



Conservation Genetics of the Endangered Del Mar Manzanita (*Arctostaphylos glandulosa* subsp. *Crassifolia*) Based On Rad Sequencing Data

Authors: Burge, Dylan O., Parker, V. Thomas, Mulligan, Margaret, and García Valderamma, César

Source: Madroño, 65(3) : 117-130

Published By: California Botanical Society

URL: <https://doi.org/10.3120/0024-9637-65.3.117>

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

CONSERVATION GENETICS OF THE ENDANGERED DEL MAR MANZANITA
(*ARCTOSTAPHYLOS GLANDULOSA* SUBSP. *CRASSIFOLIA*) BASED ON RAD
SEQUENCING DATA

DYLAN O. BURGE

University of California, Davis, One Shields Avenue, Davis, CA, 95616
dylan.orion.burge@gmail.com

V. THOMAS PARKER

Department of Biology, San Francisco State University, 1600 Holloway Ave, San Francisco,
CA, 94132

MARGARET MULLIGAN

Department of Botany, San Diego Natural History Museum, Balboa Park, 1788 El Prado,
San Diego, CA, 92101

CÉSAR GARCÍA VALDERAMMA

Nativ@s de las Californias, Cuarto Balcón 15901, Balcón Las Huertas, Tijuana, Baja
California, México

ABSTRACT

Del Mar manzanita (*Arctostaphylos glandulosa* Eastw. subsp. *crassifolia* (Jeps.) P.V. Wells) is a federally listed endangered shrub found in San Diego County, California and Baja California, Mexico. This manzanita forms part of the imperiled southern maritime chaparral of southwestern California and adjacent Baja California, Mexico. Del Mar manzanita is problematic to identify because of morphological intergradation with other subspecies of *A. glandulosa*. Such intergradation could result from biological phenomena, such as gene flow among subspecies. Alternatively, it could be that the current circumscription of the Del Mar manzanita is not correct, and that the morphological characters used to diagnose this subspecies are inaccurate indicators of underlying genetics. This situation leads to problems for conservation planning, where accurate identification of individual plants is essential. Here, we used high-throughput sequencing of restriction-site associated DNA markers (RADseq) to develop single nucleotide polymorphism (SNP) data for a large sample of putative Del Mar manzanita, and a small sample of closely related subspecies of *A. glandulosa*. We analyzed genetic relationships using a total of 65,964 SNPs, with the aim of testing whether morphological traits used to identify Del Mar manzanita are an accurate reflection of underlying genetic patterns. We conclude that vegetative morphology is a poor predictor of genetic patterns, and that the current morphology-based circumscription of Del Mar manzanita is probably in need of some change. However, due to the limited sampling of *A. glandulosa* subspecies in this study, it is not possible to determine the taxonomic limits of Del Mar manzanita using our SNP data.

Key Words: California, DNA, genome, military, polyploid, rare, San Diego County.

Arctostaphylos glandulosa Eastw. subsp. *crassifolia* (Jeps.) P.V. Wells (hereafter referred to by its common name, Del Mar manzanita), is an endangered, burl-forming, tetraploid ($n = 26$) shrub found in coastal chaparral habitats of San Diego County, California, and Baja California, Mexico (Federal Register 1996). Like all subspecies of *A. glandulosa*, Del Mar manzanita can re-sprout from a buried lignotuber following fire, and is therefore considered to be strongly fire-adapted (Wells 1969). While pollination biology research has not been carried out on the Del Mar manzanita specifically, at least one close relative is essentially self-incompatible (Fulton and Carpenter 1978), so Del Mar manzanita is assumed to employ the same strategy.

Del Mar manzanita is generally found on sandstone-derived terraces near the sea (Federal Register 1996; Parker et al. 2012) and is one of several key

indicators of the imperiled southern maritime chaparral plant community (Hogan et al. 1996). Throughout its small geographic range, Del Mar manzanita faces threats from urban development and other human activities, which have led to the loss of many populations and the fragmentation of remaining ones, especially in far western San Diego County (Federal Register 1996; USFWS 2010). Although Del Mar manzanita has been identified in northern Baja California, Mexico (USFWS 2010; Parker et al. 2012), the taxon has no special legal status according to Mexico's list of protected taxa, the Norma Oficial Mexicana (Mexican Official Standard) NOM-059 (SEMARNAP 2002).

Along the eastern margin of its geographic range, Del Mar manzanita comes into contact with its close relative, Eastwood manzanita (*A. glandulosa* Eastw. subsp. *glandulosa*; including *A. glandulosa* Eastw.

TABLE 1. COLLECTING LOCALES AND SAMPLING INFORMATION. *Taxon* represents *A. glandulosa* individuals identified according to dichotomous keys in the Jepson Manual (Parker et al. 2012). *Code* indicates a field collection code for each locale used as a short-hand to refer to individual locales; each code corresponds to an herbarium voucher specimen housed at the UC Davis Center for Plant Diversity or the San Diego Natural History Museum (only one herbarium voucher was collected per site). M. Mulligan sampled locales 3318 and 3321; D. Burge sampled all other locales. *Samples* correspond to the number of individuals sequenced. Latitude and longitude are given in the WGS 84 datum.

| Taxon | Code | Locale name | Samples | Latitude | Longitude |
|-------------------------------|------------------------------|----------------------------|---------------|----------|-----------|
| <i>A. g. ssp. atumescens</i> | 2006 | Cerro Buenavista | 1 | 31.6737 | -116.6314 |
| <i>A. g. ssp. crassifolia</i> | 1694 | Cerro del Coronel | 1 | 32.2830 | -116.9300 |
| | 1701 | Cañon San Isidro | 2 | 31.2920 | -116.3434 |
| | 1703 | Mesa de Descanso | 4 | 32.1718 | -116.8898 |
| | 1709 | Encinitas Community Center | 3 | 33.0444 | -117.2669 |
| | 1717 | Torrey Pines | 3 | 32.9406 | -117.2471 |
| | 1719 | MCAS 1 | 2 | 32.9164 | -117.0398 |
| | 1720 | MCAS 2 | 3 | 32.8897 | -117.0640 |
| | 1722 | MCAS 3 | 3 | 32.8787 | -117.0659 |
| | 1723 | MCAS 4 | 3 | 32.8649 | -117.0690 |
| | 1725 | MCAS 5 | 3 | 32.8932 | -117.0757 |
| | 1729 | Crest Canyon | 4 | 32.9501 | -117.2538 |
| | 2071 | Encinitas | 3 | 33.0360 | -117.2487 |
| | 3318 | MCAS 18 | 3 | 32.8938 | -117.0398 |
| | 3321 | MCAS 17 | 3 | 32.8915 | -117.0516 |
| | <i>A. g. ssp. glandulosa</i> | 1523 | Cavedale Road | 2 | 38.3624 |
| 1688 | | Newhall Pass | 3 | 34.3471 | -118.5102 |
| 1746 | | Mount Tamalpais | 4 | 37.9110 | -122.5775 |
| 2033 | | Bolinas Ridge | 2 | 37.9468 | -122.6665 |
| 2087 | | Viejas Mountain | 1 | 32.8540 | -116.7421 |
| 2090 | | Santa Ana Mountains | 1 | 33.6535 | -117.4455 |

subsp. *zacaensis* Eastw.; Parker et al. 2012). Eastwood manzanita is a plant of interior habitats in the Coast and Peninsular Ranges from central California to northern Baja California, Mexico. Although the extent to which genetic exchange among wild populations takes place is not known, the close geographic proximity of the two subspecies in San Diego County suggests that there are opportunities for such exchange (Keeley et al. 2007).

Like many *Arctostaphylos*, Del Mar manzanita is difficult to identify, with morphological traits used to define this taxon intergrading into other subspecies of *A. glandulosa*, particularly Eastwood manzanita. In the past, this continuity was assumed to result from gene flow among taxa (Keeley et al. 2007). However, it is also possible that the morphological traits used to define taxa are simply poor indicators of underlying genetic patterns. In the latter case, new circumscriptions may be required to bring taxonomy into line with biology. As with any rare plant, taxonomic ambiguity and resulting problems with identification can lead to difficulties for conservation and recovery. Therefore, tools are needed that facilitate objective means of identification. Genetic analyses provide such a tool (Ogden et al. 2009; Kelly 2011). In the Del Mar manzanita and its closest relatives, combined genetic and morphometric analysis could reveal whether circumscriptions based on morphology are reflected by genetic patterns.

We employed restriction site associated DNA (RAD) sequencing (Miller et al. 2007; Davey and Blaxter 2010) to develop a large single nucleotide

polymorphism (SNP) dataset for putative populations of Del Mar manzanita (including its type locality in the city of Del Mar, San Diego County) and a sample of its close relatives, especially those found in southern California and northern Baja California, Mexico. In recent years, RAD sequencing has become a common strategy to obtain genome-level information on genetic variants for use in both animal and plant conservation genetics (Allendorf et al. 2010; Hohenlohe et al. 2011; Lozier 2014; Jennings et al. 2016; Torres-Martínez and Emery 2016; Lozier and Zayed 2017).

We also collected vegetative morphometric data from leaves and stems to compare to genetic patterns. We then analyzed the morphological and genetic data, with the aim of testing whether morphological traits used to circumscribe Del Mar manzanita match genetic patterns. Overall, the results of this study will provide a genetic perspective for future management and conservation work on the Del Mar manzanita, and could have implications for the conservation, taxonomy, and systematics of other manzanitas.

MATERIALS AND METHODS

Taxon Sampling

For sampling of the Del Mar manzanita, we visited 14 locations in California and Baja California, Mexico (Table 1, Fig. 1, Appendix 1). In some large populations, we sampled from several subpopula-

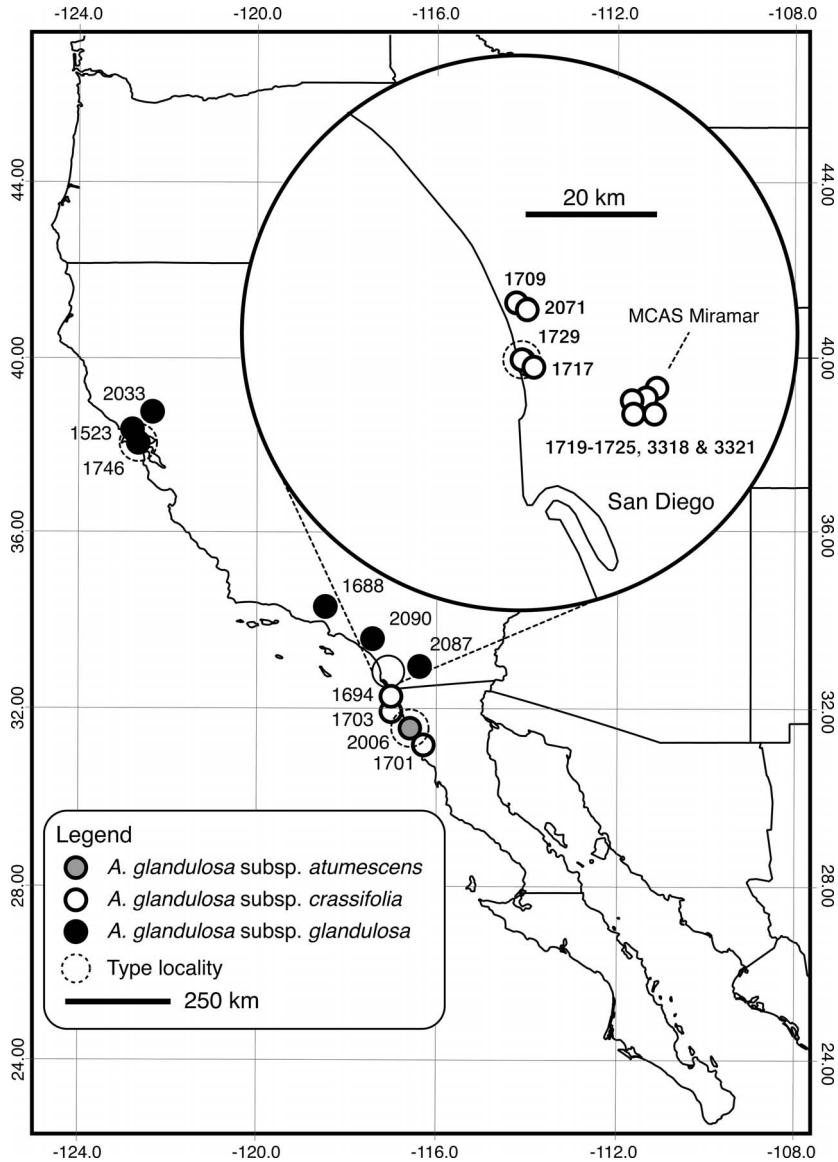


FIG. 1. Sampling map. Grid dimensions are in degrees of latitude and longitude (WGS84 datum). See Table 1 for more information on locales.

tions. To prevent confusion between populations and subpopulations, we refer to plant collecting locations as *locales*, which we define as groups of plants that are discontinuous from one another, with centers at least one km apart. In every case, the edges of the locales were at least 0.5 km apart. We were not able to obtain estimates of population size. However, the spatial density and size of populations/subpopulations vary dramatically across the range of the Del Mar manzanita (AECOM 2015), with consequences for population genetics.

We sampled intensively in central San Diego County, especially at the Marine Corps Air Station

(MCAS) Miramar, a federally owned property managed by the United States Department of Defense (Fig. 1). MCAS Miramar is thought to contain large populations of Del Mar manzanita, and is also a zone of potential contact between this taxon and Eastwood manzanita (Rebman and Dossey 2002).

To aid in determining the correspondence of genetics with morphometric taxon concepts, we sampled from the taxon thought to be most closely related to Del Mar manzanita, the Eastwood manzanita (six locales, Table 1, Fig. 1), and from *Arctostaphylos glandulosa* subsp. *atumescens*

J.E.Keeley, M.C.Vasey & V.T.Parker, a Mexican endemic that is the only subspecies of *A. glandulosa* other than the Del Mar manzanita and the Eastwood manzanita known to occur close to the coast in southern California and Baja California, Mexico (Table 1, Fig. 1). This latter species is a narrow endemic found only on Punta Banda Peninsula in Baja California, Mexico (Fig. 1), where it is the only manzanita. Thus, while this species does occur in the same region as Del Mar manzanita, they are not known or expected to come into close contact.

Here we use scientific names that are based on the morphology of the plants prior to DNA sequencing. Because the purpose of our study was to find out whether plants identified as Del Mar manzanita using morphological criteria could be reliably grouped together based on genetic similarity, we began by using names based on the morphology of the plants that we sampled (Table 1). Plants were identified according to the taxonomic keys and descriptions in the Jepson Manual (Parker et al. 2012), supplemented by relevant taxonomic revisions (Keeley et al. 2007). In both Parker et al. (2012) and Keeley et al. (2007), the presence of glandular hairs on leaves, twigs, and inflorescences, as well as color of the leaves and shape of the fruit are used to discern Del Mar manzanita from other taxa. Using this means of identification allowed us to begin with a hypothesis as to the correct identification of the plants, which was then tested using DNA. In the results presented below, we call attention to groups of plants in which the DNA results do not agree with the morphology-based names.

We took special care to sample multiple plants from the type localities of Del Mar manzanita (Del Mar, CA; Jepson 1922) and Eastwood manzanita (Mount Tamalpais, CA; Eastwood 1897). We included the type localities in the analysis so that we could make progress toward establishing the genetic limits of these taxa. We recognize that this is not ideal in the eyes of some, and that it might be preferable to sample DNA directly from the type specimens for such work. However, preliminary tests showed that DNA is not recoverable from older herbarium specimens in sufficient quantities to allow RAD DNA sequencing, the genomic method employed here. Therefore, sampling living plants at the type localities, rather than from the type specimens themselves, represents a next-best approach. This may also have the advantage of identifying relevant variation in morphology and genetics currently present in the population of the type locality.

For all sampled plants, we collected young leaves and flower buds, which were frozen on dry ice within 48 hours of collection. Up to 10 individuals were sampled from each locale, but not all collected individuals were sequenced (Table 1). We collected more individuals than required due to the low rate of successful DNA isolation; only high-quality samples were used.

Morphometrics

Morphologically, Del Mar manzanita is distinguished from other subspecies of *A. glandulosa* by qualitative characteristics of the reproductive and vegetative parts (Parker et al. 2012; Keeley et al. 2007; Table 2), many of which are subjective, particularly in the case of leaf color. Because most of the sites were visited in the spring, to obtain fresh flower buds for DNA extraction (see below), we were unable to collect material that included fruit and were therefore unable to use fruit characters in our morphometric analysis. We also did not collect data on leaf color. Though leaf color is often used to help distinguish subspecies of *A. glandulosa* from one another (Table 2), we found that it was difficult to assess this character in our dried herbarium specimens, due to differences in drying conditions. Instead, we focused on the hairs of the leaves and youngest stems (Table 2). If present at all, these hairs are typically in two layers, one of longer, stiff hairs that sometimes are tipped with glands, and another of shorter hairs that are typically appressed to the stem below the level of the long hairs.

For each of the specimens from which we obtained genetic data, we also collected data on the following traits: (A) *long hair density*, the density of long hairs on the young stem (0, absent; 1, sparse; 2, intermediate; 3, very dense), (B) *long hair type*, the texture of the long hairs (1, stiff; 2, soft), (C) *long hair glands*, the presence of glands at the tips of the long hairs (0, absent; 1, present), (D) *short hair presence*, the presence of short hairs below the layer of the long hairs (0, absent; 1, present), (E) *short hair type*, the type of the short hairs (1, stiff and erect; 2, soft and appressed), (F) *margin hairs*, the presence or absence of hairs on the margin of the leaf, (0, absent; 1, present), (G) *margin hair type*, type of the margin hairs, if present (1, long, glandular; 2, short, nonglandular; 3, stiff, long, nonglandular), (H) *quantitative leaf shape* (leaf length (mm) divided by width (mm); average for three mature leaves). Raw morphological data were deposited at the Dryad Digital Repository (Appendix S1; doi: 10.5061/dryad.kv573c7).

The vegetative morphometric data were treated in a multivariate framework. We visualized data using principal components analysis (PCA) and tree reconstruction, carried out in R, version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria). For PCA, we used the *prcomp* function in R to model the seven categorical variables and the single quantitative variable (H, leaf shape). Quantitative leaf shape was transformed into a Z-score before analysis. Also, due to the needs of the model, all individual plants with “missing” data (cases in which characters were not applicable to the individual in question) were excluded from analysis. The first two principal components were visualized in bivariate space to examine relationships. The contribution of each morphological character to the

TABLE 2. KEY MORPHOLOGICAL TRAITS TRADITIONALLY USED TO DISTINGUISH SUBSPECIES OF *A. GLANDULOSA*. Traits are taken from Keeley et al. (2007) and Parker et al. (2012).

| Taxon | Leaf color | Leaf vestiture | Nascent inflorescence shape | Nascent inflorescence bract vestiture | Fruit shape | Basal burl | Branchlet vestiture |
|--------------------------------|-------------------------|---------------------------|-----------------------------|---------------------------------------|---------------------------------------|------------|---------------------------------|
| <i>A. g. ssp. adamsii</i> | Strongly white-glaucous | Glabrous | Pendant | Non-glandular | Spheric to slightly depressed spheric | Present | Short-nonglandular |
| <i>A. g. ssp. atunescens</i> | Green to gray-green | Glabrous | Pendant | Glandular | Spheric to slightly depressed spheric | Absent | Long-glandular |
| <i>A. g. ssp. crassifolia</i> | Dark green | Glabrous to pubescent | Pendant | Non-glandular | Strongly depressed-spheric | Present | Short-nonglandular |
| <i>A. g. ssp. cushingiana</i> | Green to gray-green | Glabrous | Pendant | Non-glandular | Spheric to slightly depressed spheric | Present | Short-nonglandular |
| <i>A. g. ssp. erecta</i> | Yellowish green | Pubescent | Erect | Glandular | Spheric to slightly depressed spheric | Present | Short-nonglandular |
| <i>A. g. ssp. gabrielensis</i> | Lustrous green | Glabrous | Pendant | Non-glandular | Spheric | Present | Short-nonglandular |
| <i>A. g. ssp. glandulosa</i> | Green to gray-green | Glandular-puberulent | Pendant | Glandular | Spheric to slightly depressed spheric | Present | Both short- & long-glandular |
| <i>A. g. ssp. howellii</i> | Green to gray-green | Glandular to nonglandular | Pendant | Glandular | Spheric to slightly depressed spheric | Present | Short-nonglandular |
| <i>A. g. ssp. leucophylla</i> | Strongly white-glaucous | Glandular to puberulent | Pendant | Glandular | Spheric to slightly depressed spheric | Present | Both short- & long-nonglandular |
| <i>A. g. ssp. mollis</i> | Bright green | Glabrous | Pendant | Non-glandular | Spheric to slightly depressed spheric | Present | Both short- & long-nonglandular |

principal components was determined based on vector loadings. Based on preliminary results of the PCA analysis, it was determined that only A, C, D, and G contributed significantly to the model (absolute value of maximum vector loading on first and second principal component < 0.09).

Based on the seven categorical variables, we constructed a classification tree using the *rpart* R package, version 4.1-10. The *rpart* package implements recursive partitioning (Therneau and Atkinson 1997; De'ath 2002). In recursive partitioning, a tree linking all of the observations (individual plants in this case) is built by a simple process: a categorical variable is selected that "best" splits the individuals into two groups. This process is applied until the subgroups either reach a minimum size, or all individuals are grouped. We set the "minsplit" parameter of *rpart* to 2, the "minbucket" parameter of *rpart* to 1, and the "cp" parameter to 0.001. We did not prune the resulting tree, because cross-validation suggested that no improvement could be made.

In addition to recursive partitioning, we constructed a distance-based tree to visualize relationships among individual plants based on their overall dissimilarity across the categorical variables that we scored. We constructed a UPGMA (Unweighted Pair Group Method with Arithmetic Mean) tree using the *vegan* package of R, version 2.3-4. We used the *hclust* function of *vegan*, with 'method' set to average. Based on the preliminary results of the PCA, we excluded traits that did not contribute strongly to the preliminary PCA analysis, as these would likely add noise to the UPGMA tree; the UPGMA tree used traits A, C, D and G (see above).

DNA Sequencing & Variant Detection

Total genomic DNA was extracted from flower buds using the DNeasy Plant Mini Kit (Qiagen, Germantown, MD) according to the manufacturer's instructions. Total genomic DNA was checked for degradation on a 1% agarose electrophoresis gel; samples with high molecular-weight DNA (~20-50 kb) were standardized to 30 ng/ μ L (diluted to a volume of 150 μ L in TE) and sent for quality control, library construction, and sequencing at Floragenex, Inc. (Portland, OR). Methods for this process generally followed those described by Lozier (2014), with the exception that our work used the restriction enzyme PstI (Hipp et al. 2014). In brief, sequence identifier barcodes (a unique one for each plant) and sequence adapters were added to genomic DNA after digestion by the endonuclease PstI. The barcoded samples were then combined and the fragments sequenced outwards from restriction-sites using single-end DNA sequencing reads that were 100 base pairs long. All DNA sequencing was done on a HiSeq 2500 DNA sequencing instrument (Illumina, San Diego, CA). Following sequencing, DNA sequences from individual plants were separated

based on their unique barcodes using the program *fastq-multx* (Aronesty 2013); the barcodes were then removed from the sequences.

Due to the lack of an existing reference genome for *Arctostaphylos* or a closely related member of the Ericaceae, detection of variation among individual plants was done using a "RAD reference" approach (Lozier 2014), in which a kind of reference genome is constructed using sequence data from the plant with the greatest number of unique RAD clusters. We employed this strategy to ensure the best reference genome "target" for subsequent alignment of sequences from other plants. In developing the RAD reference, custom methods developed by Floragenex (Lozier 2014) were used to cluster identical sequences that had 5–500 \times sequencing coverage, which produced a preliminary assembly. The assembly was then collapsed back to separate sequences and these sequences were realigned against the preliminary RAD reference genome using the program BWA [Li and Durbin 2009; *aln function, edit distance (-n) 3, -N (disable iterative search)*], allowing at most four mismatches among reads within a cluster. The purpose of self-alignment was to identify and remove repetitive DNA regions.

Reads for each individual plant were aligned to the RAD reference using BOWTIE (Langmead et al. 2009), relying on sequence quality information to aid in the process of match-making, allowing a maximum of three mismatched bases per read, and permitting alignment of each read to no more than one region of the reference. SAMTOOLS (Li et al. 2009) was used to detect SNPs and call genotypes. Filters for SNP calling required a minimum phred score of 20, a minimum of 15 \times sequence coverage and a maximum of 10% missing data across samples.

Following SNP calling, we also excluded: 1) all positions in which data were missing for any individual plant (Arnold et al. 2013; Davey et al. 2013), 2) all invariant sites (where every individual had the same allele call), and 3) all sites in which more than two allelic variants were detected. The final filtering step restricted the dataset to only biallelic SNPs, thus rendering the data diploid. We used this filter because most population genetic software are not able to deal with polyploid data (where more than two allelic variants are allowed per locus). By rendering our data as biallelic SNPs, we are able to apply population genetic methods that assume diploid loci. We recognize that this is not ideal, as the underlying genetics of the organisms is tetraploid. We think that this filtering method is the best option given the limitations of available SNP calling methods. We also note that this method of dealing with polyploids in population genetic analyses has been employed in other studies (Qi et al. 2015; reviewed by Dufresne et al. 2014). The rendering of loci as diploid when the underlying genetics is polyploid can lead to conflicts with assumptions of software, mainly due to the issue of calculating allele dosage in polyploids (Dufresne et

al. 2014). Methods are available to deal with polyploids in genomic data by calling alleles using a draft reference genome (Garrison and Marth 2012). Unfortunately, the lack of a manzanita reference genome makes it impossible to apply such methods at this time.

The data resulting from the above filtering process was used to create two datasets, one using all of the *A. glandulosa* samples (54 individuals; Table 1), and the second focused on just the locales classified as Del Mar manzanita, and for which we had more than one sample (39 individuals; Table 1). Hereafter, we refer to the former as the “Complete” and to the latter as the “Reduced” dataset. In both datasets, we excluded invariant sites, as well as sites with a minor allele frequency of less than 0.02 (Bradbury et al. 2007).

Population Genetics

To objectively identify statistically meaningful groups of individuals, we used a combination of parametric and non-parametric approaches. These approaches were applied in exactly the same way to all datasets. For the non-parametric approach, we carried out principal coordinates analysis (PCoA) using an identity by site genetic distance matrix. For the parametric approach, we used the program STRUCTURE version 2.3.4 (Pritchard et al. 2000; Falush et al. 2003, 2007). For each STRUCTURE run, we used 50,000 Markov chain Monte Carlo replicates, applying an admixture ancestry model and assuming diploid loci. In each of the runs, the first 10,000 replicates were discarded as “burnin”, a method that reduces noise in the final sample by excluding samples that were taken while the model was still unstable (Hubisz et al. 2009). For each dataset, we did 10 replicate runs at each level of K from one to thirteen. The appropriate K for each dataset was inferred using the software program STRUCTURE HARVESTER (Earl and vonHoldt 2012), according to the method of Evanno et al. (2005).

Morphology Versus Genetics

To test whether genetic patterns match groupings of plants based on morphology, we used a Mantel test, carried out in the R package *vegan* (version 2.3-4). For the genetic data, we used the same identity by state distance matrix used for PCoA (see above). For the morphological data, we created a Euclidean distance matrix using the eight traits used for the morphometric analyses described above. We ran a partial Mantel test using both Pearson and Spearman correlation coefficients, and controlling for geographic distance.

RESULTS

Vegetative Morphometric Patterns

Patterns of morphological variation across subspecies, locales, and individuals were complex, as

suggested by visualizations of multivariate analysis (Fig. 2A). In PCA, preliminary analysis suggested that only four variables contributed strongly to the model (margin hair type, short hair type, long hair density, long hair glands), and so only these variables were retained for the final analysis. In the final model (Fig. 2A), the first two principal components (those plotted) account for more than 84% of the variance in the data. Long hair density is strongly negatively correlated with the first principal component (vector loading: -0.93), followed closely by margin hair type, which is positively correlated (vector loading: 0.32). Long hair glands and margin hair type both contribute very strongly to the second principal component, but in opposite directions (vector loadings of 0.62 and -0.68, respectively).

The classification tree (Fig. 2B) suggests some of the same patterns as the PCA, indicating that a large number of putative Del Mar manzanita individuals, including some from the type locality, have a lower density of long hairs than other sampled plants. As in the PCA, other sampled individuals are also differentiated by margin hair type and short hair presence; with the exception of long hair glands, the most informative characters are the same in both the PCA and the classification tree.

The UPGMA tree (Fig. 2C) allows for examination of relationships among individual plants. This tree shows that most groupings of individual plants do not conform to expectations based on the morphology of the plants, or even the locality where they were collected (Table 1). Subspecies do not form cohesive groups at any level of the tree, and in many cases individuals from the same locale are not cohesive. Finally, collections from the type locality of Del Mar manzanita are not strongly cohesive and are found in multiple portions of the tree. Overall, these results demonstrate the morphological heterogeneity of the sample, with morphologically distinct plants co-occurring at the same locale.

Genetic Variants

A total of 54 samples were successfully sequenced (Table 1). Quality was high for all samples; none were excluded due to low DNA sequencing coverage or excessive missing data. Sequencing rates averaged 6,903,818 reads per individual ($SD = 3,184,037$). For information on sequencing coverage for each plant, see Appendix S2 (Dryad Digital Repository). Raw DNA sequence data is on the NCBI BioProject database under accession PRJNA396085.

A total of 16,083,660 reads were available for the sample used to construct the RAD reference genome (D. Burge 2071_4, *A. glandulosa* subsp. *crassifolia*, Table 1). 15,606,332 reads passed quality filters and were assembled into 25,321 contigs 92 of bp in length, for an assembled RAD reference genome length of 2,329,532 bp.

Using the RAD reference genome as a guide, we called a total of 163,793 loci. The complete variant

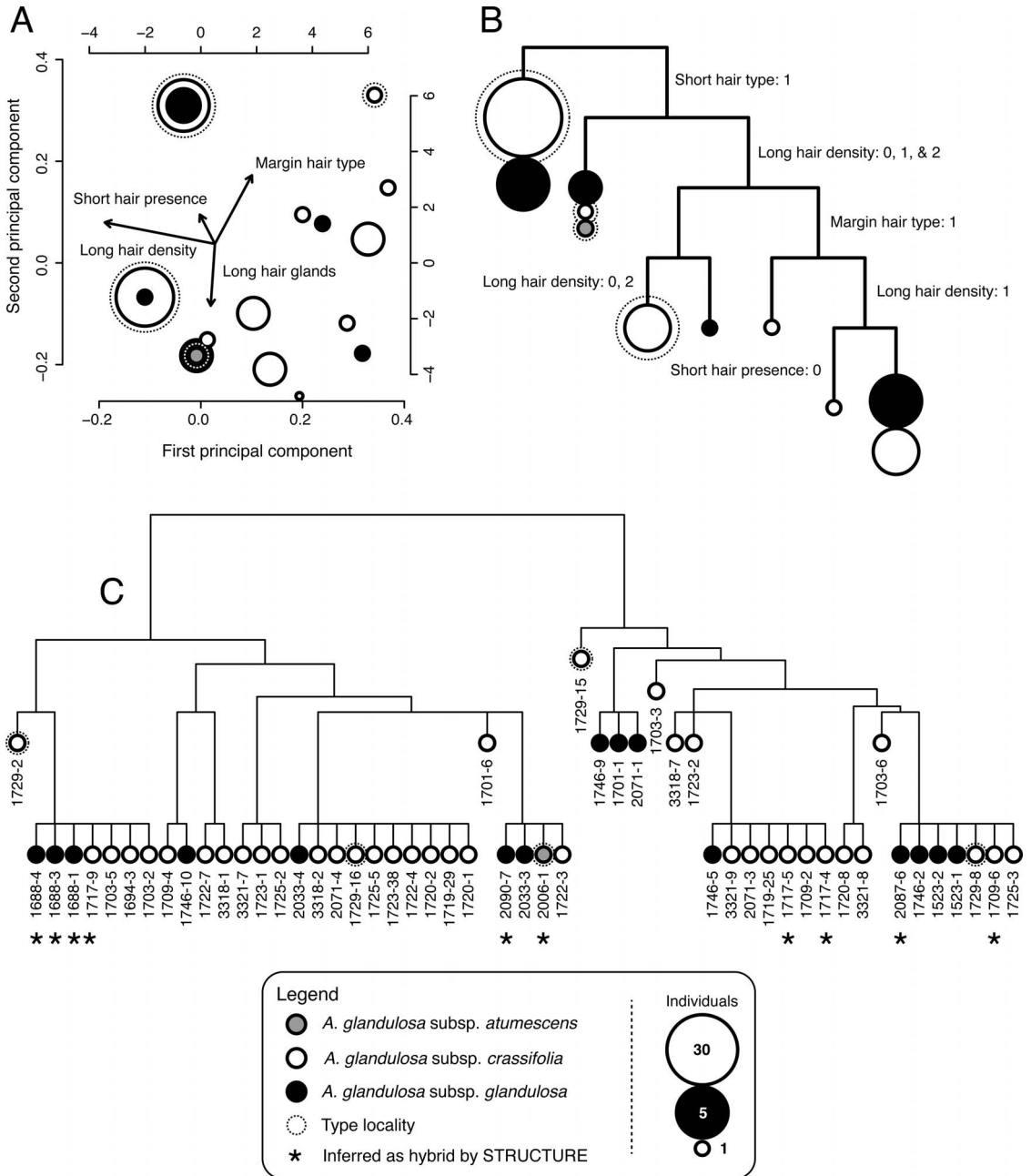


FIG. 2. Results of morphometric analysis. A, results of principal components analysis (PCA); size of circle indicates the number of individuals. B, classification tree for manzanita individuals; character states at left define all individuals below that branch; the character 'long hair density' appears in this tree twice because some sub-groups of plants are distinguished by very dense hairs (3) and some by sparse hairs (1). C, UPGMA tree for manzanita individuals; branch length represents morphological distance.

set is available at the Dryad Digital Repository (Appendix S3; doi: 10.5061/dryad.kv573c7). After the application of post-variant calling filters, the Complete dataset contained 60,865 loci for 54

individuals; the Reduced dataset contained 53,850 loci for 39 individuals. These datasets are available at the Dryad Digital Repository (Appendices S4 & S5; doi:10.5061/dryad.kv573c7).

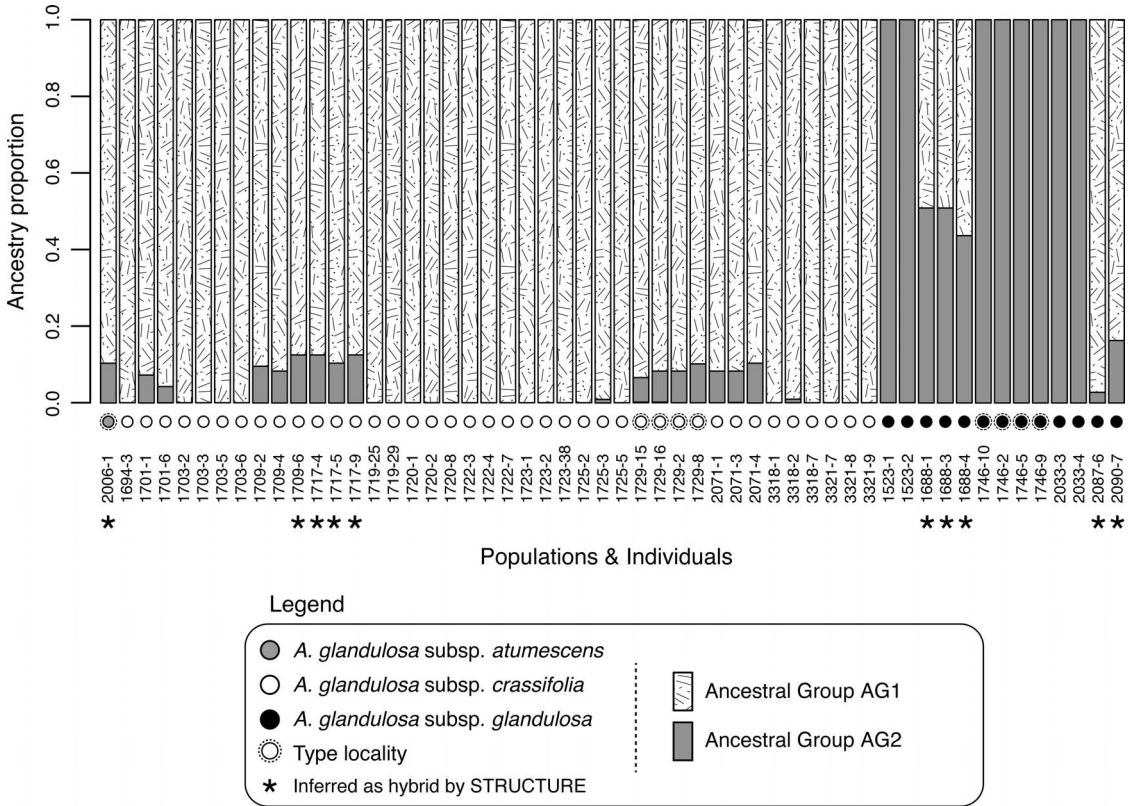


FIG. 4. Proportional genetic ancestry based on analysis of all SNP data using STRUCTURE. Results of analysis with the program STRUCTURE, assuming two ancestral groups (see Methods). Under the assumption that two ancestral groups are present, STRUCTURE estimated the proportion of genetic variation in each sampled individual that was assignable to each of these groups. Proportions are indicated by shaded columns.

Ancestral Groups AG1 and 2). STRUCTURE results also suggest some genetic admixture between these ancestral groups, represented by six plants that were inferred to be of hybrid origin ($P \ll 0.01$; Fig. 4, 1719-25, 1719-29, 1722-3, 1722-4, 1722-7, and 1725-3). All of these plants are from MCAS Miramar.

Genetics Versus Morphology

The Mantel test using the Pearson correlation coefficient indicated a significant relationship between morphology and genetics (Mantel statistic $r = 0.094$; $P = 0.005$), as did the test using the Spearman correlation coefficient (Mantel statistic $r = 0.069$; $P = 0.022$).

DISCUSSION

Does Morphology Predict Genetic Relationships?

Overall, our Mantel test results suggest that morphology is a reliable predictor of the underlying genetic groups. However, there is a very weak association between genetic groups and the names

that were assigned to the sampled plants based on current morphological circumscriptions of *A. glandulosa* subspecies (Figs. 3 and 4). This lack of correspondence suggests that the current circumscriptions should be modified. Unfortunately, our sampling of *A. glandulosa* subspecies and populations is not broad enough to provide the evidence necessary to support such changes. Future studies should aim to expand on the sampling employed in this study, ideally including a geographically broad sample of *A. glandulosa* populations from all of the known subspecies.

Our morphological analyses did not use fruit characters, due to the time of year in which we sampled. Fruit and seed characters are frequently used in manzanita systematics, and some putative Del Mar manzanita populations are known to have a distinctive (within *A. glandulosa*), flattened fruit (Parker et al. 2012), especially near the type locality, close to the sea (V. T. Parker, San Francisco State University, personal observation). Future research should aim to measure a broader suite of characters than those analyzed in this study, especially those relating to fruit and inflorescence morphology.

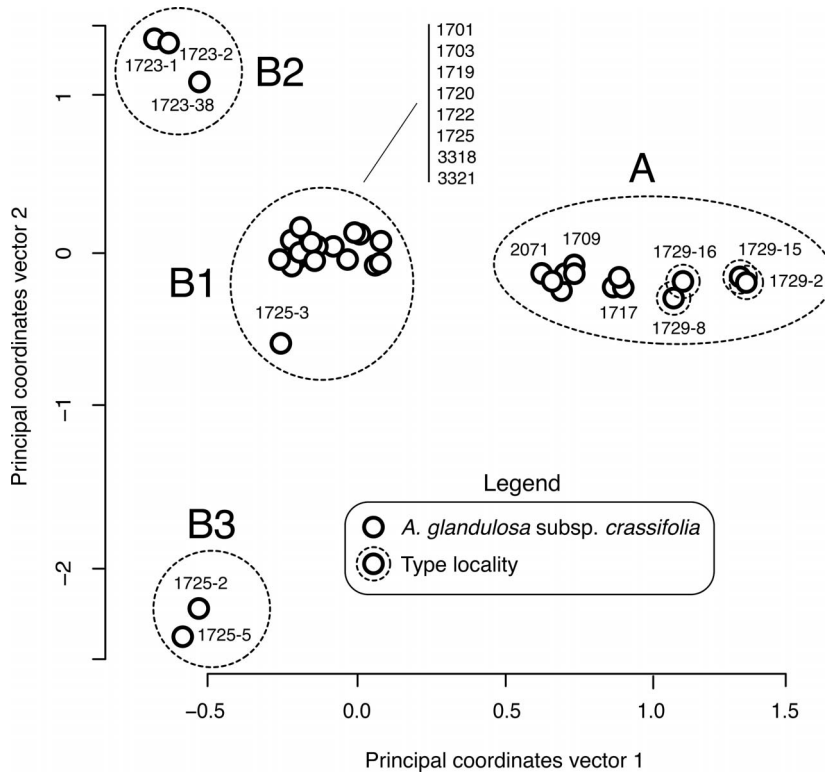


FIG. 5. Genetic relationships based on PCoA of SNP data for only putative Del Mar manzanita. Individuals from the type locality of Del Mar manzanita (1729; Del Mar) are indicated with a dashed halo. For clarity of presentation, some of the plants are not labeled individually.

How Widespread is Del Mar Manzanita?

At the broadest scale, our genetic data suggest that the Del Mar manzanita forms part of a very widespread genetic lineage (Fig. 4, AG1) found mainly in southern California and northern Baja California, Mexico. This result argues in favor of a greatly expanded circumscription of Del Mar manzanita. In this case, Del Mar manzanita would be the dominant burl-forming manzanita in coastal portions of southern California and northern Baja California, Mexico (Fig. 4, AG1). On the other hand, the more focused analysis of only the samples from coastal San Diego County suggests that plants from the type locality and nearby locations close to the coast fall into one genetic group (Fig. 6, AGC1), while plants from more distant locations, especially more inland locations like MCAS Miramar, do not form part of this group. This result suggests that a more restrictive view of Del Mar manzanita might be possible; the circumscription could be modified to include only those plants from the second analysis that fall into the same genetic group as the plants from the type locality (Fig. 6, AGC1), which are from sandstone derived soils very close to the sea. Unfortunately, we did not obtain a broad enough sample of *A. glandulosa* subspecies and populations

to reliably test taxonomic concepts in the group based on DNA, and so we are not able to modify circumscriptions according to genetic patterns. More research is needed to determine taxonomic limits in the subspecies of *A. glandulosa*. Ideally, future research will expand on the present work by adding samples of more subspecies and populations from especially from coastal San Diego County and northern Baja California, Mexico.

Conservation and Recovery Implications

Management of rare plants demands consistent, objective tools to identify these plants to the exclusion of other taxa. Identification is usually done using morphological features. However, many rare plants are difficult to identify morphologically, or only display their diagnostic features during a particular phenological period (e.g., during flowering or fruiting). There are a variety of methods available that seek to overcome this limitation using genetic markers (reviewed by Hollingsworth et al. 2016). However, many of these techniques, particularly DNA barcoding methods (Kress et al. 2005), fall short in groups with low levels of genetic variation. This situation limits objective identification of rare plants, which in turn limits management, where

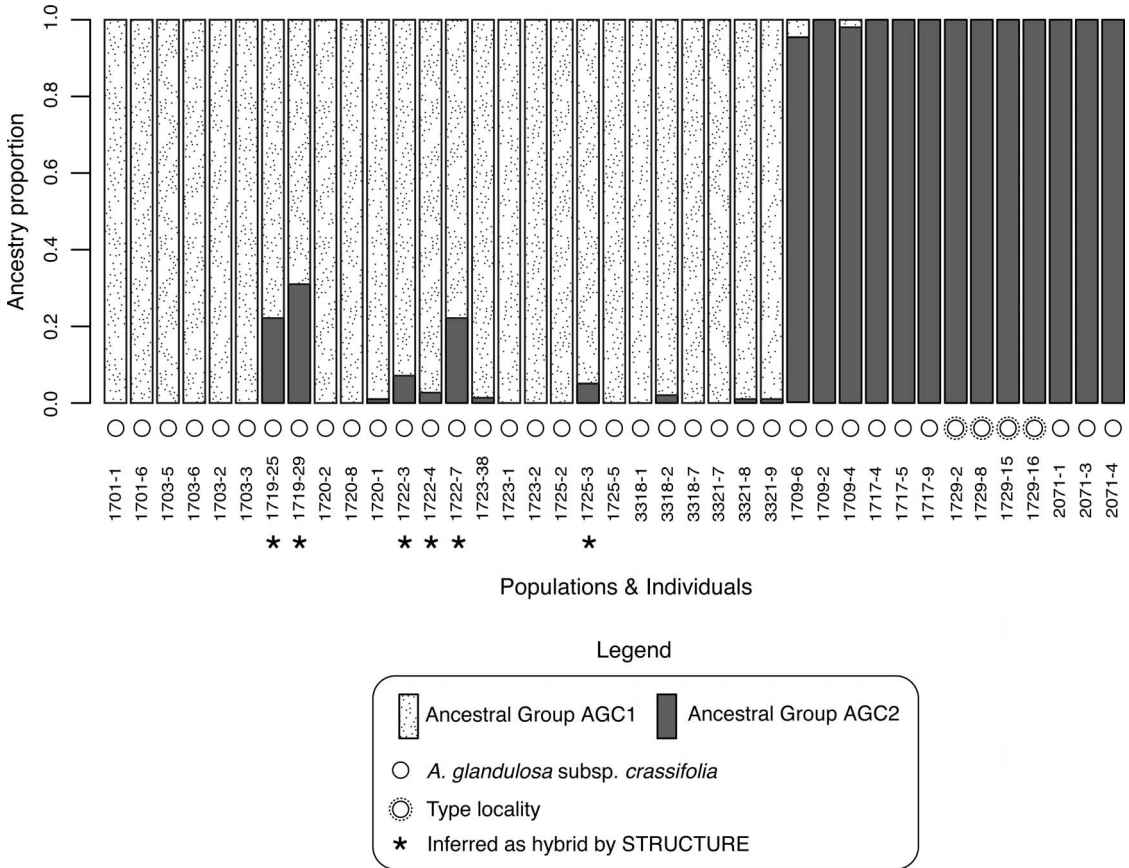


FIG. 6. Proportional genetic ancestry for putative Del Mar manzanitas based on STRUCTURE. Results of analysis with the program STRUCTURE, assuming two ancestral groups (see Methods). Under the assumption that two ancestral groups are present, STRUCTURE estimated the proportion of genetic variation in each sampled individual that was assignable to each of these groups. Proportions are indicated by shaded columns. Individuals marked with an asterisk are considered to be meaningfully admixed at the $P < 0.05$ threshold.

precise identification of rare plants is critical for effective conservation. As demonstrated here and in other recent work (Hollingsworth et al. 2016; Andrews et al. 2016), high-throughput DNA sequencing, particularly RADseq, can be used to target large numbers of variable SNPs without any prior knowledge of the SNPs or the genome of the target organism, allowing for the rapid development of useful panels of genetic markers for precise and repeatable conservation work.

Our results show that in the case of Del Mar manzanita and its close relatives, there is a mismatch between genetic groups and the groups based on current taxonomic concepts. As we explained above, our genetic results could be used equally effectively to argue in favor of either a greatly reduced or a greatly expanded circumscription of Del Mar manzanita. Clearly, these two alternatives have dramatically different implications for conservation and recovery; if we were to choose the first option, this already rare

taxon would become even more rare; if we were to choose the second option, the taxon would become very widespread, and would probably not require conservation measures. As outlined above, future research should expand upon our work by sampling more populations and taxa, in order to thoroughly test the circumscriptions of the *A. glandulosa* subspecies. Such work will lead to conservation outcomes that are consistent with the underlying genetics of the plants being conserved, rather than the names applied to them.

ACKNOWLEDGMENTS

Assistance with field work was provided by A. Ash, C. Grassa, J. Keeley, and M. Schwager. This research was primarily funded by the United States Marine Corps on behalf of the MCAS Miramar Environmental Management Department. Additional funding was provided by the JiJi Foundation.

LITERATURE CITED

- AECOM, INCORPORATED. 2015. Upland endangered plant species census and monitoring for Del Mar Manzanita, MCAS Miramar, Final Report (Contract Number N62470-13-D-8017). Prepared for Marine Corps Air Station Miramar, Natural Resources Division. San Diego, CA.
- ALLENDORF, F. W., P. A. HOHENLOHE AND G. LUIKART. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics* 11:697–709.
- ANDREWS, K. A., J. M. GOOD, M. R. MILLER, G. LUIKART, AND P. A. HOHENLOHE. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics* 17:81–92.
- ARNOLD, B. R., B. CORBETT-DETIG, D. HARTL, AND K. BOMBLIES. 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Molecular Ecology* 22:3179–90.
- ARONESTY, E. 2013. Comparison of sequencing utility programs. *The Open Bioinformatics Journal*. DOI:10.2174/1875036201307010001.
- BRADBURY, P. J., Z. ZHANG, D. E. KROON, T. M. CASSTEVENS, Y. RAMDOSS, AND E. S. BUCKLER. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23:2633–2635.
- DAVEY, J. W., AND BLAXTER, M. L. 2010. RADSeq: next-generation population genetics. *Briefings in Functional Genomics* 9:416–423.
- , T. CEZARD, P. FUENTES-UTRILLA, C. ELAND, K. GHARBI, AND M. L. BLAXTER. 2013. Special features of RAD Sequencing data: implications for genotyping. *Molecular Ecology* 22:3151–3164.
- DE'ATH, G. 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* 83:1105–1117.
- DUFRESNE, F., M. STIFT, R. VERGILINO, AND B. K. MABLE. 2014. Recent progress and challenges in population genetics of polyploid organisms: an overview of current state-of-the-art molecular and statistical tools. *Molecular Ecology* 23:40–69.
- EARL, D. A., AND B. M. VONHOLDT. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361.
- EASTWOOD, A. 1897. Studies in the herbarium and the field. I. *Proceedings of the California Academy of Sciences* 3:71–88.
- EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14:2611–2620.
- FALUSH, D., M. STEPHENS, AND J. K. PRITCHARD. 2003. Inference of population STRUCTURE using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
- , ———, AND ———. 2007. Inference of sampling locale STRUCTURE using multilocus genotype data: Dominant markers and null alleles. *Molecular Ecology Notes* 7:574–578.
- FEDERAL REGISTER. 1996. Rules and Regulations—Endangered and Threatened Wildlife and Plants; Determination of Endangered or Threatened Status for Four Southern Maritime Chaparral Plant Taxa from Coastal Southern California and Northwestern Baja California, Mexico. *Federal Register* 61:52370.
- FULTON, R. E. AND L. CARPENTER. 1979. Pollination, reproduction, and fire in California *Arctostaphylos*. *Oecologia* 38:147–157.
- GARRISON, E., AND G. MARTH. 2012. Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907.
- HIPP, AL. L., D. A. R. EATON, J. CAVENDER-BARES, E. FITZEK, R. NIPPER, AND P. S. MANOS. 2014. A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE* 9: e93975.
- HOGAN, D. C., J. O. SAWYER, AND C. SAUNDERS. 1996. Southern maritime chaparral. *Fremontia* 24:3–7.
- HOHENLOHE, P. A., S. J. AMISH, J. M. CATCHEN, F. W. ALLENDORF, AND G. LUIKART. 2011. Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources* 11:117–122.
- HOLLINGSWORTH, P. M., D.-Z. LI, M. VAN DER BANK, AND A. D. TWYFORD. 2016. Telling plant species apart with DNA: from barcodes to genomes. *Philosophical Transactions of the Royal Society B* 371:20150338.
- HUBISZ, M. J., D. FALUSH, M. STEPHENS, AND J. K. PRITCHARD. 2009. Inferring weak population STRUCTURE with the assistance of sample group information. *Molecular Ecology Resources* 9:1322–1332.
- JENNINGS, H., K. WALLIN, J. BRENNAN, A. DEL VALLE, A. GUZMAN, D. HEIN, S. HUNTER, A. LEWANDOWSKI, S. OLSON, H. PARSONS, S. SCHEIDT, Z. WANG, A. WERRA, R. Y. KARTZINEL, AND T. J. GIVNISH. 2016. Inbreeding, low genetic diversity, and spatial genetic structure in the endemic Hawaiian lobeliads *Clermontia fauriei* and *Cyanea pilosa* ssp. *longipedunculata*. *Conservation Genetics Resources* 8:145–158.
- JEPSON, W. L. 1922. Revision of the California species of the genus *Arctostaphylos*. *Madroño* 1:76–96.
- KEELEY, J. E., M. C. VASEY, AND V. T. PARKER. 2007. Subspecific variation in the widespread burl-forming *Arctostaphylos glandulosa*. *Madroño* 54:42–62.
- KELLY, R. P. 2011. The use of population genetics in Endangered Species Act listing decisions. *Ecology Law Quarterly* 37:1107–1159.
- KRESS, J. W., K. J. WURDACK, E. A. ZIMMER, L. A. WEIGT, AND D. H. JANZEN. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences* 102:8369–8374.
- LANGMEAD, B., C. TRAPNELL, M. POP, AND S. L. SALZBERG. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10:R25.
- LI, H. AND R. DURBIN. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25:1754–1760.
- , B. HANDSAKER, A. WYSOKER, T. FENNELL, J. RUAN, N. HOMER, G. MARTH, G. ABECASIS, R. DURBIN, AND THE 1000 GENOME PROJECT DATA PROCESSING SUBGROUP. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.
- LOZIER, J. D. 2014. Revisiting comparisons of genetic diversity in stable and declining species: assessing genome-wide polymorphism in North American bumblebees using RAD sequencing. *Molecular Ecology* 23:788–801.
- AND A. ZAYED. 2017. Bee conservation in the age of genomics. *Conservation Genetics* 18:713–729.
- MILLER, M. R., J. DUNHAM, A. AMORES, W. CRESKO, AND E. JOHNSON. 2007. Rapid and cost-effective

- polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* 17:240–288.
- OGDEN, R., N. DAWNAY, AND R. MCEWING. 2009. Wildlife DNA forensics—bridging the gap between conservation genetics and law enforcement. *Endangered Species Research* 9:179–195.
- PARKER, V. T., M. C. VASEY, AND J. E. KEELEY. 2012. *Arctostaphylos*. In B. G. Baldwin et al. (eds.), *The Jepson Manual: Vascular Plants of California*. University of California Press, Berkeley, California.
- PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of sampling locale STRUCTURE using multilocus genotype data. *Genetics* 155:945–959.
- QI, Z.-C., Y. YU, X. LIU, A. PAIS, T. RANNEY, R. WHETTEN, AND Q.-Y. XIANG. 2015. Phylogenomics of polyploid *Fothergilla* (Hamamelidaceae) by RAD-tag based GBS—insights into species origin and effects of software pipelines. *Journal of Systematics and Evolution* 53:432–447.
- REBMAN, J. AND R. DOSSEY. 2002. Final Report. East Miramar, Focused Rare and Endangered Plant Survey, Marine Corps Air Station Miramar, Training Areas 1-4. San Diego Natural History Museum. Prepared for Marine Corps Air Station Miramar.
- RUDI, K., M. KROKEN, O. J. DAHLBERG, A. DEGGERDAL, K. S. JAKOBSEN, AND F. LARSEN. 1997. Rapid, universal method to isolate PCR-ready DNA using magnetic beads. *Biotechniques* 22:506–511.
- SEMARNAP (SECRETARIA DE MEDIA AMBIENTE, RECURSOS NATURALES Y PESCA). 2002. Norma Oficial Mexicana NOM-059-ECOL-20011, Protección ambiental-Especies nativas de México de flora y fauna silvestre-Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio-Lista de especies en riesgo. *Diario Oficial de la Nación*, 6 de Marzo del 2002. México, D.E.
- THERNEAU, T. M. AND E. J. ATKINSON. 1997. An introduction to recursive partitioning using the rpart routines. *Division of Biostatistics* 61, Mayo Clinic.
- TORRES-MARTÍNEZ, L. AND N. C. EMERY. 2016. Genome-wide SNP discovery in the annual herb, *Lasthenia fremontii* (Asteraceae): genetic resources for the conservation and restoration of a California vernal pool endemic. *Conservation Genetics Resources* 8:145–158.
- WELLS, P. V. 1969. The relation between mode of reproduction and extent of speciation in woody genera of the California chaparral. *Evolution* 23:264–267.
- UNITED STATES FISH AND WILDLIFE SERVICE. 2010. *Arctostaphylos glandulosa* subsp. *crassifolia* (Del Mar manzanita), 5-Year Review: Summary and Evaluation. Carlsbad Fish and Wildlife Office, Carlsbad, California.

APPENDIX I

DETAILED COLLECTION INFORMATION FOR SAMPLED LOCALES

For each sampled locale (Table 1), the format is as follows: collector name and number (herbarium of voucher deposition), description of locality, (GPS coordinates), political region (county and state in the case of California; country and state in the case of Mexico).

Arctostaphylos glandulosa subsp. *atumescens*—D. Burge 2006 (SD), 18 Jan 2016, Cerro Buenavista (GPS [NAD84]: 31.6737, -116.6314), Baja California, Mexico.

Arctostaphylos glandulosa subsp. *crassifolia*—D. Burge 1694 (DAV), 18 Mar 2015, Cerro del Coronel (GPS [NAD84]: 32.283, -116.93), Baja California, Mexico. D. Burge 1701 (DAV), 19 Mar 2015, Cañon San Isidro (GPS [NAD84]: 31.292, -116.3434), Baja California, Mexico. D. Burge 1703 (DAV), 20 Mar 2015, Mesa de Descanso (GPS [NAD84]: 32.1718, -116.8898), Baja California, Mexico. D. Burge 1709 (DAV), 21 Mar 2015, Encinitas Community Center (GPS [NAD84]: 33.0444, -117.2669), San Diego Co., CA. D. Burge 1717 (DAV), 3 Apr 2015, Torrey Pines (GPS [NAD84]: 32.9406, -117.2471), San Diego Co., CA. D. Burge 1719 (DAV), 4 Apr 2015, MCAS 1 (GPS [NAD84]: 32.9164, -117.0398), San Diego Co., CA. D. Burge 1720 (DAV), 4 Apr 2015, MCAS 2 (GPS [NAD84]: 32.8897, -117.064), San Diego Co., CA. D. Burge 1722 (DAV), 4 Apr 2015, MCAS 3 (GPS [NAD84]: 32.8787, -117.0659), San Diego Co., CA. D. Burge 1723 (DAV), 4 Apr 2015, MCAS 4 (GPS [NAD84]: 32.8649, -117.069), San Diego Co., CA. D. Burge 1725 (DAV), 4 Apr 2015, MCAS 5 (GPS [NAD84]: 32.8932, -117.0757), San Diego Co., CA. D. Burge 1729 (DAV), 4 Apr 2015, Crest Canyon (GPS [NAD84]: 32.9501, -117.2538), San Diego Co., CA. D. Burge 2071 (SD), 21 Feb 2016, Encinitas (GPS [NAD84]: 33.036, -117.2487), San Diego Co., CA. M. Mulligan 3318 (SD), 12 Feb 2016, MCAS 18 (GPS [NAD84]: 32.8938, -117.0398), San Diego Co., CA. M. Mulligan 3321 (SD), 12 Feb 2016, MCAS 17 (GPS [NAD84]: 32.8915, -117.0516), San Diego Co., CA.

Arctostaphylos glandulosa subsp. *glandulosa*—D. Burge 1523 (DAV), 23 Mar 2014, Cavedale Road (GPS [NAD84]: 38.3624, -122.4719), Sonoma Co., CA. D. Burge 1688 (DAV), 17 Mar 2015, Newhall Pass (GPS [NAD84]: 34.3471, -118.5102), Los Angeles Co., CA. D. Burge 1746 (DAV), 5 Jun 2015, Mount Tamalpais (GPS [NAD84]: 37.911, -122.5775), Marin Co., CA. D. Burge 2033 (DAV), 31 Jan 2016, Bolinas Ridge (GPS [NAD84]: 37.9468, -122.6665), Marin Co., CA. D. Burge 2087 (SD), 22 Feb 2016, Viejas Mountain (GPS [NAD84]: 32.854, -116.7421), San Diego Co., CA. D. Burge 2090 (DAV), 23 Feb 2016, Santa Ana Mountains (GPS [NAD84]: 33.6535, -117.4455), Orange Co., CA.