Pronounced Genetic Separation among Varieties of the Primula cusickiana Species Complex, a Great Basin Endemic

Authors: Koontz, Austin, Pearse, William D., and Wolf, Paul G.

Source: Systematic Botany, 47(3) : 887-897

Published By: The American Society of Plant Taxonomists

URL: https://doi.org/10.1600/036364422X16573019348175

Systematic Botany (2022), 47(3): pp. 887–897
ISSN 0363-6445 (print); ISSN 1548-2324 (online)

DOI 10.1600/036364422X16573019348175
Date of publication September 16, 2022

Pronounced Genetic Separation Among Varieties of the *Primula cusickiana* Species Complex, a Great Basin Endemic

Austin Koontz,[1,2,5] William D. Pearse,[2,3] and Paul G. Wolf[4]

[1]Center for Tree Science, Morton Arboretum, 4100 IL-53, Lisle, Illinois 60532, USA; akoontz@mortonarb.org
[2]Department of Biology, Utah State University, Logan, Utah 84322, USA
[3]Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK
[4]Department of Biological Sciences, University of Alabama in Huntsville, Huntsville, Alabama 35899, USA
[5]Author for correspondence

Communicating Editor: Diana Jolles

Abstract—Distinguishing between populations with strong genetic structure and unique species is a common challenge in systematics, especially for taxa occurring in fragmented habitats where allopatric speciation may be widespread and distinct groups may be morphologically similar. Such is often the case with species complexes across sky island environments. In these scenarios, biogeography may help to explain the taxonomic relations between species complex members, and restriction site-associated DNA (RAD) sequencing methods are commonly used to compare closely related taxa across thousands of loci. Here we use RADseq to clarify the boundaries separating the geographically distinct but morphologically similar varieties of the *Primula cusickiana* species complex, and to contextualize past findings of strong genetic structure among populations within varieties. Our genetic analyses demonstrate pronounced separation between isolated populations of this Great Basin endemic, indicating that the current varietal classification of complex members is inaccurate, and emphasizing their conservation importance. We discuss how these results correspond to recent biogeographical models used to describe the distribution of other sky island taxa in western North America. Our findings also fit into a wider trend observed for alpine *Primula* species complexes, and we consider how edaphic specialization and heterostylous breeding systems may be contributing to frequent diversification via allopatric speciation in this genus.

Keywords—Allopatry, biogeography, cryptic speciation, edaphic endemism, heterostyly, *Primula*, RADseq, sky island.

A canonical driver of biological diversification is allopatry, whereby geographic barriers lead to population isolation and, eventually, speciation. Sky islands are places where sharp changes in elevation lead to pronounced ecological differences over relatively short distances, providing the types of barriers required for allopatric speciation. Historically, climatic fluctuations have determined the presence and distribution of sky island environments for mountain ranges across the world, and this in turn is reflected by the genetic patterns seen in montane species today (Hewitt 1996). However, in this biogeographic context, distinguishing between closely related species and genetically structured populations may prove challenging (Huang 2020), especially if similar niches across mountain ranges maintain phenotypic similarities (e.g. Yang et al. 2019). Additionally, in the short-term, genetic patterns will be influenced by particular aspects of a species' biology, such as dispersal and breeding systems, which may facilitate or hinder reproductive isolation between genetically distinct entities. Here, we examine the genetic relations between the sky island populations of members of the *Primula cusickiana* species complex, a group of plants endemic to the Great Basin region of the western United States.

The *Primula cusickiana* species complex (Holmgren and Kelso 2001) is a group of herbaceous, perennial plants that fall within *Primula* section *Parryi* (Wright Smith and Fletcher 1949). When including the phylogenetically nested *Dionysia* Fenzl., *Dodecatheon* L., and *Cortusa* L. (Trift et al. 2002; Schmidt-Lebuhn et al. 2012), the genus is composed of roughly 500 species, distributed worldwide but largely found in temperate and montane habitats in the northern hemisphere, with the Himalayas as the center of diversity (Richards 2003). *Primula* section *Parryi* (five species, eight taxa; Kelso et al. 2009) is one of 37 sections of the genus (Richards 2003) and one of three parts of *Primula* subgenus *Auriculastrum* Schott (Schott 1852), which also includes *Primula* sections *Auricula* Duby and *Auneifolia* Balfour (Zhang and Kadereit 2004), as well as *Dodecatheon* L. (Mast and Reveal 2007) and *Primula suffrutescens* (Gray 1888). *Primula* subgenus *Auriculastrum* makes up many of the *Primula* taxa endemic to western North America, with *Primula* section *Parryi* consisting of those found in the Intermountain West, the region between the Sierra Nevada range to the west and the Rocky Mountains to the east. The members of *Primula* section *Parryi* are largely found in alpine to subalpine habitats, but while species *Primula parryi* and *P. rusbyi* Green are widely distributed within their respective mountain ranges, *P. angustifolia* Torrey, *P. capillaris* (Holmgren and Holmgren 1974), and members of the *P. cusickiana* species complex (*P. cusickiana* varieties *maguirei*, *domensis*, *nevadensis*, and *cusickiana* (see Fig. 1)) are more restricted to patches of anomalously cool habitat in otherwise xeric environments. For the *P. cusickiana* species complex, this is particularly true of *P. cusickiana* var. *maguirei* (Williams 1936), found growing on cliffs and small ledges of dolomitic limestone substrates in a 20 km stretch of Logan Canyon within the Bear River Range in northern Utah. *Primula cusickiana* varieties *domensis* (Kass and Welsh 1985) and *nevadensis* (Holmgren 1967) also occur in very small pockets on limestone substrates, but at higher elevations past treeline, with *P. cusickiana* var. *domensis* being found in the House Range (west central Utah) and *P. cusickiana* var. *nevadensis* having documented populations in the Snake Range (just west of the House Range) and the Grant Range of Nevada. Finally, *P. cusickiana* var. *cusickiana* (Gray 1888) is more widespread, found in cool, wet microsites at relatively low elevations along the Snake River Plain in Idaho, but with documented populations as far west as eastern Oregon (on the Owyhee High Desert Plain) and as far south as the Cougar Mountains (near Jarbidge) of northeastern Nevada (Mansfield 2010).

The geographic separation and ecological differences of *Primula cusickiana* varieties *maguirei*, *nevadensis*, and *cusickiana* contributed to their original classification as unique species

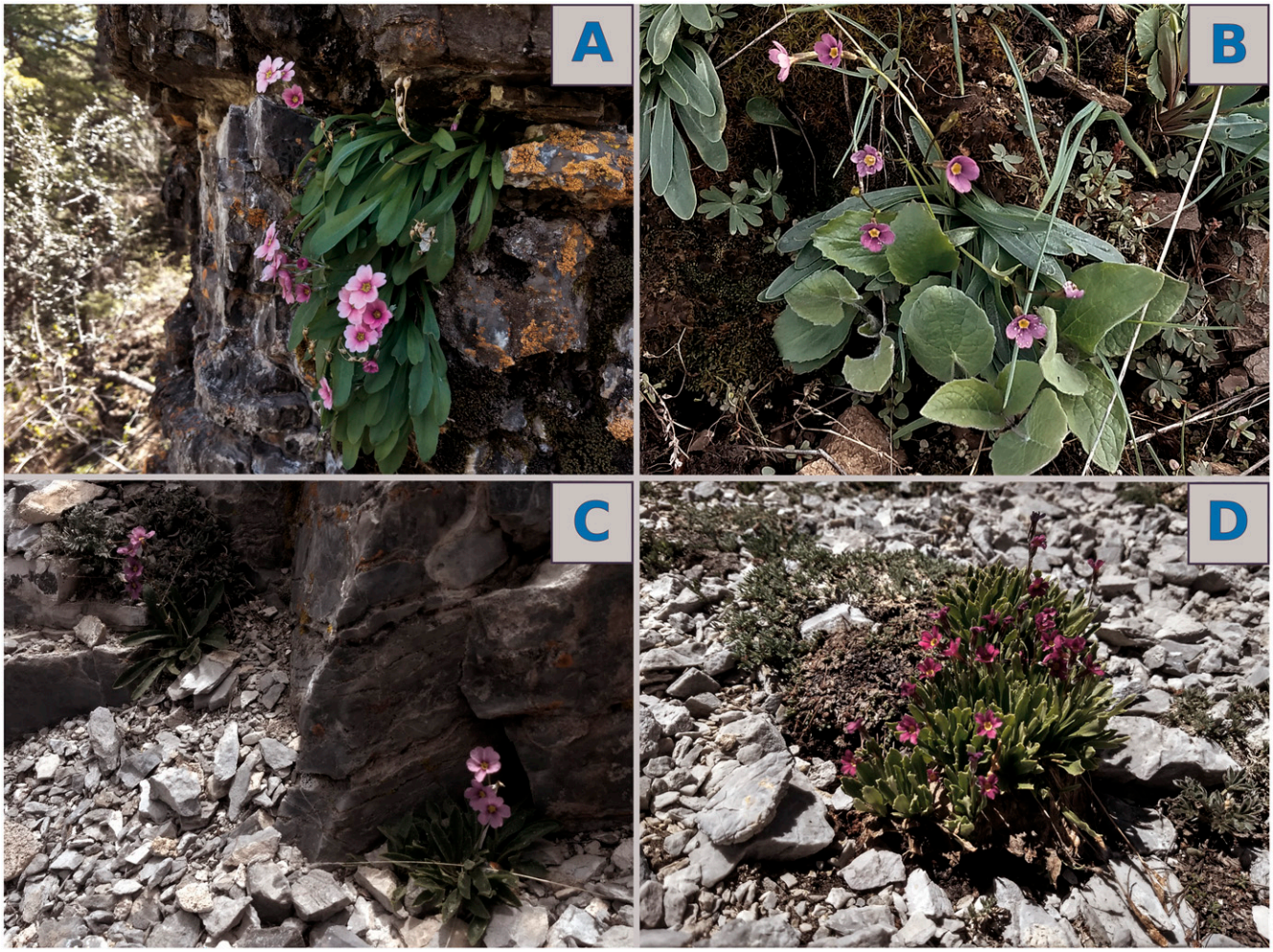

FIG. 1. Four members of the *Primula cusickiana* species complex. A. *P. c.* var. *maguirei*, in Right Hand Fork of Logan Canyon. B. *P. c.* var. *cusickiana*, near Cougar Point in Jarbidge, Nevada. C. *P. c.* var. *domensis*, at Notch Peak in the House Range, Utah. D. *P. c.* var. *nevadensis*, on Mount Washington in the Snake Range (Great Basin National Park), in Nevada.

(i.e. *P. maguirei*, *P. nevadensis*, and *P. cusickiana*) upon initial description. The discovery and publication of *P. domensis* in 1985, along with the continued collection of the other varieties, began to cast doubt on the species distinction for each complex member. Morphologically, the differences among the four varieties are subtle: *P. cusickiana* vars. *maguirei* and *cusickiana* are entirely glabrous, and distinguished from one another by relative calyx length, while in *P. cusickiana* vars. *nevadensis* and *domensis*, plants are pubescent, with *P. cusickiana* var. *nevadensis* having a shorter corolla tube length than *P. cusickiana* var. *domensis* (Holmgren et al. 2005). A review in 2001 considered these morphological distinctions "meager at best," and concluded that "recognition at the varietal level would be most appropriate" (Holmgren and Kelso 2001).

At the time of this shift, no genetic data was available to justify classification at the variety level. However, a 1997 analysis of *P. cusickiana* var. *maguirei* used allozyme marker genes to uncover a significant degree of genetic structure between the relatively proximate (~10 km) populations (Wolf and Sinclair 1997) within this one taxon. A later analysis of the same populations using amplified fragment length polymorphism (AFLP) loci confirmed this finding, and found similar levels of polymorphism between the upper and lower canyon groups, suggesting this genetic structure is not the result of a past bottleneck event (Bjerregaard and Wolf 2008). A further analysis of AFLP and chloroplast DNA from members of *Primula* section *Parryi* showed *P. cusickiana* var. *maguirei* and the other *P. cusickiana* complex members as being monophyletic, but relationships within the complex were incongruent, with only weak support of a clade containing *P. cusickiana* vars. *nevadensis* and *domensis* being sister to a clade made up of *P. cusickiana* vars. *maguirei* and *cusickiana* (Kelso et al. 2009). To better resolve the relationships between varieties, the authors suggested an analysis utilizing more populations from across the range of this species complex. Restriction site-associated DNA sequencing (RADseq) methods available today, with their ability to generate reads over many sequence regions of closely related individuals, are well-suited to provide the data required for such an analysis.

In addition to clarifying the genetic relations between geographically distinct varieties, a more detailed analysis of the *P. cusickiana* species complex can meaningfully contribute to ongoing conservation efforts. *Primula cusickiana* var. *maguirei* was listed as Threatened in 1985, due to its unique habitat in Logan Canyon and threats of habitat loss due to development (US Fish and Wildlife Service 1985). Given the strong genetic structure between *P. cusickiana* var. *maguirei*'s populations, either population may be more closely related to populations

of a different complex variety than the neighboring Logan Canyon population, a finding which would have significant implications for the protection of this variety. More broadly, an understanding of the genetic relations in this species complex will determine whether the varietal classification properly reflects the extent of divergence of each complex member, and thus the extent of unique evolutionary history. This understanding can direct management of the narrow-range endemics included in this species complex, such as *P. cusickiana* var. *maguirei*, but also *P. cusickiana* vars. *nevadensis* and *domensis*, and also inform the identification of potential evolutionarily significant units (Coates et al. 2018).

We sought to clarify the relatedness of the *P. cusickiana* complex members by using a RADseq approach to genotype all four varieties located at distinct populations scattered throughout the Great Basin. In addition to contextualizing the genetic structure between the upper and lower Logan Canyon *P. cusickiana* var. *maguirei* populations, this analysis provides insights into the biogeographic history of this species complex, and could have important conservation implications for this rare endemic plant.

Materials and Methods

Sampling—Samples from all four *P. cusickiana* species complex members were gathered in the field. Additionally, we collected samples from *P. parryi*, a species outside of the *P. cusickiana* species complex, in order to compare genetic differences among complex members to those extending outside of the complex. Finally, because past research has shown variable relations between *P. capillaris* and the *P. cusickiana* species complex (Kelso et al. 2009), we also tried to collect *P. capillaris* in the field. However, we were unable to locate any wild *P. capillaris* individuals in the Ruby Mountains: at one location suggested by past herbaria data, a population of *P. parryi* was found instead. To compensate, two *P. capillaris* samples were sourced from herbaria (see Appendix 1). By sampling all four species complex varieties, *P. capillaris*, and *P. parryi*, we captured 75% of the taxa in the *Primula* section *Parryi*. The number of collected samples for each population of each variety, as well as *P. parryi*, along with population coordinates, are provided in Table 1. Sampling sites were determined by examining all populations documented through specimens in herbaria across the United States, and selecting the populations of each variety that would maximize the geographic extent of our genetic survey. For *P. cusickiana* vars. *maguirei*, *domensis*, and *nevadensis*, this meant collecting from previously documented locations within each variety's mountain ranges. For *P. cusickiana* var. *cusickiana*, we collected from populations documented at the most westerly (Craters of the Moon National Monument), northerly (Bear, Idaho), southerly (Jarbidge, Nevada), and easterly (Owyhee High Desert, Oregon) extent of the variety's known range.

At each population location, an individual plant was removed as completely as possible as a voucher specimen. For DNA samples, two leaves from each of ten plants were removed and placed in labeled paper envelopes, which were stored on silica crystals to keep samples dry. Vouchers were deposited at the Intermountain Herbarium (UTC); *P. cusickiana* var. *nevadensis* voucher specimens collected from Mt. Washington were additionally deposited at the Great Basin National Park herbarium.

Leaf tissue from 89 samples (87 silica-dried field collections representing all sample sites, and two herbarium specimens of *P. capillaris*) were placed into Qiagen Collection Microtubes (catalog number 19560) and sent to the University of Wisconsin-Madison Biotechnology Center, for DNA extraction, library prep, and DNA sequencing (described below). Seven replicate samples were also included to assess the quality of sequencing results, and were distributed across all four *P. cusickiana* varieties, as well as *P. parryi*.

DNA Extraction—DNA was extracted using the Qiagen Dneasy mericon 96 QIAcube HT kit. DNA was then quantified using the Quant-iT^{TM} PicoGreenR© dsDNA kit (Life Technologies, Grand Island, New York).

Library Preparation and Sequencing—Libraries were prepared following Elshire et al. (2011). ApekI (New England Biolabs, Ipswich, Massachusetts) was used to digest 100 ng of DNA. Following digestion, Illumina adapter barcodes were ligated onto DNA fragments using T4 ligase (New England Biolabs, Ipswich, Massachusetts). Size selection was run on a PippinHT (Sage Science, Inc., Beverly, Massachusetts) to subset samples down to 300–450 bp fragments, after which samples were purified using a SPRI bead cleanup. To generate quantities required for sequencing, adapter-ligated samples were pooled and then amplified, and a post-amplification SPRI bead cleanup step was run to remove adapter dimers. Final library qualities were assessed using the Agilent 2100 Bioanalyzer and High Sensitivity Chip (Agilent Technologies, Inc., Santa Clara, California), and concentrations were determined using the Qubit© dsDNA HS Assay Kit (Life Technologies, Grand Island, New York). Libraries were sequenced on an Illumina NovaSeq 6000, generating paired-end data for fragments of approximately 150 bp in length.

Data Processing—Raw FASTQ data files were demultiplexed and processed using steps 1–7 of the ipyrad software, v. 0.9.31 (Eaton and Overcast 2020). Single nucleotide polymorphisms (SNPs) recognized by ipyrad were used as the basis for variation between individuals for downstream analyses, and assemblies were generated de novo. All ipyrad and STRUCTURE parameter files, as well as R scripts used for analysis and data visualization, can be found in the Supplementary Materials available on Dryad (Koontz et al. 2022), as well as on GitHub (github.com/akoontz11/Primula/). Raw, demultiplexed sequencing data can also be accessed on the NCBI Sequence Read Archive (SRA; accession number PRJNA705310).

COMPLEX-WIDE GENETIC SURVEY—For our complex-wide genetic survey, we ran ipyrad twice: we used the results from our initial run to confirm sequencing consistency for replicate samples, and to identify samples with low coverage. For both runs, demultiplexed sequences were paired

Table 1. Number of samples and locations per taxon.

Variety	Site Description	Number of Samples	Coordinates	Elevation
Primula cusickiana var. *cusickiana*	Bogus Basin, Boise, Ada County, Idaho	5	N 43°41.07′, W 116°11.27′	1093 m
	South of Bear, Adams County, Idaho	8	N 44°58.87′, W 116°40.8′	1297 m
	Lazy K10 Ranch, Camas Prairie, Elmore County, Idaho	6	N 43°18.65′, W 115°5.47′	1587 m
	Tom Cat Hill, Craters of the Moon National Monument, Butte County, Idaho	5	N 43°24.46′, W 113°37.60′	1757 m
	North of Cougar Point, Jarbidge, Elko County, Nevada	5	N 41°57.51′, W 115°18.96′	2045 m
	5 miles south of Jackson Summit, Owyhee High Desert, Malheur County, Oregon	5	N 42°10.20′, W 117°36.93′	1939 m
Primula cusickiana var. *maguirei*	Second Practice Wall, Logan Canyon, Cache County, Utah	5	N 41°44.73′, W 111°45.14′	1534 m
	Greenhouse Wall, Logan Canyon, Cache County, Utah	3	N 41°44.56′, W 111°45.70′	1496 m
	Right Hand Fork of the Logan River, Logan Canyon, Cache County, Utah	5	N 41°46.62′, W 111°37.97′	1865 m
	Wood Camp (Seed Source), Logan Canyon, Cache County, Utah	5	N 41°47.90′, W 111°38.55′	1655 m
Primula cusickiana var. *domensis*	Sawtooth Canyon, House Range, Millard County, Utah	5	N 39°8.47′, W 113°23.27′	2467 m
	Notch Peak trail, House Range, Millard County, Utah	5	N 39°8.05′, W 113°23.72′	2438 m
	Notch Peak summit, House Range, Millard County, Utah	5	N 39°8.57′, W 113°24.34′	2814 m
Primula cusickiana var. *nevadensis*	Mt. Washington summit, Snake Range, Great Basin National Park, White Pine County, Nevada	5	N 38°54.74′, W 114°18.52′	3489 m
	Troy Peak summit, Grant Range, Nye County, Nevada	4	N 38°19.32′, W 115°29.87′	3272 m
Primula parryi	Thomas Creek headwaters, Ruby Mountains, Elko County, Nevada	5	N 40°36.61′, W 115°24.36′	2966 m

and merged, and low quality bases, adapters, and primers were filtered prior to SNP calling. Default values were used for the ipyrad parameters in these steps, as well as for the clustering threshold (clust_threshold; 0.85) and minimum sequencing depth (mindepth_statistical; 6) parameters.

For our initial run, we specified a minimum number of samples per locus (min_samples_locus) parameter of 10, in order to obtain loci shared between two to three sample locations for any taxa. Using the results from this run, we used a Python script (vcf2Jaccard.py) to compare samples with replicates by calculating the mean Jaccard similarity coefficients between all samples. We found that all replicates matched highly with their corresponding samples (Fig. S1, Koontz et al. 2022).

After merging replicates and removing low coverage (generally, less than 30 loci in the final assembly) samples from the dataset, 82 of our 87 original samples remained for our complex-wide analysis. We reran ipyrad (steps 1–7) using these 82 samples to select for loci specific to this subset. We used a min_samples_locus parameter of 32 for this second run, to cluster using loci shared across the species complex; ipyrad default values were used otherwise. Because very low numbers of loci were retrieved for both herbarium specimens of *P. capillaris* (possibly due to the age of these specimens), we were unable to include *P. capillaris* in downstream clustering analyses.

Variety Specific Clustering—In addition to our complex-wide survey, we were interested in exploring population structure within *P. cusickiana* var. *maguirei* which could not be resolved using loci shared across all species complex members. To do so, we ran ipyrad on just the 18 *P. cusickiana* var. *maguirei* samples used in our complex-wide survey. Because five samples from each of the upper Logan canyon sampling sites were included in our ipyrad assembly, we specified a min_samples_locus parameter of 5; ipyrad default parameter values were used otherwise.

Population Analyses—STRUCTURE—To visualize relations between complex members across their geographic range, and to determine the number of identifiable genetic clusters within the complex, we used the program STRUCTURE v. 2.3 (Pritchard et al. 2000). STRUCTURE uses Bayesian clustering analysis to probabilistically assign individuals to one or more of K source populations, where the loci within each population are assumed to be in Hardy-Weinberg proportions and linkage equilibrium. While the presence of an outgroup is not required for the STRUCTURE software, our preliminary findings suggested that including *P. parryi* did not significantly alter our clustering results, and we chose to include this species for this and our other population level analyses (see below). For all STRUCTURE runs, we used a burn-in length of 50,000, and 100,000 MCMC reps after burn-in. For our complex-wide survey, we ran STRUCTURE for K values of 2–16, with 50 replicates per K value. For our *P. cusickiana* var. *maguirei*-only analyses, we ran STRUCTURE for K values of 2–6, with 50 replicates per K value. We used the CLUMPAK server (Kopelman et al. 2015) to summarize results across replicates for each K value, and to build STRUCTURE plots.

For all of our STRUCTURE analyses, we used two different methods to determine the "optimal" K value: the method described in the STRUCTURE manual (Pritchard et al. 2000), which identifies the K value with the greatest likelihood (the natural logarithm of the posterior probability of the genetic data at a specific K value); and the method described in Evanno et al. (2005), which measures the second order rate of change in the posterior probability of the genetic data between successive K values (ΔK). While the ΔK method described by Evanno was designed to better accommodate scenarios in which unequal migration is occurring between putative source populations, it is worth noting that both techniques are ad hoc and require interpretation (as described by their authors). Furthermore, there is an inherent difficulty in inferring an unambiguous number of genetic clusters from any given set of populations (Novembre 2016). Therefore, following recommendations in the original STRUCTURE manual (Pritchard et al. 2000; Novembre 2016), we generated STRUCTURE outputs within a range of K values, and examined how relationships changed or remained constant across values. Because robust species boundaries consider genetic data in concert with other types of data (Carstens et al. 2013), we took the geographic distribution of each population into consideration when determining a biologically reasonable value of source populations. We were particularly interested in visualizing the extent of genetic differences between geographically distinct populations within varieties across K values, and how different values of K illustrated the extent of admixture (as indicated by individuals having portions of their genomes assigned to different sources) between adjacent populations. Finally, given evidence for strong genetic structure between the populations of *P. cusickiana* var. *maguirei* (Wolf and Sinclair 1997; Bjerregaard and Wolf 2008), we were interested to see if either *P. cusickiana* var. *maguirei* population in Logan Canyon was more closely related to another species complex population than it was to the neighboring Logan Canyon population.

SplitsTree—In addition to our STRUCTURE analysis, we used the NeighborNet split network algorithm (Bryant and Moulton 2002) to visualize the taxa sampled in this survey using a phylogenetic network. NeighborNet works by iteratively grouping pairs of taxa together based on similarities, in the same way that a neighbor joining tree is built. However, rather than a tree, the end result is a split network, in which splits in the taxa are represented by parallel lines, and conflicting signals in relations of taxa are represented by boxes. The ability to represent different phylogenetic hypotheses simultaneously allows this method to accommodate processes which undermine traditional phylogenetic analyses, such as scenarios involving recombination and hybridization. This makes the NeighborNet algorithm useful for summarizing RADseq data, where patterns of loci across individuals may imply different genealogies. For our complex-wide survey, we passed the output from our ipyrad de novo assembly to the software SplitsTree v. 4.17.1 (Huson and Bryant 2006), which we used to construct the NeighborNet split network.

Discriminant Analysis of Principal Components—In addition to STRUCTURE and the NeighborNet algorithm, we analyzed the results of our complex-wide survey using Discriminant Analysis of Principal Components (DAPC; Jombart et al. 2010) in the package adegenet in R v. 3.6.3 (R Core Team 2020). DAPC is a statistical technique designed to accommodate the size of genomic data sets and capable of differentiating within-group variation from between-group variation. SNP data is first transformed using a principal components analysis (PCA), and then k-means clustering is run to generate models and likelihoods corresponding to each number of population clusters. The best-fitting model, and so the best-supported number of populations, is assessed using the models' Bayesian Information Criterion (BIC) scores. We chose to utilize DAPC to visualize population clusters in a PCA format, and to determine whether the supported number of clusters was congruent with our STRUCTURE and SplitsTree results, indicating a more robust determination of the number of unique taxa contained within the complex (Carstens et al. 2013).

Fst Estimates—Because we wanted to measure the extent of genetic variance within the groups analyzed, we used the VCFtools software (Danecek et al. 2011) to generate weighted F_{ST} estimates (Weir and Cockerham 1984). We generated an F_{ST} estimate for our complex-wide analysis (across all populations of all *P. cusickiana* varieties, excluding *P. parryi*) as well as for the samples included in our *P. cusickiana* var. *maguirei*-only analysis.

Results

Complex-Wide Genetic Survey—We retrieved, on average, 2.04×10^6 reads per sample, and our complex-wide ipyrad run identified 1277 loci that were used in our subsequent STRUCTURE analysis. Using the Evanno et al. (2005) method yielded an optimal K value of K = 5; using the method described in the STRUCTURE manual (Pritchard et al. 2000) identified the K value with the greatest likelihood as K = 14 (Fig. 2). Given the limited differentiation between the makeup of individuals for K values beyond K = 7 (demonstrated by the addition of minor genetic proportions in individuals composed largely of a single genetic group), we felt K = 14 was too large a number of source populations to describe these taxa, and that distinctions at this level of clustering were observable at much lower values of K. While the value recommended by the Evanno method provides a more conservative hypothesis, in that it recommends a lower number of putative source populations, this level of clustering fails to illustrate the genetic distinction between the geographically separated Nevada (Jarbidge) and Oregon (Owyhee) populations of *P. cusickiana* var. *cusickiana*, which is illustrated at values greater than K = 5. Similarly, we felt that the clustering at K = 6 failed to properly distinguish the distinct populations of *P. cusickiana* var. *nevadensis* in Great Basin National Park (GRBA) and the Grant Range (Troy), a distinction which is quite notable at higher values of K. Therefore, following the recommendations provided in Pritchard et al. (2000) by examining how relationships between our sampled

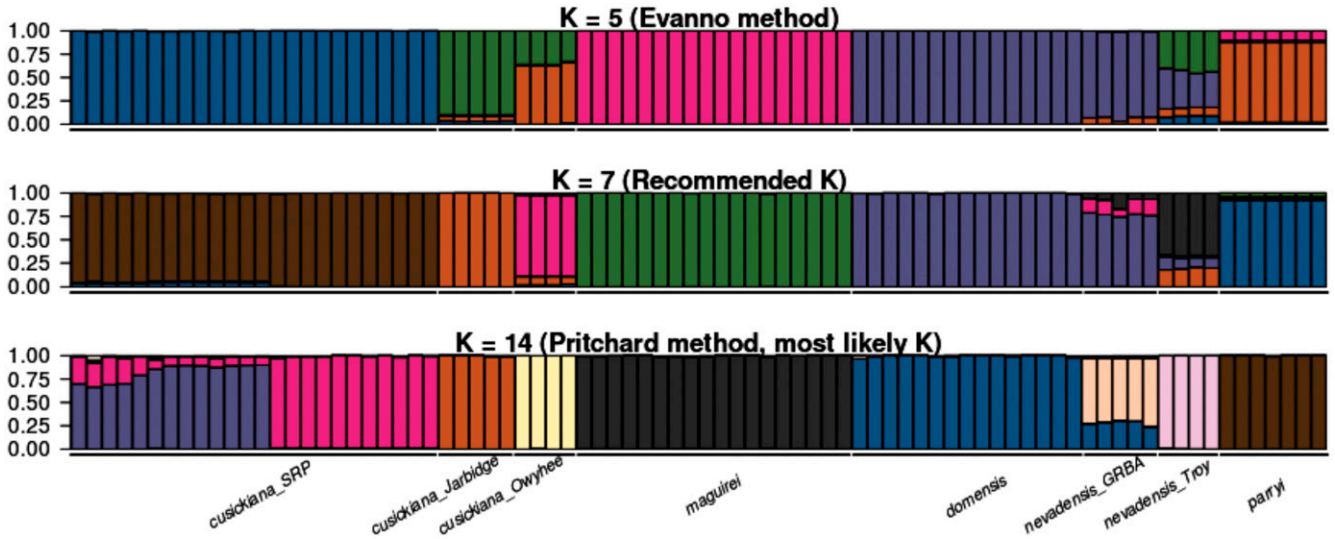

FIG. 2. Sample STRUCTURE plots at K = 5 (the optimal K value determined by the Evanno ΔK method), K = 7 (the recommended value of K based on all STRUCTURE outputs), and K = 14 (the value of K with the greatest likelihood). Bars and names on the bottom indicate populations grouped based on K = 7 clustering. While groupings remain largely coherent across K values, *nevadensis* individuals at Troy peak appear more admixed and less unique at K = 5, while separation between *cusickiana* populations along the Snake River plain is more observable at K = 14.

populations change across K values ranging from K = 2 to K = 16 (Figs. S2–S4, Koontz et al. 2022), we determined K = 7 to be the most biologically relevant value to describe this group of plants. This level of source populations best captures the geographic isolation of all sampled populations, with *P. cusickiana* vars. *domensis* and *maguirei* being clearly delineated, *P. cusickiana* var. *nevadensis* showing distinctions between its two populations, and *P. cusickiana* var. *cusickiana* split into three groups composed of populations from the Snake River Plain in Idaho (SRP), Nevada (Jarbidge), and Oregon (Owyhee). Since higher K values emphasize the divisions seen at this level, and further subdivide isolated populations of *P. cusickiana* vars. *cusickiana* and *nevadensis*, K = 7 is an estimate which reflects the strong divisions within this complex while allowing for more nuanced groupings of unique populations to be made in light of future evidence.

Our NeighborNet split network analysis supports this clustering value, and divisions seen in our STRUCTURE analysis are similarly reflected in this split network (Fig. 3). *Primula cusickiana* var. *maguirei* is separated from all other varieties, while *P. cusickiana* var. *domensis* and the Great Basin National Park population (GRBA) of *P. cusickiana* var. *nevadensis* group together, and separately from the Grant Range population of *P. cusickiana* var. *nevadensis* (Troy). Finally, *P. cusickiana* var. *cusickiana* is split into the same three groups shown in our STRUCTURE analysis: populations from the Snake River Plain in Idaho (SRP), Nevada (Jarbidge), and Oregon (Owyhee), with the last population splitting from a branch shared with *P. parryi*. Our DAPC analysis, however, revealed that the greatest supported number of clusters (i.e. the value with the lowest BIC score) was eleven (data not shown), a value incongruent with our SplitsTree and STRUCTURE results, suggesting that boundaries within this complex are elaborate. At this level of genetic clusters, several groups were quite small (consisting of only one or two samples), and groupings were incoherent with the spatial distribution of populations. To provide a clearer comparison to our STRUCTURE results, and to examine relations strictly within the species complex, we removed *P. parryi* outgroup samples from our dataset and ran our DAPC with a specification of six clusters (Fig. S5, Koontz et al. 2022). At this level of clustering, the population of *P. cusickiana* var. *nevadensis* in the Snake Range of Great Basin National Park (GRBA) is shown as a unique cluster, while the *P. cusickiana* var. *nevadensis* population further south in the Grant Range groups with the *P. cusickiana* var. *cusickiana* population sampled from Oregon (Owyhee). *Primula cusickiana* var. *domensis* is a unique cluster which groups closely to both of these. Thus, while our population level analyses do not point to an unambiguous number of "true" genetic clusters, some patterns are shared across techniques.

The extreme level of divergence between the sky island populations in this species complex is reflected not only in our population level analyses, but also in our relatively large F_{ST} estimate across all complex populations, which was 0.72. Figure 4 illustrates proportions of sample membership to clusters based on our STRUCTURE analysis at K = 7 for all populations in their geographic context across the Great Basin.

Variety Specific Clustering—In our complex-wide analysis, all *P. cusickiana* var. *maguirei* samples grouped as a single cluster, distinct from all other populations of all other varieties, indicating that neither Logan Canyon population is more closely related to any populations of another variety. Even at values of K = 16, the upper and lower Logan Canyon populations of *P. cusickiana* var. *maguirei* were not resolved from one another.

However, reducing our sample set to only *maguirei* samples allowed us to retain loci informative to this variety but unshared with other complex member populations. Our *P. cusickiana* var. *maguirei*-only ipyrad run generated an assembly with 68,492 loci, indicating a large number of loci specific to *P. cusickiana* var. *maguirei* and not shared with the wider species complex. To speed up processing times, we ran STRUCTURE on a 17,988 loci subset of *maguirei*-specific markers. Using the CLUMPAK server, we found optimal K values of K = 4 (using the Evanno method) and K = 3 (using the likelihood method described in the STRUCTURE manual).

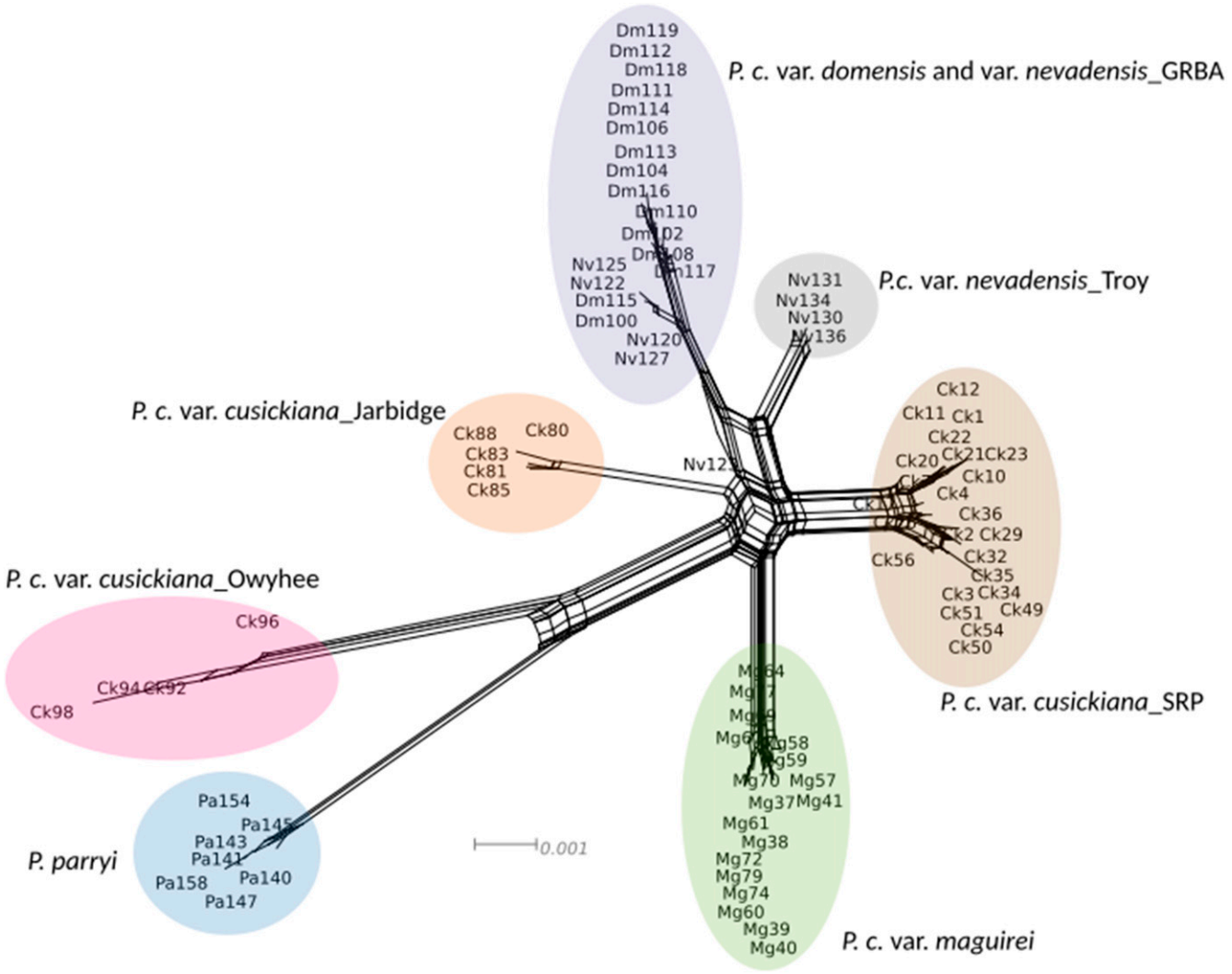

FIG. 3. Split network analysis generated using the NeighborNet algorithm. Tips indicate individuals, and groups are labeled by variety and populations, with coloration matching the groupings used in the K = 7 clustering of Fig. 2. Splits between taxa are represented by parallel edges (lines), and the lengths of edges are proportional to the weight of the associated split.

Figure 5 shows the STRUCTURE plot at K = 3, which resolves similar groupings of *maguirei* populations supported in (Bjerregaard and Wolf 2008), and the distinctions between upper and lower canyon populations. We also estimated an F_{ST} value of 0.33 among these three populations, which is comparable to previous estimates in Bjerregaard and Wolf (2008).

DISCUSSION

Our analysis of RADseq data from *Primula cusickiana* complex members demonstrates that the disjunct geographical distribution of populations across the Great Basin is reflected by pronounced genetic divergences. While the results of our clustering analyses largely coincide with the divisions of current varietal classifications, there are notable exceptions. Distinctions between isolated populations within varieties, as well as similarities between neighboring populations of different varieties, can be observed in our STRUCTURE plots of low K values (i.e. ranging from 2–6; see Figs. S2–S4) and in our split network analysis. For instance, we found Great Basin National Park (GRBA) *P. cusickiana* var. *nevadensis* populations to be admixed, with genetic contributions coming from *P. cusickiana* var. *domensis* to the east and (to a lesser extent) Grant Range *P. cusickiana* var. *nevadensis* populations to the south. This is in accordance with past analyses of AFLP and chloroplast DNA from the *Primula* section *Parryi*, which found these two varieties to be extremely close (Kelso et al. 2009).

Our results also suggest a more nuanced understanding of *P. cusickiana* var. *cusickiana*. Populations of this variety are split into distinct clusters in our analysis, with Jarbidge (Nevada) and Owyhee (Oregon) populations appearing unique from each other and the remaining Snake River Plain (SRP) populations in Idaho. That these distinctions are seen in our STRUCTURE and SplitsTree analyses imply the robustness of this result. Given the relatively wide distribution of this variety (growing in moist soils at lower elevations than other complex members), our findings of genetic divergence between its populations are noteworthy, and support past evidence of phenotypic differences in different portions of its range. For instance, past morphological research of Idaho *P. cusickiana* var. *cusickiana* populations has suggested dividing this taxa into three unique species (Mansfield 1993),

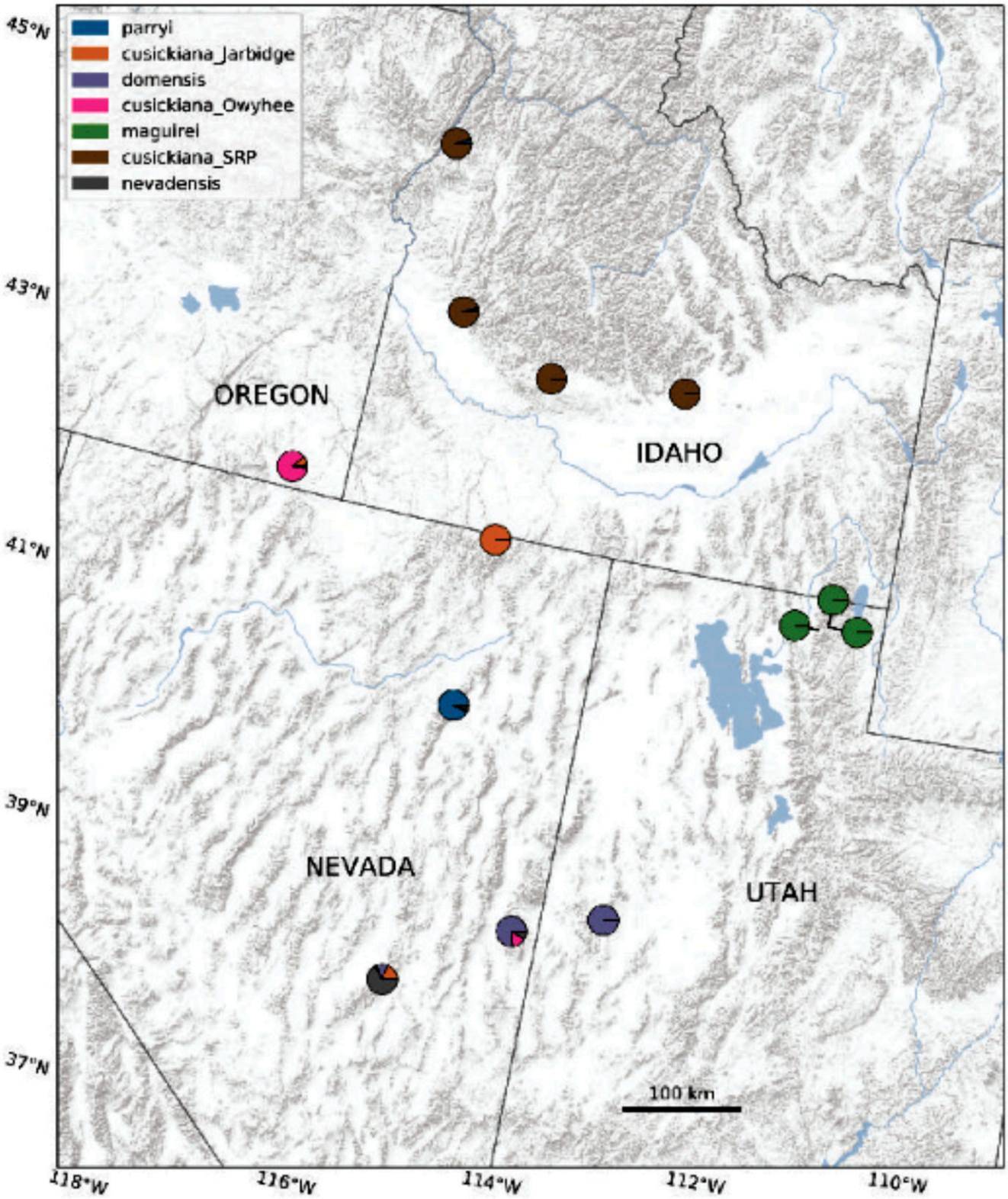

FIG. 4. Map of sample locations with cluster membership. Sampling locations are represented by pie charts indicating percentage of population membership to clusters determined at K = 7 STRUCTURE clustering threshold. With exception to *nevadensis*, most samples fall almost entirely within a specified cluster.

with Owyhee populations being classified as *P. wilcoxiana*. Interestingly, our SplitsTree analysis suggests this population has a close and possible reticulated relationship with *P. parryi*, perhaps belonging outside of the *P. cusickiana* species complex entirely. While all interpretations of this group implicate its distinction from the rest of *P. cusickiana* var. *cusickiana*, a more detailed phylogenetic and morphological description of the Owyhee populations of *P. cusickiana* var. *cusickiana* would be required before species recognition is warranted.

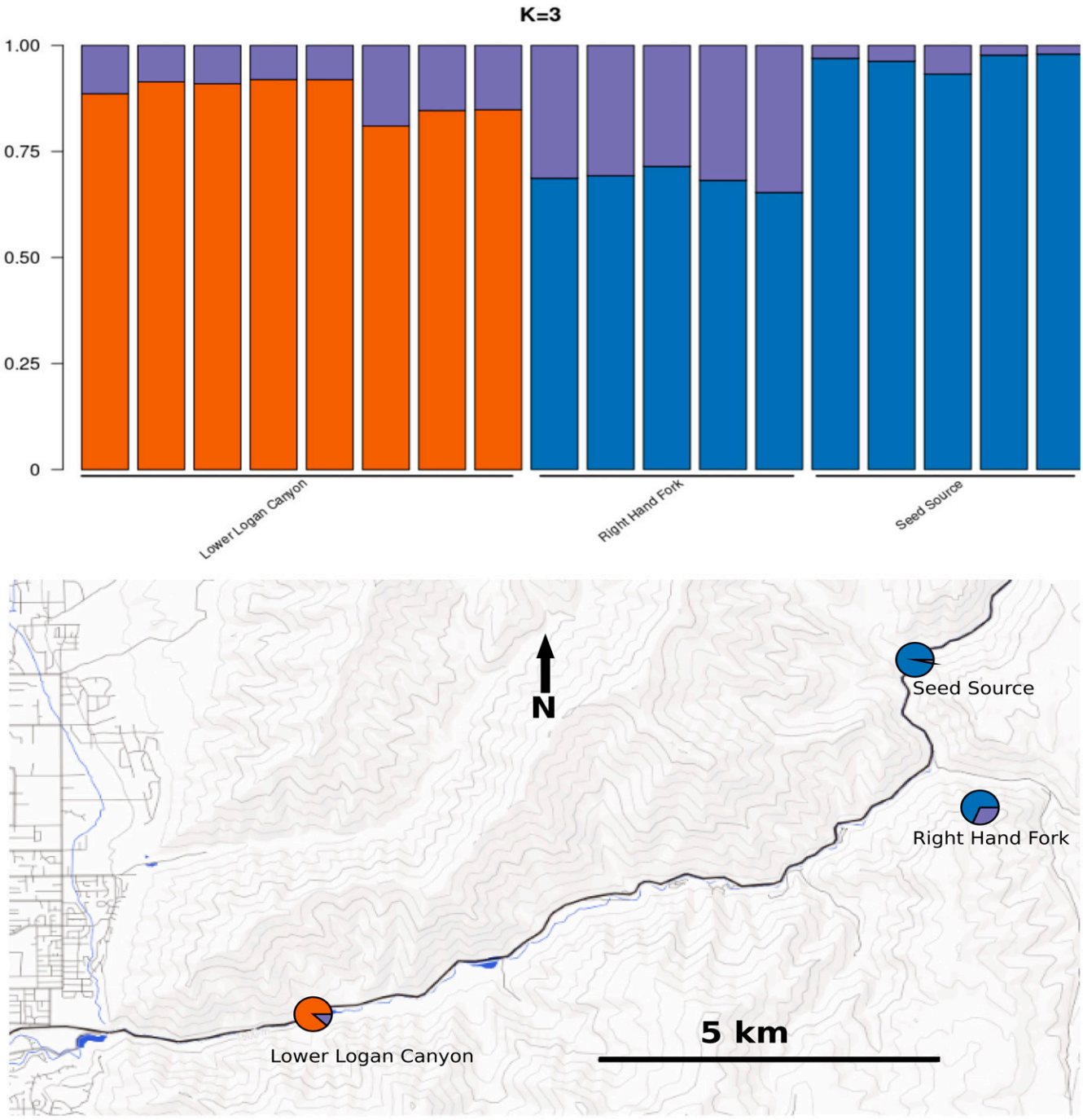

FIG. 5. STRUCTURE plot for *maguirei* samples at a clustering threshold of K = 3, with a map depicting the lower and upper canyon collection sites. While *maguirei* clustered together in the complex-wide analysis, *maguirei*-only analysis was able to resolve divisions suggested in past studies.

The separation between populations within *P. cusickiana* var. *cusickiana*, as well as our support of past findings of significant genetic distances between the proximate populations of *P. cusickiana* var. *maguirei*, underscore our discovery of profound genetic divergences between all members of this species complex, despite their distribution over a relatively small geographic area. This trend is reflected not only in our clustering analyses, but also in our weighted F_{ST} estimate of 0.72 across complex populations, a high value compared to similar estimates for other plant taxa (for instance, the mean F_{ST} for plant taxa in a meta-analysis by Leinonen et al. (2008) was estimated to be 0.24). Our sampling sites were determined by maximizing sampling across the known geographic distribution of all populations of *P. cusickiana* var. *cusickiana*, in addition to collecting from all documented populations of the more geographically restricted varieties (*P. cusickiana* vars. *maguirei*, *nevadensis*, and *domensis*). Therefore, short of the discovery of new populations (e.g. in different sky island habitats in the Great Basin, or outside *P. cusickiana* var. *cusickiana*'s known locations), we believe this is an exhaustive geographic survey of this complex. Although these results support the historical designation of species for these complex members, rather than variety, we restrain from asserting that designation here, based on our reported population genetic comparisons

alone. In accordance with using several types of nongenetic data, in addition to more standard genetic tests, to support species definitions (Carstens et al. 2013), we believe more thorough morphological and phylogenetic analyses (ideally including both nuclear and cpDNA), as well as breeding surveys, are necessary before unambiguous species designations can be made. Strong genetic structure between groups, while typically indicative of separate species, is not sufficient on its own for establishing species identification (Sukumaran and Knowles 2017), especially given the potential for genetically separate groups to generate fertile offspring, thereby violating the central criterion of the biological species concept (Huang 2020). However, regardless of formal classification, these findings of strong genetic separation are worthy of interest simply due to their environmental and biogeographical context. Below, we consider how three phenomena (biogeographical trends in the Great Basin, edaphic endemism, and reproductive traits specific to *Primula*) may contribute to the significant genetic divergence observed across this species complex.

Great Basin Sky Island Biogeography—Members of the *P. cusickiana* complex are found at relatively high (2000–3300 m) elevations throughout the Great Basin. Many of these are sky island locations associated with strong ecological shifts as habitat transitions from lower sagebrush steppe to cooler, more forested regions dominated by pinyon and juniper. Now separated by arid basins due to climatic warming in the Holocene, these sky islands are understood to be the fragmented remnants of a continuous region of cool, moist habitat which once extended across the Great Basin (Thompson and Mead 1982). This has led to their characterization as refugia for various taxa, particularly mammals (Brown 1971; Badgley et al. 2014), but also butterflies (Boggs and Murphy 1997) and plants (Harper et al. 1978; Nowak et al. 1994; Charlet 2007).

However, research regarding the populations unique to these sky island habitats has noted that many species distribution patterns among Great Basin mountaintops do not follow a strictly island biogeographical model (Lawlor 1998; Fleishman et al. 2001), in that neither island surface area nor proximity to "mainland" source populations (typically identified as the western Sierra Nevada or eastern Rocky Mountains) is predictive of species abundance (Fleishman et al. 2001). And in some taxa, there is evidence for regular, modern dispersal between Great Basin ranges (Floyd et al. 2005). An alternative scenario is that this complex has followed what has been described as an "expanding-contracting archipelago" (ECA) model, in response to Quaternary glacial cycles (DeChaine and Martin 2005a). The ECA model has been used to describe the divergence between Rocky Mountain sky island plant taxa (Dechaine and Martin 2005b; Hodel et al. 2021), and provides a framework for explaining the genetic structure observed between isolated montane populations on a broad spatial scale. In this model, populations are assumed to become fragmented as they contract up-slope during warmer interglacials; during glacial periods, populations expand down-slope as moist, cool habitat becomes widespread, leading to hybrid zones and possible admixture. Given the degree of fragmentation between *P. cusickiana*'s populations in today's climate (which resembles past interglacial periods), and the admixture between the relatively proximate populations of *P. cusickiana* vars. *domensis* and *nevadensis* revealed in our analysis, this model offers a viable explanation for the trends observed in this species complex. In addition to determining areas of refugia and secondary contact between these varieties, Quaternary glacial cycles may have influenced the colonization of the Intermountain West region more broadly. In the late Tertiary, *Primula* section *Parryi* is considered to have derived from an ancestral lineage within the *Auriculastrum* clade in east Asia (Zhang and Kadereit 2004; Zhang et al. 2004), with *Primula parryi* being the first to diverge from a common North American ancestor (Kelso et al. 2009). Given the prevalence and diversification of *Primula* in the Arctic (Kelso 1992), climatic fluctuations have likely played a role in the north-south dispersal of this genus across the North American continent.

Edaphic Endemism—In conjunction with climatic niche preferences, complex *P. cusickiana* vars. *maguirei*, *domensis*, and *nevadensis* are found on the cliffs and crevices of exclusively limestone substrates. Edaphic endemism is known to play a role in the diversification of plant species, both globally (Rajakaruna 2018; Hulshof and Spasojevic 2020) and within the Great Basin region specifically (De Queiroz et al. 2012; Brown and Mansfield 2017), and we might consider whether the trends uncovered in our genetic survey are a result of edaphic preferences of these populations. Within *Primula* section *Parryi*, the only species that is as narrowly restricted to calcareous substrates is *P. capillaris*, which is endemic to the Ruby Mountains of Nevada and has been shown to be closely related to *P. cusickiana* complex members (Kelso et al. 2009). *P. rusbyi*, which has a range from the southern Rocky Mountains to northern Mexico, is also found in limestone soils, but has been associated with granitic habitats as well. Given the wider edaphic niches of members of *Primula* section *Parryi* outside of the *P. cusickiana* species complex, it seems more likely that the edaphic preferences of *P. cusickiana* vars. *maguirei*, *domensis*, and *nevadensis* are the result of a tendency towards moisture retaining substrates, rather than specific mineral or pH constraints (Kelso et al. 2009). If so, we would expect the available niche space of the populations within this species complex to be contingent on the historical climatic fluctuations of the Great Basin, as described above, and not the edaphic heterogeneity of the region alone. Regardless of the ultimate cause, allopatry across distinctive climatic and edaphic niches seems to contribute to the genetic divergences in *P. cusickiana*'s populations, a trend observed in other sections of Primulaceae, as well (Boucher et al. 2016).

Speciation and Heterostyly in Primula—Recent research has shown several different alpine *Primula* species complexes to contain previously undescribed cryptic species, in China (Huang et al. 2019; Ren et al. 2020) and in Europe (Schorr et al. 2013; Theodoridis et al. 2019). Our findings on the *P. cusickiana* species complex resonate with these trends, and raise the question of what unique traits *Primula* possesses which might cause such frequent diversification via allopatric speciation. In addition to edaphic preferences in this genus, it has been argued that heterostyly, a widespread breeding system in angiosperms to promote outcrossing, may be a driving force leading to speciation in *Primula* (He et al. 2021). In heterostyly, "pin" and "thrum" floral morphologies prevent self-fertilization via insect pollination (Darwin 1877), and are associated with a sporophytic-incompatibility system which follows a Mendelian pattern of inheritance (Li et al. 2016). Changes in the prevalence of heterostyly across populations could lead to the divergence between distylous and homostylous populations, and ultimately speciation. Within populations, the possible deleterious effects of reduced effective population size (due to exclusive mating with individuals of the opposite morphology)

seem to be counterbalanced by the genetic advantages of outcrossing, leading to greater net diversification over evolutionary time (De Vos et al. 2014). While the extent of distyly within populations has been well documented in *P. cusickiana* var. *maguirei* (Davidson et al. 2014), observation of distylous morph ratios in other species complex populations would be required to determine if these dynamics are driving the divergences seen at the species complex level.

Conclusion—The results of our genetic survey of *Primula cusickiana* fit into a wider trend demonstrating abundant allopatric speciation despite little niche divergence in other alpine *Primula* species complexes. Our findings support the historical classification of each of these complex members as unique species, rather than the varietal classification taken in (Holmgren and Kelso 2001). Furthermore, these results warrant a more detailed understanding of the isolated and genetically unique populations in this complex (such as *P. cusickiana* var. *cusickiana* populations in Nevada and Oregon), and of the admixture observed in the populations of *P. cusickiana* var. *nevadensis*. Similarly, updated morphological comparisons between varieties, breeding experiments across the groups described in this study, and observations into the levels of heterostyly in disjunct populations, would offer a clearer understanding of the mechanisms of speciation occurring within this complex. Further genetic comparisons including *P. capillaris* would be able to clarify the currently ambiguous relationship between this species and the *P. cusickiana* species complex, and may bolster or conflict with the trends of allopatry seen in our results. Finally, the endemic species with narrow niches included in this study, such as *P. cusickiana* var. *maguirei*, but also *P. cusickiana* vars. *nevadensis* and *domensis*, and *P. capillaris*, warrant concern of extinction, and more work needs to be done to better understand the threats faced by each of these taxa in order to ensure their survival in an increasingly arid Great Basin.

Acknowledgments

This research was supported by funds provided by the Margaret Williams Research Grant offered by the Native Nevada Plant Society, the Lawrence Piette Graduate Scholarship offered by the Utah State University (USU) College of Science, the Dr. Ivan J. Palmblad Graduate Research Award offered by the USU Department of Biology, and the Graduate Research Award offered by the USU Ecology Center. WDP and the Pearse Lab are funded by NSF EF-1802605, NSF ABI-1759965, and UKRI-NERC NE/V009710/1.

The authors would like to thank Dr. Carol Rowe for her help with the genetic analyses, her Python script used to calculate Jaccard similarities from SNP data, and her Python scripts for creating the complex-wide and *maguirei* map. Thanks also to Dr. Barbara Ertter and especially Dr. Don Mansfield at the College of Idaho, for their assistance in collecting *P. cusickiana* var. *cusickiana*. Trish Winn and Jennifer Lewihnsohn in the United States Forest Service provided the collection permit for variety *maguirei* from the Uinta-Wasatch-Cache National Forest; Todd Stefanic and Gretchen Baker with the National Park Service coordinated collection from Craters of the Moon National Monument and Great Basin National Park, respectively. Michael Piep, Elizabeth Makings, and Jerry Tiehm allowed for *Primula* specimens to be sampled from the Intermountain Herbarium, Arizona State Vascular Plant Herbarium, and University of Nevada, Reno Herbarium, respectively. Thanks also to Kris Valles at the Intermountain Herbarium. Dr. Leila Shultz assisted with identification of the Owyhee *P. cusickiana* var. *cusickiana* population. The authors would also like to thank Jean Howerton, Buddy Smith, and Noel and Pat Holmgren. Finally, all collections were made on the ancestral lands of the Western Shoshone, Eastern Shoshone, Shoshone-Bannock, Southern Paiute, Goshute, and Nez Perce Native American tribes.

Author Contributions

AK determined sample locations, performed the majority of sample collection, and ran genetic analyses. WDP contributed to study design and assisted with genetic analyses and manuscript writing. PW guided study design and assisted with genetic analyses, sample collection, and manuscript writing.

Literature Cited

Badgley, C., T. M. Smiley, and J. A. Finarelli. 2014. Great Basin mammal diversity in relation to landscape history. *Journal of Mammalogy* 95: 1090–1106.

Bjerregaard, L. and P. G. Wolf. 2008. Strong genetic differentiation among neighboring populations of a locally endemic primrose. *Western North American Naturalist* 68: 66–75.

Boggs, C. L. and D. D. Murphy. 1997. Community composition in mountain ecosystems: Climatic determinants of montane butterfly distributions. *Global Ecology and Biogeography Letters* 6: 39–48.

Boucher, F. C., N. E. Zimmermann, and E. Conti. 2016. Allopatric speciation with little niche divergence is common among alpine Primulaceae. *Journal of Biogeography* 43: 591–602.

Brown, J. H. 1971. Mammals on mountaintops: Nonequilibrium insular biogeography. *American Naturalist* 105: 467–478.

Brown, B. J. and D. H. Mansfield. 2017. Edaphic and geographic origins of varietal differentiation in *Eriogonum calcareum*. *Madroño* 64: 22–31.

Bryant, D. and V. Moulton. 2002. NeighborNet: An agglomerative method for the construction of planar phylogenetic networks. *Algorithms in Bioinformatics* 375–391.

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.

Charlet, D. A. 2007. Distribution patterns of Great Basin conifers: Implications of extinction and immigration. *Aliso: A Journal of Systematic and Floristic Botany* 24: 31–61.

Coates, D. J., M. Byrne, and C. Moritz. 2018. Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution* 6: 165.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, and 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics* 27: 2156–2158.

Darwin, C. 1877. *The Different Forms of Flowers on Plants of the Same Species*. London: John Murray.

Davidson, J. B., S. L. Durham, and P. G. Wolf. 2014. Breeding system of the threatened endemic *Primula cusickiana* var. *maguirei* (Primulaceae). *Plant Species Biology* 29: E55–E63.

De Queiroz, T. F., C. Baughman, O. Baughman, M. Gara, and N. Williams. 2012. Species distribution modeling for conservation of rare, edaphic endemic plants in White River Valley, Nevada. *Natural Areas Journal* 32: 149–158.

De Vos, J. M., C. E. Hughes, G. M. Schneeweiss, B. R. Moore, and E. Conti. 2014. Heterostyly accelerates diversification via reduced extinction in primroses. *Proceedings. Biological Sciences* 281: 20140075.

DeChaine, E. G. and A. P. Martin. 2005a. Historical biogeography of two alpine butterflies in the Rocky Mountains: Broad-scale concordance and local-scale discordance. *Journal of Biogeography* 32: 1943–1956.

Dechaine, E. G. and A. P. Martin. 2005b. Marked genetic divergence among sky island populations of *Sedum lanceolatum* (Crassulaceae) in the Rocky Mountains. *American Journal of Botany* 92: 477–486.

Eaton, D. A. R. and I. Overcast. 2020. ipyrad: interactive assembly and analysis of RADseq datasets. *Bioinformatics* 36: 2592–2594.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Pol, K. Kawamoto, E. S. Buckler, and E. Sharon. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* 6: e19379.

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.

US Fish and Wildlife Service. 1985. *Endangered and Threatened Wildlife and Plants: Final rule to determine Primula maguirei (Maguire's primrose) to be a threatened species* (Vol. 50, pp. 49–52). Department of the Interior.

Fleishman, E., G. T. Austin, and D. D. Murphy. 2001. Biogeography of Great Basin butterflies: Revisiting patterns, paradigms, and climate change scenarios. *Biological Journal of the Linnean Society. Linnean Society of London* 74: 501–515.

Floyd, C. H., D. H. Van Vuren, and B. May. 2005. Marmots on great basin mountaintops: Using genetics to test a biogeographic paradigm. *Ecology* 86: 2145–2153.

Gray, A. 1888. *Synoptical Flora of North America*. Ivison, Blakeman, Taylor & Co.

Harper, K. T., C. D. Freeman, K. W. Ostler, and L. G. Klikoff. 1978. The flora of Great Basin mountain ranges: Diversity sources and dispersal ecology. *Great Basin Naturalist Memoirs* 2: 81–103.

He, X., J. J. Cao, W. Zhang, Y. Q. Li, C. Zhang, X. H. Li, G. H. Xia, and J. W. Shao. 2021. Integrative taxonomy of herbaceous plants with narrow fragmented distributions: A case study on *Primula merrilliana* species complex. *Journal of Systematics and Evolution*: DOI: 10.1111/jse.12726.

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

Hodel, R. G. J., R. Massatti, S. G. D. Bishop, and L. L. Knowles. 2021. Testing which axes of species differentiation underlie covariance of phylogeographic similarity among montane sedge species. *Evolution* 75: 349–364.

Holmgren, N. H. 1967. A new species of primrose from Nevada. *Madroño* 19: 27–29.

Holmgren, N. H. and A. H. Holmgren. 1974. Three new species from the Great Basin. *Brittonia* 26: 309–315.

Holmgren, N. H., P. K. Holmgren, and A. Cronquist. 2005. *Intermountain Flora: Vascular Plants of the Intermountain West, USA, vol. 2, Part B: Subclass Dilleniidae*. New York: New York Botanical Garden.

Holmgren, N. H. and S. Kelso. 2001. *Primula cusickiana* (Primulaceae) and its varieties. *Brittonia* 53: 154–156.

Huang, J. 2020. Is population subdivision different from speciation? From phylogeography to species delimitation. *Ecology and Evolution* 36: 303.

Huang, Y. F., L. N. Dong, and W. B. Xu. 2019. *Lysimachia fanii*, a new species of Primulaceae from limestone area of Guangxi, China. *PhytoKeys* 130: 75–84.

Hulshof, C. M. and M. J. Spasojevic. 2020. The edaphic control of plant diversity. *Global Ecology and Biogeography* 29: 1634–1650.

Huson, D. H. and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11: 94.

Kass, R. and S. Welsh. 1985. New species of *Primula* (Primulaceae) from Utah. *The Great Basin Naturalist* 45: 548–550.

Kelso, S. 1992. The genus *Primula* as a model for evolution in the Alaskan flora. *Arctic and Alpine Research* 24: 82–87.

Kelso, S., P. M. Beardsley, and K. Weitemier. 2009. Phylogeny and Biogeography of *Primula* sect. *Parryi* (Primulaceae). *International Journal of Plant Sciences* 170: 93–106.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. CLUMPAK: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources* 15: 1179–1191.

Koontz, A., W. D. Pearse, and P. Wolf. 2022. Data from: Pronounced genetic separation among varieties of the *Primula cusickiana* species complex, a Great Basin endemic. Dryad Digital Repository. https://doi.org/10.5061/dryad.12jm63xzv.

Lawlor, T. E. 1998. Biogeography of Great Basin mammals: Paradigm lost? *Journal of Mammalogy* 79: 1111–1130.

Leinonen, T., R. B. O'Hara, J. M. Cano, and J. Merilä. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology* 21: 1–17.

Li, J., J. M. Cocker, J. Wright, M. A. Webster, M. McMullan, S. Dyer, D. Swarbreck, M. Caccamo, C. van Oosterhout, and P. M. Gilmartin. 2016. Genetic architecture and evolution of the S locus supergene in *Primula vulgaris*. *Nature Plants* 2: 16188.

Mansfield, D. 1993. The status of "*Primula wilcoxiana*" and the *Primula cusickiana* complex in Southwestern Idaho. Bureau of Land Management, U.S. Department of the Interior.

Mansfield, D. 2010. Vascular flora of the Owyhee River watershed in Oregon. *Journal of the Idaho Academy of Science* 46: 1–127.

Mast, A. R. and J. L. Reveal. 2007. Transfer of *Dodecatheon* to *Primula* (Primulaceae). *Brittonia* 59: 79–82.

Novembre, J. 2016. Pritchard, Stephens, and Donnelly on population structure. *Genetics* 204: 391–393.

Nowak, C. L., R. S. Nowak, R. J. Tausch, and P. E. Wigand. 1994. Tree and shrub dynamics in northwestern Great Basin woodland and shrub steppe during the Late-Pleistocene and Holocene. *American Journal of Botany* 81: 265–277.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

R Core Team. 2020. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Rajakaruna, N. 2018. Lessons on evolution from the study of edaphic specialization. *Botanical Review* 84: 39–78.

Ren, G., R. G. Mateo, E. Conti, and N. Salamin. 2020. Population genetic structure and demographic history of *Primula fasciculata* in Southwest China. *Frontiers in Plant Science* 11: 986.

Richards, A. J. 2003. *Primula*. Portland, Oregon: Timber Press.

Schmidt-Lebuhn, A. N., J. M. de Vos, B. Keller, and E. Conti. 2012. Phylogenetic analysis of *Primula* section *Primula* reveals rampant non-monophyly among morphologically distinct species. *Molecular Phylogenetics and Evolution* 65: 23–34.

Schorr, G., P. B. Pearman, A. Guisan, and J. W. Kadereit. 2013. Combining palaeodistribution modelling and phylogeographical approaches for identifying glacial refugia in alpine *Primula*. *Journal of Biogeography* 40: 1947–1960.

Schott, H. W. 1852. Oesterreichische Primeln. *Österreichisches Botanisches Wochenblatt* 2: 35–36.

Sukumaran, J. and L. L. Knowles. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences USA* 114: 1607–1612.

Theodoridis, S., D. Nogués-Bravo, and E. Conti. 2019. The role of cryptic diversity and its environmental correlates in global conservation status assessments: Insights from the threatened bird's-eye primrose (*Primula farinosa* L.). *Diversity & Distributions* 25: 1457–1471.

Thompson, R. S. and J. I. Mead. 1982. Late Quaternary environments and biogeography in the Great Basin. *Quaternary Research* 17: 39–55.

Trift, I., M. Källersjö, and A. A. Anderberg. 2002. The monophyly of *Primula* (Primulaceae) evaluated by analysis of sequences from the chloroplast gene *rbcL*. *Systematic Botany* 27: 396–407.

Weir, B. S. and C. C. Cockerham. 1984. Estimating F-Statistics for the analysis of population structure. *Evolution* 38: 1358–1370.

Williams, L. O. 1936. Revision of the Western Primulas. *American Midland Naturalist* 17: 741–748.

Wolf, P. G. and R. B. Sinclair. 1997. Highly differentiated populations of the narrow endemic plant Maguire Primrose (*Primula maguirei*). *Conservation Biology* 11: 375–381.

Wright Smith, W. and H. R. Fletcher. 1949. The genus *Primula*: Sections *Cuneifolia, Floribundae, Parryi*, and *Auricula*. *Transactions of the Royal Society of Edinburgh* 61: 631–686.

Yang, L., H. Kong, J. P. Huang, and M. Kang. 2019. Different species or genetically divergent populations? Integrative species delimitation of the *Primulina hochiensis* complex from isolated karst habitats. *Molecular Phylogenetics and Evolution* 132: 219–231.

Zhang, L., H. P. Comes, and J. W. Kadereit. 2004. The temporal course of Quaternary diversification in the European high mountain endemic *Primula* sect. *Auricula* (Primulaceae). *International Journal of Plant Sciences* 165: 191–207.

Zhang, L. B. and J. W. Kadereit. 2004. Classification of *Primula* sect. *Auricula* (Primulaceae) based on two molecular data sets (ITS, AFLPs), morphology and geographical distribution. *Botanical Journal of the Linnean Society* 146: 1–26.

Appendix 1. List of voucher specimens included in this study. Order of data is as follows: Species, Voucher, Herbarium. Institutional barcodes or accession numbers are included as parenthetical values following the voucher, when available.

Ingroup: *Primula cusickiana* var. *cusickiana*, 25330978, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25330990, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331045, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331062, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331021, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331015, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331018, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331034, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25331004, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25330994, Intermountain Herbarium; *Primula cusickiana* var. *cusickiana*, 25330991, Intermountain Herbarium; *Primula cusickiana* var. *maguirei*, 25331026, Intermountain Herbarium; *Primula cusickiana* var. *maguirei*, 25331039, Intermountain Herbarium; *Primula cusickiana* var. *maguirei*, 25331041, Intermountain Herbarium; *Primula cusickiana* var. *nevadensis*, 25331101, Intermountain Herbarium; *Primula cusickiana* var. *nevadensis*, 25331106, Intermountain Herbarium; *Primula cusickiana* var. *nevadensis*, 25331092, Intermountain Herbarium; *Primula cusickiana* var. *domensis*, 25331066, Intermountain Herbarium; *Primula cusickiana* var. *domensis*, 25331070, Intermountain Herbarium; *Primula cusickiana* var. *domensis*, 25331077, Intermountain Herbarium; *Primula cusickiana* var. *domensis*, 25331083, Intermountain Herbarium.

Outgroups: *Primula capillaris*, 770850 (ASU0020421), Arizona State University Vascular Plant Herbarium; *Primula capillaris*, 3025822 (UTC00138833), Intermountain Herbarium; *Primula parryi*, 25331110, Intermountain Herbarium; *Primula parryi*, 25331112, Intermountain Herbarium.