

Phylogenomics in the Hard Pines (*Pinus* subsection *Ponderosae*; Pinaceae) Confirms Paraphyly in *Pinus ponderosa*, and Places *Pinus jeffreyi* with the California Big Cone Pines

Ann Willyard,^{1,8} David S. Gernandt,² Blake Cooper,¹ Connor Douglas,¹ Kristen Finch,^{1,3} Hassan Karemera,¹ Erik Lindberg,⁴ Stephen K. Langer,⁵ Julia Lefler,¹ Paula Marquardt,⁶ Dakota L. Pouncey,¹ and Frank Telewski⁷

¹Hendrix College, Biology Department, 1600 Washington Avenue, Conway, Arkansas 72032, USA

²Universidad Nacional Autónoma de México, Departamento de Botánica, Instituto de Biología, Ciudad de México 04510, Mexico; dgernandt@ib.unam.mx

³Present address: Oregon State University, Department of Wood Science and Engineering, 3180 SW Jefferson Way, Corvallis, Oregon 97331, USA; Kristen.Finch@oregonstate.edu

⁴Texas Tech University, Department of Biological Sciences, Lubbock, TX 79409, USA; elindberg@treecarescience.com

⁵620 28th Street, Ogden, UT 84403, USA; langerbotany@yahoo.com

⁶USDA Forest Service, Northern Research Station, 5985 Highway K, Rhinelander, WI 54501, USA; marca17@msu.edu

⁷W. J. Beal Botanical Garden, Department of Plant Biology, Michigan State University, East Lansing, MI 48824, USA, telewski@msu.edu

⁸Author for correspondence (willyard@hendrix.edu)

Communicating Editor: James I. Cohen

Abstract—We sampled 130 individuals (2 to 25 per taxon) of *Pinus* subsections *Ponderosae* and *Sabinianae*. Nucleotide sequences were obtained by targeting 703 low copy nuclear genes. From the unenriched portion of the short reads, we assembled nearly complete plastome nucleotide sequences. We used 600 nuclear genes and the plastome sequences to create phylogenies and species trees that we compared to evaluate cytonuclear concordance and reticulation. We found that *Pinus jeffreyi* belongs with *Pinus* subsect. *Sabinianae* based on morphological synapomorphies as well as strong molecular phylogenetic support. *Pinus ponderosa* sensu lato is paraphyletic, and we suggest treatment as three species: *P. ponderosa* sensu stricto (with *P. ponderosa* var. *ponderosa*, *P. ponderosa* var. *benthamiana*, and *P. ponderosa* var. *washoensis*), *P. scopulorum*, and *P. brachyptera*. The persistence of lineages with the footprints of ancient nuclear introgression (labeled bpw in clade N4) and chloroplast capture (labeled bpw in clade P1) should caution species identification in *Pinus* subsection *Ponderosae* based on limited molecular data. The hybrid frequency was low based on cytonuclear discordance, and the persistence of an ancient P1 plastid clade is a better explanation than hybridization between *P. ponderosa* and *P. jeffreyi* for unexpected plastid associations in the western Sierra Nevada, USA. We identified a new potential zone of ancient admixture between *P. ponderosa* and *P. scopulorum* in Idaho, USA. Some populations of *P. arizonica*, *P. brachyptera*, *P. engelmannii*, and *P. scopulorum* in the USA are more closely related to taxa with distributions limited to Mexico than they are to each other. To integrate phylogeny and taxonomy, future work should sample widely in Mexico and the USA, score morphological characters (including seedling characters from the known seed parent), on the same individual as used for molecular data, and use methods that are based on individuals rather than population frequencies.

Keywords—Cytonuclear discordance, low copy nuclear genes, Jeffrey pine, ponderosa pine.

The ponderosa pine species of North and Central America were first noticed by European botanists nearly 200 years ago, yet the relationships among these important tree species that dominate and maintain the structure and function of vast forests remain unclear. The geographic distribution of *Pinus* subsect. *Ponderosae* Loudon (Pinaceae, *Pinus* L. subg. *Pinus* sect. *Trifoliae*; Table 1) extends from Canada into Central America (Price et al. 1998). In the USA, their distribution extends from the Pacific coast eastward to Nebraska, USA and in Mexico, there are members of *P.* subsect. *Ponderosae* that grow from the Pacific coast to the Sierra Madre Oriental (Price et al. 1998). Because of their ecological importance and the economic value of some taxa that are in the early stages of domestication for reforestation, hundreds of experiments have studied their ancestry, their growth characteristics, and their interactions with symbionts and pathogens. Unfortunately, these studies have often been designed and interpreted using a taxonomic system that is probably wrong (Farjon and Styles 1997; Willyard et al. 2009; López-Reyes et al. 2015). Many of the morphological diagnoses of the 22 accepted *P.* subsect. *Ponderosae* species do not fit well with the diversity observed in the field. Using the currently accepted taxonomy, researchers have attributed measurable differences among populations of species of *P.* subsect. *Ponderosae* to ecological adaptation, phenotypic plasticity, introgressive

hybridization, or a species complex lacking coalescence (Willyard et al. 2017). However, growth trials and molecular data for the ponderosa pines show strong and sometimes conflicting patterns of genetic differentiation within and among currently accepted species (Weidman 1939; Wells 1964b; Read 1980; Kitzmiller 1990; Rehfeldt 1993, 1999a; Cregg 1994; Callahan 2013a; Potter et al. 2013; Willyard et al. 2017). Because these growth and molecular results show geographic patterns, we think that some biologically meaningful species of *P.* subsect. *Ponderosae* may exist that differ from their current taxonomic treatments. A better understanding of the ancestry of ponderosa pine populations, some of which may be growing in unrecognized sympatry or parapatry because they descended from recent migrants, might change the interpretation of differences among populations. In this paper, we will argue that the testing and refining of taxonomic definitions is a valuable goal for building species-level molecular phylogenies.

Our understanding of the shared ancestry of life on earth has been greatly improved using molecular phylogenies, but interpreting species-level relationships closer to the tips of phylogenetic trees for plant species remains tricky. Because questions of how species are related focuses near the intersection of phylogeny and tokogeny (Hennig 1966), comparing closely related taxa often violates the assumptions of

TABLE 1. Taxa sampled (Pinaceae, *Pinus* subgenus *Pinus* section *Trifoliae*) in *Pinus* subsections *Ponderosae* and *Sabinianae*. *Names accepted in World Flora Online (<http://www.worldfloraonline.org/>) on 6 May 2020.

Taxon	Selected synonyms	Countries: states
<i>Pinus</i> subsection <i>Ponderosae</i>		
<i>Pinus arizonica</i> Engelm [*]		Mexico: Chih, Coa, Dur, NL, Sin, Son, Zac; USA: AZ, NM
<i>Pinus arizonica</i> var. <i>stormiae</i> Martínez [*]	<i>P. ponderosa</i> var. <i>stormiae</i> (Martínez) Silba	Mexico: Coa, NL, Tam, SLP; USA: NM, TX
<i>Pinus benthamiana</i> Hartweg	<i>P. ponderosa</i> var. <i>scopulorum</i> Engelm [*] ; <i>P. p.</i> var. <i>benthamiana</i> (Hartweg) Vasey; <i>P. p.</i> var. <i>pacifica</i> Haller; <i>P. p.</i> var. <i>critchfieldiana</i> Callahan	USA: CA, OR, WA
<i>Pinus brachyptera</i> Engelm [*]	<i>P. ponderosa</i> var. <i>scopulorum</i> Engelm [*] ; <i>P. p.</i> var. <i>brachyptera</i> (Engelm [*]) Lemmon	USA: AZ, CO, NV, NM, UT, WY
<i>Pinus cooperi</i> C.E.Blanco	<i>P. arizonica</i> var. <i>cooperi</i> (C.E.Blanco) Farjon; <i>P. arizonica</i> Engelm [*]	Mexico: Chih, Dur
<i>Pinus devoniana</i> Lindley [*]	<i>P. michoacana</i> Martínez	Guatemala; Mexico: Agu, Chia, DF, Gue, Hid, Jal, Méx, Mic, Mor, Nay, Oax, Pue, Que, SLP, Sin, Tla, Ver, Zac
<i>Pinus douglasiana</i> Martínez [*]		Mexico: Dur, Gue, Jal, Méx, Mic, Mor, Nay, Oax, Sin
<i>Pinus durangensis</i> Martínez [*]		Mexico: Chih, Dur, Jal, Mic, Son, Zac
<i>Pinus engelmannii</i> Carrière [*]	<i>P. macrophylla</i> Engelm. in Wislizenus	Mexico: Chih, Dur, NL, Sin, Son, Zac; USA: AZ, NM
<i>Pinus hartwegii</i> Lindley [*]		Mexico: Chia, Chih, Coa, Col, DF, Dur, Gue, Hid, Jal, Méx, Mic, Mor, NL, Oax, Pue, Tam, Tla, Ver
<i>Pinus maximinoi</i> H.E. Moore [*]	<i>P. douglasiana</i> var. <i>maximinoi</i> (H.E. Moore) Silba	El Salvador; Guatemala; Honduras; Mexico: Chia, Col, Dur, Gue, Hid, Jal, Méx, Mic, Mor, Nay, Oax, Pue, Sin, Tlax, Ver; Nicaragua
<i>Pinus montezumae</i> Lambert [*]		Guatemala; Mexico: Chia, Coa, DF, Gue, Hid, Jal, Mex, Mic, Mor, NL, Oax, Pue, Tam, Tla, Ver
<i>Pinus ponderosa</i> P.&C.Lawson [*] <i>sensu stricto</i>		USA: CA, MT, NV, OR, ID, WA
<i>Pinus pseudostrobus</i> Lindley [*]		El Salvador; Guatemala; Honduras, Mexico: Chia, Coa, DF, Dur, Gua, Gue, Hid, Jal, Mic, Méx, Mor, NL, Oax, Pue, Sin, Tla, Ver
<i>Pinus pseudostrobus</i> var. <i>apulcensis</i> (Lindley) Shaw [*]	<i>P. apulcensis</i> Lindley; <i>P. oaxacana</i> Mirov; <i>P. nubicola</i> J.P.Perry	El Salvador; Guatemala; Mexico: Chia, Gue, Hid, Méx, Oax, Pue, Tla, Ver
<i>Pinus scopulorum</i> (Engelm [*]) Lemmon	<i>P. ponderosa</i> var. <i>scopulorum</i> Engelm [*]	USA: CO, MT, NE, SD, WY
<i>Pinus washoensis</i> H.Mason & Stockwell	<i>P. ponderosa</i> var. <i>washoensis</i> Haller & Vivrette; <i>P. p.</i> P.&C.Lawson [*]	USA: CA, NV
<i>Pinus yecorensis</i> Debreczy & I.Rác	<i>P. pseudostrobus</i> Lindley [*]	Mexico: Dur, Chih, Sin, Son
<i>Pinus</i> subsection <i>Sabinianae</i>		
<i>Pinus coulteri</i> D.Don [*]		Mexico: BC; USA: CA
<i>Pinus jeffreyi</i> A.Murray bis [*]		Mexico: BC; USA: CA, NV, OR
<i>Pinus sabiniana</i> Douglas [*]		USA: CA
<i>Pinus torreyana</i> Parry ex Carrière [*]		USA: CA
Outgroup: <i>Pinus</i> subsection <i>Australes</i>		
<i>Pinus taeda</i> L. [*]		
Outgroup: <i>Pinus</i> subsection <i>Contortae</i>		
<i>Pinus contorta</i> Douglas ex Loudon [*] var. <i>contorta</i>		
<i>Pinus contorta</i> Douglas ex Loudon var. <i>latifolia</i> Engelm [*]		

phylogenetic theory due to incomplete lineage sorting (ILS). The coalescence of genetic regions and speciation itself are slow and uneven processes over evolutionary time and even clearly diagnosable plant species are not monophyletic for all genic regions. The presence of ILS is well known in *Pinus* (Syring et al. 2007) because many species are relatively young (Germandt et al. 2008) with long generation times and large effective population sizes (Willyard et al. 2007). On the flip side, ILS can create a biased interpretation with anomalous results when based on limited genetic information (especially when using only organellar data; Willyard et al. 2009) and gene flow can also cause the inference of species trees that are inconsistent with ancestry (Solís-Lemus et al. 2016). Plastid data are vital for phylogeography because of their faster

coalescence that can reveal shared patterns that may infer prior sympatry, but plastid data are not always reliable indicators of species because of chloroplast capture. In conifers, the paternal inheritance of plastids (Neale and Sederoff 1989) and wind-dispersed pollen potentially contribute to more geographically widespread plastid lineages than in other plants where plastids are inherited only via seeds. Mitochondria, dispersed in all plants only by seed, can potentially reveal more fine-scale structure within and among ponderosa pine populations than either plastid or nuclear data (Potter et al. 2013). However, it may be even more likely for sympatric species to share these mitochondrial markers than plastid lineages (Zhou et al. 2010). Even phylogenetic trees inferred from biparentally inherited nuclear genes may show a lack of

resolution due to insufficient data as well as ILS. A distortion caused by exemplar plant sampling in the midst of ILS can potentially be problematic. For example, early publication of low copy nuclear gene trees for *P. subsect. Ponderosae* yielded strong support for alternative relationships that cannot each reflect the species tree (Willyard et al. 2009). There are models to analyze gene trees to infer a species tree, but methods that are being used for some types of organisms (e.g., fungi and insects) to delimit species and detect genetic structure in cryptic species remains a formidable challenge in plants where many more unlinked gene trees may be needed for comparable inferences (Fujita et al. 2012; Zhang et al. 2013; Fujisawa et al. 2016). Nevertheless, inferences from hundreds of nuclear gene sequences and organellar data can create a critical foundation to help sort out the trickiest of the species delimitation questions in plants. Species-level molecular phylogenies can be used in concert with taxonomy in several ways. First, a clade on a phylogenetic tree is critical for testing hypotheses of how functional traits evolved (Nock et al. 2016). Second, the integration of molecular and other taxonomic data can be very powerful when used iteratively to explore which putative morphological characters (and/or growth and physiology, biochemical, or ecological factors) might diagnose them and which putative characters were misleading (Dayrat 2005). This iterative process may be critical in some taxonomic groups where useful diagnostic characters have eluded earlier workers despite other indications (e.g., geographic distribution or divergent ecological settings) that they may be forming phylogenetic species (Baum and Donoghue 1995). Third, interspecific relationships revealed on phylogenetic trees can be invaluable by providing an outside perspective on intraspecific variation. For example, a divergent plastid haplotype within part of the geographic range of *P. lambertiana* Douglas was found to be similar to a distantly related and geographically distant group of taxa (Liston et al. 2007). Fourth, even tentative taxonomic names and either putative morphological diagnostic characters or molecular phylogenetic clades can help guide experimental design as well as conservation initiatives. Researchers in other disciplines often use currently accepted taxonomic names as a starting assumption when they compare ecological or other traits. Using a phylogenetic species definition (De Queiroz 2007), monophyly is not required for a species. But, an iterative examination of phylogenetic resolution and putative morphological diagnostic characters can be useful. For example, well-supported paraphyly can be used to recognize improperly “lumped” species, especially intraspecific varieties that do not share ancestry with their basionym. Updating erroneous taxonomic names makes an important contribution to future research, conservation, and land management and it is an important ultimate goal when using morphology and molecules in concert. In addition, well-resolved, yet unnamed, clades towards the tips of a phylogenetic tree may be useful on their own because of their shared ancestry. These clades can be used for phylogeographic modeling to infer migration and refugia, for example. Some of these clades might also warrant being tested as candidate species (Padial et al. 2010), laying a foundation for later taxonomic recognition (Reeves and Richards 2011). Better taxonomic names and well-supported clades in the ponderosa pines will support other research, for example, the important three-way relationship of ponderosa pine trees, bark beetles, and fungal pathogens (Alamouti et al. 2011); the multi-faceted relationship of

ponderosa pine trees and the birds and rodents that consume and disperse their seeds (Vander Wall 2003); predictions of where ponderosa pine forests might survive in a changing climate (Shinneman et al. 2016; Vickers et al. 2019); and adaptive silviculture strategies specific to the ponderosa pine ecosystem with the goal of integrating climate change adaptation into silvicultural planning (Nagel et al. 2017).

Pinus subsection *Ponderosae* is either circumscribed to contain about 18 taxa (Little and Critchfield 1969) or is expanded to subsume three species of California big cone pines (*P. sabiniana*, *P. coulteri*, and *P. torreyana*) that are elsewhere treated as *P. subsect. Sabinianae* Loudon (Price et al. 1998). A fourth species that has large but not heavily armed cones, *P. jeffreyi*, has consistently resolved into a molecular phylogenetic clade with the California big cone pines, and this four-species clade resolved as sister to most other samples of *P. subsect. Ponderosae* (Krupkin et al. 1996; Gernandt et al. 2009; Willyard et al. 2009; Parks et al. 2012; Hernandez-Leon et al. 2013). Although the timing and geographic origin of this lineage remain unclear (Millar 1993), *P. subsect. Sabinianae* is a useful taxonomic group that is well supported by phylogenies. Using previously published phylogenies and diagnostic characters, we began with a working hypothesis that *P. jeffreyi* best fits in *P. subsect. Sabinianae*. We will formalize this taxonomy in the Discussion. For simplicity, we refer to the species in *P. subsect. Ponderosae* as the “ponderosa pines” and to the four species of *P. subsect. Sabinianae* as “California big cone pines”.

Current floristic treatments of the ponderosa pines have taken dramatically opposing approaches to species recognition in the USA compared to Mexico and Central America. In the USA, one species name (*P. ponderosa*) was generally assigned to the vast majority of populations. There has been a long-standing consensus that there are at least two taxa of *P. ponderosa* (Weidman 1939; Wells 1964a, 1964b; Read 1980; Kral 1993; Potter et al. 2013, 2015; Willyard et al. 2017). Within *P. ponderosa*, two or three varieties are usually recognized, currently accepted as *P. ponderosa* var. *ponderosa* and *P. ponderosa* var. *scopulorum* (<http://www.worldfloraonline.org/>) and *P. ponderosa* var. *washoensis* (Baldwin et al. 2012). The two widely accepted varieties (*P. ponderosa* var. *ponderosa* and *P. ponderosa* var. *scopulorum*) have extremely limited mitochondrial introgression where their geographic ranges meet in central Montana (Latta and Mitton 1999; Johansen and Latta 2003). Research findings spanning more than 80 years have confirmed an enormous ecological amplitude within this broadly defined species. *Pinus ponderosa* var. *ponderosa* (including *P. ponderosa* var. *washoensis*) primarily grows west of the Great Basin in climates where most of the precipitation occurs during the winter and spring. *Pinus ponderosa* var. *scopulorum* grows in climates that receive less overall precipitation, but the precipitation occurs during the spring and summer growing season (Conkle and Critchfield 1988; Shinneman et al. 2016). This ecological distinction, along with growth traits and many other types of data, have supported suggestions that these two varieties are not conspecific (Conkle and Critchfield 1988; Willyard et al. 2009). The unrealistic taxonomic stability of the name *P. ponderosa* can largely be attributed to the widespread use of the name because of the ecological and economic importance of these trees. Part of the reticence to change the taxonomy is also due to confusion that originated in the naming frenzy during the early days of rushing new North American species back to Europe to

satisfy the demand for novel plants that were being discovered (Willyard et al. 2017). Affected by this checkered past and by a lack of easily recognizable diagnostic characters, most research publications in recent decades chose to not apply taxonomic names to the distinctive populations in the USA that their results suggested. Rather, they called them races or ecotypes, or assigned new names to varieties within *P. ponderosa* (Haller and Vivrette 2011; Callahan 2013b). Following some recent work (Lauria 1996b; Willyard et al. 2017) we resurrected the early published species names (*P. benthamiana*, *P. brachyptera*, *P. scopulorum*, and *P. washoensis*) for our working hypothesis (Table 1). Conspicuously contrasting their taxonomic treatment in the USA, 11 taxa of *P. subsect. Ponderosae* are currently accepted in Mexico and Central America (Table 1; <http://www.worldfloraonline.org/>) and some botanists recognize additional taxa (Martínez 1948; Stead 1983; Farjon and Styles 1997; Debreczy et al. 2011). Many of these 11 recognized taxa are considered problematic to identify using either morphological or molecular data (Matos 1995; López-Reyes et al. 2015). Only four *P. subsect. Ponderosae* taxa have populations both north and south of the Rio Grande River. *Pinus ponderosa* var. *scopulorum* inhabits the sky islands in Sonora, Mexico just south of Arizona, USA (Kral 1993). *Pinus arizonica* and *P. engelmannii* were both published with type localities in Arizona, USA and then later had populations in Mexico assigned to them (Martínez 1948). *Pinus arizonica* var. *stormiae* was published with a type locality in Coahuila, Mexico, and later had populations in Arizona, USA, assigned to it (Silba 1990).

We accounted for the possibility that individuals with different ancestry are currently growing undetected in comingle populations by using an experimental design that infers separate nuclear and plastome species trees based on individual samples, rather than using frequencies within a population as needed to interpret nSSR (Potter et al. 2015) and cpSSR (Willyard et al. 2017) data. We sampled 2 to 25 individuals each of the accepted *P. subsect. Ponderosae* and *P. subsect. Sabinianae* taxa (Table 1; Appendix 1), and used high-throughput sequencing (HTS) to target a large number of nuclear genes. We assembled nearly complete plastome nucleotide sequences from the same set of HTS short reads, allowing intra-individual comparisons of nuclear and plastome lineages. Using diagnostic characters (including geographic location) as currently understood, we made a prior assignment of each individual to a taxon (Appendix 1). We used these data to visualize the support for taxonomic assignments and to assess their level of monophyly and paraphyly.

MATERIALS AND METHODS

Samples—We selected 130 individuals representing 19 taxa of *P. subsect. Ponderosae* and the four taxa of *P. subsect. Sabinianae* (Table 1; Appendix 1). Samples were identified to a taxon based on morphology and geographical location (Kearney and Peebles 1960; Kral 1993; Farjon and Styles 1997; Arevalo and Gonzalez Elizondo 1998; Eckenwalder 2009; Baldwin et al. 2012; Meyers et al. 2015). Most populations were represented by one random sample, but we included two individuals from each of five locations: ari3 and ari4 (Mt. Lemmon, Arizona, USA), bra06 and bra07 (Spring Mountains, Nevada, USA), max3 and max4 (Lagunas de Montebello, Chiapas, Mexico), mon3 and mon4 (San Salvador el Seco, Puebla, Mexico), and sto2 and sto3 (Davis Mountains, Texas, USA). At Big Bend National Park, Texas, USA, we sampled one individual from each of three stands that are within about 4 km of each other (sto4, sto5, and sto6). Herbarium vouchers for the populations sampled were deposited at the Critchfield Memorial Herbarium (IFGP), Oregon State University (OSC),

and/or Herbario Nacional de México (MEXU). Maps of sample locations (Appendix S1) and Appendices S2–S9 were deposited with Dryad Digital Repository (Willyard et al. 2021).

Guided by published phylogenies (Syring et al. 2005; Willyard et al. 2007; Gernandt et al. 2009; Hernandez-Leon et al. 2013; Gernandt et al. 2018), we chose six samples from two species as outgroups: *P. taeda* (*P. subsect. Australes* Loudon) and *P. contorta* (*P. subsect. Contortae* Little and Critchfield) (Table 1; Appendix 1). To assess whether these taxa were distant enough for proper rooting, we separately conducted trial phylogenies using more distantly related samples from *P. sect. Pinus*. For the plastome, we included GenBank samples of *P. taiwanensis* Hayata (NC027415) and *P. thunbergii* Parl. (NC001631). For the nuclear phylogeny, we included unpublished sequences for the same targeted gene set of *P. pinex* L. and *P. thunbergii*. DNA was isolated as previously reported (Willyard et al. 2017; Gernandt et al. 2018), and quantitated using a Qubit fluorometer (Life Technologies, Carlsbad, California).

Library Preparation—We targeted a set of putatively orthologous low copy nuclear genes (LCNG; Gernandt et al. 2018), most of them *Pinus taeda* unigenes that were successfully captured via hybridization (Neves et al. 2013). One of three custom arrays of genes was used for each sequencing lane to target 711, 1058, or 1416 genes, with 703 genes common to the three arrays (Table 2; Appendix S2). Sheared DNA was ligated to TruSeq barcode adapters (Illumina, San Diego, California) and enriched using 120 bp biotinylated baits designed for our custom arrays (MYBaits; Arbor Biosciences, Ann Arbor, Michigan). For each sample, enriched libraries were combined with unenriched libraries in ratios of 50:50, 60:40, 80:20, or 70:30 to obtain plastid reads as well as the nuclear gene targets (Table 2; Weitemier et al. 2014). Samples were multiplexed at 48, 72, or 96 samples per lane on Illumina HiSeq 2500 (Arbor Biosciences) to yield 100, 125, or 150 bp paired-end reads (Table 2). Reads were archived at the GenBank SRA (PRJNA540071, PRJNA542874, and PRJNA475761).

Assembly, Alignment, and Filtering—Reads were paired into fastq files using their barcode adapters and then quality trimmed with Trimmomatic v. 0.36 (Bolger et al. 2014) using a 4 bp sliding window with a quality threshold of Q20, and retaining sequences with a minimum length of 30 bps. Each of the 703 targeted nuclear genes was assembled using the HybPiper v. 1.2 pipeline (Johnson et al. 2016) using the nucleotide sequences of the LCNGs from the probes as the references. We used a subset of 46 samples to identify genes for which HybPiper inferred potential paralogy. Our first filtering step excluded 78 genes with inferred paralogy. After independently aligning each of the remaining nuclear genes using MAFFT v. 7.3 (Katoh et al. 2005) in Geneious Prime v. 2019 (<http://www.geneious.com>; Kearse et al. 2012), we excluded an additional 17 genes with data for less than 77 of the 130 samples, 11 genes with pairwise identity less than 0.80, and one gene that was determined to be mitochondrial. We used the 600 genes that passed these filtering criteria for phylogenetic analysis. A concatenated nuclear alignment was archived (Appendix S1).

Nearly-complete plastomes were assembled in an iterative fashion using map-to-reference and de novo assembly steps. For each sample, the paired and trimmed reads were first mapped to a plastome reference sequence from *P. section Trifoliae* subsect. *Australes* (*Pinus greggii*; KY963967; Aguirre-Dugua and Gernandt 2017) in Geneious (medium-low sensitivity, 5 iterations). The presumably plastid reads that mapped to this reference plastome were then de novo assembled using SPAdes v. 3.10.0 (Bankevich et al. 2012), and the resulting scaffolds were mapped to the *P. greggii* reference sequence in Geneious (high sensitivity, 5 iterations) to create a temporary chimeric sequence, using the option in Geneious while building a consensus sequence to use the reference nucleotide if no coverage. The original paired reads were mapped to this temporary chimeric sequence in Geneious (medium-low sensitivity, 5 iterations), then chimeric place holders were removed using the option to call N if no

TABLE 2. Sequencing lane conditions. Parameters used for 8 lanes of Illumina sequencing.

Lane	Lane	Samples per lane	Probe version	Number of loci targeted	Illumina read length	Percent captured
1	GP01	48	v1	711	100	50%
2	GP02	48	v1	711	100	80%
3	GP04	48	v2	1058	100	70%
4	GP05	72	v2	1058	125	60%
5	GP07	96	v3	1416	150	60%
6	WP01	48	v1	711	100	80%
7	WP02	48	v1	711	100	70%
8	WP03	72	v2	1058	100	60%

coverage. For some samples, this assembly and the original reads were processed through ABySS v. 2.0 (Jackman et al. 2017) to create our final, nearly-complete plastome assembly for each sample. Annotations were transferred from the reference sequence in Geneious after aligning the final assemblies for each sample. Numerous minor manual adjustments were made to the alignment of plastomes, especially in regions with tandem near-repeats. Each assembled plastome sequence was archived at GenBank (Appendix 1) and the plastome alignment was archived (Appendix S4).

Phylogenetic Trees—Separate trees were inferred from LCNG and plastome data for every method. Gaps were treated as missing data for all analyses. We used Modeltest-NG (Darriba et al. 2019) on the CIPRES science gateway (<http://www.phylo.org/>) to infer the best nucleotide substitution model for LCNG and plastome data. To create Maximum Likelihood (ML) trees, we used RAxML v. 8.2.12 (Stamatakis 2014) on CIPRES using the GTRGAMMA plus proportion of invariable sites model with 1000 rapid bootstrap inferences followed by a thorough ML search to choose the best tree. Although estimating the proportion of invariable sites is not recommended for RAxML, we also ran it without this parameter for comparison. The nuclear tree was based on a concatenated alignment of 600 LCNG and the plastome tree on the entire plastome alignment. We used Gblocks v. 0.91b (Castresana 2000) on the plastome alignment, first using the program's calculation of default block parameters for our dataset: a minimum of 65 conserved sequences, 108 minimum flanking sequences, 8 maximum nonconserved sequences, a minimum block length of 10 nucleotides, and no gaps. We repeated the Gblocks analysis allowing gaps and reducing the minimum number of flanking sequences to 65 (the smallest allowed by the program) and the minimum block length to 5. Phylogenetic trees were estimated on these reduced alignments using RAxML as described above.

We used Dendroscope v. 3.5.9 (Huson and Scornavacca 2012) to visualize the congruence of the resolution of each sample between the nuclear and plastid ML trees in a tanglegram. The invert tree option in iTOL (Letunic and Bork 2019) was used to create a static view of an inverted tree to show the sample-to-sample connections. We also created trees for these nuclear and for these plastome data with maximum parsimony (MP) using PAUP* v. 4.0a167 (Swofford 2003) with 1000 bootstrap replicates, heuristic searches, maxtrees set at 100, and saved 50% majority-rule consensus trees.

To create species trees, we first inferred a gene tree using RAxML as described above for each of the 600 LCNG. Because plastid genes may not always behave as a single locus (Gonçalves et al. 2019), we repeated this process and created a gene tree for each of 100 plastid genes extracted from the annotated plastome alignment. Separately for the nuclear and the plastid sets of gene trees, we used ASTRAL-III v. 5.6.2 (Zhang et al. 2018) to infer species trees by assigning samples to their morphological taxonomic assignment (Appendix 1). Two samples were pruned from the nuclear species after tree construction because assignment to original species could not be confirmed. We visualized congruence between species resolutions on the nuclear and plastid trees with a tanglegram as above.

We also selected eight nuclear clades, seven plastome clades, and one plastome grade that were consistently recovered across inference methods, labeled N to signify nuclear clade 1–8 (i.e., N1–N8) and P to signify plastome clade or grade 1–8 (i.e., P1–P8) on all trees where they were found. Samples that resolved within the eight nuclear clades (N1–N8) were used in ASTRAL as a species map for the 600 LCNG gene trees and samples that resolved within the eight plastome clades or grades (P1–P8) were used for the 100 plastid gene trees. We included ben01 with N1, bra07 with N5, and pon3 with P2 for clade mapping. Because they were consistently monophyletic, the samples for the *P. subsect. Sabinianae* species and the two outgroup taxa were mapped to their respective species in these clade trees. We created a simplified view of the congruence between these major nuclear and plastome clades by counting the number of samples in each clade on the all-sample tanglegram to manually show the pairings between nuclear and plastome clades on the clade trees. A set of maps were created to show the sometimes-differing geographic pattern for these nuclear and plastome clades. These maps were color-coded by species, using symbols to show the nuclear and the plastome clade for each sample.

Monophyly Tests—For the 12 taxa with more than three samples (Appendix 1) and for both of the ingroup subsections, we applied the Approximate Unbiased (AU) test in CONSEL v. 0.02 (Shimodaira and Hasegawa 2001; Shimodaira 2002) using the set of bootstrap trees from 600-gene LCNG data analyzed in RAxML with the topology constrained to monophyly for each taxon.

TABLE 3. Comparison of nuclear and plastome alignments used for analyses.

Dataset	Characters (bp)	Variable sites (bp)	Proportion of variable sites	Parsimony-informative sites (bp)	Proportion of sites PI
600 low copy nuclear loci	603,747	100,918	0.17	51,043	0.085
Plastomes	141,779	32,694	0.23	11,298	0.080

RESULTS

Sequencing, Assembly, and Alignment—We obtained a mean of 5,882,699 reads per sample for 130 samples (min = 1,639,156; max = 13,791,580; Appendix S2). A mean of 566,942 of these reads (min = 12,068; max = 4,456,880) were plastid sequences based on the number of reads that mapped to the reference plastid genome (Appendix S2). For the LCNG, we aligned 126 samples that met our quality criteria for 600 of the 703 targeted genes with a mean length per gene of 1006 bps (StDev = 503.6) for a mean of 109.5 (StDev = 15.9) samples per gene (Appendix S5). This yielded a concatenated alignment of 603,747 bps with 100,918 variable sites and 51,043 parsimony informative sites (Table 3). Our plastome alignment was successful for 124 samples, with a mean of 121,030 bps (Stdev = 590) per sample (Appendix S2). This yielded an alignment of 141,779 bps with 32,694 variable sites and 11,298 parsimony informative sites (Table 3). The plastome alignment allowing only the conserved blocks as inferred by Gblocks yielded 85,677 bps (60%) using the program's calculated parameters and 121,950 bps (86%) using less stringent parameters. Trees tested with *P. subsect. Pinus* samples as outgroups resolved the same ingroup topologies as those rooted with *P. contorta* and *P. taeda*. Because we had more nuclear information for this dataset, we presented all of our results rooted with the branch leading to *P. contorta* and *P. taeda*.

Modeltest—For LCNG and plastomes, the AIC and AICc criteria favored GTR+I+G4, while the BIC criterion favored TIM2+I+G4 for the LCNG alignment and TPM1uf+I+G4 for the plastome alignment.

All-Sample Trees—Except for nodes with poor support, the phylogenies were mostly robust to method: nuclear ML trees were similar with or without estimation of the proportion of invariable sites (Fig. 1; Appendix S6) and these ML trees were similar to the MP tree (Appendix S7). The plastome ML tree inferred using only the conserved blocks was a rake for most clades and the plastome ML tree inferred using a less stringent estimation of conserved blocks had fewer resolved nodes and shorter branches than the tree inferred with the total alignment (results not shown). However, plastome trees using the entire alignment were similar to each other comparing these same methods and we used these trees for subsequent comparisons (Fig. 1; Appendix S8).

We labeled three nodes on all trees to highlight congruence and a few well-supported conflicts: A, the divergence of *P. subsect. Ponderosae* from *P. subsect. Sabinianae*; B, divergence of the majority of *P. benthamiana*, *P. ponderosa*, and *P. washoeensis* samples (labeled bpw) from all other taxa; C, two major lineages of non-bpw taxa. Nuclear and plastome trees were congruent for the relationship of *P. subsections Ponderosae* and *Sabinianae* (Node A) except for the placement of clade P1 with five *P. benthamiana* samples as sister to *P. subsect.*

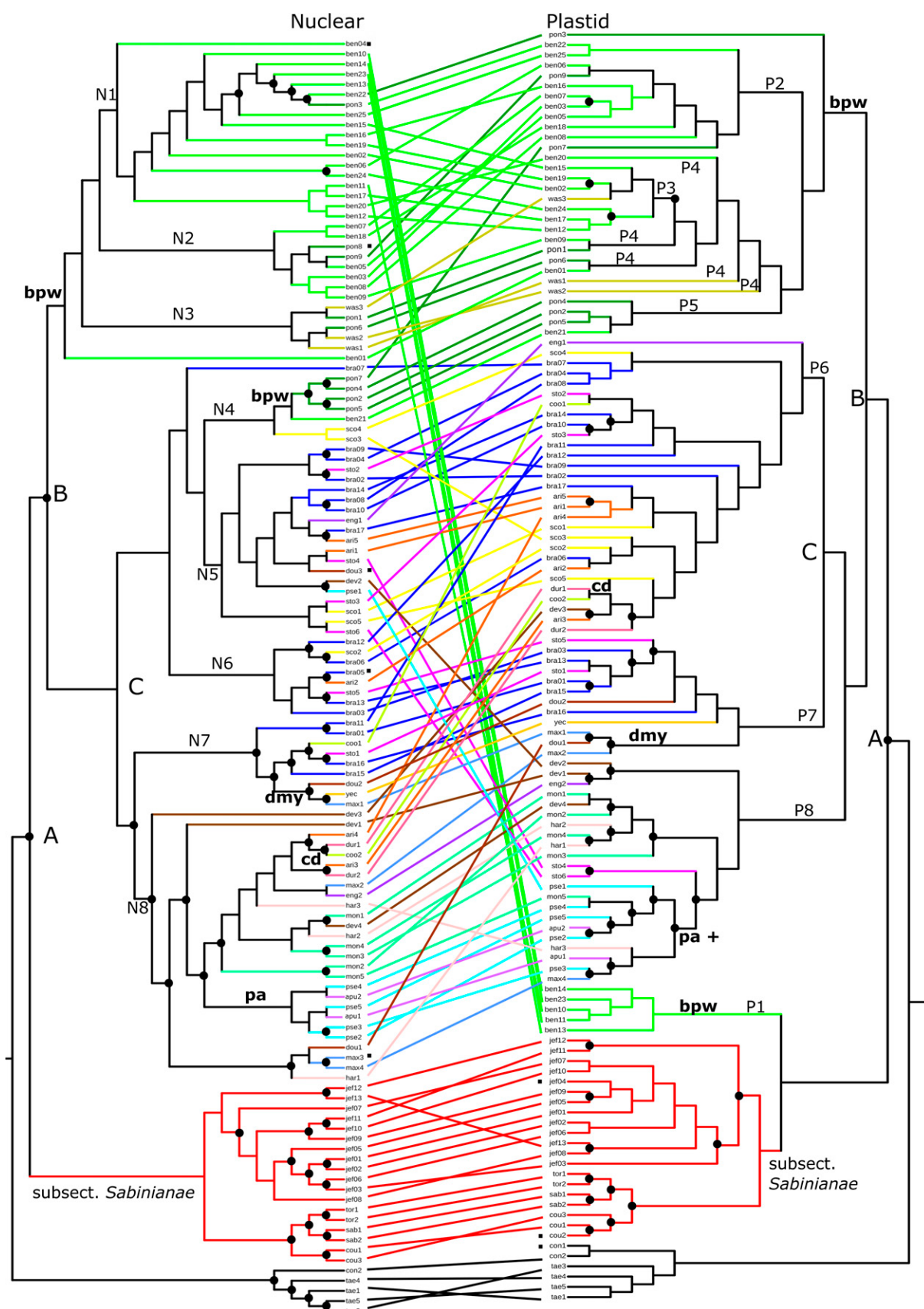


FIG. 1. Congruence and conflicts between nuclear and plastid phylogenies inferred with RAxML using a concatenation of 600 low copy nuclear genes (left) and the whole plastid alignment (right). Trees were rooted on the branch leading to outgroup samples. Colors correspond to taxa, with all subsect. *Sabinianae* in red (see Fig. 2). *A priori* sample assignments to taxa are given in Appendix 1. Node A (divergence of subsect. *Ponderosae* and subsect. *Sabinianae*, including a clade of five *P. benthamiana* samples on the plastome tree); node B (divergence of most bpw from all other subsect. *Ponderosae* taxa); node C (two major subclades in non-bpw group); four clades consistently recovered: bpw (*P. benthamiana*, *P. ponderosa*, *P. washoensis*), cd (*P. cooperi* and *P. durangensis*); dmy (*P. douglasiana*, *P. maximinoi*, *P. yecorensis*), pa (*P. pseudostrobus*, *P. pseudostrobus* var. *apulcensis*); labels N1–N8 and P1–P8 are clades (or grades) selected for further analysis. Nodes with circle symbols were supported by bootstraps of 80% or greater.

Sabinianae on the plastome tree (Fig. 1). Within *P.* subsect. *Ponderosae*, nuclear and plastome trees were also mostly congruent in separating the bpw group from all other taxa, except that four *P. ponderosa* samples and a *P. benthamiana* sample resolved in a clade with two *P. scopulorum* samples (N4; Fig. 1).

Of our six locations where more than one individual was sampled, only the *P. maximinoi* (max3 and max4 from Lagunas de Montebello, Chiapas, Mexico) and the *P. montezumae* samples (mon3 and mon4 from San Salvador el Seco, Puebla, Mexico) resolved as sisters. The same-population samples of

P. arizonica (ari3 and ari4 from Mt. Lemmon, Arizona, USA), *P. brachyptera* samples (bra06 and bra07 from Spring Mountains, Nevada, USA), *P. arizonica* var. *stormiae* samples (sto2 and sto3 from Davis Mountains, Texas, USA), and *P. arizonica* var. *stormiae* samples (sto4, sto5 and sto6 from Big Bend National Park, Texas, USA) populations resolved into different clades (Fig. 1).

Species Trees—The ASTRAL species tree built from 600 nuclear gene trees and the ASTRAL species tree built from 100 plastid gene trees showed a mostly congruent resolution of the basal nodes and some species groups (Fig. 2) and with

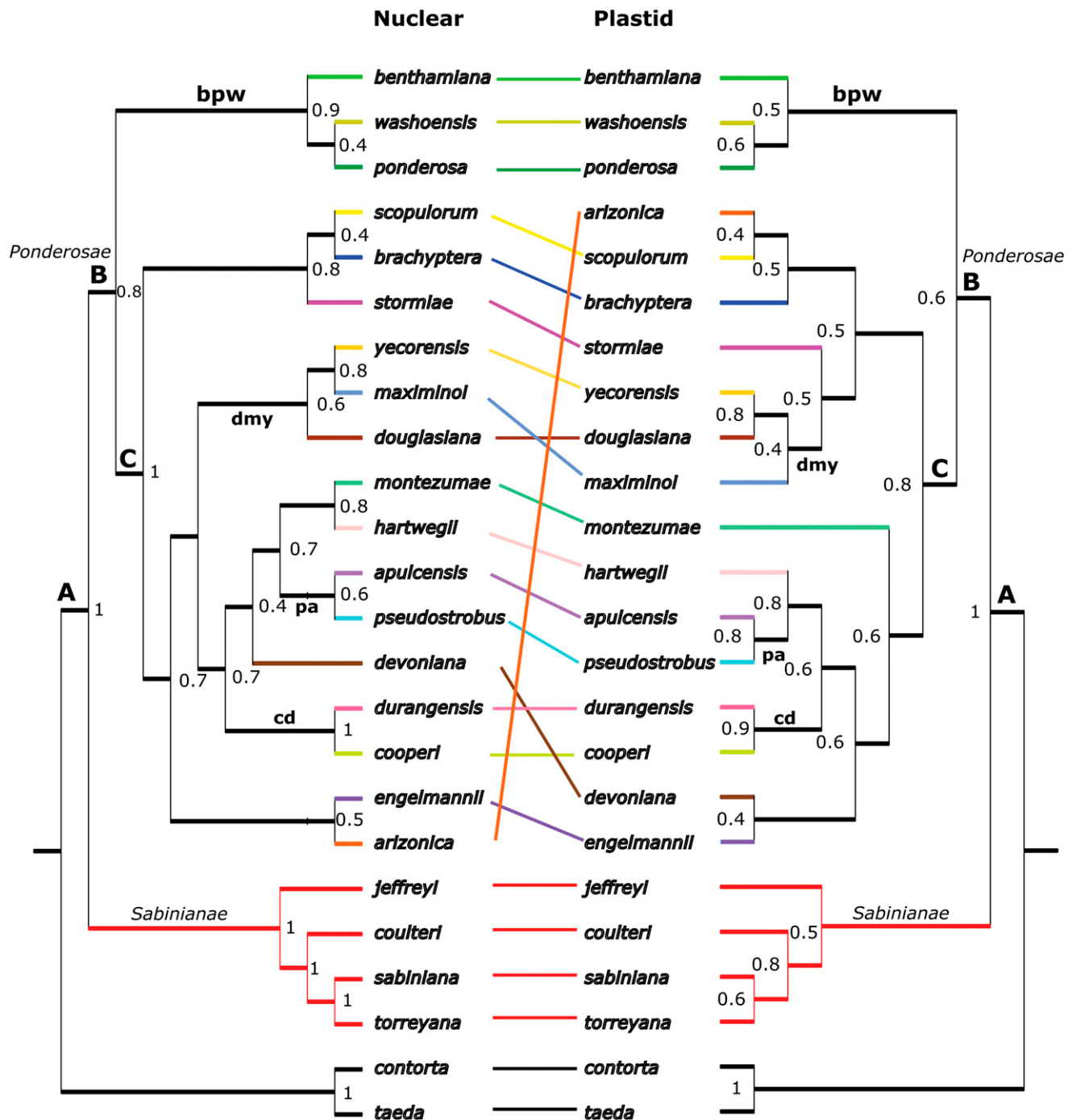


FIG. 2. Species tree congruence and conflicts between nuclear (600 low copy nuclear gene trees) and plastid (100 plastid gene trees), each inferred using ASTRAL with samples assigned to *a priori* taxa (Appendix 1). Trees were rooted on the branch leading to outgroup taxa; numbers are posterior probabilities; node labels A, B, C, bpw, dmy, np, and cd were defined in Fig. 1.

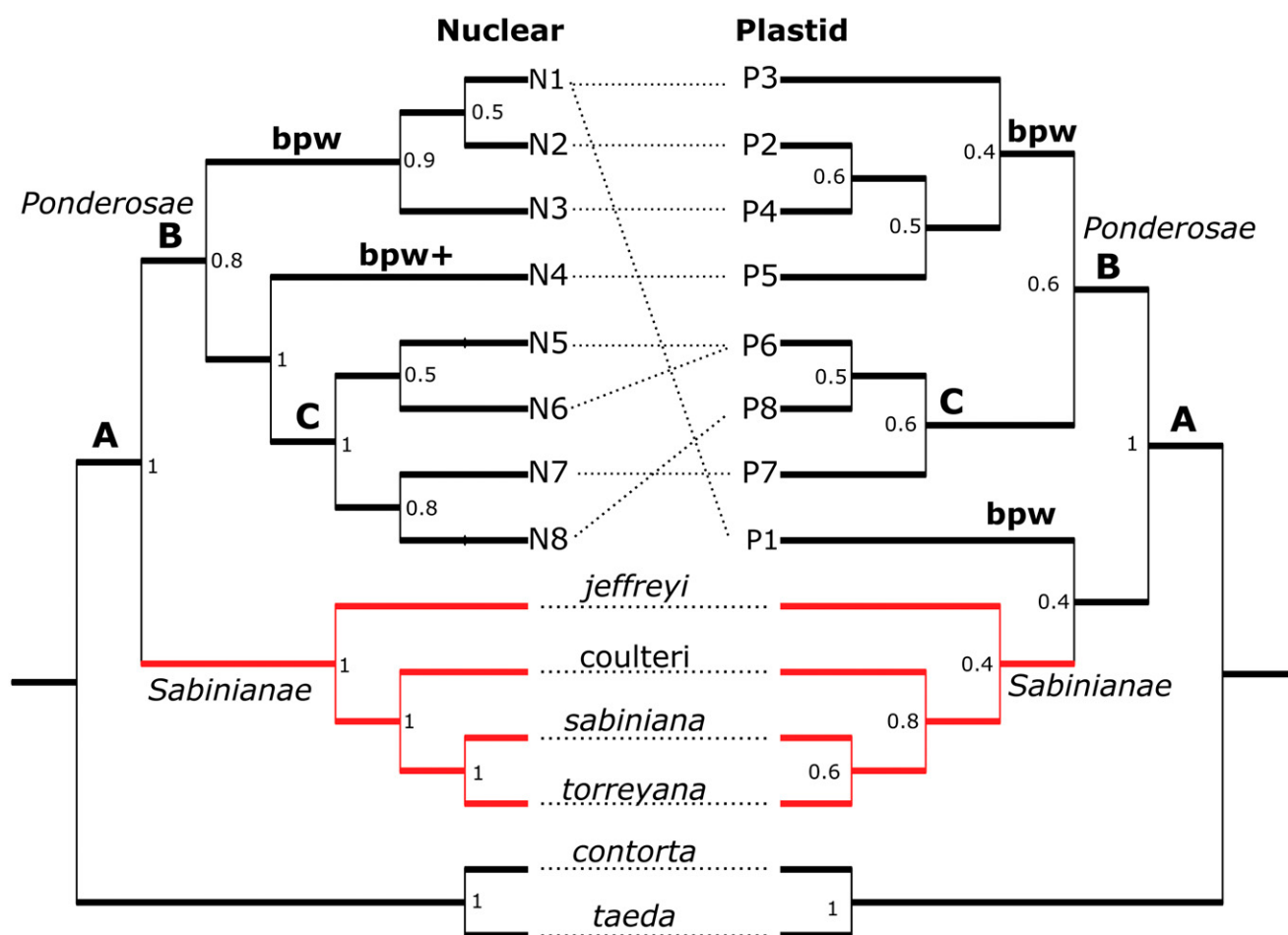


FIG. 3. Congruence and conflict between nuclear and plastid trees built using the same method as Fig. 2, except with samples assigned to clades N1—N8 and P1—P8 based on their resolution in Fig. 1. Dotted lines were drawn manually to connect the nuclear and plastid clades where most samples were resolved (Fig. 1). Trees were rooted on the branch leading to outgroup taxa; numbers are posterior probabilities; node labels A, B, C, and bpw were defined in Fig. 1.

the all-samples trees (Fig. 1). *Pinus* subsections *Ponderosae* and *Sabinianae* resolved with posterior probability = 1 on both species trees at Node A. Within *P.* subsect. *Ponderosae*, the nuclear and plastid species trees supported the separation of the bpw clade (*P. benthamiana*, *P. ponderosa*, and *P. washoensis*) from the other taxa with BS = 0.8 at Node B. These species trees also had consistent resolutions for the groups labeled dmy (*P. douglasiana* + *P. maximinoi* + *P. yecorensis*, but with different branching order), pa (*P. pseudostrobus* + *P. pseudostrobus* var. *apulcensis*), and cd (*P. cooperi* + *P. durangensis*). Placement of taxa differed dramatically between nuclear and plastid species trees for *P. arizonica* (sister to *P. engelmannii* on the nuclear tree and to *P. devoniana* on the plastid tree) and *P. arizonica* var. *stormiae* (sister to *P. scopulorum* and *P. brachyptera* on the nuclear tree and the dmy clade on the plastid tree).

Monophyly—The confidence intervals for the AU values were broad, with each taxon and the subsections having a $p < 0.05$ rejecting monophyly for at least one of the BS trees as well as at least one BS tree with a p -value supporting monophyly (Appendix S9). Although every taxon and the subsections had a majority of BS trees with a $p < 0.05$ that rejected monophyly, six taxa (*P. jeffreyi*, *P. devoniana*, *P. douglasiana*, *P. brachyptera*, *P. benthamiana*, and *P. arizonica* var. *stormiae*) and each of the subsections had mean p -values that did not reject monophyly.

Clade Trees—The ASTRAL trees built by mapping samples to the eight nuclear or the eight plastid clades selected from all-sample phylogenies (Fig. 1) highlighted the three previously-identified nuclear-plastid incongruences (Fig. 3). First, the five *P. benthamiana* samples in plastome P1 were also resolved sister to *P.* subsect. *Sabinianae* in this plastid tree, and their nuclear N1 was sister to N2. The remaining samples in N1 with plastome P3 resolved as expected. Second, although nuclear N8 and plastome P8 share most individuals, these groups resolved with different sisters on nuclear and plastome clade trees. Third, plastome P6 contained individuals that were resolved in either nuclear N5 or N6 (Figs. 1, 3). The geographic pattern of these major nuclear and plastid clades differed somewhat (Figs. 4–6). The nuclear and plastome clades of individual samples also differed. For example, some *P. benthamiana* individual samples carried N1 and P1, while other *P. benthamiana* individuals had N1 and P2 (Fig. 4), and there were other examples of this phenomenon for most taxa.

DISCUSSION

Monophyly—Despite the expected impact of ILS in *Pinus* (Syring et al. 2007) and the nuclear/plastid incongruence that has been reported in *P.* subsect. *Ponderosae* (Willyard et al.

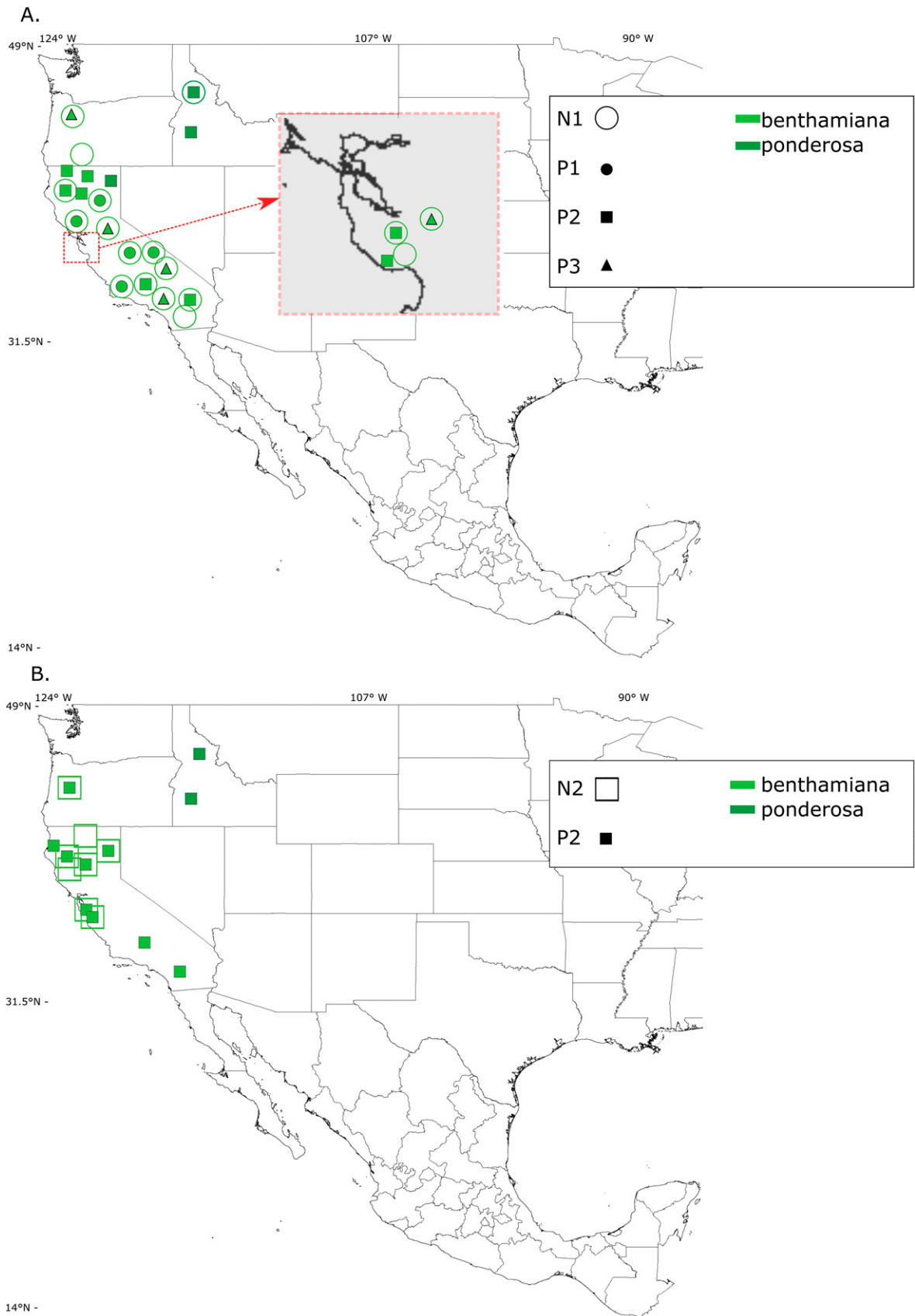


FIG. 4. Phylogenetically related samples (Fig. 1) at their approximate collection site (Appendix 1; Appendix S1). Colors show the prior taxon assignment; large symbols show the nuclear clade; and nested smaller symbols show the plastome clade for each individual. A. Nuclear clade N1 and plastid clades P1, P2, and P3. B. Nuclear clade N2 and plastid clade P2.

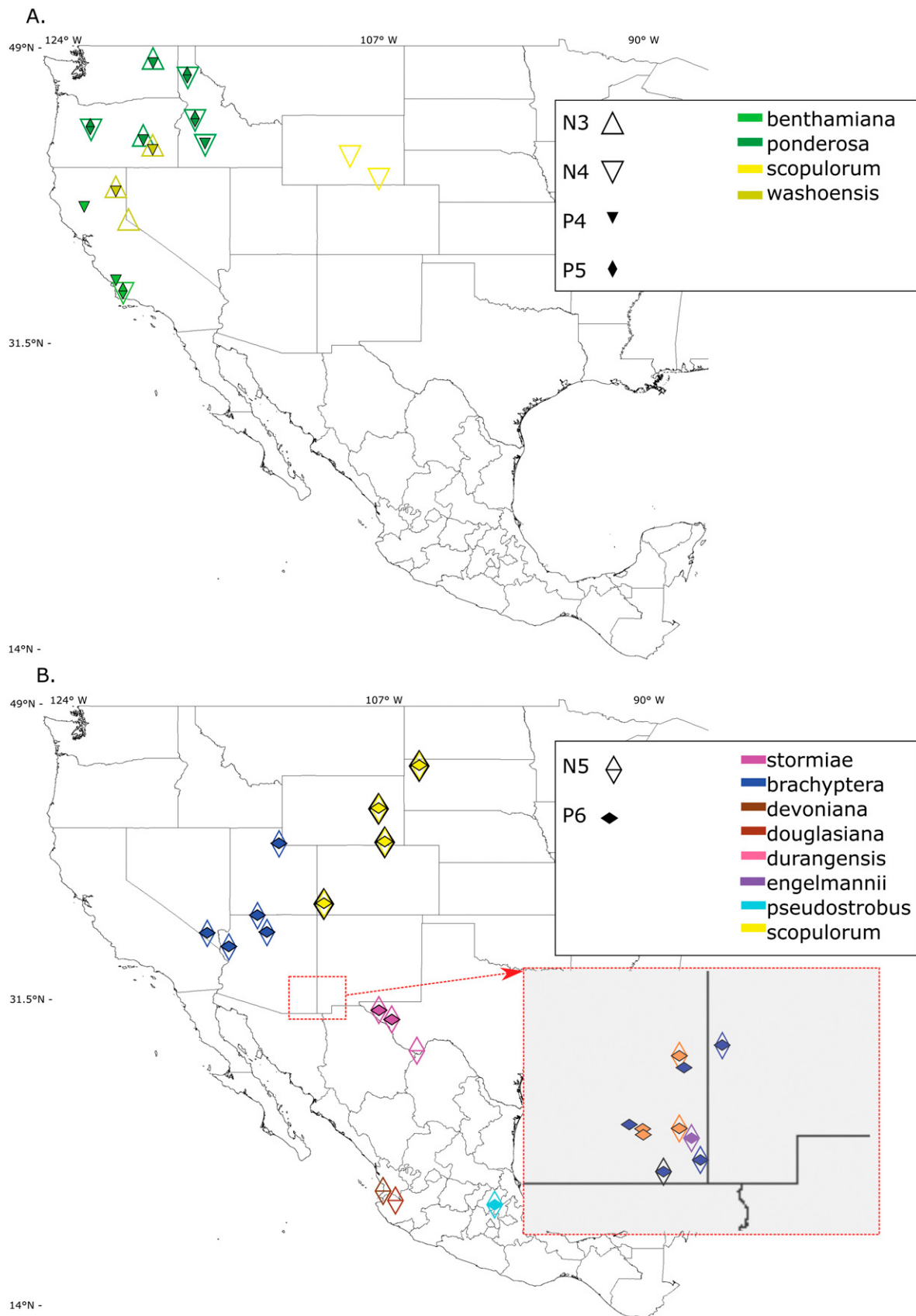


FIG. 5. Phylogenetically related samples (Fig. 1) at their approximate collection site (Appendix 1; Appendix S1). Colors show the prior taxon assignment; large symbols show the nuclear clade; and nested smaller symbols show the plastome clade for each individual. A. Nuclear clades N3 and N4, and plastid clades P4 and P5. B. Nuclear clade N5 and plastid clade P6.

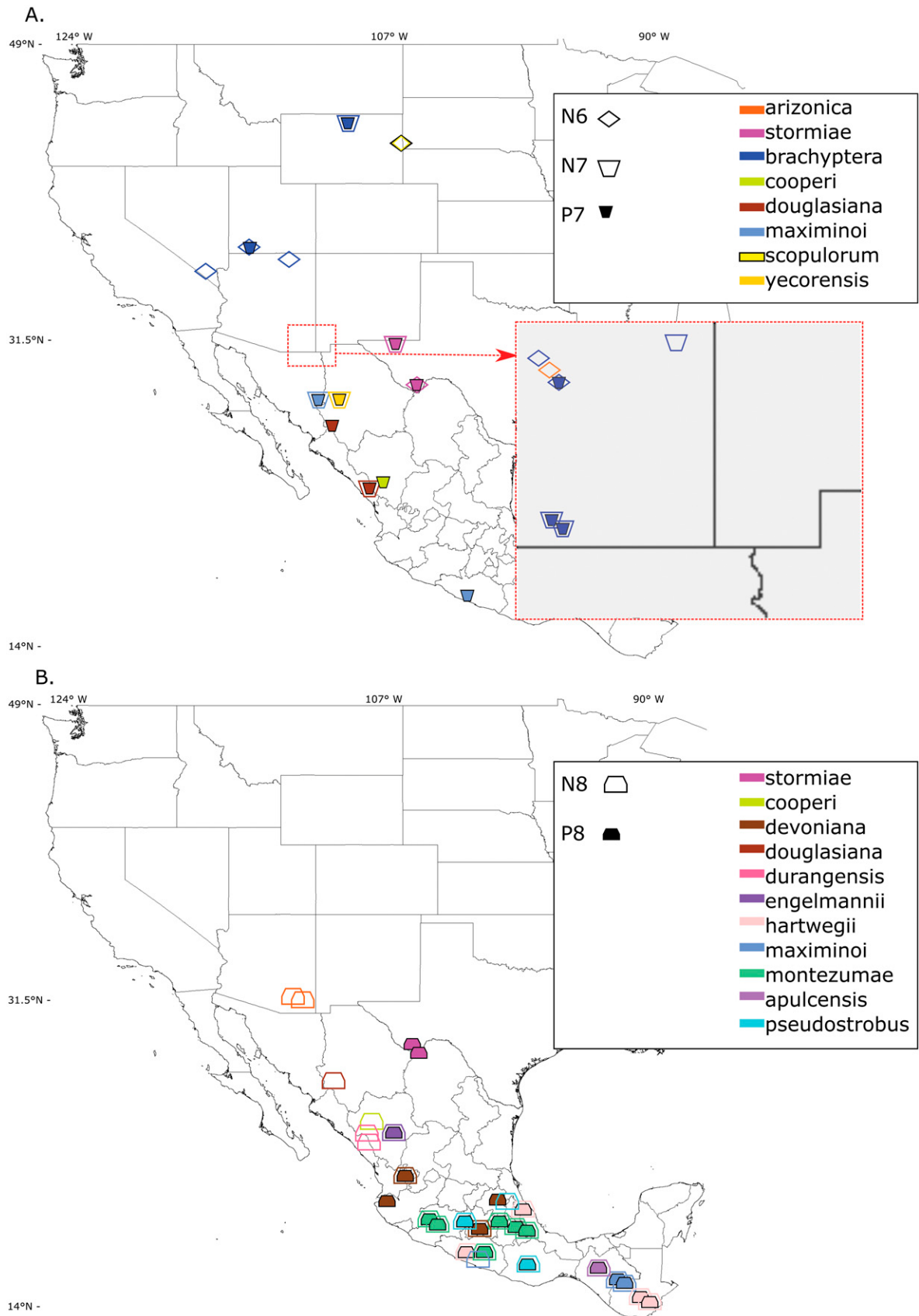


FIG. 6. Phylogenetically related samples (Fig. 1) at their approximate collection site (Appendix 1; Appendix S1). Colors show the prior taxon assignment; large symbols show the nuclear clade; and nested smaller symbols show the plastome clade for each individual. A. Nuclear clades N6 and N7, and plastid clade P7. B. Nuclear clade N8 and plastid clade P8.

2009), our phylogenies were useful in resolving some relationships in this group of hard pines using several samples per species with 600 nuclear genes and nearly complete plastomes. Although node support tended to be low, the topologies were robust to method and largely (but not exclusively) congruent between nuclear and plastome. This allowed us to focus on a few fairly well supported nuclear/plastome conflicts and on how well the clades that were recovered matched named taxa. We evaluated the monophyly of two taxonomic subsections and each named taxon in two ways: 1) by their resolution on phylogenetic trees, and 2) by their support in the AU test based on LCN. When rooted on the branch leading to outgroups in *P. subsect. Contortae* and *P. subsect. Australes*, our trees consistently showed *P. subsect. Sabinianae* (including *P. jeffreyi*) to be reciprocally monophyletic to *P. subsect. Ponderosae* (node A) except for the plastome clade labeled P1. Despite the placement of P1, the monophyly of these two subsections was not rejected by the AU test (Appendix S9).

As found in another group of pine species (Syring et al. 2007), samples rarely formed a clade with all other samples assigned to the same taxon. This result was comparable to previous estimates of monophyly in *P. subsect. Ponderosae* using plastid data, where about 0.31 of taxa were monophyletic (Hernandez-Leon et al. 2013). Nevertheless, we only statistically rejected monophyly for one-half of the named taxa for which we had three or more samples using the AU test (Appendix S9). There were only three groups of taxa that were recovered as monophyletic on most phylogenetic trees, labeled dmy, pa, and cd (Fig. 1).

Although ASTRAL has been shown to perform well (Zhang et al. 2018), our lack of evidence for monophyletic resolution of assigned taxa (Fig. 1) should inspire a cautious interpretation of the species tree (Fig. 2) because it infers the species topology based on our prior assignment of samples to each nominal taxon. The genetic structure revealed in the clades resolved using multispecies coalescence are not necessarily delimiting taxa (Sukumaran and Knowles 2017), but these clades are worthy of investigation. As a comparison to the species tree using prior taxonomic assignments, we investigated the consistently recovered clades as an alternative “species tree” to begin an investigation of whether these consistently recovered clades might contain previously unrecognized taxa (Fig. 3). This also allowed a simpler view of the conflict between nuclear and plastome clades. Based on the number of samples in common, nuclear clades N5 and N6 corresponded to plastome clade P6; N7 corresponded to P7, and N8 corresponded to P8 (Fig. 3). Importantly, because we inferred nuclear and plastome phylogenies from the same individual and the same lane of nucleotide sequencing reads, we were able to study intra-individual variation. We found that the nuclear and plastome histories of many individual samples were incongruent. The nuclear-plastome conflict could be attributed to ILS or the ancient acquisition and persistence of different plastome lineages. For example, nuclear N5 included 13 samples from the USA plus one sample each of *P. devoniana*, *P. douglasiana*, and *P. pseudostrobus* (Fig. 1). Nuclear N6 contained only samples of *P. arizonica*, *P. brachyptera*, *P. scopulorum*, and *P. arizonica* var. *stormiae* from the USA, but both N5 and N6 samples had plastomes that were placed in P6. We observed a geographic pattern for each clade (Fig. 4–6), but no obvious correspondence to morphological characters that are currently in use to diagnose taxa.

For example, N4, N5, and N7 each included specimens with needles per fascicle that varied widely. We discuss the differing geographic pattern for nuclear and plastome clades below.

Paraphyly—Our phylogenetic trees also allowed us to identify well-supported examples of paraphyly. At node B, the branch labeled bpw on all of our phylogenetic trees contained *P. benthamiana*, *P. ponderosa* sensu stricto (s. s.), and *P. washoensis*, and the other branch encompassed all of the remaining taxa, including *P. brachyptera* and *P. scopulorum*, which are currently accepted as varieties of *P. ponderosa* sensu lato (s. l.; Figs. 1, 2, 3). This makes *P. ponderosa* s. l. paraphyletic at node B, and this topology was found for both nuclear and plastome phylogenies regardless of method. Because of this paraphyly, we strongly suggest that future experiments not start with an assumption that *P. ponderosa* s. l. is monophyletic. Some experiments, e.g., (Roberts and Hamann 2015) have even continued to use the abandoned taxonomic assignment of *P. ponderosa* var. *arizonica* to include a geographic distribution for *P. ponderosa* in the Sierra Madre Occidental (Little 1971) for decades after this varietal assignment was abandoned. Future experiments that analyze *P. ponderosa* var. *ponderosa* and *P. scopulorum* as one widespread taxon may further muddle reconstructions of their genetic variability as well as the geographic ranges occupied by these taxa in the past and where they might grow under future climate scenarios. More useful ecological niches have been modeled for *P. ponderosa* s. l. that correspond to mitochondrial haplotypes (Shinneman et al. 2016; Maguire et al. 2018). Because the geographic distribution of mitochondrial haplotypes does not correspond to either the nuclear or the plastid clades that we recovered here in any simple way, ecological niche models for the nuclear and plastid lineages would make an important comparison.

On the “non-bpw” branch, there were 15 putative taxa that diverged into two clades at node C. Except for *P. arizonica*, taxon placements on the two subclades of node C were congruent between nuclear and plastome species trees (Fig. 2). However, resolution nearer the tips showed a poor agreement of samples with taxonomic assignments (Fig. 1). In some cases, this lack of species-level monophyly can be attributed to unnatural taxonomic assignments that have been reported before. For example, many researchers have struggled to find consistent morphological, genetic, or ecological distinctions for all of the populations treated as *P. ponderosa* var. *scopulorum* (Rehfeldt 1999a; Willyard et al. 2017). We confirmed some of the broad geographic associations reported in mitochondrial haplotypes (Potter et al. 2013) and plastid phylogeny (Gernandt et al. 2009). Our sample from the Bighorn Mountains, Wyoming, USA (bra01), a sample from the Guadalupe Mountains, Texas, USA (sto1) and three samples from the Sky Islands of Arizona, USA (bra 11 (Pinalenos), bra15 (Mt. Wrightson), and bra16 (Mt. Hopkins)) resolved in the nuclear clade N7 with the dmy clade (most of the *P. douglasiana*, *P. maximinoi*, and *P. yecorensis* samples) plus a sample of *P. cooperi* (coo1). Most of the N7 samples were also resolved in the plastome clade P7. These previous findings and our results strongly suggest that the *P. brachyptera*/*P. scopulorum* question needs to be investigated in the context of closer relationships to taxa only recognized in Mexico. Previous international taxonomic investigations have concluded that the ponderosa pines in the sky islands of southern Arizona and southern Texas are related to those in the nearby Sierra

Madre Occidental in Mexico, assigning *P. arizonica* and *P. engelmannii* to populations in Mexico and *P. arizonica* var. *stormiae* to populations in the USA (Martínez 1948; Silba 1990; Farjon and Styles 1997; Arevalo and Gonzalez Elizondo 1998; Eckenwalder 2009). Our phylogenies suggested that there was shared ancestry for these samples with some accessions from Mexico (N5), but not with their currently assigned conspecifics (Fig. 5B). Because our sampling of *P. arizonica*, *P. engelmannii*, and *P. arizonica* var. *stormiae* in Mexico was limited, our finding that one sample of *P. engelmannii* from the Chiricahuas, Arizona, USA (eng1) did not resolve with its putative conspecific (eng2, Durango, Mexico) on any phylogenetic tree is preliminary. We also could not confirm or strongly reject the placement of *P. arizonica* var. *stormiae* and *P. cooperi* as varieties of *P. arizonica*. Based on these findings, we suggest that an extension of this taxonomic investigation is needed much more widely in the USA and Mexico. Ultimately, denser sampling of individuals (especially in Mexico), with morphological characters on the same individual used for molecular data, will be needed to reveal whether the 15 taxa below node C (including samples that we assigned to *P. brachyptera* or *P. scopulorum*) are a variable species complex or whether we simply have not yet identified diagnostic morphological characters.

Reticulation and Introgression—Because of the greater ease of obtaining plastid data and the faster coalescent times of plastid lineages compared to nuclear lineages, it has been suggested that plastid DNA would better delimit pine taxa (e.g., Zhou et al. 2010). However, this is clearly not always the case. Chloroplast capture has been widely documented in *Pinus* (Matos and Schaal 2000; Willyard et al. 2009). Localized mitochondrial movement among taxa has only been detected in one location across a narrow band of sympatry between *P. ponderosa* var. *ponderosa* and *P. scopulorum* (Latta and Mitton 1999; Johansen and Latta 2003), but few other comparative datasets have been available to test *Pinus* mitochondrial sharing. The expectation for *Pinus* is that mitochondrial lineages will reveal maternal clusters based on seed dispersal, while plastid lineages will reveal less genetic structure because of wider pollen dispersal as well as through seed dispersal (Latta et al. 1998). In fact, some ponderosa pine populations were found to have nuclear genetic clusters within about 10 m apart that varied genetically from nearby clusters because of localized maternal seed distribution patterns (Linhart et al. 1981). Experiments that have measured organelle lineages in the ponderosa pines have generally supported this expectation, with most populations nearly fixed for a mitochondrial haplotype (Potter et al. 2013), but each population more variable for cpSSRs (Willyard et al. 2017). Although pine pollen has the potential to disperse long distances (Bruczyk et al. 2004), some local stand densities can restrict the effective pollen movement substantially (Dyer and Sork 2001). This means that the nuclear, plastid, and mitochondrial genetic structure of ponderosa pine populations are still in need of study to evaluate whether any incongruent phylogenies (Willyard et al. 2009) were due to reticulate evolution. As discussed above under monophyly, the intra-individual conflict between nuclear and plastome ancestry was widespread in our experiment. However, some of these conflicts may be explained by different levels of lineage sorting in these data sets. For example, the samples in nuclear clade N1 that have either of the sister plastid clades P2 or P3 could be attributed to different plastome phylogenetic resolution (Figs. 1, 4A).

Reticulate evolution caused by the introgression of nuclear genes is even harder to detect in the presence of ILS, but our dataset had some evidence for nuclear reticulation. Because our nuclear and plastome trees were largely concordant, we were able to detect four categories of reticulate evolutionary history. Recent intra-individual plastid capture is supported in only one geographic location. In the second category of reticulate evolution, plastid lineage P1 has persisted after an ancient plastid transfer. In the third type of reticulation, the placement of five samples of *P. ponderosa* s. s. in nuclear N4 despite morphological and plastome (P5) assignment to bpw, indicates ancient nuclear introgression. The fourth category was our observation that there was a widespread pattern of nuclear and plastid lineages that have different ancient histories. We discuss each category in turn.

CHLOROPLAST CAPTURE—Recent plastid transfer within the bpw clade is the simplest explanation for the nuclear resolution of *P. washoensis* with *P. ponderosa* s. l. (N3) in contrast to their plastome placement with sympatric *P. benthamiana* (P4; Figs. 1, 7). Introgression in the Sky Islands of Arizona has been reported to be more complicated (Peloquin 1984) and we did not identify any clear examples of hybrid origin. Rather than recent first-generation hybrid individuals, these incongruences potentially involved both nuclear and plastome lineages shared with taxa currently only recognized in Mexico. Ironically, we consider the three sympatric species (*P. arizonica*, *P. brachyptera*, and *P. engelmannii*) to be morphologically diagnosable in the Sky Islands of Arizona, USA. Our interpretation of these taxa in this study was limited by our incomplete sampling of both taxa in Mexico. We sampled seven *P. brachyptera* individuals in the Sky Islands of Arizona (bra11, bra12, bra13, bra14, bra15, bra16, and bra17), where some introgression between *P. brachyptera* (a 3-needle pine) and *P. arizonica* (a 5-needle pine) may occur. In the Huachuca Mountains, Arizona, USA, an individual clearly identified morphologically as *P. brachyptera* (bra17) resolved with *P. arizonica* samples in both nuclear and plastome phylogenies (Fig. 1). On Mt. Lemmon, Arizona, USA, our two samples of *P. brachyptera* (bra12, bra13) resolved as expected with other *P. brachyptera*, or sister to *P. arizonica* var. *stormiae* or with *P. scopulorum* (Fig. 1), supporting previous conclusions that introgression is uncommon among sympatric *P. arizonica* and *P. brachyptera* (Epperson et al. 2009; Marquardt et al. 2018).

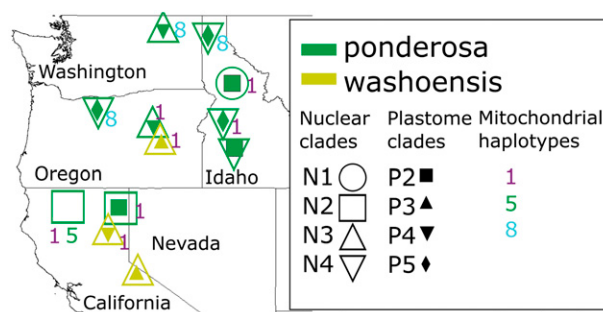


Fig. 7. For nine samples of *P. ponderosa* (green) and 3 samples of *P. washoensis* (tan) determined morphologically (Table 4), the nested symbols show their nuclear and plastome clades (Fig. 1) and adjacent numbers give the predominant mitochondrial haplotype found in this geographic area (Potter et al. 2013).

TABLE 4. The nine samples of *P. ponderosa* and three samples of *P. washoensis* shown in Fig. 7. Nuclear and plastome clades show this sample's phylogenetic resolution in Fig. 1; mitochondrial haplotypes are for the predominant mitochondrial haplotype found in this geographic area (Potter et al. 2013).

Sample	Location	Nuclear clade	Plastome clade	Mitochondrial haplotype(s)
pon1	USA: Washington: Bisbee Mtn.	N3	P4	8
pon2	USA: Idaho: Coeur D'Alene	N4	P5	8
pon3	USA: Idaho: Kooskia Rd	N1	P2	1
pon4	USA: Idaho: Last Chance	N4	P5	1
pon5	USA: Oregon: Santiam Pass	N4	P5	8
pon6	USA: Oregon: Blue Mtns.	N3	P4	1
pon7	USA: Idaho: Boise	N4	P2	n/a
pon8	USA: California: McCloud	N2	n/a	1, 5
pon9	USA: California: Likely	N2	P2	n/a
was1	USA: Oregon: Blue Mtns.	N3	P4	1
was2	USA: California: Likely	N3	P4	1
was3	USA: Nevada: Mt. Rose	N3	P3	n/a

ANCIENT CHLOROPLAST CAPTURE—In the second category, our phylogenies indicated that the ancient capture of a plastid lineage has persisted in some geographical regions. The five bpw samples that resolved in plastid clade P1, sister to *P. subsect. Sabinianae* (Fig. 1) were a clear example. Previously, a few samples of *P. benthamiana* from the Sierra Nevada, California, USA were resolved sister to *P. jeffreyi* using plastid genes, leading to a suggestion that *P. benthamiana* was hybridizing with (or at least capturing chloroplasts from) sympatric *P. jeffreyi* (Gernandt et al. 2009, their sample CAL2; Parks et al. 2012, their additional file 4; Willyard et al. 2009, their Fig. 6). Because our present study used both nuclear and plastome data from the same individuals and compared them with samples of other species of *P. subsect. Sabinianae*, our results provided a much stronger alternative explanation. Some individuals that resolved as expected in nuclear N1 had the unexpected plastome P1 (Figs. 1, 3) that is sister to all of *P. subsect. Sabinianae*. In addition to the western Sierra Nevada samples (ben10, ben13, and ben14), a sample from Clearlake, California, USA (ben11), and one from the Transverse Range, California, USA (ben23) carried plastid P1 (Fig. 4A). This strongly supported result is yet another reason to not infer taxonomic origin or introgression in this group of hard pines based solely on plastid data.

INTROGRESSION—In the third category, we make the first report that a footprint of nuclear introgression between *P. ponderosa* and *P. scopulorum* persists in the Bitterroot Mountains, Idaho, USA despite no present-day sympatry. This geographic region was identified in a “hot spot” cluster of contact zones and phylogeographic breaks reported for various species of plants and animals (Swenson and Howard 2005). Previously, the only gene flow that had been identified between *P. ponderosa* s. s. and *P. scopulorum* was in a narrow contact zone in Montana, USA (Latta and Mitton 1999; Johansen and Latta 2003). Based on consistent phylogenetic resolution among different methods, the nuclear lineage N4 apparently involved ancient reticulation between bpw and *P. scopulorum*. Interestingly, these individuals retained plastids that resolved with bpw in P5 or P2 (Fig. 1), in agreement with cpSSR results (Willyard et al. 2017). The Bitterroot Mountains of Idaho, USA also had mitochondrial haplotype 8 related to bpw (Potter et al. 2013). When these clade placements were used in lieu of taxon assignments to create the ASTRAL clade tree, N4 again resolved as sister to the non-bpw clades and plastome P5 was again placed within bpw (Fig. 3). The bpw samples in the N4 clade, like the plastome P1 discussed above, had a substantial branch length

(Appendix S6). Clade N4 contained five bpw, along with two *P. scopulorum* samples from Wyoming (sco3 and sco4; Fig. 1). Although these *P. ponderosa* samples are among the easternmost bpw, there is a large geographical distance (ca. 450 km) to the nearest present-day *P. scopulorum* and our *P. scopulorum* samples that resolved in N4 are about 800 km away (Fig. 5A). Based on clear paraphyly of *P. ponderosa* s. l. (discussed above), their geography, and on the conflicting nuclear and plastome resolution of this clade placement, ancient admixture during a time period in which there was secondary contact with *P. scopulorum* is the simplest explanation for this topology. This reticulation example included three samples from Idaho, USA (pon2, pon4, and pon7), but not pon3 (Kooskia Rd, Idaho, USA). What's more, one sample from Oregon, USA (Santiam Pass; pon5) resolved here. The fifth sample was a *P. benthamiana* from near the Ventana Wilderness in coastal California, USA (ben21). This agreed with a previous study where cpSSR frequencies from this population placed it with a *P. ponderosa* rather than a *P. benthamiana* group (Willyard et al. 2017). In the present work, ben21 also had a plastome placement in P5 with *P. ponderosa* samples. This site is only about 20 km from the Big Creek (ben20) population that resolved as expected in nuclear N1 with *P. benthamiana* (Figs. 4A, 5A). Alternative explanations for this coastal California site are an unpublished offsite planting (Hipkins and Krutovsky 2005) or long-distance dispersal. Based on our results from three of our four samples from Idaho, USA, we suggest that there is a footprint remaining from an ancient *P. ponderosa* - *P. scopulorum* admixture zone in Idaho and perhaps the admixture has extended (or introgressed individuals have migrated) as far west as central Oregon.

NUCLEAR AND PLASTOME COALESCENT PATTERNS DIFFER—Other conflicts between the nuclear and plastome phylogenies, such as P6 that mostly contains a large grade of individuals that have either N5 or N6, may simply be due to better nuclear than plastome resolution in this particular experiment, along with ILS in both nuclear and plastome phylogenies. The story is not quite that simple, though. The plastome and nuclear lineages that are robust to the phylogenetic method have different histories. For example, most individuals in nuclear N8 have plastome P8 (Fig. 1), yet these clades have very different shared ancestries (Fig. 3). There is a geographic pattern to the nuclear clades and a geographic pattern for the plastome clades, but they are different from each other (Figs. 4–6). These maps show symbols for the nuclear and the plastome clade nested in each individual. Importantly, there are many

instances that if an experiment sampled only plastome data, the patterns would not reflect the nuclear (and presumably the species) resolution. For example, individuals that carry nuclear N1 could have either plastome P2 or P3 or the very distantly related plastome P1 (Figs. 3–6). Although the geographic pattern of these plastid clades is roughly congruent with the pattern shown with cpSSRs (Willyard et al. 2017), there are some major differences with mtHaps (Potter et al. 2013) and with ecotypes (Weidman 1939; Wells 1964a; Read 1980; Callahan 2013a). A selective advantage to this regional sharing of plastome lineages would be interesting to test, perhaps starting with the correspondence of these plastome clades to ecological niches (Maguire et al. 2018).

Despite these examples, our results indicate that introgression among the ponderosa pines is probably infrequent. Hybridization among *Pinus* taxa and even the origin of the diploid hybrid species *P. densata* (Wang et al. 2001) have been extensively documented. Although previous studies often point to introgressive hybridization as the likely explanation for individual ponderosa pine trees that seem to have intermediate morphologies or overlap in putative diagnostic characters among populations, hybridization among sympatric taxa usually occurs at a low frequency, and several experiments have concluded that backcrossing does not seem to be affecting the genetic integrity of each parental taxon (Peloquin 1984; Rehfeldt 1999a). In artificial crosses, many of the ponderosa pines (including taxa with huge geographic disjunctions) yielded viable seed, although only a limited number of tests were successful when crossing *P. jeffreyi* with *P. ponderosa* (Conkle and Critchfield 1988). There is evidence for common natural hybridization between *P. engelmannii* and *P. arizonica* using AFLPs (Ávila-Flores et al. 2016) and between *P. montezumae* and *P. pseudostrobus* using plastid simple sequence repeats (cpSSRs; Delgado et al. 2007) and using morphology (Caballero Deloya 1967). However, hybrid frequency was rare to uncommon between *P. hartwegii* and *P. montezumae* based on plastid haplotypes (Matos and Schaal 2000), between *P. jeffreyi* and *P. ponderosa* based on morphology (Haller 1962), between *P. coulteri* and *P. jeffreyi* based on monoterpenes (Smith 1977) and based on morphology (Zobel 1951), and between all combinations of *P. arizonica*, *P. brachyptera*, and *P. engelmannii* using morphology (Peloquin 1971).

Sympatry—Most workers describe sympatry in only three groups of *P. subsect. Ponderosae* species in the USA: *P. jeffreyi*, *P. benthamiana*, *P. ponderosa*, *P. washoensis*, and *P. coulteri* (plus limited parapatry in *P. sabiniana*) in California, USA; *P. arizonica*, *P. engelmannii*, and *P. brachyptera* in Arizona, USA; and the contact zone between *P. ponderosa* and *P. scopulorum* in central Montana, USA. In Mexico, *P. arizonica* and *P. engelmannii* are sympatric, as are *P. douglasiana*, *P. maximinoi*, *P. pseudostrobus*, and *P. yecorensis* (Martínez 1948; López-Reyes et al. 2015). Despite the examples of putative hybridization discussed above, there is evidence for the long-term maintenance of *P. subsect. Ponderosae* taxa in sympatry (Peloquin 1984; Rehfeldt 1999a). Our sampling for this study was generally limited to one collection in a geographic region per taxon, using our prior assignments (Appendix 1). Nevertheless, we found four examples of individuals from geographically proximal locations that were resolved into different clades. The best explanation for these four examples is that different taxa are growing in sympatry (or at least parapatry) that have remained unrecognized.

The first example of sympatry is *P. washoensis*. Some have attributed the differences between *P. ponderosa* and *P. washoensis* to growing conditions. Under this explanation, individuals assigned to Washoe pine look different because they are growing on colder and windier higher elevation sites (as occur at their type locality on Mt. Rose, Nevada, USA) than *P. ponderosa* (Rehfeldt 1999b; Sorensen et al. 2001). An alternative is that at some locations, *P. washoensis* individuals are growing among the *P. ponderosa*, contributing to high population-averaged variability (Sturgeon 1979; Smith 1981; Critchfield 1984; Niebling and Conkle 1990). Failure to recognize an intermingled sympatry has been suggested as a contributing factor to the confusion that persists to the present day regarding the description of *P. ponderosa* (Haller and Vivrette 2011). The original ponderosa pine collection by David Douglas in 1836 near Spokane, Washington, USA (likely near our pon2) was a branch that was filed in the herbarium with Viscaceae to document a new species of *Arceuthobium*. If this sterile specimen is designated as a lectotype for *P. ponderosa*, the cones would now be known only from original drawings, which fit *P. ponderosa* (Lauria 1996a). However, the seeds that Douglas used to document the new pine that he named *P. ponderosa* were collected at Douglas' request by trapper John Work near Kettle Falls, Washington, USA (likely near our pon1). Trees grown from this seed still exist in European pinetums. These now-mature trees display a mixture of cone sizes and shapes, some of which would be attributed to *P. washoensis*. This suggests that Work collected a mixture of seeds from sympatric taxa. Haller and Vivrette (Haller and Vivrette 2011) noted that these pinetums have created a long-term, albeit unofficial, common garden experiment that is now demonstrating that cone and needle characters for Washoe pine have a genetic rather than ecological basis. In our field work in 2011, we observed that there was a large diversity of cone sizes and shapes on Bisbee Mountain near Kettle Falls, Washington, USA near where at least some of this legacy seed was likely collected. Our molecular results might also support two sympatric taxa, possibly with introgression between them (Fig. 7). We consistently resolved nuclear clade N3 with our three samples that represent *P. washoensis*, but this clade also included two samples with morphological characters that fit *P. ponderosa*: one from the presumed collection site of the seed for the European pinetums (pon1, Bisbee Mountain, Washington, USA) and one from the Blue Mountains, Oregon, USA (pon6). All of the other *P. ponderosa* samples resolved in nuclear clade N2 or N4 (Figs. 1, 7). Our plastome clades and published mitochondrial haplotypes revealed that these populations contain incongruent organellar lineages as well (Fig. 7). Along with the mixed morphological characters observed in the pinetums, this suggests that at least two ancestral lineages are persisting in sympatry with organellar (and probably some nuclear) introgression and that the other individuals that have *P. ponderosa*-style morphological characters are either related to *P. benthamiana* (the ones in N2) or unexpectedly to *P. scopulorum* (the ones in N4 discussed above under Reticulation). Future experiments that sample more individuals for genomic and morphological inferences will be needed to resolve this question.

The second example where we documented previously unrecognized sympatry is in the Chisos Mountains of southern Texas, USA. These isolated populations are of conservation interest due to projected climate change because their

limited sky island habitat is surrounded by desert. We assigned these ponderosa pines in Big Bend National Park (Texas, USA) to *P. arizonica* var. *stormiae*, a taxon with its main distribution in Coahuila, Nuevo León, Tamaulipas, and San Luis Potosí, Mexico based on the suggestion of potential occurrence in the Chisos, Guadalupe, and Davis mountains (Silba 1990) as adopted in local hypotheses (Wauer and Fleming 2002; Alex 2004). There are only three stands of ponderosa pines in Big Bend National Park, all within about 4 km of each other. A previous experiment that sampled only the USA *P.* subsect. *Ponderosae* included populations from two of those stands in their cpSSR analysis (Willyard et al. 2017). We included one individual from each of the three populations in this study. The sample from Crown Mountain (sto4) was in N5, sister to a *P. arizonica* (Pinalenos Arizona, USA), despite not sharing the 5-needle character of that taxon. Sample sto4 resolved in plastome P8, in contrast to its previous cpSSR clustering with the *P. scopulorum* group. The sample from Pine Canyon (sto5) was in nuclear N6 and plastome P7, sister to *P. brachyptera* samples on both trees. This is congruent with previous cpSSR clustering of the Pine Canyon population with *P. brachyptera*. The sample from Boot Canyon (sto6) was in N5 with the Crown Mountain sample, but sister to *P. scopulorum*. In contrast, the sto6 plastome resolved in P8 near a clade of pa samples. From this modest sampling, it appears that migration into the Chisos Mountains has occurred from three different ancestral populations or at least from a very genetically diverse population that had different nuclear and plastome lineages.

The Spring Mountains of southeastern Nevada are also comparatively disjunct from other ponderosa pines, and the population of ponderosa pines there harbored two different mitochondrial haplotypes, neither of which was identified elsewhere in the USA (Potter et al. 2013). We included two samples from this contiguous population, which both fit with *P. brachyptera* based on the morphological characters. In our nuclear tree, one sample resolved as sister to N5 and the other one in N6. These two samples were only distantly related to each other in the large plastome clade P6. Although ILS appears to persist deeply in the subclades of node C, it is noteworthy that the Spring Mountains have now been shown to harbor different nuclear, plastome, and mitochondrial lineages in one apparently continuous stand. This finding should caution future experiments to consider the possibility that even isolated stands of ponderosa pines may not be homogeneous.

Our fourth example is in the Bighorn Mountains, Wyoming, USA. Previous reports found that the ponderosa pines there belonged to a cpSSR group and a mitochondrial haplotype similar to *P. brachyptera* (Potter et al. 2013; Willyard et al. 2017), despite being in the geographical area historically included with *P. scopulorum* (Little 1971). Our phylogenies placed a sample from the Bighorn Mountains (bra01), in nuclear N7 rather than with the *P. scopulorum* that we sampled that are within about 180 km (sco1, sco2, sco3; Figs. 5B, 6A). Our nearest *P. brachyptera* was a sample from Utah (bra02) that is over 500 km away from the Bighorn Mountains (Fig. 5B). What's more, this Utah sample resolved distantly based on nuclear data (N5; Fig. 1). Our fieldwork noted different immature cone colors on adjacent individuals in the Bighorn Mountains, as has been noted in other regions (Smith 1981). It is also interesting that individuals in a newly established stand of ponderosa pines in the grasslands west of the

Bighorn Mountains contained individuals with high genetic diversity, suggesting the seed was dispersed from different parent populations (Lesser et al. 2013). As with the Chisos and Spring Mountains, the Bighorn Mountains are worthy of a future experiment that examines each individual to test the possibility of even finer-scaled sympatry within the Bighorn Mountains.

Together, these results strongly suggest that future experiments on the ponderosa pines should be designed to test the correlation of molecular with morphological (Stead 1983; López-Reyes et al. 2015), and anatomical (Whang et al. 2001; Whang et al. 2004) characters among individuals rather than using frequencies within and among populations. These morphological studies will need to address several challenges. First, some characters are ephemeral, e.g., the color of immature ovulate cones, especially in non-masting years; seed sizes and shapes because seeds take two years to mature, and are then quickly dispersed and/or damaged by insects and animals; ovulate cone scales that are not mature or that have apices that have been abraded; and leaf anatomy that cannot easily be observed in preserved specimens. Seedling characters (e.g., number and color of cotyledons) either have to rely on an assumption that a seedling in its natural setting germinated from a nearby seed parent, or seed must successfully be obtained and germinated from the same parent tree being sampled for other characters. The character state variability within an individual ponderosa pine tree, within a population, and among populations is also very complicated. This has been attributed to many factors, and it seems likely that alternative explanations each explain some of the variability among different groups of pine taxa. For example, the number of needles per fascicle was heritable in a large common garden experiment involving *P. ponderosa* s. l., supporting geographically identifiable "forms or races" (Weidman 1939) despite some observed year-to-year variability (Haller 1965). In contrast, a strong cline in the number of needles per fascicle was identified along an elevational gradient for *P. engelmannii*, and progeny grown from 5-needled *P. engelmannii* adult trees only rarely (2%) produced 5 needles per fascicle in a common garden (Yeaton et al. 1983). In the sky islands of Arizona, USA, there is an elevational cline in the number of needles per fascicle that was hypothesized to represent hybridization. Instead, this difference can be attributed to a 5-needled taxon (*P. arizonica*) at the lower elevations that overlaps the distribution of a variable-needle taxon (*P. brachyptera* with 3, 4, or 5 needles per fascicle) whose distribution extends to higher elevations (Epperson et al. 2009; Marquardt et al. 2018). Another example of complex variation is the resin ducts in the needles. The position of the resin ducts may help distinguish some taxa, e.g., the ponderosa pines of Hidalgo, Mexico (Espinoza-Pelcastre et al. 2018), even though the number of resin ducts is highly variable among needles of some taxa (Martínez 1963).

Local Adaptation and Plastic Growth Responses—The identification of morphological synapomorphies for the ponderosa pines is complicated by several factors that are of ecological interest. For example, some growth characteristics and morphological variation within and among *P. ponderosa* s. l. populations and among *P. jeffreyi* populations has been attributed to local adaptation. California tree seed zones dictate that progeny for reforestation be grown from seed sources in elevational bands within geographic zones (Buck et al. 1970). This scheme was based on evidence for differential growth

by source elevation, especially on the western slope of the Sierra Nevada, California, USA (Callaham and Liddicoet 1961; Wells 1964a; Rehfeldt 1999b). Similar local adaptation by elevation exists in the ponderosa pine forests in Arizona and New Mexico, USA (Dixit and Kolb 2020). Local adaptation was found in *P. jeffreyi* growing on ultramafic soils (Furnier and Adams 1986). An abrupt transition in growth traits of ponderosa pine forests west and east of the crest of the Cascade Range in southern Oregon, USA was attributed to local adaptation for temperature and precipitation regimes (Sorensen et al. 2001). This study noted that an abrupt change in immature cone color and resistance to bark beetles (*Dendroctonus brevicornis* LeConte) also occurs in the same transition zone, leading to their conclusion that the populations on either side have adapted to their local conditions. Coevolution with bark beetles (directional selection for limonene) was proposed for differences in terpenoid chemistry in this geographic region (Sturgeon 1979).

Some research has concluded that differences among populations are due to a plastic growth response. Measurable differences in growth characteristics and morphology were maintained in progeny grown under drought, but gas exchange and carbon allocation showed a plastic response to their common growing conditions (Cregg 1994; Zhang and Cregg 2005). Other studies have found no difference in growth for putative taxonomic categories. For example, despite diagnostic purple-black immature ovulate cones and deep purple pollen cones in *P. washoensis*, the lack of significant growth differences led to a suggestion that this taxon was synonymous with *P. ponderosa* (Rehfeldt 1999b). Some differences among populations have been suggested to be entirely explained by phenotypic plasticity, with characters observed in their original setting not being manifest in seedlings grown in a common garden (Yeaton et al. 1983; Maher et al. 2002).

Other researchers have concluded that differences among ponderosa pine populations were due to genetics. Based on growth characteristics of progeny in large, replicated common garden experiments, genetic variation among populations was found to form geographic clusters that correspond to source populations (Weidman 1939; Wells 1964a; Read 1980). Although the variation in the types and proportions of monoterpenes present in the ponderosa pines does not correspond directly with their shared ancestry because of differential expression (Smith 1982; Keefover-Ring and Linhart 2010; Adams and Wright 2012), some terpenoid comparisons have added evidence for genetic differences among putatively distinctive groups of ponderosa pine populations (Smith 1977; Adams and Edmunds 1989; Adams et al. 2015). Genetic differences among populations of *P. subsect. Ponderosae* have also been identified recently using plastid (Willyard et al. 2017), mitochondrial (Potter et al. 2013), and nuclear evidence (Potter et al. 2015). The presence of isolated ponderosa pine populations in several unique ecological conditions has been hypothesized to represent cryptic taxa which may be unique. As discussed in turn below, most of the seven isolated or putatively distinctive populations that we sampled did not gain support from our results for any taxonomic distinction. Nevertheless, we added evidence for one published (but currently not accepted) taxon (*P. jeffreyi* var. *baja-californica* Silba 1990) and that one population (ben01, Fort Lewis, Washington, USA) is more genetically distant than would be expected based on previous data.

The geographically isolated population at Ft. Lewis, Washington, USA is also ecologically isolated in the Puget Sound lowlands. This population had different biochemical concentrations than other populations in the Pacific Northwest, USA (Gerson et al. 2009). Our phylogenies were in agreement with this result, showing a potential distinctiveness for our one sample from Ft. Lewis, Washington, USA (ben01). It resolved as sister to all other samples in the main bpw nuclear clade using ML (Fig. 1). This sample had a plastome that was sister to a *P. ponderosa* from the Blue Mountains, Oregon, USA (pon6), suggesting the possibility of an ancient chloroplast capture. No diagnostic morphological characters have been presented and no taxonomic name has been published for this isolated group of ponderosa pines that is currently protected on the Joint Base Lewis-McChord, but our results suggest that it is worth investigating whether the trees in this stand are genetically as well as ecologically distinct. The ponderosa pines in the Willamette Valley, Oregon, USA are somewhat isolated from populations in the Cascade Range to the east, and from the Coastal and Klamath Ranges to the south. These ponderosa pine forests were more widespread before European settlement (Hibbs et al. 2002; Bouffier et al. 2003; Gerson et al. 2009; Kerr et al. 2015). Our results placed two samples from these populations (ben02 and ben03) as expected with other *P. benthamiana* in the bpw clade, in agreement from growth trials (Wells 1964a) and with the most recent floristic treatment (Meyers et al. 2015). There are ponderosa pines in the redwood forest on patches of Arnold series sandhill soil (Griffin 1964) in the Santa Cruz Mountains, California, USA, that are somewhat isolated from other ponderosa pine populations in the general region that do not attract as much attention because they are not in the midst of the redwood forest. These sandhill soil populations may be the type locality for *P. benthamiana* Hartweg (but see discussion below under Systematics). Although some have suggested that their needles and their cones are longer than other ponderosa pines (Lauria 1996b), our results for three samples (ben16, ben17, and ben18) as well as other systematic evaluations (Haller and Vivrette 2011) did not find evidence for their distinctiveness from other *P. benthamiana*. We did find potential support on our nuclear phylogeny for *P. jeffreyi* var. *baja-californica*. Two samples, jef12 (Parque Nacional Constitucion de 1857) and jef13 (Parque Nacional Sierra de San Pedro Martir), originated in Baja California, Mexico as described for this intraspecific taxon (Silba 1990). These two samples formed a nuclear clade that was sister to our other *P. jeffreyi* samples, although they were not monophyletic on the plastome tree (Fig. 1).

The *P. brachyptera* populations in the sky islands of Arizona, USA, are isolated on mountains that are surrounded by the Chihuahuan desert, with each mountain range providing a limited amount of suitable ponderosa pine habitat. In addition to curiosity regarding how those populations were able to migrate to their current locations (Anderson 1989) and how they will adapt to climate change (Rehfeldt et al. 2014; Marquardt et al. 2018, 2019), there is morphological and genetic variation that has inspired some to suggest that the ponderosa pines on these sky islands might form a unique taxon (Rehfeldt et al. 1996; Rehfeldt 1999a). What's more, we observed in our fieldwork that the ponderosa pines on Mt. Hopkins and Mt. Wrightson were hard to diagnose using the traits that we had chosen to distinguish *P. brachyptera* from *P. engelmannii*. We sampled seven populations in the Arizona

sky islands (bra 11–bra17) and found them resolved in divergent clades. Our southernmost USA collections (bra14, Chiricahuas and bra17, Huachucas) resolved in N5 with *P. brachyptera* from Hualapai (bra09) and bra08 (South Rim Grand Canyon), as well as *P. scopulorum* and taxa identified only in Mexico (Figs. 1, 5B). Our two samples from Mt. Lemmon (bra12 and bra13) resolved in N6 with taxa mostly from the southwestern USA. Our sample from Mt. Hopkins (bra15) and our sample from Mt. Wrightson (bra16) resolved in clade N7, along with bra01 (Bighorn Mountains, Wyoming, USA), bra 11 (Pinalinos, Arizona, USA), and the Mexican dmy group (Figs. 1, 6A). As with some other geographic regions that we sampled, this contributes to the evidence that the *P. brachyptera* taxonomic assignments that we used represent a paraphyletic species. As discussed above in the section on Sympatry, the *P. brachyptera* in the Spring Mountains of Nevada, USA may harbor ponderosa pine trees that are unique, but their heterogeneity remains to be explored.

Samples from central Montana, the Black Hills of South Dakota, and the Niobrara region of Nebraska, USA were grouped into a Northern Interior ecotype using results from progeny tests in a common garden (Wells 1964a). With different sampling for a common garden, a High Plains ecotype was separated from a Black Hills ecotype (Read 1980). Neither of these large studies included samples from the Bighorn Mountains, Wyoming, USA. Populations in the northern plains region (including the Bighorn Mountains) had a mitochondrial haplotype that extended southwards through the Rocky Mountains almost to New Mexico, USA (Potter et al. 2013). Our phylogenies placed sco1 from South Dakota, USA in N5 and P6; sco2 from Black Hills, South Dakota, USA in N6 and P6, and bra01 from Bighorn Mountains, Wyoming, USA in N7 and P7 (Figs. 5B, 6A). Because different sampling schemes and different types of data have yielded different affinities, this geographic region is likely not homogeneous and will also benefit from additional sampling that includes Mexican taxa. We did not sample the island populations of *P. torreyana* var. *insularis* J.R.Haller, an intraspecific taxon that is distinctive and has current taxonomic acceptance (Whittall et al. 2010).

Systematics—While the results presented here did not clarify many of the species-level questions that have been asked about these taxa, we were able to support two important findings: a revised *P.* subsect. *Sabinianae* that includes *P. jeffreyi* and the split of *P. ponderosa* s. l. into three species that do not form sister clades on phylogenies. These topics are discussed in turn below.

SUBSECTIONS—When our trees were rooted on the branch leading to *P.* subsects. *Contortae* and *Australes*, as supported by previous phylogenies, we consistently recovered node A, the most recent common ancestor of *P.* subsect. *Ponderosae*, and an amended *P.* subsect. *Sabinianae* that included *P. jeffreyi*. We confirmed node A by rooting with just the branch leading to *P.* subsect. *Contortae*, by rooting with the branch leading to *P.* subsect. *Australes*, and by rooting with samples from *P.* subg. *Pinus* (see Methods). The only exception to the reciprocal monophyly at node A was one ancient reticulation (plastome P1; Figs. 1, 3). Despite its superficial vegetative similarity to the partially sympatric *P. ponderosa*, there are morphological characters as well as molecular phylogenetic evidence that clearly unite *P. jeffreyi* with *P.* subsect. *Sabinianae*. There are four synapomorphies for a four-taxon *P.* subsect. *Sabinianae*: cones greater than 13 cm long (Kral 1993; Baldwin et al. 2012),

a seed-wing-length to seed-length ratio that is less than two (Johnson et al. 2003; Salazar-Tortosa et al. 2019), a seed weight greater than 0.14 g (Johnson et al. 2003), and heptane as a component of xylem and leaf monoterpenes (Smith 1982; Adams and Wright 2012). *Pinus jeffreyi* lacks the commonly observed California big cone feature of an irregular trunk with spreading branches, its ovulate cones are not as heavy, and its cone scale apices are not reflexed. But, *P. jeffreyi* has a heavy seed with a relatively short wing that animals gather directly from the cone or after the seed has fallen a short distance from the tree. The hoarding of *P. jeffreyi* seeds by rodents and jays is vital for Jeffrey pine regeneration (Vander Wall 1992, 1993, 2002; Briggs et al. 2009; Vander Wall et al. 2012) in the same way that *P. coulteri*, *P. sabiniana*, and *P. torreyana* rely on animals rather than wind for effective regeneration (Borchert et al. 2003; Johnson et al. 2003). *Pinus ponderosa* seeds are also hoarded by rodents (Vander Wall 2003) and cached by pinyon jays and Clark's nutcrackers (Balda and Bateman 1971; Lorenz and Sullivan 2009) and birds are the presumed long-distance dispersers that have facilitated the establishment of new populations of *P. ponderosa* (Parchman and Benkman 2007; Lesser and Jackson 2013). Although secondary scatter-hoarding of wind-dispersed seed by rodents may also make a substantial contribution to the germination of *P. ponderosa* seeds in some cases (Keyes et al. 2007), wind is still a viable dispersal mechanism for their relatively long-winged seeds. In contrast, wind is likely ineffective at dispersing *P. jeffreyi* and other *P.* subsect. *Sabinianae* seeds which are over 0.1 g (Benkman 1995). Excluding the distantly related Mediterranean *P. pinea*, all other species in subg. *Pinus* are primarily wind-dispersed (Tomback and Linhart 1990). This suggests that the evolution of dependence on animal dispersal may have arisen in *P.* subsect. *Sabinianae*, making large seeds, a feature of increased reliance on animal dispersal, an important synapomorphy for a four-species *P.* subsect. *Sabinianae*. This story is complicated by the different animals that compete for seed in different habitats (Siepielski and Benkman 2010). Our finding raises questions regarding the selective advantage for the ancestor of the currently higher-elevation *P. jeffreyi* (as well as lower elevation *P. coulteri*, *P. sabiniana*, and *P. torreyana*) to rely more heavily on animal rather than wind dispersal of seeds. Another question is whether *P. jeffreyi* converged to look superficially more like *P. ponderosa* in some cone and vegetative characters as it adapted to higher elevations or ultramafic soils (Furnier and Adams 1986). Alternatively, *P. jeffreyi* may have retained some plesiomorphic characters it shared with the ancestors of *P.* subsect. *Ponderosae*, while *P. coulteri*, *P. sabiniana*, and *P. torreyana* developed heavier cones with larger outcurved prickles and other vegetative characters that are useful in frequent-fire foothills (Schwilk and Keeley 2006) or on maritime islands (Haller 1986) with increasing reliance on the caching behavior of different rodent species for regeneration (Johnson et al. 2003). The perspective provided by our phylogenies also implies that the occasional hybrids between *P. jeffreyi* and *P. benthamiana* span a broader evolutionary distance than we previously understood and are likely to be much less common than we supposed (Conkle and Critchfield 1988).

Pinus subsection *Sabinianae* was originally proposed to distinguish *P. sabiniana*, *P. coulteri*, and *P. roxburghii* based on their dramatically larger cones that have elongated and hooked cone scale apices (Loudon 1838). Early molecular phylogenies clarified that *P. roxburghii* belongs to *P.* sect. *Pinus* rather than *P.* sect. *Trifoliae*, indicating that the large

cones of *P. roxburghii* are not due to shared ancestry (Germandt et al. 2005; Willyard et al. 2007). The publication of the narrowly endemic *P. torreyana* Parry ex Carrière 1855 added a third species of California big cone pine, and Torrey pine was included in a three-species California big cone pine group that Shaw (1914) called *Macrocarpae*. Although our recommendation of amending Loudon's subsectional description will leave the shared ancestor of *P.* subsect. *Ponderosae* and *P.* subsect. *Sabinianae* an unnamed clade, it would be inaccurate and misleading to leave *P. jeffreyi* in a paraphyletic *P.* subsect. *Ponderosae*.

PINUS PONDEROSA—Within *P.* subsect. *Ponderosae*, we observed a consistent divergence between two clades in both nuclear and plastome phylogenies, on ASTRAL species trees built with prior taxonomic assignments, and on ASTRAL species trees built with clade assignments (node B in all trees). Most of the samples of bpw (*P. benthamiana*, *P. ponderosa*, and *P. washoensis*) resolved on one branch. The other branch contained samples of all of the remaining taxa, including those currently accepted as *P. ponderosa* var. *scopulorum* and 11 more accepted taxa (Table 1). The intriguing reticulations that we observed in the placement of nuclear N4 (but not their plastomes) and plastome P1 (but not their nuclear genes) were discussed above. These two limited reticulations do not conflict with a strong conclusion that *P. ponderosa* s. l. is paraphyletic. The continued treatment of *P. ponderosa* s. l. as one species that includes *P. ponderosa* var. *scopulorum* would contradict a lineage concept of species (De Queiroz 2007).

Our results for resolution within the bpw group supported the morphological and ecological traits that have been suggested to diagnose three varieties within *P. ponderosa* s. s. (Haller 1961; Haller and Vivrette 2011). As discussed above under Sympatry, we think that there is evidence for mixed populations of *P. ponderosa* and *P. washoensis* (including near the type locality of *P. ponderosa*) and that some level of introgression, at least in some of these locations, is likely. As discussed above, we did not find molecular phylogenetic evidence to support resurrecting *P. washoensis* by overturning the combination *P. ponderosa* var. *washoensis* (Haller and Vivrette 2011) as it is currently treated (Baldwin et al. 2012). Although "Pacific variety" has been frequently used as a common name, the publication of *P. ponderosa* var. *benthamiana* (Hartweg) Vasey in 1876 takes precedence over the name *P. ponderosa* var. *pacifica* Haller 2011. The rejection of the original name based on an argument that in 1847 Hartweg collected the type specimen of *P. benthamiana* from a localized population of ponderosa pines on isolated pockets of sandhill soil among the redwood forest in the Santa Cruz Mountains does not hold up in two ways. First, Hartweg's journal noted that he had been to the redwood forests near Santa Cruz before, and that while his ship was docked at Santa Cruz harbor in 1847, he "went a different direction" (Lauria 1996b). Hartweg's journal does not mention the redwoods as an associated species with the ponderosa pine collection and there are other ponderosa pine stands that he might have accessed, for example those near oak woodlands in present-day Henry Coe State Park. Hartweg also noted that in June 1847, he saw *P. benthamiana* again in the Sierra Nevada foothills. The second argument is that there is no published genetic or morphological evidence that the individuals from the sandhill soil form a unique entity, including in this study where ben16, ben17, and ben18 were not resolved together. Thus, Hartweg's 1847 collection and the publication of *P.*

benthamiana take precedence over the "Pacific variety." Although our species trees supported *P. benthamiana* as a separate entity (BS = 0.93 in nuclear and BS = 0.54 in plastome) as sister to a clade of *P. ponderosa* plus *P. washoensis* (Fig. 2), our *P. benthamiana* and *P. ponderosa* samples were not resolved as mutually monophyletic (Fig. 1). What's more, the morphological characters available now to diagnose these taxa are mostly overlapping and geographic/ecological distributions are also unclear (Haller and Vivrette 2011). It may be that future evidence will support resurrecting *P. benthamiana* and/or *P. washoensis* as separate species, but based on the evidence at hand, we recommend accepting *P. ponderosa* var. *benthamiana* Vasey and *P. ponderosa* var. *washoensis* Haller & Vivrette.

There remains the challenge of choosing appropriate names for the *P. ponderosa* s. l. samples that are within node C (Figs. 1, 2). The *scopulorum* epithet (published in 1880) has been widely used for the Rocky Mountain ponderosa pines with mostly two needles per fascicle. Descriptions of this taxon were confused by the expansion of the ranges of characters (especially by the inclusion of three longer and more slender needles per fascicle) to include the southwestern USA populations (Kral 1993). The name *P. brachyptera* Engelmann was validly published in 1848 with a type locality in northern New Mexico, USA, and morphological characters that tend to fit these southwestern USA populations. Because of our conclusion that some of these ponderosa pine populations are more closely related to Mexican taxa with early publication dates (Table 1) than they are to each other, our preliminary recommendation is to assign *P. ponderosa* var. *scopulorum* individuals to either *P. scopulorum* or *P. brachyptera*. We were unable to reconcile putative diagnostic characters (e.g., needles per fascicle, needle length, cone size, number of stomata, or number of resin ducts) with any molecular clades. Nevertheless, because these taxa are clearly not conspecific with *P. ponderosa*, the names *P. scopulorum* and *P. brachyptera* are preferred over varietal names.

Similarly, our results for *P. arizonica*, *P. engelmannii*, and *P. arizonica* var. *stormiae* show that the species-level taxonomic assignments that we used for populations in the USA would render these taxa paraphyletic with their supposedly conspecific samples from Mexico (Fig. 1). We predict that individual samples from these populations in the USA along with a thorough sampling of other *Ponderosae* in Mexico with nuclear, plastome, and morphological data on the same individuals will be needed to address these taxa.

Conclusions—*Pinus jeffreyi* belongs with *P.* subsect. *Sabinianae* based on morphological synapomorphies as well as strong molecular phylogenetic support from both nuclear and plastome trees. We suggest that larger seeds in a four-taxon *P.* subsect. *Sabinianae* were an important innovation that facilitated their present-day reliance on animals for effective seed dispersal. *Pinus ponderosa* s. l. is paraphyletic, as is *P. ponderosa* var. *scopulorum*. We suggest taxonomic treatment as *P. ponderosa* s. s. (with *P. ponderosa* varieties *benthamiana*, *washoensis*, and typical), *P. scopulorum*, and *P. brachyptera*. The persistence of lineages with the footprints of ancient nuclear introgression (the bpw in N4) and chloroplast capture (the bpw in P1) should caution species identification in *P.* subsect. *Ponderosae* based on limited molecular data. Based on only two supported incongruences that give evidence of recent reticulate evolution, we found that hybrid frequency was low. Instead of previous conclusions of hybrid origin, we

showed that the persistence of an ancient P1 plastid clade is a better explanation than hybridization or chloroplast capture between *P. ponderosa* and *P. jeffreyi* for unexpected plastid associations in the western Sierra Nevada, USA. A new potential zone of ancient admixture between *P. ponderosa* and *P. scopulorum* was identified in Idaho, USA, where there is no present-day sympatry. Some populations of *P. arizonica*, *P. brachyptera*, *P. engelmannii*, and *P. scopulorum* in the USA are more closely related to taxa only distributed in Mexico than they are to each other. Even our limited within-population sampling suggested that several populations contain individuals with different ancestries that represent currently unrecognized sympatry. These potentially include *P. ponderosa* with *P. ponderosa* var. *wasoensis* and several taxa mixed in the Chisos Mountains, Texas, USA, in the Spring Mountains, Nevada, USA, and in the Sky Islands, Arizona, USA. To integrate phylogeny and taxonomy, future work should sample widely in Mexico and the USA, score morphological characters (including seedling characters from the known seed parent), on the same individual that is used for molecular data, and choose nuclear data that are useful based on individual samples rather than frequencies in populations.

ACKNOWLEDGMENTS

This experiment was not preregistered with an independent, institutional registry. We thank the editors and two anonymous reviewers for thoughtful comments on an earlier version of the manuscript and we thank Diana Ramos Dorantes, Alejandro López, and Stephen Meyers for contributing samples and Adam Bigott, Jack Finney, Payton Lea, and Kevin Spatz for field and lab assistance. We thank Bruce Baldwin, Arnaldo Ferreira, Kevin Potter, Stephen Vander Wall, and Keith Woeste for comments on an earlier version of the manuscript, David Charlet and George Ferguson for helpful advice, Steve Karafit and Matthew Moran for help in the field, and Gordon Burleigh for sharing an unpublished phylogeny. Land-owners and managers who facilitated our collections in the USA: Big Bend National Park, California State Parks, Fred Lawrence Whipple Observatory, Guadalupe Mountains National Park, Hendrix College, Joint Base Lewis-McChord, Lander-Hills Big Creek Reserve, Mohave County AZ Parks, Navajo Nation, North Sierra Tree Improvement, San Jacinto Mountains Reserve, Santa Cruz County CA Parks, Sierra Pacific Industries, and University of California Santa Cruz Arboretum, USDA Forest Service (Bighorn, Black Hills, Boise, Cleveland, Colville, Coronado, Custer, Deschutes, Eldorado, Gila, Humboldt-Toiyabe, Inyo, Kaibab, Klamath, Los Padres, Medicine Bow, Mendocino, Modoc, Nez Perce, Ochoco, Payette, San Bernardino, San Juan, Sequoia, Shasta-Trinity, Sierra, Six Rivers, Uinta, Umpqua, and Wasatch National Forests). Collections in Mexico were conducted under permits SGPA/DGGFS/712/3716/12, and SGPA/DGGFS/712/1768/16 emitted by the Subsecretaría de Gestión Para la Protección Ambiental. Funding was provided by the California Native Plant Society, Hendrix College Odyssey Grants, and Arkansas Student Undergraduate Research Fellowships.

AUTHOR CONTRIBUTIONS

AW and DSG designed and conducted the experiment, analyzed data, and wrote the manuscript; BC, CD, KF, HK, JL, and DP contributed undergraduate research funds, did field and lab work, and contributed to the manuscript; EL, SKL, PM, and FT contributed samples and contributed to the manuscript.

LITERATURE CITED

- Adams, R. P., and G. F. J. Edmunds. 1989. A re-examination of the volatile leaf oils of *Pinus ponderosa* Dougl. ex P. Lawson using ion trap mass spectroscopy. *Flavour and Fragrance Journal* 4: 19–23.
- Adams, R. P. and J. W. Wright. 2012. Alkanes and terpenes in wood and leaves of *Pinus jeffreyi* and *P. sabiniana*. *The Journal of Essential Oil Research* 24: 435–440.
- Adams, R. P., G. M. Ferguson, and D. Thornburg. 2015. First comprehensive report on the composition of the leaf volatile terpenoids of *Pinus arizonica* Engelm. and *P. ponderosa* var. *brachyptera* (Engelm.) Lemmon. *Phytologia* 97: 45–50.
- Aguirre-Dugua, X. and D. S. Gernandt. 2017. Complete plastomes of three endemic Mexican pine species (*Pinus* subsection *Australes*). *Mitochondrial DNA Part B: Resources* 2: 562–565.
- Alamouti, S. M., V. Wang, S. DiGuistini, D. L. Six, J. Bohlmann, R. C. Hamelin, N. Feau, and C. Breuil. 2011. Gene genealogies reveal cryptic species and host preferences for the pine fungal pathogen *Grossmannia clavigera*. *Molecular Ecology* 20: 2581–2602.
- Alex, B. 2004. Mapping and modeling sensitive plants in Big Bend National Park. Pp. 101–116 in *Proceedings of the Sixth Symposium on the Natural Resources of the Chihuahuan Desert Region*, eds. C. A. Hoyt and J. Karges. Fort Davis, Texas, USA: Chihuahuan Desert Research Institute.
- Anderson, R. S. 1989. Development of the southwestern ponderosa pine forests: What do we really know? Pp 15–22 in *Multiresource Management of Ponderosa Pine Forests. General Technical Report RM-185*. USDA Forest Service: Rocky Mount Forest and Range Experiment Station.
- Arealo, A. G. and M. S. Gonzalez Elizondo. 1998. *Pinaceae de Durango*. Veracruz, Mexico: Instituto de Ecología, A.C.
- Ávila-Flores, I. J., J. Hernández-Díaz, M. S. González-Elizondo, J. Á. Prieto-Ruiz, and C. Wehenkel. 2016. Degree of hybridization in seed stands of *Pinus engelmannii* Carr. in the Sierra Madre Occidental, Durango, Mexico. *PLOS One* 11: 1–13.
- Balda, R. P. and G. C. Bateman. 1971. Flocking and annual cycle of the piñon jay, *Gymnorhinus cyanocephalus*. *The Condor* 73: 287–302.
- Baldwin, B. G., D. H. Goldman, D. J. Keil, R. Patterson, T. J. Rosatti, and D. H. Wilken. 2012. *The Jepson Manual*, ed. 2. Berkeley: University of California Press.
- Bankevich, A., S. Nurk, D. Antipov, A. A. Gurevich, M. Dvorkin, A. S. Kulikov, V. M. Lesin, S. P. Nikolenko, A. D. Pribelski, A. V. Pyshkin, A. V. Sirotkin, N. Vyahhi, G. Tesler, M. A. Alekseyev, and P. Pevzner. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19: 455–477.
- Baum, D. A. and M. J. Donoghue. 1995. Choosing among alternative "Phylogenetic" species concepts. *Systematic Botany* 20: 560–573.
- Benkman, C. 1995. Wind dispersal capacity of pine seeds and the evolution of different seed dispersal modes in pines. *Oikos* 73: 221–224.
- Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Borchert, M., M. Johnson, D. S. Schreiner, and S. B. Vander Wall. 2003. Early postfire seed dispersal, seedling establishment and seedling mortality of *Pinus coulteri* (D. Don) in central coastal California, USA. *Plant Ecology* 168: 207–220.
- 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.
- Briggs, J. S., S. B. Vander Wall, and S. H. Jenkins. 2009. Forest rodents provide directed dispersal of Jeffrey pine seeds. *Ecology* 90: 675–687.
- Bruczyk, J., S. P. DeFazio, and W. T. Adams. 2004. Gene flow in forest trees: How far do genes really travel? *Forest Genetics* 11: 179–192.
- Buck, J. M., R. S. Adams, J. Cone, M. T. Conkle, W. J. Libby, C. J. Eden, and M. J. Knight. 1970. *California Tree Seed Zones*. San Francisco: USDA Forest Service.
- Caballero Deloya, M. 1967. Estudio comparativo de dos especies de pinos Mexicanos (*Pinus pseudostrobus* y *Pinus montezumae* Lamb.) con base en característica de plántula y semilla. *Instituto Nacional de Investigaciones Forestales, Mexico*: 1–40.
- Callahan, R. Z. 2013a. *Pinus ponderosa*: Geographic races and subspecies based on morphological variation. *USDA Forest Service, Pacific Southwest Research Station PSW-RP-265*: 1–54.
- Callahan, R. Z. 2013b. *Pinus ponderosa*: A taxonomic review with five subspecies in the United States. *USDA Forest Service, Pacific Southwest Research Station PSW-RP-264*: 1–53.
- Callahan, R. Z. and A. R. Liddicoet. 1961. Altitudinal variation at 20 years in ponderosa and Jeffrey pines. *Journal of Forestry* 59: 814–820.
- Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17: 540–552.

- Conkle, M. and W. B. Critchfield. 1988. Genetic variation and hybridization of ponderosa pine. Pp. 27–44 in *Ponderosa Pine: The Species and its Management*, eds. D. Baumgartner and J. Lotan. Pullman, Washington: Washington State University.
- Clegg, B. M. 1994. Carbon allocation, gas exchange, and needle morphology of *Pinus ponderosa* genotypes known to differ in growth and survival under imposed drought. *Tree Physiology* 14: 883–898.
- Critchfield, W. B. 1984. Crossability and relationships of Washoe pine. *Madroño* 31: 144–170.
- Darriba, D., D. Posada, A. M. Kozlov, A. Stamatakis, B. Morel, and T. Flouri. 2019. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution* 37: 291–294.
- Dayrat, B. 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85: 407–415.
- De Queiroz, K. 2007. Species concepts and species delimitation. *Systematic Biology* 56: 879–886.
- Debreczy, Z., I. Rácz, and K. Musial. 2011. *Conifers Around the World*. Wellesley, Massachusetts: International Dendrological Research Institute.
- Delgado, P., R. Salas-Lizana, A. Vázquez-Lobo, A. Wegier, M. Anzidei, E. R. Alvarez-Buylla, G. G. Vendramin, and D. Piñero. 2007. Introgressive hybridization in *Pinus montezumae* Lamb and *Pinus pseudostrobus* Lindl. (Pinaceae): Morphological and molecular (cpSSR) evidence. *International Journal of Plant Science* 168: 861–875.
- Dixit, A. and T. Kolb. 2020. Variation in seedling budburst phenology and structural traits among southwestern ponderosa pine provenances. *Canadian Journal of Forest Research* 999: 1–8.
- Dyer, R. and V. Sork. 2001. Pollen pool heterogeneity in shortleaf pine, *Pinus echinata* Mill. *Molecular Ecology* 10: 859–866.
- Eckenwalder, J. E. 2009. *Conifers of the World*. Portland, Oregon: Timber Press, Inc.
- 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.
- Espinoza-Pelcastre, C., S. Hernández-León, D. S. Gernandt, O. Arce-Cervantes, R. Rodríguez-Laguna, and J. González-Ávalos. 2018. Taxonomic identification key with leaf anatomical characters for *Pinus* L. species in Hidalgo. *Revista Mexicana de Ciencias Forestales* 9: 1–19.
- Farjon, A. and B. Styles. 1997. *Flora Neotropica, Monograph 75, Pinus (Pinaceae)*. New York: The New York Botanical Garden.
- Fujisawa, T., A. Aswad, and T. G. Barraclough. 2016. A rapid and scalable method for multilocus species delimitation using Bayesian model comparison and rooted triplets. *Systematic Biology* 65: 759–771.
- Fujita, M. K., A. D. Leaché, F. T. Burbrink, J. A. McGuire, and C. Moritz. 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology and Evolution* 27: 480–488.
- 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., G. Gaeda López, S. O. García, and A. Liston. 2005. Phylogeny and classification of *Pinus*. *Taxon* 54: 29–42.
- Gernandt, D. S., S. Magallón, G. Gaeda López, O. Zerón Flores, A. Willyard, and A. Liston. 2008. Use of simultaneous analyses to guide fossil-based calibrations of Pinaceae phylogeny. *International Journal of Plant Science* 169: 1086–1099.
- Gernandt, D. S., S. Hernández-León, E. Salgado-Hernández, J. A. Pérez de la Rosa, and L. Jardón-Barbolla. 2009. Phylogenetic relationships of *Pinus* subsection *Ponderosae* inferred from rapidly evolving cpDNA regions. *Systematic Botany* 34: 481–491.
- Gernandt, D. S., X. Aguirre Dugua, A. Vázquez-Lobo, A. Willyard, A. Moreno Letelier, J. Pérez de la Rosa, D. Piñero, and A. Liston. 2018. Multi-locus phylogenetics, lineage sorting, and reticulation in *Pinus* subsection *Austroales*. *American Journal of Botany* 105: 1–15.
- Gerson, E. A., R. G. Kelsey, and J. B. St Clair. 2009. Genetic variation of piperidine alkaloids in *Pinus ponderosa*: a common garden study. *Annals of Botany* 103: 447–457.
- Gonçalves, D. J. P., B. B. Simpson, E. M. Ortiz, G. H. Shimizu, and R. K. Jansen. 2019. Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. *Molecular Phylogenetics and Evolution* 138: 219–232.
- Griffin, J. R. 1964. Isolated *Pinus ponderosa* forests on sandy soils near Santa Cruz, California. *Ecology* 45: 410–412.
- Haller, J. R. 1961. Some recent observations on ponderosa, Jeffrey, and Washoe pines in Northeastern California. *Madroño* 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. 1965. The role of 2-needle fascicles in the adaptation and evolution of ponderosa pine. *Brittonia* 17: 354–382.
- Haller, J. R. 1986. Taxonomy and relationships of the mainland and island populations of *Pinus torreyana* (Pinaceae). *Systematic Botany* 1: 39–50.
- Haller, J. R. and N. J. Vivrette. 2011. Ponderosa pine revisited. *Aliso* 29: 53–57.
- Hennig, W. 1966. *Phylogenetic Systematics*. Urbana: University of Illinois Press.
- Hernández-León, S., D. S. Gernandt, J. A. Pérez de la Rosa, and L. Jardón-Barbolla. 2013. Phylogenetic relationships and species delimitation in *Pinus* section *Trifoliae* inferred from plastid DNA. *PLOS One* 8: e70501.
- Hibbs, D. E., M. V. Wilson, and A. L. Bower. 2002. Ponderosa pine of the Willamette Valley, western Oregon. *Northwest Science* 76: 80–84.
- Hipkins, V. and K. Krutovsky. 2005. Ponderosa pine (*Pinus ponderosa*) source identification using molecular genetics. Pp. 4–71 in *USDA Forest Service National Forest Genetics Laboratory (NFGEL) Annual Report*. Placerville, California: USDA Forest Service.
- Huson, D. H. and C. Scornavacca. 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees and networks. *Systematic Biology* 61: 1061–1067.
- Jackman, S. D., B. P. Vandervalk, H. Mohamadi, J. Chu, S. Yeo, S. A. Hammond, G. Jahesh, L. Coombe, R. L. Warren, and I. Birol. 2017. ABYSS 2.0: Resource-efficient assembly of large genomes using a Bloom filter. *Genome Research* 27: 768–777.
- Johansen, A. D. and R. G. Latta. 2003. Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Molecular Ecology* 12: 293–298.
- Johnson, M., S. B. Vander Wall, and M. Borchert. 2003. A comparative analysis of seed and cone characteristics and seed-dispersal strategies of three pines in the subsection *Sabinianae*. *Plant Ecology* 168: 69–84.
- Johnson, M. G., E. M. Gardner, Y. Liu, R. Medina, B. Goffinet, A. J. Shaw, N. J. C. Zerega, and N. J. Wickett. 2016. HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1–7.
- Katoh, K., K.-i. Kuma, H. Toh, and T. Miyata. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research* 33: 511–518.
- Kearney, T. H. and R. H. Peebles. 1960. *The Flora of Arizona, With Supplement*. Berkeley: University of California Press.
- Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Keefe-Ring, K. and Y. B. Linhart. 2010. Variable chemistry and herbivory of ponderosa pine cones. *International Journal of Plant Science* 171: 293–302.
- Kerr, K. L., F. C. Meinzer, K. A. McCulloh, D. R. Woodruff, and D. E. Marias. 2015. Expression of functional traits during seedling establishment in two populations of *Pinus ponderosa* from contrasting climates. *Tree Physiology* 35: 535–548.
- Keyes, C. R., D. A. Maquire, and J. C. Tappeiner. 2007. Observed dynamics of ponderosa pine (*Pinus ponderosa* var. *ponderosa* Dougl. ex Laws.) seedling recruitment in the Cascade Range, USA. *New Forests* 34: 95–107.
- Kitzmler, J. H. 1990. Managing genetic diversity in a tree improvement program. *Forest Ecology and Management* 35: 131–149.
- Kral, R. 1993. *Pinus*. Pp. 373–398 in *Flora of North America*, vol. 2. New York: Oxford University Press.
- Krupkin, A. B., A. Liston, and S. H. Strauss. 1996. Phylogenetic analysis of the hard pines (*Pinus* subgenus *Pinus*, Pinaceae) from chloroplast DNA restriction site analysis. *American Journal of Botany* 83: 489–498.
- 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.
- Latta, R. G., Y. B. Linhart, D. Fleck, and M. Elliot. 1998. Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine. *Evolution* 52: 61–67.
- Lauria, F. 1996a. The identity of *Pinus ponderosa* Douglas ex. C. Lawson (Pinaceae). *Linzer biologische Beitrag* 28: 999–1052.

- Lauria, F. 1996b. Typification of *Pinus benthamiana* Hartw. (Pinaceae), a taxon deserving renewed botanical examination. *Annalen des Naturhistorischen Museums in Wien* 98 B Suppl.: 427–446.
- Lesser, M. R. and S. T. Jackson. 2013. Contributions of long-distance dispersal to population growth in colonising *Pinus ponderosa* populations. *Ecology Letters* 16: 380–389.
- Lesser, M. R., T. L. Parchman, and S. T. Jackson. 2013. Development of genetic diversity, differentiation and structure over 500 years in four ponderosa pine populations. *Molecular Ecology* 22: 2640–2652.
- Letunic, I. and P. Bork. 2019. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Research* 47: W256–W259.
- Linhardt, Y. B., J. B. Mitton, K. B. Sturgeon, and M. L. Davis. 1981. Genetic variation in space and time in a population of ponderosa pine. *Heredity* 46: 407–426.
- 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.
- Little, E. L. Jr. 1971. *Atlas of United States Trees*, vol. 1. Washington, DC: USDA Forest Service.
- Little, E. L. Jr. and W. B. Critchfield. 1969. Subdivisions of the Genus *Pinus* (Pines), Misc. Pub. 1144. Washington, DC: USDA Forest Service.
- López-Reyes, A., J. Pérez 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.
- Lorenz, T. J., and K. A. Sullivan. 2009. Seasonal differences in space use by Clark's nutcrackers in the Cascade Range. *The Condor* 111: 326–340.
- Loudon, J. 1838. *Arboretum et Fruticetum Britannicum*. London: James Ridgeway and Sons.
- Maguire, K. C., D. J. Shinneman, K. M. Potter, and V. D. Hipkins. 2018. Intraspecific niche models for ponderosa pine (*Pinus ponderosa*) suggest potential variability in population-level response to climate change. *Systematic Botany* 67: 965–978.
- Maherali, H., B. L. Williams, K. N. Paige, and E. H. Delucia. 2002. Hydraulic differentiation of ponderosa pine populations along a climate gradient is not associated with ecotypic divergence. *Functional Ecology* 16: 510–521.
- Marquardt, P. E., B. R. Miranda, S. Jennings, G. Thurston, and F. W. Telewski. 2018. Variable climate response differentiates the growth of Sky Island ponderosa pines. *Trees* 33: 317–332.
- Marquardt, P. E., B. R. Miranda, and F. W. Telewski. 2019. Shifts in climate-growth relationships of sky island pines. *Forests* 10: 1011.
- Martínez, M. 1948. *Los Pinos Mexicanos*. Mexico City: Ediciones Botas.
- Martínez, M. 1963. Las pináceas mexicanas. *Anales del Instituto de Biología Universidad Nacional de México* 16: 284–286.
- Matos, J. A. 1995. *Pinus hartwegii* and *P. rudis*: A critical assessment. *Systematic Botany* 20: 6–21.
- Matos, J. A. and B. A. Schaal. 2000. Chloroplast evolution in the *Pinus montezumae* complex: A coalescent approach to hybridization. *Evolution* 54: 1218–1233.
- Meyers, S. C., T. Jaster, K. E. Mitchell, and L. K. Hardison. 2015. *Flora of Oregon*. Fort Worth, Texas: The Botanical Research Institute of Texas.
- Millar, C. I. 1993. Impact of the Eocene on the evolution of *Pinus* L. *Annals of the Missouri Botanical Garden* 80: 471–498.
- Nagel, L. M., B. J. Palik, M. A. Battaglia, A. W. D'Amato, J. M. Guldin, C. W. Swanston, M. K. Janowiak, M. P. Powers, L. A. Joyce, C. I. Millar, D. L. Peterson, L. M. Ganio, C. Kirschbaum, and M. R. Roske. 2017. Adaptive silviculture for climate change: a national experiment in manager-scientist partnerships to apply an adaptation framework. *Journal of Forestry* 115: 167–178.
- Neale, D. and R. Sederoff. 1989. Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theoretical and Applied Genetics* 77: 212–216.
- Neves, L. G., J. M. Davis, W. B. Barbazuk, and M. Kirst. 2013. Whole-exome targeted sequencing of the uncharacterized pine genome. *The Plant Journal* 75: 146–156.
- 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.
- Nock, C. A., R. J. Vogt, and B. E. Beisner. 2016. Functional traits. In *eLS*. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0026282.
- Padial, J. M., A. M. Miralles, I. De la Riva, and M. Vences. 2010. The integrative future of taxonomy. *Frontiers in Zoology* 7: 1–14.
- Parchman, T. L. and C. Benkman. 2007. The geographic selection mosaic for ponderosa pine and crossbills: A tale of two squirrels. *Evolution* 62: 348–360.
- 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.
- Pelouquin, R. L. 1971. *Variation and Hybridization Patterns in Pinus ponderosa and Pinus engelmannii*. Ph.D. dissertation. Santa Barbara: University of California.
- Pelouquin, R. L. 1984. The identification of three-species hybrids in the ponderosa pine complex. *The Southwestern Naturalist* 29: 115–122.
- 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: 1–23.
- Price, R. A., A. Liston, and S. H. Strauss. 1998. Phylogeny and systematics of *Pinus*. Pp 49–68 in *Ecology and Biogeography of Pinus*, ed. D. M. Richardson. Cambridge, UK: Cambridge University Press.
- Read, R. 1980. *Genetic Variation in Seedling Progeny of Ponderosa Pine Provenances*. Forest Science Monograph 23. Washington, DC: Society of American Foresters.
- Reeves, P. A. and C. M. Richards. 2011. Species delimitation under the general lineage concept: An empirical example using wild North American hops (Cannabaceae: *Humulus lupulus*). *Systematic Biology* 60: 45–59.
- Rehfeldt, G. E. 1993. Genetic variation in the *Ponderosae* of the Southwest. *American Journal of Botany* 80: 330–343.
- 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.
- Rehfeldt, G. E., B. C. Wilson, S. P. Wells, and R. M. Jeffers. 1996. Phyto-geographic, taxonomic, and genetic implications of phenotypic variation in the *Ponderosae* of the Southwest. *The Southwestern Naturalist* 41: 409–418.
- Rehfeldt, G. E., B. C. Jaquish, J. Lopez-Upton, C. Saenz-Romero, J. B. St Clair, L. P. Leites, and D. G. Joyce. 2014. Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotsuga menziesii*: Realized climate niches. *Forest Ecology and Management* 324: 126–137.
- Roberts, D. R. and A. Hamann. 2015. Glacial refugia and modern genetic diversity of 22 western North American tree species. *Proceedings of the Royal Society B* 282: 1–9.
- Salazar-Tortosa, D., J. Castro, B. Saladin, N. E. Zimmermann, and R. R. De Casas. 2019. Arid environments select for larger seeds in pines (*Pinus* spp.). *Evolutionary Ecology* 34: 11–26.
- Schwilld, D. W. and J. E. Keeley. 2006. The role of fire refugia in the distribution of *Pinus sabiniana* (Pinaceae) in the southern Sierra Nevada. *Madroño* 53: 364–372.
- Shaw, G. R. 1914. *The Genus Pinus*. Boston: Publications of the Arnold Arboretum.
- Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* 51: 492–508.
- Shimodaira, H. and M. Hasegawa. 2001. CONSEL: For assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17: 1246–1247.
- 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: e0151811.
- Siepielski, A. M. and C. W. Benkman. 2010. Conflicting selection from an antagonist and a mutualist enhances phenotypic variation in a plant. *Evolution* 64: 1120–1128.
- Silba, J. 1990. A supplement to the international census of the Coniferae, II. *Phytologia* 68: 7–78.
- Smith, R. H. 1977. *Monoterpenes of Ponderosa Pine Xylem Resin in Western United States*. USDA Forest Service Technical Bulletin 1532. USDA Forest Service.
- 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. 1982. Xylem monoterpenes of some hard pines of western North America: Three studies. *USDA Forest Service Pacific Southwest Forest and Range Experiment Station PSW-160*: 1–12.
- Solís-Lemus, C., M. Yang, and C. Ané. 2016. Inconsistency of species tree methods under gene flow. *Systematic Biology* 65: 843–851.
- Sorensen, F. C., N. L. Mandel, and J. E. Aagaard. 2001. Role of selection versus historical isolation in racial differentiation of ponderosa pine in southern Oregon: an investigation of alternative hypotheses. *Canadian Journal of Forest Research* 31: 1127–1139.
- Stamatakis, A. 2014. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Stead, J. 1983. Studies in Central American pines V: A numerical study of variation in the *Pseudostrobus* group. *Silvae Genet* 32: 101–115.
- Sturgeon, K. B. 1979. Monoterpene variation in ponderosa pine xylem resin related to western pine beetle predation. *Evolution* 33: 803–814.
- Sukumaran, J. and L. L. Knowles. 2017. Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences, USA* 114: 1607–1612.
- Swenson, N. G. and D. J. Howard. 2005. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist* 166: 581–591.
- Swofford, D. 2003. PAUP* Phylogenetic analysis using parsimony (* and other methods), v. 4. Sunderland: Sinauer Associates.
- Syring, J., A. Willyard, R. Cronn, and A. Liston. 2005. Evolutionary relationships among pine (Pinaceae) subsections inferred from multiple low-copy nuclear loci. *American Journal of Botany* 92: 2086–2100.
- Syring, J., K. Farrell, R. Businsky, R. Cronn, and A. Liston. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. *Systematic Biology* 56: 163–181.
- Tombback, D. F. and Y. B. Linhart. 1990. The evolution of bird-dispersed pines. *Evolutionary Ecology* 4: 185–219.
- Vander Wall, S. B. 1992. The role of animals in dispersing a "wind-dispersed" pine. *Ecology* 73: 614–621.
- Vander Wall, S. B. 1993. Cache site selection by chipmunks (*Tamias* spp.) and its influence on the effectiveness of seed dispersal in Jeffrey pine (*Pinus jeffreyi*). *Oecologia* 96: 246–252.
- Vander Wall, S. B. 2002. Masting in animal-dispersed pines facilitates seed dispersal. *Ecology* 83: 3508–3516.
- Vander Wall, S. B. 2003. Effects of seed size of wind-dispersed pines (*Pinus*) on secondary seed dispersal and the caching behavior of rodents. *Oikos* 100: 25–34.
- Vander Wall, S. B., M. S. Enders, S. Barga, C. Moore, A. Seaman, and R. Perea. 2012. Jeffrey pine seed dispersal in the Sierra San Pedro Mártir, Baja California, Mexico. *Western North American Naturalist* 72: 534–542.
- Vickers, L. A., J. Houser, J. Rooni, and J. M. Guldin. 2019. Some lessons learned on early survival and growth of containerized, locally-sourced ponderosa pine seedlings in the Davis Mountains of western Texas, US. *Forests* 10: 1–16.
- Wang, X.-R., A. Szmidi, and O. Savolainen. 2001. Genetic composition and diploid hybrid speciation of a high mountain pine, *Pinus densata*, native to the Tibetan Plateau. *Genetics* 159: 337–346.
- Wauer, R. H. and C. M. Fleming. 2002. *Naturalist's Big Bend*. College Station: Texas A & M University Press.
- Weidman, R. H. 1939. Evidences of racial influence in a 25-year test of ponderosa pine. *Journal of Agricultural Research* 59: 855–887.
- Weitemier, K., S. C. K. Straub, R. C. Cronn, M. Fishbein, R. Schmickl, A. McDonnell, and A. Liston. 2014. Hyb-Seq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Science* 2: 1–7.
- 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. S., J.-H. Pak, R. S. Hill, and K. Kim. 2001. Cuticle micromorphology of leaves of *Pinus* (Pinaceae) from Mexico and Central America. *Botanical Journal of the Linnean Society* 135: 349–373.
- Whang, S. S., K. Kim, and R. S. Hill. 2004. Cuticle micromorphology of leaves of *Pinus* (Pinaceae) from North America. *Botanical Journal of the Linnean Society* 144: 303–320.
- Whittall, J., J. Syring, M. Parks, J. Buenrostro, C. Dick, A. Liston, and R. Cronn. 2010. Finding a (pine) needle in a haystack: Chloroplast genome sequence divergence in rare and widespread pines. *Molecular Ecology* 19: 100–114.
- Willyard, A., J. Syring, D. S. Gernandt, A. Liston, and R. Cronn. 2007. Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. *Molecular Biology and Evolution* 24: 90–101.
- 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.
- Willyard, A., D. S. Gernandt, K. Potter, V. Hipkins, P. Marquardt, M. F. Mahalovich, S. K. Langer, F. W. Telewski, B. Cooper, C. Douglas, K. Finch, H. H. Karemera, J. Lefler, P. Lea, and A. Wofford. 2017. *Pinus ponderosa*: A checkered past obscured four species. *American Journal of Botany* 104: 161–181.
- Willyard, A., D. S. Gernandt, B. Cooper, C. Douglas, K. Finch, H. Karemera, E. Lindberg, S. K. Langer, J. Lefler, P. Marquardt, D. L. Pouncey, and F. Telewski. 2021. Data from: Phylogenomics in the hard pines (*Pinus* subsection *Ponderosae*; Pinaceae) confirms paraphyly in *Pinus ponderosa*, and places *Pinus jeffreyi* with the California big cone pines. Dryad Digital Repository. <https://doi.org/10.5061/dryad.05qfttflq>
- Yeaton, R. I., R. W. Yeaton, and J. P. I. Waggoner. 1983. Changes in morphological characteristics of *Pinus engelmannii* over an elevational gradient in Durango, Mexico. *Madroño* 30: 168–175.
- Zhang, C., M. Rabiee, E. Sayyari, and S. Mirarab. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19 (Suppl 6): 15–30.
- 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.
- Zhang, J., P. Kapli, P. Pavlidis, and A. Stamatakis. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29: 2869–2876.
- Zhou, Y. F., R. J. Abbott, Z. Y. Jiang, F. K. Du, R. I. Milne, and J. Q. Liu. 2010. Gene flow and species delimitation: A case study of two pine species with overlapping species distributions in southeast China. *International Journal of Organic Evolution* 64: 2342–2352.
- Zobel, B. 1951. The natural hybrid between Coulter and Jeffrey pines. *Evolution* 5: 405–413.

APPENDIX 1. Samples in alphabetical order: Ingroup/Outgroup: *Species*: Sample, Voucher/lab name, Location, Latitude degrees, Longitude degrees, Plastome GenBank Accession.

Ingroup: *Pinus pseudostrabus* var. *apulcensis*: apu1, DSG889, Mexico: Chiapas: San Cristóbal de las Casas, 16.49, -92.30, MW315147; apu2, DSG901, Guatemala: San Jose Pinula, 14.54, -90.41, MW315148; *Pinus arizonica*: ari1, DSG874/AM19, USA: Arizona: Pinaleno: Mt Graham, 32.63, -109.82, MT071003; ari2, PALA22/PA44, USA: Arizona: Palisades White Tail Campground, 32.41, -110.71, MT071052; ari3, DSG783, USA: Arizona: Catalina: Hwy to Mt. Lemmon, 32.40, -110.69, MW315119; ari4, DSG784, USA: Arizona: Catalina Hwy to Mt. Lemmon, 32.40, -110.69, MW315120; ari5, AMW1080/CH10, USA: Arizona: Chiricahua, 31.95, -109.31, MT071015; *Pinus benthamiana*: ben01, AMW1099/F111, USA: Washington: Fort Lewis, 47.07, -122.48, MT071025; ben02, Meyers/VW10, USA: Oregon: Willamette Valley, 45.35, -122.77, MT071074; ben03, Meyers/VB26, USA: Oregon: Willamette Valley, 44.38, -122.91, MT071073; ben04, AMW1176/UM17, USA: Oregon: Tucker Gap: north of Tiller, 42.90, -122.71, MT071069; ben05, AMW1179/KL13, USA: California: Happy Camp, 41.88, -123.41, MT071046; ben06, AMW1182/TR03, USA: California: Friday Ridge: Willow Creek, 40.84, -123.68, MT071066; ben07, AMW1017/PS17, USA: California: Shasta County, 40.67, -122.70, MT071055; ben08, AMW1183/RL01, USA: California: Hayfork Summit, 40.60, -123.02, MT071059; ben09, AMW1155/ST25, USA: California: Basin Gulch Campground, 40.35, -122.96, MT071065; ben10, AMW1098/PC01, USA: California: Paynes Creek, 40.34, -121.73, MT071054; ben11, AMW1186/CL16, USA: California: Clearlake: east of Lucerne, 39.12, -122.79, MT071017; ben12, AMW1187/W204, USA: California: east of Pollock Pines, 38.78, -120.46, MT071076; ben13, AMW1158/CW11, USA: California: Wawona, 37.59, -119.70, MT071019; ben14, AMW1159/CG11, USA: California: Bishop Creek, 37.29, -118.56, MT071014; ben15, AMW1107/HE12, USA: California: Henry Coe State Park, 37.19, -121.54, MT071029; ben16, AMW1102/QH16, USA: California: Quail Hollow County Park, 37.09, -122.06, MT071057; ben17, AMW1103/HL19, USA: California: Henry Cowell Redwoods State Park, 37.03, -122.05, MT071030; ben18, AMW1104/UZ25, USA: California: UC Santa Cruz Arboretum, 36.98, -122.06, MT071070; ben19, AMW1161/CB05, USA: California: Onion Valley, 36.79, -118.29, MT071011; ben20, AMW1105/VL19, USA: California: Landels-Hill Big Creek Reserve,

- 36.06, -121.57, MT071071; ben21, AMW1106/V201, USA: California: Santa Lucia near Ventana Wilderness, 35.89, -121.44, MT071072; ben22, AMW1163/CE02, USA: California: Lake Isabella, 35.78, -118.56, MT071012; ben23, AMW1156/CF15, USA: California: Figueroa Mountain, 34.74, -119.98, MT071013; ben24, AMW1165/CR09, USA: California: San Bernardino, 34.25, -117.30, MT071018; ben25, AMW1166/CJ03, USA: California: San Jacinto Mountains, 33.81, -116.77, MT071016; *Pinus brachyptera*: bra01, AMW1143/C203, USA: Wyoming: Bighorn Mountains, 44.31, -106.81, MT071010; bra02, AMW1141/U201, USA: Utah: Uinta Mountains, 40.63, -111.17, MT071068; bra03, AMW1140/U116, USA: Utah: Red Canyon, 37.73, -112.25, MT071067; bra04, AMW1139/A426, USA: Arizona: Jacob Lake, 36.80, -112.26, MT070999; bra05, AMW1136/A118, USA: Arizona: Navajo Nation, 36.55, -110.47, MT070996; bra06, LangerRand2011/SM14, USA: Nevada: Spring Mountains, 36.32, -115.68, MT071062; bra07, LangerRand2011/SM26, USA: Nevada: Spring Mountains, 36.32, -115.68, MT071063; bra08, AMW1137/A229, USA: Arizona: South Rim, 35.38, -111.96, MT070997; bra09, AMW1138/A319, USA: Arizona: Hualapai, 35.10, -113.88, MT070998; bra10, AMW1083/AC11, USA: New Mexico: Mogollon Mountains: Pinos Altos, 32.91, -108.23, MT071001; bra11, AMW1077/LB05, USA: Arizona: Pinalenos: Lady Bug Saddle, 32.62, -109.83, MT071048; bra12, Marquardt/LA03, USA: Arizona: Mt. Lemmon, 32.44, -110.79, MT071047; bra13, PALA15/PT09, USA: Arizona: Palisades White Tail Campground, 32.41, -110.71, MT071056; bra14, AMW1078/BF30, USA: Arizona: Chiricahuas: Bar Foot Park, 31.92, -109.28, MT071007; bra15, Langer2010/SL11, USA: Arizona: Santa Ritas: Mt. Wrightson, 31.70, -110.85, MT071061; bra16, AMW1082/SR03, USA: Arizona: Santa Ritas Mt Hopkins, 31.69, -110.88, MT071064; bra17, AMW1081/HS19, USA: Arizona: Huachuclas: Pat Scott Peak, 31.43, -110.35, MT071031; *Pinus cooperi*: coo1, DSG1436, Mexico: Durango: west of Ejido Chavarria, 23.73, -105.55, MT071023; coo2, DSG929, Mexico: Durango: Mpio El Salto, 23.76, -105.51, MW315123; *Pinus coulteri*: cou1, DSG990, USA: California: Palomar Mountain Observatory, 33.36, -116.87, MW315124; cou2, AMW1167/J305, USA: California: Fry Creek Campground, 33.34, -116.89, MT071036; cou3, DSG1175, USA: California: north of Julian, 33.10, -116.61, MW315125; *Pinus deoniana*: dev1, DSG727, Mexico: Jalisco: Mpio Mezquitic: north of Bolanos, 21.89, -103.87, MW315127; dev2, DSG1465, Mexico: Jalisco: east of Mascota, 20.48, -104.77, MW315128; dev3, DSG721, Mexico: Hidalgo: Omitlán de Juárez, 20.20, -98.64, MW315126; dev4, DSG770, Mexico: Morelos: Tres Marias & Huitzilac, 19.04, -99.26, MW315145; *Pinus douglasiana*: dou1, DSG1076, Mexico: Chihuahua: Mirador Cerro Gallego, 27.23, -107.94, MW315129; dou2, ALR217, Mexico: Sinaloa: Puente el Carrizo, 23.52, -105.84, MT071002; dou3, DSG1470, Mexico: Jalisco: west of Puerto La Campana, 20.38, -104.60, n/a; *Pinus durangensis*: dur1, DSG1437, Mexico: Durango: west of Ejido Chavarria, 23.74, -105.55, MW315131; dur2, DSG1434, Mexico: Durango: west of Ejido Chavarria, 23.74, -105.55, MW315130; *Pinus engelmannii*: eng1, AMW1079/EC17, USA: Arizona: Chiricahuas, 31.95, -109.31, MT071024; eng2, DSG927, Mexico: Durango: west of Durango, 23.93, -104.91, MW315132; *Pinus hartwegii*: har1, DSG1182, Mexico: Veracruz: west of Cinco Palos, 19.54, -97.10, MW315135; har2, DSG765, Mexico: Guerrero: Mpio Tlacotepec Puerto de Gallo, 17.46, -100.14, MW315133; har3, DSG905, Guatemala: Sacatepéquez: Volcan de Agua, 14.47, -90.74, MW315134; *Pinus jeffreyi*: jef01, AMW1177/J520, USA: Oregon: Tiller: on serpentine soil, 42.89, -122.95, MT071037; jef02, AMW1180/J604, USA: California: Happy Camp, 41.95, -123.49, MT071038; jef03, AMW1018/JL05, USA: California: Likely, 41.23, -120.41, MT071042; jef04, AMW1181/J703, USA: California: Horse Mountain Botanical Area, 40.87, -123.73, MT071039; jef05, AMW1184/J808, USA: California: The Lassics Botanical Area, 40.35, -123.55, MT071040; jef06, AMW1000/JT05, USA: Nevada: Thomas Creek, 39.39, -119.84, MT071044; jef07, AMW1185/J910, USA: California: Snow Mountain; Lake County, 39.36, -122.74, MT071041; jef08, AMW998/JR01, USA: Nevada: Mt. Rose, 39.33, -119.87, MT071043; jef09, AMW1162/J212, USA: California: Mammoth Lakes, 37.64, -118.93, MT071035; jef10, AMW1286/J103, USA: California: Onion Valley, 36.77, -118.34, MT071034; jef11, DSG1193, USA: California: Wrightwood, 34.34, -117.60, MW315138; jef12, DG06172, Mexico: Baja: Sierra Juárez, 32.01, -115.95, MW315136; jef13, DSG949, Mexico: Baja: Mpio Ensenada, 31.04, -115.46, MW315137; *Pinus maximinoi*: max1, DSG1056, Mexico: Sonora: Mpio Yecora, 28.38, -109.04, MW315159; max2, DSG742, Mexico: Guerrero: south of San Vicente de Benitez, 17.28, -100.30, MW315140; max3, ALR116, Mexico: Chiapas: Lagunas de Montebello, 16.13, -91.73, n/a; max4, ALR117, Mexico: Chiapas: Lagunas de Montebello, 16.10, -91.73, MW315139; *Pinus montezumae*: mon1, DSG792, Mexico: Michoacan: Paracho de Verduzco, 19.58, -102.09, MW315146; mon2, DRD2, Mexico: Puebla: north of San Rafael Ixtapalucan, 19.31, -98.55, MW315141; mon3, DRD20, Mexico: Puebla: near San Salvador el Seco, 19.09, -97.67, MW315142; mon4, DRD21, Mexico: Puebla: near San Salvador el Seco, 19.09, -97.67, MW315143; mon5, DSG763, Mexico: Guerrero: Puerto de Gallo, 17.47, -100.16, MW315144; *Pinus ponderosa* s. s.: pon1, AMW1111/IH07, USA: Washington: Bisbee Mountain, 48.62, -118.17, MT071032; pon2, AMW1108/IW21, USA: Idaho: Coeur D'Alene, 47.71, -116.86, MT071033; pon3, AMW1175/G217, USA: Idaho: Kooskia Rd, 46.14, -115.75, MT071027; pon4, AMW1174/G116, USA: Idaho: Last Chance Campground, 44.98, -116.20, MT071026; pon5, AMW1115/KF01, USA: Oregon: Santiam Pass, 44.42, -121.77, MT071045; pon6, AMW1015/PB01, USA: Oregon: Blue Mountains, 44.07, -118.79, MT071053; pon7, AMW1173/BO21, USA: Idaho: Hwy 21 near Boise, 43.78, -115.89, MT071008; pon8, AMW1178/MC15, USA: California: Happy Camp Outplants from McCloud, 41.28, -121.95, MT071051; pon9, AMW1021/LP18, USA: California: Likely, 41.23, -120.41, MT071049; *Pinus pseudostrobus*: pse1, DSG1040, Mexico: Hidalgo: Apulco, 20.30, -98.35, MW315151; pse2, DSG1485, Mexico: Michoacán: se of Naranja de Tapia Tirindaro, 19.71, -101.60, MW315153; pse3, DSG999, Mexico: Mexico: north of Valle de Bravo, 19.31, -100.11, MW315150; pse4, DSG1089, Mexico: Oaxaca: Sierra de Juarez, 17.19, -96.60, MW315152; pse5, DSG890, Mexico: Chiapas: San Cristóbal de las Casas, 16.50, -92.31, MW315149; *Pinus sabiniana*: sab1, DSG1205, USA: California: north of Fresno, 37.11, -119.74, MW315155; sab2, DSG1204, USA: California: Keene: along Hwy 58, 35.22, -118.57, MW315154; *Pinus scopulorum*: sco1, Langer2011/RG28, USA: South Dakota: Slim Buttes: Reva Gap, 45.52, -103.16, MT071058; sco2, AMW1145/S122, USA: South Dakota: Black Hills, 43.85, -104.05, MT071060; sco3, AMW1142/W124, USA: Wyoming: Casper Mountains, 42.76, -106.33, MT071075; sco4, AMW1172/MB22, USA: Wyoming: Vedawoo Campground, 41.16, -105.38, MT071050; sco5, AMW1135/C101, USA: Colorado: Dolores, 37.73, -108.21, MT071009; *Pinus arizonica* var. *stormiae*: sto1, Langer2010/GU03, USA: Texas: Guadalupe Mountains, 31.89, -104.85, MT071028; sto2, ELD063/DM14, USA: Texas: Davis Mountains, 30.65, -104.18, MT071020; sto3, ELD063/DM22, USA: Texas: Davis Mountains, 30.65, -104.18, MT071021; sto4, AMW1047/AS15, USA: Texas: Big Bend National Park: Crown Mountain, 29.26, -103.26, MT071006; sto5, AMW1048/AP01, USA: Texas: Big Bend National Park: Pine Canyon, 29.26, -103.25, MT071005; sto6, AMW1044/AB02, USA: Texas: Big Bend National Park: Boot Canyon, 29.26, -103.29, MT071000; *Pinus torreyana*: tor1, DSG969, USA: California: San Diego: North Torrey Pines Road, 32.94, -117.26, MW315157; tor2, DSG970, USA: California: San Diego: North Torrey Pines Road, 32.94, -117.26, MW315158; *Pinus washoensis*: was1, AMW1025/WB21, USA: Oregon: Blue Mountains, 44.07, -118.79, MT071077; was2, AMW1002/WL12, USA: California: Likely, 41.19, -120.22, MT071078; was3, AMW999/WR01, USA: Nevada: Mt. Rose, 39.33, -119.87, MT071079; *Pinus yecorensis*: yec, DSG1063, Mexico: Sonora: Highway 16 east of Yecora, 28.38, -108.87, MT071022;
- Outgroup: *Pinus contorta* var. *contorta*: con1, DSG973, USA: Oregon: South of Newport: South Bay, 44.61, -124.07, MW315121; *Pinus contorta* var. *latifolia*: con2, DSG1018, USA: Utah: Pine Valley Campground, 40.60, -111.12, MW315122; *Pinus taeda*: tae1, tae01s2, USA: Virginia: Pembroke Rd: Virginia Beach, 36.85, -75.98, MW315156; tae3, DOB1291, USA: North Carolina: Duke University, 36.00, -78.95, n/a; tae4, AMW1043, USA: Arkansas: Hendrix College, 35.09, -92.44, MT071004; tae5, DSG605, USA: Texas: Bastrop: Lost Pines Golf Course, 30.11, -97.29, n/a.