RESEARCH ARTICLE

AMERICAN JOURNAL OF BOTANY

Pinus ponderosa: A checkered past obscured four species[1]

Ann Willyard[2,10], David S. Gernandt[3], Kevin Potter[4], Valerie Hipkins[5], Paula Marquardt[6], Mary Frances Mahalovich[7], Stephen K. Langer[8], Frank W. Telewski[9], Blake Cooper[2], Connor Douglas[2], Kristen Finch[2], Hassani H. Karemera[2], Julia Lefler[2], Payton Lea[2], and Austin Wofford[2]

PREMISE OF THE STUDY: Molecular genetic evidence can help delineate taxa in species complexes that lack diagnostic morphological characters. *Pinus ponderosa* (Pinaceae; subsection *Ponderosae*) is recognized as a problematic taxon: plastid phylogenies of exemplars were paraphyletic, and mitochondrial phylogeography suggested at least four subdivisions of *P. ponderosa*. These patterns have not been examined in the context of other *Ponderosae* species. We hypothesized that putative intraspecific subdivisions might each represent a separate taxon.

METHODS: We genotyped six highly variable plastid simple sequence repeats in 1903 individuals from 88 populations of *P. ponderosa* and related *Ponderosae* (*P. arizonica*, *P. engelmannii*, and *P. jeffreyi*). We used multilocus haplotype networks and discriminant analysis of principal components to test clustering of individuals into genetically and geographically meaningful taxonomic units.

KEY RESULTS: There are at least four distinct plastid clusters within *P. ponderosa* that roughly correspond to the geographic distribution of mitochondrial haplotypes. Some geographic regions have intermixed plastid lineages, and some mitochondrial and plastid boundaries do not coincide. Based on relative distances to other species of *Ponderosae*, these clusters diagnose four distinct taxa.

CONCLUSIONS: Newly revealed geographic boundaries of four distinct taxa (*P. benthamiana*, *P. brachyptera*, *P. scopulorum*, and a narrowed concept of *P. ponderosa*) do not correspond completely with taxonomies. Further research is needed to understand their morphological and nuclear genetic makeup, but we suggest that resurrecting originally published species names would more appropriately reflect the taxonomy of this checkered classification than their current treatment as varieties of *P. ponderosa*.

KEY WORDS Pinaceae; *Pinus*; plastid microsatellites; ponderosa pine; *Ponderosae*

Conflicting morphological data can make it challenging to assign populations of related plants to meaningful taxonomic units. Although not the panacea we once anticipated, molecular genetic evidence has helped clarify some species delineation questions. These data have contributed greatly to our understanding of many taxonomic groups, including species complexes that were affected by hybridization (Manos et al., 1999). Many species of *Pinus* (Pinaceae) have clear morphological diagnostic characters, but ponderosa pine taxonomy remains unsettled despite decades of research (Lauria, 1996a). Five taxonomic varieties are commonly accepted within the very broad geographic range of *Pinus ponderosa* Douglas ex C.Lawson. Two of these varieties are clearly distinct using a combination of genetic and climate differences, along with overlapping ranges of morphological character states: *Pinus ponderosa* var. *ponderosa* and *Pinus ponderosa* var. *scopulorum* Engelm. (Conkle and Critchfield, 1988; Potter et al., 2013, 2015). The distinctions are less clear for *Pinus ponderosa* var. *benthamiana* (Hartw.) Vasey, *Pinus ponderosa* var. *brachyptera* (Engelm.) Lemmon, and *Pinus ponderosa* var. *washoensis* (H.Mason & Stockw.) J.R.Haller & Vivrette. Four of these taxa were originally published as unique species and *Pinus scopulorum* (Engelm.) Lemmon was elevated to species rank in 1897, just 17 years after its publication as a variety of

[1] Manuscript received 16 September 2016; revision accepted 15 November 2016.

[2] Hendrix College, 1600 Washington Ave, Conway, Arkansas 72032 USA;
[3] Departmento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, A.P. 70-233, Mexico City 04510 México;
[4] Department of Forestry and Environmental Resources, North Carolina State University, Research Triangle Park, North Carolina 27709 USA;
[5] National Forest Genetics Laboratory, USDA Forest Service, Placerville, California 95667 USA;
[6] USDA Forest Service, Northern Research Station, Institute for Ecosystem Studies, 5985 Highway K, Rhinelander, Wisconsin 54501 USA;
[7] Genetic Resource Program, Northern, Rocky Mountain, Southwestern, and Intermountain Regions, USDA Forest Service, Moscow, Idaho 83843 USA;
[8] 301 North Main #214, Layton, Utah 84041 USA; and
[9] W. J. Beal Botanical Garden, Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824 USA
[10] Author for correspondence (e-mail: willyard@hendrix.edu)

doi:10.3732/ajb.1600336

P. ponderosa. Reviews and treatments have reached very different conclusions regarding whether they should be accepted and, if so, whether they are varieties of *P. ponderosa*, where each is distributed geographically, and which morphological characters reliably distinguish them (Lauria, 1991, 1996a, 1997; Millar and Libby, 1991; Kral, 1993; Haller and Vivrette, 2011; Baldwin et al., 2012; Callaham, 2013a; Meyers et al., 2015). The question of taxonomic delineation also occurs more widely across *Pinus* subsection *Ponderosae* Loudon (section *Trifoliae*; subgenus *Pinus*). For example, it is unclear whether subsection *Sabinianae* Loudon—the California big-coned pines, *Pinus coulteri* D.Don (Coulter pine), *Pinus jeffreyi* A.Murray bis (Jeffrey pine), *Pinus sabiniana* Douglas (gray pine), and *Pinus torreyana* Parry ex Carrière (Torrey pine)—is sister to subsection *Ponderosae* Loudon (Willyard et al., 2009) or whether these four species are nested within subsection *Ponderosae* (Gernandt et al., 2009; Parks et al., 2012). Molecular phylogenies of *Ponderosae* either cannot rule out incomplete lineage sorting or resolve exemplars of most species as not monophyletic (Gernandt et al., 2009; Willyard et al., 2009). For example, full plastome nucleotide sequences found three samples of *P. ponderosa* to be paraphyletic: (1) *P. ponderosa* var. *ponderosa* from Montana was sister to the remaining *Ponderosae*; (2) *P. ponderosa* var. *scopulorum* from South Dakota was sister to a clade that included *Pinus douglasiana* Martínez, *Pinus engelmannii* Carrière, *Pinus cooperi* C.E.Blanco, and *Pinus arizonica* Engelm.; and (3) *P. ponderosa* var. *benthamiana* from Butte County, California, was sister to subsection *Sabinianae* (Parks et al., 2012). Unfortunately, these gene trees included only a few exemplars of each named species and could not test the status of the named varieties of *P. ponderosa* (Gernandt et al., 2009; Willyard et al., 2009). However, it seems very likely that some populations that are currently treated as varieties are actually more closely related to other *Ponderosae*.

Despite some acceptance of the named varieties of *P. ponderosa*, conflicting intraspecific delineations have been inferred from growth, isozyme, terpene, and other types of data (Weidman, 1939; Smith, 1964; Wells, 1964a; Read, 1980; Conkle and Critchfield, 1988; Callaham, 2013b). A unified interpretation of these results is difficult because different populations were sampled (often emphasizing one portion of the geographic range), and data were published as mean values within hypothesized groupings, masking differences within the tested groups or for different boundaries. Recently published range-wide experiments for *P. ponderosa* revealed the mitochondrial haplotype distribution (Potter et al., 2013), nuclear simple sequence repeat (nSSR) patterns (Potter et al., 2015), and climate niches for the same populations (Shinneman et al., 2016). Together, these data strongly suggest western and eastern subdivisions within *P. ponderosa*, supporting a well-known contact zone in Montana (Latta and Mitton, 1999) and a lesser-studied contact area in southern California. But these new data also conflict with previous delineations on a finer scale. For example, populations in California, Oregon, and Washington have mitochondrial diversity that does not correspond to previous hypotheses for the range of *P. ponderosa* var. *benthamiana* (Lauria, 1996b), for *Pinus ponderosa* var. *pacifica* J.R.Haller&Vivrette (Haller and Vivrette, 2011), or for *Pinus ponderosa* subsp. *critchfieldiana* Callaham (Callaham, 2013a). In the eastern part of the range, a widespread mitochondrial haplotype occurs in populations that have been assigned to *P. ponderosa* var. *brachyptera* (the "southwestern form"), but this mitochondrial haplotype also extends far to the north in populations traditionally assigned to *P. ponderosa* var. *scopulorum* (the "Rocky Mountain form") (Potter et al., 2013). Unique mitochondrial haplotypes were also identified in a region of southeastern Nevada, southwestern Utah, and far northwestern Arizona for which no taxon has been published (Potter et al., 2013).

Resolution of this issue has been complicated by reports of low levels of introgressive hybridization (introgression) with sympatric species. In California, plastid transfers between *P. ponderosa* and *P. jeffreyi* (Willyard et al., 2009) are found at roughly the same frequency as morphological intermediates (Haller, 1962). There is also evidence for infrequent hybridization in southern Arizona between *P. ponderosa* and *P. arizonica* (Epperson et al., 2009) and between *P. ponderosa* and *P. engelmannii* (Peloquin, 1984; Rehfeldt, 1999a). Sampling to date has been inadequate to test the impact of introgression on the genetic patterns observed in *P. ponderosa*. Experiments have either sampled only a few *P. ponderosa* exemplars along with other *Ponderosae*, or they have sampled *P. ponderosa* widely but did not include sympatric *P. jeffreyi*, *P. arizonica*, or *P. engelmannii*. Thus, it is still unknown whether any of the unexpected mitochondrial haplotypes observed in *P. ponderosa* (Potter et al., 2013) may have been acquired from introgression with *P. jeffreyi*, *P. arizonica*, or *P. engelmannii*.

The taxonomy of *P. ponderosa* was aptly described as "chequered" (Lauria, 1996a, p. 1023) using one definition of this word as having a history of varied fortune or discreditable incidents. In addition to species publications with limited peer review in seed magazines, the type specimens for *P. ponderosa*, *P. benthamiana* Hartw., *P. brachyptera* Engel., and *P. ponderosa* var. *scopulorum* Engelm are all unavailable (Lauria, 1996a), and type localities were only described broadly or have been reconstructed more than a century later. For example, the typification of *P. ponderosa* could be based on David Douglas' collection in the spring of 1826 of a sterile branch to which a specimen of *Arceuthobium* was attached or to the trees growing in present-day England from seed that Douglas asked John Work (a trader for the Hudson Bay Company) to collect later in the fall of 1826 (Lauria, 1996a). In another example, it has been argued that because George Engelmann probably used specimens from across the Rocky Mountains to describe *P. ponderosa* var. *scopulorum*, there are 10 syntypes from seven different states in the United States that support his description of this taxon (Lauria, 1996a). This absence of type specimens has inspired several attempts to assign neotypes (Haller and Vivrette, 2011; Callaham, 2013a). One suggestion was to use specimens collected from legacy trees growing in European gardens (Lauria, 1991) despite the lack of documentation for their seed source. Some important contributions have avoided nomenclature altogether, e.g., publishing a putatively unique taxon in the Sky Islands of Arizona as "Taxon X" (Rehfeldt, 1999a). The lack of clear type specimens created a nomenclatural conundrum in this species complex overlaying the lack of clear morphological and genetic differences. Because the geographic extent of any potentially unique taxon is unclear, the application of published names is problematic. One example is that the *benthamiana* epithet could be applied to just the ponderosa pines of the Santa Cruz Mountains if the populations growing on sand hill formations interspersed within redwood forests were found to be unique (Griffin, 1964), or to the populations of the Santa Cruz Mountains and the Klamath Mountains if they formed a biologically meaningful unit (Lauria, 1996b), or to all of the coastal ponderosa pines (including Oregon's Willamette Valley and the populations at Fort Lewis, WA) (Meyers et al., 2015). Alternatively, this name could be abandoned in favor of *P. ponderosa*

var. *pacifica* with a much wider definition to include the ponderosa pines on the western flank of the Sierra Nevada (Haller and Vivrette, 2011). These taxonomic decisions have been stalemated by the conflicting subdivisions suggested in the studies described above. When viewed on maps of the western United States, these published delineations reminded our research team of a many-layered mosaic sculpture. We hypothesized that if plastid lineages do not match the mitochondrial distributions or any of the previous treatments, this species complex might have a checkered present as well as a checkered past, using the other definition of this word—a pattern of alternating squares of different colors. Perhaps these populations are evolving as a genetic mosaic (i.e., carrying different plastid, mitochondrial, and nuclear lineages). Migrations, enhanced by occasional introgression with rather distantly related species such as *P. jeffreyi*, *P. arizonica*, and *P. engelmannii*, may have strongly affected genome distribution over the landscape, as previously reported in pines (Liston et al., 2007; Willyard et al., 2009). It may be that any one population (and possibly any one individual) is carrying disparate lineages of plastid, mitochondrial, and nuclear genes. This genomic mosaic could contribute to the long-recognized within-population variation and plastic growth responses of individual plants in subsection *Ponderosae* (Zhang and Cregg, 2005; Callaham, 2013a) that have so far stymied attempts to recognize morphologically distinct taxa.

We here report a range-wide assessment of plastid diversity in *P. ponderosa* and related taxa using a criterion of genotypic clusters to infer the existence of taxa that have existed as a lineage for some period of time (Mallet, 1995). We evaluated the relative separation of genotypic clusters and the relative distances among those clusters to judge meaningful assignments to species or to intraspecific varieties using plastid data from *P. ponderosa* and from sympatric *Ponderosae* and *Sabinianae* populations. Operational taxonomic units (OTUs) were used to test subdivisions within *P. ponderosa*, to test how *P. arizonica* var. *stormiae* Martínez fits within the typical variety, and to test whether *P. jeffreyi* from serpentine soils in the Klamath Mountains has a plastid lineage that is distinct from *P. jeffreyi* in the Sierra Nevada. The mosaic idea was tested by comparing these plastid results with published mitochondrial haplotype patterns.

We used plastid simple sequence repeats (cpSSRs) that are highly variable in *P. ponderosa* (Wofford et al., 2013). Because they are so variable, we could not rely on private haplotypes to define an OTU. However, we were able to examine whether each OTU is supported by differing frequencies of plastid haplotypes, and we used haplotype relationships from a minimum spanning network (MSN) to infer phylogeographic relationships. By also sampling sympatric species, we could infer which haplotypes were possibly retained from ancestors as opposed to haplotypes that may have been introduced to the population via admixture. It is important to note that estimations of admixture are always maximum values because the possibility cannot be ruled out that some individuals are carrying a plastid lineage from a shared ancestor. Their high level of incomplete lineage sorting makes this an important factor between *P. ponderosa* and these sympatric taxa (Willyard et al., 2009). Our sampling scheme also allowed us to use relative distances to evaluate putative subdivisions of *P. ponderosa*: if an OTU is more distant to other *P. ponderosa* OTUs than it is to a species diagnosed by published criteria (e.g., *P. arizonica*, *P. engelmannii*, or *P. jeffreyi*), then it makes no biological sense to lump that OTU within *P. ponderosa*. Our final goal was to evaluate which taxonomic units inferred by other types of evidence gain support from distinctive plastid haplotype clusters. By evaluating the plastid genetic structure across the entire range of ponderosa pine and related taxa, we confirmed that *P. ponderosa*, as currently treated, does not form a single genotypic cluster. Instead, a comparison of plastid genotypic clusters and the plastid MSNs with mitochondrial phylogeography support at least four distinct lineages.

MATERIALS AND METHODS

Plant material—We collected 1903 samples from 88 populations (Table 1; Fig. 1). Leaf or terminal bud tissue was collected from trees spaced at least 100 m apart within each population. Tissue was either dried immediately on silica gel or kept chilled until frozen. We collected one herbarium specimen per population, and these were vouchered at the Institute of Forest Genetics, Pacific Southwest Research Station, USDA Forest Service (IFGP), Universidad Nacional Autónoma de Mexico (MEXU), Oregon State University (OSC), and the Sul Ross Herbarium (SRSC) (Appendix 1).

OTU assignments—Each population was originally categorized as one of 16 prior OTUs (Table 1; Fig. 1). We assigned 73 populations to 11 OTUs within *P. ponderosa*. The remaining 15 populations represented five OTUs that were identified morphologically as belonging to taxa other than *P. ponderosa*. Three of these OTUs represented other species of *Ponderosae* that are partially sympatric with *P. ponderosa*: *P. engelmannii*, *P. arizonica*, and *P. arizonica* var. *stormiae*. Two OTUs represented *P. jeffreyi*, which is more distantly related (subsection *Sabinianae*) but is also partially sympatric and capable of hybridizing with *P. ponderosa*. Except for *P. coulteri*, we included samples of all sympatric species that could potentially hybridize with *P. ponderosa*. An isolated case of introgression between *P. coulteri* and *P. ponderosa* was suggested based on intermediate terpene composition (Smith, 1967) but was not confirmed with other data, and artificial crosses between *P. coulteri* and *P. ponderosa* were not successful (Conkle and Critchfield, 1988). In a preliminary phase of our experiment, we amplified these cpSSR loci in a population of *P. coulteri* but found them to be very divergent (data not shown).

OTU A (Pacific Northwest) combines the relatively isolated ponderosa pine populations from the Willamette Valley, Oregon and the Puget Sound Basin, Washington, which have been proposed to inhabit a distinctive ecological niche and to possibly have distinctive characters (Wells, 1964a; Gooding, 1998; Bouffier et al., 2003; Gerson and Kelsey, 2004). The Puget Sound Basin population at Fort Lewis, Washington was reported to share a mitochondrial haplotype with trees from the Klamath Range in California, whereas a Willamette Valley, Oregon collection shared a mitochondrial haplotype with trees from the Blue Mountains, Oregon (Potter et al., 2013). We chose the OTU name to reflect geography because no named variety encompasses just these populations.

OTU B (Klamath) was defined to test a mitochondrial haplotype observed only in the Klamath Range, California (Potter et al., 2013). The ponderosa pines of this region were reported to vary in monoterpenes (Smith et al., 1969). A distinct species in this region might correspond to *P. beardsleyi* A.Murray from Scott Mountain, California. Because this epithet has not been in general use, we chose the OTU name to reflect geography.

TABLE 1. Geographic locations, prior, and posterior OTU assignments of 88 *Pinus* populations analyzed with cpSSRs.

							All 88 populations									73 populations		
							k = 16			*k* = 15			*k* = 9			*k* = 11		
Pop.	Location	Latitude	Longitude	*h*	*N*	nMLH	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU
Pinus ponderosa (OTU A—OTU K)																		
01	CA, Larabee Valley	40.43	−123.68	0.138	21	7	K	D	86	K	D	86	K	C+D+E	90	K	D	86
02	CA, Willow Creek	40.84	−123.68	0.503	24	16	B	—	33	B	—	33	B	F+G+J	42	B	—	46
03	CA, Happy Camp	41.88	−123.41	0.126	24	8	B	—	75	B	—	75	B	—	75	B	—	96
04	OR, Eugene	43.92	−123.13	0.237	21	8	C	D	76	C	D	76	C+D+E	—	81	C	D	76
05	CA, Hayfork Summit	40.60	−123.02	0.189	24	11	B	—	75	B	—	75	B	—	75	B	—	88
06	CA, Basin Gulch Campground	40.35	−122.96	0.180	14	6	B	E	93	B	E	93	B	C+D+E	93	B	E	93
07	OR, Willamette Valley	44.38	−122.91	0.210	25	4	A	—	72	A	—	72	A	—	72	A	—	68
08	CA, Clearlake	39.12	−122.79	0.170	24	6	B	—	88	B	—	88	B	—	88	B	—	100
09	OR, Willamette Valley	45.35	−122.77	0.156	25	4	A	—	80	A	—	80	A	—	80	A	—	84
10	OR, Tiller	42.90	−122.71	0.130	18	6	C	B	89	C	B	89	C+D+E	B	67	C	B	89
11	CA, Shasta County	40.67	−122.70	0.087	14	4	B	A	86	B	A	86	B	A	86	B	A	86
12	WA, Fort Lewis	47.07	−122.48	0.132	25	7	A	—	80	A	—	80	A	—	80	A	—	84
13	CA, UC Santa Cruz Arboretum	36.98	−122.06	0.175	19	4	K	A	68	K	A	68	K	A	65	K	A	89
14	CA, Quail Hollow	37.09	−122.06	0.118	22	5	K	O	77	K	B	77	K	B	77	K	B	77
15	CA, Henry Cowell Redwoods	37.03	−122.05	0.153	25	3	K	A	92	K	A	92	K	A	92	K	A	68
16	CA, McCloud (Happy Camp OP)	41.28	−121.95	0.180	22	7	C	—	82	C	—	82	C+D+E	—	83	C	—	86
17	OR, Santiam Pass	44.42	−121.77	0.159	25	7	C	A	76	C	A	76	C+D+E	A	76	C	A	72
18	CA, Paynes Creek	40.34	−121.73	0.228	20	9	C	K	55	C	K	55	C+D+E	K	55	C	A	60
19	CA, Big Creek	36.06	−121.57	0.130	25	8	K	A	84	K	A	84	K	A	84	K	A	92
20	CA, Henry Coe State Park	37.19	−121.54	0.213	23	9	K	—	65	K	—	65	K	—	65	K	A	70
21	CA, Santa Lucia	35.89	−121.44	0.266	23	10	K	D	74	K	D	74	K	C+D+E	78	K	D	78
22	CA, Pollock Pines	38.78	−120.46	0.503	21	16	C	—	67	C	—	67	C+D+E	—	63	C	—	62
23	CA, Likely	41.23	−120.41	0.468	12	9	C	E	75	C	E	67	C+D+E	—	75	C	E	75
24	CA, Likely	41.19	−120.22	0.346	17	11	D	—	71	D	—	71	C+D+E	—	71	D	—	71
25	CA, Babbitt Peak	39.61	−120.11	0.272	20	9	D	—	65	D	—	65	C+D+E	—	71	D	—	65
26	CA, Figueroa Mtn.	34.74	−119.98	0.200	22	7	E	—	95	E	—	95	C+D+E	—	95	E	—	95
27	NV, Mt. Rose	39.33	−119.87	0.290	19	10	D	E	79	D	E	79	C+D+E	—	79	D	E	79
28	CA, Wawona	37.59	−119.70	0.283	23	14	C	E	91	C	E	91	C+D+E	—	91	C	E	91
29	OR, Blue Mtns.	44.07	−118.79	0.147	21	6	C	A	81	C	A	81	C+D+E	A	81	C	A	81
30	OR, Blue Mtns.	44.07	−118.79	0.313	22	12	C	D	45	C	D	45	C+D+E	—	50	C	D	59
31	CA, Breckenridge Mtn.	35.49	−118.57	0.315	22	14	C	E	91	C	E	91	C+D+E	—	91	C	E	91
32	CA, Lake Isabella	35.78	−118.56	0.379	18	17	C	E	78	C	E	78	C+D+E	—	78	C	E	78
33	CA, Bishop Creek	37.29	−118.56	0.236	23	9	E	—	100	E	—	100	C+D+E	—	100	E	—	100
34	CA, Onion Valley	36.79	−118.29	0.268	23	10	E	—	100	E	—	100	C+D+E	—	100	E	—	100
35	WA, Bisbee Mtn.	48.62	−118.17	0.051	25	4	C	A	88	C	A	88	C+D+E	A	88	C	A	100
36	CA, San Bernardino	34.25	−117.30	0.279	19	10	E	—	95	E	—	95	C+D+E	—	95	E	—	95
37	ID, Cour D'Alene	47.71	−116.86	0.246	19	7	C	E	58	C	E	58	C+D+E	—	58	C	E	63
38	CA, San Jacinto Mtns.	33.81	−116.77	0.242	19	7	E	—	89	E	—	89	C+D+E	—	89	E	—	89
39	ID, Last Chance Campground	44.79	−116.20	0.231	24	12	C	—	79	C	—	79	C+D+E	—	83	C	—	83
40	ID, Boise	43.78	−115.89	0.432	22	15	C	—	45	C	—	45	C+D+E	—	63	C	—	45
41	ID, Kooskia Rd.	46.14	−115.75	0.330	24	13	C	B	46	C	B	46	C+D+E	B	46	C	B	50
42	NV, Spring Mtns.	36.32	−115.68	0.410	18	8	J	G	50	J	G	50	F+G+J	—	83	J	G	50
43	NV, Kyle Canyon	36.26	−115.61	0.320	19	12	J	—	74	J	—	74	F+G+J	—	90	J	—	79
44	NV, Grant Range	38.44	−115.44	0.340	16	11	F	—	75	F	—	75	F+G+J	—	80	F	—	69

continued

TABLE 1, *continued*

							All 88 populations									73 populations		
							k = 16			*k* = 15			*k* = 9			*k* = 11		
Pop.	Location	Latitude	Longitude	*h*	*N*	nMLH	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU
45	AZ, Hualapai Co. Park	35.10	−113.88	0.298	25	13	H	—	72	H	—	72	H	—	72	H	—	72
46	AZ, Black Rock Mtns.	36.80	−113.75	0.444	14	13	F	—	50	F	—	50	F+G+J	—	65	F	J	50
47	AZ, Jacob Lake	36.80	−112.26	0.319	24	14	H	I	75	H	I	75	H	I+N	71	H	I	71
48	UT, Red Canyon	37.73	−112.25	0.307	18	11	F	G	89	F	G	89	F+G+J	—	94	F	G	94
49	AZ, South Rim	35.38	−111.96	0.379	25	17	I	—	52	I	—	52	I+N	—	52	I	—	52
50	UT, Uinta Mtns.	40.63	−111.17	0.339	25	11	H	—	68	H	—	68	H	—	68	H	—	72
51	UT, Price Cyn. Rec. Area	39.76	−110.93	0.412	26	16	H	I	69	H	I	69	H	I+N	65	H	I	65
52	AZ, Mt. Hopkins	31.69	−110.88	0.590	21	14	I	J	43	I	J	33	I+N	F+G+J	86	I	J	52
53	AZ, Mt. Wrightson	31.70	−110.85	0.503	14	10	I	—	50	I	—	50	I+N	—	71	I	—	50
54	AZ, Mt. Lemmon	32.44	−110.79	0.391	22	416	I	—	86	I	—	77	I+N	—	86	I	—	82
55	AZ, Whitetail Campground	32.41	−110.71	0.343	24	12	I	—	46	I	N	46	I+N	—	88	I	H	54
56	AZ, Navajo Nation	36.55	−110.47	0.363	20	10	H	—	55	H	—	55	H	—	55	H	—	55
57	AZ, Huachuca Mtns.	31.26	−110.21	0.377	28	14	I	—	57	I	—	57	I+N	—	68	I	—	61
58	AZ, Pinaleno Mtns.	32.62	−109.83	0.354	29	11	I	—	69	I	—	69	I+N	—	62	I	—	69
59	MT, Judith Mtns.	47.13	−109.36	0.570	18	18	G	F	50	G	F	56	F+G+J	—	62	G	F	44
60	AZ, Barfoot Park	31.92	−109.28	0.310	26	14	I	—	88	I	—	88	I+N	—	81	I	—	88
61	NM, Pinos Altos	32.91	−108.23	0.365	19	12	I	—	58	I	—	58	I+N	—	58	I	—	63
62	CO, Dolores	37.73	−108.21	0.331	22	13	H	—	59	H	—	59	H	—	64	H	—	68
63	NM, Mt. Taylor	35.25	−107.73	0.417	24	15	H	—	46	H	—	46	H	—	46	H	—	50
64	WY, Bighorn Mtns.	44.31	−106.81	0.499	23	19	G	—	83	G	—	83	F+G+J	—	91	G	—	83
65	WY, Casper Mtns.	42.76	−106.33	0.524	25	21	G	I	36	G	I	36	F+G+J	I+N	40	G	H	36
66	NM, Santa Fe	35.73	−105.86	0.348	29	16	H	I	45	H	—	45	H	—	48	H	—	52
67	NM, Mescalero Apache Res.	33.07	−105.38	0.398	14	12	H	F	57	H	F	57	H	F+G+J	67	H	F	50
68	WY, Vedawoo Campground	41.16	−105.38	0.334	24	13	G	F	92	G	F	92	F+G+J	—	100	G	F	79
69	TX, Guadalupe Mtns.	31.89	−104.85	0.367	13	8	H	G	85	H	G	85	H	F+G+J	58	H	G	85
70	SD, Black Hills	43.85	−104.05	0.446	24	16	G	—	54	G	—	54	F+G+J	—	58	G	—	42
71	SD, Reva Gap	45.52	−103.16	0.495	12	9	G	J	33	G	J	33	F+G+J	—	44	G	H	25
72	NE, Chadron	42.70	−103.01	0.522	25	19	G	H	40	G	H	40	F+G+J	—	76	G	H	44
73	NE, Niobrara	42.79	−100.02	0.385	25	14	G	—	68	G	—	68	F+G+J	I+N	84	G	—	64
P. arizonica (OTU N)																		
74	AZ, Whitetail Campground	32.41	−110.71	0.111	25	8	N	—	84	N	—	84	I+N	—	95	n/a	n/a	n/a
75	AZ, Mt. Graham	32.63	−109.82	0.270	20	11	N	—	75	N	—	80	I+N	—	92	n/a	n/a	n/a
76	AZ, Chiricahua Mtns.	31.95	−109.31	0.072	12	3	N	—	75	N	—	92	I+N	F+G+J	100	n/a	n/a	n/a
P. arizonica var. *stormiae* (OTU M)																		
77	TX, Big Bend N. P., Crown Mtn.	29.26	−103.26	0.335	13	8	M	G	54	G	—	54	F+G+J	—	54	n/a	n/a	n/a
78	TX, Big Bend N. P., Pine Canyon	29.26	−103.25	0.415	22	13	M	H	45	H	—	45	H	—	45	n/a	n/a	n/a
P. engelmannii (OTU L)																		
79	AZ, Chiricahua Mtns.	31.95	−109.31	0.222	22	10	L	—	68	L	—	68	L	—	68	n/a	n/a	n/a
P. jeffreyi (subsect. *Sabinianae*; OTU O and OTU P)																		
80	CA, Horse Mtn. Botanical Area	40.87	−123.73	0.184	22	8	P	—	91	P	—	86	O+P	—	63	n/a	n/a	n/a
81	CA, Lassics Botanical Area	40.35	−123.55	0.417	23	14	P	—	74	P	—	74	O+P	F+G+J	46	n/a	n/a	n/a
82	CA, Happy Camp	41.95	−123.49	0.179	15	7	P	—	73	P	—	73	O+P	—	38	n/a	n/a	n/a
83	OR, Tiller	42.89	−122.95	0.148	18	4	P	C	61	P	C	61	O+P	C	48	n/a	n/a	n/a

continued

TABLE 1, *continued*

Pop.	Location	Latitude	Longitude	*h*	*N*	nMLH	All 88 populations								73 populations			
							k = 16			*k* = 15			*k* = 9			*k* = 11		
							Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU	Prior OTU	Posterior OTU if different	% to posterior OTU
84	CA, Snow Mtn.	39.36	−122.74	0.171	24	7	P	—	96	P	—	96	O+P	F+G+J	54	n/a	n/a	n/a
85	CA, Likely	41.23	−120.41	0.281	15	9	O	P	60	O	P	60	O+P	—	60	n/a	n/a	n/a
86	NV, Mt. Rose	39.33	−119.87	0.218	12	7	O	P	75	O	P	75	O+P	—	75	n/a	n/a	n/a
87	NV, Thomas Creek	39.39	−119.84	0.388	15	12	O	—	47	O	—	47	O+P	—	53	n/a	n/a	n/a
88	CA, Mammoth Lakes	37.64	−118.93	0.239	22	8	O	—	50	O	—	59	O+P	—	59	n/a	n/a	n/a
	Mean			0.297	21	10.5			69.8			69.9			72.2			71.7
	Total				1849	467												

Notes: *h*, diversity; *N*, number of samples analyzed; nMLH, number of multilocus haplotypes; for four hypotheses (*k* = 16, *k* = 15, *k* = 9, and *k* = 11), the prior OTU, posterior OTU where majority of individuals were assigned by discriminant analysis of principal components and percentage of individuals assigned to the majority OTU.

OTU C (Ponderosa) represents the general area for the type locality of *P. ponderosa*. It encompasses most of the geographic range of *P. ponderosa* var. *ponderosa* (Kral, 1993) except for populations that we assigned to OTU A (Pacific Northwest), OTU B (Klamath), or OTU E (Transverse). Other collection sites in the geographic region of our OTU C have been reported to contain two different mitochondrial haplotypes (Potter et al., 2013).

OTU D (Washoe) was based on the Washoe pine, currently treated as *P. washoensis* H.Mason & Stockw. (Kral, 1993) or as *P. ponderosa* var. *washoensis* (Haller and Vivrette, 2011). Although many doubts have been raised about the validity of this taxon, its status remains an open question (Wells, 1964a; Haller, 1965a; Smith, 1967, 1981; Critchfield, 1984; Niebling and Conkle, 1990; Sorensen, 1994; Lauria, 1997; Rehfeldt, 1999b; Patten and Brunsfeld, 2002).

Separating OTU E (Transverse) from OTU C was inspired by a mitochondrial haplotype in the ponderosa pines from the Transverse Range, California that differed from the nearby southern Sierra Nevada haplotypes. This haplotype was shared with isolated populations in southern Nevada and in southeastern New Mexico (Potter et al., 2013). Different monoterpene profiles were reported in the ponderosa pines of this geographic region (Smith et al., 1969; Smith, 1977), and other disjunct species occurrences in this area have been found (Major and Bamberg, 1967). We chose the OTU name to reflect geography because no named variety encompasses just these populations.

OTU F (Canyonlands) was created to test a unique mitochondrial haplotype identified in this region of southeastern Nevada and southwestern Utah (Potter et al., 2013). We chose the OTU name to reflect geography because no named variety encompasses just these populations.

Populations in Montana, South Dakota, Wyoming, and Nebraska were assigned to OTU G (Scopulorum), an area that encompasses the likely type locality of *P. ponderosa* var. *scopulorum* and much of the geographic distribution of *P. ponderosa* var. *scopulorum* as presently treated (Kral, 1993; Latta and Mitton, 1999). We tested a separate OTU H (Brachyptera) by assigning populations in northern Utah and all populations from Colorado, Arizona (except the Sky Islands of Arizona described below), and New Mexico (except *P. arizonica* var. *stormiae* described below) to this "southwest form" (Callaham, 2013b). This region includes the type locality for *P. brachyptera* east of Santa Fe, New Mexico. In drawing the line between OTU G and OTU H, we considered climatic regions from a growth experiment (Weidman, 1939), previously inferred ecotypes (Wells, 1964a; Haller, 1965b; Millar and Libby, 1991), and mitochondrial haplotypes (Potter et al., 2013).

OTU I (Sky Island) was based on two different studies that concluded the ponderosa pines in southern Arizona with three needles per fascicle were distinct. Growth in common gardens showed the pines in the Sky Islands of southern Arizona to be distinct from sympatric *P. engelmannii* and *P. arizonica* as well as from *P. ponderosa* (Rehfeldt, 1999a). A plastid haplotype and two low-copy nuclear gene trees placed three-needle pine samples from Mt. Lemmon, Arizona in a clade with pines of Mexico rather than with samples that would represent *P. ponderosa* var. *scopulorum* or *P. ponderosa* var. *brachyptera* (Epperson et al., 2009). We chose the Sky Island OTU name to reflect geography because no named variety encompasses just these populations.

OTU J (Spring Mountains) was based on a unique mitochondrial haplotype observed in an isolated group of populations in the

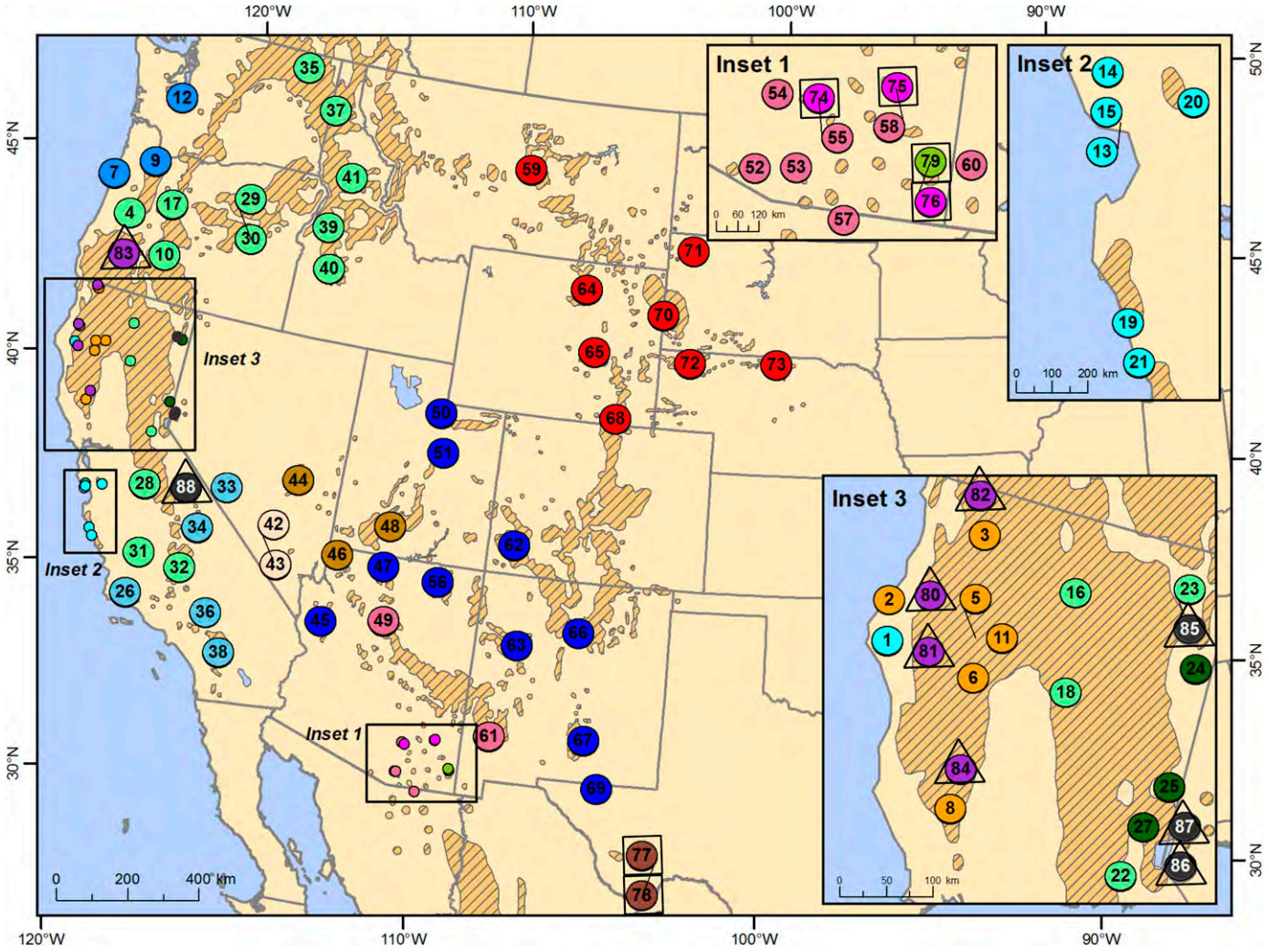

FIGURE 1 Geographic locations of the 88 *Pinus* populations sampled (see Table 1). Colors represent 16 prior OTU hypotheses (see legend for Fig. 2A). Squares are *P. engelmannii* or *P. arizonica*; triangles are the more distantly related *P. jeffreyi*.

far southern tip of Nevada (Potter et al., 2013). We are unaware of any other study reporting these populations as morphologically or genetically distinctive.

OTU K (Benthamiana) was limited to the ponderosa pines near Santa Cruz, CA. This geographic region includes the type locality for *P. benthamiana*, and most of the populations assigned to this OTU are from sand hill formations (Griffin, 1964). This much narrower definition (Lauria, 1996b) differs from a recent treatment of *P. ponderosa* var. *benthamiana* that includes coastal populations in Oregon and Washington (Meyers et al., 2015). Because a unique mitochondrial haplotype was observed at Henry Coe State Park and at Larabee Valley, California (Potter et al., 2013), we included these two populations in OTU K, despite being slightly inland and farther north, respectively, from the Santa Cruz Mountains.

OTU L (*P. engelmannii*) has been considered by most botanists to be a distinct species since its description in 1854, although hybrid offspring from natural crosses with *P. ponderosa* and with *P. arizonica* have been documented (Peloquin, 1984). *Pinus engelmannii* has a much wider distribution in Mexico, but our sampling was limited to one population in the United States where it is sympatric with *P. ponderosa* and *P. arizonica*.

The *stormiae* taxon (OTU M, *P. arizonica* var. *stormiae*) was published as a variety of *P. arizonica* (Martínez, 1945). Lingering dispute over its taxonomy is due in part to confusion that Martínez (1948) introduced with a subsequent extension of the *P. arizonica* name to some three-needled pines in Mexico. An earlier suggestion to reassign *P. arizonica* as a variety of *P. ponderosa* (Shaw, 1914) apparently affected a suggestion that the *stormiae* taxon was also a variety of *P. ponderosa* (Silba, 1990). Although Silba examined only one specimen from Nuevo León, his statement that the taxon was "possibly" also in the Chisos Mountains of Texas led to the inclusion of *P. arizonica* var. *stormiae* in the USDA Plants Database with a distribution in Texas (http://plants.usda.gov/core/profile?symbol=PIARS2) despite it not being accepted as present north of Mexico in the *Flora of North America north of Mexico* (Kral, 1993). We included two collections from Big Bend National Park (which encompass the Chisos Mountains) in our study, but it is important to note that there is no reason to believe these isolated stands in the USA adequately represent the wide distribution of the *stormiae* taxon in Mexico.

OTU N (*P. arizonica*), treated as *P. ponderosa* var. *arizonica* (Engelm.) Shaw in *Flora of North America north of Mexico* (Kral,

1993), is included in this study based on its limited distribution in the United States. Our three populations are not adequate samples of its full geographic distribution, which is much wider in Mexico. Although hybrid offspring with *P. engelmannii* and with *P. ponderosa* have been reported (Peloquin, 1971, 1984), *P. arizonica* is genetically distinct in a contact zone with *P. ponderosa* (Epperson et al., 2009) and is resolved closer to *P. cooperi* and to *P. durangensis* Martínez than to *P. ponderosa* (Gernandt et al., 2009).

OTU O and OTU P represent the more distantly related *P. jeffreyi*. Although *P. jeffreyi* is vegetatively similar to *P. ponderosa* (Baldwin et al., 2012), it is resolved with subsection *Sabinianae* (California big-cone pines) (Willyard et al., 2009) and produces heptane, a distinctive secondary compound (Mirov, 1961). *Pinus jeffreyi* occupies higher altitude or drier sites (Kral, 1993) or grows in harsher serpentine soils than *P. ponderosa* (Baldwin et al., 2012). There is some evidence that the populations of Jeffrey pine growing on serpentine soils in the Klamath Range are genetically distinct from those growing in the Sierra Nevada (Furnier and Adams, 1986), but no intraspecific names have been published for the serpentine populations. The dry habitats in southern California and in Baja California were named *P. jeffreyi* var. *peninsularis* Lemmon, but these were not part of this study. We divided our *P. jeffreyi* collections into two OTUs, with four populations representing high-altitude sites in the Sierra Nevada (OTU O) and five populations growing in serpentine soils (OTU P).

On the basis of the results from testing these 16 OTUs, we reassigned two unsupported populations and repeated the analyses for 15 prior OTUs as described below. A further reduction was tested that collapsed these assignments into nine prior OTUs. Our final test analyzed only the 11 OTUs within *P. ponderosa*. We refer to these independent analyses as the $k = 16$, $k = 15$, $k = 9$, and $k = 11$ tests.

DNA isolation and cpSSR genotyping—We chose six loci that represent different SSR regions of the plastome (Wofford et al., 2013). Each locus had a variable-length fragment with a single base pair (mononucleotide) repeat. DNA was isolated from each sample, and fragment lengths for each locus were obtained using multiplexed PCR with fluorescently labeled primers for capillary electrophoresis (Applied Biosystems, Foster City, California, USA) as previously described (Wofford et al., 2013). Ten populations were genotyped at University of Arkansas on an ABI 3130xe using Genescan 500 LIZ size standard (Thermo Fisher Scientific, Waltham, Massachusetts, USA), and 78 populations were genotyped at University of Missouri on an ABI 3730xl using Genescan 600 LIZ size standard (Thermo Fisher Scientific). We regenotyped 29 of the 218 samples (13%) that were originally analyzed using 500 LIZ on the 600 LIZ system and calculated a slight adjustment for each locus that we applied to the remaining samples.

Data analyses—We analyzed 1849 samples that had no missing data for these six loci (Appendix S1, see Supplemental Data with the online version of this article). Data analyses were performed using packages available for R v. 3.2.2 (R Core Team, 2015) or using GenAlEx v. 6.501 (Peakall and Smouse, 2006, 2012). Scripts with examples of the commands used for R analyses are provided in Appendix S2 (see online Supplemental Data). We counted the number of alleles for each locus using the loc.n.all function in the program POPPR v. 2.0.2 (Kamvar et al., 2014, 2015) and the number of private alleles for each population using the PAS function in GenAlEx. Because the plastome is haploid and essentially nonrecombining, we categorized samples into unique combinations of length variants across six loci to create multilocus haplotypes (MLHs) using the mlg function in POPPR. For each population, we counted the number of MLHs (nMLH) and calculated h = Diversity as 1 – Sum p_i^2, where p_i is the frequency of the ith allele for the population using the HDP function in GenAlEx. An analysis of molecular variance was estimated with 999 permutations to partition variance within and among populations using the AMOVA function in GenAlEx. For an initial estimate of how many groups were supported by our data, we used the find.clusters function in adegenet v. 2.0.0 (Jombart, 2008; Jombart and Ahmed, 2011) to compare the Bayesian information criterion (BIC) for varying numbers of clusters. We repeated the find.clusters function with the maximum number of clusters set at 5, 10, 15, and 20, looking for an elbow in the curve of each graph.

Discriminant analysis of principal components (DAPC)—We examined the cohesiveness of our OTUs using the DAPC function (Jombart et al., 2010) in adegenet. This multivariate technique uses principal component analysis (PCA) to transform the data, then discriminant analysis (DA) to maximize between-cluster differences. This two-step process helps identify complex clusters by partitioning out within-cluster variation starting with the assignment of each individual to a prior group. In the process, the DA calculates a probabilistic assignment of each *individual* to an OTU that may or may not be the same as the prior OTU. Importantly for our purposes, each of our four DAPC analyses began with the raw data matrix for samples: the lengths of each of the six loci (Wofford et al., 2013). Using this raw allele data provided a view of the patterns that is independent of the contracted MLHs described below that use multilocus haplotypes. The first DAPC run for each test ($k = 16$, $k = 15$, $k = 9$, or $k = 11$) began with an assignment of *all individuals in a population* to a prior OTU (Table 1); the original $k = 16$ assignments are shown in Fig. 1. Then (unlike Wofford et al., 2013), we used DAPC results reassigning some *individuals* to other OTUs. We used this reassignment of *individuals* from each DAPC run and repeated the DAPC analyses using the most recently inferred prior assignment until the proportion of the correct posterior assignment of individuals (the assign.per.pop statistic in the summary.dapc function) to each cluster was above 95%. In other words, for subsequent DAPC runs, we used the reassigned OTU from the previous run for each *individual's* prior assignment for the next run. This approach should perform better for populations that contain individuals with plastid lineages from more than one OTU than a model that requires every individual in a population to belong to one cluster. For each repeated DAPC analysis, we retained the number of principal components suggested by an alpha-spline interpolation and retained all linear discriminants (Jombart et al., 2010). Scatter plots with an inertia ellipse for each OTU were used to visualize the final results of each hypothesis. We mapped the frequency of the most abundant final OTU assignment for each population using colors from Fig. 1 and combined all of the other OTU assignments as a gray slice in each population pie chart to show only the geographic distribution of the most abundant OTU assignments. In the DAPC analysis with $k = 16$, all of the individuals from Pop77 and Pop78 (*P. arizonica* var. *stormiae* from Big Bend National Park) were reassigned to other OTUs. We used the majority-rule assignments for these two populations to create a hypothesis with $k = 15$ prior OTUs (Table 1). There were three lines of

evidence suggesting that fewer clusters might better explain the data: the BIC graph allowing a maximum of 15 clusters showed an elbow at $k = 10$ (online Appendix S3); the scatter plots with $k = 16$ and $k = 15$ showed substantial overlap among some OTUs, and these OTUs were grouped into nodes with the contracted multilocus haplotypes (see below). For this DAPC hypothesis, we started with a prior assignment of all individuals in each population to $k = 9$ clusters: OTU A, OTU B, OTU C+D+E (Ponderosa merged), OTU G+F+J+Pop77 (Scopulorum merged), OTU H+Pop78, OTU K, OTU L, OTU N+I (Sky Island merged with *P. arizonica*), and OTU O+P (*P. jeffreyi*). Finally, we repeated the DAPC analysis for a $k = 11$ hypothesis using only the 73 populations within *P. ponderosa* (i.e., excluding all *P. arizonica*, *P. engelmannii*, and *P. jeffreyi* populations).

OTU statistics—We used the poppr function in POPPR to calculate the Simpson lambda (Simpson, 1949) and evenness (Pielou, 1975; Ludwig and Reynolds, 1988; Grünwald et al., 2003) for each OTU, starting with membership of individuals assigned to each OTU by the final DAPC run for each of the four hypotheses. The corrected Simpson lambda (1 minus the sum of squared genotype frequencies) accounts for differences in sample size by multiplying lambda by $N / (N - 1)$. On this scale, a corrected Simpson lambda of 0 indicates that no genotypes are different; 1 indicates that all genotypes are different. An evenness statistic of 1 indicates that all cMLHs are present in equal abundance; an evenness value close to 0 indicates that the OTU is dominated by a single cMLH.

Contracted multilocus haplotypes (cMLHs)—We grouped similar MLHs with the mlg.filter command in POPPR using a distance matrix that sums the number of length differences at each locus. A minimum spanning network (MSN) was created with the msn function in POPPR to show similarity among cMLHs at thresholds of 4 through 10 differences. We colored nodes on each MSN with the frequency of individuals assigned by DAPC to each OTU. It is important to note that frequencies within MSN nodes are the *posterior* DAPC OTU assignments of *individuals*, not the prior assignments of entire populations (Fig. 1).

RESULTS

Samples, alleles, haplotypes, and populations—We obtained a genotype for all six loci in 1849 samples (97% of 1903 samples attempted), yielding a mean of 21 and a minimum of 12 samples per population (Table 1). There were 53 alleles, with a mean of 8.8 (SD = 1.8) alleles per locus (Table 2). The subset of 1569 individuals in 73 populations of *P. ponderosa* carried 51 of the total alleles, lacking the length = 172 allele in Pt87268 and the length = 267 allele in PcI2T1, which were only observed in *P. jeffreyi* (Table 2). For each locus, we observed all of the lengths expected from 1-bp indels with two exceptions. For PcL2T1, there was a 9-bp gap between length = 267 (observed in four individuals of *P. jeffreyi*) and the next length (276) and a 7-bp gap between length = 288 (observed in one individual of Pop64; Bighorn Mtns., WY) and the next shorter length (281). Only five populations had a private allele: Pop02 (Willow Creek, CA): Pt71936, length = 156; Pop64 (Bighorn Mtns., WY): Pcl2T1, length = 281 and length = 288; Pop81 (*P. jeffreyi*; Lassics Botanical Area, CA): Pt87268, length = 172; and Pop22 (Pollock Pines, CA): Pc10, length = 212.

TABLE 2. For six cpSSR loci, the number of alleles and observed fragment lengths in all 88 populations of *Pinus ponderosa* and related taxa and in just the subset of 73 populations of *P. ponderosa*.

Locus	Number of alleles: All populations	Number of alleles: *P. ponderosa* populations	Observed fragment lengths
Pc10	9	9	204, 205, 206, 207, 208, 209, 210, 211, 212 [a]
PcG2R1	7	7	102, 103, 104, 105, 106, 107, 108
Pcl2T1	8	7	267 [b,c], 276, 277, 278, 279, 280, 281 [a,c], 288 [a]
Pt100183	7	7	123, 124, 125, 126, 127, 128, 129
Pt71936	10	10	156 [a], 157, 158, 159, 160, 161, 162, 163, 164, 165
Pt87268	12	11	172 [a,b], 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183
Total	53	51	
Mean	8.8	8.5	
SD	1.8	1.6	

[a] Five private alleles
[b] Two fragment lengths that were only observed in *P. jeffreyi*
[c] Intervening allele size not observed

The combination of six loci created 467 MLHs, with 245 (52.5%) of those haplotypes found only in a single individual. The subset of 73 *P. ponderosa* populations contained 404 (86.5%) of the MLHs. Populations contained an average of 10.5 MLHs. Pop15 (Henry Cowell Redwoods S.P., CA) and Pop76 (*P. arizonica*, Chiricahua Mtns., AZ) tied for the fewest with three MLHs each. Pop65 (Casper Mtns., WY) had the most with 21 MLHs (Table 1). The mean diversity per population was $h = 0.297$ (Table 1). Pop35 (Bisbee Mtn., WA) had the lowest diversity with $h = 0.051$ and Pop52 (Mt. Hopkins, AZ) had the highest with $h = 0.590$. The AMOVA showed 32% of variation within populations and 68% among populations. Depending on the maximum number of clusters specified in the *find.clusters* function, we found subtle elbows in graphs of BIC against number of clusters at $k = 6$, $k = 7$, $k = 10$, $k = 12$, and $k = 15$ (Appendix S3).

DAPC: Scatter plots—The first two linear discriminant functions together explained 87.4%, 87.5%, 89.5%, and 98.8% of the variation for the $k = 16$, $k = 15$, $k = 9$, and $k = 11$ tests, respectively (Fig. 2). It is important to note that these scatter plots reflect the DAPC reassignment of presumably introgressed individuals to their best-matching cluster and, thus, plot the distances between plastid lineages even if the lineage was found in a morphologically distinctive species (e.g., *P. jeffreyi*). The scatter plots for $k = 16$ and $k = 15$ were largely similar (Fig. 2A, 2B). As expected by its separate taxonomic recognition, OTU L (*P. engelmannii*) was well defined on the first two axes. However, the *P. jeffreyi* OTUs (O and P), were close to each other but not as distant from other clusters as expected from a classification as subsection *Sabinianae*. OTU A (Pacific Northwest), OTU B (Klamath), and OTU K (Benthamiana) were each well separated from the other clusters. The other *P. ponderosa* var. *ponderosa* OTUs (C, D, and E) were clustered close to each other. OTU F (Canyonlands), OTU G (Scopulorum), and OTU J (Springs Mtns.) also were close to each other. OTU H (Brachyptera) and OTU I (Sky Island) were adjacent to the morphologically distinct OTU N (*P. arizonica*).

The scatter plot for $k = 9$ retained the separation for OTU L (*P. engelmannii*) and placed OTU C+D+E (Ponderosa merged) and

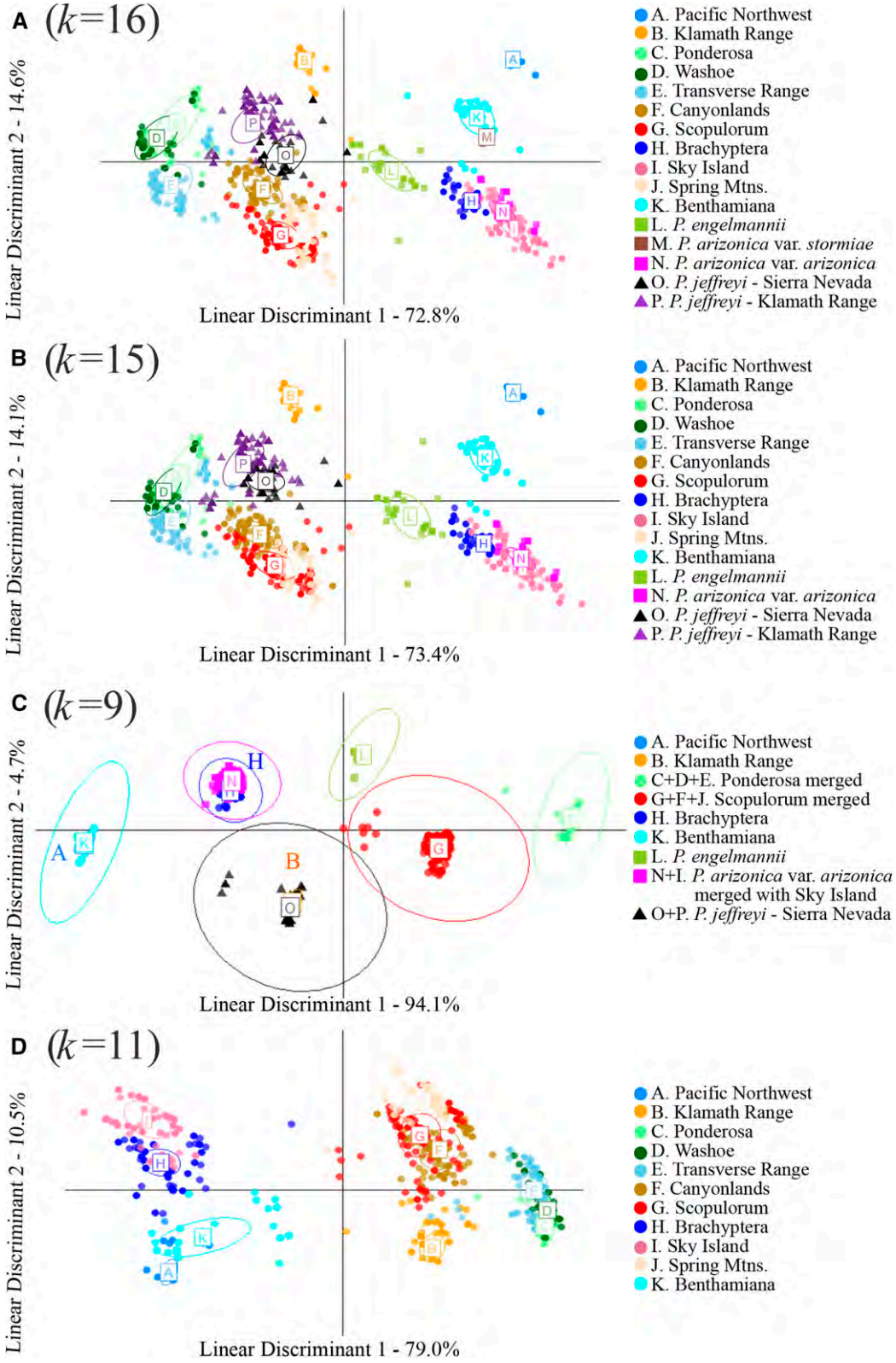

FIGURE 2 Scatter plots of the discriminant analysis of principal components of the first two linear discriminants using the final reassignment of individuals to each OTU (see methods and Tables 1, 3). Dots are individuals; ovals are inertia ellipses. (A) k = 16 prior clusters; (B) k = 15 prior clusters; (C) k = 9 prior clusters; (D) k = 11 prior clusters (*Pinus ponderosa* only; *P. engelmannii*, *P. arizonica*, and *P. jeffreyi* samples were excluded from the k = 11 test).

OTU G+F+J (Scopulorum merged) as separate clusters (Fig. 2C). There were also three sets of overlapping OTU clusters: the combined OTU N+I (Sky Island merged with *P. arizonica*) surrounded OTU H (Brachyptera); OTU O (*P. jeffreyi*) surrounded OTU B (Klamath); and OTU K (Benthamiana) surrounded OTU A (Pacific Northwest). The six obvious clusters of plastid lineages shown in this scatter plot belie the recognizable taxonomic diversity—*P. ponderosa* from the Klamath clustered within *P. jeffreyi*; *P. ponderosa* var. *brachyptera* clustered within morphologically distinct *P. arizonica*; and *P. arizonica* var. *stormiae* sampled from two isolated populations in geographic close proximity in Big Bend National Park were assigned to different OTUs (one to OTU G [Scopulorum] and one to OTU H [Brachyptera]). As the number of clusters was reduced from $k = 16$ to $k = 9$, a group of OTU G (Scopulorum) points near the center of the graph was placed even farther from the center of this cluster (Fig. 2). As described below under cMLHs, these OTU G (Scopulorum) outliers show up as satellite nodes attached to the main western nodes on the MSNs (online Appendix S4). These OTU G+F+J (Scopulorum merged) outliers in the $k = 9$ hypothesis appeared in four disjunct geographic areas: on the west coast, including two *P. jeffreyi* populations; in southeastern Nevada/southwestern Utah/northwestern Arizona; in southern New Mexico/western Texas; and Pop52 on Mt. Hopkins, AZ (Fig. 3). This geographic perspective suggests that the clustering of OTU G+F+J (Scopulorum merged) in the $k = 9$ test is likely an artifact of homoplasy in fast-evolving cpSSRs, with four unrelated lineages grouping together. We also note that the DAPC scatter plot for $k = 16$ that placed OTU M (*P. arizonica* var. *stormiae*) close to OTU K (Benthamiana) was misleading, as most individuals assigned to this cluster were NOT from the two prior populations (Pop77 and Pop78; Fig. 2A). The DAPC scatter plots for $k = 16$, $k = 15$, and the *P. ponderosa*-only $k = 11$ (Fig. 2A, 2B, 2D) could be divided into a western group (OTUs A, B, C, D, E, K plus *P. jeffreyi* OTUs O and P) vs. an eastern group (OTUs F, G, H, I, J plus *P. engelmannii* OTU L and *P. arizonica* OTU N). This western–eastern pattern was not evident on the $k = 9$ scatter plot (Fig. 2C).

With our hypothesis of $k = 16$, the starting number of individuals with prior assignments to an OTU varied from 22 in OTU L (*P. engelmannii*) to 358 in OTU C (Ponderosa), with a mean of 115.6 (SD = 90.1; Table 3). When we collapsed OTUs to test $k = 9$,

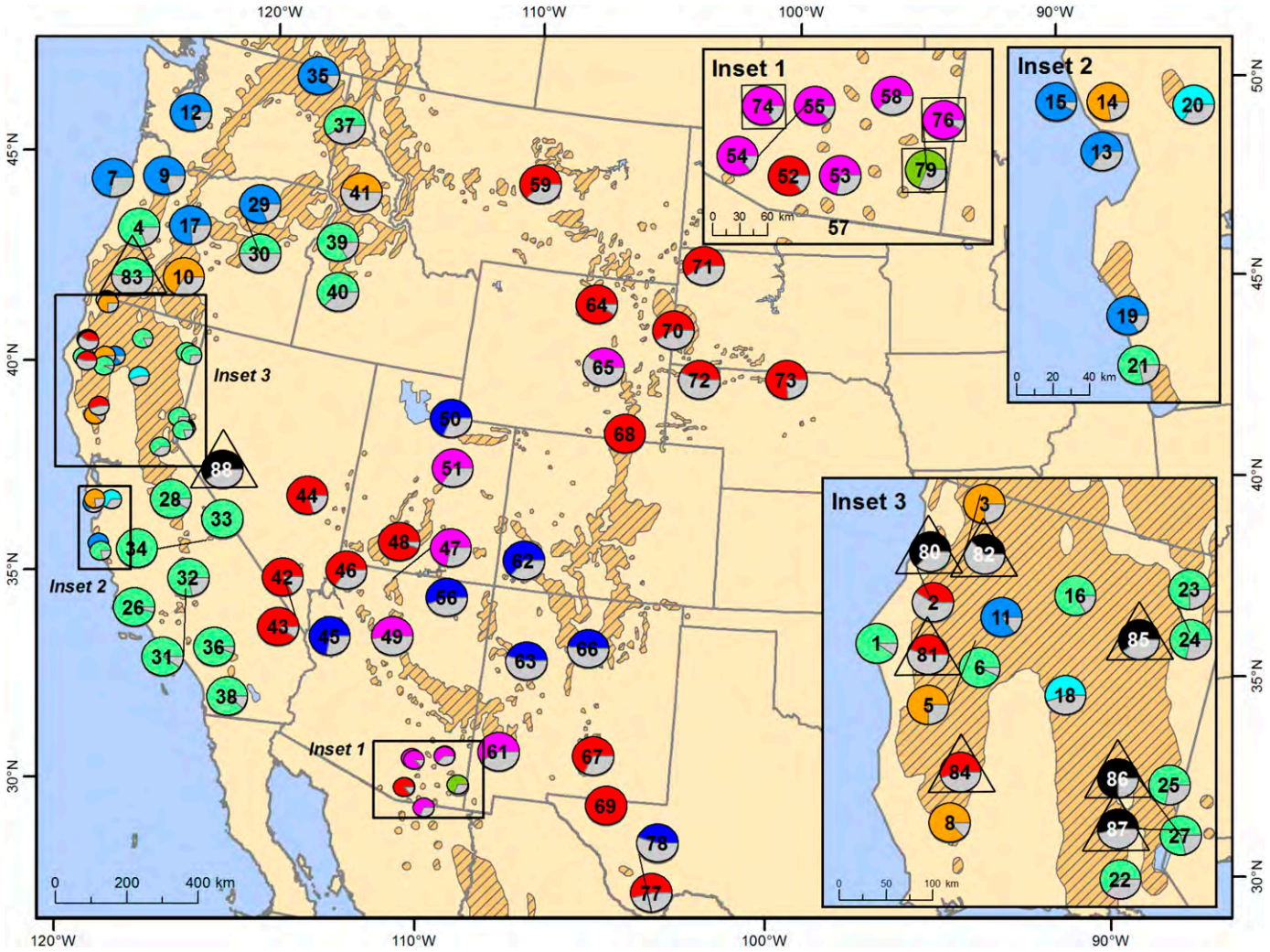

FIGURE 3 The proportion of individuals assigned to the most frequent OTU by discriminant analysis of principal components for each population using the $k = 9$ prior assignments (see legend for Fig. 2C and Table 1). Proportions assigned to all other OTUs are shown in gray.

TABLE 3. OTU statistics for four hypotheses (k = 16, k = 15, k = 9, and k = 11): the number of individuals in prior assignment; the number of individuals in posterior DAPC assignment, number of multilocus haplotypes (nMLH), Simpson lambda, and evenness index.

OTU	OTU name	Number of samples		nMLH	Corrected Simpson lambda	Evenness
		Prior	Posterior			
k = 16						
A	Pacific Northwest	75	205	5	0.292	0.494
B	Klamath	124	110	13	0.593	0.470
C	Ponderosa	358	103	21	0.873	0.658
D	Washoe	56	107	20	0.800	0.510
E	Transverse	106	220	39	0.940	0.663
F	Canyonlands	48	103	61	0.982	0.761
G	Scopulorum	176	159	58	0.942	0.490
H	Brachyptera	222	164	24	0.805	0.528
I	Sky Island	209	223	58	0.955	0.630
J	Spring Mtns.	37	70	36	0.973	0.816
K	Benthamiana	158	76	24	0.937	0.755
L	*P. engelmannii*	22	36	20	0.948	0.765
M	*P. arizonica* var. *stormiae*	35	5	3	0.800	0.950
N	*P. arizonica*	57	70	14	0.696	0.569
O	*P. jeffreyi*, Sierra Nevada	64	61	24	0.904	0.593
P	*P. jeffreyi*, Klamath Range	102	137	47	0.918	0.469
	Total	1849	1849	467	0.984	0.338
	Mean	115.6	115.6	29.2	0.835	0.633
	SD	90.1	64.7	18.6	0.2	0.1
k = 15						
A	Pacific Northwest	75	205	5	0.292	0.494
B	Klamath	124	128	16	0.685	0.496
C	Ponderosa	358	105	23	0.878	0.641
D	Washoe	56	107	20	0.800	0.510
E	Transverse	106	219	38	0.940	0.667
F	Canyonlands	48	107	64	0.984	0.763
G	Scopulorum (incl. Pop77)	189	160	59	0.943	0.488
H	Brachyptera (incl. Pop78)	244	167	25	0.812	0.520
I	Sky Island	209	214	57	0.952	0.618
J	Spring Mtns.	37	69	36	0.973	0.810
K	Benthamiana	158	80	26	0.943	0.751
L	*P. engelmannii*	22	32	17	0.935	0.761
N	*P. arizonica*	57	76	14	0.731	0.612
O	*P. jeffreyi*, Sierra Nevada	64	46	23	0.945	0.744
P	*P. jeffreyi*, Klamath Range	102	134	44	0.914	0.478
	Total	1849	1849	467	0.984	0.338
	Mean	123.3	123.3	31.1	0.848	0.624
	SD	93.1	59.6	17.9	0.2	0.1
k = 9						
A	Pacific Northwest	75	205	5	0.292	0.290
B	Klamath	124	131	16	0.699	0.693
C+D+E	Ponderosa merged	520	431	81	0.965	0.963
G+F+J	Scopulorum merged	274	380	179	0.985	0.983
H	Brachyptera merged	244	196	41	0.863	0.858
K	Benthamiana	158	58	16	0.906	0.890
L	*P. engelmannii*	22	31	16	0.931	0.901
N+I	Sky Island + *P. arizonica*	266	284	66	0.954	0.951
O+P	*P. jeffreyi*	166	133	47	0.927	0.920
	Total	1849	1849	467	0.984	0.984
	Mean	205.4	205.4	51.9	0.836	0.828
	SD	145.7	137.2	54.1	0.2	0.2
k = 11						
A	Pacific Northwest	75	243	14	0.475	0.401
B	Klamath	124	160	35	0.817	0.398
C	Ponderosa	358	88	19	0.844	0.623
D	Washoe	56	100	21	0.776	0.480
E	Transverse	106	209	41	0.941	0.642
F	Canyonlands	48	107	64	0.984	0.763
G	Scopulorum (incl. Pop77)	176	160	64	0.944	0.461
H	Brachyptera (incl. Pop78)	222	224	47	0.891	0.473
I	Sky Island	209	167	43	0.937	0.615
J	Spring Mtns.	37	68	35	0.972	0.813
K	Benthamiana	158	43	21	0.947	0.810
	Total	1569	1569	404	0.980	0.317
	Mean	142.6	142.6	36.7	0.866	0.589
	SD	96.1	66.2	17.3	0.1	0.2

the mean prior number per OTU was 205.4 (SD =145.7), with a maximum of 520 in OTU C+D+E (Ponderosa merged; Table 3). After the final DAPC run for k = 16, 39 of 88 populations (44.3%) had a majority of individuals reassigned to a different OTU than the prior (Table 1). For k = 15, k = 9, and k = 11 there were 43.2%, 29.5%, and 50.7% of populations, respectively, where a majority of individuals were reassigned to a different OTU than the prior. Despite this high frequency of population reassignments, there were only five populations with inconsistent majority assignments in different prior-clustering scenarios. A putatively hybrid population Pop55 (Whitetail Campground, AZ) had a majority assigned to either OTU I (Sky Island), OTU N (*P. arizonica*), or OTU H (Brachyptera); Pop65 (Casper Mtns., WY) had a majority reassigned to either OTU I (Sky Island) or OTU H (Brachyptera). In the k = 9 hypothesis, Pop02 (Willow Creek, CA), Pop81 (*P. jeffreyi*; Lassics, CA), and Pop84 (*P. jeffreyi*; Snow Mtn., CA) were unexpectedly reassigned to the combined OTU F+G+I (Scopulorum merged). Other populations had a majority of individuals assigned to the same OTU regardless of scenario (Table 1).

There were three consistent population reassignments that can be categorized as consolidating OTU K (Benthamiana) with OTU A (Pacific Northwest): Pop13 (UC Santa Cruz Arboretum, CA), Pop15 (Henry Cowell Redwoods S.P., CA), and Pop19 (Big Creek, CA). Despite OTU K (Benthamiana) receiving high support after the final run in each test, only Pop18 (Paynes Creek, CA) and Pop20 (Henry Coe State Park, CA) had a majority of individuals assigned to OTU K. This assignment for Pop18 was unexpected because its location in the Sierra Nevada guided our prior assignment to OTU C (Ponderosa). Eight population reassignments in the k = 9 hypothesis could be categorized as consolidating OTU C (Ponderosa) with OTU D (Washoe) and OTU E (Transverse Range): Pop04 (Eugene, OR), Pop06 (Basin Gulch Campground, CA), Pop21 (Santa Lucia, CA), Pop23 (Likely, CA), Pop30 (Blue Mtns., OR), Pop31 (Breckenridge Mtn., CA), Pop32 (Lake Isabella, CA), and Pop37 (Cour D'Alene, ID).

There were seven populations where the consistent reassignment of a majority of individuals to an OTU would greatly increase the expected geographic range of the OTU. Pop17 (Santiam Pass, OR), Pop29 (Blue Mountains, OR), and Pop35 (Bisbee Mtn., WA) were assigned to OTU A (Pacific Northwest), expanding the geographic range far inland from our expectations based on mitochondrial haplotypes (Fig. 1 vs. Fig. 3). Pop11 (Shasta County, CA) could be interpreted as either a disjunct inland population of OTU A (Pacific Northwest) or a northward extension along the Sacramento Valley of California coastal populations (Pop13, Pop15, and Pop19). In contrast, the assignment of coastal California Pop01 (Larabee Valley, CA) to either OTU D (Washoe) or to OTU C+D+E (Ponderosa merged) is unexpected geographically and could reflect either a disjunct geographic range or homoplasy. The reassignment of a majority of individuals in Pop10 (Tiller, OR) and Pop14 (Quail Hollow, CA) to OTU B (Klamath Range) could be an expansion of the geographic range for this plastid lineage.

Five population reassignments (6%), if they were truly reflecting shared ancestry rather than homoplasy, would cause large disjunctions in an OTU: Pop41 (Kooskia Rd., ID) to OTU B (Klamath Range); Pop47 (Jacob Lake, AZ) and Pop51 (Price Cyn. Rec. Area, UT) to OTU I (Sky Island); Pop67 (Mescalero Apache Res., NM) to OTU F (Canyonlands); and Pop69 (Guadalupe Mtns., TX) to OTU G (Scopulorum).

OTU assignments by DAPC—At each of the clustering levels (k = 16, k = 15, k = 9, and k = 11), most populations had individuals that DAPC assigned to multiple OTUs (Table 1). Although the mean percentage assigned to one OTU was roughly 70% regardless of prior clustering, some populations had only about 33% of their individuals assigned to a majority OTU. Because of this intrapopulation variation, we looked for broad geographic patterns by mapping only a pie chart piece for each population that was color-coded for the OTU to which the most individuals were assigned (Fig. 3; Appendices S3, S4). The geographic pattern in posterior OTUs was clearest in the collapsed k = 9 test (Fig. 3), and that pattern had some important differences from our prior hypothesis (Fig. 1).

Based on mitochondrial haplotypes (Potter et al., 2013), we expected plastid OTU A (Pacific Northwest) to be limited to coastal Washington and Oregon. Instead, we found plastid OTU A extending into eastern Washington, central and eastern Oregon, down the California coast, and in Shasta County, California (Pop11; Fig. 3). Also contrary to our hypothesis from mitochondrial haplotypes, OTU B (Klamath Range) was not the predominant plastid lineage in the Klamath Mountains, but this OTU *was* found northward in central Oregon and in a disjunct population in the Santa Cruz Mountains. The broad geographic range of OTU C+D+E (Ponderosa merged) in the k = 9 test yielded a plastid lineage extending from southern California through the Sierra Nevada, across the Blue Mountains in Oregon, and representing a major component of populations in Idaho and in two populations in the Klamath Range (Fig. 3). However, there are other populations interspersed within this geographic range that harbor a majority of individuals assigned to OTU A (Pacific Northwest) and to OTU B (Klamath). In the k = 16, k = 15, and k = 11 scenarios, plastid OTU D (Washoe) was retained in two prior populations, Pop24 (Likely, CA) and Pop25 (Babbitt Peak, CA; online Appendices S5, S6). However, this OTU lost a majority of individuals at the type locality of *P. washoensis* (Pop27, Mt. Rose, NV) and unexpectedly gained a majority in two populations in Oregon (Pop04, Eugene, OR and Pop30, Blue Mountains, OR) and in one coastal California population (Pop01, Larabee Valley, CA). The geographic range of plastid OTU E (Transverse) was greatly extended from our hypothesis into the Sierra Nevada (including the type locality of *P. washoensis*, Mt. Rose, NV) and in a disjunct population in the Klamath Range.

In the k = 16, k = 15, and *P. ponderosa* only k = 11 hypotheses, plastid OTU F (Canyonlands) was only assigned to four very distantly scattered populations in Nevada, Montana, Wyoming, and New Mexico. Similarly, OTU J (Spring Mtns.) was assigned to two very distant populations—one in South Dakota and one in southern Arizona. When OTU F (Canyonlands) and OTU J (Spring Mtns.) were collapsed into OTU G+F+J (Scopulorum merged) in the k = 9 hypothesis, the geographic pattern became simpler. Nonetheless, the geographic ranges of plastid lineages for OTU G (Scopulorum) and OTU H (Brachyptera) in Nebraska, South Dakota, Wyoming, and Utah differed from expectations based on mitochondrial patterns. Whether we tested the plastid OTU I (Sky Island) by itself (Appendices S5, S6) or combined with OTU N (*P. arizonica*) in the k = 9 test because of overlapping scatter plot placement (Fig. 2A, 2B, 2D), this cluster was assigned to two populations on the Mogollon Rim that are close to the Sky Islands of southern Arizona (Fig. 3). This cluster was also assigned to three distant populations in Utah and Wyoming (Fig. 3). Based on mitochondrial haplotypes, we expected plastid OTU J (Spring Mtns.) to be limited to the isolated mountains in southeastern Nevada. One of those two prior populations (Pop42) was assigned to OTU G (Scopulorum), but plastid OTU J was unexpectedly assigned to a very disjunct population (Pop52; Mt. Hopkins, AZ). Despite our sampling representing only one population of *P. engelmannii* (Pop79), plastid OTU L was cohesive across different clustering tests (Table 1; Fig. 3; Appendices S3, S4).

Although there was some apparent support for the distinctness of plastid lineages of *P. jeffreyi* in the Sierra Nevada (OTU O) from this species in the Klamath Range (OTU P; Fig. 2A), this is misleading. Two populations—Pop85 (Likely, CA) and Pop86 (Mt. Rose, NV) from the Sierra Nevada—were part of OTU P (Klamath Range) and Pop82 (Tiller, OR), which we had identified morphologically as *P. jeffreyi*, was assigned to OTU C (Ponderosa).

Contracted multilocus haplotypes (cMLHs)—Simplifying at thresholds (t) of 4, 5, 6, 7, 8, 9, and 10 differences yielded 35, 21, 14, 11, 9, 8, and 5 cMLHs, respectively. None of the thresholds yielded contracted nodes that corresponded to our OTU hypotheses. We show an MSN with four or five nodes (t = 10) color-coded by the frequency of DAPC-assigned individuals to OTUs for k = 16, k = 15, k = 9, and k = 11 in Appendix S4. For each of these MSNs, there was one major node representing most of the OTUs from the western part of the distribution (including *P. jeffreyi*) and another major node representing most of the OTUs from the eastern part of the distribution (including *P. engelmannii* and *P. arizonica*).

OTU statistics—The variability statistics by OTU are based on the DAPC posterior assignment of *individuals* to OTUs (see methods), which allowed individuals within populations to be assigned to different plastid OTUs (Table 3). As expected from a reduction

in the number of clusters, the mean number of MLHs per OTU increased from 29.2 in the k = 16 to 51.9 in the k = 9 hypothesis. The minimum number of MLHs observed in an OTU was five in OTU A (Pacific Northwest). The most variable was OTU F (Canyonlands) using k = 16 or k = 15 priors. OTU F (Canyonlands) and OTU G (Scopulorum) were tied as the most variable OTUs in the k = 11 scenario. In the k = 9 test, OTU G+F+I (Scopulorum merged) was the most variable, with 179 MLHs observed in 274 individuals (Table 3). Corrected Simpson lambda values were generally high, suggesting genotypes varied substantially within each cluster. OTU A (Pacific Northwest) had the lowest value, indicating fewer genotypes were different among individuals assigned to this OTU (Table 3). None of the evenness values were low, indicating that none of the OTUs were dominated by a single cMLH (Table 3).

Correspondence of organelle lineages—We collapsed our plastid lineages into six related groups: OTU A+K, OTU C+D+E, OTU F, OTU G, OTU H+I, and OTU M. A map showing the generalized distribution of these six collapsed plastid lineages along with a generalized distribution of mitochondrial haplotypes (Potter et al., 2013) revealed large areas of correspondence and other areas with a genetic mosaic (Fig. 4).

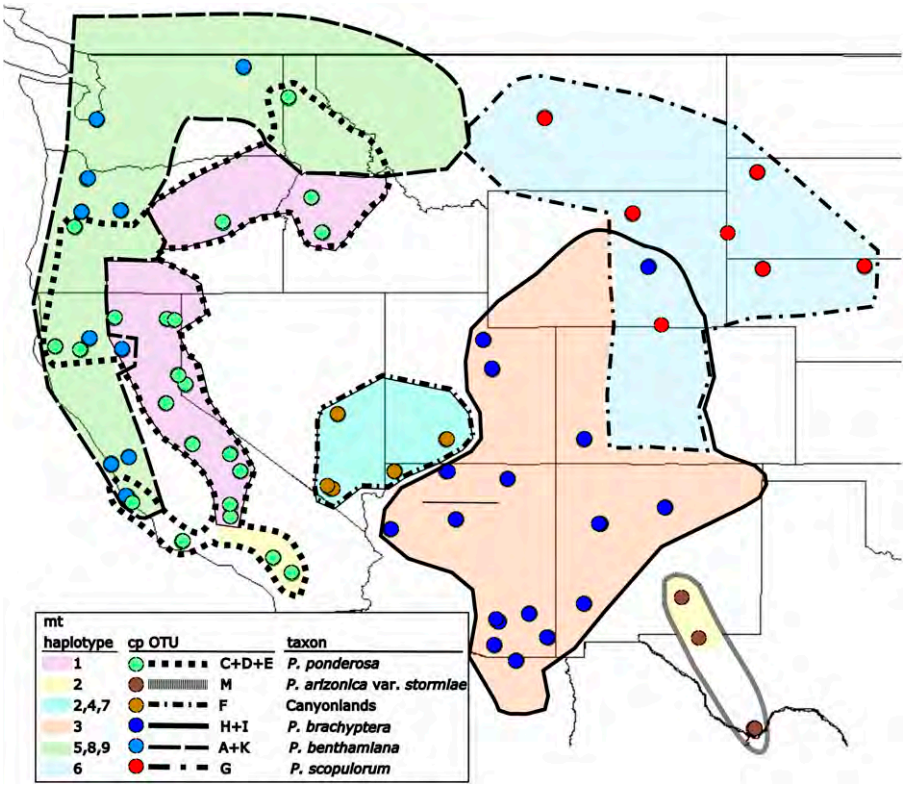

FIGURE 4 The generalized geographic distribution of mitochondrial haplotypes (Potter et al., 2013) and six plastid (cp) OTUs that roughly correspond to them: a reduced concept of *Pinus ponderosa* represented by mitochondrial haplotype 1 and cp OTU C+D+E; *P. arizonica* var. *stormiae* represented by mitochondrial haplotype 2 and cp OTU M; an unpublished taxon represented by mitochondrial haplotypes 2, 4, and 7 and cp OTU F (Canyonlands); *P. brachyptera* represented by mitochondrial haplotype 3 and cp OTU H+I; *P. benthamiana* represented by mitochondrial haplotypes 5, 8, and 9 and cp OTU A+K; *P. scopulorum* represented by mitochondrial haplotype 6 and cp OTU G.

DISCUSSION

Main findings—The phylogeography for six plastid OTUs corresponded in a general way with that of their mitochondria (Potter et al., 2013), and there is some support for distinctive climate niches for these mitochondrial haplotypes (Shinneman et al., 2016). Five of these organellar OTUs correspond to previously published species (Fig. 4). Although an organelle lineage might fail to track species ancestry due to introgression (Willyard et al., 2009; Ran et al., 2015), the concordance of independent genetic patterns provides a strong predictive model for inferring relatedness of species. By including samples from related taxa, we were able to show that five OTUs within *P. ponderosa* s.l. carry plastid lineages more distantly related to each other than they are to other species of subsection *Ponderosae* and that the plastid lineage for a sixth OTU (*P. arizonica* var. *stormiae*) carries a plastid lineage more distantly related to *P. arizonica* than to other species. Together, these patterns support resurrecting three species that have been lumped into *P. ponderosa*: *P. benthamiana* is supported by OTU A+K (Pacific Northwest and Benthamiana) and mitochondrial haplotypes 5, 8, and 9; a reduced concept of *P. ponderosa* by OTU C+D+E (Ponderosa merged) and mitochondrial haplotype 1; *P. brachyptera* by OTU H+I (Brachyptera and Sky Island) and mitochondrial haplotype 3; and *P. scopulorum* by the main northeastern group assigned to OTU G+F+J (Scopulorum merged) and mitochondrial haplotype 6 (Potter et al., 2013) (Fig. 4). The original publications of four separate species better describe evolutionary history than a broadly defined *P. ponderosa* with four intraspecific varieties. Organellar patterns also support the distinctiveness of *P. arizonica* var. *stormiae* from *P. arizonica*: disjunct assignments to OTU G and mitochondrial haplotype 2 (Potter et al., 2013) (Fig. 4). We discuss some potentially confounding issues in our data, then present evidence (or lack thereof) for each of our prior OTUs.

Homoplasy—Because our study design included more than one species, convergent evolution may have generated enough homoplasy in these highly variable cpSSR loci to be a confounding factor. We looked for evidence that homoplasy among more distantly related *P. arizonica*, *P. engelmannii*, *P. jeffreyi*, and *P. ponderosa* distorted our results. We took two approaches (DAPC and cMLH MSNs) to analyzing these multilocus data and observed clusters of each of these species. At the least, their separate clustering suggests that homoplasy is not overwhelming this data set. We also observed very similar results for the k = 11 test when *P. ponderosa* OTUs were analyzed without *P. arizonica*, *P. engelmannii*, and *P. jeffreyi* (Appendix S6). What is more, we observed a low level of presumably introgressed individuals to and from other species where it was expected in areas of sympatry (online Appendix S7). Again, this result suggests

that homoplasy is not rampant and that the relative distances inferred between clusters in DAPC and the arrangement of cMLH nodes in the MSNs were biologically meaningful. However, homoplasy is the most likely explanation for the clustering of four highly disjunct regions (three populations in the Klamath Range, five populations near the Nevada–Arizona–Utah border, three populations in southeastern New Mexico and Texas, and one population in the Sky Islands) with the main northeastern range of OTU G+F+J (Scopulorum merged) in the $k = 9$ hypothesis (Fig. 3). Because of the potential for homoplasy, we avoided attributing minor variation within populations as evidence for migrants from other OTUs.

OTUs with small sample sizes—If there were too few representatives of a highly diverged lineage in the DAPC, aberrant results could occur, with the too-different group "pulling" toward an unrelated heterogeneous cluster (Jombart et al., 2010). We were aware that because of the limited geographic distribution of *P. arizonica*, *P. arizonica* var. *stormiae*, and *P. engelmannii* in the United States where we collected, our small sample sizes of these species might be problematic (Table 1). We found that *P. engelmannii*, with only one population, consistently formed a cluster and that the few reassigned individuals to and from sympatric OTUs are likely due to introgression. In contrast, two populations of OTU M (*P. arizonica* var. *stormiae*), geographically very close to each other in Big Bend National Park, TX, did not form a recognizable cluster in DAPC. Implications of the assignments for *P. arizonica* var. *stormiae* are discussed below. It may be that our weak plastid differentiation between morphologically distinct *P. arizonica*, OTU H (Brachyptera), and OTU I (Sky Island) was partly due to the limited sample size of the former. This factor may also have contributed to our failure to support OTU F (Canyonlands) and OTU J (Spring Mtns.). In the collapsed $k = 9$ test, five populations retained a prior assignment of OTU G+F+J (Scopulorum merged), but were very widely separated from the main range of this cluster (Fig. 3). Support for distinctive genotypes in three of the regions assigned to OTU G+F+J (Scopulorum merged) is given below. Our study also had fairly strong support for a cryptic taxon that might have been unrecognized because there were too few samples for DAPC to form a cluster. Because the pines at Mt. Hopkins, AZ (Pop52) seemed distinctive from other Sky Island *P. ponderosa* populations that we had collected, we were not surprised that their plastid lineage was not assigned to OTU I (Sky Island; Table 1; Fig. 3, inset 1).

Introgression—The disparate placement of some *P. ponderosa* samples on phylogenetic trees (Gernandt et al., 2009; Willyard et al., 2009; Parks et al., 2012) might have been explained by exemplars in those studies that represented the infrequent individuals with genes introgressed from sympatric *P. jeffreyi* (Haller, 1961, 1962). Our population-level sampling allowed us to investigate the extent of this introgression empirically. Some of our samples provided evidence for plastid introgression, although as discussed above, we cannot distinguish introgression from homoplasy or incomplete lineage sorting in most cases. In several areas where there are morphological distinctions that have been used to recognize other species, there are *individuals* whose plastid lineage was assigned by DAPC to a sympatric species. For example, *P. jeffreyi* in both the Klamath Range (Pop83 Tiller, OR; Fig. 3) and in the Sierra Nevada (Pop87 Thomas Creek, NV and Pop88 Mammoth Lakes, CA) were assigned to nearby *P. ponderosa* populations (Appendix S7). The *P. engelmannii* population had 7 of 15 individuals (32%) that carried plastid lineages of the morphologically distinct *P. arizonica* (Appendix S7).

Four taxa rather than one P. ponderosa s.l.—The support for resurrecting *P. benthamiana*, *P. brachyptera*, and *P. scopulorum* as distinct from *P. ponderosa* is based on the genetic distinctness of mitochondrial and plastid lineages and the important finding reported here that these four taxa are relatively more genetically distant to each other than they are to *P. jeffreyi*, to *P. arizonica*, or to *P. engelmannii* (Fig. 2). In the far western part of the range, there are at least two unrelated taxa—OTU A+K (Pacific Northwest plus Benthamiana) that includes the type locality of *P. benthamiana* and OTU C+D+E (Ponderosa merged) that includes one of the two possible type localities for *P. ponderosa*. Plastid OTU A (Pacific Northwest) and OTU K (Benthamiana) consistently resolved near each other in DAPC scatter plots (Fig. 2). Importantly, in every run that includes other subsection *Ponderosae* taxa (i.e., $k = 16$, $k = 15$, $k = 9$), OTU A and OTU K are more distant to OTU C, to OTU D, and to OTU E than they are to *P. arizonica*, to *P. engelmannii*, and to *P. jeffreyi*. This finding strongly indicates that plastid OTU A+K (Pacific Northwest combined with Benthamiana) is not conspecific with OTU C+D+E (Ponderosa merged) (Fig. 2) and that geographic regions where they live in proximity represent secondary contact between nonsister taxa. Our analyses purposefully kept OTU K (Benthamiana) separate to test its correspondence to geographic regions that have distinctive mitochondrial haplotype 9 (Potter et al., 2013) or the ecologically recognizable sand hill populations interspersed in redwood forests (Griffin, 1964). Neither pattern was evident in plastid data. Only one unexpected population (Pop18; Paynes Creek, CA) in the $k = 15$ hypothesis had a majority of individuals assigned to OTU K (Benthamiana; Appendix S5), and this pattern was lost in the $k = 9$ test (Fig. 3). Similarly, the Pacific Northwest (OTU A) had support in these data but not for a plastid lineage unique to Fort Lewis, Washington or limited to the Willamette Valley, Oregon, or for those two regions combined (Fig. 3). Nonetheless, the geographic distribution of assignments to a combination of these two OTUs does roughly correspond to the geographic range of related mitochondrial haplotypes 5, 8, and 9 (Potter et al., 2013): the California coast (including the sand hill region which is the type locality for *P. benthamiana* near Santa Cruz, CA), some populations in the Klamath Range and in the Sierra Nevada northeast of the Sacramento Valley, Oregon's Willamette Valley, the isolated coastal population at Fort Lewis, and parts of central and eastern Oregon and Washington (Fig. 3). Plastid OTU C+D+E (Ponderosa merged) occurs in the Sierra Nevada, the Transverse Range, some populations in the Klamath Range, the Blue Mountains of southern Oregon, and in southern Idaho (Fig. 3). A distant relationship between these two taxa was also found in a plastid nucleotide sequence phylogeny where an exemplar of *P. ponderosa* collected near Chico, California, resolved sister to *Sabinianae* and a collection from western Montana was sister to other *Ponderosae* (Parks et al., 2012). The geographic distribution of OTU C+D+E (Ponderosa merged) roughly corresponds to that of mitochondrial haplotype 1 (Potter et al., 2013) (Fig. 4).

Based on the sampling in our study and the mitochondrial study, the organelle phylogeography appears to be a mosaic in some areas. For example, the Klamath Range in southwestern Oregon has plastid OTU C+D+E (Ponderosa merged) but mitochondrial haplotype 5. We also observed plastid OTU C+D+E (Ponderosa merged) in

the Cascade Mountains of southern Oregon near where mitochondrial haplotypes 5 and 8 were observed. The Transverse Range in southern California also has plastid OTU C+D+E (Ponderosa merged) but the distantly related mitochondrial haplotype 2 (Potter et al., 2013). There are areas that display a mosaic of plastid OTUs in proximity. For example, the Klamath Range in California and the northwest flank of the Sierra Nevada (Pop18) have populations assigned to OTU A+K (Pacific Northwest combined with Benthamiana), to OTU C+D+E (Ponderosa merged), and to a possibly divergent OTU B (Klamath Range) (Fig. 3, Inset 3). A mosaic is apparent inland as well. Pop35 (Bisbee Mtn., WA) was assigned to plastid OTU A (Pacific Northwest), but Pop37 (Cour D'Alene, ID) was assigned to OTU C+D+E (Ponderosa merged), despite these two sites being only about 150 km apart (Fig. 3; Appendices S5, S6). Interestingly, these two collection sites were selected to be near the two potential *P. ponderosa* type localities—Douglas' 1826 seed collection area near Kettle Falls, Washington, and his *Arceuthobium*-bearing sterile branch near Spokane, Washington, respectively (Lauria, 1996a). Thus, the apparently continuous ponderosa pine forests in eastern Washington and northern Idaho that were the basis for naming *P. ponderosa* may be a previously unrecognized contact zone. The plastid mosaic is also apparent on the California coast, where Pop21 (Santa Lucia, CA) was assigned to OTU C+D+E (Ponderosa merged) despite being only about 25 km from Pop19 (Big Creek, CA) that was assigned to OTU A+K (Pacific Northwest combined with Benthamiana; Fig. 3, inset 2). The existence of a "Pacific" species or variety of ponderosa pine has been suggested numerous times. Deep-green leaves and yellowish-brown bark have been used to compare this taxon with the typical variety's grayish green leaves and reddish bark. However, the geographic range revealed by plastid and mitochondria strongly conflicts with the geographic ranges inferred from other data (Weidman, 1939; Wells, 1964b; Callaham, 2013b) and only partially coincides with taxonomic treatments because none of them include the inland region of central Oregon, central Washington, and southern Idaho in a "Pacific" taxon (Kral, 1993; Haller and Vivrette, 2011; Baldwin et al., 2012; Callaham, 2013a; Meyers et al., 2015). Our suggested treatment of *P. benthamiana* has very different boundaries than *P. ponderosa* var. *pacifica* and *P. ponderosa* subsp. *critchfieldana*. Based on our results, the concept of *P. ponderosa* is reduced to the Sierra Nevada, some southern California coastal populations, and the Blue Mountains of Oregon. The contact zones for *P. benthamiana* and *P. ponderosa* in the Klamath Range and across central Oregon and Idaho (Fig. 4) correspond to "hot spot clusters" of hybrid zones identified for multiple species of trees, birds, and mammals (Swenson and Howard, 2005). When *P. benthamiana* and *P. ponderosa* were tested as a single species, glacial refugia were inferred along the California coast, the Klamath Mountains, the Sierra Nevada, and southern California mountains (Roberts and Hamann, 2015). In light of our genetic patterns (Fig. 4), it would be interesting to test whether *P. benthamiana* dominated the first two refugia and *P. ponderosa* occupied the latter two.

The eastern part of the *P. ponderosa* geographic range also has at least two taxa. It is now clear from an accumulation of evidence that they are not conspecific with *P. ponderosa* from the western part of the range. Rather, they are more closely related to each other than they are to the western ponderosa pines. The plastid MSNs consistently show separate major nodes for western (OTUs A+B+C+D+E+ *P. jeffreyi*) vs. eastern (OTUs F+G+H+I+J+ *P. arizonica* + *P. engelmannii*) groups (Appendix S4). The eastern OTUs collapse into one node with plastid OTU H+I (Brachyptera plus Sky Island) and another with plastid OTU G+F+J (Scopulorum merged; Appendix S4). Each of our DAPC plots show these two major eastern clusters separated from each other and from the two separate western groups described above (Fig. 2). Their respective geographic regions also carry two different mitochondrial haplotypes. The plastid lineage that encompasses OTU H+I (Brachyptera plus Sky Island) likely has shared ancestry with *P. arizonica* (OTU N; Fig. 2), and these populations occupy part of the wide geographic range of mitochondrial haplotype 3 (Potter et al., 2013). OTU H+I (Brachyptera plus Sky Island) and mitochondrial haplotype 3 are both present near the type locality of *P. brachyptera* (near Pop66, Santa Fe, NM; Fig. 3). These results confirm the placement of a few exemplars on plastid nucleotide sequence phylogenies: a collection from South Dakota near our Pop70 (Parks et al., 2012) and two collections from the Sky Islands of Arizona (Gernandt et al., 2009) resolved sister to clades containing *P. arizonica*. The main region of populations assigned to OTU G+F+J (Scopulorum merged) in Montana, Wyoming, South Dakota, Nebraska, and Colorado (Fig. 3) corresponds fairly closely to that of mitochondrial haplotype 6 (Potter et al., 2013). This distribution is a reasonable fit for the syntypes suggested for *P. ponderosa* var. *scopulorum* that include the northern Rocky Mountains (Lauria, 1996a), but *P. scopulorum* as supported by organellar lineages would be confined to a smaller and more northerly range than the broader definition used in treatments that did not recognize *P. brachyptera*. Morphological distinctions published for the *scopulorum* taxon and for *P. brachyptera* are subtle. The *scopulorum* taxon was described as having two to three needles per fascicle (Lemmon, 1897; Kral, 1993), and *P. brachyptera* was described with three needles (rarely two to four) per fascicle and slightly larger cones (Wislizenus, 1848), but it is difficult to assess which geographic range of individuals were used to support each range of morphological variation. The contact zone suggested by our plastid phylogeography is in Colorado, coinciding with a small part of a broad swath of "hot spot clusters" for tree, bird, and mammal hybrid zones (Swenson and Howard, 2005). Postglacial expansion into this region may have converged species from refugia in the southern Arizona mountains, the southwestern tablelands, and the Sierra Madre (Swenson and Howard, 2005; Roberts and Hamann, 2015).

Other OTUs within P. ponderosa s.l.—The existence of a distinct lineage of ponderosa pines in southern Nevada and nearby parts of Utah and far northeastern Arizona remains a possibility worth exploring. Despite the failure of OTU F (Canyonlands) and OTU J (Spring Mtns.) to gain support as separate clusters in DAPC, there was a strong correspondence between the geographic distribution of Pop42, Pop43, Pop44, and Pop46 in the collapsed $k = 9$ hypothesis (Fig. 3) and the combined distribution of mitochondrial haplotypes 2, 4, and 7 (Potter et al., 2013) (Fig. 4). As noted above, having too few samples of a heterogeneous taxon might explain their failure to form a cohesive cluster in DAPC. These populations have not been published as a distinct taxon, but it is possible that they belong to a cryptic species for which morphological characters are yet to be identified.

Although there might appear to be some support for OTU D (Washoe) in the $k = 16$ and $k = 15$ hypotheses, the population from the type locality (Pop27) at Mt. Rose was not assigned to it (Table 1; Appendix S5), and the scatter plot showed heavy overlap with OTU C (Ponderosa) and OTU E (Transverse Range) (Fig. 2). The inclusion

of Pop01 (Larabee Valley, CA) with Washoe does not make sense geographically and is likely due to homoplasy in a population that also had mixed mitochondrial haplotypes (Potter et al., 2013). Thus, our data did not lend any support for a separate plastid lineage for *P. ponderosa* var. *washoensis*.

There was no support for a separate plastid lineage in the Transverse Range of southern California (Fig. 3; Appendices S5, S6). This pattern suggests that there is an organelle mosaic in the ponderosa pines of southern California, with some populations carrying the plastid lineage of nearby Sierra Nevada pines and mitochondrial haplotype 2, that was only reported in southern Nevada and in southern New Mexico (Potter et al., 2013). The Transverse Range has been suggested to be a suture zone for hybrid interactions among many species (Remington, 1968), and our findings may add another example.

We considered whether the Sky Island pines (OTU I) were divergent enough to be treated as a separate variety of *P. brachyptera* and concluded that this is another open question worthy of investigation. DAPC assignment of individuals (except the putative hybrid Pop55, Whitetail Campground, AZ) reliably placed a majority of individuals from the Sky Island populations into OTU I (Table 1; Appendices S5, S6). In a previous study, cpSSR patterns suggested that the three-needled pines from Mt. Lemmon, AZ were distinct from *P. arizonica* and a plastid haplotype and two low-copy nuclear gene trees placed samples in a clade that did not include *P. ponderosa* (Epperson et al., 2009). However, the only two Sky Island populations where mitochondrial haplotypes have been reported share haplotype 3 with Brachyptera (Potter et al., 2013). Inertia ellipses for OTU I (Sky Island) overlapped those of OTU N (*P. arizonica*) rather than OTU H (Brachyptera) (Fig. 2). This pattern suggests that despite being morphologically distinguishable from *P. arizonica*, the Sky Island pines have a plastid lineage that is more closely related to *P. arizonica* than to OTU H (Brachyptera; Fig. 2). When we merged OTU N+I (*P. arizonica* plus Sky Island), populations in central Arizona, the Rocky Mountains of Utah, and the Casper Mountains in Wyoming formed a mosaic pattern among those assigned to OTU H (Brachyptera) (Fig. 3). A mosaic of mitochondrial haplotypes 3 and 6 was also observed in northern Colorado and northern Wyoming. Thus, the phylogeography of the ponderosa pines in this entire region may well represent a broad zone of secondary contact and differential admixture of organelle lineages. If so, this group of taxa may share a more recent common ancestor than the disparate plastid and mitochondrial lineages of *P. benthamiana* and *P. ponderosa*.

After DAPC reassignments, OTU B (Klamath Range) clustered nearer to *P. jeffreyi* than to OTU A (Pacific Northwest) or to OTU K (Benthamiana), and when we simplified to $k = 9$, the DAPC scatter plots placed *P. jeffreyi* overlapping *P. ponderosa* OTU B (Klamath Range) (Fig. 2). Unexpectedly, populations assigned to this cluster had a distribution that starts in the Klamath Range of California and bends across southern Oregon to reach central Idaho (Fig. 3). The easternmost populations cannot be due to recent introgression with *P. jeffreyi* because the latter species is absent from this area, but introgression followed by dispersal is a possible explanation. The geographic range for OTU B (Klamath) did not correspond to any mitochondrial haplotype patterns (Potter et al., 2013), and populations that carry this lineage are intermixed with populations that carry other plastid lineages. What is more, three populations in the Klamath Range (including two *P. jeffreyi*) had a highly disjunct majority assignment to a combined OTU G+F+J (Scopulorum; Fig. 3). Together, these patterns suggest that the DAPC support for OTU B as a separate cluster may reflect heterogeneity where *P. ponderosa* and *P. benthamiana* are sympatric. Alternatively, these genotypes may represent remnants of an ancestral lineage related in some way to the ancestors of *P. jeffreyi* and the other *Sabinianae*. Although the four species of California big-coned pines in *Sabinianae* were monophyletic in plastid genealogies (Gernandt et al., 2009; Parks et al., 2012), we note that to our knowledge a ponderosa pine representing OTU B has yet to be included in a published phylogeny.

Stormiae pine—Two isolated and nearby populations from Big Bend National Park (Pop77 and Pop78) have been assigned to *P. arizonica* var. *stormiae*, which has a much wider distribution in Mexico. Other populations in southern New Mexico and western Texas have been suspected to belong to this taxon as well. Three of our populations—Pop67 (Mescalero Apache Res, NM), Pop69 (Guadalupe Mtns., TX), and Pop77 (Big Bend National Park, TX)—form one of the disjunct clusters in the plastid OTU G+F+J (Scopulorum merged) scenario (Figs. 3, 4). Our data clearly show that these populations do not have plastid lineages that belong to OTU N (*P. arizonica*) (Fig. 2). Of these four populations, only a majority of the individuals in the highly heterogeneous Pop78 are assigned to the geographically proximal OTU H (Brachyptera) (Fig. 3; Appendix S7). The failure of these four populations to be recognized as a plastid cluster by DAPC may be due to the small sample size of heterogeneous individuals, or the plastid haplotype frequencies may be exhibiting some admixture from OTU H (Brachyptera). It has been suggested that this taxon belongs as a variety of *P. ponderosa* rather than *P. arizonica* (Silba, 1990), but this is not supported by mitochondrial haplotypes. Trees from southern New Mexico carry mitochondrial haplotype 2, which is distantly related to haplotype 3 that is found in Brachyptera to the west and to the north (Potter et al., 2013). The genetic relations between *P. arizonica* var. *stormiae* where it is widely distributed in Mexico and these United States stands remain to be explored, but it is clear from their relative placement in DAPC scatter plots that the ponderosa pines from Big Bend National Park are not a variety of *P. arizonica* (Callaham, 2013a).

Jeffrey pine—We considered whether there is genetic structuring within this species that corresponds to ecological niches. Although *P. jeffreyi* populations on serpentine soils in the Klamath Range are somewhat diverged from the Sierra Nevada high altitude populations, the pattern is not strong. Our $k = 16$ and $k = 15$ DAPC hypotheses show only a very weak subdivision between the *P. jeffreyi* of the Klamath Range (where they grow mostly on serpentine soils) and the *P. jeffreyi* of the Sierra Nevada (where they grow mostly at higher altitudes). The DAPC scatterplot clustering is not as strong a support for subdivision as it might appear because some individuals assigned to OTU P (*P. jeffreyi* in the Klamath Range) were collected in the Sierra Nevada (e.g., Pop85 and Pop86; Fig. 3).

Further study—Our results suggest several fruitful areas that would warrant further study. The relationship between *P. arizonica* and *P. arizonica* var. *stormiae* deserves a fresh consideration across the entire geographic range. The origin of the ponderosa pines on Mt. Hopkins will require a comparison with Mexican taxa that includes morphological data as well as information from mitochondrial and nuclear genomes. Are there further subdivisions that our data were not powerful enough to observe? For example, lack of clear support

in these plastid data (e.g., for OTU B (Klamath), OTU I (Sky Island), OTU J (Spring Mountain), or a more narrowly defined OTU K (Benthamiana) that corresponds to mitochondrial haplotype 9), is not proof against genetic structuring. How are *P. jeffreyi* and the other California big-coned pines (subsection *Sabinianae*) related to *P. benthamiana*? Finally, the extensive zones of sympatry among ponderosa pine taxa—both where plastid OTUs are intermingled and where plastid and mitochondrial ancestry conflict—were not previously recognized. Major areas of overlap are in California (the Transverse Range and where the Klamath Range meets the Cascade Range), in Oregon (where the northern Siskiyou region meets the Willamette Valley), in Washington and Idaho (the northern Rockies), and in Wyoming (the western edge of the Rocky Mountains) (Fig. 4). These suture zones will offer exciting opportunities to clarify the nature of the genetic mosaic in the ponderosa pines. Do intermingled taxa account for some of the morphological and growth variability described within *P. ponderosa*? Do some individual trees carry organelle or nuclear lineages with different ancestry? If so, does either of these organelle lineages correspond to past, present, or projected future climate factors? This study provides hypotheses to further confirm or refute, and it leads to many more questions regarding the mysterious past of the ponderosa pines.

CONCLUSIONS

The phylogeographic patterns and a biologically useful taxonomic classification of the *Ponderosae* are complex problems, yet most research on the ecologically and economically valued *P. ponderosa* in the United States have assumed it to be one species or at least one species complex. These researchers were often constrained to consider this putative species in isolation from other *Ponderosae* in North America, which obscured important genetic differences among populations. We found a rough agreement (as well as intriguing regions where populations appear to be a genetic mosaic in contact zones) between our plastid results and recently published mitochondrial and nuclear microsatellite patterns. Importantly, the plastid results presented here were able to show that the relative genetic distance among some subdivisions of *P. ponderosa* is large compared with the distance to other *Ponderosae*. We suspect that robust nuclear evidence using exemplars from all of the *Ponderosae* will solidify the intuitive understanding that the United States–Mexico border was an unfortunate choice for species delimitation. We think that a species tree built using a coalescent model from the gene trees of a large number of low copy nuclear loci and plastome sequences will be needed to clarify evolutionary relationships among these taxa, and efforts to do that are underway. Although a species tree will provide a critical framework, it will not allow us to elucidate the genetic mosaic that we described here because within-population variation will not be measured using our current experimental plan for high-throughput sequencing that relies on choosing a limited number of exemplar samples. Nor will our current plan help determine which morphological characters can be used to support the genetic divisions that are evident. Those latter two goals will require a comprehensive sampling with morphological and molecular data measured for the same individuals. It is clear that the characters previously used to diagnose these taxa are inadequate, but many others could be investigated. For example, ovulate cone scale, leaf morphology and anatomy, leaf cuticle micromorphology, and seedling characters have been described (Harlow, 1947; Mirov, 1967; Stead, 1983; Whang et al., 2004; López-Reyes et al., 2015). Gathering these data widely will be challenging due to phenological constraints coupled with interannual differences and the need to sample many individuals because of variability among individuals. At present, some geographic boundaries remain fuzzy because their plastid and mitochondrial haplotypes have not been sampled. Nevertheless, we suggest that a classification based on four published species (Fig. 4) would reflect the genetic history better than current classifications of four varieties within the *P. ponderosa* species complex.

ACKNOWLEDGEMENTS

We thank the Hendrix College Odyssey Professorship for funding. Adam Bigott, Richard C. Cronn, Bryan Epperson, Jack Finney, J. Robert Haller, Dalton Hoose, Brandon Linz, Aaron Liston, Bob Means, Stephen Meyers, Glenn Lunak, Hoai Trang Nguyen, Matthew Parks, Thomas Patterson, Dakota Pouncey, Brian Schumacher, Nicole Segear, Joshua Smith, Kevin Spatz, and Justen Whittall contributed plant material, laboratory work, or helpful suggestions. We are grateful to the following for permission to collect plant material: Arboretum at the University of California–Santa Cruz, Big Bend National Park, Bighorn National Forest, Black Hills National Forest, Boise National Forest, Bureau of Land Management, California State Parks, Cibola National Forest, Colville National Forest, Coronado National Forest, Custer National Forest, Eldorado National Forest, Fred Lawrence Whipple Observatory, Gila National Forest, Guadalupe National Park, Hualapai Mountain Park (Mohave County, AZ), Humboldt-Toiyabe National Forest, Idaho Panhandle National Forest, Inyo National Forest, James San Jacinto Mountains Reserve (University of California), Joint Base Lewis-McChord, Kaibab National Forest, Klamath National Forest, Landel-Hills Big Creek Reserve (University of California), Los Padres National Forest, Medicine Bow-Routt National Forest, Mendocino National Forest, Mescalero Apache Tribe, Modoc National Forest, Nez Perce-Clearwater National Forest, North Sierra Tree Improvement Association, Niobrara Valley Preserve (The Nature Conservancy), Price Canyon Recreation Area, Quail Hollow Park (Santa Cruz County, CA), San Bernardino National Forest, San Juan National Forest, Santa Fe National Forest, Sequoia National Forest, Shasta-Trinity National Forest, Sierra National Forest, Sierra Pacific Industries, Six Rivers National Forest, Tahoe National Forest, The Navajo Nation, Uinta-Wasatch Cache National Forest, Umpqua National Forest, Willamette National Forest. Comments on a previous version of the manuscript from Daniel Potter and two anonymous reviewers were very helpful.

LITERATURE CITED

Baldwin, B. G., D. H. Goldman, D. J. Keil, R. Patterson, T. J. Rosatti, and D. H. Wilken. 2012. The Jepson manual, 2nd ed., University of California Press, Berkeley, California, USA.

Bouffier, L. A., B. L. Gartner, and J.-C. Domec. 2003. Wood density and hydraulic properties of ponderosa pine from the Willamette Valley vs. the Cascade mountains. *Wood and Fiber Science* 35: 217–233.

Callaham, R. Z. 2013a. *Pinus ponderosa*: A taxonomic review with five subspecies in the United States. USDA Forest Service, Pacific Southwest Research Station PSW-RP-264. Washington, D.C., USA.

Callaham, R. Z. 2013b. *Pinus ponderosa:* Geographic races and subspecies based on morphological variation. USDA Forest Service, Pacific Southwest Research Station PSW-RP-265. Washington, D.C., USA.

Conkle, M., and W. B. Critchfield. 1988. Genetic variation and hybridization of ponderosa pine. *In* D. Baumgartner and J. Lotan [eds.], Ponderosa pine: The species and its management. Washington State University, Pullman, Washington, USA.

Critchfield, W. B. 1984. Crossability and relationships of Washoe pine. *Madroño* 31: 144–170.

Epperson, B., F. W. Telewski, and A. Willyard. 2009. Chloroplast diversity in a putative hybrid swarm of *Ponderosae. American Journal of Botany* 96: 707–712.

Furnier, G. R., and W. T. Adams. 1986. Geographic patterns of allozyme variation in Jeffrey pine. *American Journal of Botany* 73: 1009–1015.

Gernandt, D. S., S. Hernández-León, E. Salgado-Hernández, and J. Pérez de la Rosa. 2009. Phylogenetic relationships of *Pinus* subsection *Ponderosae* inferred from rapidly evolving cpDNA regions. *Systematic Botany* 34: 481–491.

Gerson, E. A., and R. G. Kelsey. 2004. Piperidine alkaloids in North American *Pinus* taxa: Implications for chemosystematics. *Biochemical Systematics and Ecology* 32: 63–74.

Gooding, G. D. 1998. Genetic variation and mating system of ponderosa pine in the Willamette Valley of Oregon. M.S. thesis, Oregon State University, Corvallis, Oregon, USA.

Griffin, J. R. 1964. Isolated *Pinus ponderosa* forests on sandy soils near Santa Cruz, California. *Ecology* 45: 410–412.

Grünwald, N. J., S. B. Goodwin, M. G. Milgroom, and W. E. Fry. 2003. Analysis of genotypic diversity data for populations of microorganisms. *Phytopathology* 93: 738–746.

Haller, J. R. 1961. Some recent observations on ponderosa, Jeffrey, and Washoe pines in Northeastern California. *Madrono* 16: 126–132.

Haller, J. R. 1962. Variation and hybridization in ponderosa and Jeffrey pines. *University of California Publications in Botany* 34: 123–165.

Haller, J. R. 1965a. *Pinus washoensis* in Oregon: Taxonomic and evolutionary implications. *American Journal of Botany* 52: 646.

Haller, J. R. 1965b. The role of 2-needle fascicles in the adaptation and evolution of ponderosa pine. *Brittonia* 17: 354–382.

Haller, J. R., and N. J. Vivrette. 2011. Ponderosa pine revisited. *Aliso* 29: 53–57.

Harlow, W. 1947. The identification of the pines of the United States, native and introduced, by needle structure. New York Technical Publication 32, New York State College of Forestry, Syracuse University, Syracuse, New York, USA.

Jombart, T. 2008. adegenet: An R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.

Jombart, T., and I. Ahmed. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.

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: 94–108.

Kamvar, Z., J. Brooks, and N. Grünwald. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Frontiers in Genetics* 6: 208.

Kamvar, Z., J. Tabima, and N. Grünwald. 2014. Poppr: An R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2: e281.

Kral, R. 1993. *Pinus. In* Flora of North America Editorial Committee [ed.], Flora of North America north of Mexico, vol. 2, 373–398. Oxford University Press, New York, New York, USA.

Latta, R. G., and J. B. Mitton. 1999. Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution* 53: 769–774.

Lauria, F. 1991. Taxonomy, systematics, and phylogeny of *Pinus*, subsection *Ponderosae* Loudon (Pinaceae). *Linzer biologische Beitrag* 23: 129–202.

Lauria, F. 1996a. The identity of *Pinus ponderosa* Douglas ex. C. Lawson (Pinaceae). *Linzer biologische Beitrag* 28(2): 999–1052.

Lauria, F. 1996b. Typification of *Pinus benthamiana* Hartw. (Pinaceae), a taxon deserving renewed botanical examination. *Annalen des Naturhistorischen Museums in Wien* 98B: 427–446.

Lauria, F. 1997. The taxonomic status of *Pinus washoensis* H. Mason & Stockw. (Pinaceae). *Annalen des Naturhistorischen Museums in Wien* 99B: 655–671.

Lemmon, J. G. 1897. Three West-American conifers. *Garden and Forest* 10: 183–184.

Liston, A., M. Parker-Defeniks, J. V. Syring, A. Willyard, and R. Cronn. 2007. Interspecific phylogenetic analysis enhances intraspecific phylogeographic inference: A case study in *Pinus lambertiana*. *Molecular Ecology* 16: 3926–3937.

López-Reyes, A., J. P. de la Rosa, E. Ortiz, and D. S. Gernandt. 2015. Morphological, molecular, and ecological divergence in *Pinus douglasiana* and *P. maximinoi*. *Systematic Botany* 40: 658–670.

Ludwig, J., and J. Reynolds. 1988. Statistical ecology: A primer on methods and computing. John Wiley, New York, New York, USA.

Major, J., and S. A. Bamberg. 1967. Some Cordilleran plants disjunct in the Sierra Nevada of California, and their bearing on Pleistocene ecological conditions. *In* H. E. Wright and W. H. Osburn [eds.], Arctic and alpine environments. Indiana University, Bloomington, Indiana, USA.

Mallet, J. 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution* 10: 294–299.

Manos, P. S., J. J. Doyle, and K. C. Nixon. 1999. Phylogeny, biogeography, and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution* 12: 333–349.

Martínez, M. 1945. Las pinaceas mexicanas, vol. I. *Anales del Instituto de Biológia de la Universidad Nacional de México* 16: 284–286.

Martínez, M. 1948. Los pinos mexicanos, 2nd ed, Ediciones Botas, Mexico City, Mexico.

Meyers, S. C., T. Jaster, K. E. Mitchell, and L. K. Hardison. 2015. Flora of Oregon. Botanical Research Institute of Texas, Fort Worth, Texas, USA.

Millar, C. I., and W. J. Libby. 1991. Strategies for conserving clinal, ecotypic, and disjunct population diversity in widespread species. *In* D. A. I. Falk and K. E. Holsinger [eds.], Genetics and conservation of rare plants. Oxford University Press, New York, New York, USA.

Mirov, N. T. 1961. Composition of gum turpentines of pines. USDA Forest Service Technical Bulletin 1239, Washington, D.C., USA.

Mirov, N. T. 1967. The genus *Pinus*, Ronald Press, New York, New York, USA.

Niebling, C. R., and M. T. Conkle. 1990. Diversity of Washoe pine and comparisons with allozymes of ponderosa pine races. *Canadian Journal of Forest Research* 20: 298–308.

Parks, M., R. Cronn, and A. Liston. 2012. Separating the wheat from the chaff: Mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L.(Pinaceae). *BMC Evolutionary Biology* 12: 100.

Patten, A., and S. Brunsfeld. 2002. Evidence of a novel lineage within the *Ponderosae. Madroño* 49: 189–192.

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

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

Peloquin, R. L. 1971. Variation and hybridization patterns in *Pinus ponderosa* and *Pinus engelmannii*. Ph.D. dissertation, University of California Santa Barbara, Santa Barbara, California, USA.

Peloquin, R. L. 1984. The identification of three-species hybrids in the ponderosa pine complex. *Southwestern Naturalist* 29: 115–122.

Pielou, E. 1975. Ecological diversity. John Wiley, New York, New York, USA.

Potter, K. M., V. D. Hipkins, M. F. Mahalovich, and R. E. Means. 2013. Mitochondrial DNA haplotype distribution patterns in *Pinus ponderosa* (Pinaceae): Range-wide evolutionary history and implications for conservation. *American Journal of Botany* 100: 1562–1579.

Potter, K. M., V. D. Hipkins, M. F. Mahalovich, and R. E. Means. 2015. Nuclear genetic variation across the range of ponderosa pine (*Pinus ponderosa*): Phylogeographic, taxonomic, and conservation implications. *Tree Genetics & Genomes* 11: 38–60.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria USA. Website https://www.R-project.org/.

Ran, J.-H., T.-T. Shen, W.-J. Liu, P.-P. Wang, and X.-Q. Wang. 2015. Mitochondrial introgression and complex biogeographic history of the genus *Picea*. *Molecular Phylogenetics and Evolution* 93: 63–76.

Read, R. 1980. Genetic variation in seedling progeny of ponderosa pine provenances, Forest Science, Monograph 23, Society of American Foresters, Washington, D.C., USA.

Rehfeldt, G. E. 1999a. Systematics and genetic structure of *Ponderosae* taxa (Pinaceae) inhabiting the mountain islands of the Southwest. *American Journal of Botany* 86: 741–752.

Rehfeldt, G. E. 1999b. Systematics and genetic structure of Washoe pine: Applications in conservation genetics. *Silvae Genetica* 48: 167–173.

Remington, C. L. 1968. Suture-zones of hybrid interaction between recently joined biotas. *In* T. Dobzhansky, M. K. Hecht, and W. D. Steer [eds.], Evolutionary biology, 321–428. Springer US, New York, New York, USA.

Roberts, D. R., and A. Hamann. 2015. Glaical refugia and modern genetic diversity of 22 western North American tree species. *Proceedings of the Royal Society, B, Biological Sciences* 282: 20142903–20142912.

Shaw, G. R. 1914. The genus *Pinus*. Publications of the Arnold Arboretum no. 5, Cambridge, Massachusetts, USA.

Shinneman, D. J., R. E. Means, K. M. Potter, and V. D. Hipkins. 2016. Exploring climate niches of ponderosa pine (*Pinus ponderosa* Douglas ex Lawson) haplotypes in the western United States: Implications for evolutionary history and conservation. *PLOS ONE* 11: e051811.

Silba, J. 1990. A supplement to the international census of the Coniferae, II. *Phytologia* 68: 7–78.

Simpson, E. 1949. Measurement of diversity. *Nature* 163: 688.

Smith, R. H. 1964. Variation in the monoterpenes of *Pinus ponderosa* Laws. *Science* 143: 1337–1338.

Smith, R. H. 1967. Variations in the monoterpene composition of the wood resin of Jeffrey, Washoe, Coulter and lodgepole pines. *Forest Science* 13: 1967.

Smith, R. H. 1977. Monoterpenes of ponderosa pine xylem resin in Western United States. USDA Forest Service Technical Bulletin no. 1532, Washington, D.C., USA.

Smith, R. H. 1981. Variation in immature cone color of ponderosa pine (Pinaceae) in northern California and southern Oregon. *Madroño* 28: 272–275.

Smith, R. H., R. L. Peloquin, and P. C. Passof. 1969. Local and regional variation in the monoterpenes of ponderosa pine wood oleoresin. USDA Forest Service Research Paper PSW-56, Washington, D.C., USA.

Sorensen, F. C. 1994. Genetic variation and seed transfer guidelines for ponderosa pine in Central Oregon. USDA Forest Service Research Paper PNW-RP-472, Washington, D.C., USA.

Stead, J. 1983. Studies in Central American pines V: A numerical study of variation in the Pseudostrobus group. *Silvae Genetica* 32: 101–115.

Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *American Naturalist* 166: 581–591.

Weidman, R. H. 1939. Evidences of racial influence in a 25-year test of ponderosa pine. *Journal of Agricultural Research* 59: 855–887.

Wells, O. O. 1964a. Geographic variation in ponderosa pine I. The ecotypes and their distribution. *Silvae Genetica* 13: 89–103.

Wells, O. O. 1964b. Geographic variation in ponderosa pine II. Correlations between progeny performance and characteristics of the native habitat. *Silvae Genetica* 13: 126–132.

Whang, S., K. Kyungsik, and R. Hill. 2004. Cuticle micromorphology of leaves of *Pinus* (Pinaceae) from North America. *Botanical Journal of the Linnean Society* 144: 303.

Willyard, A., R. Cronn, and A. Liston. 2009. Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution* 52: 498–511.

Wislizenus, A. 1848. Memoir of a tour to northern Mexico, connected with Col. Doniphan's expedition. Senate Miscellaneous Publication No. 26, Washington, D.C., USA.

Wofford, A. M., K. Finch, A. Bigott, and A. Willyard. 2013. A set of plastid loci for use in multiplex fragment length genotyping for intraspecific variation in *Pinus* (Pinaceae). *Applications in Plant Sciences* 2 (5) 1400002: 1–10.

Zhang, J., and B. M. Cregg. 2005. Growth and physiological responses to varied environments among populations of *Pinus ponderosa*. *Forest Ecology and Management* 219: 1–12.

APPENDIX 1 **Population vouchers.**

Taxon; Population code, *Voucher specimen*, Herbaria.

Pinus ponderosa; Pop01, *Potter s.n.*, IFGP. Pop02, *AMW1182*, IFGP. Pop03, *AMW1179*, IFGP. Pop04, *Potter s.n.*, IFGP. Pop05, *AMW1183*, IFGP. Pop06, *AMW1155*, IFGP. Pop07, *Meyers s.n.*, IFGP. Pop08, *AMW1186*, IFGP. Pop09, *Meyers s.n.*, IFGP. Pop10, *AMW1176*, IFGP. Pop11, *AMW1017*, IFGP. Pop12, *AMW1099*, IFGP. Pop13, *AMW1104*, IFGP. Pop14, *AMW1102*, MEXU. Pop15, *AMW1103*, MEXU. Pop16, *AMW1178*, IFGP. Pop17, *AMW1115*, MEXU. Pop18, *AMW1098*, MEXU, OSC. Pop19, *AMW1105*, MEXU, OSC. Pop20, *AMW1107*, IFGP. Pop21, *AMW1106*, MEXU. Pop22, *AMW1187*, IFGP. Pop23, *AMW1021*, MEXU, OSC. Pop24, *AMW1002*, MEXU. Pop25, *Potter s.n.*, IFGP. Pop26, *AMW1156*, IFGP. Pop27, *AMW999*, IFGP. Pop28, *AMW1158*, IFGP. Pop29, *AMW1015*, IFGP. Pop30, *AMW1025*, MEXU. Pop31, *AMW1164*, IFGP. Pop32, *AMW1163*, IFGP. Pop33, *AMW1159*, IFGP. Pop34, *AMW1161*, IFGP. Pop35, *AMW1111*, MEXU. Pop36, *AMW1165*, IFGP. Pop37, *AMW1108*, MEXU, OSC. Pop38, *AMW1166*, IFGP. Pop39, *AMW1174*, IFGP. Pop40, *AMW1173*, IFGP. Pop41, *AMW1175*, IFGP. Pop42, *Langer s.n.*, MEXU. Pop43, *Potter s.n.*, IFGP. Pop44, *Potter s.n.*, IFGP. Pop45, *AMW1138*, MEXU, OSC. Pop46, *Potter s.n.*, IFGP. Pop47, *AMW1139*, MEXU, OSC. Pop48, *AMW1140*, MEXU, OSC. Pop49, *AMW1137*, MEXU, OSC. Pop50, *AMW1141*, MEXU, OSC. Pop51, *DSG1029b*, IFGP. Pop52, *AMW1082*, MEXU. Pop53, *Langer s.n.*, MEXU. Pop54, *Marquardt s.n.*, IFGP. Pop55, *Marquardt s.n.*, IFGP. Pop56, *AMW1136*, MEXU, OSC. Pop57, *AMW1081*, MEXU. Pop58, *AMW1077*, MEXU. Pop59, *Potter s.n.*, IFGP. Pop60, *AMW1078*, MEXU, OSC. Pop61, *AMW1083*, MEXU, OSC. Pop62, *AMW1135*, MEXU, OSC. Pop63, *Langer s.n.*, MEXU, OSC. Pop64, *AMW1143*, MEXU, OSC. Pop65, *AMW1142*, MEXU, OSC. Pop66, *AMW1073*, MEXU, OSC. Pop67, *Potter s.n.*, IFGP. Pop68, *AMW1172*, IFGP. Pop69, *Langer s.n.*, SRSC. Pop70, *AMW1145*, MEXU, OSC. Pop71, *Langer s.n.*, MEXU. Pop72, *AMW1146*, MEXU, OSC. Pop73, *AMW1147*, MEXU, OSC.

P. arizonica; Pop74, *Marquardt s.n.*, IFGP. Pop75, *DSG874*, IFGP. Pop76, *AMW1080*, IFGP.

P. arizonica* var. *stormiae; Pop77, *AMW1047*. MEXU, OSC. Pop78, *AMW1048*, IFGP.

P. engelmannii; Pop79, *AMW1079*, IFGP.

P. jeffreyi; Pop80, *AMW1181*, IFGP. Pop81, *AMW1184*, IFGP. Pop82, *AMW1180*, IFGP. Pop83, *AMW1177*, IFGP. Pop84, *AMW1185*, IFGP. Pop85, *AMW1018*, MEXU. Pop86, *AMW998*, IFGP. Pop87, *AMW1000*, IFGP. Pop88, *AMW1162*, IFGP.