11-17-2020

Genome-wide RAD sequencing resolves the evolutionary history of serrate leaf Juniperus and reveals discordance with chloroplast phylogeny

Kathryn A. Uckele

Robert P. Adams
Baylor University

Andrea E. Schwarzbach
The University of Texas Rio Grande Valley

Thomas L. Parchman

Recommended Citation

Uckele, K.A., Adams, R.P., Schwarzbach, A.E., Parchman, T.L., Genome-wide RAD sequencing resolves the evolutionary history of serrate leaf Juniperus and reveals discordance with chloroplast phylogeny, Molecular Phylogenetics and Evolution (2020), doi: https://doi.org/10.1016/j.ympev.2020.107022

Title: Genome-wide RAD sequencing resolves the evolutionary history of serrate leaf *Juniperus* and reveals discordance with chloroplast phylogeny

Authors: Kathryn A. Uckele,[a,*] Robert P. Adams,[b] Andrea E. Schwarzbach,[c] and Thomas L. Parchman[a]

[a] Department of Biology, MS 314, University of Nevada, Reno, Max Fleischmann Agriculture Building, 1664 N Virginia St., Reno, NV 89557, USA

[b] Baylor University, Utah Lab, 201 N 5500 W, Hurricane, UT 84790, USA

[c] Department of Health and Biomedical Sciences, University of Texas - Rio Grande Valley, 1 W University Drive, Brownsville, TX 78520, USA

E-mail address: kuckele@unr.edu (K. A. Uckele)

E-mail address: Robert_Adams@baylor.edu (R. P. Adams)

E-mail address: andrea.schwarzbach@utrgv.edu (A. E. Schwarzbach)

E-mail address: tparchman@unr.edu (T. L. Parchman)

[*]**Address for correspondence:** Kathryn Uckele, 1664 N Virginia Street, MS 314, Reno, NV 89557, USA, *E-mail address:* kuckele@unr.edu

Abstract

Juniper (*Juniperus*) is an ecologically important conifer genus of the Northern Hemisphere, the members of which are often foundational tree species of arid regions. The serrate leaf margin clade is native to topologically variable regions in North America, where hybridization has likely played a prominent role in their diversification. Here we use a reduced-representation sequencing approach (ddRADseq) to generate a phylogenomic data set for 68 accessions representing all 22 species in the serrate leaf margin clade, as well as a number of close and distant relatives, to improve understanding of diversification in this group. Phylogenetic analyses using three methods (SVDquartets, maximum likelihood, and Bayesian) yielded highly congruent and well-resolved topologies. These phylogenies provided improved resolution relative to past analyses based on Sanger sequencing of nuclear and chloroplast DNA, and were largely consistent with taxonomic expectations based on geography and morphology. Calibration of a Bayesian phylogeny with fossil evidence produced divergence time estimates for the clade consistent with a late Oligocene origin in North America, followed by a period of elevated diversification between 12 and 5 Mya. Comparison of the ddRADseq phylogenies with a phylogeny based on Sanger-sequenced chloroplast DNA revealed five instances of pronounced discordance, illustrating the potential for chloroplast introgression, chloroplast transfer, or incomplete lineage sorting to influence organellar phylogeny. Our results improve understanding of the pattern and tempo of diversification in *Juniperus*, and highlight the utility of reduced-representation sequencing for resolving phylogenetic relationships in non-model organisms with reticulation and recent divergence.

Keywords: diversification, juniper, RADseq, reticulation, western North America

1. Introduction

The complex geologic and climatic history of western North America played an important role in the diversification of many plant groups throughout the Cenozoic (Axelrod, 1948, 1950). Tectonic uplift, climate change, transcontinental land bridges, and glacial cycles created opportunity for range shifts, geographic barriers to admixture, and allopatric speciation (Hewitt, 1996; Calsbeek et al., 2003; Hewitt, 2004; Weir and Schluter, 2007). Hybridization has also been prominent in the evolutionary history of Nearctic plant taxa, as glacial cycles allowed periods of isolation and subsequent secondary contact (Swenson and Howard, 2005; Hewitt, 2011). The interactions among topography, climate, and reticulation have shaped diversification and challenged phylogenetic analyses for many plant genera in western North America (e.g., Rieseberg et al., 1991; Kuzoff et al., 1999; Bouillé et al., 2011; Xiang et al., 2018; Shao et al., 2019). However, improved genomic sampling enabled by high-throughput sequencing data has recently increased phylogenetic resolution for many young and reticulated groups (e.g., Stephens et al., 2015; Massatti et al., 2016; McVay et al., 2017; Moura et al., 2020) and generally stands to enhance our understanding of diversification for plant taxa in this region.

Junipers (*Juniperus*, Cupressaceae) are ecologically and economically important conifers of arid and semi-arid landscapes throughout the Northern Hemisphere (Farjon, 2005; Adams, 2014). Unlike other genera in Cupressaceae, the juniper lineage evolved a fleshy female cone, functionally resembling a berry, which is an important food source for many birds and small mammals (Phillips, 1910; Santos et al., 1999). The serrate junipers, distinguished by the presence of microscopic serrations on their scale leaf margins, are particularly resistant to water stress compared with other juniper groups (Willson et al., 2008) and often represent the dominant trees in arid habitats of the western United States and Mexico (West et al., 1978; Romme et al., 2009).

A number of species in this clade are expanding their range in North America, and while the main causes of these expansions are unclear for some taxa (Miller and Wigand, 1994; Weisberg et al., 2007; Romme et al., 2009), fire suppression, over-grazing by cattle, and under-browsing by native herbivores appear to be the dominant factors underlying *J. ashei* and *J. pinchotii* range expansion in west Texas (Taylor, 2008). Despite several attempts to resolve phylogenetic relationships in this ecologically important clade (Mao et al., 2010; Adams and Schwarzbach, 2013a,b), its complex evolutionary history including recent divergence, long generation times, and hybridization have likely obfuscated phylogenetic signal in previous molecular data sets.

The juniper lineage likely originated in Eurasia during the Eocene and subsequently split into three major monophyletic sections (Mao et al., 2010; Adams and Schwarzbach, 2013a): sect. Caryocedrus (1 sp., *J. drupacea*, eastern Mediterranean); sect. Juniperus (14 spp., Asia and the Mediterranean except *J. jackii* and *J. communis*); and the largest clade, sect. Sabina (approximately 62 spp., Northern Hemisphere except *J. procera*). Section Sabina contains three main monophyletic clades (Mao et al., 2010; Adams and Schwarzbach, 2013a): the turbinate, single-seeded, entire leaf margin junipers of the Eastern Hemisphere (16 spp.); the multi-seeded, entire leaf margin junipers of both the Eastern and Western Hemispheres (23 spp.); and the serrate leaf margin junipers (serrate junipers hereafter) of western North America (22 spp.), which are the focus of this study. The ancestral serrate juniper lineage likely arrived in North America from Eurasia via the North Atlantic Land Bridge (NALB) or Bering Land Bridge (BLB) (Mao et al., 2010). Extant serrate junipers are largely restricted to North America, inhabiting arid and semi-arid regions of the western United States, Mexico, and the high, dry mountains of Guatemala (*J. standleyi*; Adams, 2014) (Fig. 1).

A previous phylogenetic analysis based on Sanger sequencing data with complete species-level sampling of the serrate juniper clade was highly biased towards chloroplast DNA (cpDNA), utilizing four cpDNA regions and one nuclear DNA (nrDNA) region [full data set representing 4,411 base pairs (bp), referred to as nr-cpDNA hereafter; Adams and Schwarzbach, 2013b]. Hybridization and discordance between cpDNA and nrDNA based phylogenies have been reported across *Juniperus* (Adams, 2016; Adams et al., 2016) and within the serrate junipers in particular (Adams et al., 2017) and may have contributed to unexpected topologies in the previous predominantly cpDNA based phylogeny (Adams and Schwarzbach, 2013b). Incomplete lineage sorting due to long generation times and recent divergence may have also contributed to paraphyletic and unresolved relationships in the nr-cpDNA analyses of Adams and Schwarzbach (2013b). Multi-locus data encompassing larger genealogical variation should reduce topological uncertainty in this clade, while also allowing for insight into nuclear-chloroplast discordance and its potential causes. Mao et al. (2010) estimated divergence times, diversification rates, and geographic origins of all major juniper clades; however, limited sampling of the serrate juniper clade precluded dating for many of its internal nodes. Divergence time estimation for a complete serrate juniper phylogeny stands to elucidate patterns of diversification at more recent time scales which appear to be important for diversification across the genus (Mao et al., 2010).

High-throughput sequencing technologies have rapidly improved our ability to apply genome-wide information to phylogenetic inference (McCormack et al., 2013; Leaché and Oaks, 2017; Bravo et al., 2019). Data from whole genomes (e.g., Kimball et al., 2019; Allio et al., 2020), whole transcriptomes (e.g., Leebens-Mack et al., 2019), targeted capture (e.g., de La Harpe et al., 2019; Liu et al., 2019; Karimi et al., 2020), and genome-skimming approaches (e.g.,

112 Liu et al., 2020; Nevill et al., 2020) have resolved evolutionary relationships complicated by
113 incomplete lineage sorting and reticulate evolution (Faircloth et al., 2013; Alexander et al., 2017;
114 Carter et al., 2019). Methods using restriction enzyme digest to reduce genome complexity [e.g.,
115 restriction site-associated DNA sequencing (RADseq; Miller et al., 2007; Baird et al., 2008)]
116 have been particularly valuable for phylogenetic applications in non-model organisms due to
117 their ability to sample large numbers of informative polymorphisms without requiring prior
118 genomic resources (Takahashi et al., 2014; Leaché and Oaks, 2017; Near et al., 2018; Salas-
119 Lizana and Oono, 2018; Hipp et al., 2020). RADseq data have improved the resolution of many
120 groups that have been recalcitrant to phylogenetic analysis with small numbers of Sanger-
121 sequenced loci due to rapid, recent, or reticulate evolution (Wagner et al., 2013; Massatti et al.,
122 2016; Paetzold et al., 2019; Rancilhac et al., 2019; Léveillé-Bourret et al., 2020). Although
123 allelic dropout (i.e., the nonrandom absence of sequence data at a locus due to restriction site
124 mutations) can result in larger amounts of missing data across more strongly diverged lineages,
125 analyses of empirical and simulated RADseq data have illustrated its effectiveness for resolving
126 even relatively deep divergences (e.g., up to 60 Mya, Rubin et al., 2012; Cariou et al., 2013;
127 Eaton et al., 2017; Lecaudey et al., 2018; Du et al., 2020).

128 Here we utilized a double-digest RADseq approach (ddRADseq; Parchman et al., 2012;
129 Peterson et al., 2012) to generate a phylogenomic data set for all extant species of serrate
130 junipers (*Juniperus* sect. Sabina) as well as several close and distant relatives. As methods for
131 phylogenetic inference utilizing multi-locus data make different assumptions about genealogical
132 variation among lineages, we inferred phylogenetic trees using three distinct approaches
133 (SVDquartets, maximum likelihood, and Bayesian). Our results produce consistent and highly
134 resolved topologies, reveal discordance with phylogenies inferred with cpDNA alone, and

illustrate variation in diversification rates consistent with the climatic and geologic history of western North America.

2. Materials & Methods

2.1 Taxon sampling and ddRADseq library prep

We sampled leaf material from 68 individuals representing all 22 serrate juniper species and six outgroup species (Table S1). Most serrate juniper taxa and two outgroup taxa (*Hesperocyparis bakeri* and *H. arizonica*, Cupressaceae; Zhu et al., 2018) were either the same individuals or different individuals collected from the same populations as those analyzed previously by Adams and Schwarzbach (2013b). Thus, analyses of the data presented here have 50 samples (73.5%) in common with Adams and Schwarzbach (2013b) and 18 samples (26.5%) which are unique to this study. Five additional outgroup taxa [*Juniperus drupacea* (*Juniperus* sect. Caryocedrus); *J. communis* (*Juniperus* sect. Juniperus); *J. virginiana, J. sabina* var. *sabina*, and *J. sabina* var. *balkanensis* (smooth leaf junipers of sect. Sabina)] were added to better understand evolutionary divergence at deeper time scales in this genus. Two additional *J. poblana* var. *poblana* localities (Nayarit, MX, and Amozoc de Mota, Puebla, MX), one additional *J. poblana* variety (*J. poblana* var. *decurrens*), and an additional *J. durangensis* locality (Sierra de Gamón, Durango, MX) were included to investigate the potential for recent evolutionary divergence in these taxa. Finally, we substituted *J. ashei* samples from Waco, TX, with *J. ashei* samples from nearby Tarrant County, TX, for this study.

DNA was extracted from dried leaf tissue with Qiagen DNeasy Plant Mini Kits and quantified with a Qiagen QIAxpert microfluidic analyzer prior to library preparation (Qiagen Inc., Valencia, CA, USA). Reduced-representation libraries for Illumina sequencing were

158 constructed using a ddRADseq method (Parchman et al., 2012; Peterson et al., 2012) in which
159 genomic DNA was digested with two restriction enzymes, *Eco*RI and *Mse*I, and custom oligos
160 with Illumina base adaptors and unique barcodes (8, 9 or 10 bases in length) were ligated to the
161 digested fragments. Ligated fragments were PCR amplified with a high-fidelity proofreading
162 polymerase (Iproof polymerase, BioRad Inc., Hercules, CA, USA) and subsequently pooled into
163 a single library. Libraries were size-selected for fragments between 350 and 450 bp in length
164 with the Pippin Prep System (Sage Sciences, Beverly, MA) at the University of Texas Genome
165 Sequencing and Analysis Facility. Two lanes of single-end 100-base sequencing were executed
166 at the University of Wisconsin-Madison Biotechnology Center using an Illumina HiSeq 2500
167 platform.

168

169 **2.2 Preparation, filtering, and assembly of ddRADseq data**

170 To identify and discard Illumina primer/adapter sequences and potential biological
171 sequence contaminants (e.g., PhiX, *E. coli*), we used the `tapioca` pipeline
172 (https://github.com/ncgr/tapioca), which uses `bowtie2` (v. 2.2.5; Langmead and Salzberg, 2012)
173 to identify reads which align to a database of known contaminant sequences. To ensure that
174 cpDNA did not influence our analyses, we used the same approach to discard all reads which
175 aligned to the *Juniperus squamata* chloroplast genome (GenBank Accession Number
176 MK085509; Xie et al., 2019). To demultiplex reads to individual, we used a custom Perl script
177 that corrects one or two base sequencing errors in barcoded regions, parses reads according to
178 their associated barcode sequence, and trims restriction site-associated bases. Files with the read
179 data for each individual are available at Dryad (https://doi.org/10.5061/dryad.qbzkh18df).

To process the raw data into a matrix of putatively orthologous aligned loci, we utilized `ipyRAD` (v. 0.9.16; Eaton, 2014) which was designed to process reduced-representation data for phylogenetic workflows and allows for indel variation across samples during clustering (Eaton, 2014; Razkin et al., 2016). We largely used default values, as these settings produced multiple alignments of tractable size which led to highly resolved, supported, and consistent topologies across inference methods. First, nucleotide sites with phred quality scores less than 33, which represent base calls with an error probability greater than 0.0005%, were considered missing and replaced with an ambiguous nucleotide base ("N"). Next, sequences were *de novo* clustered within individuals using `vsearch` (v. 2.14.1; Rognes et al., 2016) and aligned with `muscle` (v. 3.8.155; Edgar, 2004) to produce stacks of highly similar reads. A similarity clustering threshold (*clust_threshold*) of 85% was applied during this and a later clustering step because it produced a thorough yet tractable number of loci and a highly supported topology with the `TETRAD` (SVDquartets) inference method. To ensure accurate base calls, all stacks with a read depth less than 6 were discarded. Observed base counts across all sites in all stacks informed the joint estimation of the sequencing error rate and heterozygosity, which informed statistical base calls according to a binomial model. At this step, each stack within each individual was reduced to one consensus sequence with heterozygote bases represented by IUPAC ambiguity codes, and any consensus sequences with more than 5% ambiguous bases (*max_Ns_consens*) or heterozygous sites (*max_Hs_consens*) were discarded to remove poor alignments. The remaining consensus sequences from all individuals were clustered again, this time across individuals, using the same assembly method and similarity threshold as used in the previous within-sample clustering step. The resulting clusters, which represent putative ddRADseq loci shared across individuals, were discarded if they contained more than 8 indels (*max_Indels_locus*) or 20%

variable sites (*max_SNPs_locus*), as an excess of either could indicate poor alignment. To detect potential paralogs, consensus sequences were removed if they contained one or more heterozygous sites shared across more than 50% of all samples (*max_shared_Hs_locus*) or more than 2 haplotypes (Eaton, 2014). We retained all loci that were present in a minimum of four samples (*min_samples_locus*) to prevent over-filtering of missing data, which can negatively affect downstream inference (Rubin et al., 2012; Wagner et al., 2013; Huang and Knowles, 2014; Takahashi et al., 2014). Two sequence alignment formats, `ipyRAD`'s database file and a `phylip` file of concatenated loci, were used as input for SVDquartets (`TETRAD`) and maximum likelihood (`RAxML`) phylogenetic analyses, respectively. The database file contains the clustered sequence data as well as linkage information for each locus. We used a python script (http://github.com/btmartin721/raxml_ascbias/) to remove all invariant sites from the `phylip` sequence alignment prior to analysis with `RAxML`.

To understand the timing and tempo of diversification within the serrate juniper clade, we utilized fossil evidence to inform divergence time estimates in a Bayesian phylogenetic inference framework. For this analysis, we included one sample per serrate juniper species, including three outgroup samples from the closely related smooth leaf juniper clade (*J. virginiana*, *J. sabina* var. *sabina*, and *J. sabina* var. *balkanensis*), with priority given to juniper samples with higher sequencing coverage depth. Sequencing reads for this subset of 25 samples were *de novo* assembled with default `ipyRAD` parameter values except for the *min_samples_locus* parameter, which was increased from 4 to 20, and the *clust_threshold* parameter, which was increased from 85% to 90%. Increasing these parameters effectively reduced both the proportion of missing data and the size of the sequence alignment to ensure tractable computation time with Bayesian inference methods. However, one caveat of excluding missing data in RADseq data sets is that it

can bias the distribution of mutation rates represented across loci and lower the accuracy of downstream phylogenetic inference (Huang and Knowles, 2014). The resulting `nexus` sequence alignment of concatenated loci was utilized as input for Bayesian analysis (`RevBayes`). Complete information on parameter settings for this and the aforementioned assembly, as well as the sequence alignment files, are archived at Dryad (https://doi.org/10.5061/dryad.qbzkh18df).

2.3 Phylogenetic analyses

After removing invariant sites, the `phylip` formatted sequence alignment for all taxa, including outgroups, was analyzed with maximum likelihood as implemented by `RAxML` (v. 8.2.12; Stamatakis, 2014) under the GTR+Γ model of nucleotide substitution corrected for ascertainment bias (`-m ASC_GTRGAMMA`). Support was assessed with 100 rapid bootstrap replicates (`-N 100`), followed by a thorough maximum likelihood search for the best-scoring tree (`-f a`). Although `RAxML` is fast and often used for analysis of concatenated RADseq loci (Lemmon and Lemmon, 2013), phylogenetic inference with concatenated data necessarily ignores genealogical variation among loci and is statistically inconsistent as the number of genes increases (Kubatko and Degnan, 2007; Roch and Steel, 2015).

To account for genealogical variation among sampled loci and to incorporate coalescent stochasticity into analyses, we also conducted species tree inference using a site-based approach, SVDquartets (Chifman and Kubatko, 2014), as implemented by `TETRAD` (Eaton et al., 2017). `TETRAD` is included with `ipyRAD` and implements the SVDquartets algorithm, using information on genotype calls and linkage to sample unlinked SNPs. Briefly, SVDquartets uses the multi-species coalescent model to generate a probability distribution on the data patterns at the tips of a species tree which can be used to compute a score on a quartet of taxa and infer the true quartet

relationship (Chifman and Kubatko, 2014, 2015). These quartet relationships can be inferred for all or a subset of all possible quartets, and a quartet amalgamation software (in this case, QMC v. 2.10; Snir and Rao, 2012) joins the inferred quartets into the species tree. Here, we used TETRAD's default number of quartets, which is the number of samples to the power of 2.8, which yielded 135,215 quartets (16.6% of total possible). To quantify support for the nodes of the species tree, we implemented a standard nonparametric bootstrapping procedure for 100 replicates. The inferred tree was manually rooted with the clade containing *Hesperocyparis bakeri* and *H. arizonica*.

To enable comparison of topologies produced with ddRADseq and cpDNA Sanger sequencing data, we repeated the methods of Adams and Schwarzbach (2013b) on the same individuals or different individuals collected from the same populations as those analyzed in the ddRADseq analysis for a total of 66 individual samples. Thus, the cpDNA analysis presented here has 59 samples (89.4%) in common with the aforementioned ddRADseq analyses and 7 substitutional samples (10.6%). DNA extractions, PCR amplifications, and Sanger sequencing of the four chloroplast loci (petN-psbM, trnS-trnG, trnD-trnT, and trnL-trnF) were conducted using the methods described in Adams and Schwarzbach (2013b). The GTR+Γ+I nucleotide substitution model provided the best fit to the cpDNA data according to Akaike's information criterion in Modeltest (v.3.7; Posada and Crandall, 1998), and analysis was conducted with Mr.Bayes (v.3.1; Ronquist and Huelsenbeck, 2003). Two rounds of four chains were run for a total of 10 million generations, sampling every 1000 generations after an initial burn in of 25% of generations.

To understand diversification rate variation and the timing of divergence events across the serrate juniper clade, we inferred a time-calibrated phylogeny for a subset of individuals

representing all serrate juniper taxa and three closely related outgroup samples from the smooth leaf juniper clade (*J. virginiana*, *J. sabina* var. *sabina*, and *J. sabina* var. *balkanensis*) with a Bayesian method (`RevBayes` v. 1.0.12; Höhna et al., 2017). First, we implemented a model-selection procedure to compare the relative fits with Bayes factors of the JC, HKY, GTR, GTR+Γ, and GTR+Γ+I models of nucleotide substitution. Second, the `nexus` sequence alignment of concatenated loci generated with `ipyRAD` was modeled under the best fit substitution model given a topology modeled with a constant-rate birth-death process, which was parameterized with a sampling fraction of 0.39 due to incomplete sampling of the smooth leaf juniper clade (Kendall, 1948; Nee et al., 1994; Höhna, 2015). We relaxed the assumption of a global molecular clock by allowing each branch-rate variable to be drawn from a lognormal distribution. Eight independent MCMC chains were run for 400,000 generations with a burn-in of 10,000 generations and sampled every 10 generations. Chains were visually assessed for convergence with `Tracer` (v. 1.7.1; Rambaut et al., 2018) and quantitatively assessed with effective sample sizes (ESS) and the Gelman-Rubin convergence diagnostic (Gelman and Rubin, 1992) using the *gelman.diag* function in R (`CODA` package; Plummer at al., 2006).

Fossil calibration points and node age prior distributions can influence estimates of divergence times (Graur and Martin, 2004; Sauquet et al., 2012; Wang and Mao, 2016). We used three fossil calibration points: one at the root node for the serrate juniper clade (not shown in Fig. 4A) and two at internal nodes (asterisks, Fig. 4A) representing the MRCA (Most Recent Common Ancestor) of all extant serrate leaf junipers and the MRCA of the western U.S. clade (*J. californica*, *J. osteosperma*, *J. occidentalis*, and *J. grandis*). Fossil assignments were based on morphology and coincided with those made by a previous phylogenetic analysis of *Juniperus* (Mao et al., 2010). Justifications for these assignments can be found in Table S2. A fossil

specimen of *J. creedensis* (23 Mya; Axelrod, 1987), representing the first appearance of a serrate juniper in the fossil record, provided the minimum age constraints for both the root node (representing the MRCA of the serrate leaf juniper clade and the smooth leaf juniper outgroup taxa) and the internal node representing the MRCA of the serrate junipers. The maximum age constraint for the root node, specified with a uniform prior distribution, was the estimated age of the crown lineage of Cupressoideae (134 Mya; Mao et al., 2012), a subfamily of Cupressaceae which contains *Thuja*, *Cupressus*, *Juniperus*, and other genera. A fossil specimen of *J. desatoyana* (16 Mya; Axelrod, 1991), representing a stem ancestor of a subclade containing *J. osteosperma*, *J. occidentalis*, and *J. grandis*, provided the minimum age constraint of 16 Mya for the divergence of this subclade from *J. californica* (i.e., the MRCA of the western U.S. clade). For the internal nodes representing the MRCA of the serrate leaf junipers and the MRCA of the western U.S. clade, the ages of the fossil specimens were modelled as exponential distributions with means of 23 Mya + 1 and 16 Mya + 1, respectively, divided by λ, the parameter of the exponential distribution. The maximum clade credibility tree was inferred from the burned distribution of posterior trees, and the smooth leaf juniper outgroup samples were pruned in R with the *drop.tip* function (`ape` package; Paradis and Schliep, 2019) prior to subsequent visualization and analyses.

The inferred Bayesian chronogram was used to generate a lineage through time plot with the *ltt.plot* function in R (`ape` package; Paradis and Schliep, 2019). To determine whether the rate of lineage diversification was constant through time, we used the *diversi.gof* function in R (`ape` package; Paradis and Schliep, 2019) to compute the Cramér-von Mises and Anderson-Darling goodness-of-fit tests (Stephens, 1974; Paradis, 1998).

To estimate the probability of all possible ancestral ranges at each ancestral node, we utilized the `BioGeoBEARS` package (v. 1.1.2; Matzke, 2013a,b) and its dependencies, `rexpokit` (Matzke et al., 2019) and `cladoRcpp` (Matzke, 2018), in R. This package permits statistical selection of six competing historical biogeographical models (DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, and BAYAREALIKE+J) and includes an additional cladogenetic event, founder-event speciation, represented by the +J notation in DEC+J, DIVALIKE+J, and BAYAREALIKE+J models (Matzke, 2014). While these six methods similarly assume that anagenetic dispersal and extinction occur along branches, they allow for different subsets of cladogenetic range-changing processes. The `BioGeoBEARS` supermodel incorporates all of these different processes, treating them as free parameters which can be excluded or estimated from the data.

Five operational geographic areas (A, western U.S.; B, central U.S.; C, eastern U.S.; D, northern/central MX; E, southern MX; Fig. 5) were defined by both geopolitical and ecologically-relevant boundaries (Level I Ecoregions of North America; see https://www.epa.gov/eco-research/ecoregions). To determine the contemporary geographic range of each species, we referenced U.S. tree species range maps when available (Little, 1971) and juniper range maps otherwise (Adams, 2014) (Table S3). This matrix of distribution information for each species, as well as the maximum clade credibility tree inferred with `RevBayes`, was used as input for ancestral range estimation. We used plotting functions provided by `BioGeoBEARS` to visualize estimates of ancestral range for the model with the lowest AIC.

3. RESULTS

3.1 Assembly of ddRADseq data for phylogenetic inference

Two Illumina HiSeq lanes generated approximately 460 million reads, of which 373,596,722 remained after quality and contaminant filtering. `Bowtie2` aligned 4,007,039 reads (1.07%) to the *J. squamata* chloroplast genome, which we subsequently removed prior to read assembly and SNP calling. Three samples were removed prior to assembly due to low read count relative to other samples, providing 68 samples for `ipyRAD` input. The full data set of 68 samples was initially assembled into 307,146 loci, of which 130,581 remained after filtering, providing 929,267 SNPs (344,189 parsimony informative) for phylogenetic inference with `RAxML` and `TETRAD`. Each individual possessed, on average, approximately five million raw reads which were assembled, on average, into 19,417 loci (14.9% of total loci). Similar to other RADseq phylogenetic data sets (Cariou et al., 2013; Eaton et al., 2017), the resulting sequence alignments provided as input for `RAxML` and `TETRAD` exhibited a large proportion of missing data (84.69% and 83.51% of sites contained missing values, respectively). 10,461,968 invariant sites were removed from the `phylip` formatted sequence alignment prior to analysis with `RAxML`. `TETRAD` sampled 124,530 unlinked SNPs for its analysis.

For the Bayesian analysis, increasing the *min_samples_locus* and *clust_threshold* parameters for assembly of the 22 serrate juniper and 3 outgroup samples effectively diminished the effect of allelic dropout and reduced the proportion of missing data at the expense of incorporating fewer loci for phylogenetic inference. An initial set of 479,143 loci were reduced to 2,390 after filtering steps, providing 18,436 SNPs (7,894 parsimony informative) for phylogenetic inference. On average, each individual possessed 5.7 million raw reads which were assembled into 2,078 loci (86.9% of total loci). Only 14.72% of sites contained missing values in the resulting `nexus` sequence alignment.

3.2 Phylogenetic analyses

The maximum likelihood and SVDquartets analyses of ddRADseq data (hereafter referred to as the ddRADseq phylogenies) recovered high support (>95%) for most nodes in the phylogeny, with few exceptions (Fig. 2). The maximum likelihood phylogeny identified nine monophyletic clades within the serrate junipers (Fig. 2 left), which are colored accordingly in Figs. 2-4. The SVDquartets phylogeny resolved the same nine clades (Fig. 2 right), although two were less supported: 1) the Cerro Petosí clade (*J. zanonii* and *J. saltillensis*, which are sympatric on Cerro Petosí, MX) and 2) the subalpine-alpine clade (*J. jaliscana*, *J. standleyi*, and *J. monticola*, which are collectively found in subalpine/alpine environments). The ddRADseq phylogenies consistently recovered deeper relationships among three main monophyletic clades: 1) the western U.S. clade (*J. californica*, *J. osteosperma*, *J. occidentalis*, and *J. grandis*); 2) the *ashei* clade (*J. comitana*, *J. ovata*, and *J. ashei*), the *J. deppeana* species complex, the one-seeded serrate junipers (*J. arizonica*, *J. monosperma*, *J. coahuilensis*, *J. pinchotii*, and *J. angosturana*, which largely exhibit 1 seed per cone); and 3) the Cerro Petosí clade, the *J. durangensis* clade (*J. martinezii* and *J. durangensis* subspp.), the subalpine-alpine clade, *J. flaccida*, and the *J. poblana* species complex. The ddRADseq phylogenies were consistent in their relationships among the three high-level clades, including the placement of the western U.S. clade as basal to the other serrate juniper clades (Fig. 2). Although nearly all relationships were strongly supported and consistent across both phylogenies (Fig. 2), three were inconsistently resolved. In the maximum likelihood phylogeny, the outgroup taxa *J. drupacea* and *J. communis* are in distinct clades, whereas they are sister to one another in the SVDquartets phylogeny (Fig. 2). In the maximum likelihood phylogeny, the *ashei* clade is basal to the *J. deppeana* species complex and the one-seeded group with high support; whereas, in the SVDquartets phylogeny,

386 the *J. deppeana* species complex is basal, with high support (Fig. 2). Finally, although both
387 placements had low support, the maximum likelihood phylogeny placed *J. flaccida* as sister to
388 the *J. poblana* complex, whereas the SVDquartets phylogeny placed *J. flaccida* as basal to the
389 subalpine-alpine clade (Fig. 2).

390 Aside from the few conflicts above, the topologies inferred across multiple approaches
391 (maximum likelihood, SVDquartets, and Bayesian) were consistent, highly supported, and
392 congruent with established taxonomy based on morphological and chemical characters (Figs. 2,
393 4A). Whereas Adams and Schwarzbach (2013b) inferred a paraphyletic relationship for *J. sabina*
394 in which *J. virginiana* was sister to *J. sabina* var. *sabina* (Fig. 1 from Adams and Schwarzbach,
395 2013b), the ddRADseq phylogenies recovered a monophyletic relationship for the two *J. sabina*
396 varieties (Fig. 2). In addition, three of the nine monophyletic clades recovered with generally
397 high support in the ddRADseq phylogenies (Fig. 2) were paraphyletic in the nr-cpDNA
398 phylogeny of Adams and Schwarzbach (2013b): 1) the western U.S. clade; 2) the *J. ashei* clade;
399 and 3) the subalpine-alpine clade. First, the western U.S. clade was paraphyletic in the nr-
400 cpDNA tree of Adams and Schwarzbach (2013b) and is not basal to the other serrate juniper
401 clades, except for *J. californica*. Second, the *J. ashei* clade was paraphyletic in the nr-cpDNA
402 tree, with *J. comitana* basal to the western U.S. clade, *J. ovata* basal to the Cerro Petosí clade,
403 and *J. ashei* sister to *J. deppeana* (Fig. 1 from Adams and Schwarzbach, 2013b). Third, the nr-
404 cpDNA tree of Adams and Schwarzbach (2013b) placed *J. flaccida* and *J. poblana* in the
405 subalpine-alpine clade, causing the subalpine-alpine clade to be paraphyletic.

406 Sanger-sequenced data spanning four cpDNA regions (petN-psbM, trnS-trnG, trnL-trnF,
407 trnD-trnT), originally generated by Adams and Schwarzbach (2013b), was reanalyzed here with
408 additional samples to produce a phylogeny for detection of cyto-nuclear discordance when

compared with ddRADseq phylogenies (both analyses were largely based on the same sets of individuals, or individuals from the same populations). The cpDNA phylogeny inferred here had less resolution and a distinctly different topology than that of the combined nr-cpDNA analysis of Adams and Schwarzbach (2013b). Figure 3 illustrates five areas of discordance between the maximum likelihood ddRADseq and Bayesian cpDNA phylogenies. First, the cpDNA phylogeny inferred a sister relationship between *J. sabina* var. *balkinensis* and *J. virginiana* (Fig. 3 right), whereas the maximum likelihood phylogeny inferred a sister relationship between *J. sabina* var. *balkinensis* and *J. sabina* var. *sabina* (Fig. 3 left), consistent with taxonomic expectations. Second, the western U.S. clade is paraphyletic in the cpDNA tree, and *J. californica* is sister to *J. comitana* rather than grouped with the other western U.S. serrate junipers (Fig. 3 right). Third, the cpDNA tree placed *J. zanonii* sister to *J. ovata* and nested within a clade with *J. ashei* (Fig. 3 right), rather than sister to *J. saltillensis* as it is in the maximum likelihood tree (Fig. 3). Fourth, the cpDNA tree also included *J. arizonica* in this highly supported clade, making the one-seeded group (*J. arizonica*, *J. monosperma*, *J. coahuilensis*, *J. pinchotii*, and *J. angosturana*) paraphyletic (Fig. 3 right). Finally, in the cpDNA tree, *J. flaccida* is nested within *J. poblana*, which causes this complex to be paraphyletic (Fig. 3 right).

3.3 Diversification history of the serrate junipers

The GTR+ Γ model of nucleotide substitution provided the best fit to the sequence alignment generated for the subset of serrate juniper samples, including three outgroup samples (*J. virginiana*, *J. sabina* var. *sabina*, and *J. sabina* var. *balkanensis*). The Bayesian topology was largely consistent with the maximum likelihood and SVDquartets phylogenies, with an exception being the paraphyletic relationship among the one-seeded junipers (Fig. 4A). The other eight of

the nine monophyletic clades and all three high-level clades recovered by the ddRADseq phylogenies (Fig. 2) were likewise recovered by the Bayesian phylogeny (Fig. 4A) with high support (>99% posterior support for all nodes; Figure 4A). Our Bayesian calibration suggests that the serrate juniper clade arose during the late Oligocene (crown age 23.73 Mya, 95% highest posterior density [HPD]: 23 – 25.15 Mya), which is slightly younger but not inconsistent with previous estimates of 25.82 (23.00 – 31.20) and 29.43 Mya (23.25 – 41.72) inferred from cpDNA data with BEAST and MULTIDIVTIME, respectively (Mao et al. 2010). According to our analysis, the western U.S. clade (*J. californica*, *J. osteosperma*, *J. occidentalis*, and *J. grandis*) arose in the early Miocene (crown age 17.20 Mya, HPD: 16.00 – 19.32 Mya), which is slightly younger but not inconsistent with previous estimates of 19.16 (16.00 – 25.44) and 24.39 (15.88 – 36.64) Mya inferred from cpDNA with BEAST and MULTIDIVTIME, respectively (Mao et al., 2010).

The Bayesian phylogenetic model estimated a mean speciation rate for the serrate juniper and closely related smooth leaf juniper clades of 0.14 sp/Ma (HPD: 2.46E-5 – 0.21 sp/Ma), and an extinction rate of 0.03 sp/Ma (HPD: 7.18E-8 – 0.11 sp/Ma), resulting in a mean net diversification rate (speciation rate – extinction rate) of 0.11 sp/Ma (HPD: -0.07 – 0.20 sp/Ma). A lineage through time plot (Fig. 4B) suggests deviations from a constant rate of diversification over time, which was confirmed quantitatively with the Cramér-von Mises and Anderson-Darling goodness-of-fit tests, both of which rejected the null model of constant diversification rate and exponentially distributed branching times (Cramér-von Mises: $W2 = 2.326$, $p < 0.01$; Anderson-Darling GOF: $A2 = 3.189$, $p < 0.01$). Comparing lineage origination over time with a constant rate of diversification reveals a period of notably elevated diversification from ~12-5 Mya (Fig. 4B).

Comparison of AIC and AICc values for each of the six historical biogeographical models with `BioGeoBEARS` suggested that the DIVALIKE model provided the best fit to the data (AICc weight = 0.62). According to this model, the most probable ancestral range for the serrate juniper clade is a combined range of the Western U.S. and northern/central MX (Fig. 5). The ancestral range of the western U.S. serrate junipers was estimated as the western U.S., but the ancestral range of the remaining serrate junipers was estimated as northern/central MX (Fig. 5).

4. Discussion

Junipers are considered foundational plants throughout arid regions of North America, where they provide habitat and food resources for numerous animal species (Poddar and Lederer, 1982; Gottfried, 1992; Adams, 2014). The serrate juniper clade is endemic and adapted to arid environments of North America, yet lack of phylogenetic resolution has precluded thorough understanding of how geography and climate may have influenced diversification in this relatively young group. Compared with previous work on limited numbers of serrate juniper taxa and Sanger-sequenced cp and nr loci (Mao et al., 2010; Adams and Schwarzbach, 2013b), the phylogenies inferred here with ddRADseq data offer greater resolution and support, and are largely consistent with longstanding taxonomy. Our results provide insight into the evolutionary history of the serrate junipers, including variation in the tempo of diversification, and reveal notable instances of discordance among phylogenies inferred from nuclear and chloroplast variation.

4.1 Diversification history of the serrate leaf margin junipers

Our results are consistent with the hypothesis (Mao et al. 2010) that the ancestral serrate juniper lineage originated during the Oligocene epoch in North America (Fig. 4). During the Eocene-Oligocene transition (~33.9 Mya), decreasing temperatures and increasing seasonality occurred in many regions globally, potentially favoring the expansion of arid-adapted juniper populations (Kennett, 1977; Buchardt, 1978; Wolfe, 1978). As suggested by Mao et al. (2010), the serrate juniper ancestor may have first reached North America via the North Atlantic Land Bridge (NALB) or the Bering Land Bridge (BLB). The NALB, which provided an Atlantic connection through Greenland, was beginning to fragment during the Eocene, but fossil evidence suggests that it continued to facilitate the transatlantic migration of tree species well into the Miocene (Donoghue et al., 2001; Grímsson and Denk, 2005; Denk et al., 2010; Helmstetter et al., 2019). The BLB, which provided a Pacific connection across the Bering Strait, likely facilitated numerous transcontinental migrations during the Cenozoic (Hopkins, 1959, 1967; Donoghue et al., 2001; Wang and Ran, 2014), including other North American tree genera (e.g., *Fagus* and *Quercus*, Manos and Stanford, 2001; *Hesperocyparis* + *Callitropsis*, Terry et al., 2016; *Pinus*, Badik et al., 2018; *Picea*, Shao et al., 2019).

We inferred a combined ancestral range for the serrate juniper clade which included the western U.S. and northern/central MX (Fig. 5). Two lines of evidence suggest that the common ancestor of the serrate junipers established in the western United States after migrating from Eurasia and potentially before expanding into northern and central Mexico. First, our results generally suggest that the western U.S. clade is basal to all other serrate juniper clades (Figs. 2). Second, the earliest appearances of serrate junipers in the fossil record date to the late Oligocene and early Miocene in the western United States, and feature characteristics similar to extant western U.S. junipers (Axelrod, 1956, 1987, 1991; Wolfe, 1964). During the Oligocene, the

western United States was characterized by drier climates, expanding sclerophyll vegetation, and the origin of many contemporary tree species (Axelrod, 1976; Reveal, 1980). Moderate temperatures during this time shifted mixed conifer and subalpine forests coastward (Axelrod, 1976), which, alongside increasingly xeric conditions throughout the region, may have provided ecological opportunity for serrate juniper establishment.

Divergence time estimates suggest that approximately one-third of all divergence events occurred relatively recently in the serrate juniper clade. Elevated diversification rates occurred from approximately 12 to 5 Mya during the late Miocene and early Pliocene (Fig. 4B). Notably, this period coincided with enhanced diversification rates across juniper generally, which was attributed by Mao et al. (2010) to global cooling and uplift of the Qinghai-Tibetan plateau, though the latter is not relevant for North America. In western North America, uplift of the American Cordillera during the late Miocene (~12-5 Mya) induced a rain shadow effect and the expansion of arid habitats (Axelrod, 1950, 1985; Leopold and Denton, 1987; Wilson and Pitts, 2010), causing population extirpation and the evolution of drought adapted flora (Reveal, 1980). The serrate junipers are particularly tolerant to water stress (Willson et al., 2008) and may have persisted or expanded into newly vacant habitats during this period. Furthermore, increased fire and the expansion of grassland habitat at lower elevations may have restricted junipers to higher elevations, causing range disjunctions between mountain chains and allopatric divergence across altitudinal zones (Retallack, 1997; Wilson and Pitts, 2010). Indeed, some extant sister species exhibit geographical associations with adjacent mountain ranges, with one example being *J. occidentalis* and *J. grandis*, which diverged around the Miocene-Pliocene boundary: *Juniperus occidentalis* inhabits low to intermediate elevations associated with the Cascade range and *J. grandis* occupies mid to high elevation alpine environments associated with the Sierra Nevada

range (Terry et al., 2000). Miocene diversification has also been observed in other temperate trees (*Pinus*; Willyard et al., 2007; *Cupressus*; Xu et al., 2010; *Abies*; Aguirre-Planter et al., 2012; *Quercus* section *Lobatae*, series *Agrifoliae*; Hauser et al., 2017) and has been similarly attributed to falling global temperatures and mountain uplift.

4.2 Utilizing ddRADseq data to resolve relationships among the serrate junipers

Our analyses were highly consistent across different inference approaches and recapitulated many of the general patterns suggested by previous analyses, including the monophyly of the "one-seeded", "Cerro Petosí", and "*durangensis*" clades (Adams and Schwarzbach, 2013b) and recognition of two *J. deppeana* varieties, var. *gamboana* and var. *deppeana* (Mao et al., 2010; Adams and Schwarzbach, 2013b). However, ddRADseq analyses, based on more extensive genomic sampling, provided enhanced resolution of early divergences in the serrate juniper clade by consistently recovering three major groups with high support: 1) the western U.S. clade; 2) the *J. ashei* clade, *J. deppeana* species complex, and one-seeded clade [also suggested by Mao et al. (2010)]; and 3) the Cerro Potosí clade, *J. durangensis* clade, the subalpine-alpine clade, *J. flaccida*, and *J. poblana* species complex. Our analyses additionally recovered some relationships which were previously unresolved due to incomplete sampling, predominantly cpDNA-based inference, or analyses being based on limited genomic sampling (e.g., Mao et al., 2010; Adams and Schwarzbach, 2013b). We highlight noteworthy examples of these results below.

Members of the western U.S. clade (*J. occidentalis*, *J. grandis*, *J. osteosperma*, and *J. californica*) are morphologically cohesive (see Vasek, 1966) and occur along a north-south moisture gradient from the montane zone of the eastern Cascade and Sierra Nevada ranges (*J.*

occidentalis and *J. grandis*, respectively), through the pinyon-juniper woodlands of the Great Basin and Colorado Plateau (*J. osteosperma*), to the Mojave Desert (*J. californica*). Nonetheless, both Mao et al. (2010) and Adams and Schwarzbach (2013b) inferred paraphyletic placements of *J. californica* relative to other members of the group. In contrast, our analyses inferred *J. californica* as the most basal member of a monophyletic western U.S. clade (Figs. 2, 4A), consistent with previous taxonomic classification. Our analyses additionally resolved relationships among *J. osteosperma*, *J. occidentalis*, and *J. grandis*, which hybridize in western Nevada (Terry et al., 2000; Terry, 2010; Adams, 2013a,b). *Juniperus grandis* and *J. occidentalis* were previously classified as *J. occidentalis* varieties based on morphological similarities which exhibit clinal variation (Vasek, 1966); however, they were not sister to one another in the analysis of Adams and Schwarzbach (2013b). Our analyses assigned them as sister taxa and placed *J. osteosperma* basal to them (Figs. 2, 4A), consistent with expectations based on morphology and geography.

Juniperus ashei and *J. ovata* (previously *J. ashei* var. *ovata*; Adams and Baker, 2007) hybridize extensively where they occur parapatrically in the trans-Pecos region of Texas, and were considered subspecies until recent phylogenetic analysis merited the recognition of *J. ovata* at the specific level (Adams and Schwarzbach, 2013b). In contrast to Adams and Schwarzbach (2013b), our analyses indicate a sister relationship for *J. ashei* and *J. ovata*, which is supported by morphology and the geographical proximity of these taxa (Figs. 2, 4A). The inference of *J. comitana* as the basal member of this clade (Figs. 2, 4A), however, is not supported by morphology and chemistry (Adams, 2000) and merits additional research.

We included new collections of *J. durangensis* from Sierra Gamon, Durango, in our analyses due to their atypical morphology relative to type localities of *J. durangensis* (Socorro

570 Gonzales, pers., comm.). Our analyses suggest phylogenetic distinctness of *J. durangensis* from
571 Sierra Gamon, despite growing only 150 km northeast of the type locality near El Salto, Durango
572 (Fig. 2). Ongoing morphological and phytochemical analyses may help determine whether *J.*
573 *durangensis* from Sierra Gamon merits recognition as a new variety. Similarly, new *J. poblana*
574 accessions were analyzed from Nayarit, Oaxaca, and Puebla, as potential cases of intraspecific
575 divergence. The only additional variety suggested by our analyses besides the previously
576 recognized *J. poblana* var. *decurrens* is represented by samples from Oaxaca, which formed a
577 monophyletic group in both the maximum likelihood and SVDquartet analyses (Fig. 2).

578 In contrast to Adams and Schwarzbach (2013b), the ddRADseq maximum likelihood
579 analysis placed *J. jaliscana*, *J. monticola*, and *J. standleyi* in a highly-supported monophyletic
580 clade, and *J. flaccida* and *J. poblana* in a distinct sister clade with low support (Fig. 2 left). The
581 SVDquartets and Bayesian trees likewise indicate monophyly of *J. jaliscana*, *J. monticola*, and *J.*
582 *standleyi*, but with lower support (Figs. 2 right, 4A). We refer to this group as the "subalpine-
583 alpine clade" because they occur at mid-high elevations. *Juniperus monticola* is widespread in
584 Mexico and occupies subalpine and alpine habitats at elevations of 2400-4500 m (Adams, 2014),
585 while *J. jaliscana* occupies pine-oak forests at elevations of 1335-2670 in southern Durango and
586 northwest Jalisco (Zanoni and Adams, 1979). *Juniperus standleyi* is found in extreme southeast
587 Mexico and Guatemala at elevations of 3000-4250 m (Adams, 2014). Phylogenetically adjacent
588 taxa, *J. flaccida* and *J. poblana*, likewise occur in subalpine habitats but are distinguished
589 morphologically from the subalpine-alpine clade by branches which are flaccid at the tips so that
590 their foliage appears to be drooping (Adams, 2014).

591 The relationship between *J. flaccida* and *J. poblana* (previously *J. flaccida* var. *poblana*)
592 has been taxonomically challenging due to the paucity of distinguishing morphological features

593 and their ability to hybridize (Zanoni and Adams, 1976; Adams et al., 2018c). Our analyses
594 suggest a distinct taxonomic status for *J. poblana*, but disagree on the relationship between *J.*
595 *poblana* and *J. flaccida*. Consistent with taxonomic expectations, maximum likelihood and
596 Bayesian phylogenies support a sister relationship between *J. flaccida* and *J. poblana* (although
597 poorly supported in the former) (Figs. 2 left, 4A); however, the SVDquartets tree suggests a
598 more distant placement of *J. flaccida* basal to the subalpine-alpine clade (Fig. 2 right). An
599 affinity of *J. flaccida* towards the subalpine-alpine clade was suggested by the Adams and
600 Schwarzbach (2013b) phylogeny, which recovered a sister relationship between *J. flaccida* and
601 *J. standleyi*. A potential explanation for this, and for conflicting phylogenetic signal in the
602 ddRADseq data, could be introgression from *J. standleyi* into *J. flaccida*.

603 While the maximum likelihood and SVDquartets analyses produced predominantly
604 consistent results, there were three instances of discordance which highlight areas where gene
605 tree variation may have influenced inference (Maddison, 1997; Huang et al., 2010; Tonini et al.,
606 2015). As incomplete lineage sorting (ILS) is a major source of gene tree-species tree
607 discordance, phylogenetic inference under the multi-species coalescent (e.g., SVDquartets) may
608 perform more accurately under high ILS conditions compared with concatenation approaches
609 (e.g., RAxML) (Chou et al., 2015). Shallow divergences may be especially prone to ILS, which
610 may explain the discordance between the *J. ashei* clade, the *J. deppeana* complex, and the one-
611 seeded clade (Figs. 2, 4A). Alternatively, hybridization is widely reported throughout *Juniperus*
612 (e.g., Adams, 1994; Terry et al. 2000; Adams et al., 2020) and may have contributed to
613 topological discordance in areas of low support, e.g., the relationship of *J. flaccida* (Fig. 2).
614 Finally, allelic dropout in reduced-representation data may complicate the resolution of older
615 splits, and may have played a role in the discordance observed among two outgroup samples, *J.*

616 *communis* and *J. drupacea* (Fig. 2). Overall, differences in model assumptions and conflicting
617 phylogenetic signal likely influenced the few points of discordance observed among our different
618 inference methods.

619

620 **4.3 Discordance between phylogenies inferred with nuclear and chloroplast DNA**

621 Discordance among nr and cpDNA is common, and can arise from processes including
622 incomplete lineage sorting (Degnan and Rosenburg, 2009), hybridization (Rieseberg and Soltis,
623 1991; Rieseberg et al., 1996), and lateral transfer of organellar genomes (Stegemann et al.,
624 2012). In angiosperms prone to hybridization, discordance among nr and cpDNA gene trees has
625 often been attributed to introgression and chloroplast capture (e.g., Acosta and Premoli, 2010;
626 Lee-Yaw et al., 2019; Liu et al., 2020). When maternally inherited in angiosperms, cpDNA
627 exhibits more intraspecific population divergence and higher introgression across species
628 boundaries than nrDNA (Petit and Excoffier, 2009; Du et al., 2009). However, conifer cpDNA is
629 usually paternally inherited through pollen (Neale and Sederoff, 1989; Mogensen, 1996),
630 typically exhibits weaker population differentiation than nr or mtDNA, and is expected to move
631 less readily across species boundaries (e.g., Petit et al., 2005; Gerardi et al., 2010; Godbout et al.,
632 2010). Thus, chloroplast introgression should generally be less likely in conifers, although
633 potential examples of chloroplast introgression and capture have been described (e.g., Liston et
634 al., 2007; Gernandt et al., 2018). Interestingly, theoretical work suggests cp capture may be
635 driven by mitochondrial based cytoplasmic male sterility (Frank, 1989) in hybridizing
636 angiosperms with maternal co-inheritance of mt and cp genomes (Tsitrone et al., 2003). This
637 mechanism couldn't operate in most conifers (e.g., *Picea* and *Pinus*) which inherit mt (maternal)
638 and cp (paternal) genomes separately. However, Cupressaceae (including *Juniperus*) have

639 paternal inheritance of both mt and cp genomes (Mogensen, 1996; Adams, 2019), which could
640 increase the probability of chloroplast capture via cytoplasmic interactions (Tsitrone et al.,
641 2003). Alternatively, lateral transfer of chloroplast through natural grafting during periods of
642 sympatry could lead to apparent chloroplast capture in the absence of hybridization (Stegeman et
643 al., 2012).

644 As in other conifers (Petit and Hampe, 2006), reproductive isolation is often weak among
645 *Juniperus*, and hybridization has been documented among serrate juniper species including *J.*
646 *occidentalis* and *J. osteosperma* (Terry et al. 2000; Terry 2010), *J. ashei* and *J. ovata* (Adams et
647 al., 2020), and *J. angosturana* and *J. coahuilensis* (Adams, 1994). Potential cases of
648 introgression or horizontal transfer of cpDNA have also been noted in the group (Adams, 2016;
649 Adams et al., 2016, 2017). For example, *J. occidentalis* and *J. osteosperma* hybridize extensively
650 in northwestern Nevada, and a cpDNA haplotype fixed in *J. occidentalis* appears to have
651 introgressed through the western range of *J. osteosperma* (Terry et al., 2000, Terry, 2010). A
652 potential case of chloroplast capture occurred in the closely related smooth leaf juniper clade
653 (Fig. 3 right) between *J. thurifera* (chloroplast donor, not shown) and *J. sabina* var. *sabina*
654 (chloroplast recipient), giving rise to the allotetrapoloid *J. sabina* var. *balkanensis* (Adams et al.,
655 2016, 2018a,b; Farhat et al., 2019). The cpDNA tree indicates notable discordance consistent
656 with this idea, placing *J. sabina* var. *balkanensis* in a clade with *J. virginiana* (Fig. 3 right),
657 while ddRADseq analyses inferred the expected monophyletic relationship for the *J. sabina*
658 varieties (Figs. 2, 3 left). As the ddRADseq phylogenies are congruent with taxonomic
659 expectations based on morphology and geography, several strong instances of discordance in the
660 cpDNA phylogeny suggest the potential for chloroplast introgression or transfer, although
661 incomplete lineage sorting remains plausible for several of these cases.

Clear instances of discordance involve species from diverged lineages inferred with nuclear data that unexpectedly share cpDNA variation (Fig. 3). ddRADseq data inferred a western U.S. clade containing *J. californica* (Fig. 3 left), as expected based on morphology and geography; however, the cpDNA tree placed *J. californica* in a well-supported clade with *J. comitana*, which is restricted to southern Mexico/northern Guatemala (Fig. 3 right). Introgression or transfer of a *J. comitana*-type chloroplast from an ancestral *J. comitana* lineage into *J. californica* could underly such discordance (Fig. 3 right). Second, cpDNA placed *J. zanonii*, a sub-alpine plant that grows at the 3550 m summit of Cerro Potosí, NL, Mexico, within a clade with *J. ashei* and *J. ovata*, sibling species that grow on limestone in Central Texas (Adams, 2008) (Fig. 3 right). The *ashei* clade is substantially diverged from *J. zanonii* in ddRADseq analyses, which placed *J. zanonii* with *J. saltillensis* (Fig. 3 left), consistent with *J. zanonii* and *J. saltillensis* exhibiting altitudinal zonation at Cerro Petosí, Mexico. This discordance could have arisen from chloroplast introgression or transfer from an ancestral *J. ovata*/*J. ashei* into ancestral *J. zanonii*, as these lineages likely experienced sympatry during the Pleistocene (Adams and Baker, 2007). Third, *J. arizonica* and *J. coahuilensis* occur parapatrically, but the two taxa are highly similar morphologically and hybridize in the Trans-Pecos, Texas region (Adams, 2014, 2017). ddRADseq analyses placed *J. arizonica* in the one-seeded group with *J. coahuilensis* (Fig. 3 left), as expected, while cpDNA placed *J. arizonica* within the *J. ashei* clade (Fig. 3 right). Chloroplast introgression or transfer from *J. ashei* to *J. arizonica* could underly such discordance (Fig. 3 right), although incomplete lineage sorting is also possible for these closely related clades. These discordances suggest that nr and cpDNA histories can vary prominently in *Juniperus*, and while evidence for chloroplast capture or horizontal transfer is scarce in conifers, these processes may deserve further study in Cupressaceae.

5. Conclusion

Our analyses of ddRADseq data produced highly resolved and largely consistent phylogenies depicting the evolutionary history of the serrate junipers of western North America. While these phylogenies were strongly consistent with taxonomic expectations based on morphology and ecology, cpDNA phylogenies illustrated several pronounced cases of discordance, suggesting the potential for processes to differentially influence the evolutionary history of the chloroplast genome. An improved understanding of the timing and tempo of diversification, including the age of origin of the serrate juniper clade and its elevated rate of diversification during the late Miocene, illustrates how the interaction between geologic, geographic, and climatic processes may have influenced patterns of diversification in this group. This study contributes to a growing body of research demonstrating the effectiveness of reduced-representation sequencing data for resolving the phylogenies of non-model organisms (e.g., Eaton and Ree, 2013; Herrera and Shank, 2016; Massatti et al., 2016; Eaton et al., 2017; Near et al., 2018; Paetzold et al., 2019) and the complex evolutionary histories of western North American taxa characterized by reticulate evolution and recent divergence.

6. Acknowledgements

We would like to acknowledge María Socorro González-Elizondo, Jim Bartel, Lucio Caamaño Onofre, Allen Coombes, and Alex Tashev for field assistance and sample collection. We are very grateful to Lauren Im, who provided technical support in ArcGIS. We would like to acknowledge and thank Michael May for his invaluable support utilizing RevBayes. We also thank Joshua Jahner and Matthew Forister for their valuable comments on the manuscript. This

709 work was funded by Baylor University (Project No. 032512) to R.P.A., a National Science
710 Foundation Graduate Research Fellowship Award to K.A.U. (Award No. 1650114), and a 2018
711 EECG award from the American Genetics Association to K.A.U. The authors declare no
712 conflicts of interests.

713

Literature Cited

Acosta, M.C., Premoli, A.C., 2010. Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). Mol. Phylogenet. Evol. 54 (1), 235-242, https://doi.org/10.1016/j.ympev.2009.08.008

Adams, R.P., 1994. Geographic variation and systematics of monospermous *Juniperus* (Cupressaceae) from the Chihuahua Desert based on RAPDs and terpenes. Biochem. Syst. Ecol. 22 (7), 699-710, https://doi.org/10.1016/0305-1978(94)90056-6

Adams, R.P., 2000. The serrate leaf margined *Juniperus* (Section Sabina) of the western hemisphere: systematics and evolution based on leaf essential oils and Random Amplified Polymorphic DNAs (RAPDs). Biochem. Syst. Ecol. 28 (10), 975-989, https://doi.org/10.1016/S0305-1978(00)00022-3

Adams, R.P., 2008. Distribution of *Juniperus ashei* var. *ashei* and var. *ovata* around New Braunfels, Texas. Phytologia. 90 (1), 97-102.

Adams, R.P., 2013a. Hybridization between *Juniperus grandis*, *J. occidentalis* and *J. osteosperma* in northwest Nevada I: Terpenes, Leviathan mine, Nevada. Phytologia. 95 (1), 58-69.

Adams, R.P., 2013b. Hybridization between *Juniperus grandis*, *J. occidentalis* and *J. osteosperma* in northwest Nevada II: Terpenes, Buffalo Hills, Northwestern Nevada. Phytologia. 95 (1), 107-114.

Adams, R.P., 2014. Junipers of the World: The genus *Juniperus*. Trafford Publishing, Bloomington.

Adams, R.P., 2016. Two new cases of chloroplast capture in incongruent topologies in the *Juniperus excelsa* complex: *J. excelsa* var. *turcomanica* comb. nov. and *J. excelsa* var. *seravschanica* comb. nov. Phytologia. 98 (3), 219-231.

738 Adams, R.P., 2017. Multiple evidences of past evolution are hidden in nrDNA of *Juniperus*
739 *arizonica* and *J. coahuilensis* populations in the trans-Pecos, Texas region. Phytologia. 99
740 (1), 38-47.

741 Adams, R.P., 2019. Inheritance of chloroplasts and mitochondria in Conifers: A review of
742 paternal, maternal, leakage and facultative inheritance. Phytologia. 101 (2), 134-138.

743 Adams, R.P., Baker, L.E., 2007. Pleistocene infraspecific evolution in *Juniperus ashei* Buch.
744 Phytologia. 89 (1), 8-23.

745 Adams, R.P., Johnson, S.T., Schwarzbach, A.E., 2020. Long distance gene flow facilitated by
746 bird-dispersed seeds in wind-pollinated species: A story of hybridization and introgression
747 between *Juniperus ashei* and *J. ovata* told by nrDNA and cpDNA. Phytologia. 102 (2), 55-
748 74.

749 Adams, R.P., Boratynski, A., Marcysiak, K., Roma-Marzio, F., Peruzzi, L., Bartolucci, F.,
750 Conti, F., Mataraci, T., Tashev, A.N., Siljak-Yakovlev, S., 2018a. Discovery of *Juniperus*
751 *sabina* var. *balkanensis* R. P. Adams & Tashev in Macedonia, Bosnia-Herzegovina, Croatia
752 and southern Italy and relictual polymoal. rphisms found in nrDNA. Phytologia. 100 (2),
753 117-127.

754 Adams, R.P., Farhat, P., Shuka, L., Silak-Yakovlev, S., 2018b. Discovery of *Juniperus sabina*
755 var. *balkanensis* R. P. Adams and A. N. Tashev in Albania and relictual polymorphisms
756 found in nrDNA. Phytologia. 100 (3), 187-194.

757 Adams, R.P., Johnson, S., Coombes, A.J., Caamaño, L., González-Elizondo, M.S., 2018c.
758 Preliminary examination of hybridization and introgression between *Juniperus flaccida* and
759 *J. poblana*: nrDNA and cpDNA sequence data. Phytologia. 100 (2), 145-152.

760 Adams, R.P., Schwarzbach A.E., 2013a. Phylogeny of *Juniperus* using nrDNA and four cpDNA
761 regions. Phytologia. 95 (2), 179-187.

762 Adams, R. P., Schwarzbach, A.E., 2013b. Taxonomy of the serrate leaf *Juniperus* of North
763 America: Phylogenetic analyses using nrDNA and four cpDNA regions. Phytologia. 95 (2),
764 172-178.

765 Adams, R.P., Schwarzbach, A.E., Tashev, A.N., 2016. Chloroplast capture by a new variety,
766 *Juniperus sabina* var. *balkanensis* R. P. Adams and A. N. Tashev, from the Balkan
767 peninsula: A putative stabilized relictual hybrid between *J. sabina* and ancestral *J. thurifera.*
768 Phytologia. 98 (2), 100-111.

769 Adams, R.P., Socorro González-Elizondo M., González-Elizondo M., Ramirez Noy D.,
770 Schwarzbach A.E., 2017. DNA sequencing and taxonomy of unusual serrate *Juniperus* from
771 Mexico: Chloroplast capture and incomplete lineage sorting in *J. coahuilensis* and allied
772 taxa. Phytologia. 99 (1), 62-73.

773 Aguirre-Planter, É., Jaramillo-Correa, J.P., Gómez-Acevedo, S., Khasa, D.P., Bousquet, J.,
774 Eguiarte, L.E., 2012. Phylogeny, diversification rates and species boundaries of
775 Mesoamerican firs (*Abies*, Pinaceae) in a genus-wide context. Mol. Phylogenet. Evol. 62
776 (1), 263-274, https://doi.org/10.1016/j.ympev.2011.09.021

777 Alexander, A.M., Su, Y.C., Oliveros, C.H., Olson, K.V., Travers, S.L., Brown, R.M., 2017.
778 Genomic data reveals potential for hybridization, introgression, and incomplete lineage
779 sorting to confound phylogenetic relationships in an adaptive radiation of narrow-mouth
780 frogs. Evolution. 71 (2), 475-488, https://doi.org/10.1111/evo.13133

781 Allio, R., Scornavacca, C., Nabholz, B., Clamens, A.L., Sperling, F.A., Condamine, F.L., 2020.
782 Whole genome shotgun phylogenomics resolves the pattern and timing of swallowtail
783 butterfly evolution. Syst. Biol. 69 (1), 38-60, https://doi.org/10.1093/sysbio/syz030

784 Axelrod, D.I., 1948. Climate and evolution in western North America during middle Pliocene
785 time. Evolution. 2, 127-144, https://doi.org/10.2307/2405373

786 Axelrod, D.I., 1950. Evolution of desert vegetation in western North America. Carnegie Inst.
787 Wash. Publ. 590, 215-306.

788 Axelrod, D.I., 1956. Mio-Pliocene floras from west-central Nevada. Vol. 33. University of
789 California Press.

790 Axelrod, D.I., 1976. History of the coniferous forests, California and Nevada. University of
791 California Press, Berkeley.

792 Axelrod, D.I., 1985. Rise of the grassland biome, central North America. Bot. Rev. 51 (2), 163-
793 201.

794 Axelrod, D.I., 1987. The late Oligocene Creede flora, Colorado. Vol. 130. University of
795 California Press.

796 Axelrod, D.I., 1991. The Early Miocene buffalo canyon flora of western Nevada. Vol. 135.
797 University of California Press, Berkeley.

798 Badik, K.J., Jahner, J.P., Wilson, J.S., 2018. A biogeographic perspective on the evolution of
799 fire syndromes in pine trees (*Pinus*: Pinaceae). R. Soc. Open Sci. 5 (3), 172412,
800 http://dx.doi.org/10.1098/rsos.172412

801 Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U.,
802 Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using
803 sequenced RAD markers. PLoS One. 3 (10), e3376,
804 http://doi.org/10.1371/journal.pone.0003376

805 Bouillé, M., Senneville, S., Bousquet, J., 2011. Discordant mtDNA and cpDNA phylogenies
806 indicate geographic speciation and reticulation as driving factors for the diversification of

807 the genus *Picea*. Tree Genet. Genomes. 7 (3), 469-484, http://doi.org/10.1007/s11295-010-
808 0349-z

809 Bravo, G.A., Antonelli, A., Bacon, C.D., Bartoszek, K., Blom, M.P., Huynh, S., Jones, G.,
810 Knowles, L.L., Lamichhaney, S., Marcussen, T. Morlon, H., 2019. Embracing
811 heterogeneity: coalescing the Tree of Life and the future of phylogenomics. PeerJ. 7, e6399.

812 Buchardt, B., 1978. Oxygen isotope palaeotemperatures from the Tertiary period in the North
813 Sea area. Nature. 275 (5676), 121-123.

814 Calsbeek, R., Thompson, J.N., Richardson, J.E., 2003. Patterns of molecular evolution and
815 diversification in a biodiversity hotspot: the California Floristic Province. Mol. Ecol. 12 (4),
816 1021-1029, https://doi.org/10.1046/j.1365-294X.2003.01794.x

817 Cariou, M., Duret, L., Charlat, S., 2013. Is RAD-seq suitable for phylogenetic inference? An in
818 silico assessment and optimization. Ecol. Evol. 3 (4), 846-852,
819 https://doi.org/10.1002/ece3.512

820 Carter, K.A., Liston, A., Bassil, N.V., Alice, L.A., Bushakra, J.M., Sutherland, B.L., Mockler,
821 T.C., Bryant, D.W., Hummer, K.E., 2019. Target capture sequencing unravels *Rubus*
822 evolution. Front. Plant Sci. 10, 1615, https://doi.org/10.3389/fpls.2019.01615

823 Chifman, J., Kubatko, L., 2014. Quartet inference from SNP data under the coalescent model.
824 Bioinformatics. 30 (23), 3317-3324, https://doi.org/10.1093/bioinformatics/btu530

825 Chifman, J., Kubatko, L., 2015. Identifiability of the unrooted species tree topology under the
826 coalescent model with time-reversible substitution processes, site-specific rate variation, and
827 invariable sites. J. Theor. Biol. 374, 35-47, https://doi.org/10.1016/j.jtbi.2015.03.006

828 Chou, J., Gupta, A., Yaduvanshi, S., Davidson, R., Nute, M., Mirarab, S., Warnow, T., 2015. A
829 comparative study of SVDquartets and other coalescent-based species tree estimation
830 methods. BMC Genomics, 16 (10), 1-11

831 Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the
832 multispecies coalescent. Trends Ecol. Evol. 24 (6), 332-340,
833 https://doi.org/10.1016/j.tree.2009.01.009

834 de La Harpe, M., Hess, J., Loiseau, O., Salamin, N., Lexer, C., Paris, M., 2019. A dedicated
835 target capture approach reveals variable genetic markers across micro- and macro-
836 evolutionary time scales in palms. Mol. Ecol. Resour. 19 (1), 221-234,
837 https://doi.org/10.1111/1755-0998.12945

838 Denk, T., Grímsson F., Zetter, R., 2010. Episodic migration of oaks to Iceland: Evidence for a
839 North Atlantic "land bridge" in the latest Miocene. Am. J. Bot. 97 (2), 276-287,
840 https://doi.org/10.3732/ajb.0900195

841 Donoghue, M.J., Bell, C.D., Li, J., 2001. Phylogenetic patterns in Northern Hemisphere plant
842 geography. Int. J. Plant Sci. 162 (S6), S41-S52, https://doi.org/10.1086/323278

843 Du, F.K., Petit, R.J., Liu, J.Q., 2009. More introgression with less gene flow: chloroplast vs.
844 mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other
845 conifers. Mol. Ecol. 18 (7), 1396-1407, https://doi.org/10.1111/j.1365-294X.2009.04107.x

846 Du, Z.Y., Harris, A.J., Xiang, Q.Y.J., 2020. Phylogenomics, co-evolution of ecological niche
847 and morphology, and historical biogeography of buckeyes, horsechestnuts, and their
848 relatives (Hippocastaneae, Sapindaceae) and the value of RAD-seq for deep evolutionary
849 inferences back to the Late Cretaceous. Mol. Phylogenet. Evol. 145, 106726,
850 https://doi.org/10.1016/j.ympev.2019.106726

851 Eaton, D.A.R., 2014. PyRAD: assembly of de novo RADseq loci for phylogenetic analyses.
852 Bioinformatics. 30 (13), 1844-1849, https://doi.org/10.1093/bioinformatics/btu121

853 Eaton, D.A.R, Ree, R.H., 2013. Inferring phylogeny and introgression using RADseq data: an
854 example from flowering plants (*Pedicularis*: Orobanchaceae). Syst. Bio. 62 (5), 689-706,
855 https://doi.org/10.1093/sysbio/syt032

856 Eaton, D.A., Spriggs, E.L., Park, B., Donoghue, M.J., 2017. Misconceptions on missing data in
857 RAD-seq phylogenetics with a deep-scale example from flowering plants. Syst. Bio. 66 (3),
858 399-412, https://doi.org/10.1093/sysbio/syw092

859 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high
860 throughput. Nucleic Acids Res. 32 (5), 1792-1797, https://doi.org/10.1093/nar/gkh340

861 Faircloth, B.C., Sorenson, L., Santini, F., Alfaro, M.E., 2013. A phylogenomic perspective on
862 the radiation of ray-finned fishes based upon targeted sequencing of ultraconserved
863 elements (UCEs). PloS One. 8 (6), e65923,
864 https://dx.doi.org/10.1371%2Fjournal.pone.0065923

865 Farhat, P., Siljak-Yakovlev, S., Adams, R.P., Robert, T., Dagher-Kharrat, M.B., 2019. Genome
866 size variation and polyploidy in the geographical range of *Juniperus sabina* L.
867 (Cupressaceae). Bot. Lett. (in press), https://doi.org/10.1080/23818107.2019.1613262

868 Farjon, A., 2005. Monograph of Cupressaceae and *Sciadopitys*. Royal Botanic Gardens, Kew,
869 Richmond.

870 Frank, S.A., 1989. The evolutionary dynamics of cytoplasmic male sterility. Am. Nat. 133 (3),
871 345-376, https://doi.org/10.1086/284923

872 Gelman, A., Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences.
873 Stat. Sci. 7 (4), 457-472, https://www-jstor-org.unr.idm.oclc.org/stable/2246093

874 Gerardi, S., Jaramillo-Correa, J.P., Beaulieu, J. and Bousquet, J., 2010. From glacial refugia to
875 modern populations: new assemblages of organelle genomes generated by differential

876 cytoplasmic gene flow in transcontinental black spruce. Mol. Ecol. 19 (23), 5265-5280,
877 https://doi.org/10.1111/j.1365-294X.2010.04881.x

878 Gernandt, D.S., Aguirre Dugua, X., Vázquez-Lobo, A., Willyard, A., Moreno Letelier, A.,
879 Pérez de la Rosa, J.A., Piñero, D., Liston, A., 2018. Multi-locus phylogenetics, lineage
880 sorting, and reticulation in *Pinus* subsection *Australes*. Am. J. Bot. 105(4), 711-725,
881 https://doi.org/10.1002/ajb2.1052

882 Gernandt, D.S., López, G.G., García, S.O., Liston, A., 2005. Phylogeny and classification of
883 *Pinus*. Taxon. 54 (1), 29-42, https://doi.org/10.2307/25065300

884 Godbout, J., Beaulieu, J. and Bousquet, J., 2010. Phylogeographic structure of jack pine (*Pinus*
885 *banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern
886 North America. Am. J. Bot. 97 (11), 1903-1912, https://doi.org/10.3732/ajb.1000148

887 Gottfried, G.J., 1992. Ecology and management of the southwestern pinyon-juniper woodlands.
888 In: Ffolliott, P.F., Gottfried, G.J., Bennett, D.A., Hernandez, V.M., Ortega-Rubio, C.A.,
889 Hamre, R.H. [Tech Coords.], Ecology and management of oak and associated woodlands:
890 perspectives in the southwestern United States and northern Mexico. United States
891 Department of Agriculture, Forest Service, Washington, D.C., pp. 78-85.

892 Graur, D., Martin, W., 2004. Reading the entrails of chickens: molecular timescales of evolution
893 and the illusion of precision. Trends Genet. 20 (2), 80-86.

894 Grímsson, F., Denk, T., 2005. *Fagus* from the Miocene of Iceland: systematics and
895 biogeographical considerations. Rev. Palaeobot. Palynol. 134 (1-2), 27-54,
896 https://doi.org/10.1016/j.revpalbo.2004.11.002

897 Hauser, D.A., Keuter, A., McVay, J.D., Hipp, A.L., Manos, P.S., 2017. The evolution and
898 diversification of the red oaks of the California Floristic Province (*Quercus* section *Lobatae*,
899 series *Agrifoliae*). Am. J. Bot. 104 (10), 1581-1595, https://doi.org/10.3732/ajb.1700291

900 Helmstetter, A.J., Buggs, R.J., Lucas, S.J., 2019. Repeated long-distance dispersal and
901 convergent evolution in hazel. Sci. Rep. 9 (1), 1-12, https://doi.org/10.1038/s41598-019-
902 52403-2

903 Herrera, S., Shank, T.M., 2016. RAD sequencing enables unprecedented phylogenetic
904 resolution and objective species delimitation in recalcitrant divergent taxa. Mol. Phylogenet.
905 Evol. 100, 70-79, https://doi.org/10.1007/s10709-011-9547-3

906 Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and
907 speciation. Biol. J. Linn. Soc. 58 (3), 247-276, https://doi.org/10.1111/j.1095-
908 8312.1996.tb01434.x

909 Hewitt, G.M., 2004. Genetic consequences of climatic oscillations in the Quaternary. Philos.
910 Trans. R. Soc. Lond., B, Biol. Sci. 359 (1442), 183-195,
911 https://doi.org/10.1098/rstb.2003.1388

912 Hewitt, G.M., 2011. Quaternary phylogeography: the roots of hybrid zones. Genetica. 139 (5),
913 617-638.

914 Hipp, A.L., Manos, P.S., Hahn, M., Avishai, M., Bodénès, C., Cavender-Bares, J., Crowl, A.A.,
915 Deng, M., Denk, T., Fitz-Gibbon, S., Gailing, O., 2020. Genomic landscape of the global
916 oak phylogeny. New Phytol. 226 (4), 1198-1212, https://doi.org/10.1111/nph.16162

917 Höhna, S., 2015. The time-dependent reconstructed evolutionary process with a key-role for
918 mass-extinction events. J. Theor. Biol. 380, 321-331,
919 https://doi.org/10.1016/j.jtbi.2015.06.005

920 Höhna, S., Landis, M.J., Heath, T.A., 2017. Phylogenetic inference using RevBayes. Curr.
921 Protoc. Bioinform. 57 (6.16), 1-34, https://doi.org/10.1002/cpbi.22

922 Hopkins, D.M., 1967. The Bering land bridge. Vol. 3. Stanford University Press, Palo Alto.

Hopkins, D.M., 1959. Cenozoic history of the Bering land bridge. Science. 129 (3362), 1519-1528, https://www.jstor.org/stable/1757656

Huang, H., He, Q., Kubatko, L.S., Knowles, L.L., 2010. Sources of error inherent in species-tree estimation: impact of mutational and coalescent effects on accuracy and implications for choosing among different methods. Syst. Biol. 59 (5), 573-583, https://doi.org/10.1093/sysbio/syq047

Huang, H., Knowles L.L., 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Syst. Biol. 65 (3), 357-365, https://doi.org/10.1093/sysbio/syu046

Karimi, N., Grover, C.E., Gallagher, J.P., Wendel, J.F., Ané, C., Baum, D.A., 2020. Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). Syst. Biol. 69 (3), 462-478, https://doi.org/10.1093/sysbio/syz073

Kendall, D.G., 1948. On the generalized "birth-and-death" process. Ann. Math. Stat. 19 (1), 1-15, https://www.jstor.org/stable/2236051

Kennett, J.P., 1977. Cenozoic evolution of Antarctic glaciation, the circum-Antarctic Ocean, and their impact on global paleoceanography. J. Geophys. Res. 82 (27), 3843-3860, https://doi.org/10.1029/JC082i027p03843

Kimball, R.T., Oliveros, C.H., Wang, N., White, N.D., Barker, F.K., Field, D.J., Ksepka, D.T., Chesser, R.T., Moyle, R.G., Braun, M.J., Brumfield, R.T., 2019. A phylogenomic supertree of birds. Diversity, 11 (7), 109, https://doi.org/10.3390/d11070109

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56 (1), 17-24, https://doi.org/10.1080/10635150601146041

947 Kuzoff, R.K., Soltis, D.E., Hufford, L., Soltis, P.S., 1999. Phylogenetic relationships within
948 *Lithophragma* (Saxifragaceae): hybridization, allopolyploidy, and ovary diversification.
949 Syst. Bot. 24 (4), 598-615, https://doi.org/10.2307/2419645

950 Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods.
951 9 (4), 357, http://www.nature.com/doifinder/10.1038/nmeth.1923

952 Leaché, A.D., Oaks, J.R., 2017. The utility of single nucleotide polymorphism (SNP) data in
953 phylogenetics. Annu. Rev. Ecol. Evol. Syst. 48, 69-84, https://doi.org/10.1146/annurev-
954 ecolsys-110316-022645

955 Lecaudey, L.A., Schliewen, U.K., Osinov, A.G., Taylor, E.B., Bernatchez, L., Weiss, S.J., 2018.
956 Inferring phylogenetic structure, hybridization and divergence times within Salmoninae
957 (Teleostei: Salmonidae) using RAD-sequencing. Mol. Phylogenet. Evol. 124, 82-99,
958 https://doi.org/10.1016/j.ympev.2018.02.022

959 Leebens-Mack, J.H., Barker, M.S., Carpenter, E.J., et al., 2019. One thousand plant
960 transcriptomes and the phylogenomics of green plants. Nature. 574, 679-685,
961 https://doi.org/10.1038/s41586-019-1693-2

962 Lee-Yaw, J.A., Grassa, C.J., Joly, S., Andrew, R.L., Rieseberg, L.H., 2019. An evaluation of
963 alternative explanations for widespread cytonuclear discordance in annual sunflowers
964 (*Helianthus*). New Phytol. 221 (1), 515-526, https://doi.org/10.1111/nph.15386

965 Lemmon, E.M., Lemmon A.R., 2013. High-throughput genomic data in systematics and
966 phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99-121, https://doi.org/10.1146/annurev-
967 ecolsys-110512-135822

968 Leopold, E.B., Denton, M.F., 1987. Comparative age of grassland and steppe east and west of
969 the northern Rocky Mountains. Ann. Mo. Bot. Gard. 74, 841-867,
970 https://doi.org/10.2307/2399452

971 Léveillé-Bourret, É., Chen, B.H., Garon-Labrecque, M.É., Ford, B.A., Starr, J.R., 2020. RAD
972 sequencing resolves the phylogeny, taxonomy and biogeography of Trichophoreae despite a
973 recent rapid radiation (Cyperaceae). Mol. Phylogenet. Evol. 145, 106727,
974 https://doi.org/10.1016/j.ympev.2019.106727

975 Liston, A., Parker-Defeniks, M., Syring, J.V., Willyard, A. and Cronn, R., 2007. Interspecific
976 phylogenetic analysis enhances intraspecific phylogeographical inference: a case study in
977 *Pinus lambertiana*. Mol. Ecol. 16 (18), 3926-3937, https://doi.org/10.1111/j.1365-
978 294X.2007.03461.x

979 Little Jr, E.L., 1971. Atlas of United States trees. Volume 1. Conifers and important hardwoods.
980 Miscellaneous publication 1146. US Department of Agriculture, Forest Service,
981 Washington, DC.

982 Liu, B.B., Campbell, C.S., Hong, D.Y., Wen, J., 2020. Phylogenetic relationships and
983 chloroplast capture in the *Amelanchier-Malacomeles-Peraphyllum* clade (Maleae,
984 Rosaceae): evidence from chloroplast genome and nuclear ribosomal DNA data using
985 genome skimming. Mol. Phylogenet. Evol. 147, 106784,
986 https://doi.org/10.1016/j.ympev.2020.106784

987 Liu, Y., Johnson, M.G., Cox, C.J., Medina, R., Devos, N., Vanderpoorten, A., Hedenäs, L.,
988 Bell, N.E., Shevock, J.R., Aguero, B., Quandt, D., 2019. Resolution of the ordinal
989 phylogeny of mosses using targeted exons from organellar and nuclear genomes. Nat.
990 Commun. 10 (1), 1-11, https://doi.org/10.1038/s41467-019-09454-w

991 Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46 (3), 523-536,
992 https://doi.org/10.1093/sysbio/46.3.523

993 Manos, P.S., Stanford, A.M., 2001. The historical biogeography of Fagaceae: tracking the
994 tertiary history of temperate and subtropical forests of the Northern Hemisphere. Int. J. Plant
995 Sci. 162 (S6), S77-S93, https://doi.org/10.1086/323280

996 Mao, K., Hao, G., Liu, J., Adams, R.P., Milne, R.I., 2010. Diversification and biogeography of
997 *Juniperus* (Cupressaceae): variable diversification rates and multiple intercontinental
998 dispersals. New Phytol. 188 (1), 254-272, https://doi.org/10.1111/j.1469-8137.2010.03351.x

999 Mao, K., Milne, R.I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R.R., Renner, S.S., 2012.
1000 Distribution of living Cupressaceae reflects the breakup of Pangea. Proc. Natl. Acad. Sci.
1001 U.S.A. 109 (20), 7793-7798.

1002 Massatti, R., Reznicek, A.A., Knowles, L.L., 2016. Utilizing RADseq data for phylogenetic
1003 analysis of challenging taxonomic groups: A case study in *Carex* sect. *Racemosae*. Am. J.
1004 Bot. 103 (2), 337-347, https://doi.org/10.3732/ajb.1500315

1005 Matzke, N.J., 2013a. BioGeoBEARS: BioGeography with Bayesian (and likelihood)
1006 evolutionary analysis in R Scripts. University of California, Berkeley, Berkeley, CA.

1007 Matzke, N.J., 2013b. Probabilistic historical biogeography: new models for founder-event
1008 speciation, imperfect detection, and fossils allow improved accuracy and model-testing.
1009 Front. Biogeogr. 5 (4), 242-248.

1010 Matzke, N.J., 2014. Model selection in historical biogeography reveals that founder-event
1011 speciation is a crucial process in island clades. Syst. Biol. 63 (6), 951-970.

1012 Matzke, N.J., 2018. cladoRcpp v0.15.1: C++ Implementations of Phylogenetic Cladogenesis
1013 Calculations. University of Auckland, New Zealand.

1014 Matzke, N.J., Sidje, R., Schmidt, D., 2019. rexpokit v0.26.6.6: R wrappers for EXPOKIT; other
1015 matrix functions. School of Biological Sciences, University of Auckland, New Zealand.

1016 McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013.
1017 Applications of next-generation sequencing to phylogeography and phylogenetics. Mol.
1018 Phylogenet. Evol. 66 (2), 526-538, https://doi.org/10.1016/j.ympev.2011.12.007

1019 McVay, J.D., Hauser, D., Hipp, A.L., Manos, P.S., 2017. Phylogenomics reveals a complex
1020 evolutionary history of lobed-leaf white oaks in western North America. Genome. 60 (9),
1021 733-742, https://doi.org/10.1139/gen-2016-0206

1022 Miller, M.R., Dunham, J.P., Amores, A., Cresko, W.A., Johnson, E.A., 2007. Rapid and cost-
1023 effective polymorphism identification and genotyping using restriction site associated DNA
1024 (RAD) markers. Genome Res. 17 (2), 240-248, https://doi.org/10.1101/gr.5681207

1025 Miller, R.F., Wigand, P.E., 1994. Holocene changes in semiarid pinyon-juniper woodlands:
1026 response to climate, fire, and human activities in the US Great Basin. BioScience. 44 (7),
1027 465-474, https://doi.org/10.2307/1312298

1028 Mogensen, H.L., 1996. Invited special paper: the hows and whys of cytoplasmic inheritance in
1029 seed plants. Am. J. Bot. 83 (3), 383-404, https://doi.org/10.1002/j.1537-
1030 2197.1996.tb12718.x

1031 Moura, A.E., Shreves, K., Pilot, M., Andrews, K.R., Moore, D.M., Kishida, T., Möller, L.,
1032 Natoli, A., Gaspari, S., McGowen, M., Gray, H., 2020. Phylogenomics of the genus
1033 *Tursiops* and closely related Delphininae reveals extensive reticulation among lineages and
1034 provides inference about eco-evolutionary drivers. Mol. Phylogenet. Evol. 146, 106756.

1035 Neale, D.B., Sederoff, R.R., 1989. Paternal inheritance of chloroplast DNA and maternal
1036 inheritance of mitochondrial DNA in loblolly pine. Theor. Appl. Genet. 77 (2), 212-216.

1037 Near, T.J., MacGuigan, D.J., Parker, E., Struthers, C.D., Jones, C.D., Dornburg, A., 2018.
1038 Phylogenetic analysis of Antarctic notothenioids illuminates the utility of RADseq for

resolving Cenozoic adaptive radiations. Mol. Phylogenet. Evol. 129, 268-279, https://doi.org/10.1016/j.ympev.2018.09.001

Nee, S., May, R.M., Harvey, P.H., 1994. The reconstructed evolutionary process. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 344 (1309), 305-311, https://doi.org/10.1098/rstb.1994.0068

Nevill, P.G., Zhong, X., Tonti-Filippini, J., Byrne, M., Hislop, M., Thiele, K., van Leeuwen, S., Boykin, L.M., Small, I., 2020. Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. Plant Methods. 16 (1), 1-8, https://doi.org/10.1186/s13007-019-0534-5

Paetzold, C., Wood, K.R., Eaton, D., Wagner, W.L., Appelhans, M.S., 2019. Phylogeny of Hawaiian *Melicope* (Rutaceae): RAD-Seq resolves species relationships and reveals ancient introgression. Front. Plant Sci. 10, 1074, https://doi.org/10.3389/fpls.2019.01074

Paradis, E., 1998. Testing for constant diversification rates using molecular phylogenies: a general approach based on statistical tests for goodness of fit. Mol. Biol. Evol. 15 (4), 476-479, https://doi.org/10.1093/oxfordjournals.molbev.a025946

Paradis, E., Schliep, K., 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 35 (3), 526-528, https://doi.org/10.1093/bioinformatics/bty633

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W., Buerkle, C.A., 2012. Genome-wide association genetics of an adaptive trait in lodgepole pine. Mol. Ecol. 21 (12), 2991-3005, https://doi.org/10.1111/j.1365-294X.2012.05513.x

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PloS One. 7 (5), e37135, https://dx.doi.org/10.1371%2Fjournal.pone.0037135

1063 Petit, R.J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., Vendramin, G.G., 2005. Invited
1064 review: comparative organization of chloroplast, mitochondrial and nuclear diversity in
1065 plant populations. Mol. Ecol. 14(3), 689-701, https://doi.org/10.1111/j.1365-
1066 294X.2004.02410.x

1067 Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24 (7),
1068 386-393, https://doi.org/10.1016/j.tree.2009.02.011

1069 Petit, R.J., Hampe, A., 2006. Some evolutionary consequences of being a tree. Annu. Rev. Ecol.
1070 Evol. Syst. 37, 187-214, https://doi.org/10.1146/annurev.ecolsys.37.091305.110215

1071 Phillips, F.J., 1910. The dissemination of junipers by birds. J. For. 8 (1), 60-73.

1072 Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output
1073 analysis for MCMC. R news, 6 (1), 7-11.

1074 Poddar, S., Lederer, R.J., 1982. Juniper berries as an exclusive winter forage for Townsend's
1075 Solitaires. Am. Midl. Nat. 108 (1), 34-40, https://doi.org/10.2307/2425289

1076 Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA
1077 substitution. Bioinformatics, 14 (9), 817-818,
1078 https://doi.org/10.1093/bioinformatics/14.9.817

1079 Rancilhac, L., Goudarzi, F., Gehara, M., Hemami, M.R., Elmer, K.R., Vences, M., Steinfarz, S.,
1080 2019. Phylogeny and species delimitation of near Eastern *Neurergus* newts (Salamandridae)
1081 based on genome-wide RADseq data analysis. Mol. Phylogenet. Evol. 133, 189-197,
1082 https://doi.org/10.1016/j.ympev.2019.01.003

1083 Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior
1084 summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67 (5), 901,
1085 https://dx.doi.org/10.1093%2Fsysbio%2Fsyy032

1086 Razkin, O., Sonet, G., Breugelmans, K., Madeira, M.J., Gómez-Moliner, B.J., Backeljau, T.,
1087 2016. Species limits, interspecific hybridization and phylogeny in the cryptic land snail
1088 complex *Pyramidula*: the power of RADseq data. Mol. Phylogenet. Evol. 101, 267-278,
1089 https://doi.org/10.1016/j.ympev.2016.05.002

1090 Retallack, G.J., 1997. Neogene expansion of the North American prairie. Palaios. 12 (4), 380-
1091 390, https://doi.org/10.2307/3515337

1092 Reveal, J.L., 1980. Intermountain biogeography—a speculative appraisal. Mentzelia. 4, 1-92.

1093 Rieseberg, L.H., Beckstrom-Sternberg, S.M., Liston, A., Arias, D.M., 1991. Phylogenetic and
1094 systematic inferences from chloroplast DNA and isozyme variation in *Helianthus* sect.
1095 *Helianthus* (Asteraceae). Syst. Bot. 50-76, https://doi.org/10.2307/2418973

1096 Rieseberg, L.H., Soltis, D. E., 1991. Phylogenetic consequences of cytoplasmic gene flow in
1097 plants. Evol. Trends Plants. 5, 65-84.

1098 Rieseberg, L.H., Whitton, J., Linder, C.R., 1996. Molecular marker incongruence in plant
1099 hybrid zones and phylogenetic trees. Acta Bot. Neerl. 45 (3), 243-262.

1100 Roch, S., Steel, M., 2015. Likelihood-based tree reconstruction on a concatenation of aligned
1101 sequence data sets can be statistically inconsistent. Theor. Popul. Biol. 100, 56-62,
1102 https://doi.org/10.1016/j.tpb.2014.12.005

1103 Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open
1104 source tool for metagenomics. PeerJ. 4, e2584, https://dx.doi.org/10.7717%2Fpeerj.2584

1105 Romme, W.H., Allen, C.D., Bailey, J.D., Baker, W.L., Bestelmeyer, B.T., Brown, P.M.,
1106 Eisenhart, K.S., Floyd, M.L., Huffman, D.W., Jacobs, B.F., Miller, R.F., Muldavin, E.H.,
1107 Swetnam, T.W., Tausch, R.J., Weisberg, P.J., 2009. Historical and modern disturbance
1108 regimes, stand structures, and landscape dynamics in pinon–juniper vegetation of the

1109 western United States. Rangeland Ecol. Manag. 62 (3), 203-222, https://doi.org/10.2111/08-
1110 188R1.1

1111 Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under
1112 mixed models. Bioinformatics. 19 (12), 1572-1574,
1113 https://doi.org/10.1093/bioinformatics/btg180

1114 Rubin, B.E.R., Ree, R.H., Moreau, C.S., 2012. Inferring phylogenies from RAD sequence data.
1115 PloS One. 7 (4), e33394, https://dx.doi.org/10.1371%2Fjournal.pone.0033394

1116 Salas-Lizana, R., Oono, R., 2018. Double-digest RAD seq loci using standard Illumina indexes
1117 improve deep and shallow phylogenetic resolution of *Lophodermium*, a widespread fungal
1118 endophyte of pine needles. Ecol. Evol. 8 (13), 6638-6651, https://doi.org/10.1002/ece3.4147

1119 Santos, T., Tellería, J.L., Virgós, E., 1999. Dispersal of Spanish juniper *Juniperus thurifera* by
1120 birds and mammals in a fragmented landscape. Ecography. 22 (2), 193-204,
1121 https://doi.org/10.1111/j.1600-0587.1999.tb00468.x

1122 Sauquet, H., Ho, S.Y., Gandolfo, M.A., Jordan, G.J., Wilf, P., Cantrill, D.J., Bayly, M.J.,
1123 Bromham, L., Brown, G.K., Carpenter, R.J. and Lee, D.M., 2012. Testing the impact of
1124 calibration on molecular divergence times using a fossil-rich group: the case of *Nothofagus*
1125 (Fagales). Syst. Biol. 61 (2), 289-313.

1126 Shao, C.C., Shen, T.T., Jin, W.T., Mao, H.J., Ran, J.H., Wang, X.Q., 2019.
1127 Phylotranscriptomics resolves interspecific relationships and indicates multiple historical
1128 out-of-North America dispersals through the Bering Land Bridge for the genus *Picea*
1129 (Pinaceae). Mol. Phylogenet. Evol. 141, 106610,
1130 https://doi.org/10.1016/j.ympev.2019.106610

1131 Snir, S., Rao, S., 2012. Quartet MaxCut: a fast algorithm for amalgamating quartet trees. Mol.
1132 Phylogenet. Evol. 62 (1), 1-8, https://doi.org/10.1016/j.ympev.2011.06.021

1133 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1134 large phylogenies. Bioinformatics. 30 (9), 1312-1313,
1135 https://doi.org/10.1093/bioinformatics/btu033

1136 Stegemann, S., Keuthe, M., Greiner, S., Bock, R., 2012. Horizontal transfer of chloroplast
1137 genomes between plant species. Proc. Natl. Acad. Sci. U.S.A. 109 (7), 2434-2438,
1138 https://doi.org/10.1073/pnas.1114076109

1139 Stephens, J.D., Rogers, W.L., Mason, C.M., Donovan, L.A., Malmberg, R.L., 2015. Species
1140 tree estimation of diploid *Helianthus* (Asteraceae) using target enrichment. Am. J. Bot. 102
1141 (6), 910-920, https://doi.org/10.3732/ajb.1500031

1142 Stephens, M.A., 1974. EDF statistics for goodness of fit and some comparisons. J. Am. Stat.
1143 Assoc. 69 (347), 730-737, https://doi.org/10.2307/2286009

1144 Swenson, N.G., Howard, D.J., (2005). Clustering of contact zones, hybrid zones, and
1145 phylogeographic breaks in North America. Am. Nat. 166 (5), 581-591,
1146 https://doi.org/10.1086/491688

1147 Takahashi, T., Nagata, N., Sota, T., 2014. Application of RAD-based phylogenetics to complex
1148 relationships among variously related taxa in a species flock. Mol. Phylogenet. Evol. 80,
1149 137-144, https://doi.org/10.1016/j.ympev.2014.07.016

1150 Taylor, C.A., 2008. Ecological consequences of using prescribed fire and herbivory to manage
1151 *Juniperus* encroachment. In: Van Auken, O.W. (Ed.), Western North American Juniperus
1152 Communities. Springer, New York, pp. 239-252.

1153 Terry, R.G., 2010. Re-evaluation of morphological and chloroplast DNA variation in *Juniperus*
1154 *osteosperma* Hook and *Juniperus occidentalis* Torr. Little (Cupressaceae) and their putative
1155 hybrids. Biochem. Syst. Ecol. 38 (3), 349-360, https://doi.org/10.1016/j.bse.2010.03.001

1156 Terry, R.G., Nowak, R.S., Tausch, R.J., 2000. Genetic variation in chloroplast and nuclear
1157 ribosomal DNA in Utah juniper (*Juniperus osteosperma*, Cupressaceae): evidence for
1158 interspecific gene flow. Am. J. Bot. 87 (2), 250-258, https://doi.org/10.2307/2656913

1159 Terry, R.G., Pyne, M.I., Bartel, J.A., Adams, R.P., 2016. A molecular biogeography of the New
1160 World cypresses (*Callitropsis*, *Hesperocyparis*; Cupressaceae). Plant Syst. Evol. 302 (7),
1161 921-942.

1162 Tonini, J., Moore, A., Stern, D., Shcheglovitova, M., Ortí, G., 2015. Concatenation and species
1163 tree methods exhibit statistically indistinguishable accuracy under a range of simulated
1164 conditions. PLoS Curr. 7.

1165 Tsitrone, A., Kirkpatrick, M., Levin, D.A., 2003. A model for chloroplast capture. Evolution. 57
1166 (8), 1776-1782, https://doi.org/10.1111/j.0014-3820.2003.tb00585.x

1167 Vasek, F.C., 1966. The distribution and taxonomy of three western junipers. Brittonia. 18 (4),
1168 350-372, https://doi.org/10.2307/2805152

1169 Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., Sivasundar, A.,
1170 Seehausen, O., 2013. Genome-wide RAD sequence data provide unprecedented resolution
1171 of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol.
1172 Ecol. 22 (3), 787-798, https://doi.org/10.1111/mec.12023

1173 Wang, Q., Mao, K.S., 2016. Puzzling rocks and complicated clocks: how to optimize molecular
1174 dating approaches in historical phytogeography. New Phytol. 209 (4), 1353-1358.

1175 Wang, X.Q., Ran, J.H., 2014. Evolution and biogeography of gymnosperms. Mol. Phylogenet.
1176 Evol. 75, 24-40, https://doi.org/10.1016/j.ympev.2014.02.005

1177 Weir, J.T., Schluter, D., 2007. The latitudinal gradient in recent speciation and extinction rates
1178 of birds and mammals. Science. 315 (5818), 1574-1576,
1179 https://doi.org/10.1126/science.1135590

1180 Weisberg, P.J., Lingua, E., Pillai, R.B., 2007. Spatial patterns of pinyon–juniper woodland
1181 expansion in central Nevada. Rangeland Ecol. Manag. 60 (2), 115-124,
1182 https://doi.org/10.2111/05-224R2.1

1183 West, N.E., Tausch, R.J., Rea, K.H., Tueller, P.T., 1978. Phytogeographical variation within
1184 juniper-pinyon woodlands of the Great Basin. Great Basin Naturalist Memoirs. 8 (2), 119-
1185 136, https://www.jstor.org/stable/23376562

1186 Willson, C.J., Manos, P.S., Jackson, R.B., 2008. Hydraulic traits are influenced by phylogenetic
1187 history in the drought-resistant, invasive genus *Juniperus* (Cupressaceae). Am. J. Bot. 95
1188 (3), 299-314, https://doi.org/10.3732/ajb.95.3.299

1189 Willyard, A., Syring, J., Gernandt, D.S., Liston, A., Cronn, R., 2007. Fossil calibration of
1190 molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. Mol.
1191 Biol. Evol. 24 (1), 90-101, https://doi.org/10.1093/molbev/msl131

1192 Wilson, J.S., Pitts, J.P., 2010. Illuminating the lack of consensus among descriptions of earth
1193 history data in the North American deserts: a resource for biologists. Prog. Phys. Geogr. 34
1194 (4), 419-441, https://doi.org/10.1177%2F0309133310363991

1195 Wolfe, J.A., 1964. Miocene floras from Fingerrock wash, southwestern Nevada. US Geological
1196 Survey Professional Paper. 454-N, 1-36.

1197 Wolfe, J.A., 1978. A paleobotanical interpretation of Tertiary climates in the Northern
1198 Hemisphere: Data from fossil plants make it possible to reconstruct Tertiary climatic
1199 changes, which may be correlated with changes in the inclination of the earth's rotational
1200 axis. Am. Sci. 66 (6), 694-703, https://www.jstor.org/stable/27848958

1201 Xiang, Q.P., Wei, R., Zhu, Y.M., Harris, A.J., Zhang, X.C., 2018. New infrageneric
1202 classification of *Abies* in light of molecular phylogeny and high diversity in western North
1203 America. J. Syst. Evol. 56 (6), 562-572, https://doi.org/10.1111/jse.12458

1204 Xie, S., Jialiang, L., Jibin, M., Jingjing, X., Kangshan, M., 2019. The complete chloroplast
1205 genome of *Juniperus squamata* (Cupressaceae), a shrubby conifer from Asian Mountains.
1206 Mitochondrial DNA Part B. 4 (2), 2137-2139.

1207 Xu, T., Abbott, R.J., Milne, R.I., Mao, K., Du, F.K., Wu, G., Zhaxi, C., Liu, J., 2010.
1208 Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-
1209 Tibetan Plateau and adjacent regions. BMC Evol. Biol. 10 (1), 194.

1210 Zanoni, T.A., Adams, R.P., 1976. The genus *Juniperus* in Mexico and Guatemala: Numerical
1211 and chemosystematic analysis. Biochem. Syst. 4 (3), 147-158.

1212 Zhu, A., Fan, W., Adams, R.P., Mower, J.P., 2018. Phylogenomic evidence for ancient
1213 recombination between plastid genomes of the *Cupressus-Juniperus-Xanthocyparis*
1214 complex (Cupressaceae). BMC Evol. Biol. 18 (1), 137.

1215

Figure Legends

Figure 1: The serrate leaf junipers are distributed across arid and semi-arid regions of the western United States, Mexico, and Guatemala. Colors representing sampling localities correspond with those designating serrate juniper clades in the phylogenies of Figures 2-4. Outgroup specimens are not shown in map. Map created with ArcGIS Pro 2.4.0 (http://www.esri.com).

Figure 2: Phylogenetic analyses of ddRADseq data with maximum likelihood (left) and SVDquartets (right) provide largely consistent topologies for the serrate juniper clade and its relatives. Nine monophyletic clades resolved by both methods are indicated by colored boxes. Bootstrap support values are reported for all nodes. Branch lengths are not meaningful for the SVDquartets tree.

Figure 3: Comparison of the maximum likelihood ddRADseq tree (left) to a Bayesian cpDNA tree (right) reveals five clear instances of discordance, indicated by dashed arrows. Nine low-level clades resolved with ddRADseq data (Fig. 2) are indicated by colored boxes.

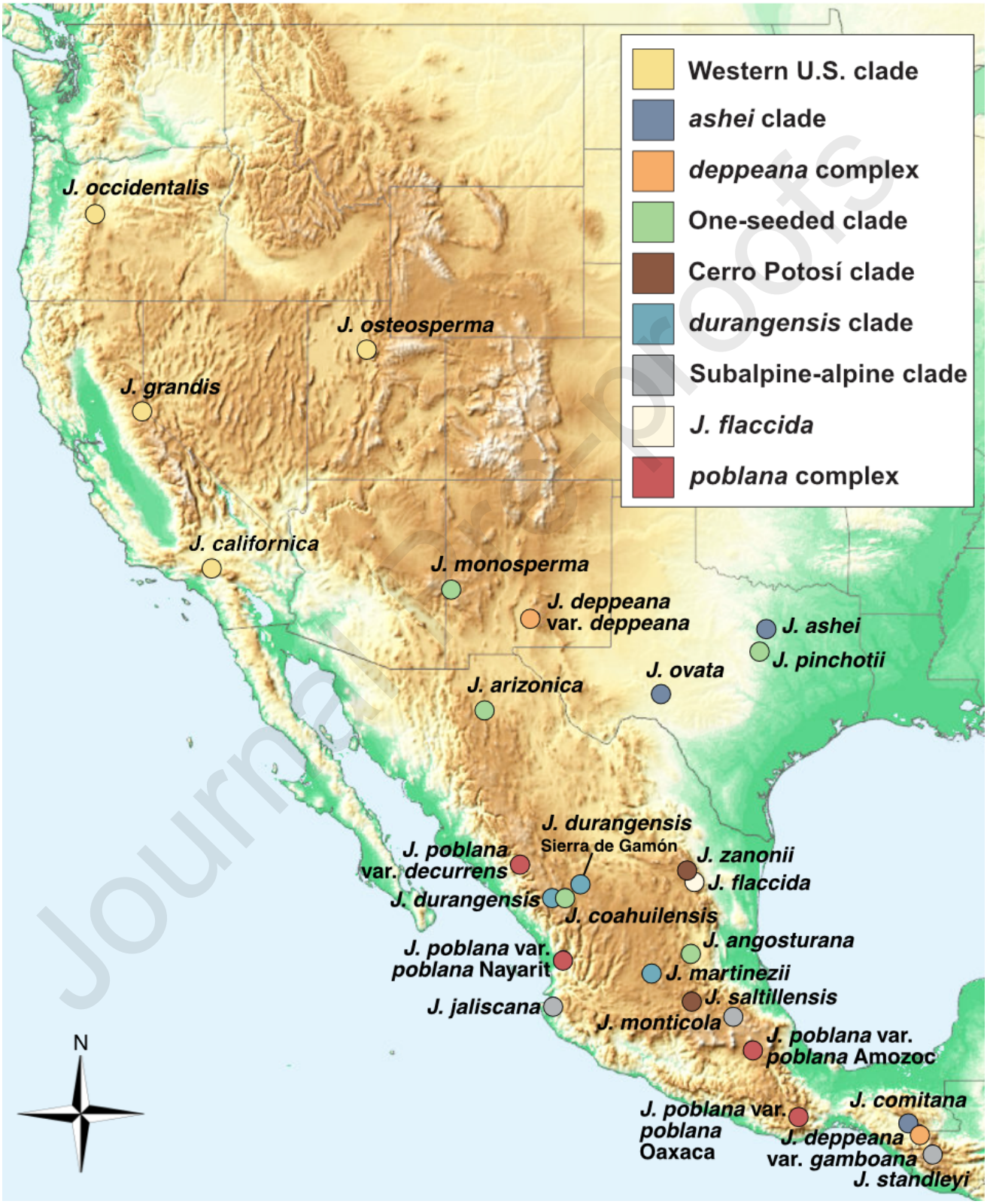
Figure 4: (A) Maximum clade credibility tree (MCC) from analyses in `RevBayes` of the serrate leaf juniper clade calibrated with fossil evidence. Smooth leaf juniper outgroup taxa were excluded from the figure for clarity. Asterisks identify two of the three calibration nodes (the calibrated root node is not shown because it was pruned prior to visualization; see Methods and Table S2 for details). All nodes received greater than 99% Bayesian posterior support. The nine low-level clades resolved in RAxML and SVDquartets phylogenetic analyses of the full set of

ddRADseq data (Fig. 2) are indicated by colored boxes. (B) Lineage through time plot for the serrate juniper clade generated with the Bayesian MCC tree in panel A. Grey dashed line represents linear diversification rate through time given the estimated crown age of the serrate clade and the extant number of species.

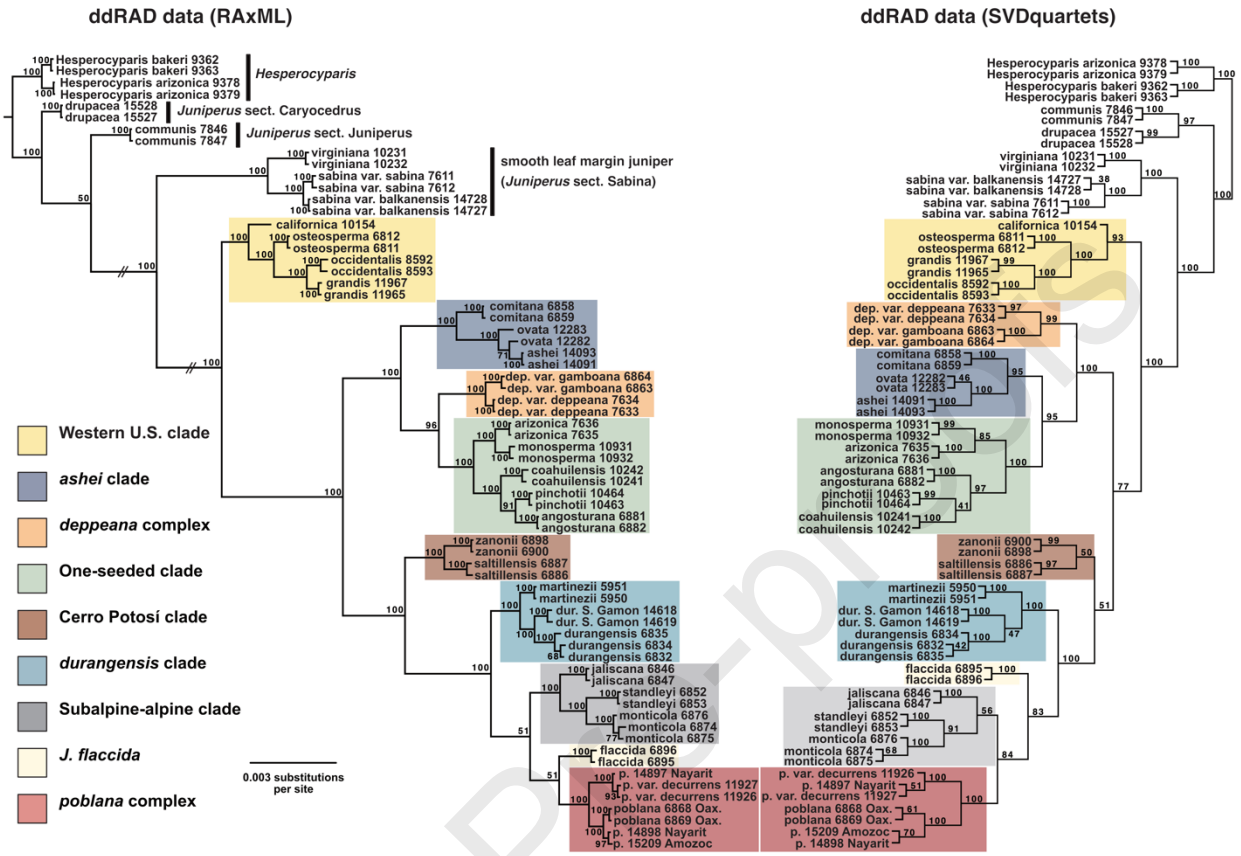
Figure 5: Ancestral ranges for the serrate junipers based on a dated phylogeny produced with `RevBayes` and the DIVALIKE model in `BioGeoBEARS`. The map inset shows the delineation of five operational areas (A, western U.S.; B, central U.S.; C, eastern U.S.; D, northern/central MX; E, southern MX), which, along with information of species distributions, informed the geographic ranges assigned to each species and model-based estimates of ancestral ranges. Pie charts at each node represent the marginal probabilities for each range estimated with maximum likelihood, where the colors of the pie sectors either represent single ancestral ranges indicated within the map inset or a possible combination of two ancestral ranges, in which case a novel color was chosen.

Figures

Figure 1



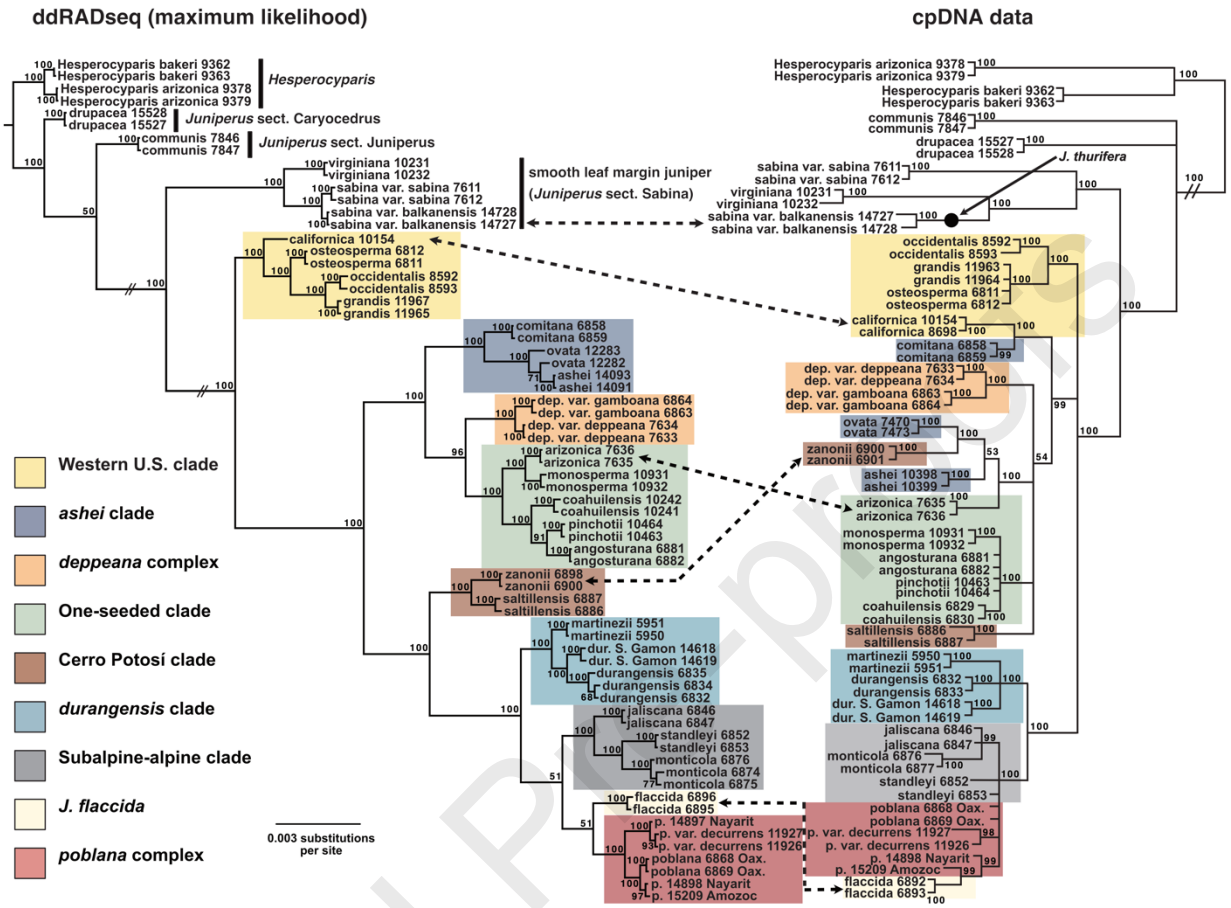
1259 **Figure 2**



1260

1261

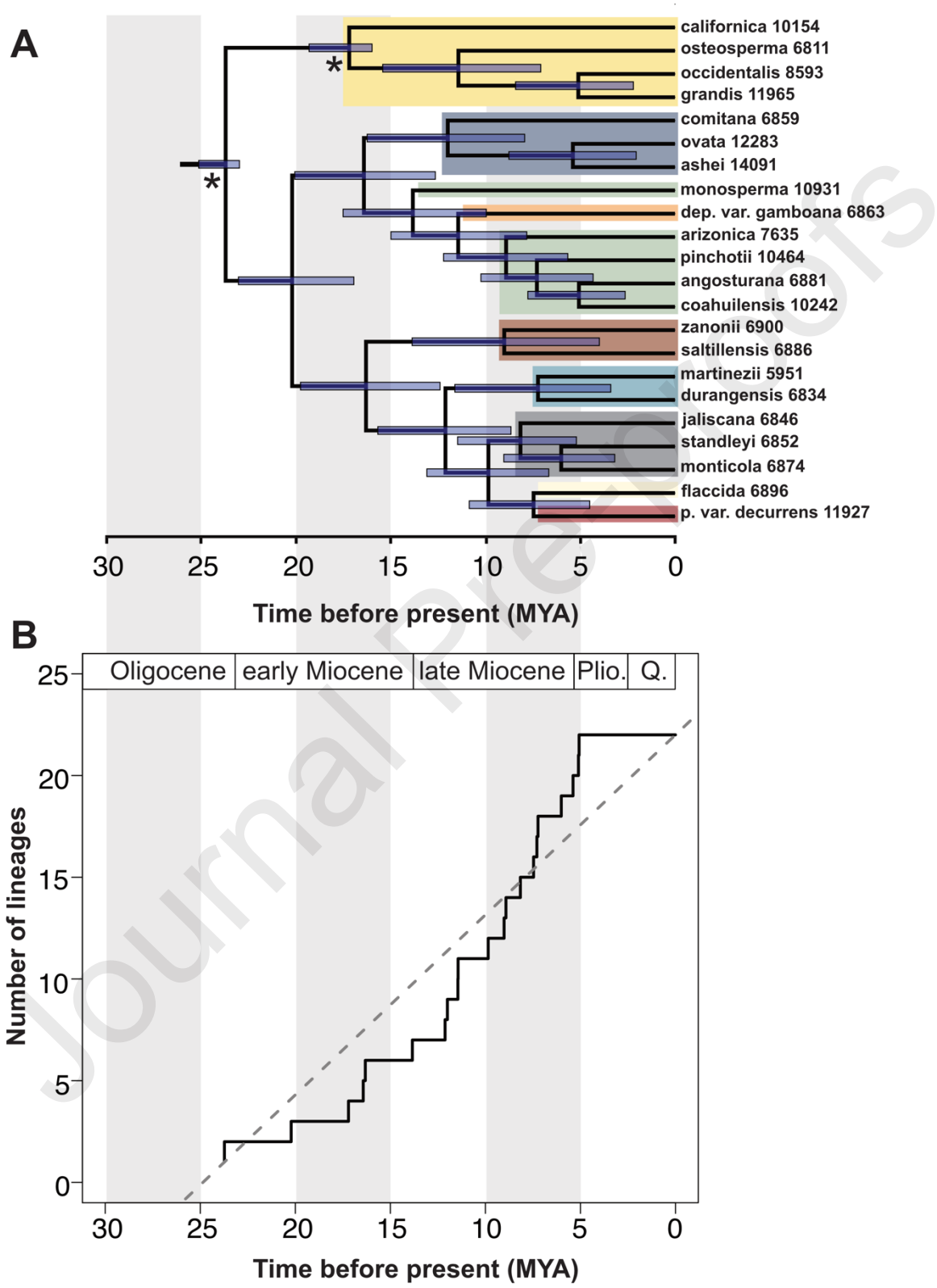
1262 **Figure 3**



1263

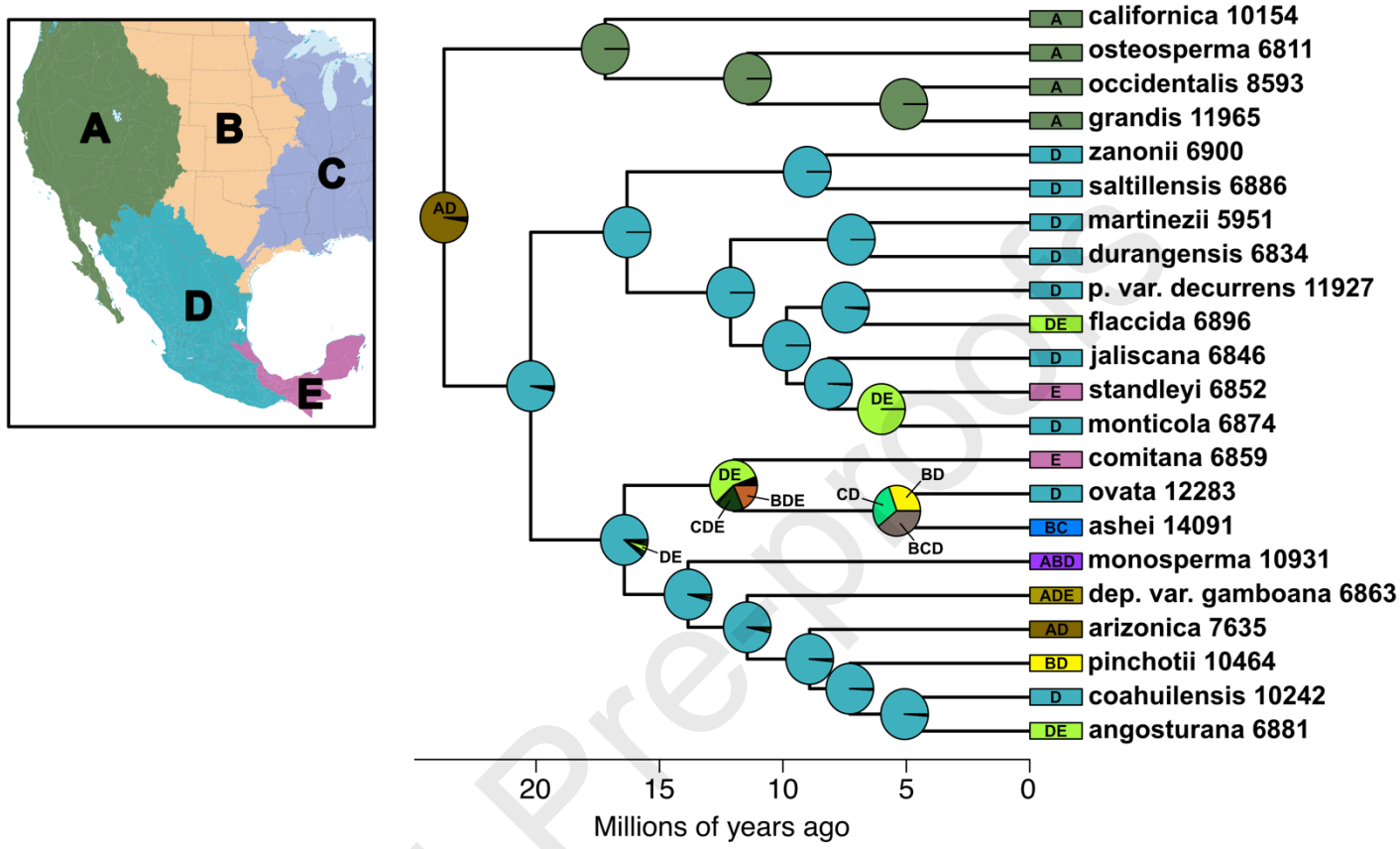
1264

1265 **Figure 4**



1266

Figure 5



Highlights

- Serrate junipers are ecologically significant trees of western North America (76 characters)
- RADseq data produced strongly resolved phylogeny for North American serrate junipers (84 characters)
- Comparison of RADseq and cp phylogenies revealed cases of strong discordance (76)
- Serrate junipers originated in Oligocene and diversified rapidly in the late Miocene (84 characters)

1282