Botanical Sciences 103 (Special Issue): S113-S131. 2025
DOI: 10.17129/botsci.3688
Received: March 4, 2025, Accepted: May 3, 2025
On line first: September 1, 2025

Taxonomy and Floristics / Taxonomía y Florística

Divergence within diversity in the endemic, long-lived conifer *Pinus greggii* from the Sierra Madre Oriental, Mexico

Xitlali Aguirre-Dugua[1,2*], David S. Gernandt[1], Alejandra Moreno-Letelier[3]

[1] Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico.
[2] Programa Investigadoras e Investigadores por México, Secretaría de Ciencia, Humanidades, Tecnología e Innovación, Mexico City, Mexico.
[3] Jardín Botánico, Instituto de Biología, Universidad Nacional Autónoma de México, Mexico City, Mexico.
*Corresponding author: xitlali.aguirre@secihti.mx

Abstract

Background: Mexico's physiography has been recognized as a driver of *Pinus* diversification. This study analyzes phylogeographical patterns of the two varieties of *P. greggii* to cast light on genomic differentiation in allopatry.

Questions: What is the role of geographic barriers in the diversification and genomic differentiation of a highland conifer?

Studied species: *Pinus greggii* Engelm. ex Parl.

Study site and dates: Sierra Madre Oriental, Mexico, 2015-2017.

Methods: We used 481 nuclear SNPs and 73 non-coding plastome regions to evaluate genetic diversity and structure. Population relationships and demography were estimated using maximum likelihood and coalescent methods under an isolation with migration model and interpreted considering ecological niche modeling under present, Last Glacial Maximum and Last Interglacial conditions.

Results: Admixture analyses based on nuclear data show that there are two region-specific clusters and one shared cluster, and that 47 % of SNPs are shared between varieties. Our data support a fragmentation event at the onset of the Last Interglacial (126ky-148ky BP), resulting in a smaller effective population size in the south. Cryptic breaks, and ecological niche models agree with secondary contact during glacial periods. Divergence has been accompanied by population equilibrium as shown by non-significant Tajima's *D* in plastome sequences.

Conclusions: Our results support that lineage divergence in *Pinus* is accompanied by the accumulation of genome-wide differences, while gene flow produced asymmetrical patterns of allele sharing. Our study highlights the effectiveness of mountain-valley systems in isolating populations and points to the need to establish conservation measures for southern *P. greggii* populations.

Keywords*:* allopatric speciation, *Australes*, phylogeography, speciation, retention of ancient polymorphisms.

Resumen

Antecedentes: México tiene una fisiografía compleja reconocida como motor de diversificación del género *Pinus*. Este estudio analiza los patrones filogeográficos de dos variedades de *P. greggii* para entender su diferenciación genómica en alopatría.

Preguntas: ¿Qué papel juegan las barreras geográficas en la diversificación y diferenciación genómica de una especie de montaña?

Especies de estudio: *Pinus greggii* Engelm. ex Parl.

Sitio y años de estudio: Sierra Madre Oriental, 2015-2017.

Métodos: Usamos 481 SNPs nucleares y 73 regiones no codificantes del plastoma para evaluar la diversidad y estructura genética. Se estimaron relaciones entre poblaciones y parámetros demográficos usando máxima verosimilitud y coalescencia bajo aislamiento con migración, y fueron interpretados considerando la modelación de nicho presente, en el Último Interglacial y Último Máximo Glacial.

Resultados: El análisis de STRUCTURE con datos nucleares reveló dos grupos específicos a cada región, y un grupo compartido; un 47 % de SNPs fueron compartidos entre variedades. Nuestros datos apoyan un evento de fragmentación a principios del último interglacial (126ky-148ky AP), que resultó en menores tamaños efectivos poblacionales en el Sur, así como contacto secundario durante periodos glaciales. La divergencia estuvo acompañada por estasis poblacional como sugiere una *D* de Tajima no significativa en secuencias de plastoma.

Conclusiones: Durante la divergencia de *P. greggii* hubo acumulación de diferencias a lo largo del genoma, mientras que el flujo génico produjo patrones asimétricos de alelos compartidos. Nuestro estudio documenta la efectividad de sistemas montañosos en aislar poblaciones, y apunta medidas de conservación para el linaje sur de *P. greggii*.

Palabras clave: *Australes*, especiación, especiación alopátrica, filogeografía, retención de polimorfismos ancestrales.

This is an open access article distributed under the terms of the Creative Commons Attribution License CCBY-NC (4.0) international.
https://creativecommons.org/licenses/by-nc/4.0/

The Sierra Madre Oriental (SMO), located in northeastern Mexico, constitutes one of the five mountain ranges of the Mexican Transition Zone, where the Nearctic and Neotropical biogeographic realms meet (Halffter & Morrone 2017). The temperate forests of the SMO are home to 17 species of *Pinus*, representing one of the main centers of pine diversity in Mexico, which is the world's center of diversity of the genus (Styles 1993).

Phylogenetic relationships and geographical distribution of *Pinus* species found in the SMO strongly suggest that this mountain range represents a corridor along which pines migrated to Mexico from the Rocky Mountains (Farjon 1996, Perry *et al.* 1998). Climatic changes of the Eocene very likely triggered the expansion and diversification of the genus in North America (Millar 1993). Later, climatic fluctuations associated with glacial cycles during the Pleistocene resulted in major distribution shifts of plant populations, leading to the modern structure of temperate plant communities in Mexico (Metcalfe 2006). The great diversity harbored by *Pinus* in the SMO can therefore be explained by the mixed presence of ancient and recently diverged taxa (Farjon 1996).

Temperate forests found in the SMO are characterized by occurring in high "island mountain ranges" surrounded by contrasting vegetation types and are characterized by high endemism (Farjon 1996, Farjon & Filer 2013), suggesting geographical isolation is an important factor enhancing allopatric speciation processes. The Pánuco River Basin is the most prominent natural barrier along the SMO (Figure 1) and has been recognized as a major boundary driving vicariant divergence processes in several plant and animal species (Bryson *et al.* 2011, Ruiz-Sanchez *et al*. 2012, Ferro *et al*. 2017, Suárez-Mota & Villaseñor 2017, Sánchez-Acevedo *et al*. 2023).

Given that the major mountain systems of the Mexican Transition Zone constitute biodiversity hotspots for highland temperate taxa (Torres-Miranda *et al*. 2011, Farjon & Filer 2013), studies addressing the effect of geographical barriers promoting lineage diversification are needed to identify the evolutionary drivers of pine diversity in this geologically and climatically complex region (Halffter & Morrone 2017). *Pinus* lineage divergence at the macro-micro evolutionary boundary remains poorly understood due to high levels of nuclear and organellar allele sharing among species (Hernández-León *et al.* 2013, Willyard *et al.* 2016, 2021, Zhou *et al.* 2017). Incomplete lineage sorting and gene flow among populations have been attributed to large effective population sizes and long lifespan (Petit & Hampe 2006) as well as wind-mediated pollination (effective pollen-mediated gene flow can attain distances up to 10^2 km; Robledo-Arnuncio 2011) and long-distance dispersal of seeds achieved by wind or by birds.

Pinus greggii Engelm. ex Parl. is a species endemic to the SMO, growing 1,500 meters above sea level. *Pinus greggii* belongs to subsection *Australes* Loudon, a monophyletic group comprising approximately 27 species exclusively found in North and Central America and the Antilles (Gernandt *et al*. 2005, Jin *et al.* 2021), whose rapid diversification has been dated to the early Miocene, which spans approximately from 5.33 to 23 million years ago (Saladin *et al.* 2017, Jin *et al.* 2021).

Populations of *Pinus greggii* are found in two disjunct areas north and south of the SMO, separated by ca. 370 km and isolated from each other by the Pánuco River Basin (Figure 1). Southern populations have been described as a separate variety, *Pinus greggii* var. *australis*, based on differences in needle and seed morphology (Donahue & Lopez-Upton 1996, 1999). Populations are commonly scattered, with individuals interspersed with other conifer and broad-leaved species (*e.g.*, *Abies*, *Quercus*). Declining population sizes related to timber exploitation, a reduced distribution range and fragmented populations have led to inclusion of *P. greggii* on the Red List from the International Union for the Conservation of Nature (IUCN), where it is considered *Vulnerable*. Because declining populations can face the erosion of their genetic variability and the lowering of fitness due to inbreeding depression, genetic data on the current diversity levels of the species, as well as the estimation of key demographic parameters -such as effective population sizes and migration rates- can assist the design of conservation and monitoring strategies (Amos & Balmford 2001, Allendorf *et al.* 2010). Identifying intraspecific gene pools is also key for the *ex situ* management and propagation of germplasm with potential for timber production, both in Mexico and in other countries (Dvorak 2012, Dvorak *et al.* 1996).

Here, we aim to analyze the phylogeographical patterns of *P. greggii* within the SMO as a case study for characterizing gene pool differentiation and diversification processes in subsection *Australes*, a rapidly evolving *Pinus*

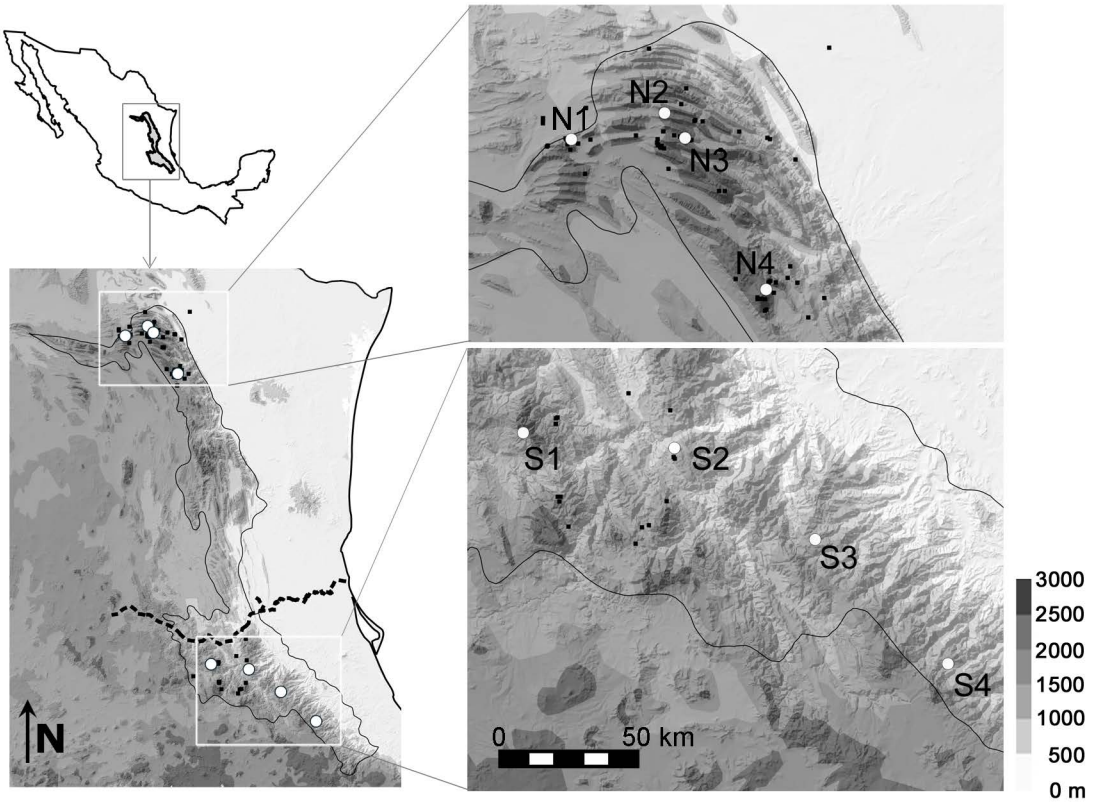

Figure 1. Location of studied *Pinus greggii* populations in Mexico (white points; Table S1). Black points show Global Biodiversity Information Facility records (GBIF.org 2025). The thin black line defines the Sierra Madre Oriental. Pointed line shows the Pánuco River and its main effluents. Elevational classes are shown in grey.

lineage. Given that the pine genome has a large size with highly repetitive elements (ca. 23 Gbp; Neale *et al*. 2014), we used an exon capture approach to obtain a high number of nuclear SNPs and several non-coding regions of the plastid genome.

Our aims are to: 1) Cast light on genomic differentiation during lineage divergence of pines, comparing nuclear to plastid DNA data to identify contrasting patterns of gene flow and retention of ancient polymorphisms; 2) evaluate the role of geographical barriers as diversity drivers in highland taxa from the SMO, under a historical framework of distribution shifts and demographic attributes of populations, and 3) provide baseline data on genetic diversity of this species for the purposes of conservation and timber production.

Materials and methods

Study site. The SMO pine and mixed forests develop in the highlands, whereas eastern slopes support humid montane forests, and western slopes support deciduous tropical forests and xerophytic shrublands (Espinosa *et al.* 2004). Northern *P. greggii* sites have average annual temperatures near 14 °C, annual rainfalls of 650 mm, and neutral to slightly alkaline soils, whereas southern sites have an average annual temperature of 17 °C, annual rainfall of 1,200 mm, and mainly acidic soils (Donahue & Lopez-Upton 1996).

Sampling. A total of 39 individuals from eight populations were analyzed (Figure 1, Table S1). Although this number of individuals might seem low, Aguirre-Liguori *et al.* (2017) showed that parameters of genetic diversity such as *Hs* and F_{ST} are not strongly affected by the number of individuals sampled but depend more heavily on the number of populations sampled. In our study, the eight populations sampled cover the entire range of distribution of the species (see Figure 1).

An individual from the state of Guerrero, outside of the natural species distribution, was also included for evaluating the potential of our data for assigning individuals to known populations. Four *Pinus patula* individuals were included to identify shared polymorphisms. Nuclear data show that *P. greggii* is sister to *P. patula* and *P. tecunumanii* (Cruz-Nicolás *et al*. 2024), whereas plastome data suggest it is sister to *P. lumholtzii* and *P. chihuahuana* (Gernandt *et al*. 2018). Voucher numbers of all specimens studied are included in Table S2.

Nuclear and plastome variant detection. To avoid the highly repetitive sequences of the *Pinus* genome, we used the Hyb-Seq approach, which consists of capturing specific exons of single-copy genes with RNA probes. The method, including selection of reference genes, total genomic DNA extraction, library preparation, exon capture, and sequencing of 713 putative single-copy genes, has been described in Gernandt *et al*. (2018).

Paired-end Illumina sequence reads were quality filtered with GENEIOUS v. 9.1.5 (Kearse *et al*. 2012): duplicate reads were removed, ends trimmed (5' and 3', error probability limit = 0.01), and reads < 50 bp filtered out, keeping only reads with mates. After quality trimming, reads of a haploid *P. greggii* sample (extracted from megagametophyte tissue) were mapped to the reference sequences to identify genes with poor assemblies, annotating intron regions and identifying genes with potential paralogs, all of which were excluded (see details in Supplementary Material 2).

Single nucleotide polymorphisms (SNPs) were called with GENEIOUS v. 9.1.5 with the following parameters: minimum coverage 20×, minimum variant frequency 0.3, maximum variant *P*-value 10^{-7}, minimum 65 % strand bias *P*-value 10^{-3}. Variant frequency was used as an indicator of heterozygous/homozygous status (0.3-0.7 *vs*. > 0.7, respectively). With this stringent procedure, we kept a final set of 415 loci, all of them represented in all the individuals sampled (*i.e.*, no missing data).

SNPs under selection were identified (and then excluded) with *F*-statistics, under a hierarchical island model of two groups (*i.e.*, two varieties), and 20,000 simulations with ARLEQUIN v. 3.5.1.3 (Excoffier *et al.* 2005). Diploid genotypes were phased using PHASE v. 2.1 (Stephens & Scheet 2005) with default parameter values, using haploid samples for specifying known phases. Our final dataset consisted of 481 SNPs (mean = 1.16 SNPs/locus).

The size of our SNP dataset might seem low when compared to most current genomic studies based on high throughput genome sequencing (such as *genotyping by sequencing* or GBS), which are characterized by the sequencing of hundreds of thousands of anonymous loci dispersed across the genome. In the case of pines, genome-wide high-throughput sequencing is unfeasible because of their huge genome size and the highly repetitive sequences that encompass 70 to 80 % of the entire genome (Niu *et al.* 2022). Exon capture approaches, focused on a smaller set of specific but highly informative genes, have therefore been developed as an alternative for studies on *Pinus* phylogeography and systematics (Peláez *et al.* 2020, Willyard *et al.* 2021, Cruz-Nicolás *et al.* 2024); they also represent an interesting avenue for future genomic research in general (Kesälahti *et al*. 2024). In a final stage, SNP coding was modified to reflect ancestral/derived states using *P. pringlei* as an outgroup (Gernandt *et al*. 2018, Cruz-Nicolás *et al.* 2024, Supplementary Material 2).

Plastid reads were assembled over the *P. greggii* reference plastome NC_035947.1 (Aguirre-Dugua & Gernandt 2017). Variant detection in 73 non-coding regions was based on a high-quality filter with the same parameters as previously described, excluding microsatellite length variants (see details in Supplementary Material 2). The final matrix had a total length of 38,301 bp. Unique plastid haplotypes were identified using DnaSP v. 5.10 (Librado & Rozas 2009), excluding invariant sites and treating indels of more than 1 bp as one single mutational event (Müller 2006).

Population diversity and historical demography analysis. For nuclear SNP data, we assessed genetic diversity through the expected heterozygosity index (*Hs;* Nei 1987) and population structure through Wright's F_{IS} and F_{ST} indices with an Analysis of Molecular Variance (AMOVA) in ARLEQUIN v. 3.5.1.3 (Excoffier *et al.* 2005).

Genetic differentiation was investigated using STRUCTURE v. 2.3 (Pritchard *et al*. 2000). We performed 10 replicate runs (100,000 burnin, 100,000 steps) for each K from $K = 1$ to $K = 10$ based on the Admixture and Correlated Allele Frequencies models and a lambda value of 0.6181 calculated from the dataset. The suitable number of clusters

was selected according to ΔK (Evanno *et al*. 2005) and memberships across runs were integrated with CLUMPP v. 1.1.2 (Jakobsson & Rosenberg 2007). Results from this method were compared to a centered non-scaled Principal Components Analysis (PCA, in the R package '*adegenet*'; Jombart & Ahmed 2011). SNP alleles that contributed most to the PCA structure were identified as those in the fourth quartile of allele loadings, and their distribution within the species and with the outgroup was represented through a Venn diagram (with EULERAPE v. 2.0.3; Micallef & Rodgers 2014).

The best geographical grouping of populations (from k = 2 to 7) was identified with SAMOVA v. 2.0 (Spatial Analysis of MOlecular VAriance; Dupanloup *et al*. 2002) based on the maximization of the F_{CT} index. For every k, 10,000 steps were performed during the annealing process for each of 100 initial random configurations and additional parameters at default values. The best grouping was selected where F_{CT} reached a plateau, and the number of single-population groups was minimized. F_{CT} significance was tested through 1,000 permutations of populations among groups. Because we expect the SMO to act as a driver for population divergence, we tested the presence of isolation by landscape resistance (Cushman *et al.* 2006). A first null model of simple isolation by distance across a uniform landscape was evaluated with a Mantel test (Wright 1943) between the matrix of genetic distances (with Nei's (1972) distance) and the matrix of linear distances among populations (calculated with the Haversine method). Our second model proposed that the geographical barrier, expressed through topographic distance (*i.e.*, the shortest topographic path among points considering changes in elevation), exerts additional resistance to gene flow. This second model was evaluated with a partial Mantel test between the matrix of genetic distances and the matrix of topographic distances, while controlling for simple linear distance. Distance matrices and the corresponding tests (with significance assessed through 999 permutations) were obtained with R packages '*adegenet*', '*vegan*', '*topoDistance*', and '*geosphere*' (Jombart & Ahmed 2011, Oksanen *et al*. 2016, Wang 2020, Hijmans 2024).

Finally, the SNPs matrix was used to compute the joint (2D) unfolded site-frequency-spectrum (SFS) between varieties with ARLEQUIN v. 3.5.1.3. The SFS was then used under a coalescent framework to infer the main demographic parameters associated with the divergence of the two disjunct areas following an isolation with migration (IM) model. The composite-likelihood method implemented in *FASTSIMCOAL2* v. 26 (Excoffier *et al*. 2013) was applied to estimate effective population sizes, migration rates, and time to divergence. Simulations of the expected SFS under the IM model were run with the following options: -N 100,000 (number of coalescent simulations), -L 40 (number of expectation-maximization cycles), -M 0.001 (minimum relative difference in parameters between two iterations to stop EM cycles), and -C 1 (minimum SFS entry counts). The parameter values were sampled from wide search ranges with uniform and log-uniform distributions (N = 100 to 100,000 individuals, $Nm = 1 \times 10^{-2}$ to 20, $Tdiv$ = 100 to 20,000 generations ago). To determine the parameter estimates leading to the global maximum likelihood solution, 500 independent runs were performed. Estimates of time in generation units were translated to calendar years by using a generation length of 20 years. The closely related *P. elliottii*, *P. taeda*, and *P. palustris* (subsection *Australes*) can produce their first strobili at 5-10 years of age (Dorman 1976) and thus are estimated to have attained full reproductive capacity at 20 years of age.

For the plastid data, genetic diversity was quantified using Nei's haplotypic index h (Nei 1987) and nucleotide diversity at the sequence level (π; Tajima 1983, Nei 1987) with ARLEQUIN v. 3.5.1.3. Kin relationships were examined through a statistical parsimony network using TCS v. 1.21 with a 90 % connection probability, and gaps treated as a fifth character (Clement *et al.* 2000). Phylogeographic subdivision was evaluated through G_{st} and N_{st} indices (with PERMUT & CPSSR v. 2.0; Pons & Petit 1996), using 1,000 permutations to test significance.

To test if data fit a demographic expansion model, we used the distribution of pairwise differences between plastid haplotypes (Rogers & Harpending 1992) using a non-parametric bootstrap of 10,000 repetitions with the sum of squared differences (SSD) as the test statistic (Schneider & Excoffier 1999) with ARLEQUIN v. 3.5.1.3. We also calculated Fu's F_S (Fu 1997) and Tajima's D (Tajima 1989) based on full-length sequences and assessed their significance by generating 1,000 random samples under the hypothesis of selective neutrality and population equilibrium, using coalescent simulations with ARLEQUIN v. 3.5.1.3.

Ecological Niche Modelling. Climate niche models for present, Last Glacial Maximum (LGM, about 22,000 years before present) and Last Interglacial (LIG, 120,000 - 140,000 years before present) were obtained with MAXENT v. 3.4.1 (Phillips *et al*. 2006). For LGM climate layers, we used the CCSM4 circulation model; Gent *et al*. 2011). For LGM climate layers, we used the CCSM4 (Gent *et al.* 2011) circulation model because it has proven to have a better performance for Mexico (Conde *et al*. 2011), and Last Interglacial Maximum layers were based on Otto-Bliesner *et al.* (2006). All environmental layers were obtained from the WorldClim database v. 1.4 (Hijmans *et al.* 2005). Nineteen bioclimatic variables were tested for correlations and information content with a Principal Component Analysis (JMP 9; www.jmp.com) and only six variables that incorporated most of the variance were kept to avoid model overfitting. The variables used for models were: Mean annual temperature (Bio1), Isothermality (Bio3), Annual precipitation (Bio12), Precipitation in the wettest month (Bio16), Precipitation of the warmest quarter (Bio18), and Precipitation of the coldest quarter (Bio19). The models were run with 10 replicates each, using a bootstrap validation method, with no extrapolation, a regularization multiplier of 2, and 1,000 iterations per run. We used a random seed and 20 % sampling validation. The resulting rasters were cropped using a probability cutoff of 10 %, and the presence pixels were transformed into binary data. Overlap between varieties was obtained using the Raster Calculator option in QGIS v. 2.2.0 (qgis.osgeo.org).

Results

Nuclear variation. Northern and southern regions had similar expected heterozygosity (Table 1), but the south had fewer private alleles (72 *vs*. 111, respectively; Figure 2A). Significant amounts of inbreeding were detected in the species and within regions (F_{IS} = 0.34-0.35). Among-population differentiation was stronger in the north according to F_{ST}, but isolation-by-distance was not significant. In contrast, there was less structuring and a significant isolation-by-distance pattern in the south (Table 1).

When analyzed individually, 323 SNP alleles (67 %) were exclusive to *Pinus greggii* (*i.e.*, not shared with *P. patula*), and from these, 183 (57 % of *P. greggii* SNP alleles) were found exclusively in one of the two regions (Figure 2A). Approximately 33 % of all detected SNPs (158 alleles) were shared with the outgroup, most (128 alleles, 27 %) with *P. greggii* at the species level and a minor proportion with only one of the two regions (6 % of the entire SNP dataset; Figure 2A).

The STRUCTURE analysis identified $K = 3$ clusters (Figure S1). Two corresponded closely to the northern and southern regions (colored in green and red, respectively, in Figure 2B), whereas a third cluster (blue) was well represented in N4, and in some southern individuals. The *P. greggii* individual sampled in Guerrero displayed an admixed genotype, with a 0.55 membership probability to the southern red cluster (Figure 2B).

In the PCA, a clear separation of northern and southern populations was observed in the principal component (PC) 1, while the PC2 revealed within-region variability, especially for northern samples. The distribution of samples in the PCA plot agreed with the $K = 3$ clusters identified in the STRUCTURE analysis (Figure 2B, C).

From the 120 alleles with high weights in the PCA, most (57.5 %) were exclusive to the species (*i.e.*, not shared with the outgroup) and distributed both in the northern and the southern regions. Strikingly, more than one third (36.6 %) of high-loading alleles were shared with the outgroup, whereas only 11 % were private to one of the two regions (Figure 2C).

The maximum likelihood estimates of demographic parameters under the IM model (MaxEstLhood = -1067.228) included an estimated effective ancestral population size of Ne = 158,232 and, after divergence, of Ne = 65,594 for the northern area and Ne = 23,381 for the southern area. Time to divergence was estimated at 6,311 generations ago. Migration rates after the divergence were estimated at $m_{N,S} = 5.12 \times 10^{-5}$ and $m_{S,N} = 2.75 \times 10^{-7}$ in north-to-south and south-to-north directions, respectively (Figure 3A). When translated to calendar years, divergence between regions dates back well before the LGM, between 126,220- and 148,680-years BP. When considering northern population N4 as part of the southern lineage (as could be interpreted by its differential genetic composition in the STRUCTURE analysis shown in Figure 2B and intermediate position in PCA, Figure 2D), estimated demographic parameters changed marginally and maintained a similar order of magnitude (Figure 3B).

Table 1. Diversity and differentiation estimate of *Pinus greggii* based on 481 nuclear SNPs.

	Population	*n*	*Private alleles*	*Hs* (sd)	F_{IS}	F_{ST}	IBD
N1	El Diamante, Coahuila	2	1	0.079 (0.270)			
N2	La Carbonera, Coahuila	7	18	0.235 (0.234)			
N3	Los Lirios, Coahuila	6	27	0.244 (0.219)			
N4	Cerro Potosí, Nuevo León	10	11	0.189 (0.215)			
Total North		25	111	0.288 (0.150)	0.350**	0.135***	$r = 0.35$
S1	Cerro La Pingüica, Querétaro	6	5	0.202 (0.230)			
S2	Laguna Seca, Hidalgo	8	21	0.198 (0.214)			
S3	Jalamelco, Hidalgo	8	11	0.172 (0.224)			
S4	Cumbre de Muridores, Hidalgo	4	1	0.160 (0.236)			
Total South		26	72	0.283 (0.155)	0.344**	0.114***	$r = 0.74$*
Total species		51	323	0.232 (0.158)	0.347***	0.196***	$r = 0.43$**

Notes: *n*: haplotypic sample size; *Hs*: expected heterozygosity; F_{IS}: inbreeding index; F_{ST}: population differentiation index; IBD: isolation by distance. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

The analysis of landscape resistance to gene flow revealed that there is a significant effect of the SMO as a geographical barrier for promoting divergence among populations. The null model of simple isolation by distance was significant (Mantel's $r = 0.319$, $P < 0.05$), but when factoring out the effect of linear distance, our alternative model revealed a higher correlation between genetic and topographic distance (Mantel's $r = 0.6032$, $P < 0.05$).

The *SAMOVA* revealed a maximum differentiation at $k = 3$ groups ($F_{CT} = 0.101$, $P < 0.01$), which corresponded to three northern populations (N1, N2, N3), one northern population (N4), and four southern populations (S1, S2, S3, S4; Figure 4A), in close agreement with the clusters identified by STRUCTURE (Figure 2B).

Plastid haplotype variation. A total of 17 plastid haplotypes were identified in *P. greggii* (Figure 4B, C), and two in *P. patula*, based on 81 sequence variants distributed across 73 non-coding regions of the plastome (Table S3). Statistical parsimony network analysis identified two networks. Network I had 13 haplotypes, 9 exclusively found in the north (A-I), whereas Network II was formed by four haplotypes exclusively found in the south (N, O, P, Q) and the sample from Guerrero (Q) (Figure 4C). Two lineages in perfect agreement with the two networks were also recovered in a ML phylogenetic analysis of the 73 non-coding regions (Figure S2).

Similar diversity was detected between regions expressed in total number of plastid haplotypes and Nei's haplotypic diversity index, but nucleotide diversity was higher in the south (Table 2). Northern populations displayed a larger proportion of private haplotypes (six haplotypes, B-G), whereas in the south only two were found to be private to single populations (Figure 4B).

A significant phylogeographic structure was detected in the species where $N_{st} > G_{st}$ (0.410 vs. 0.140, respectively, $P < 0.001$). However, N_{st} was not significantly larger than G_{st} within each region (Tabla 2). The F_{CT} index of *SAMOVA* reached a plateau at 0.477 for $k = 4$ ($P < 0.01$). One group was formed by the four populations from the northern range (N1-N4), while southern populations S1 and S3 occurred in a second group, and populations S2 and S4 remained separated in distinct groups (Figure 4A).

The mismatch distribution analysis rejected the hypothesis of a demographic expansion at the species (*SSD* = 0.0077, $P = 0.75$) and regional levels (*SSD* = 0.0024, $P = 0.63$ for the northern region; *SSD* = 0.0119, $P = 0.77$ for the southern region). Population equilibrium was also supported by non-significant values of Tajima's *D* and Fu's *F* indices (Tabla 2).

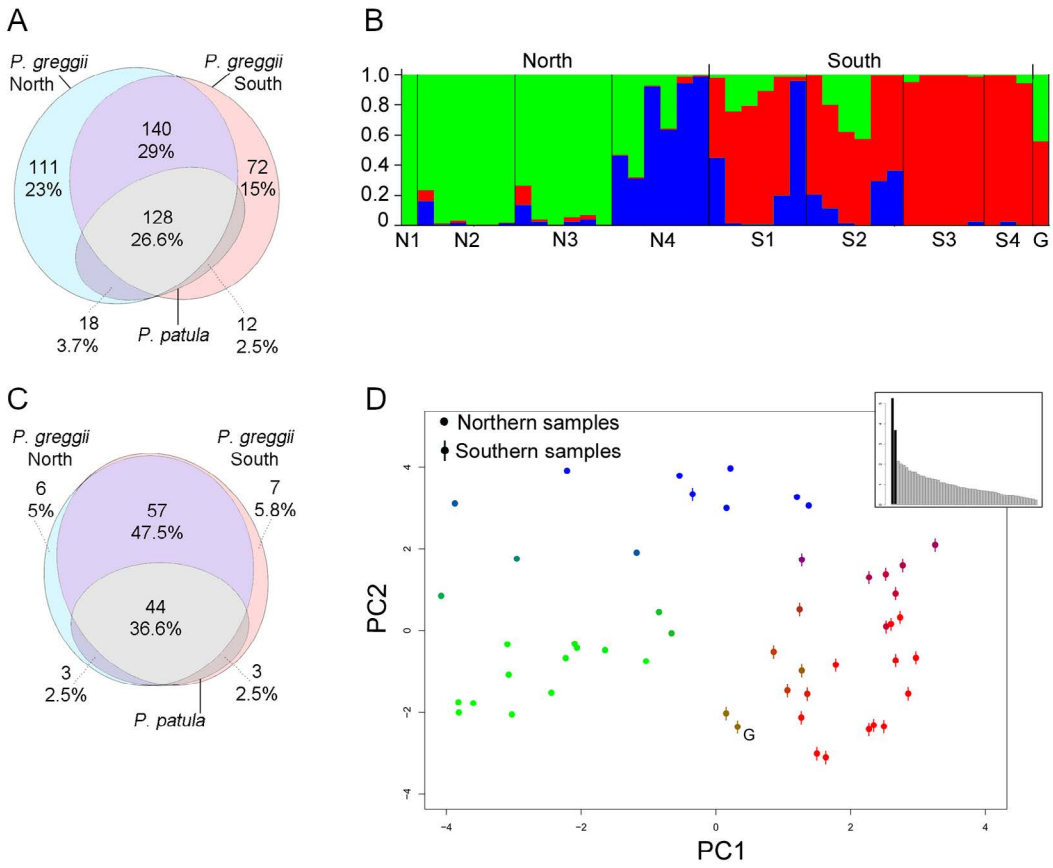

Figure 2. Nuclear variability and structure in *Pinus greggii* populations. (A) Distribution among taxa of 481 SNP alleles in northern (blue) and southern (pink) populations, in both regions (purple) and shared with *P. patula* (grey). For each set, number (top) and percentage (bottom) of alleles are shown. (B) Bayesian clustering according to STRUCTURE. The sample collected outside the natural range of the species is shown by the letter G. (C) Distribution among taxa of 120 SNP alleles that contributed the most to the structure shown by the PCA. (D) PCA plot of nuclear haplotypes. Samples were coloured according to proportional membership to each of the $K = 3$ clusters from b).

Ecological niche modelling. We found that during the Interglacial period, the distribution of the species was potentially fragmented in two areas, with *P. greggii* var. *australis* covering a broader distribution in the north, whereas *P. greggii* var. *greggii* was likely restricted to a small area in the south (Figure 5A). Similarly, *P. greggii* var. *australis* likely covered a larger area during the Last Glacial Maximum (LGM), when both varieties spread to the Mexican Plateau at the west of the SMO, where they came into contact. In fact, the Pánuco River continued acting as a barrier during the LGM, very likely due to the drastic elevational gradient found on the eastern slope of the SMO (Figure 5B), as shown by our significant analysis on isolation by landscape resistance. Finally, our modelling suggests that the potential distribution of *P. greggii* var. *australis* under current climatic conditions encompasses most of the species' range, whereas *P. greggii* var. *greggii* remains concentrated on the northern SMO (Figure 5C).

Discussion

Divergence in the presence of shared polymorphisms and gene flow. A broad scale pattern of differentiation consistent with the disjunct distribution of *P. greggii* varieties was detected both with nuclear and plastome data. Our results therefore confirm previous studies that pointed to the morphological and terpene chemistry differences that underlie the recognition of the two infraspecific taxa (Donahue & López-Upton 1996, 1999, Dvorak *et al.* 1996). A deeper

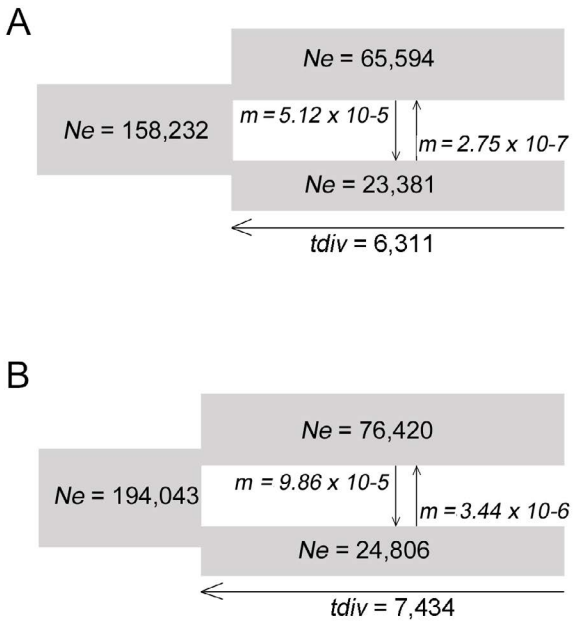

Figure 3. (A) Demographic parameters obtained with nuclear SNP data from the IM model between northern (top) and southern regions (bottom). (B) Same as A), but population N4 being included in the southern lineage (as suggested by its intermediate position in PCA, see Figure 2D). *Ne* = effective population size, *m* = migration rate; *tdiv* = time to divergence in generations.

understanding of this differentiation was attained by identifying two nuclear clusters within *P. greggii* var. *greggii* in the north and high plastome diversity in *P. greggii* var. *australis* in the south (Figure 4A). Therefore, despite an apparently small number of individuals per population, we were able to identify clear patterns of nuclear differentiation at the infraspecific level. We consider this to be the result of the exon capture approach based on the sequencing of single-copy genes, coupled to an increased statistical power through haplotype phasing and allele polarizing.

Yet, levels of shared variation between *P. greggii* regions increased from 29 to 47 % when considering those alleles with the greatest contribution to PCA structure (Figure 2C), revealing that, although present, alleles private to each region (38 % of all SNPs surveyed, including 20 % within single populations) contribute little to the general divergence pattern.

Our results support the notion that the differentiation of the two varieties is due to the accumulation of small, genome-wide differences. It has been proposed that the early stages of divergence on the speciation continuum (*i.e.*, from populations, to varieties, to species) are characterized by differentiation at multiple loci, which leads to shifted SNP allele frequencies and F_{ST} distributions (Feder *et al*. 2014). Here, divergence between nuclear gene pools was attained through shared yet strongly structured polymorphisms: high-loading PCA alleles displayed a mean F_{ST} per locus of 0.263, whereas the mean of the distribution of F_{ST} per locus was 0.145 (Figure S3). Differences in allelic frequencies between regions lead as well to an apparent deficiency of heterozygotes, producing a lower value of total *Hs* diversity at the species level (Table 1), a phenomenon known as the Wahlund effect (Hedrick 2000).

Moreover, divergent alleles were also shared with the closely related *P. patula* (37 %, Figure 2C). The phylogenetic distance between *P. patula* and *P. greggii* plastid haplotypes suggests that pollen-mediated interspecific gene flow is unlikely (Gernandt *et al*. 2018). Analyzing individuals with mixed morphological traits would be necessary to fully address this question (Donahue & Lopez-Upton 1996).

Large amounts of shared nuclear SNPs, low overall divergence at organellar markers, and a reduced number of loci with signatures of local adaptation support the strong role of extrinsic factors, such as geographical barriers and climate, in shaping ecologically driven divergent speciation in pines despite strong levels of interspecific gene flow (Menon *et al*. 2018, Wachowiak *et al*. 2018).

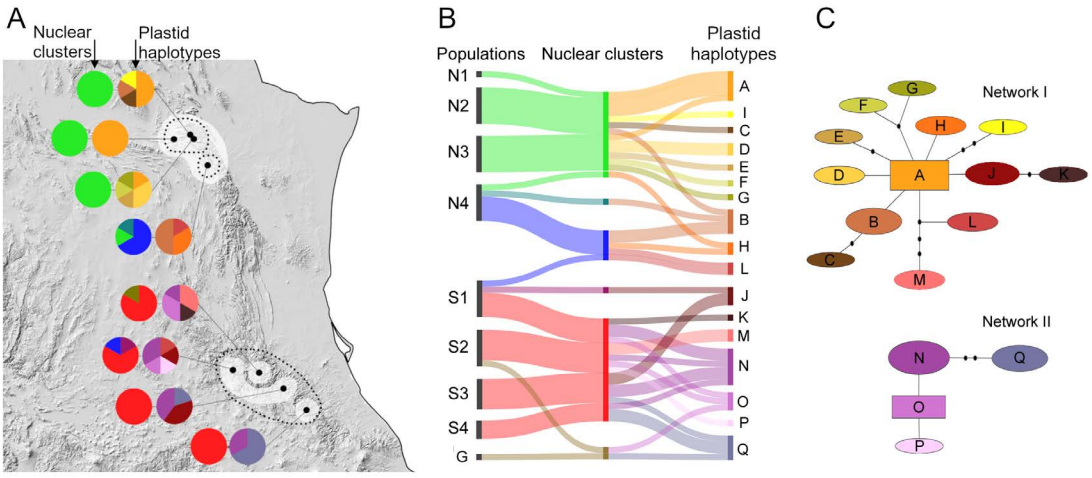

Figure 4. Relationships among nuclear and plastome composition of *Pinus greggii* individuals. (A) Distribution of nuclear clusters (left circles) and plastid haplotypes (right circles) per population. Geographically homogeneous population groups identified by SAMOVA according to nuclear data are shown with dashed lines (k = 3) whereas population groups identified with SAMOVA according to plastome data are shown in white color (k = 4). (B) Sankey diagram depicting nuclear and plastome composition of individuals within each population (admixed memberships are shown with intermediate colors), and plastid haplotypes. Each line represents an individual. (C) Statistical parsimony networks among 17 plastid haplotypes identified in *P. greggii*. Each line corresponds to one mutational step. Ellipse size is proportional to plastid haplotype frequency, black points represent unobserved missing haplotypes, squares represent the oldest haplotypes.

When compared to nuclear markers, plastid DNA showed similar structure levels (Tables 1 and 2). This result is consistent with the general pattern found in conifers, where similar estimates of genetic structure among biparentally and paternally inherited markers are observed (Petit *et al*. 2005). However, we detected that whereas region-specific red and green nuclear clusters encompassed 64 % of the samples, plastid haplotype networks were region-specific for only 50 % of the samples. The greater association between nuclear clusters and geographical region, although far from being variety-specific, can be attributed to the analysis of many loci, which reduces the stochastic variance of individual gene trees (Rosenberg & Nordborg 2002) and to the lower capacity of long-distance dispersal of seeds as compared to pollen. The drawback of nuclear alleles is their being largely shared between regions and the outgroup, making it impossible to identify diagnostic SNPs. Syring *et al.* (2007) found that the non-monophyly of single loci can occur across *Pinus* taxonomic sections.

Phylogeographical evidence suggests that allele sharing is more widespread in young lineages with broad geographic ranges (and very likely larger effective population sizes), such as the *P. ponderosa*, *P. ayacahuite-P.strobiformis-P. flexilis*, and *P. johannis-P.discolor-P.cembroides-P.culminicola* complexes in North America (Flores-Rentería *et al.* 2013, Moreno-Letelier *et al*. 2013, Willyard *et al*. 2016, Menon *et al*. 2018, Ortiz-Martínez *et al*. 2024), as well as *Pinus armandii-P. kwantungensis-P. bhutanica* from China (Liu *et al.* 2014). In contrast, representatives of *Pinus* subsection *Australes* with restricted distributions in the Caribbean basin show low levels of plastid haplotype sharing (Jardón-Barbolla *et al*. 2011).

Given the small distribution range and reduced effective population size of *P. greggii*, the pattern of nuclear vs. organellar allele sharing detected in this study (Figure 4A) reveals recent gene flow between varieties as the most relevant force in shaping their current genetic composition. Evidence supporting a scenario of gene flow includes the shared nuclear blue cluster (Figure 4B) likely caused by historical distribution shifts (see next section), the larger north-to-south migration rate, the presence of derived plastid haplotypes (J, K, L, and M) from northern Network I in the south (Figure 4A, B), and the absence of southern Network II haplotypes in northern populations (Figure 4B), contrary to theory that populations with larger effective sizes are expected to retain a greater proportion of polymorphisms due to the smaller effect of genetic drift (Hedrick 2000).

Diversification and vicariance in the SMO. The time to divergence between the two varieties (126,220-148,680 BP) matches with the last Interglacial period, which corresponds to marine isotope stage 5 (MIS5) (Lang & Wolff 2011). Despite some differences in the mechanisms behind global climate response, this interval has been considered a good past analog to our present conditions (Loutre & Berger 2003), supporting the notion that climatic regimes along the north-to-south axis of the SMO have been more contrasting during interglacial than glacial periods, as supported by our modelling of the potential distribution of the species during that time (Figure 5A).

When glacial conditions were re-established, woodlands of temperate affinity (*e.g.*, *Pinus*, *Picea*, *Abies*, *Juniperus*) developed in what is now the Chihuahuan Desert (Van Devender 1990, Metcalfe 2006). The areas potentially occupied by *P. greggii* according to our LGM distribution model (Figure 5B) agree with a glacial scenario where populations were able to occupy larger and more continuous areas than today. This scenario agrees as well with the floristic affinity between the meridional SMO and the Sierra Madre Occidental (in western Mexico, on the opposite side of the Mexican Plateau) (Espinosa *et al*. 2004), which goes hand in hand with the close phylogenetic relationships between *P. greggii* and a number of pine species found in the Sierra Madre Occidental such as *P. chihuahuana*, *P. lumholtzii*, and *P. herrerae* (Gernandt *et al*. 2018, Cruz-Nicolás *et al.* 2024).

Table 2. Plastid haplotype diversity and structure in *Pinus greggii*.

	Population	*n*	No. hap	*h* (sd)	π (sd)	*D*	F_S	G_{ST} (se)	N_{ST} (se)
N1	El Diamante	1	1	0	0	0	0		
N2	La Carbonera	6	4	0.800 (0.172)	0.10051 (0.06395)	-0.44736^{ns}	1.69856^{ns}		
N3	Los Lirios	6	5	0.933 (0.122)	0.10238 (0.06593)	-0.31466^{ns}	-0.11394^{ns}		
N4	Cerro Potosí	6	3	0.733 (0.155)	0.05595 (0.03886)	0.60031^{ns}	1.90447^{ns}		
Total North		19	10	0.895 (0.047)	0.07701 (0.04376)	-1.51684^{ns}	-1.05457^{ns}	0.129 (0.064)	0.086
S1	Cerro La Pingüica	6	5	0.933 (0.122)	0.14747 (0.09106)	1.19636^{ns}	0.77293^{ns}		
S2	Laguna Seca	6	4	0.867 (0.129)	0.18990 (0.11560)	1.24649^{ns}	3.08046^{ns}		
S3	Jalamelco	5	3	0.800 (0.164)	0.09643 (0.06535)	0.89450^{ns}	2.46126^{ns}		
S4	Cumbre de Muridores	3	2	0.667 (0.314)	0.03571 (0.03367)	0.00000^{ns}	1.60944^{ns}		
Total South		20	8	0.868 (0.047)	0.13995 (0.07501)	0.39170^{ns}	3.10735^{ns}	0.056 (0.072)	0.148 (0.179)
Total Species		39	17	0.940 (0.016)	0.11800 (0.06173)	-0.45994^{ns}	-0.57481^{ns}	0.140 (0.043)	0.410 (0.133)

Notes: *n*: sample size; *h*: haplotypic diversity index; π: nucleotide diversity index; *D*: Tajima's *D* statistic; F_S: Fu's *F*; G_{ST}: haplotypic differentiation index based on unordered alleles; N_{ST}: haplotypic differentiation based on ordered alleles. *ns*: non-significant, $P > 0.05$.

Secondary contact during the LGM resulted in cryptic breaks (nuclear in the north, plastid in the south; Figure 4A) that represent remnants of the past continuous distribution of the species (Figure 5). Similar suture zones characterized by co-distributed divergent lineages have been identified as evidence of secondary contact of refugial lineages (Gömöry *et al.* 2012, Cun & Wang 2015).

Geographical expansion of *Australes* taxa on the Caribbean islands has also been found to match glacial periods with subsequent isolation during interglacials (Jardón-Barbolla *et al.* 2011), which suggests similar climatic diversification drivers in these two areas.

Surprisingly, we found that populations show demographic stability (Table 2), and the general distribution of the species does not seem to have experienced radical reductions or expansions in the last 130,000 years (Figure 5A-C). This pattern differs from other North American and Asian conifers, including *Pinus*, *Tsuga*, *Pseudotsuga*, and *Picea*, whose history is characterized by species persistence in small, localized refugial areas, from where range expansions and demographic growth took place when temperatures changed (Gugger *et al.* 2011, Ge *et al*. 2012, Cun & Wang

2015, Willyard *et al*. 2016). In contrast, our data support a scenario of a north-south fragmentation event in the Last Interglacial (Figure 5A), accompanied by continuous *in situ* occurrence of *P. greggii* in the SMO, coupled to elevational shifts during its advance into the lower Mexican Plateau. The fragmentation event was clearly asymmetrical, as the southern region is characterized by an effective population size nearly three times smaller (Figure 3), a reduced number of private alleles (72 *vs*. 111, Figure 2A), and isolation-by-distance, despite a nuclear diversity similar to northern populations (H_S = 0.28; Table 1).

Rare alleles correlated with elevation identified in *P. ponderosa* suggest that elevation imposes adaptive constraints on pine populations (Potter *et al*. 2015). Moreover, tropical sites show stronger dissimilarities in temperature along elevational gradients as well as a more uniform monthly climatic regime compared to sites from more northern latitudes, leading to a higher effectiveness of mountain-valley systems in isolating populations due to increasing niche constraints (Janzen 1967). In this sense, the SMO represents the first of a series of mountain chains where the temperate *Pinus* started to colonize the New World's subtropical and tropical latitudes, and it encountered strong geographic barriers coupled to large climatic variance on small geographical scales. Indeed, *Australes* is characterized both by high diversification rates and the southernmost distribution of all New World pines (Farjon & Filer 2013, Gernandt *et al*. 2018, Cruz-Nicolás *et al.* 2024). Other Mexican conifers also show evidence of speciation processes linked to ecological divergence (Farjon 1996, Jaramillo-Correa *et al*. 2008, Moreno-Letelier & Piñero 2009, Martínez de León *et al.* 2022).

Implications for timber production and conservation. The identification of the two nuclear clusters in the northern area of distribution, and the presence of two distinct plastid haplotype networks can be used to improve *ex situ* management programs focused on the potential of *P. greggii* germplasm for timber production and reforestation programs. For instance, trees of *P. greggii* grown in nurseries have shown to perform to a variable degree (from low to high) in different morphological and physiological parameters, but their provenances are unknown or not included in plant quality evaluations (Rueda Sánchez *et al.* 2012), despite previous studies showing clear growth differences between trees of northern and southern provenances (Dvorak *et al.* 1996). The intraspecific variability detected in our study can therefore contribute to improving tree quality evaluation programs and to broadening the genetic basis of selected founders chosen by the forest industry (Dvorak 2012).

For instance, the individual from Guerrero (labeled *G* in Figures 2 and 4) was identified as belonging to the southern region, based both on its membership to the red cluster and its plastid haplotype *Q*. In this case, nuclear data alone could not allow for the identification of its place of origin mainly due to its admixed genotype (Figure 2B), but when considering plastid haplotype identity, it is possible to assign it to a southern provenance (Figure 4B). Yet, our data do not enable the identification of its origins to the population level.

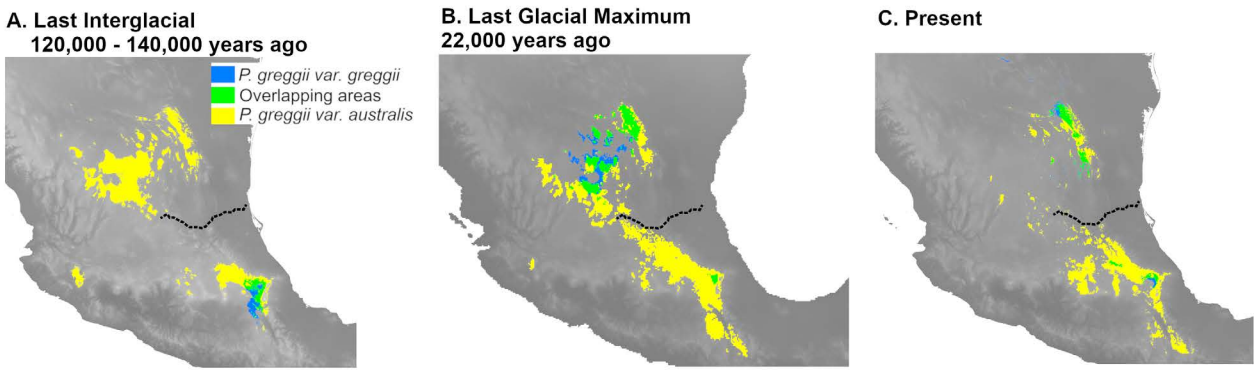

Figure 5. Ecological niche models of *Pinus greggii* intraspecific varieties in Mexico. A) Last Interglacial (ca. 130,000 years ago); B) Last Glacial Maximum (ca. 20,000 years ago); C) Present conditions (see also Figure 1). Varieties are color-coded as shown in the legend. Broken line represents the Pánuco River.

Additionally, the Hyb-Seq method used in this study could also be used for capturing genetic variation directly related to candidate genes associated with key phenotypic traits that might be of interest to the forest industry (González-Martínez *et al.* 2007, Eckert *et al.* 2013).

Regarding the conservation strategies for the species, populations were characterized by moderate levels of diversity (H_S = 0.08 - 0.244) and low amounts of structuring (F_{ST} < 0.15), in agreement with general patterns identified for long-lived tree species (Petit & Hampe 2006). However, three observations deserve attention: lower genetic diversity in rear edge populations, small effective population size in the southern region, and high levels of inbreeding.

Populations located at the limits of a species' distribution are often considered more susceptible to the effects of genetic drift and reduced gene flow caused by small population sizes and geographical isolation. Populations N1, S3, and S4 were characterized by less nuclear variation (Table 1), which could be the result of smaller effective population sizes. This result points to the urgent need to protect them, especially when we consider their smaller *Ne* despite the large area potentially suitable for their growth (Figure 5C).

Parraguirre-Lezama *et al.* (2004), using allozymes, identified moderate levels of inbreeding in *P. greggii* (F_{IS} = 0.271) and attributed it to human-driven isolation and fragmentation. Moderate F_{IS} values have also been observed with nuclear microsatellites in *P. ponderosa* (F_{IS} = 0.041-0.178; and *P. oocarpa* (F_{IS} = 0.15; Dvorak *et al.* 2009). The increased levels of inbreeding may be caused by mating among relatives due to the low density of the species and the presence of other trees that represent barriers to pollen movement (El-Kassaby & Jaquish 1996, Parraguirre-Lezama *et al*. 2004). Yet, we found very similar inbreeding in northern and southern regions, despite strong differences in effective population sizes (Table 1, Figure 3), and no evidence of past population decline (Table 2). Conifers are characterized by high levels of inbreeding depression (Kärkkäinen *et al.* 1996, Chancerel *et al*. 2013), but they also exhibit low linkage disequilibrium (Prunier *et al.* 2016), and selfing has proved to be effective in purging deleterious alleles in *P. radiata* inbred lines (Wu *et al*. 2004). One possibility is that deleterious alleles have been purged in *P. greggii*, because the species performs well in forestry programs when compared to other species (Dvorak *et al*. 2000) and produces many viable seeds (INIFAP 2003). However, it is also possible that the high inbreeding levels detected in our study are the product of current fragmentation and population decline (Ramírez-Herrera *et al.* 2005), but that the effects of deleterious alleles have not been expressed yet due to the long life-cycle of these trees. Future studies comparing the inbreeding levels of different age cohorts within the populations (*i.e.*, adult *vs*. seedling stages) would help to elucidate if inbreeding levels are increasing over succeeding generations (Lee *et al.* 2000, Pakull *et al.* 2021).

Supplementary material

Supplemental data for this article can be accessed here: https://doi.org/10.17129/botsci.3688

Data accessibility

Illumina sequencing reads are deposited at the Sequence Read Archive of NCBI (SRA PRJNA540071).

Acknowledgments

We thank A. Liston for his help during probe design and for revising a previous version of the manuscript and C. Domínguez for helping with the database of geographic coordinates. We thank the two anonymous reviewers whose insightful comments greatly improved the contents of this work.

Literature cited

Aguirre-Dugua A, Gernandt DS. 2017. Complete plastomes of three endemic Mexican pine species (*Pinus* subsection *Australes*). *Mitochondrial DNA Part B* **2**: 562-565. DOI: https://doi.org/10.1080/23802359.2017.1365637

Aguirre-Liguori JA, Tenaillon MI, Vázquez-Lobo A, Gaut BS, Jaramillo-Correa JP, Montes-Hernández S, Souza V, Eguiarte LE. 2017. Connecting genomic patterns of local adaptation and niche suitability in teosintes. *Molecular Ecology* **26**: 4226-4240. DOI: https://doi.org/10.1111/mec.14203

Allendorf FW, Hohenlohe PA, Luikart G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics* **11**: 697-709. DOI: https://doi.org/10.1038/nrg2844

Amos W, Balmford A. 2001. When does conservation genetics matter? *Heredity* **87**: 257-265. DOI: https://doi.org/10.1046/j.1365-2540.2001.00940.x

Bryson RW, Murphy RW, Lathrop A, Lazcano-Villareal D. 2011. Evolutionary drivers of phylogeographical diversity in the highlands of Mexico: A case study of the *Crotalus triseriatus* species group of montane rattlesnakes. *Journal of Biogeography* **38**: 697-710. DOI: https://doi.org/10.1111/j.1365-2699.2010.02431.x

Chancerel E, Lamy JB, Lesur I, Noirot C, Klopp C, Ehrenmann F, Boury C, Le Provost G, Label P, Lalanne C, Léger V, Salin F, Gion JM, Plomion C. 2013. High-density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. *BMC Biology* **11**: 50. DOI: https://doi.org/10.1186/1741-7007-11-50

Clement M, Posada D, Crandall KA. 2000. TCS: A computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657-1659. DOI: https://doi.org/10.1046/j.1365-294x.2000.01020.x

Conde C, Estrada F, Martínez B. 2011. Regional climate change scenarios for Mexico. *Atmósfera* **24**: 125-140.

Cruz-Nicolás J, Jaramillo-Correa JP, Gernandt, DS. 2024. Stochastic processes and changes in evolutionary rate are associated with diversification in a lineage of tropical hard pines (*Pinus*). *Molecular Phylogenetics and Evolution* **192**: 108011. DOI: https://doi.org/10.1016/j.ympev.2024.108011

Cun YZ, Wang XQ. 2015. Phylogeography and evolution of three closely related species of *Tsuga* (hemlock) from subtropical eastern Asia: Further insights into speciation of conifers. *Journal of Biogeography* **42**: 315-327. DOI: https://doi.org/10.1111/jbi.12421

Cushman SA, McKelvey KS, Hayden J, Schwartz MK. 2006. Gene flow in complex landscapes: Testing multiple hypotheses with causal modeling. *The American Naturalist* **168**: 486-499. DOI: https://doi.org/10.1086/506976

Donahue JK, Lopez-Upton J. 1996. Geographic variation in leaf, cone and seed morphology of *Pinus greggii* in native forests. *Forest Ecology and Management* **82**: 145-157. DOI: https://doi.org/10.1016/0378-1127(95)03677-6

Donahue JK, Lopez-Upton J. 1999. A new variety of *Pinus greggii* (Pinaceae) in Mexico. *SIDA* **18**: 1083-1093. DOI: https://www.jstor.org/stable/41967722

Dorman KW. 1976. *The genetics and breeding of southern pines*. Washington, DC: US. Dept. of Agriculture, Forest Service.

Dupanloup I, Schneider S, Excoffier L. 2002. A simulated annealing approach to define the genetic structure of populations. *Molecular Ecology* **11**: 2571-2581. DOI: https://doi.org/10.1046/j.1365-294X.2002.01650.x

Dvorak W. 2012. The strategic importance of applied tree conservation programs to the forest industry in South Africa. *Southern Forests: A Journal of Forest Science* **74**: 1-6. DOI: https://doi.org/10.2989/20702620.2012.683635

Dvorak WS, Kietzka JE, Donahue JK. 1996. Three-year survival and growth of provenances of *Pinus greggii* in the tropics and subtropics. *Forest Ecology and Management* **83**: 123-131. DOI: https://doi.org/10.1016/0378-1127(95)03673-3

Dvorak WS, Kietzka JE, Donahue JK, Hodge R, Stanger TK. 2000. *Pinus greggii*. *In*: *Conservation and Testing of Tropical and Subtropical Forest Tree Species by the CAMCORE Cooperative, Central America and Mexico Coniferous Resources Cooperative,* pp. 52-73. Raleigh: North Carolina State University. ISBN: 0-620-26460-8.

Dvorak WS, Potter KM, Hipkins VD, Hodge GR. 2009. Genetic diversity and gene exchange in *Pinus oocarpa*, a Mesoamerican pine with resistance to the pitch canker fungus. *International Journal of Plant Sciences* **170**: 609-626. DOI: https://doi.org/10.1086/597780

Eckert AJ, Wegrzyn JL, Liechty JD, Lee JM, Cumbie WP, Davis JM, Goldfarb B, Loopstra CA, Palle SR, Quesada T, Langley CH, Neale DB. 2013. The evolutionary genetics of the genes underlying phenotypic associations for loblolly pine (*Pinus taeda*, Pinaceae). *Genetics* **195**: 1353-1372. DOI: https://doi.org/10.1534/genetics.113.157198

El-Kassaby YA, Jaquish B. 1996. Population density and mating pattern in western larch. *Journal of Heredity* **87**: 438-443. DOI: https://doi.org/10.1093/oxfordjournals.jhered.a023034

Espinosa D, Aguilar C, Ocegueda S. 2004. Identidad geográfica de la Sierra Madre Oriental y posibles subdivisiones bióticas. *In:* Luna I, Morrone JJ, Espinosa D. eds. *Biodiversidad de la Sierra Madre Oriental*. 487-500 pp. Mexico, DF: Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Universidad Nacional Autónoma de México (UNAM), Las Prensas de Ciencias.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* **14**: 2611-2620. DOI: https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Dupanloup I, Huerta-Sanchez E, Sousa VC, Foll M. 2013. Robust demographic inference from genomic and SNP data. *Plos Genetics* **9**: e1003905. DOI: https://doi.org/10.1371/journal.pgen.1003905

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47-50.

Farjon A. 1996. Biodiversity of *Pinus* (Pinaceae) in Mexico: Speciation and palaeo-endemism. *Botanical Journal of the Linnean Society* **121**: 365-384. DOI: https://doi.org/10.1111/j.1095-8339.1996.tb00762.x

Farjon A, Filer D. 2013. *An Atlas of the World's Conifers*. Boston: Brill. ISBN: 978-90-04-21180-3.

Feder JL, Nosil P, Wacholder AC, Egan SP, Berlocher SH, Flaxman SM. 2014. Genome-wide congealing and rapid transitions across the speciation continuum during speciation with gene flow. *Journal of Heredity* **105**: 810-820. DOI: https://doi.org/10.1093/jhered/esu038

Ferro I, Navarro-Sigüenza AG, Morrone JJ. 2017. Biogeographical transitions in the Sierra Madre Oriental, Mexico, shown by chorological and evolutionary biogeographical affinities of passerine birds (Aves: Passeriformes). *Journal of Biogeography* **44**: 2145-2160. DOI: https://doi.org/10.1111/jbi.13015

Flores-Rentería L, Wegier A, Ortega Del Vecchyo D, Ortíz-Medrano A, Piñero D, Whipple AV, Molina-Freaner F, Domínguez CA. 2013. Genetic, morphological, geographical and ecological approaches reveal phylogenetic relationships in complex groups, an example of recently diverged pinyon pine species (Subsection *Cembroides*). *Molecular Phylogenetics and Evolution* **69**: 940-949. DOI: https://doi.org/10.1016/j.ympev.2013.06.010

Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**: 915-925. DOI: https://doi.org/10.1093/genetics/147.2.915

GBIF.org 2025. GBIF Occurrence Download (7 May 2025). DOI: https://doi.org/10.15468/dl.8p2n4r

Ge XJ, Hsu TW, Hung KH, Lin CJ, Huang CC, Huang CC, Chiang YC, Chiang TY. 2012. Inferring multiple refugia and phylogeographical patterns in *Pinus massoniana* based on nucleotide sequence variation and DNA fingerprinting. *Plos One* **7**: 1-13. DOI: https://doi.org/10.1371/journal.pone.0043717

Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, Lawrence DM, Neale RB, Rasch PJ, Vertenstein M, Worley PH, Yang ZL, Zhang M. 2011. The community climate system model version 4. *Journal of Climate* **24**: 4973-4991. DOI: https://doi.org/10.1175/2011JCLI4083.1

Gernandt DS, Aguirre-Dugua X, Vázquez-Lobo A, Moreno-Letelier A, Pérez-de la Rosa JA, Piñero D, Liston A. 2018. Multi-locus phylogenetics, lineage sorting and reticulation in *Pinus* subsection *Australes*. *American Journal of Botany* **105**: 1-15. DOI: https://doi.org/10.1002/ajb2.1052

Gernandt DS, Geada López G, Ortiz García S, Liston A. 2005. Phylogeny and classification of *Pinus*. *Taxon* **54**: 29-42. DOI: https://doi.org/10.2307/25065300

Gömöry D, Paule L, Krajmerová D, Romsáková I, Longauer R. 2012. Admixture of genetic lineages of different glacial origin: a case study of *Abies alba* Mill. in the Carpathians. *Plant Systematics and Evolution* **298**: 703-712. DOI: https://doi.org/10.1007/s00606-011-0580-6

González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB. 2007. Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* **175**: 399-409. DOI: https://doi.org/10.1534/genetics.106.061127

Gugger PF, González-Rodríguez A, Rodríguez-Correa H, Sugita S, Cavender-Bares J. 2011. Southward Pleistocene migration of Douglas-fir into Mexico: Phylogeography, ecological niche modeling, and conservation of "rear edge" populations. *New Phytologist* **189**: 1185-1199. DOI: https://doi.org/10.1111/j.1469-8137.2010.03559.x

Halffter G, Morrone JJ. 2017. An analytical review of Halffter's Mexican transition zone, and its relevance for evolutionary biogeography, ecology and biogeographical regionalization. *Zootaxa.* **4226**: 1-46. DOI: https://doi.org/10.11646/zootaxa.4226.1.1

Hedrick PW. 2000. *Genetics of Populations*. 2nd ed. Jones and Bartlett Publishers, Sudbury.

Hernández-León S, Gernandt DS, Pérez de la Rosa JA, Jardón-Barbolla L. 2013. Phylogenetic relationships and species delimitation in *Pinus* section *Trifoliae* inferred from plastid DNA. *Plos One* **8**: e70501. DOI: https://doi.org/10.1371/journal.pone.0070501

Hijmans R. 2024. *geosphere: Spherical Trigonometry*. R package version 1.5-20. https://github.com/rspatial/geosphere

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965-1978. DOI: https://doi.org/10.1002/joc.1276

INIFAP [Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias]. 2003. *Monografía de Pinus greggii*. Distrito Federal: INIFAP, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO).

Jakobsson M, Rosenberg NA. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* **23**: 1801-1806. DOI: https://doi.org/10.1093/bioinformatics/btm233

Janzen DH. 1967. Why mountain passes are higher in the Tropics. *The American Naturalist.* **101**: 233-249.

Jaramillo-Correa JP, Aguirre-Planter E, Khasa DP, Eguiarte LE, Piñero D, Furnier GR, Bousquet J. 2008. Ancestry and divergence of subtropical montane forest isolates: Molecular biogeography of the genus *Abies* (Pinaceae) in southern México and Guatemala. *Molecular Ecology* **17**: 2476-2490. DOI: https://doi.org/10.1111/j.1365-294X.2008.03762.x

Jardón-Barbolla L, Delgado-Valerio P, Geada-López G, Vázquez-Lobo A, Piñero D. 2011. Phylogeography of *Pinus* subsection *Australes* in the Caribbean Basin. *Annals of Botany* **107**: 229-241. DOI: https://doi.org/10.1093/aob/mcq232

Jin W-T, Gernandt DS, Wehenkel C, Xia X-M, Wei X-X, Wang X.-Q. 2021. Phylogenomic and ecological analyses reveal the spatiotemporal evolution of global pines. *Proceedings of the National Academy of Sciences* **118**: e2022302118. DOI: https://doi.org/10.1073/pnas.2022302118

Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070-3071. DOI: https://doi.org/10.1093/bioinformatics/btr521

Kärkkäinen K, Koski V, Savolainen O. 1996. Geographical variation in the inbreeding depression of Scots pine. *Evolution* **50**: 111-119. DOI: https://doi.org/10.2307/2410785

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647. DOI: https://doi.org/10.1093/bioinformatics/bts199

Kesälahti R, Kumpula TA, Cervantes S, Kujala ST, Mattila TM, Tyrmi JS, Niskanen AK, Rastas P, Savolainen O, Pyhäjärvi T. 2024. Optimizing exome captures in species with large genomes using species-specific repetitive DNA blocker. *bioRxiv* 2024.04.23.590505. DOI: https://doi.org/10.1101/2024.04.23.590505

Lang N, Wolff EW. 2011. Interglacial and glacial variability from the last 800 ka in marine, ice and terrestrial archives. *Climate of the Past* **7**: 361-380. DOI: https://doi.org/10.5194/cp-7-361-2011

Lee SL, Wickneswari R, Mahani MC, Zakri AH. 2000. Genetic diversity of a tropical tree species, *Shorea leprosula* Miq. (Dipterocarpaceae), in Malaysia: Implications for conservation of genetic resources and tree improvement. *Biotropica* **32**: 213-224.

Librado P, Rozas J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452. DOI: https://doi.org/10.1093/bioinformatics/btp187

Liu L, Hao ZZ, Liu YY, Wei XX, Cun YZ, Wang XQ. 2014. Phylogeography of *Pinus armandii* and its relatives: Heterogeneous contributions of geography and climate changes to the genetic differentiation and diversification of Chinese white pines. *Plos One* **9**: 1-12. DOI: https://doi.org/10.1371/journal.pone.0085920

Loutre MF, Berger A. 2003. Marine Isotope Stage 11 as an analogue for the present interglacial. *Global and Planetary Change* **36**: 209-217. DOI: https://doi.org/10.1016/S0921-8181(02)00186-8

Martínez de León R. Castellanos-Morales G, Moreno-Letelier A. 2022. Incipient speciation, high genetic diversity, and ecological divergence in the alligator bark juniper suggest complex demographic changes during the Pleistocene. *PeerJ* **10**: 1-29. DOI: https://doi.org/10.7717/peerj.13802

Menon M, Bagley JC, Friedline CJ, Whipple AV, Schoettle AW, Leal-Sàenz A, Wehenkel C, Molina-Freaner F, Flores-Rentería L, Gonzalez-Elizondo MS, Sniezko RA, Cushman SA, Waring KM, Eckert AJ. 2018. The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine. *Molecular Ecology* **27**: 1245-1260. DOI: https://doi.org/10.1111/mec.14505

Metcalfe SE. 2006. Late Quaternary environments of the northern deserts and central Transvolcanic Belt of Mexico. *Annals of the Missouri Botanical Garden* **93**: 258-273.

Micallef L, Rodgers P. 2014. eulerAPE: Drawing Area-Proportional 3-Venn Diagrams Using Ellipses. *Plos One* **9**: e101717. DOI: https://doi.org/10.1371/journal.pone.0101717

Millar CI. 1993. Impact of the Eocene on the evolution of *Pinus* L. *Annals of the Missouri Botanical Garden* **80**: 471-498. DOI: https://doi.org/10.2307/2399795

Moreno-Letelier A, Ortíz-Medrano A, Piñero D. 2013. Niche divergence versus neutral processes: Combined environmental and genetic analyses identify contrasting patterns of differentiation in recently diverged pine species. *Plos One* **8**: e78228. DOI: https://doi.org/10.1371/journal.pone.0078228

Moreno-Letelier A, Piñero D. 2009. Phylogeographic structure of *Pinus strobiformis* Engelm. across the Chihuahuan Desert filter-barrier. *Journal of Biogeography* **36**: 121-131. DOI: https://doi.org/10.1111/j.1365-2699.2008.02001.x

Müller K. 2006. Incorporating information from length-mutational events into phylogenetic analysis. *Molecular Phylogenetics and Evolution* **38**: 667-676. DOI: https://doi.org/10.1016/j.ympev.2005.07.011

Neale DB, Wegrzyn JL, Stevens KA, Zimin AV, Puiu D, Crepeau MW, Cardeno C, Koriabine M, Holtz-Morris AE, Liechty JD, Martínez-García PJ, Vasquez-Gross HA, Lin BY, Zieve JJ, Dougherty WM, Fuentes-Soriano S, Wu LS, Gilbert D, Marçais G, Roberts M, Holt C, Yandell M, Davis JM, Smith KE, Dean JF, Lorenz WW, Whetten RW, Sederoff R, Wheeler N, McGuire PE, Main D, Loopstra CA, Mockaitis K, deJong PJ, Yorke JA, Salzberg SL, Langley CH. 2014. Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology* **15**: R59. DOI: https://doi.org/10.1186/gb-2014-15-3-r59

Nei M. 1972. Genetic distance between populations. *American Naturalist* **106**: 283-292.

Nei M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.

Niu S, Li J, Bo W, Yang W, Zuccolo A, Giacomello S, Chen X, Han F, Yang J, Song Y, Nie Y, Zhou B, Wang P, Zuo Q, Zhang H, Ma J, Wang J, Wang L, Zhu Q, Zhao H, Liu Z, Zhang Z, Liu T, Pei S, Li Z, Hu Y, Yang Y, Li W, Zan Y, Zhou L, Lin J, Yuan T, Li W, Li Y, Wei H, Wu HX. 2022. The Chinese pine genome and methylome unveil key features of conifer evolution. *Cell* **185**: 204-217.e14. DOI: https://doi.org/10.1016/j.cell.2021.12.006

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Borman T, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J. 2025. *vegan: Community Ecology Package*. R package version 2.7-0. https://vegandevs.github.io/vegan/

Ortiz-Martínez A, Ornelas-García CP, Moo-Llanes DA, Piñero D, de la Rosa JA, Peláez P, Moreno-Letelier A. 2024. Species delimitation using multiple sources of evidence from the *Pinus strobiformis-Pinus ayacahuite* species complex. *Botanical Sciences* **102:** 482-498. DOI: https://doi.org/10.17129/botsci.3364

Otto-Bliesner BE, Marshall SJ, Overpeck JT, Miller GH, Hu A, Cape Last Interglacial Project Members. 2006. Simulating arctic climate warmth and icefield retreat in the Last Interglaciation. *Science* **311**: 1751-1753. DOI: https://doi.org/10.1126/science.1120808

Pakull B, Eusemann P, Wojacki J, Ahnert D, Liesebach H. 2021. Genetic diversity of seeds from four German Doug-

las fir (*Pseudotsuga menziesii*) seed orchards. *European Journal of Forest Research* **140**: 1543-1557. DOI: https://doi.org/10.1007/s10342-021-01419-3

Parraguirre-Lezama C, Vargas-Hernández JJ, Ramírez-Vallejo P, Ramírez-Herrera C. 2004. Sistema de cruzamiento en cuatro poblaciones naturales de *Pinus greggii* Engelm. *Agrociencia* **38**: 107-119.

Peláez P, Ortiz-Martínez A, Figueroa-Corona L, Montes JR, Gernandt DS. 2020. Population structure, diversifying selection, and local adaptation in *Pinus patula*. *American Journal of Botany* **107**: 1-12. DOI: https://doi.org/10.1002/ajb2.1566

Perry JP, Graham A, Richardson DM. 1998. The history of pines in Mexico and Central America. *In*: Richardson, DM. ed. *Ecology and Biogeography of Pinus*, 137-149. Cambridge: Cambridge University Press.

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG. 2005. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology* **14**: 689-701. DOI: https://doi.org/10.1111/j.1365-294X.2004.02410.x

Petit RJ, Hampe A. 2006. Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics* **37**: 187-214. DOI: https://doi.org/10.1146/annurev.ecolsys.37.091305.110215

Phillips SJ, Anderson RP, Schapire RE. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**: 231-259. DOI: https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pons O, Petit RJ. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **144**: 1237-1245. DOI: https://doi.org/10.1093/genetics/144.3.1237

Potter KM, Hipkins VD, Mahalovich MF, Means RE. 2015. Nuclear genetic variation across the range of ponderosa pine (*Pinus ponderosa)*: Phylogeographic, taxonomic and conservation implications. *Tree Genetics and Genomes* **11**: 38. DOI: https://doi.org/10.1007/s11295-015-0865-y

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-59. DOI: https://doi.org/10.1093/genetics/155.2.945

Prunier J, Verta JP, Mackay JJ. 2016. Conifer genomics and adaptation: At the crossroads of genetic diversity and genome function. *New Phytologist* **209**: 44-62. DOI: https://doi.org/10.1111/nph.13565

Ramírez-Herrera C, Vargas-Hernández JJ, López-Upton J. 2005. Distribución y conservación de las poblaciones naturales de *Pinus greggii*. *Acta Botanica Mexicana* **16**: 1-16.

Robledo-Arnuncio JJ. 2011. Wind pollination over mesoscale distances: An investigation with Scots pine. *New Phytologist* **190**: 222-233. DOI: https://doi.org/10.1111/j.1469-8137.2010.03588.x

Rogers AR, Harpending H. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**: 552-69. DOI: https://doi.org/10.1093/oxfordjournals.molbev.a040727

Rosenberg NA, Nordborg M. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nature Reviews Genetics* **3**: 380-90. DOI: https://doi.org/10.1038/nrg795

Rueda Sánchez A, Benavides Solorio JDD, Prieto-Ruiz JÁ, Sáenz Reyez JT, Orozco-Gutiérrez G, Molina Castañeda A. 2012. Calidad de planta producida en los viveros forestales de Jalisco. *Revista Mexicana de Ciencias Forestales* **3**: 071-082. DOI: https://doi.org/10.29298/rmcf.v3i14.475

Ruiz-Sanchez E, Rodriguez-Gomez FC, Sosa V. 2012. Refugia and geographic barriers of populations of the desert poppy, *Hunnemannia fumariifolia* (Papaveraceae). *Organisms Diversity and Evolution* **12**: 133-143. DOI: https://doi.org/10.1007/s13127-012-0089-z

Saladin B, Leslie AB, Wüest RO, Litsios G, Conti E, Salamin N, Zimmermann NE. 2017. Fossils matter: Improved estimates of divergence times in *Pinus* reveal older diversification. *BMC Evolutionary Biology* **17**: 95. DOI: https://doi.org/10.1186/s12862-017-0941-z

Sánchez-Acevedo V, González-Rodríguez A, Torres-Miranda CA, Rodríguez-Correa H, Valencia-Á. S, De-la-Cruz IM, Oyama K. 2023. Nuclear and chloroplast DNA phylogeography reveals high genetic diversity and postglacial range expansion in *Quercus mexicana*. *American Journal of Botany* **110**: e16251. DOI: https://doi.org/10.1002/ajb2.16251

Schneider S, Excoffier L. 1999. Estimation of past demographic parameters from the distribution of pairwise differ-

ences when the mutation rates vary among sites: Application to human mitochondrial DNA. *Genetics* **152**: 1079-89. DOI: https://doi.org/10.1093/genetics/152.3.1079

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing data imputation. *American Journal of Human Genetics* **76**: 449. DOI: https://doi.org/10.1086/428594

Styles BT. 1993. Genus *Pinus*: a Mexican Review. *In:* Ramamoorthy TP, Bye R, Lot A. eds. *Biological Diversity of Mexico: Origin and Distribution*, pp. 397-420. New York: Oxford University Press.

Suárez-Mota ME, Villaseñor JL. 2017. Dominios climáticos de la Sierra Madre Oriental y su relación con la diversidad florística. *Revista Mexicana de Biodiversidad* **88**: 1-10. DOI: https://doi.org/10.1016/j.rmb.2017.01.020

Syring J, Farrell K, Businský R, Cronn R, Liston A. 2007. Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. *Systematic Biology* **56**: 163-181. DOI: https://doi.org/10.1080/10635150701258787

Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**: 437-460. DOI: https://doi.org/10.1093/genetics/105.2.437

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585-595. DOI: https://doi.org/10.1093/genetics/123.3.585

Torres-Miranda A, Luna-Vega I, Oyama K. 2011. Conservation biogeography of red oaks (*Quercus*, section Lobatae) in Mexico and Central America. *American Journal of Botany* **98**: 290-305. DOI: https://doi.org/10.3732/ajb.1000218

Van Devender TR. 1990. Late Quaternary vegetation and climate of the Chihuahuan desert, United States and Mexico. *In:* Betancourt JL, Van Devender TR, Martin PS. eds. *Pack Rat Middens: The Last 40 000 Years of Biotic Change*. pp. 104-133. Tucson: University of Arizona.

Wachowiak W, Zaborowska J, Bartosz Ł, Perry A, Zucca GM, González-Martínez SC, Cavers S. 2018. Molecular signatures of divergence and selection in closely related pine taxa. *Tree Genetics & Genomes* **14**: 83. DOI: https://doi.org/10.1007/s11295-018-1296-3

Wang IJ. 2020. Topographic path analysis for modeling dispersal and functional connectivity: Calculating topographic distances using the topoDistance R package. *Methods in Ecology and Evolution* **11**: 265-272. DOI: https://doi.org/10.1111/2041-210X.13317

Willyard A, Gernandt DS, Potter K, Hipkins V, Marquardt P, Mahalovich MF, Langer SK, Telewski FW, Cooper B, Douglas C, Finch K, Karemera HH, Lefler J, Lea P, Wofford A. 2016. *Pinus ponderosa*: A checkered past obscured four species. *American Journal of Botany* **104**: 161-181. DOI: https://doi.org/10.3732/ajb.1600336

Willyard A, Gernandt DS, Cooper B, Douglas C, Finch K, Karemera H, Lindberg E, Langer SK, Lefler J, Marquardt P, Pouncy DL, Telewski F. 2021. Phylogenomics in the hard pines (*Pinus* subsection *Ponderosae*; Pinaceae) confirms paraphyly in *Pinus ponderosa*, and places *Pinus jeffreyi* with the California big cone pines. *Systematic Botany* **46**: 538-561. DOI: https://doi.org/10.1600/036364421X16312067913435

Wright S. 1943. Isolation by distance. *Genetics* **28**: 114-138. DOI: https://doi.org/10.1093/genetics/28.2.114

Wu HX, Owen JV, Abarquez A, Matheson AC. 2004. Inbreeding in *Pinus radiata* - V. The effects of inbreeding on fecundity. *Silvae Genetica* **53**: 80-87. DOI: https://doi.org/10.1515/sg-2004-0015

Zhou Y, Duvaux L, Ren G, Zhang L, Savolainen O, Liu J. 2017. Importance of incomplete lineage sorting and introgression in the origin of shared genetic variation between two closely related pines with overlapping distributions. *Heredity* **118**: 211-220. DOI: https://doi.org/10.1038/hdy.2016.72

Associate editor: Jonás Aguirre Liguori

Author contributions: XAD, conceptualization, lab work, data analysis, writing; AML, conceptualization, data analysis, field work, writing; DSG, conceptualization, data analysis, field work, funding acquisition.

Supporting agencies: SEP-CONACYT (CB-2013 / 221694) grant awarded to David S. Gernandt.

Conflict of interests: The authors declare that there is no conflict of interest, financial or personal, in the information, presentation of data and results of this article.