253442	302902	2746104	5115370	1308574	2709694	3261696
1243	1139	1173	1136	868	995	1341	1160	1217	1041
794000	999910	1115166	1101867	299649	827194	661476	1016192	1477035	927789
638.8	877.9	950.7	970	345.2	831.4	493.3	876	1213.7	891.2
8.1	10	11.1	9.7	6	8.8	18.1	9.1	15.7	9.6
230	192	229	231	80	230	200	208	229	228
4847	5722	4644	3087	3010	1596	15599	2795	8019	2345
621	870	952	980	308	850	427	871	1187	904
43	313	510	542	11	286	18	334	879	403
1128	1034	1074	1054	711	915	1121	1047	1019	946
421	409	421	415	298	360	407	408	415	383
1147	1062	1098	1072	N/A	939	1158	1079	1052	971
753	703	759	727	N/A	668	746	726	725	680

MY2600	BMEA101047	BMEA101040	MY705	UCD30	MY632	BME101021	MY3817	MY302	MY3070
Aptostichus	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus miwok*	*Aptostichus*
2674454	3955614	4254868	2823574	2189664	2629688	1677040	2956742	4152970	2174188
1184	1291	1302	1117	1260	1129	1206	1195	1428	1179
1298298	1856957	1545631	1168312	1310699	1129919	1227096	974781	1861716	1232578
1096.5	1438.4	1187.1	1045.9	1040.2	1000.8	1017.5	815.7	1303.7	1045.4
11	84.6	14.7	12.2	11.5	9.5	14.7	7.9	14.5	11.6
180	184	184	231	225	231	229	79	234	211
3905	62571	7583	4084	5812	2583	6553	2411	6897	6179
1100	1215	1163	1016	1053	1013	982.5	834	1277.5	1050
743	962	906	592	718	585	567	319	1144	673
1083	1085	1086	1019	1138	1036	1096	1117	1205	1111
418	427	420	415	422	412	421	441	441	430
1114	1115	1118	1048	1174	1062	1126	1142	1239	1138
745	770	750	724	783	730	758	774	808	772

MY216	BME101 845	AP1263	BME102 850	MY3801	BME102 241	BME102 237	MY3824	BME102 828	BME102 752
Apomastus	*Aptostichus muiri*	*Aptostichus muiri*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*	*Aptostichus*
1351228	1797994	1504088	1139858	2140166	1593040	1506586	70808	1193796	1413426
1129	921	1129	1253	1133	1354	1344	351	1193	1114
1067824	1099802	401962	962539	989443	1429568	1517872	112875	991958	1094827
945.8	1194.1	356	768.2	873.3	1055.8	1129.4	321.6	831.5	982.8
9.2	14.3	3.2	8.3	8.6	11.6	14.1	5.4	8.7	8.6
208	232	229	229	231	241	230	229	191	231
3193	4255	2140	5676	2086	4637	7288	1077	3267	2344
967	1192	340	757	892	1029	1074.5	290	831	1000
511	660	1	150	400	733	827	2	247	558
1010	830	927	1120	1058	1178	1163	275	1076	1026
378	312	375	418	428	430	432	112	418	406
1013	859	N/A	1153	1071	1216	1198	N/A	1114	1062
690	566	N/A	758	749	797	786	N/A	747	702

AUMS0081	*Myrmekiaphila*	794790	998	902596	904.4	9.6	239	2294	918	402	890	359	915	641
AUMS11404	*Entychides*	1001266	1229	1088521	885.7	8.2	237	2377	901	445	1091	417	1119	736
BME101267	*Promyrmekiaphila*	3520636	1110	900353	811.1	7.6	171	1890	831.5	272	988	395	1039	686

Table 17. UCE stats for MUS versus NC UCE data.

	MUS Specimens	**NC Specimens**	**Difference between MUS and NC**
Average Reads	356668.2	2440254.9	2083586.7
Average Contigs	509.9	1169.3	659.4
Average Loci (50p)	473.9	1041.1	567.2
Average Loci (75p)	195.8	403	207.2

Figures:

Figure 24. Distribution map of the Atomarius species group. See legend in the bottom left corner for color designation of species.

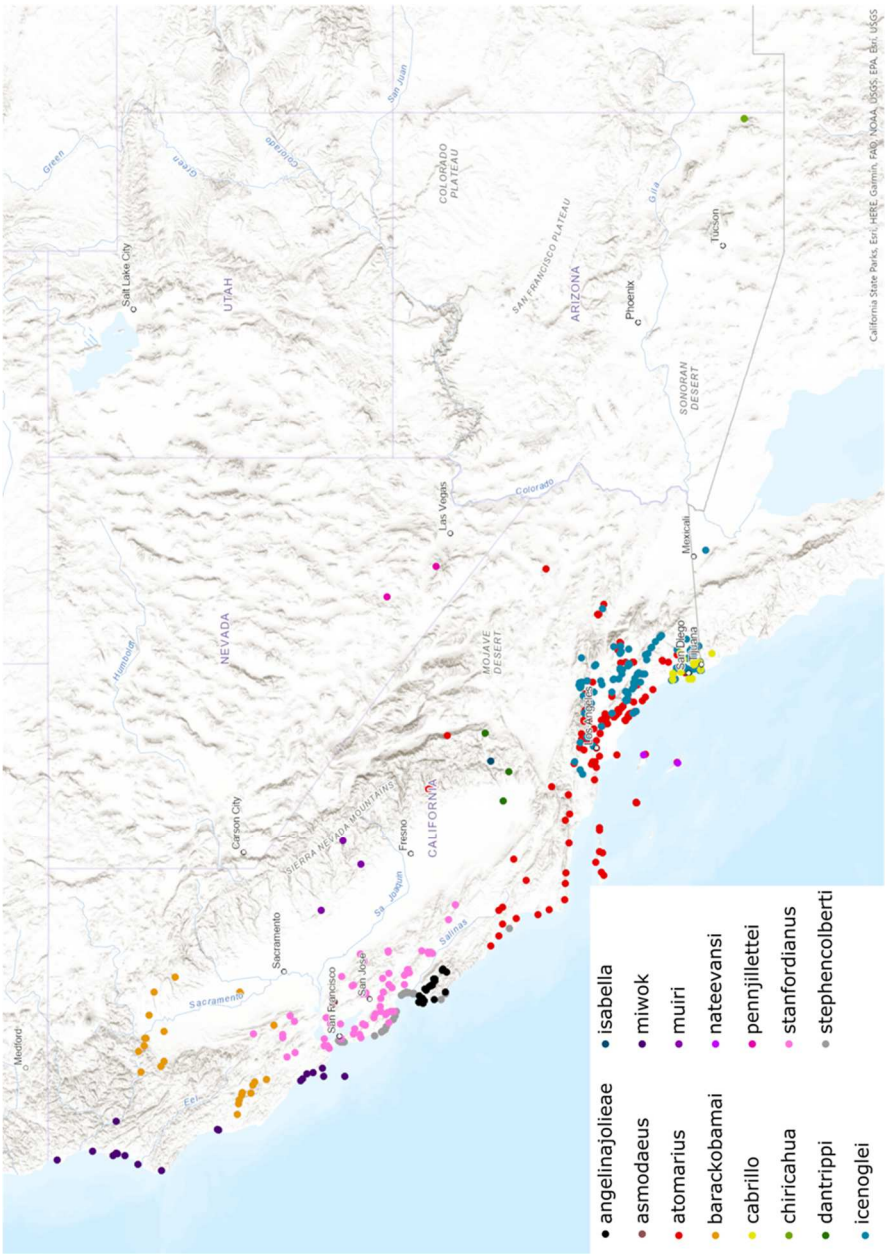


Figure 25. Distribution map of the Hesperus species group. See legend in the top right corner for color designation of species.

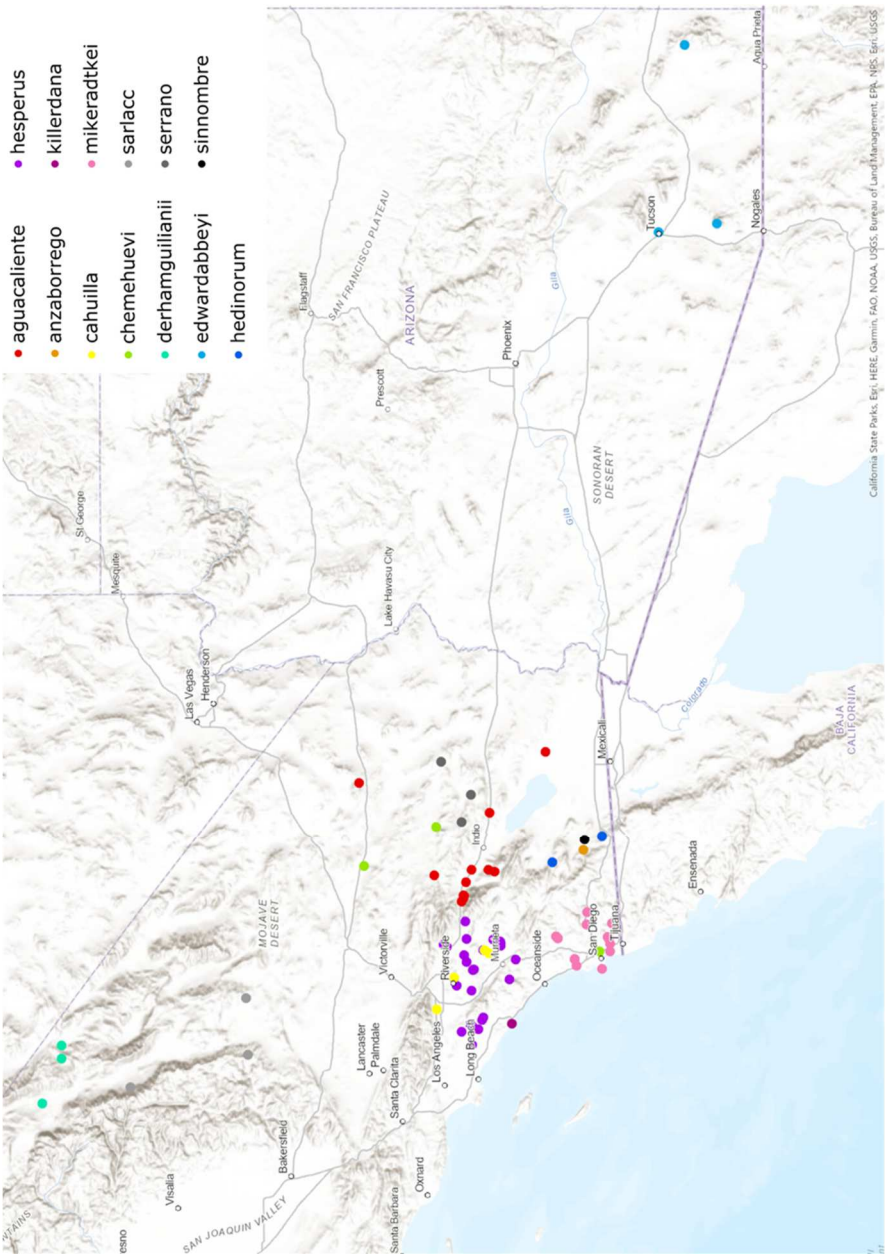


Figure 26. Distribution map of the Sierra and Simus species groups. See legend in the top right corner for color designation of species, with Simus group species in the left column and Sierra group species in the right column.

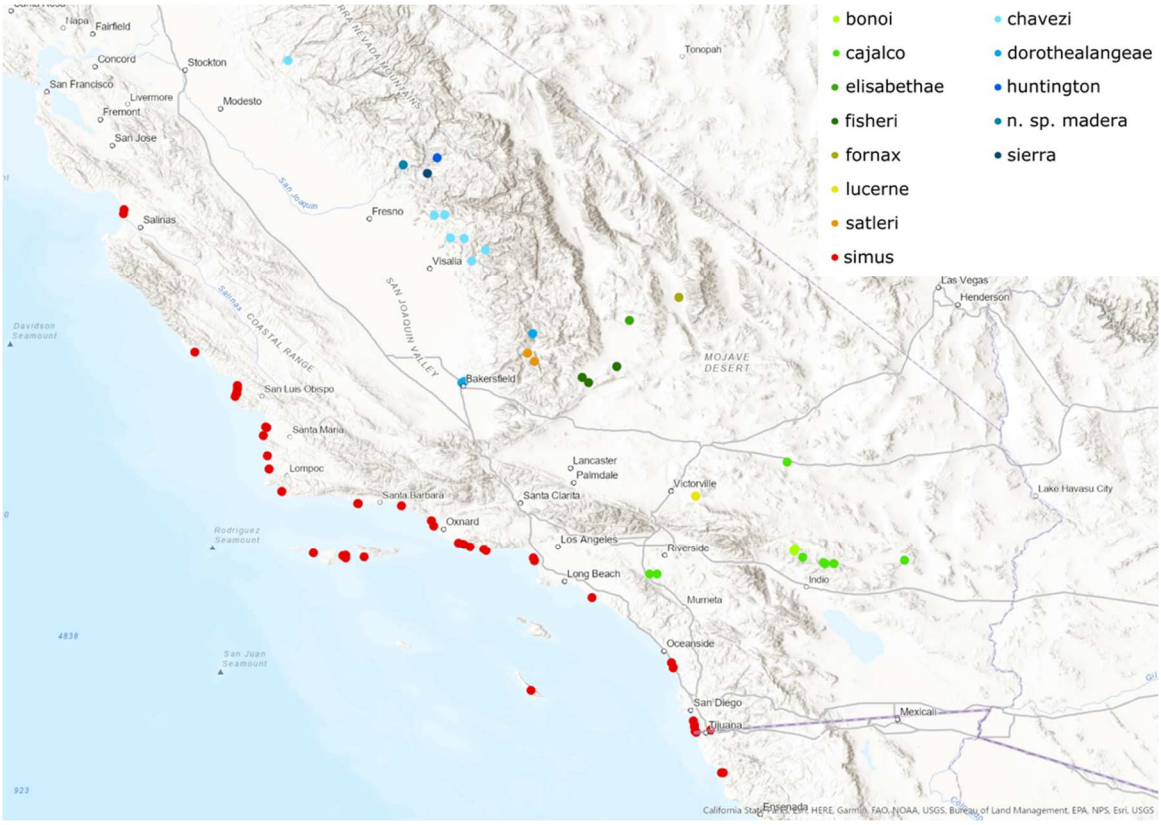


Figure 27. Photos of *Aptostichus* species representatives, habitat diversity, and a burrow. A) Photograph of one species representative from each species group, listed from top to bottom: *A. atomarius*, *A. aguacaliente*, *A. chavezi*, and *A. simus*. B) Photos representing the major habitat types of *Aptostichus* species: desert (top left), coastal dunes (top right), chaparral (bottom left), and alpine (bottom right). C) Photos of an *A. icenoglei* burrow, both closed (right) and open (left).

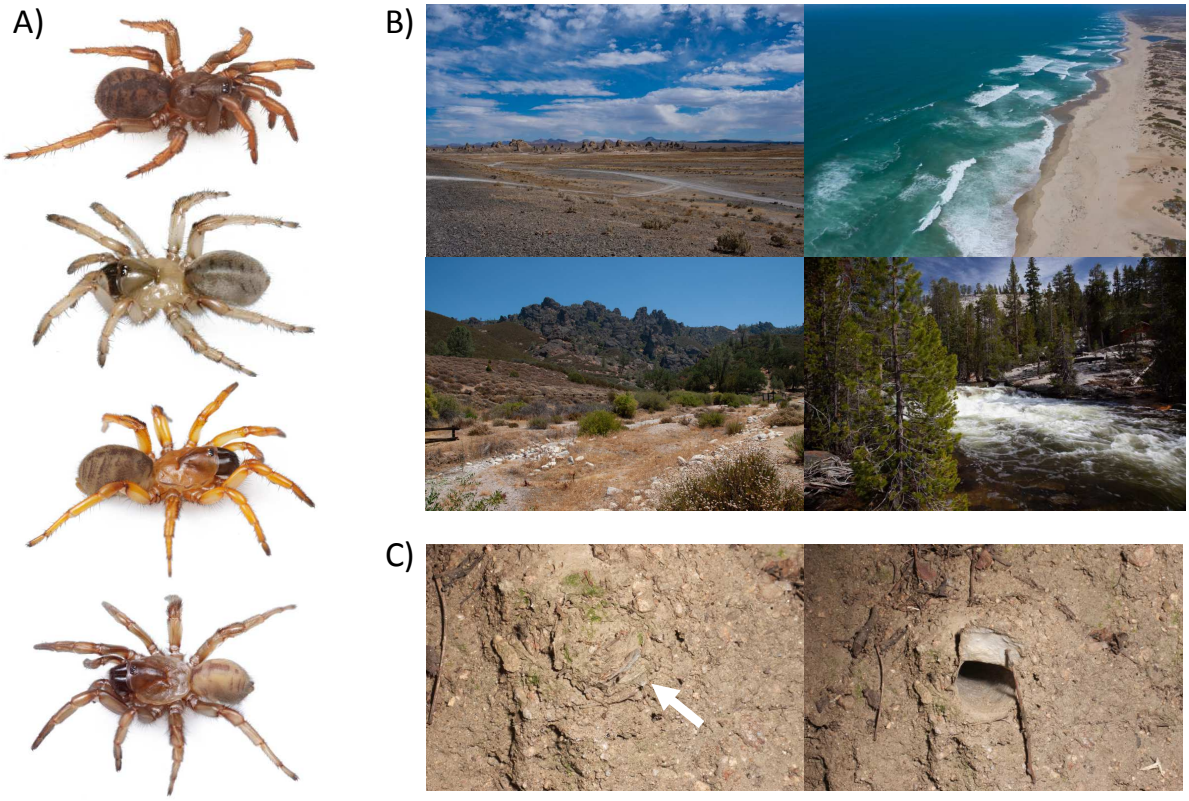


Figure 28. Phylogeny constructed using morphological data only. Black nodes denote bootstrap values greater than 70.

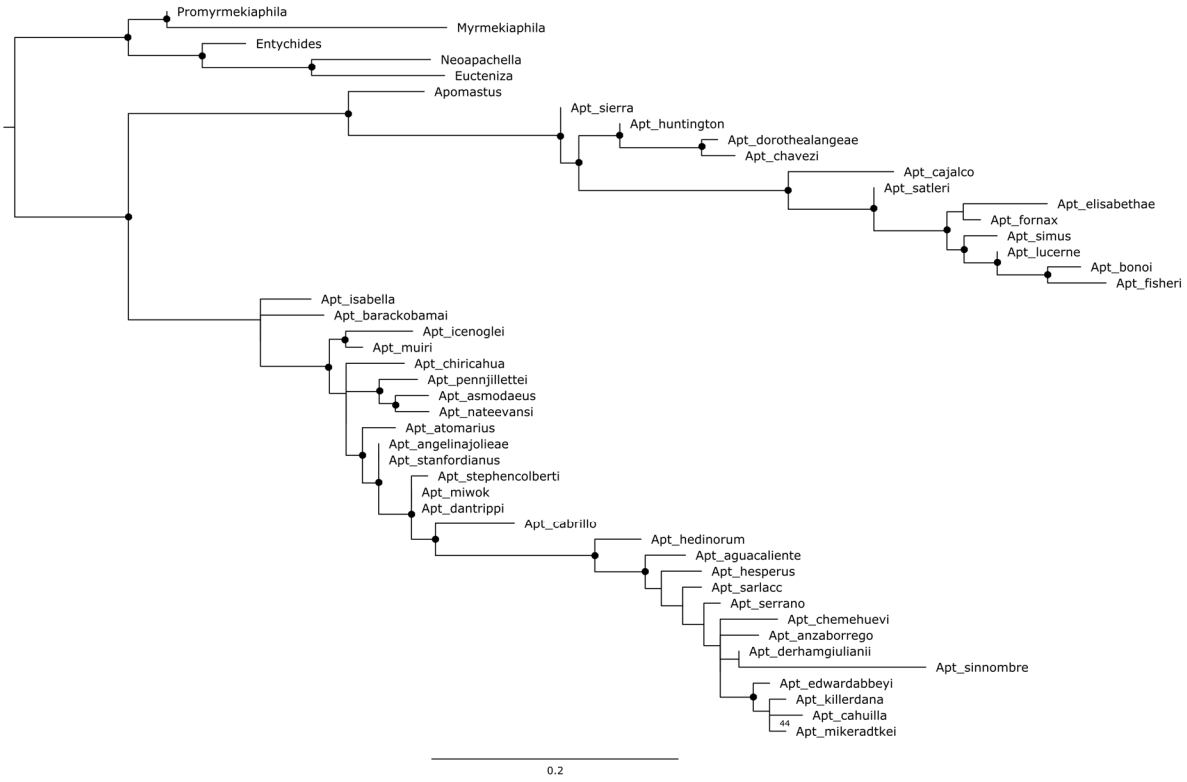


Figure 29. The preferred tree topology constructed from morphological and UCE data (both NC and MUS samples). Bootstrap values greater than 90 are denoted by black boxes and less than 75 denoted by gray boxes.

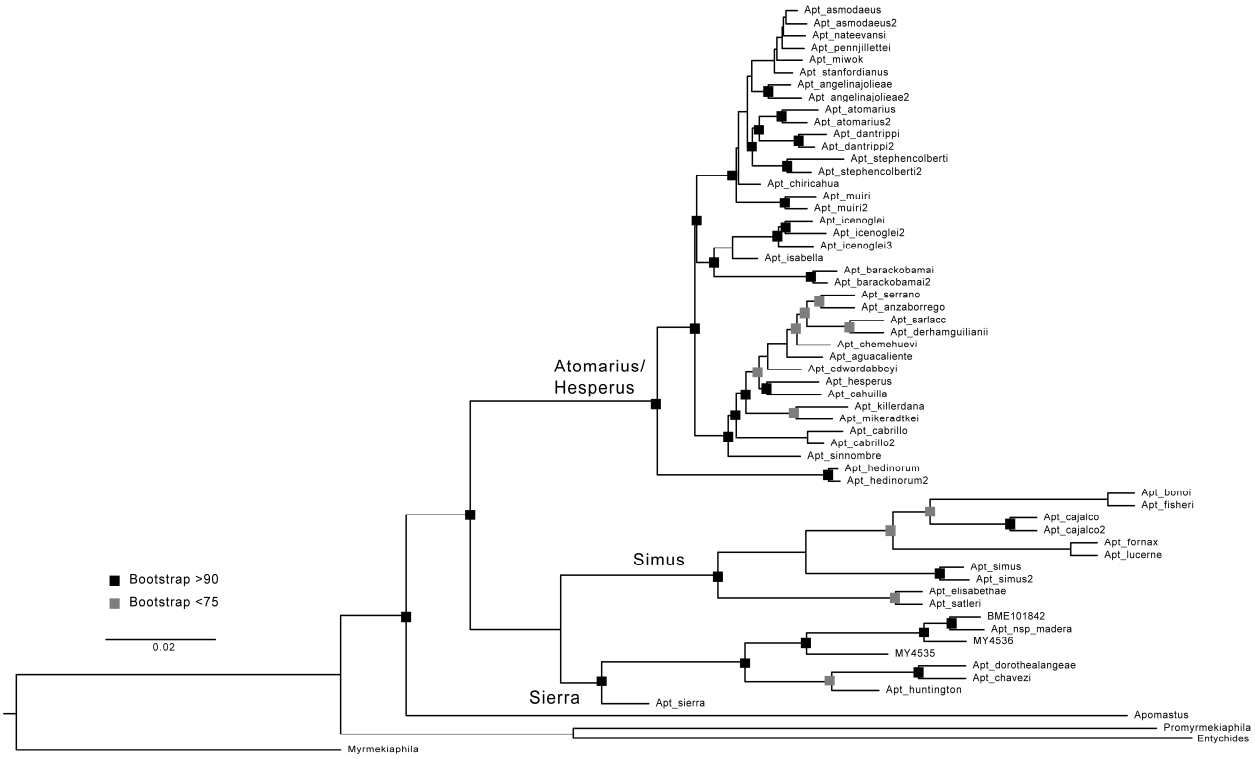


Figure 30. Raw results of the UA50p phylogeny.

Figure 31. Raw results of the UA50p_Con phylogeny.

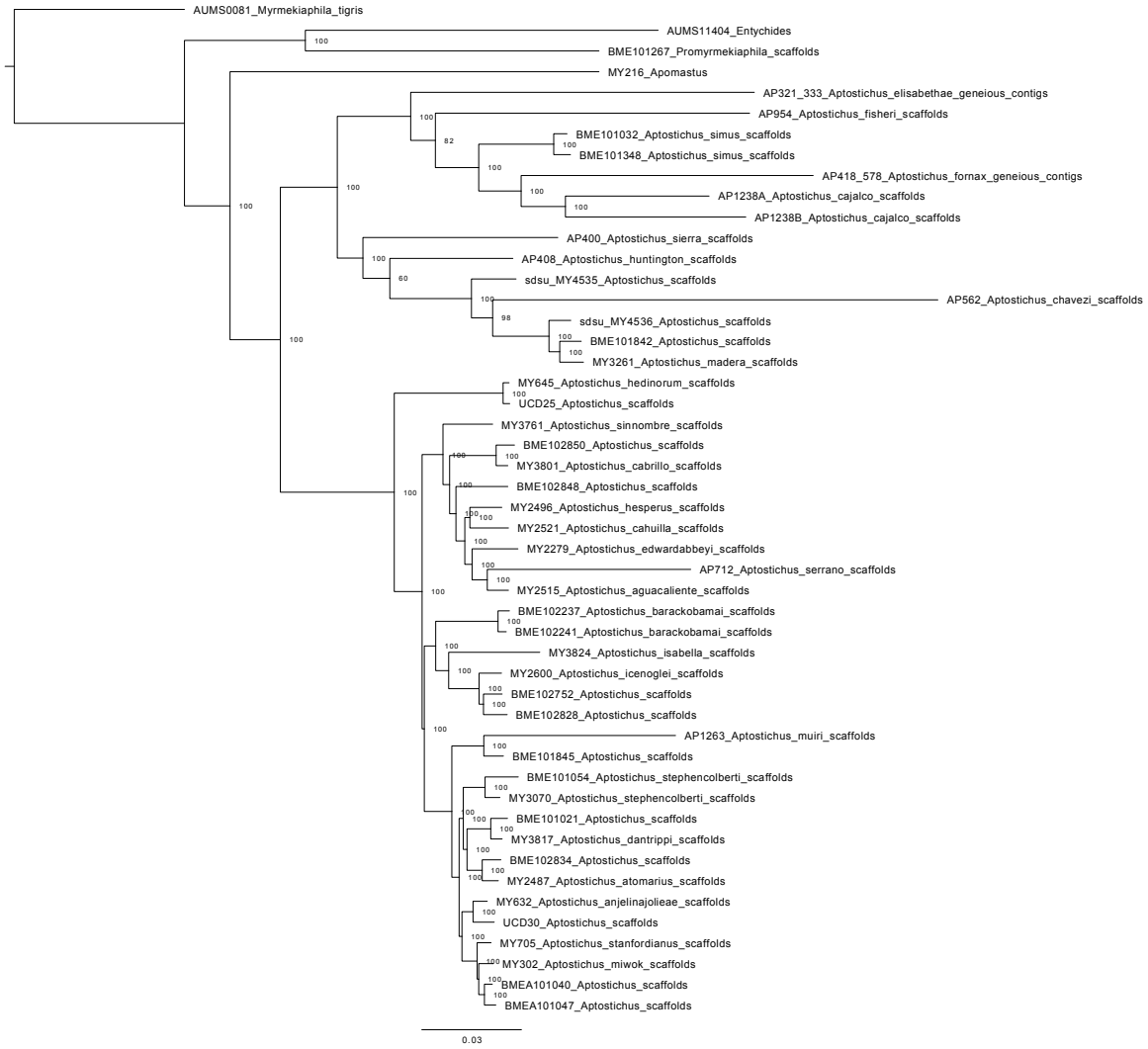


Figure 32. Raw results of the UA75p phylogeny.

Figure 33. Raw results of the UA75p_Con phylogeny.

Figure 34. Raw results of the UN50p phylogeny.

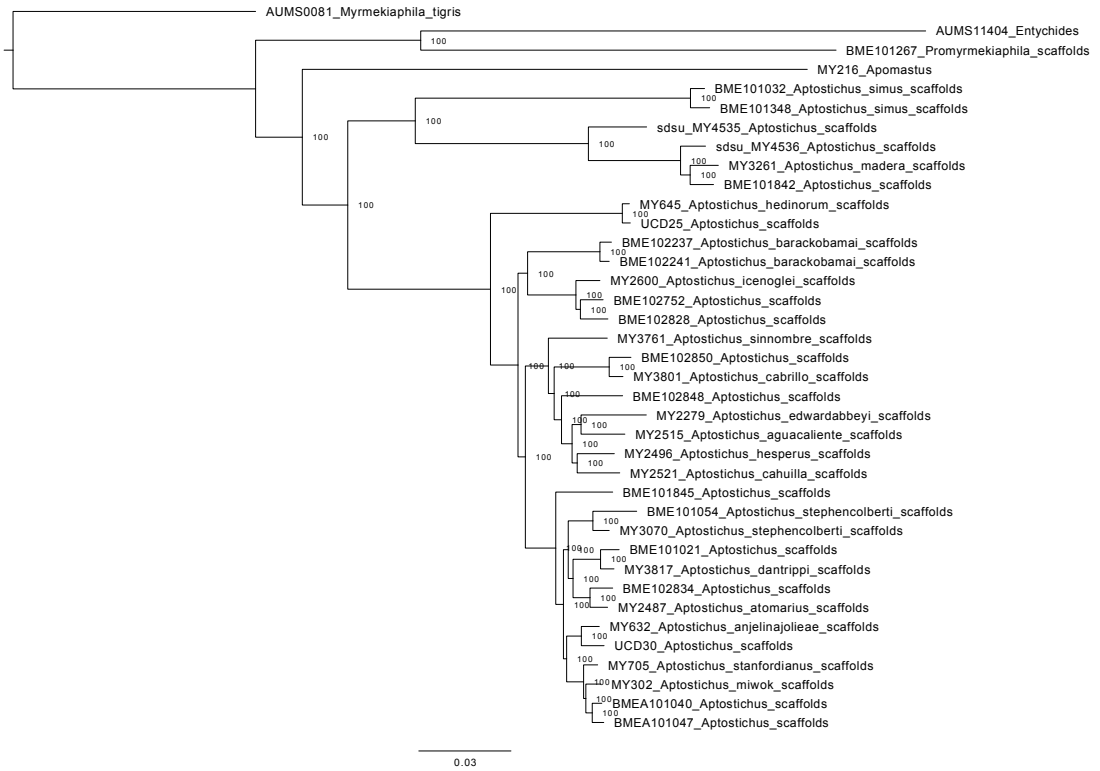


Figure 35. Raw results of the UN75p phylogeny.

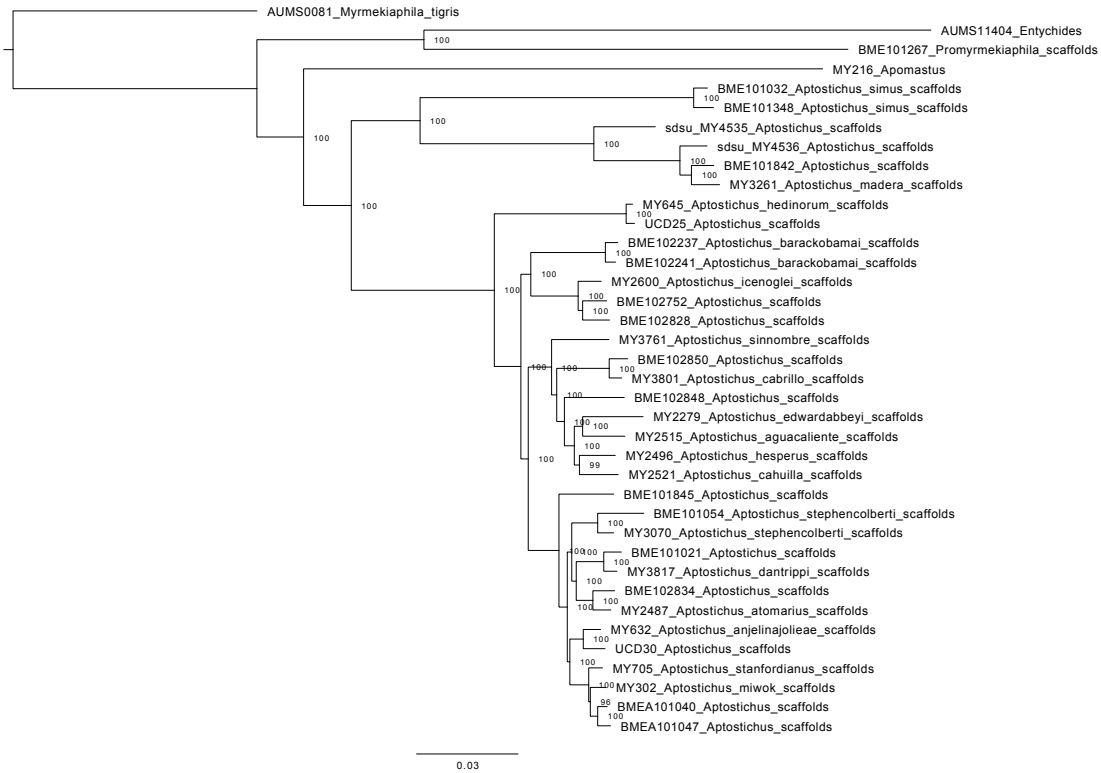


Figure 36. Raw results of the UA50p_MA phylogeny.

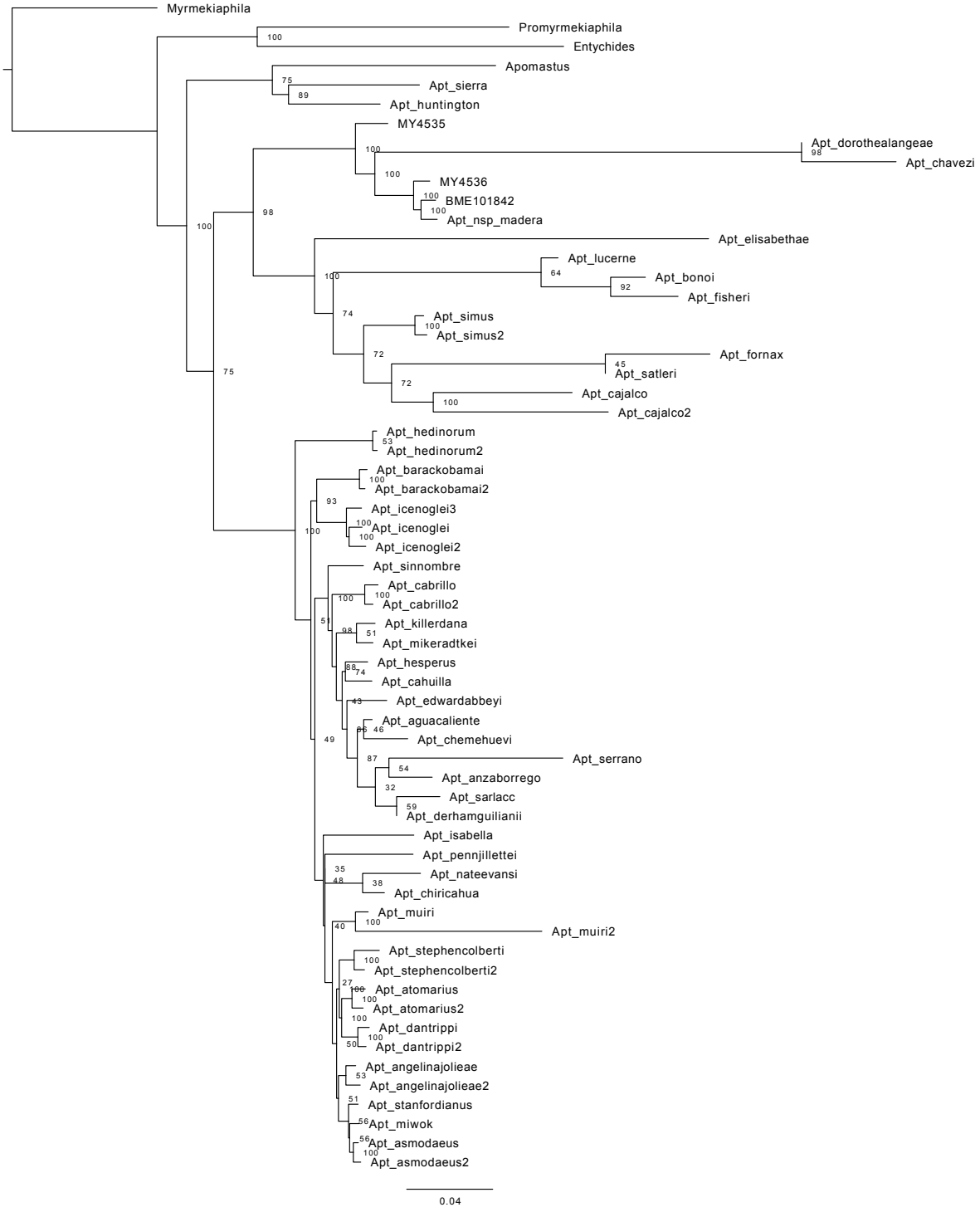


Figure 37. Raw results of the UA75p_MA phylogeny.

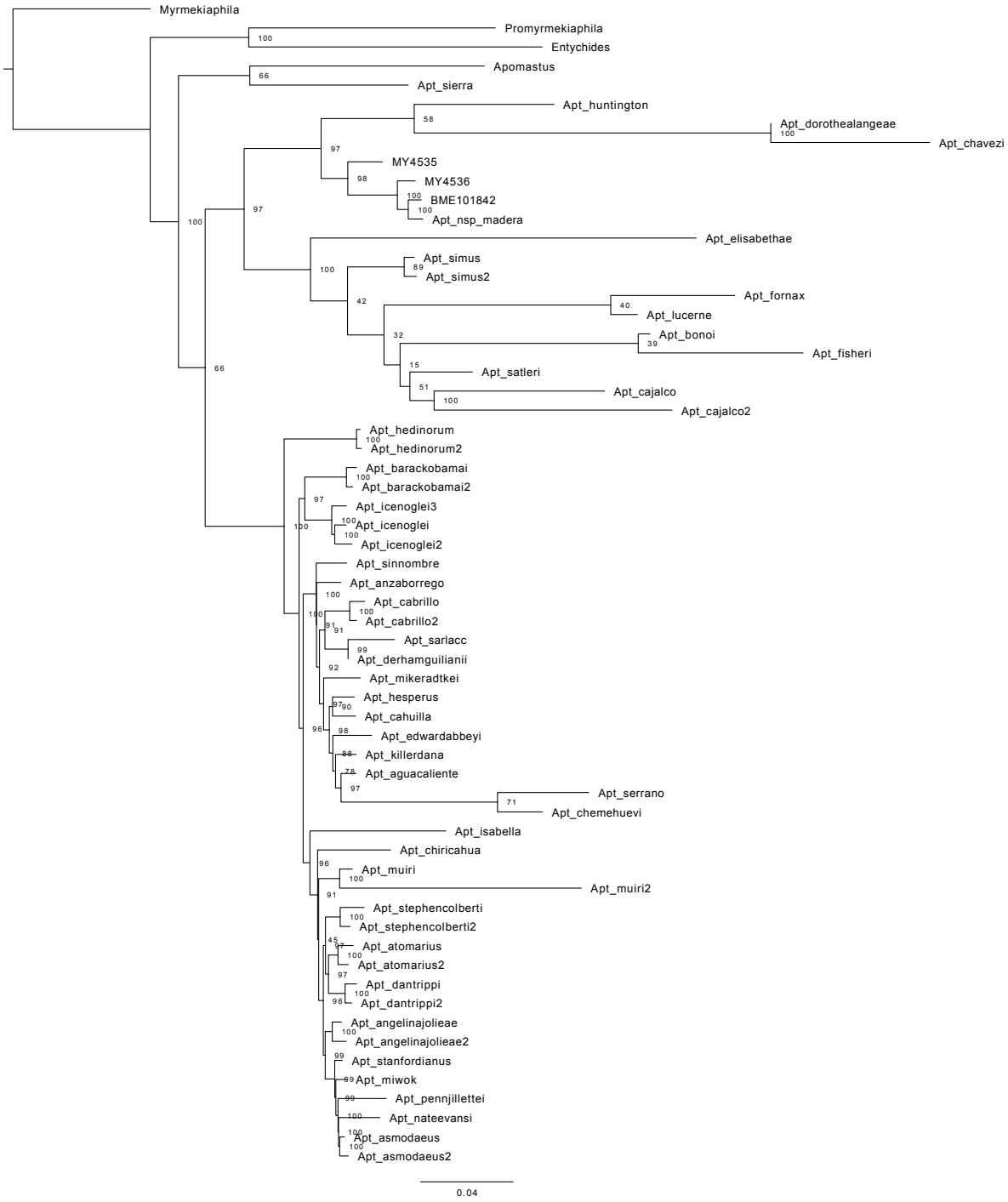


Figure 38. Raw results of the UN_MA phylogeny.

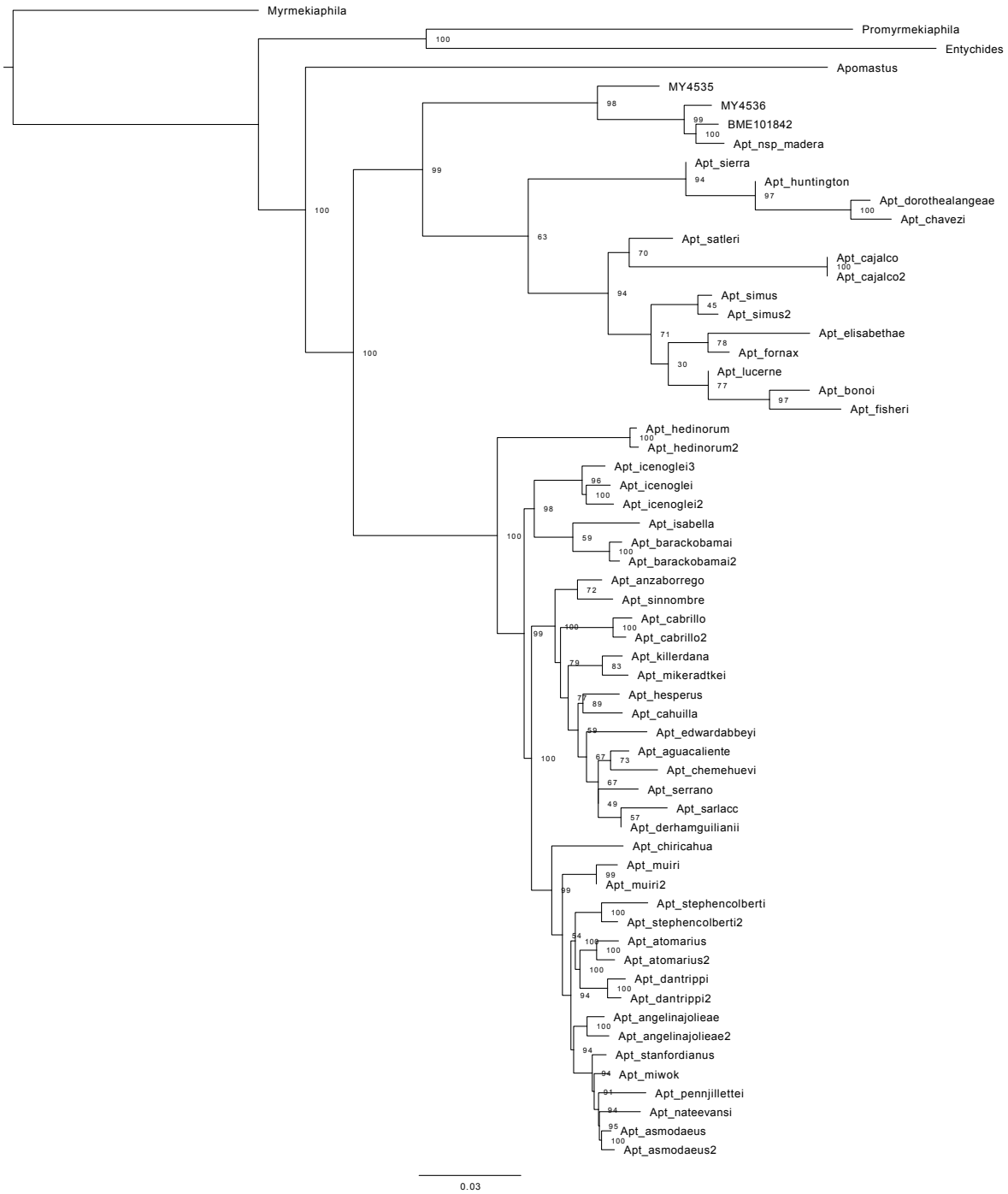


Figure 39. Raw results of the UA50p_MA_Con phylogeny.

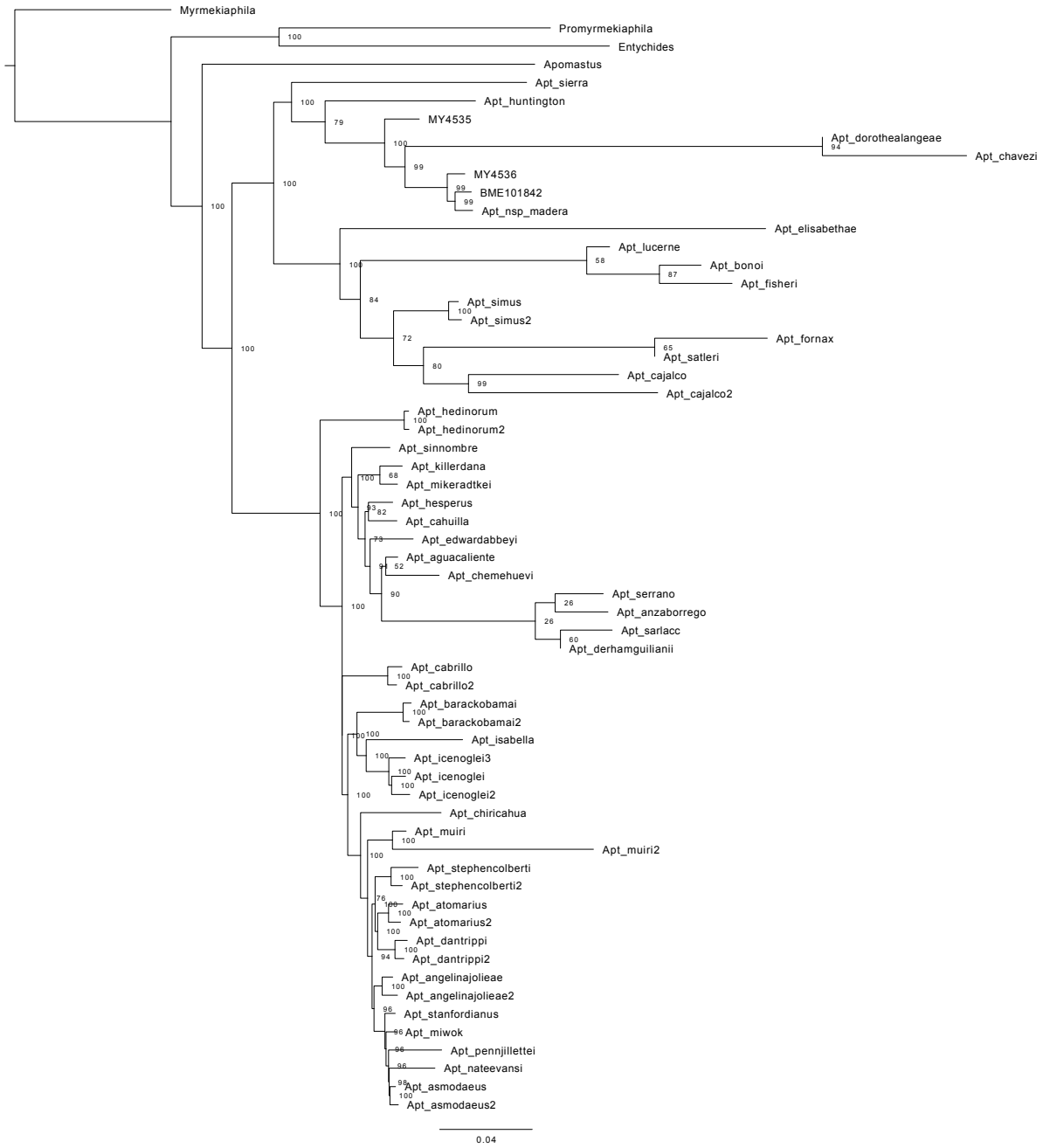


Figure 40. Raw results of the UN_MA_Con phylogeny.

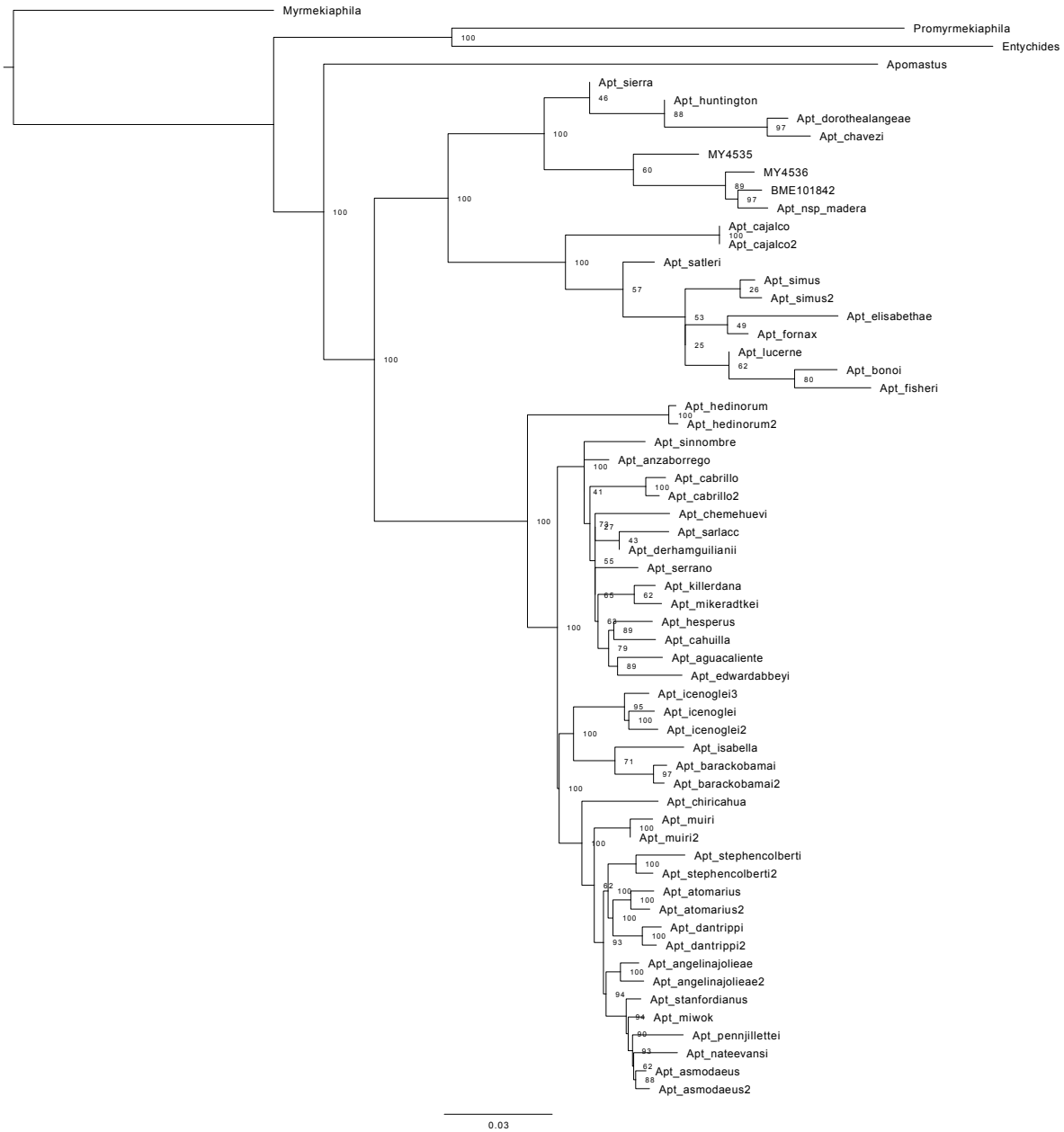


Figure 41. Raw results of the UA50p_ME phylogeny.

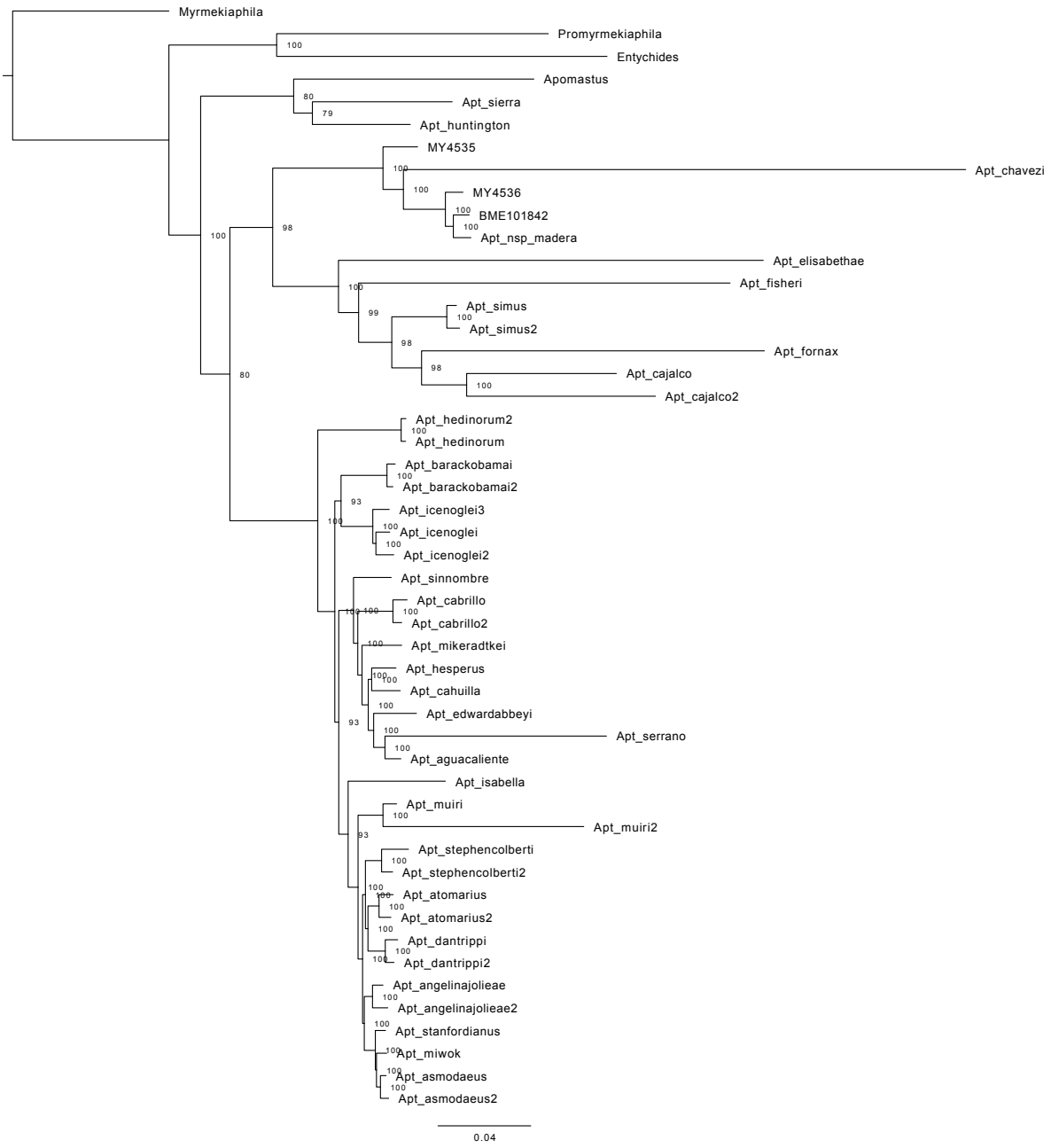


Figure 42. Raw results of the UA75p_ME phylogeny.

Figure 43. Raw results of the UN_ME phylogeny.

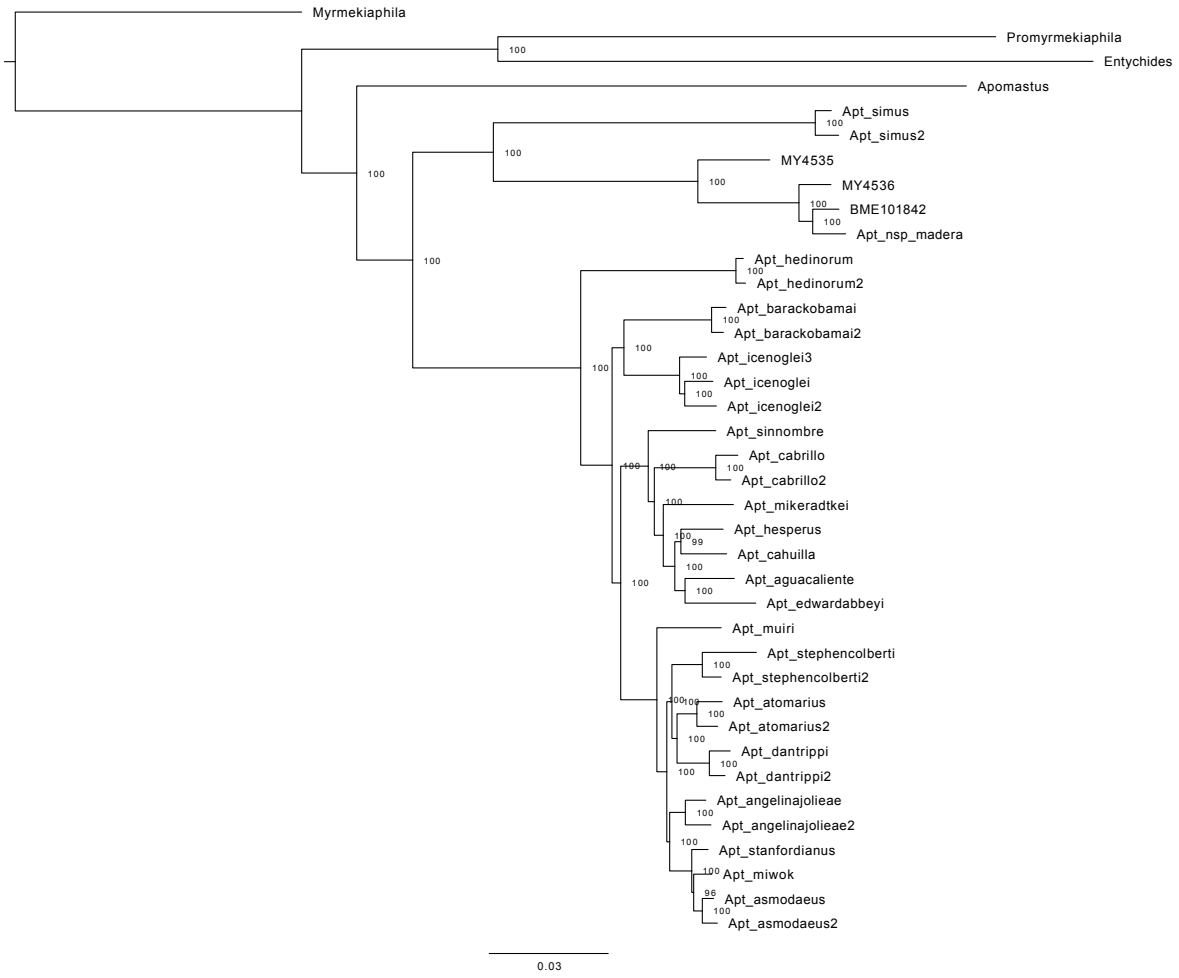
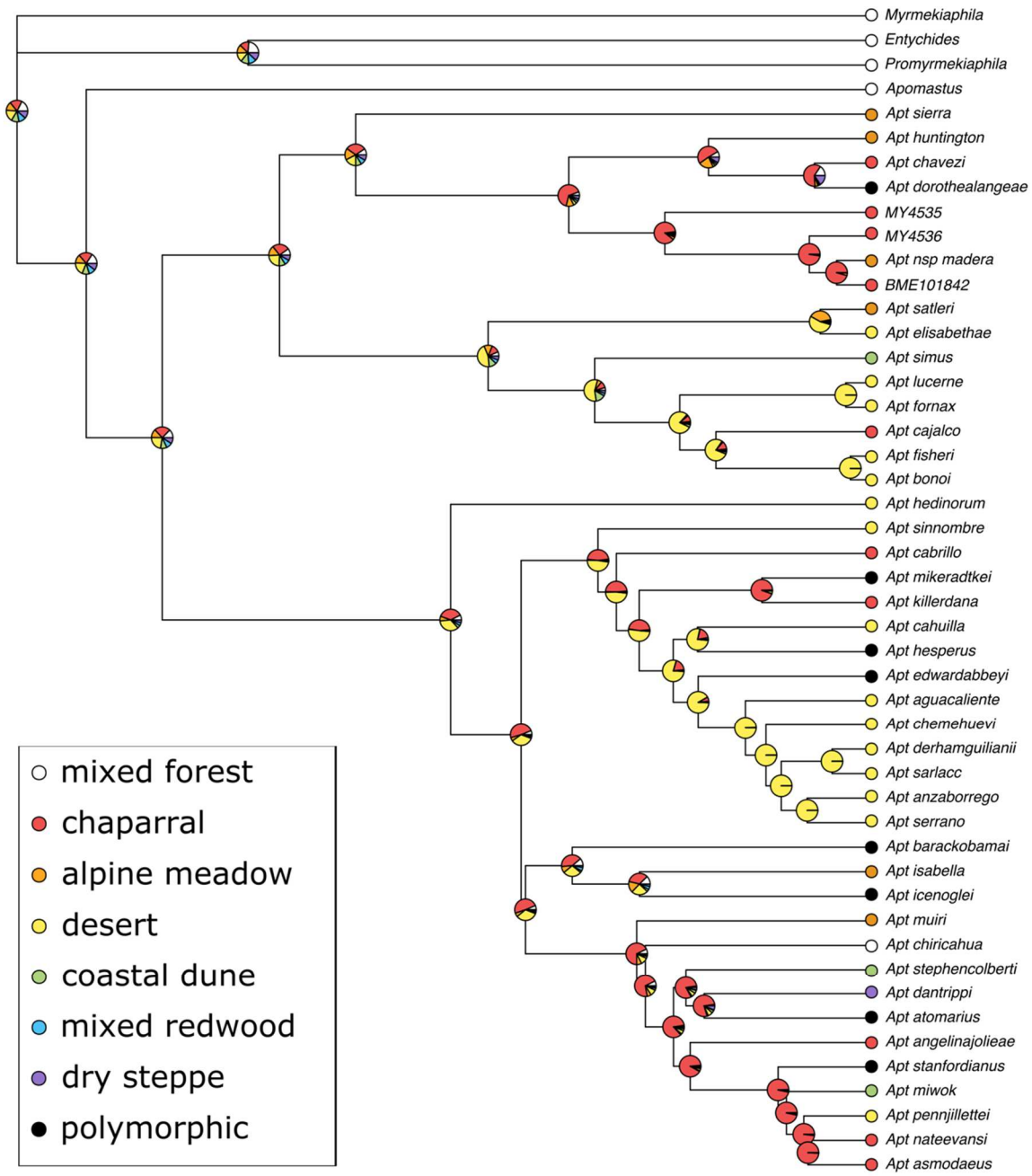


Figure 44. Ancestral state reconstruction of habitat type. See legend in the bottom left corner for habitat type color designations.


Myrmekiaphila
Entychides
Promyrmekiaphila
Apomastus
Apt sierra
Apt huntington
Apt chavezi
Apt dorothealangeae
MY4535
MY4536
Apt nsp madera
BME101842
Apt satleri
Apt elisabethae
Apt simus
Apt lucerne
Apt fornax
Apt cajalco
Apt fisheri
Apt bonoi
Apt hedinorum
Apt sinnombre
Apt cabrillo
Apt mikeradtkei
Apt killerdana
Apt cahuilla
Apt hesperus
Apt edwardabbeyi
Apt aguacaliente
Apt chemehuevi
Apt derhamguilianii
Apt sarlacc
Apt anzaborrego
Apt serrano
Apt barackobamai
Apt isabella
Apt icenoglei
Apt muiri
Apt chiricahua
Apt stephencolberti
Apt dantrippi
Apt atomarius
Apt angelinajolieae
Apt stanfordianus
Apt miwok
Apt pennjillettei
Apt nateevansi
Apt asmodaeus
mixed forest
chaparral
alpine meadow
desert
coastal dune
mixed redwood
dry steppe
polymorphic