2015-07-01

Delimiting Species and Varieties of Cycladenia humilis (Apocynaceae)

Holly Kathryn Brabazon
Brigham Young University - Provo

Delimiting Species and Varieties of *Cycladenia humilis* (Apocynaceae)

Holly Kathryn Waddel Brabazon

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Leigh A. Johnson, Chair
Robert L. Johnson
Duke S. Rogers

Department of Biology

Brigham Young University

July 2015

ABSTRACT

Delimiting Species and Varieties of *Cycladenia humilis* (Apocynaceae)

Holly Kathryn Waddel Brabazon
Department of Biology, BYU
Master of Science

Taxonomic delimitation of rare species is vital for accurate assessments of diversity and for their conservation. *Cycladenia humilis*, the sole species of *Cycladenia*, is an enigmatic perennial widely dispersed across the western United States. Within this species there are three currently recognized varieties: *C. humilis* var. *humilis* in Northern California, *C. humilis* var. *venusta* in Southern California, and *C. humilis* var. *jonesii* in Utah and Northern Arizona. Some populations occur geographically in areas between the typical distribution of each variety and the presently accepted taxonomy inadequately addresses these populations. Using five nDNA regions, we seek to clarify relationships between current varieties and assess the pattern of variation throughout the species. Analyses including K-means clustering, principle component analysis, fields for recombination, AMOVA, and ecological niche modeling were applied. Results indicate significant genetic structure between varieties and supports recognition of *C. jonesii* at the species level as distinct from *C. humilis*. Well defined intraspecific groupings are evident in the data, with evidence supporting the recognition of an additional variety in *C. humilis*, and two varieties in *C. jonesii*. Haplotype diversity and relationships between metapopulation clusters inform conservation efforts regarding diversity within *Cycladenia* and offer insights into the historical demography of this genus.

Keywords: clonal, rare plant, ecological niche modeling, fields for recombination, FFR, species delimitation, varieties, Western United States

ACKNOWLEDGEMENTS

I would like to thank my advisor, Leigh A. Johnson, for all the help he provided throughout my degree. I appreciate his willingness to help with everything from data collection to revision. I express appreciation to Duke S. Rogers and Robert L. Johnson for serving on my committee. Contributions from a Challenge Cost Share Agreement between Brigham Young University and the USDI Bureau of Land Management Utah State Office, from the Welsh Herbarium, and from the Brigham Young University Biology Department are much appreciated. Jared Brabazon provided extensive assistance with ArcGIS, final illustrations, and with thesis revision. I appreciate his patient love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Biological diversity is a valuable natural resource that can be characterized as the richness and diversity of the organisms at any geographic scale (Hawksworth, 1995). Assessing the abundance of species, or the organismal diversity, is necessary in understanding this diversity. Within species, varieties provide a more precise representation of the genetic diversity (Norse et al., 1986). Varieties represent more or less distinct lineages within a species (Templeton, 1998), can be intermediates in the process of speciation, and are often recognized by unique traits bound within a geographic location (O'Brien and Mayr, 1991). Varieties are not reproductively isolated from other populations of the same species, but may be developing geographically localized differences that create potential to become independent species in the future. Accurate recognition of varieties by their unique morphological or molecular traits provides greater specificity in conservation efforts designed to preserve diversity.

Although varieties can be distinguished by unique traits, their ability to introgress results in these unique traits often intergrading at the geographic boundaries between varieties. In a common, widespread, and geographically contiguous species, gradation of haplotypes or morphology between varieties is often easy to observe. Some species, however, have fragmented distributions with geographic gaps between varieties and, as a result, may lack apparent clues to distinguish whether variation is indicative of species or varieties. Despite this difficulty in delimiting taxonomic boundaries at a rank most appropriate to their biology, it is important to accurately identify and catalog the range of diversity across populations of such species, particularly if they are of conservation concern. Beyond interests in better managing the diversity within such species, these rare or uncommon, locally endemic species may be distantly related in comparison to other species with which they co-occur, which considerably increases

the breadth of the biological diversity within an area (Hawksworth, 1995). *Cycladenia humilis* Benth. is an uncommon plant species endemic to the Western United States. It has disjunct populations that are grouped into geographically determined varieties that encompass morphological patterns (Rosatti, 2012). These traits provide insight to methods suitable for assessing and delimiting taxonomic boundaries that reflect biological processes of divergence.

As presently understood, *Cycladenia humilis* is the sole species of the genus *Cycladenia* (Rosatti, 2012). Three varieties are currently recognized within the species, which are vars. *humilis* (Eastw.) Woodson ex. Munz, *venusta* (Eastw.) Woodson ex. Munz, and *jonesii* (Eastw.) S.L. Welsh & N.D. Atwood. Each variety possesses some differentiating morphological features with var. *humilis* having a glabrous or tomentose perianth abaxially and corolla glabrous to papillate adaxially, and vars. *venusta* and *jonesii* having a perianth more or less lacking interwoven hairs abaxially and corolla sparsely hairy adaxially. Var. *venusta* has corolla lobes 8-12 mm and var. *jonesii* has corolla lobes 4-5 mm (Last, 2009; Rosatti, 2012). These characteristics distinguish the varieties, but identification is generally determined based on geographical location (Rosatti, 2012), namely Northern California (var. *humilis*), Southern California (var. *venusta*), and Utah/Arizona (var. *jonesii*). The correspondence of morphological variation to geography is not perfectly congruent, however. Based on morphological characters, taxonomists place populations found in Western Central California into either var. *humilis*, or var. *venusta*, with populations from Eastern California being placed into var. *jonesii* (Fig. 1) (Rosatti, 2012). The taxonomic placement of the Eastern California populations in var. *jonesii* has consequences for conservation given that var. *jonesii* is federally listed as threatened; the other varieties are not federally or state listed (U. S. Fish and Wildlife Service, 1986).

Some earlier taxonomic treatments also segregated densely tomentose plants from northern California into var. *tomentosa* (A. Gray) A. Gray (e.g., Munz, 1959). Var. *tomentosa* is now regarded as part of var. *humilis,* with the presence of dense tomentum presumed to be a one-gene trait (Sipes and Wolf, 1997). Of historical note, vars. *tomentosa, jonesii*, and *venusta* have each also been recognized as separate species (Gray, 1876; Eastwood, 1902, 1942).

Geographically, *Cycladenia humilis* has a highly fragmented distribution, typically at higher elevation sites of the Desert Mountain, Southwestern, Central Western, and Northwestern floristic regions of California. It is also found in scattered sites within the Colorado Plateau floristic province of Utah and Arizona. Its nearest relatives, *Trachelospermum* and *Pinochia*, are distant both geographically (Southeastern US and Caribbean), and phylogenetically (Livshultz, 2007). *Cycladenia humilis* is likely a paleoendemic persisting in conditions unfavorable to many other species (Sipes and Wolf, 1997).

Although all *Cycladenia* populations are found in dry, gravelly to talus soils with sparse vegetation, microhabitats vary across the species range. California populations are found on high mountain slopes (Rosatti, 2012), while plants in Utah and Arizona are found predominantly in talus alongside mesas (Welsh, 1987). Reports of germination in native habitat are lacking, and sexual reproduction is infrequent or episodic (Sipes and Wolf, 1997; Pence, 2014). Due to unobserved seed germination and unknown reproduction patterns, *Cycladenia* apparently maintains populations through asexual, clonal reproduction. The combination of disjunct habitat and predominantly clonal growth make assessing the diversity within *Cycladenia* difficult.

Little information is known about population relationships among the entire species. It is questionable if the present 'one species with three varieties' treatment of *Cycladenia* best represents the levels of divergence and differentiation among lineages in this taxon and if the

hypothesis of paleoendemism is substantiated by genetic data. The only previous genetic study focused on diversity within var. *jonesii* using allozymes, and showed genetic structure and diversity within var. *jonesii* and differentiation from var. *humilis* based on one population from California used as an outgroup (Sipes and Wolf, 1997). The entire diversity of the species beyond var. *jonesii* has not been studied in-depth, leaving questions about relationships and taxonomy unanswered.

Here we test the hypothesis implicit in the current taxonomy of *Cycladenia*. Specifically, we address the following four hypotheses: 1) *Cycladenia* is comprised of one species encompassing all of California, Utah, and Arizona populations; 2) three varieties exist within *C. humilis*, including vars. *humilis, venusta,* and *jonesii*; *3*) Central California populations contain individuals from both var. *humilis* and var. *venusta*; and 4) Eastern California populations are part of var. *jonesii*. Using the General Lineage Concept (de Queiroz, 1998) as the framework for recognizing species, we sought distinctions between metapopulation clusters using multiple secondary criteria. Namely, we sought evidence of reproductive isolation and ecological distinction associated with morphological differences. We provide resources for conservation by comparing the diversity of haplotypes found in metapopulation clusters and providing more resolution on population relationships.

MATERIALS AND METHODS

We sought to recover metapopulation clusters with five low-copy nuclear markers using Principle Components Analysis (PCA), K-means Clustering (Meirmans, 2012), and fields for recombination (FFR) (Doyle, 1995; Flot et al., 2010). To validate these metapopulation clusters (Carstens et al., 2013), we performed Ecological Niche Modeling to determine if clusters exhibited ecological divergence. AMOVA in Arlequin 3.5.2.1 (Excoffier and Lischer, 2010) was used to assess the percent variation between the clusters discovered by other methods. Haplotype networks used in FFR were colored based on geographic groupings. Recombination detection analyses (RDP4) (Martin, 2015) were used to ascertain whether recombination events disrupted evolutionary signals (Templeton et al., 2000).

Sampling

Leaf tissue was collected from 26 populations covering the entire range of *Cycladenia* and encompassing nine populations of var. *humilis*, seven populations of var. *venusta* and ten populations of var. *jonesii* (Fig. 1) (Last, 2009). Voucher specimens were deposited at Brigham Young University (BRY). For each collection, leaves were sampled from ramets greater than 8–10 meters apart to reduce redundant sampling of genets (Sipes and Wolf, 1997). After collection, leaves were dried and stored in silica gel with a total of 363 individuals sampled (Appendix I). Sixteen samples were collected for each population, or in the case of populations with fewer than 16 genets, the entire population was sampled. The outgroup, *Trachelospermum difforme* (Walter) A. Gray, was determined based on phylogenetic results from Livshults et al. (2007).

DNA extractions, amplification, and sequencing

DNA extractions were performed using a modified CTAB protocol (Doyle and Doyle, 1987; Cullings, 1992). Low-copy nuclear regions were determined using primers published by

Straub et al. (2011), and by selecting regions from among their COS II loci that showed variation when sequenced across a geographically diverse sampling of *C. humilis* individuals (Appendix II). PCR cycles followed a touchdown profile decreasing in increments of 1°C with each cycle from 62°C to 52°C, with the last 20 cycles reaching 52°C. Reactions were held at 72°C for 8 minutes to ensure the addition of adenosine overhangs for subsequent TA cloning, if needed.

After amplification, PCR products were cleaned on 96-well PCR purification plates (USB Corp., Cleveland, OH) with both forward and reverse strands cycle sequenced using BigDye v.3 (Applied Biosystems, Foster City, CA). Sequences were run on the automated AB 3730x1 sequencer in Brigham Young University's DNA Sequencing Center.

Sequence alignment, phased sequence generation, and cloning

Sequences were edited based on their chromatograms in Sequencher 5.0.1 (Gene Codes Corp., Ann Arbor, MI). Due to the conserved nature of the sequences, it was possible to align by eye using Se-Al (Rambaut, 1996). Sequences that contained heterozygous sites, including indels, were manually phased into alleles based on conserved sequences in the population. If separation of heterozygotes was questionable, sequences were cloned using Invitrogen TA TOPO kits (Life Technologies, Carlsbad, CA) and plated on Ampicillin. Transformants were amplified and sequenced using M13 primers, and subsequently edited as above. Because some variation among colonies from a single cloning reaction may be attributable to PCR artifacts in the original amplicon, the phase of heterozygous sites were sometimes determined based on the consensus of sequences obtained, particularly in reference to the original sequencing results. All sequences will be deposited in GenBank prior to the formal publication of these data.

Recombination Detection

Recombination detection was performed in RDP4 using linear sequences and default settings for RDP (Martin and Rybicki, 2000), Chimaera (Posada and Crandall, 2001), BootScan (Martin et al., 2005), 3Seq (Boni et al., 2007), GENECONV (Padidam et al., 1999), MaxChi (Maynard Smith, 1992), and SiScan (Gibbs et al., 2000). The single gene that tested positive for recombination was further assessed to determine whether recombined sequences affected the structure of the phylogenetic signal by constructing four datasets. One was left unmodified, the second had recombined sequences removed, the third contained the 5' end of the sequence, and the fourth contained the 3' end of the sequence (Templeton et al., 2000). Each of the three modified datasets was re-evaluated with RDP4 and determined to contain no recombination. Each was then compared against each other and the unmodified dataset to determine the consistency between likelihood trees generated in IQ-TREE after BIC model selection within the same program (Nguyen et al., 2015) and haplotype networks generated in PopART (http://popart.otago.ac.nz).

K-means clustering and Principle Components Analysis

Nexus sequence files were imported into GenAlEx (Peakall and Smouse, 2006, 2012) to create data files representing diploid haplotypes. The haplotypes were then analyzed in GenoDive 2.0b27 (Meirmans and Van Tienderen, 2004) for K-Means Clustering, and Principal Component Analysis (PCA). For K-means clustering, parameters included simulated annealing with 500,000 steps and 20 random starts. Results were consistent over multiple runs. The best clustering was determined according to Calinski & Harabasz' (1974) pseudo-F, although each value of K was inspected for patterns. The PCA was run between populations including a covariance matrix with 99 permutations. The first three PCA axes were plotted in Plot.ly. The

five loci were resampled with a jackknife procedure to assess the level of support provided in K-means clustering and PCA data.

Haplotype Network and FFR Construction

To create a haplotype network with the TCS algorithm (Clement et al., 2000, 2002), aligned sequences were imported into PopART with coordinates designated for the 26 populations. To code the few gaps encountered in our alignments, SeqState (Muller, 2005) was used with simple coding (Simmons and Ochoterena, 2000). Coded gaps were then converted to base pairs and appended to the end of the matrices. To see if informative bases were excluded within the gaps, polymorphic loci within gaps were mapped onto haplotypes in PopART. Due to the minimal impact of polymorphic sites within gaps, these variable sites were disregarded. Consequently, all gaps were included in the final matrices as additional characters.

We created fields for recombination, or FFRs, by first generating a haplotype network of phased nuclear DNA data. This network shows mutational differences between individual alleles. Heterozygotes are a combination of two haplotypes, and in the FFR are represented as a line connecting the haplotypes. Two haplotypes within a heterozygote have potential to recombine, hence fields for recombination. The heterozygote connections form webs that imply recent genetic exchange. Partitions occur where heterozygotes do not connect haplotypes, and identical partitions in many loci indicate speciation, even when speciation is not indicated by reciprocal monophyly (Flot et al., 2010; Doyle, 1995).

Using Adobe Illustrator, connections between haplotypes shared in heterozygotes were illustrated to create a haploweb (Flot et al., 2010). The two connected haplotypes imply possibility of recombination during meiosis and the resulting haploweb designates fields for recombination or FFR (Doyle, 1995).

AMOVA

Arlequin 3.5.2.1 (Excoffier and Lischer, 2010) was used to perform a standard AMOVA (Excoffier et al., 1992) among groups, among populations within groups, among individuals within populations, and within individuals. 10,000 permutations were used. Based on results from other analyses, particularly PCA and FFR, five groups were defined at the highest hierarchical level. This consisted of Utah populations, the single Arizona population, Southern and Central California populations, Eastern California populations, and Northern California populations.

Ecological Niche Divergence

In addition to the 26 populations we sampled in this study, 14 populations from Utah, and 11 populations from California were incorporated into the analysis (Appendix I). Sample locations were determined on site via GPS. Using SDMtoolbox (Brown, 2014) in ArcGIS (ESRI, 2011), 19 WorldClim variables and an altitude layer (30 arc-seconds, about 1 km^2) (Hijmans et al., 2005; http://www.worldclim.org) were clipped to sampling bounds. Sampling locations were rarefied to 5 km. Correlated WorldClim variables (R-value greater than 0.90) were reduced to avoid redundant data. An R-value greater than 0.90 indicated correlation between WorldClim variables, and correlated variables were removed to reduce redundancy. Rarefied sampling and uncorrelated variables were input into MaxEnt v 3.3.3k (Phillips et al., 2006). Default convergence threshold value, 100 bootstrap replicates (Warren et al., 2008), and a jackknife test were specified during the run. Based on previous analyses we tested the niche divergence between the California populations and the Utah/Arizona populations to determine if ecological divergence has occurred between the metapopulation clusters.

To test whether the habitat predictions made for Utah and California populations are significantly different, Hellinger's-based *I* similarity statistic (Warren et al., 2008), which ranges from 0-1, with 1 meaning exact niche overlap, was generated in R (R Core Team 2014) using the SDMTools v 1.1-221 library (VanDerWal, 2014). One hundred Bootstrap replicates with 1000 iterations were used to calculate a null distribution of the value. Area under the curve (AUC) of the receiver operating characteristic (ROC) plot indicates the goodness of fit for the ecological niche modeling for distinguishing presence data from background (due to lack of absence data) (Warren et al., 2008). A value above 0.90 indicates excellent fit (Swets, 1988; Elith, 2002)

Consideration of Tomentum

To determine whether the previously named var. *tomentosa* merits recognition, tomentose and glabrous individuals were coded as presence/absence. This coding was placed onto the haplotype networks generated for each of the 5 loci, with patterns of distribution for the tomentose phenotype assessed relative to haplotype designation.

RESULTS

Sampling consisted of 363 diploid individuals, generating a total of 726 phased sequences. After trimming, Nc1a was 692 bp, Nc3 was 612 bp, Nc4 was 262 bp, Nc7b was 1106 bp, and Nc10a was 909 bp in length. The presence or absence of gaps, regardless of length, was recoded as a single base pair for fields for recombination, K-means Clustering, AMOVA, and Principle Component Analysis. With gaps coded, Nc1a was 672 bp, Nc7 was 1080 bp, and Nc10a was 904 bp in length.

Recombination Detection

RDP4 reported no recombination in Nc1a, Nc3, Nc4, or Nc10a. Nc7b had several recombined sequences with breaks inferred to have occurred in an area of low sequence variability which ranges between 500 and 700 bp from the 5' end of the amplified sequence. The three modified datasets, created from non-recombined sequences only (107 sequences removed), 5' 600 bp, and the 3' 480 bp retained similar overall relationships between geographic areas when compared to the entire dataset (Figs. 2–5). The impact of recombination was thus determined to not interfere with the purpose of this study, and consequently, the entire Nc7b region was used as a single locus with recombinant sequences retained. The model for the split datasets and the dataset excluding recombinants was HKY+G. The model for the complete dataset was TPM3u+I+G.

K-means Clustering and Principle Components Analysis

Using K-means clustering to determine optimal clustering of populations into a predetermined number of groups (beginning with two and increasing upward) provides insight into how haplotype clusters are patterned throughout the data. The simplest grouping (K = 2) resulted in all of the California populations grouped as distinct from a group containing all of the

Utah and Arizona populations. At K = 3, California is further divided into two clusters, with Southern and Central populations grouping together and Northern and Eastern populations grouping together. At K = 4, Eastern California populations become distinct. K = 5 separates Central California populations from the southern California populations, and K = 6 separates the Arizona population from the Utah populations (Table 1). This pattern was consistent throughout the jackknife resampling.

A 3-dimensional representation of the first three components in the PCA (Fig. 6) captured 58.69% of the variance among populations, with the fourth, unrepresented component containing 8.75% of the variation (Table 2). Groups distinguishable in three dimensions show separation between Utah/Arizona, Eastern California, Southern California, Central California, and Northern California.

Haplotype Networks and Field for Recombination

In the Nc1a dataset (Table 3, Fig. 7) 70 haplotypes with 22 singletons (haplotypes only found in a single individual) were present. Forty-six polymorphic loci and an additional four indels were recovered. The most common haplotype was found in 76 individuals with 23 individuals homozygous for that haplotype. The haploweb indicating haplotypes shared in heterozygous individuals shows ten allele pools (*sensu* Doyle, 1995). Seven of those allele pools were between two singleton haplotypes and were subsequently encompassed in the allele pool comprised of individuals in geographic and genetic proximity. Individuals from Utah represented one allele pool, individuals from Arizona were completely homozygous in a second, and all individuals from California were contained in a third allele pool. The outgroup failed to amplify due to apparent primer divergence.

Twenty-four haplotypes were present in the Nc3 data set (Fig. 8) with 23 polymorphic loci. There were seven singletons with the most common haplotype existing in 111 individuals. Eighty individuals were homozygous for that haplotype. The haploweb showed four allele pools, with two of the haplotypes consisting of only homozygous individuals. Utah and Arizona populations were encompassed in one allele pool with Northern California and Central California in a second allele pool. The Eastern California populations and the Southern California populations were composed entirely of homozygous sequences resulting in two separate allele pools. The outgroup connected to a Northern California haplotype.

The Nc4 dataset (Fig. 9) comprised ten haplotypes with nine polymorphic loci. There were 174 individuals that had the most common haplotype, and 122 individuals were homozygous for that haplotype. The haplotypes that were connected by heterozygous individuals indicated two allele pools. Northern, Central and Southern California populations grouped together in one allele pool comprised of 193 individuals, while Utah, Arizona and Eastern California grouped together in the other allele pool of 170 individuals. The outgroup connected to a haplotype from Northern California.

In the Nc7b dataset (Fig. 10) there were 74 haplotypes with 87 polymorphic loci. Five indels were present. There were 26 singletons, and 61 individuals possessed the most common haplotype. Twelve individuals were homozygous for that haplotype. Five allele pools were recovered with one encompassing Utah, a second representing Central and Southern California, another encompassing Northern California, a fourth from Arizona, and a fifth comprised of just Eastern California. The outgroup connected into a haplotype from Northern California.

Nc10a (Fig. 11) contained 32 haplotypes with 33 polymorphic loci. Four gaps were coded to base pairs. There were 11 singletons, with 113 individuals having the most common

haplotype (97 were homozygous for that haplotype). One allele pool encompassed all of Utah and Arizona, the second encompassed Eastern California. Eastern California was entirely homozygous and created an allele pool independent of other samples. The outgroup failed to amplify.

AMOVA

Results from the AMOVA indicate high levels of variation within individuals due to heterozygosity (41.80% of variation). With five groups, populations within groups expressed less variation (17.17%) than among groups (34.29%). The lowest amount of variation occurred among individuals from the same population (6.73%). (Table 4)

Ecological Niche Divergence

Rarefication of sampling locations left 36 populations from an original 51. Limiting correlation between WorldClim variables to an R-value <0.90 reduced the number of variables from 20 to 13. The AUC value for the California and Utah/Arizona populations was 0.986 (sd 0.003) and 0.968 (sd 0.014), respectively; indicating an excellent model fit. Hellinger's-based *I* statistic had a 95% confidence interval from 0.0812 – 0.183. This indicated minimal niche overlap between California and Utah/Arizona populations (Fig. 12).

Consideration of Tomentum

The presence/absence of tomentose trichome pattern did not form any type of cluster among individuals in the haplotype networks. Individuals possessing this pattern were found in many different haplotypes, and no haplotype was solely associated with a tomentose pattern (Fig. 13).

DISCUSSION

Substantial patterns of genetic differentiation are apparent within *Cycladenia*. K-means clustering, PCA, and FFR reveal a consistent pattern of two strongly differentiated metapopulation groups. These metapopulation clusters correspond to a large geographic separation between the Californian populations and those occurring in Utah and Arizona. Ecological niche modeling suggests distinct climate preferences between these two groups, and the two groups show no indication of recent genetic exchange. Geographically structured genetic differentiation also exists within these two principle metapopulations.

Overall Diversity

Although seedling germination and sexual reproduction appears infrequent in *Cycladenia* (Sipes and Wolf, 1997), the patterns of heterozygosity within populations indicate that outcrossing was important in forming the individual genets that exist today. Furthermore, multiple haplotypes within each population, haplotypes shared between populations, and unique haplotypes within populations from a larger geographic area indicate that today's disjunct populations are likely the result of fragmentation. Isolation over a substantial, yet undetermined, time frame has resulted in divergence and notable variation among recent populations. This supports the hypothesis of paleoendemism for this species (U. S. Fish and Wildlife Service, 1986). The isolated nature of populations in *Cycladenia* today likely resulted from constriction of habitat due to changes in climate.

Genetic patterns indicate Northern California populations as the likely origin of *Cycladenia* in the Western US. In the FFRs, the outgroup connected to a Northern California haplotype. Additionally, Haplotype diversity is greater here than in other metapopulation clusters even with equal population sampling from both Northern California and Utah.

Utah/Arizona vs. California

Our results indicate strong separation between California and the Utah/Arizona populations. Four of the five FFRs indicate no shared haplotypes in heterozygotes that suggest a lack of recent or current gene flow. Although the haplotype networks indicate only a few changes between these two groups, the perennial, asexual propagation of clones would slow the rate of mutation accumulation between isolated population clusters of *Cycladenia*. The K-means clustering analysis separates Utah/Arizona populations from California populations as the best clustering of the data (K=2). To determine if ecological preferences paralleled genetic differences, ecological niche prediction was performed independently for California populations and for Utah/Arizona populations. Our results provide support for this division by documenting that California populations inhabit a significantly different climate with little niche overlap compared to populations from Utah/Arizona. In accordance with the General Lineage Concept (de Queiroz, 1998) both reproductive isolation and ecological distinction provides solid evidence to re-elevate Utah/Arizona populations from *Cycladenia humilis* var. *jonesii* to the species *Cycladenia jonesii* Eastwood (1942).

Eastern California

The Eastern California populations are currently recognized as the Utah var. *jonesii*, because they do not group morphologically with either of the California varieties but are similar in some features to var. *jonesii* (a point made by Eastwood when *C. jonesii* was first described [Eastwood, 1942]) (Rosatti, 2012). In contradiction to this hypothesis, Last's (2009) chloroplast data suggests a maternal relationship with Southern California var. *venusta*. Congruent with the chloroplast data, the Nc1a locus FFR clusters the Eastern California populations with the Southern populations. This grouping, recovered in multiple datasets, indicates a closer relation

between the Eastern and Southern California populations than with those from Utah/Arizona. Three other low-copy nuclear regions amplified in this study indicate Eastern California as a cluster independent from other named varieties, even from that of Southern California var. *venusta*.

FFR, as described by Doyle (1995) and later revived by Flot et al. (2010), provides a method to delimit species based on fields for recombination. Doyle's method suggests partitions that divide species occur in 100% of the loci, Flot et al. (2010) revised this criterion by using bipartition scores to determine how each single-locus FFR either supports or contradicts each partition in the data. If one locus does not contain the partition, the partition has less support. If there is not enough support, then the multi-locus FFR will be collapsed into a single field. Our data shows a consistent partition in four of the five single-locus FFRs, which clearly illustrates a division between Utah/Arizona and California populations. Additional molecular and ecological analyses add support to this partition. However, Nc4 contradicts this pattern, and clusters the Eastern California populations with those from Utah/Arizona. By Doyle's (1995) protocol, we would collapse the multi-locus FFR into a single field losing the pattern represented in the majority of the data. This all-or-nothing approach would misrepresent the genetic structure evident throughout our analyses, and overlooks a pattern that might even be expected in a slowly evolving DNA region that preserves the signature of more ancient relationships of events relative to faster evolving DNA regions.

To best represent the pattern recovered in the genetic and ecological data, we chose to exclude Nc4 from the multi-locus FFR. Nc4 is the simplest of the loci, and with fewer polymorphic loci it is more likely to have conserved ancestral sequences due to incomplete lineage sorting. Although the results from ecological niche modeling clearly indicate different

climate preferences between California and the Utah/Arizona populations, these predicted climate preferences are most similar on the Nevada/California border where the Eastern California populations are found. Due to the potentially higher likelihood of desirable climate in this region, past genetic exchange between the Utah/Arizona populations and the California populations may have occurred at this site. Thus, Eastern California populations would have been left with haplotypes from both metapopulation clusters. If our hypothesis is correct, then haplotypes in Nc4 may indicate the location of ancestral genetic exchange between *C. humilis* and *C. jonesii*. Important barriers to gene flow in *Cycladenia* may be discovered with future research on populations throughout this region. Regardless, neither pollen flow nor seed movement is occurring between California and Utah.

Although, Eastern California populations group with other California populations in our analyses, they tend to remain distinct from fully grouping with either California variety. Eastern California populations do not possess shared haplotypes in three of the five FFR's, and they are the fourth cluster to appear in the K-means clustering. In the PCA, populations from Eastern California also are separated from other populations by the third component, which comprises approximately 12% of the total variance. Eastern California populations appear just as distinct genetically as previously named varieties within California and are also distinguished morphologically from other varieties by having a smaller flower (lobes 4-5 mm [Rosatti, 2012]). Taken together these data indicate that the Eastern California populations are best treated taxonomically by the recognition of a new variety of *C. humilis* that circumscribes the eastern California populations.

Central California

Central California populations are of interest due to the inconsistent identification of individuals to either var. *humilis* (distributed to the north) or var. *venusta* (distributed to the south). Throughout our analyses this indecisive pattern was mirrored with these populations clustering with either Northern or Southern clusters. In our FFR analyses, these populations sort with Southern California once, Northern California once, and are part of all the California populations in three of the FFRs. They are intermediate between Northern and Southern populations in the second component of the PCA. Introgression between the Northern and Southern populations would likely create a pattern equivalent to this, and such a pattern suggests the Central California populations are the result of hybridization between varieties. The occurrence of hybridization at intermediate locations adds support for keeping var. *humilis* in the north and var. *venusta* in the south as varieties of a single taxon.

Arizona

Although there is only one known population located in Arizona, this population shows some divergence from the closest populations in Utah. Independent haplotypes are represented in two of the five loci, and Arizona is the sixth (and last geographically separated) cluster to appear in K-means clustering. Morphological data from Last (2009) shows that individuals from Arizona always have prominent trichomes located at the apex of mature leaves, while other populations of *Cycladenia* lack this character.

The ecological niche modeling shows broad swaths of suitable climate throughout much of Utah and into Northern Arizona. Despite the large areas of predicted distribution, there is a channel of unsuitable elevation or climate separating the single Arizona population from the Utah populations. Haplotypes in Arizona have diverged sufficiently from other var. *jonesii*

haplotypes to show separation, but the Arizona population is relatively homogenous suggesting the origin of this population was a dispersal event from Utah. Although the Arizona population is morphologically, and ecologically, differentiated, it is clearly related to Utah genealogically. Therefore, we suggest recognizing this population as a variety within *Cycladenia jonesii*.

Recombination

Recombination is present in one of the five loci, Nc7. Although recombination can cause disagreement among datasets and misrepresentation of diversity when estimating phylogenetic trees and networks, exploration of this dataset revealed the recombination event is inconsequential. Comparing the data among an alignment with recombined sequences removed, an alignment with the 5' end of the locus, and an alignment with 3' end of the locus showed similar results to the entire dataset. The recombination only occurred in Northern California individuals, and was not influential in the patterns of interest to this study. Furthermore, recombination, as with point mutations, are a source of genetic novelty appropriately measured when assessing genetic attributes such as haplotype diversity. Studies investigating the patterns of diversity in *Cycladenia* from Northern California using this region may benefit from considering the effects of recombination on results.

Is dense tomentum a taxonomic marker in C. humilis?

Support for tomentose plants being morphological variants of var. *humilis*, rather than a distinct taxon, is provided by the scattered pattern of tomentose individuals throughout the haplotype networks. Individuals possessing dense tomentum did not form any type of cluster in the haplotype networks, and this trait was found on individuals with many different haplotypes. No haplotype was solely associated with the presence of tomentum. A simple one-gene trait

would express a similar pattern in the haplotype networks, and thus we fail to reject the hypothesis that tomentose is a one-gene trait (Rosatti, 2012, Sipes and Wolf, 1997).

Taxonomy

Following the General Lineage Concept (de Queiroz, 1998) we used multiple lines of evidence to support species delimitation. In this study, various methods were used to discover population clusters with additional methods used to test the strength of these clusters. These tests provided evidence of reproductive isolation and ecological differentiation that supported the morphological differences observed between groups. The data provide consistent support for delimiting *C. humilis* into two species, *C. humilis* and *C. jonesii*, and rejects the hypothesis of *C. humilis* being the sole species in *Cycladenia*. Within both species, additional intraspecific taxa are warranted to better represent evolutionary significant units. The hypothesis that Central California populations are the result of introgression between the varieties of *C. humilis* is not rejected, while the hypothesis that Eastern California populations are part of *C. jonesii* is rejected. *Cycladenia* has patterns of genetic diversity that fails to reject the hypothesis of paleoendemism and has likely experienced range reduction in addition to habitat fragmentation. Although *Cycladenia* is uncommon, it possesses significant genetic diversity in both species and intraspecific taxa.

Conservation

Our understanding of the genetic diversity within *Cycladenia* has been based on previous estimates from a few known populations. *C. humilis* var. *jonesii* was listed as threatened (U.S. Fish and Wildlife Service, 1986) with only three general areas known. Since the listing, more populations have been discovered throughout the region. Genetic diversity within *C. jonesii* is equivalent to diversity within of *C. humilis*, particularly the Northern California var. *humilis. C.*

humilis vars. *humilis* and *venusta* have not been federally listed as threatened nor have they been listed on the California Rare Plant Inventory (CNPS, Rare Plant Program, 2015). Further investigation into the paucity of sexual reproduction and resiliency of clones to disturbances is warranted.

Although genetic diversity and population numbers of the Utah populations are more promising than was previously understood when var. *jonesii* was listed, other evolutionary significant units are of conservation concern. The cluster of populations from Eastern California and the single population from Arizona are comprised of only a few, genetically homogenous populations isolated from other varieties. Small populations with less genetic diversity generally have a higher risk of extinction. The remote locations of these populations provide protection simply due to difficulty of access and risk of human impact on these populations is low. Despite this current refuge, these populations may not have the genetic diversity needed to survive future climate change or habitat disturbances.

LITERATURE CITED

Boni, M. F., D. Posada, and M. W. Feldman. 2007. An exact nonparametric method for inferring mosaic structure in sequence triplets. *Genetics* 176: 1035-1047.

Brown, J. L. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution* 5: 694-700.

Calinski, T., and J. Harabasz. 1974. A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods* 3: 1-27.

Carstens, B. C., T. A. Pelletier, N. M. Reid, and J. D. Satler. 2013. How to fail at species delimitation. *Molecular Ecology* 22: 4369-4383.

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

Clement, M., Q. Snell, P. Walke, D. Posada, and K. Crandall. 2002. TCS: estimating gene genealogies. *Proceedings of the 16th International Parallel & Distributed Processing Symposium* 2.

Cullings, K. W. 1992. Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology* 1: 233-240.

CNPS, Rare Plant Program. 2015. Inventory of Rare and Endangered Plants (online edition, v8-02). California Native Plant Society, Sacramento, CA. Website http://www.rareplants.cnps.org [accessed 30 June 2015].

de Queiroz, K. 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. Oxford University Press, New York.

Doyle, J. J. 1995. The Irrelevance of Allele Tree Topologies for Species Delimitation, and a Non-Topological Alternative. *Systematic Botany* 20: 574-588.

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissues. *Phytochemical Bulletin* 19: 11-15.

Eastwood, A. 1902. Cycladenia venusta, Bulletin of the Torrey Botanical Club, vol. 29(2), 77.

Eastwood, A. 1942. Cycladenia jonesii, Leaflets of Western Botany, vol. 3(7), 159-160.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions* 17: 43-57.

ESRI. 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.

Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes - Application To Human Mitochondrial-DNA Restriction Data. *Genetics* 131: 479-491.

Flot, J-F., A. Couloux, and S. Tillier. 2010. Haplowebs as a graphical tool for delimiting species: a revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evolutionary Biology* 10.

Gibbs, M. J., J. S. Armstrong, and A. J. Gibbs. 2000. Sister-Scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics* 16: 573-582.

Gray, A. 1876. Cycladenia. *In* W. H. Brewer, S. Watson, AND A. Gray [eds.], Geological Survey of California, Botany, 474.

Hawksworth, D. L. 1995. Biodiversity: Measurement and Estimation, 5-12. Chapman and Hall, London.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.

Last, M. P. 2009. Intraspecific phylogeography of *Cycladenia humilis* (Apocynaceae). Master of Science, Brigham Young University, Provo, UT.

Livshultz, T., D. J. Middleton, M. E. Endress, and J. K. Williams. 2007. Phylogeny of Apocynoideae and the APSA clade (Apocynaceae S.L.). *Ann. Missouri Bot. Gard.* 94: 324-359.

Martin, D., and E. Rybicki. 2000. RDP: detection of recombination amongst aligned sequences. *Bioinformatics* 16: 562-563.

Martin, D. P., D. Posada, K. A. Crandall, and C. Williamson. 2005. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research and Human Retroviruses* 21.

Martin, D. P., B. Murrell, M. Golden, A. Khoosal, and B. Muhire. 2015. Detection and analysis of recombination patterns in virus genomes

Maynard Smith, J. 1992. Analyzing the mosaic structure of genes. *Molecular Evolution* 34: 126-129

Meirmans, P. G. 2012. AMOVA-based clustering of population genetic data. *Journal of Heredity* 103: 744-750.

Meirmans, P. G., and P. H. Van Tienderen. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:

792-794.

Muller, K. 2005. SeqState - primer design and sequence statistics for phylogenetic DNA data sets. *Applied Bioinformatics*: 65-69.

Munz, P. A. 1959. A California flora. University of California Press, Los Angeles.

Nguyen, L. T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Molecular Biology and Evolution* 32: 268-274.

Norse, E. A., K. L. Rosenbaum, D. S. Wilcove, B. A. Wilcox, R. W. H., J. D. W., and M. L. Stout. 1986. Conserving biological diversity in our national forests. The Wilderness Society, Washington D. C.

O'Brien, S. J., and E. Mayr. 1991. Bureaucratic mischief - recognizing endangered species and subspecies. *Science* 251: 1187-1188.

Padidam, M., S. Sawyer, and C. M. Fauquet. 1999. Possible emergence of new geminiviruses by frequent recombination. *Virology* 265: 218-225.

Peakall, R., and S. P.E. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537-2539.

Pence, V., L. Finke, and R. Niedz. 2014. Reducing Hyperhydricity in Shoot Cultures of Cycladenia humilis var. jonesii, an Endangered Dryland Species. *In Vitro Cellular & Developmental Biology-Animal* 50: S62-S62.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.

Posada, D., and K. A. Crandall. 2001. Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proceedings of the National Academy of Sciences of the United States of America* 98: 13757-13762.

R Core Team 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rambaut, A. 1996. Se-Al: Sequence Alignment Editor.

Rosatti, T. J. 2012. Cycladenia. P. 206 in *The Jepson Manual: higher plants of California*. B. G. Baldwin [ed.]. University of California Press, Berkeley.

Simmons, M. P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369-381.

Sipes, S. D., and V. J. Tepedino. 1996. Pollinator Lost? Reproduction by the enigmatic Jones' Cycladenia, Cycladenia humilis var. jonesii (Apocynaceae). Proceedings of the Southwestern Rare and Endangered Plant Conference, Flagstaff, AZ: 158-166.

Sipes, S. D., and P. G. Wolf. 1997. Clonal structure and patterns of allozyme diversity in the rare endemic Cycladenia humilis var. jonesii (Apocynaceae). *American Journal of Botany* 84: 401-409.

Straub, S. C. K., M. Fishbein, T. Livshultz, Z. Foster, M. Parks, K. Weitemier, R. C. Cronn, et al. 2011. Building a model: developing genomic resources for common milkweed (Asclepias syriaca) with low coverage genome sequencing. *BMC Genomics* 12: 211.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science* 240: 1285-1293.

Templeton, A. R. 1998. Human races: A genetic and evolutionary perspective. *American Anthropologist* 100: 632-650.

Templeton, A. R., K. M. Weiss, D. A. Nickerson, E. Boerwinkle, and C. F. Sing. 2000. Cladistic

structure within the human Lipoprotein lipase gene and its implications for phenotypic association studies. *Genetics* 156: 1259-1275.

U. S. Fish and Wildlife Service 1986. Rule to determine *Cycladenia humilis* var. *jonesii* (Jones cycladenia) to be a threatened species with critical habitat. *Federal Register* 51: 16526-16530.

VanDerWal, J., L. Falconi, S. Januchowski, L. Shoo, and C. Storlie. 2014. Species Distribution Modelling Tools: Tools for processing data associated with species distribution modelling exercises, version 1.1-221.

Warren, D. L., R. E. Glor, and M. Turelli. 2008. Environmental Niche Equivalency versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution* 62: 2868-2883.

Welsh, S. L., N. D. Atwood, and S. Goodrich. 1987. A Utah Flora. Great Basin Naturalist Memoirs No. 9, Brigham Young University, Provo, UT.

TABLES AND FIGURES

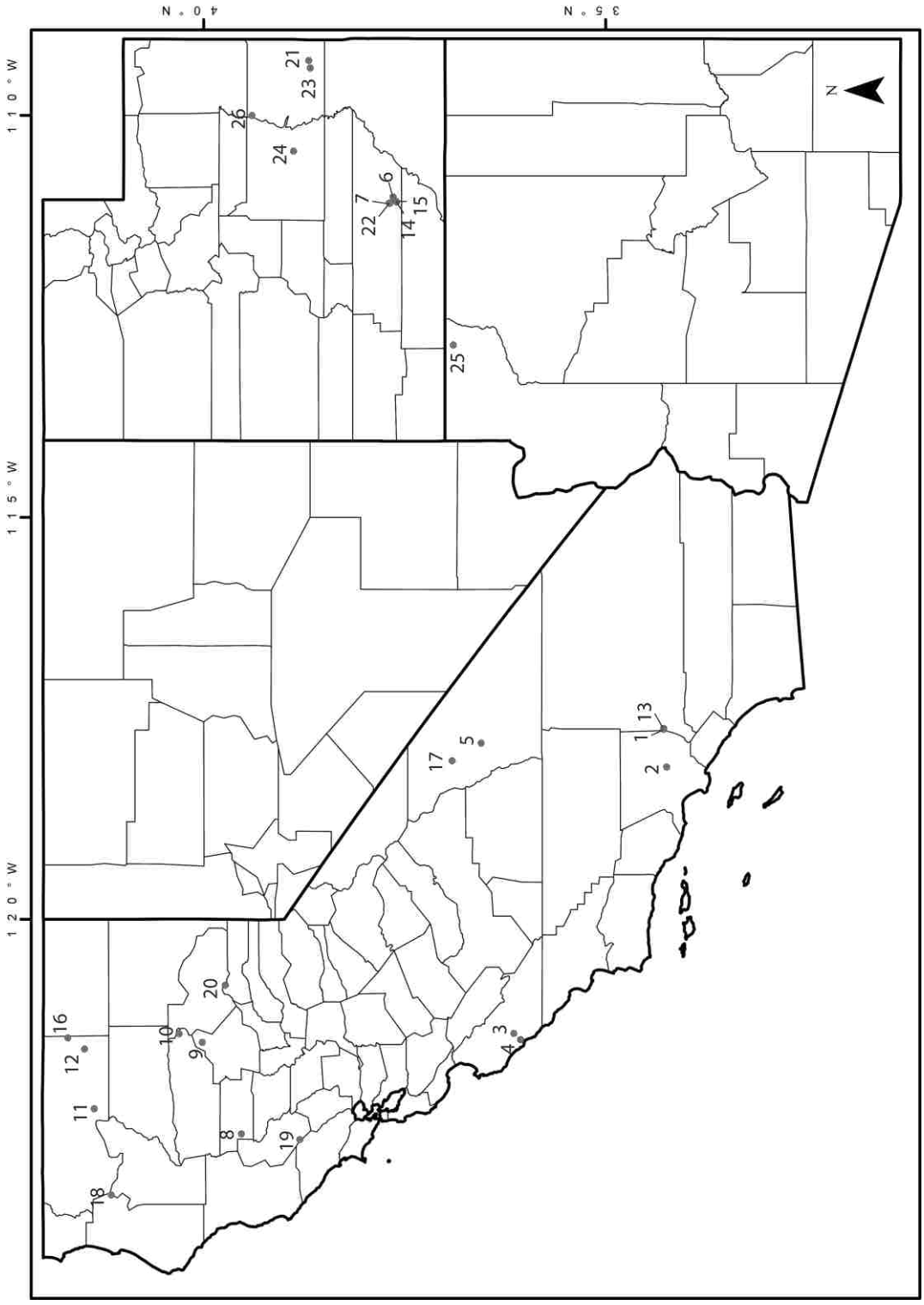

Figure 1: Populations used for genetic sampling. Northern California: 8, 9, 10, 11, 12, 16, 18, 19, 20; Central California: 3, 4; Eastern California: 5, 17; Arizona: 25; Utah: 6, 7, 14, 15, 21, 22, 23, 24, 26.

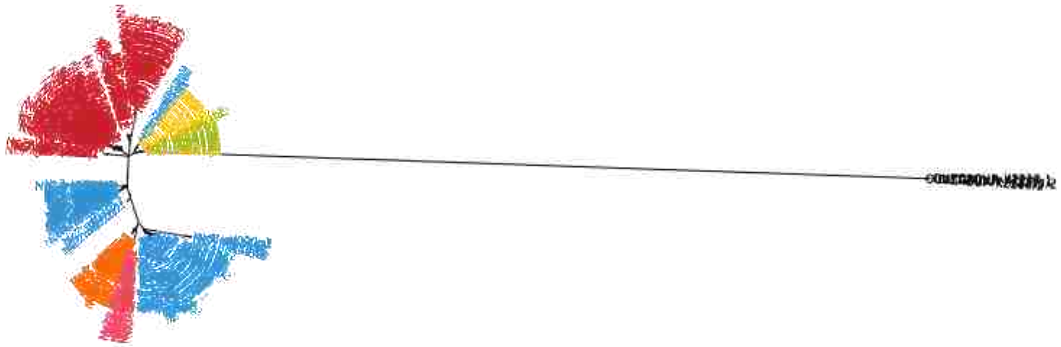

Figure 2: Likelihood tree of entire Nc7b locus including recombinants. Blue: Northern CA, Green: Central CA, Yellow: Southern CA, Red: Utah, Pink: Arizona, Black: Outgroup.

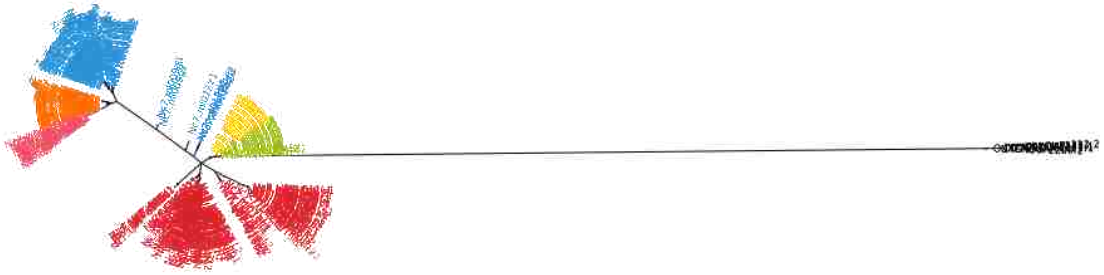

Figure 3: Likelihood tree of entire Nc7b locus excluding recombinants. Blue: Northern CA, Green: Central CA, Yellow: Southern CA, Red: Utah, Pink: Arizona, Black: Outgroup.

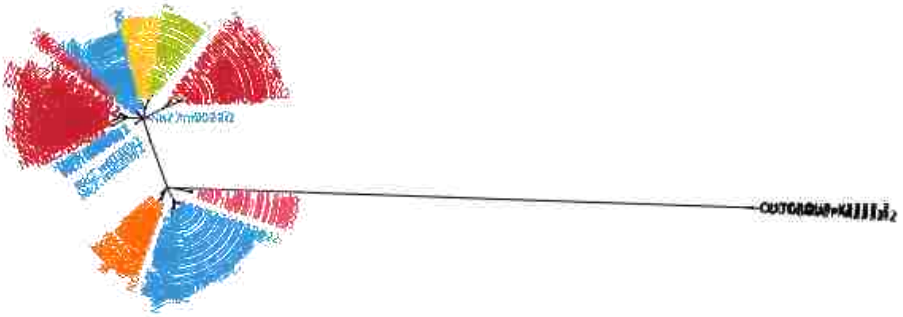

Figure 4: Likelihood tree of 3' end of Nc7b locus. Blue: Northern CA, Green: Central CA, Yellow: Southern CA, Red: Utah, Pink: Arizona, Black: Outgroup.

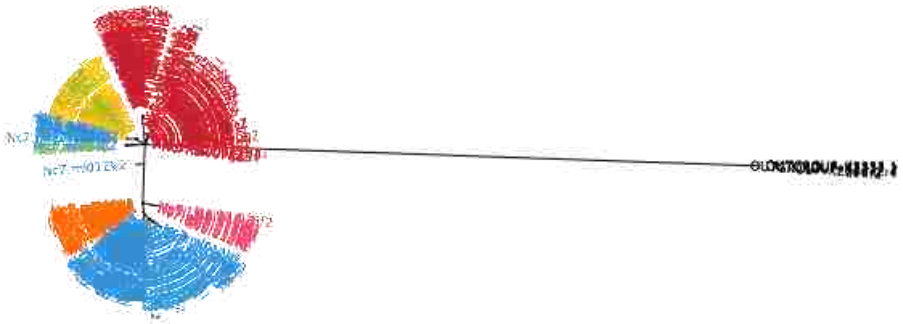

Figure 5: Likelihood tree of 5' end of Nc7b locus. Blue: Northern CA, Green: Central CA, Yellow: Southern CA, Red: Utah, Pink: Arizona, Black: Outgroup.

Table 1: K-means Clustering. Asterisk represents best clustering according to Calinski & Harabasz' pseudo-F: k=2.

Population	k2*	k3	k4	k5	k6	k7	k8	k9	k10
3: Central CA	2	3	3	5	4	6	1	4	4
4: Central CA	2	3	3	5	4	6	1	4	4
1: Southern CA	2	3	3	4	3	2	6	6	9
2: Southern CA	2	3	3	4	3	2	6	6	9
13: Southern CA	2	3	3	4	3	2	6	6	9
5: Eastern CA	2	2	2	3	1	7	4	8	2
17: Eastern CA	2	2	2	3	1	7	4	8	2
8: Northern CA	2	2	4	1	6	3	3	5	5
9: Northern CA	2	2	4	1	6	3	3	5	5
10: Northern CA	2	2	4	1	6	3	3	5	5
11: Northern CA	2	2	4	1	6	3	3	5	5
12: Northern CA	2	2	4	1	6	3	3	5	5
16: Northern CA	2	2	4	1	6	3	3	5	5
19: Northern CA	2	2	4	1	6	3	3	5	5
20: Northern CA	2	2	4	1	6	3	3	5	5
18: Northern CA	2	2	4	1	6	3	5	9	6
6: Utah	1	1	1	2	2	5	2	3	8
7: Utah	1	1	1	2	2	5	2	3	8
14: Utah	1	1	1	2	2	5	2	3	8
15: Utah	1	1	1	2	2	5	2	3	8
22: Utah	1	1	1	2	2	5	2	3	8
24: Utah	1	1	1	2	2	5	2	3	10
21: Utah	1	1	1	2	2	1	8	2	7
23: Utah	1	1	1	2	2	1	8	2	7
26: Utah	1	1	1	2	2	1	8	7	1
25: Arizona	1	1	1	2	5	4	7	1	3

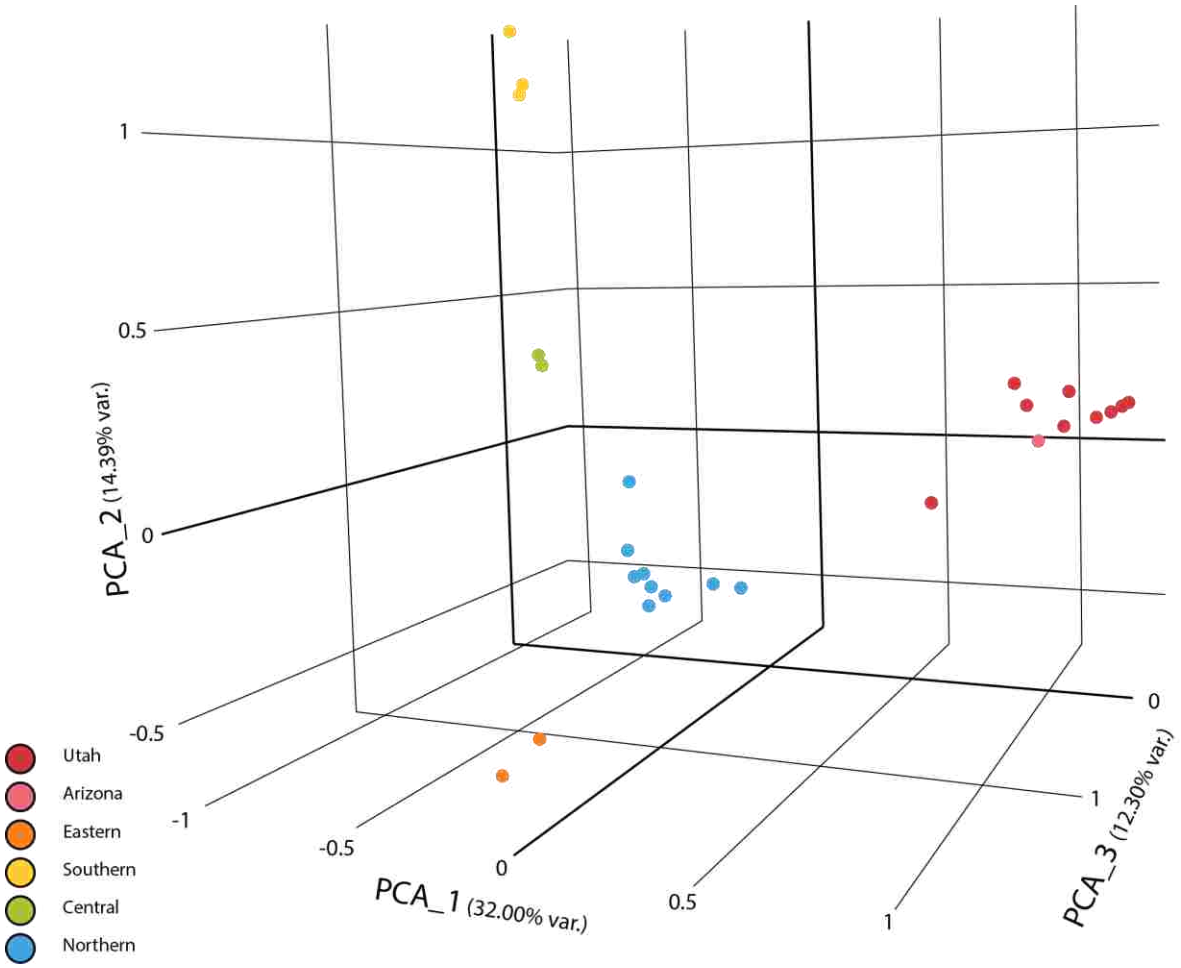

Figure 6: Principle Components Analysis. The first three components contain 58.69% of the variation between populations.

Table 2: Principle Components Analysis.

Axis	Eigenvalue	%variance	cumulative	G'st (Nei)	p-value
1	0.706	32.002	32.002	0.154	0.05
2	0.318	14.388	46.39	0.069	0.38
3	0.271	12.299	58.689	0.059	0.09
4	0.193	8.753	67.442	0.042	0.24
5	0.146	6.636	74.078	0.032	0.43
6	0.113	5.136	79.214	0.025	0.66
7	0.071	3.215	82.428	0.015	1
8	0.067	3.014	85.442	0.014	0.98
9	0.054	2.427	87.868	0.012	1
10	0.049	2.214	90.082	0.011	0.98

Table 3: Summary of fields for recombination results.

Locus	Haplotypes	Singletons	Polymorphic loci	Indels	Individuals in most common haplotype	Individuals homozygous for most common haplotype	Allele pools	Outgroup?
Nc1a	70	22	46	4	76	23	10	Failed to amplify
Nc3	24	7	23	0	111	80	4	To Northern CA
Nc4	10	0	9	0	174	122	2	To Northern CA
Nc7b	74	26	87	5	61	12	5	To Northern CA
Nc10a	32	11	33	4	113	97	2	Failed to amplify

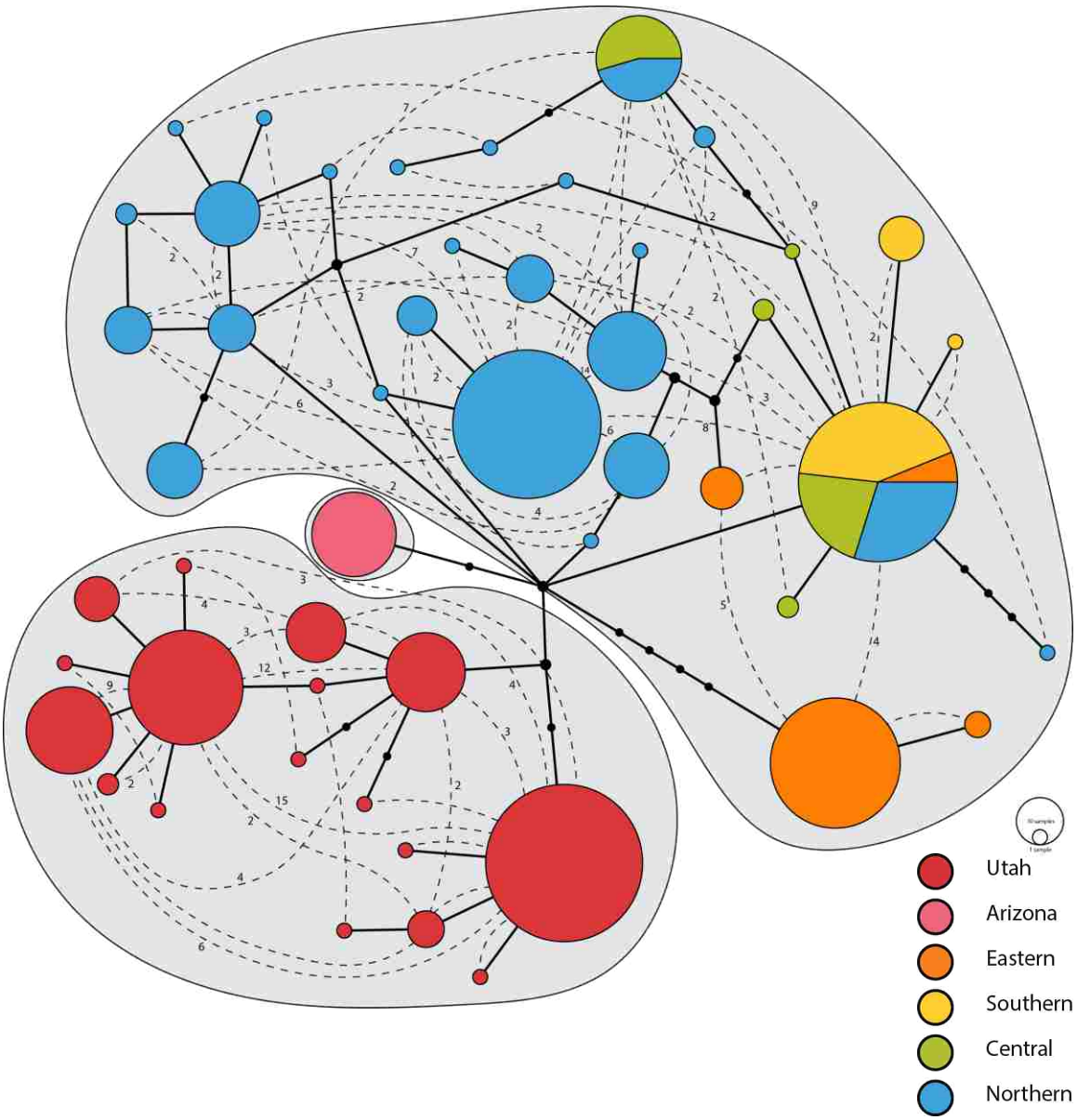

Figure 7: Nc1a Haploweb and FFR. Colored circles represent haplotypes with black circles representing un-sampled haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools.

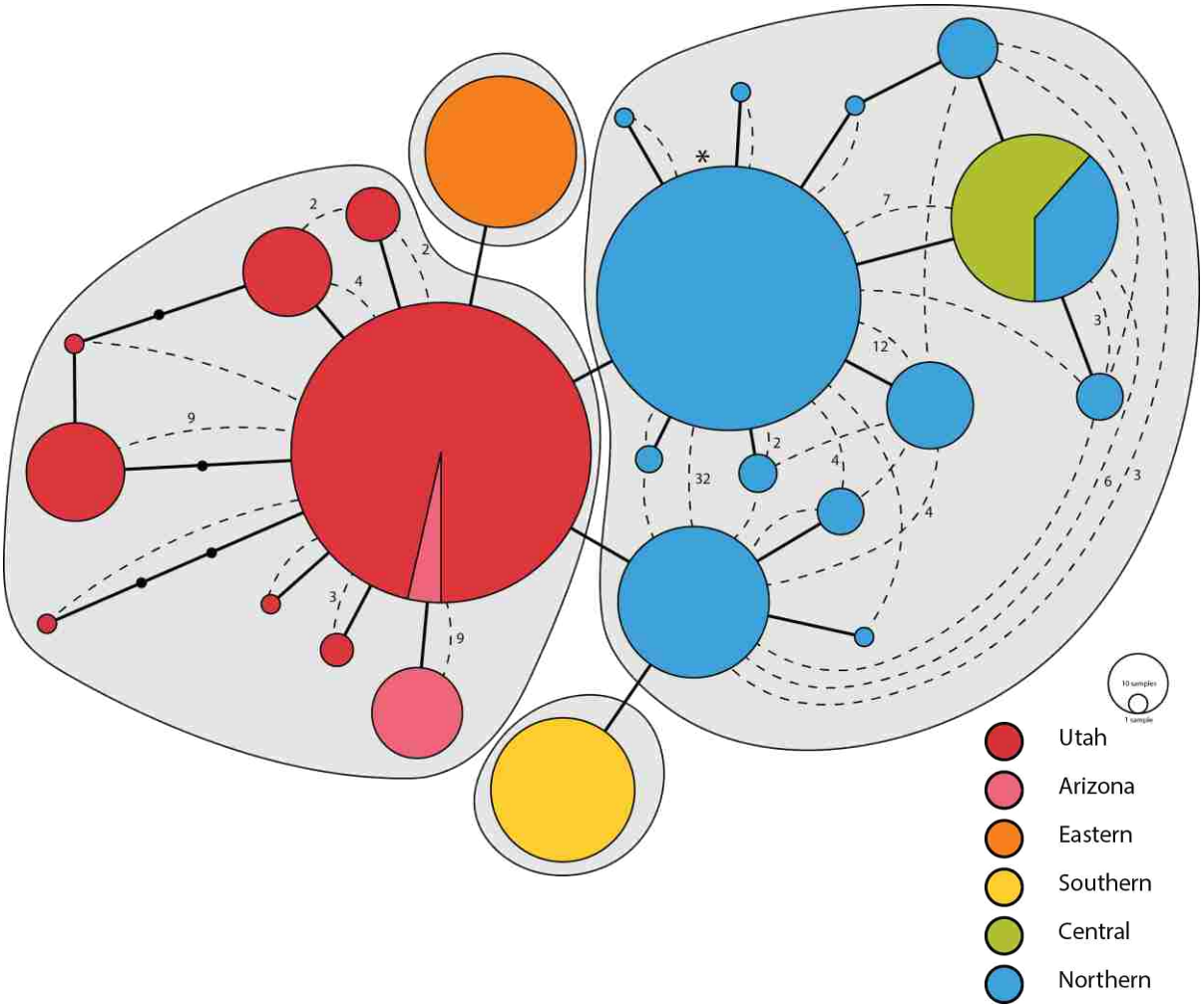

Figure 8: Nc3 Haploweb and FFR. Colored circles represent haplotypes with black circles representing un-sampled haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools. An asterisk indicates where the outgroup connects to the network.

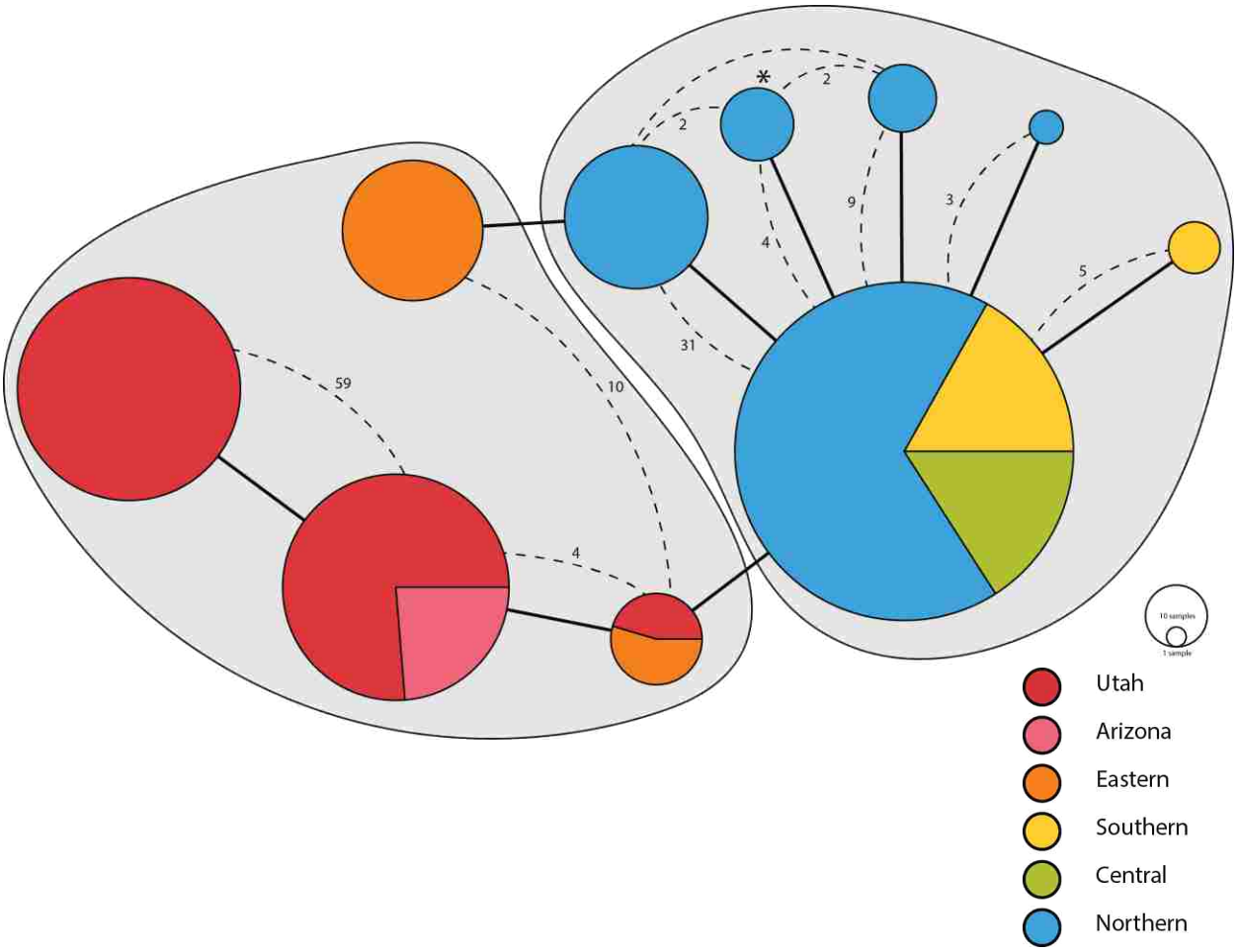

Figure 9: Nc4 Haploweb and FFR. Colored circles represent haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools. An asterisk indicates where the outgroup connects to the network.

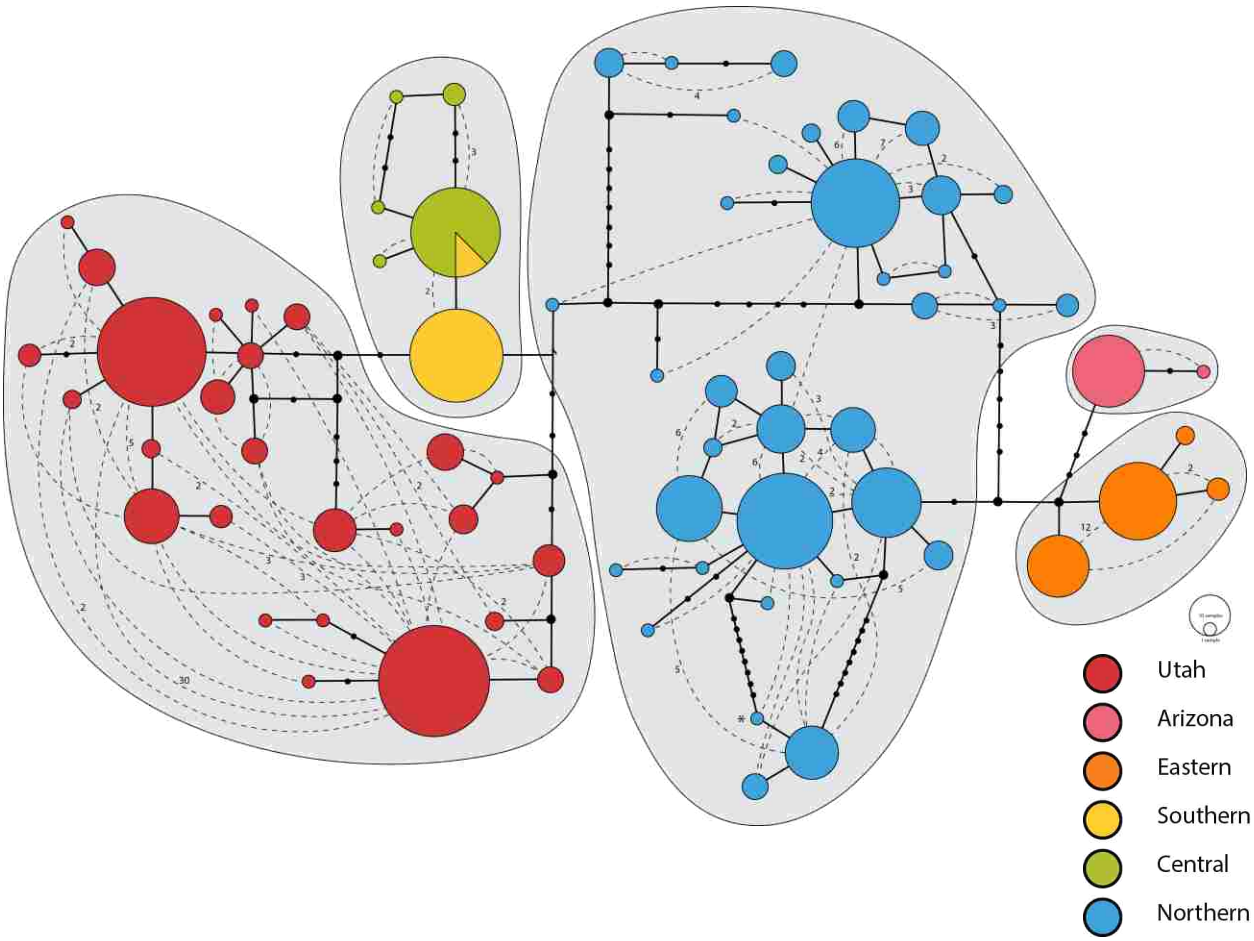

Figure 10: Nc7b Haploweb and FFR. Colored circles represent haplotypes with black circles representing un-sampled haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools. An asterisk indicates where the outgroup connects to the network.

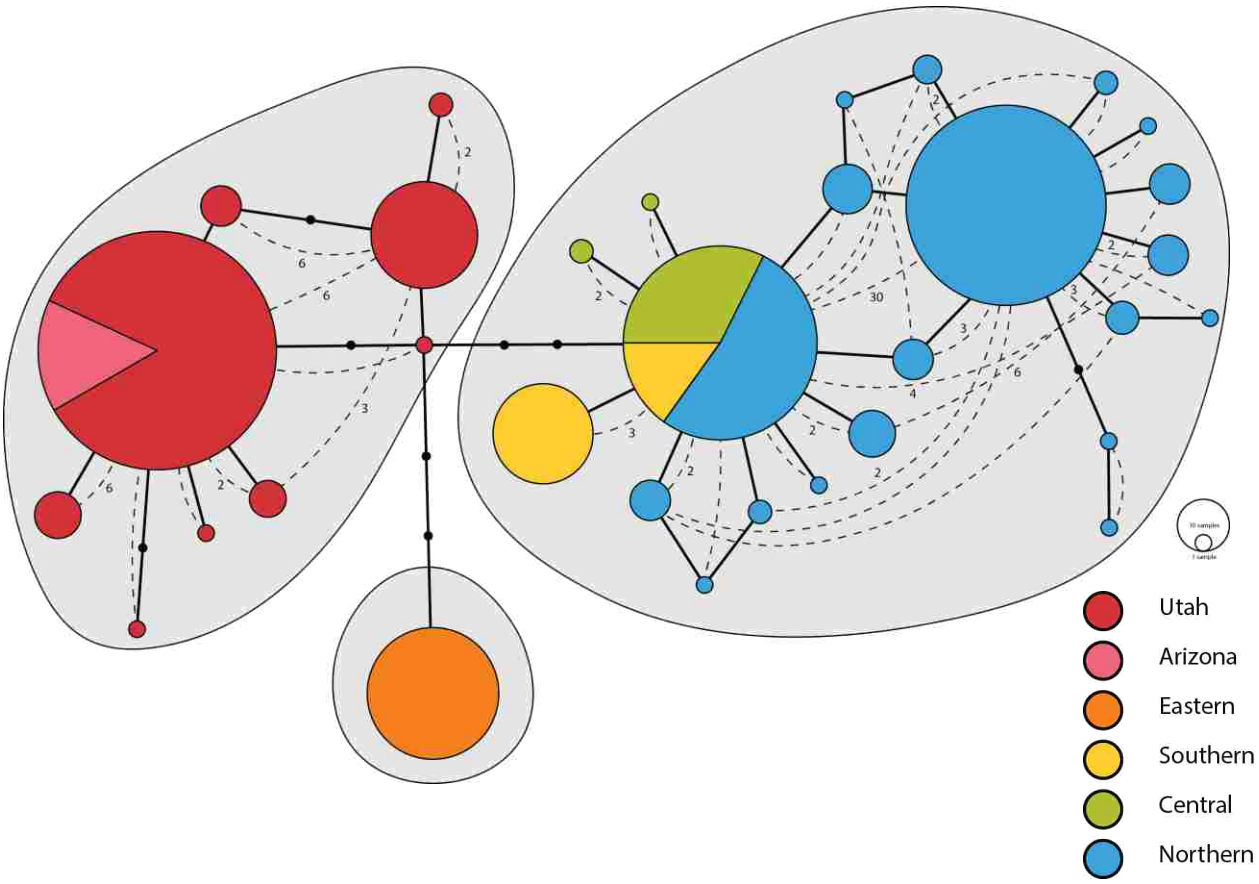

Figure 11: Nc10a Haploweb and FFR. Colored circles represent haplotypes with black circles representing un-sampled haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools.

Table 4: AMOVA results from Arlequin3.5.2.1. Groups were comprised of Utah populations, Arizona populations, Northern California populations, Eastern California populations, and Central/Southern California populations.

Source of Variation	d.f.	Sum of Squares	Variance components	Percentage of Variation	F-Statistics	P-value
Among groups	4	480.242	0.82831 Va	34.29%	F_{CT} = 0.34295	<0.0001
Among populations within groups	21	267.200	0.41470 Vb	17.17%	F_{SC} = 0.26132	<0.0001
Among individuals within populations	337	449.842	0.16260 Vc	6.73%	F_{IS} = 0.138710	<0.0001
Within individuals	363	366.500	1.00964 Vd	41.80%		
Total	725	1563.784	2.41525			

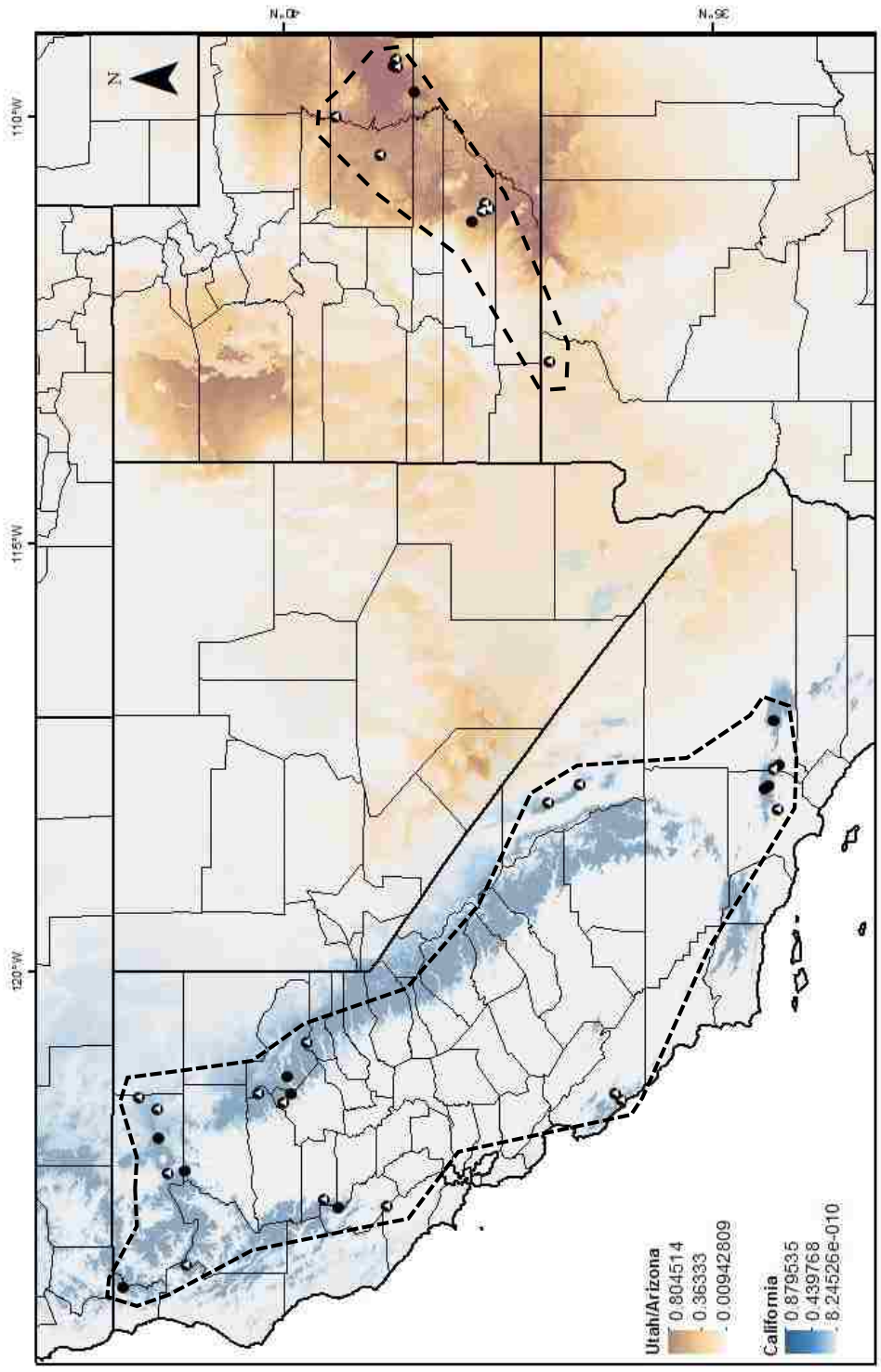

Figure 12: Ecological Niche Modeling. Predicted distribution of Utah/Arizona populations in red, and predicted distribution of California populations in blue. Black dots represent populations used in the niche modeling with white triangles representing populations genetically sampled. Dashed lines indicate sample ranges, and extrapolation beyond these lines should be considered with caution.

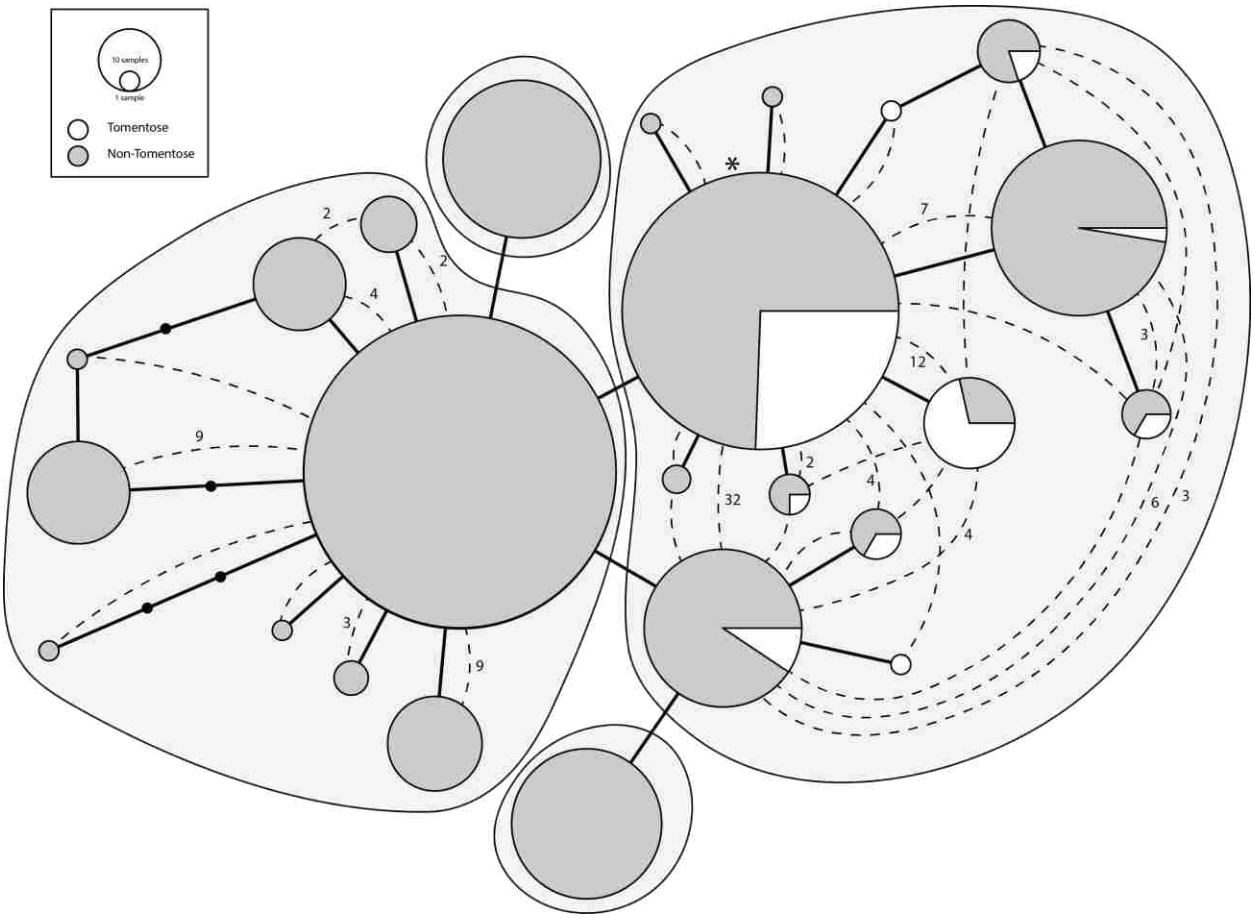

Figure 13: Tomentose individuals mapped onto the Nc3 locus. The FFR on the right is comprised of solely Northern California Individuals. Large circles represent haplotypes with black circles representing un-sampled haplotypes. Dotted lines show which haplotypes are contained in heterozygotes with numbers indicating how many individuals are heterozygous for those haplotypes. If no number is present only one individual possesses that combination. Larger gray areas represent allele pools. An asterisk indicates where the outgroup connects to the network. Similar patterns were evident in the other loci.

APPENDIX I

Collection and voucher information: all vouchers are deposited at Brigham Young University (BRY). Information includes: Site Number, locality, (number of individuals sampled), collection date, *collectors*, and collection number.

Site 1	California, San Bernardino Co., Devils Backbone near Mt. San Antonio, San Gabriel Mts. (7) 09 June 2007, *M. Last & L. Chan*, ml-001
Site 2	California, Los Angeles Co., Mt. Disappointment, San Gabriel Mts. (6) 11 June 2007, *M. Last & L. Chan*, ml-002
Site 3	CA, Monterey Co., Junipero Serra, Santa Lucia Mts. (16) 12 June 2007, *M. Last & L. Chan,* ml-003
Site 4	CA, Monterey Co., Near Cone Peak Santa Lucia Mts. (8) 12 June 2007, *M. Last & L. Chan,* ml-004
Site 5	CA, Inyo Co., Cerro Gordo Springs area, Inyo Mts. (16) 14 June 2007, *M. Last & L. Chan,* ml-004
Site 6	UT, Garfield Co., Purple Hills, Glenn Canyon Recreation Area (16) 20 June 2007, *J.M. Spence, M. Last, & L.A. Johnson,* ml-006
Site 7	UT, Garfield Co.,Horse Pasture Mesa area, Glenn Canyon Recreation Area (16) 21 June 2007, *J.M. Spence, M. Last, & L.A. Johnson,* ml-007
Site 8	CA, Glenn Co., Noel Springs, Northern Coastal Range (16) 13 July 2007, *M. Last & A. Maas,* ml-009
Site 9	CA, Butte Co., Bottle Hill, Sierra Nevada (16) 16 July 2007, *M. Last & A. Maas,* ml-010
Site 10	CA, Tehama Co., Guernsey Camp, Sierra Nevada (16) 16 July 2007, *M. Last & A. Maas,* ml-011
Site 11	CA, Siskiyou Co., Black Butte, High Cascades (12) 17 July 2007, *M. Last & A. Maas,* ml-012
Site 12	CA, Siskiyou Co., Jot Dean Ice Cave (16) 18 July 2007, *M. Last & A. Maas,* ml-013
Site 13	CA, San Bernardino Co., Mt San Antonio, San Gabriel Mts. (16) 20 July 2007, *M. Last & A. Maas,* ml-014
Site 14	UT, Garfield Co., Moody Canyon, Glenn Canyon Recreation Area (16) 17 June 2008, *J.M. Spence, M. Last, N. Laitinen, & D. Kunakeva*, ml-018
Site 15	UT, Garfield Co., Moody Canyon, Glenn Canyon Recreation Area (16) 17 June 2008, *J.M. Spence, M. Last, N. Laitinen, & D. Kunakeva*, ml-019
Site 16	CA, Siskiyou Co., Caldwell Butte, Modoc Plateau (16) 24 June 2008, *M. Last & T. Taylor*, ml-020
Site 17	CA, Inyo Co., Seep Hole Spring, Inyo Mts. (16) 26 June 2008, *M. Last & T. Taylor,* ml-021
Site 18	CA, Humboldt Co., Devils Backbone – Salmon Mts., Klamath Mountains (16) 5 July 2008, *M. Last, R. Last, & T. Taylor*, ml022
Site 19	CA, Lake Co., Cobb Mt., Sierra Nevada (16) 14 July 2008, *M. Last & T. Taylor*, ml-023
Site 20	CA, Sierra Co., Stanford Mt., Sierra Nevada (16) 14 July 2008, *M. Last & T. Taylor*, ml-024
Site 21	UT, Grand Co., Onion Creek (11) 28 June 2007, *L. Chan*, lmc-001
Site 22	UT, Garfield Co., Choprock Bench (8) 21 June 2007, *L.A. Johnson, M. Last, & J. Spence*, 07-023
Site 23	UT, Grand Co., Castle Valley (16) 21 May 2007, *L.A. Johnson, L. Chan*, 07-016
Site 24	UT, Emery Co., San Rafael Reef (16) 15 May 2007, *L.A. Johnson & H. Barnes*, 07-011
Site 25	AZ, Mohave Co., Vermillion Cliffs (16) 21 May 2008, *L.A. Johnson & C. Zanotti*, 08-018
Site 26	UT, Grand Co., Joe Hutch Canyon area (7) 26 July 2008, *N.D. Atwood*, 32467
-	CA, Del Norte Co., Ridge between Broken Rib Mt. and Wounded Knee Mt. 12 July 2012, *M.R. Mesler, L. Sloan, & I. Zacher*, 1225 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
-	CA, Colusa Co., Snow Mt. along trail to Summit Basin

8 Aug 2011, *S. Harrison* Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)

- CA, Siskiyou Co., Near Castella, 3.1 mi. north from State Park entrance up Castle Dome Trail
  5 June 2013, *R.D. Whittlesey*, 5 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Siskiyou Co., Klamath National Forest, south slope of Rainbow Mt.
  12 July 2002, *R.E. Preston,* 1917 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Butte Co., East of road sign 300 R 4, about 3 miles (air) southeast of Bald Mt.
  19 Aug 2004, *L. Ahart,* 11396 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Tehama Co., 5.5 km southeast of Childs Meadows, southwest facing slope of Lost Creek Plateau
  2 June 2013, *M. Baker,*17753 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Plumas Co., Five miles (air) north of Bucks Lake and 3 miles south of Highway 70
  11 September 2005, *B. Castro, & L.P. Janeway,* 1523 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Los Angeles Co., San Gabriel Mts., Devils Punchbowl Co. Park, north base of peak 6374 at extreme southeast corner of park
  10 June 2008, *R.G. Swinney,* 8997 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Los Angeles Co., Approx. 20 ft. north of saddle along Pacific Crest Trail
  27 June 2012, *J. Tirrell*, LS0088 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, Los Angeles Co., San Gabriel Mts. Coldwater Canyon Tributary
  16 September 2000*, R.G. Swinney,* 7867 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- CA, San Bernardino Co., Summit of Cucamonga Peak
  28 July 2004, *N. Fraga, S. De Groot, P. Morton, & A. Virgen,* 1316 Data provided by the participants of the Consortium of California Herbaria (ucjeps.berkeley.edu/consortium/)
- UT, Garfield Co., Little Brown Bench, Pioneer Mesa
  18 April 2014, *Brian Elliott*, BLM
- UT, Garfield Co., Wolverine Creek, Pioneer Mesa
  19 April 2014, *Brian Elliott*, BLM
- UT, Garfield Co., Wolverine Creek, Pioneer Mesa
  19 April 2014, *Brian Elliott*, BLM
- UT, Garfield Co., Wolverine Creek, Pioneer Mesa
  19 April 2014, *Brian Elliott*, BLM
- UT, Garfield Co., Wolverine Creek, Pioneer Mesa
  19 April 2014, *Brian Elliott*, BLM
- UT, Emery Co., North Shadscale Mesa, Spotted Wolf Canyon
  29 April 2014, *Brian Elliott*, BLM
- UT, Emery Co., North Shadscale Mesa, Spotted Wolf Canyon
  29 April 2014, *Brian Elliott*, BLM
- UT, Emery Co., North Shadscale Mesa, Spotted Wolf Canyon
  29 April 2014, *Brian Elliott*, BLM
- UT, San Juan Co., Potash Mine, Shafer Basin
  5 May 2014, *Brian Elliott*, BLM
- UT, Grand Co., Dome Plateau, Big Bend
  7 May 2014, *Brian Elliott*, BLM
- UT, Grand Co., Dome Plateau, Big Bend
  7 May 2014, *Brian Elliott*, BLM
- UT, Grand Co., Dome Plateau, Big Bend
  9 May 2014, *Brian Elliott*, BLM
- UT, Grand Co., Dome Plateau, Big Bend
  9 May 2014, *Brian Elliott*, BLM
- UT, Grand Co., Dome Plateau, Big Bend
  9 May 2014, *Brian Elliott*, BLM

Outgroup North Carolina, Johnston Co. (4)
2008, *A. Krings*, 2227, 2232, 2233, 2255

APPENDIX II

Primers and amplified regions

Primer	(Straub et al. 2011)
Nc1aF&R	At1g24360a
Nc3F&R	At1g55880a
Nc4F&R	At2g03120a
Nc7bF&R	At2g302000b
Nc10aF&R	At4g13430a
Nc3Fa1	AAGGGCGGAGGTTAAAGAAC
Nc3Ra1	AAGTGCTTGTGCCAGTCTGA
Nc10a5F	GTCTGGGACCGTGAAAAGGT
Nc10a5R	TGGAGGCACCTGAAACATAA
Nc7bRa	GACCTCCAGAAACCGCATAA
Nc7bR(internal)	GGCAAAGGCAATGTATACTGGT, TTGAAAGAAAACGCATGGA