Genomic early growth mechanisms of two endangered Mexican spruces

Carlos Alberto Segura-Sanchez
Universidad Juárez del Estado de Durango
Javier Hernández-Velasco
Universidad Intercultural de Baja California, San Quintín
Víctor Chano
University of Göttingen
Eduardo Mendoza-Maya
Universidad Nacional Autónoma de México
José Villanueva-Díaz
Instituto Nacional de Investigaciones Forestales
Pablito López-Serrano
Universidad Juárez del Estado de Durango
Christian Wehenkel
wehenkel@ujed.mx

Universidad Juárez del Estado de Durango

Research Article

Keywords: dendrogenomics, genotyping-by-sequencing, SNPs, Picea mexicana, Picea martinezii, common garden test

Posted Date: June 4th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9519860/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on July 23rd, 2026. See the published version at https://doi.org/10.1186/s12870-026-09500-y.

Abstract

This study elucidated the dendrogenomic growth mechanisms in the critically endangered endemic Mexican spruces *Picea martinezii* and *P. mexicana* by: (i) analyzing family- and population-level variations in seedling basal diameter (D), height (H) and seed weight (SW) after 12 months of growth under common garden conditions; and (ii) identifying genomic loci (SNPs) associated with these traits through genotyping-by-sequencing (GBS). Despite limited sample sizes (77 and 74 families representing all known populations of both species), we identified 32 and 10 outlier SNPs yielding 17 and six annotated candidate genes in *P. martinezii* and *P. mexicana*, respectively. These genes showed contrasting multivariate associations suggesting species-specific growth regulation: defense-oriented in *P. martinezii* and plasticity-driven in *P. mexicana*. Notably, several candidate genes encode key components of growth hormone pathways, including a gibberellin-regulated protein, a cytokinin hydroxylase and the AP2-like transcription factor ANT, providing direct mechanistic links between genetic variation and hormonal control of cell proliferation and organ size. Integration of these findings with the contrasting demographic histories of both species revealed that population bottlenecks enhance the detectability of growth-associated variants by reducing background genetic variation. These genomic resources provide actionable information for prioritizing conservation measures, implementing assisted gene flow to maintain adaptive potential under climate change and designing future breeding programs. With 80.9–99.6% sequence identity to conserved *Picea abies* homologs, these findings may extend across the genus.

Background

Plant growth is the irreversible increase in size that occurs through the combination of cell division, cell enlargement and cellular specialization. This fundamental, complex process is influenced by environmental (mainly light, water, CO_2, temperature, nutrients and hormones) and genetic factors [1]. However, genetic factors overwhelmingly determine intra and particularly interspecific variability in tree growth traits, while environmental factors explain a comparatively small proportion of such variability [2].

Dendrogenomics is an emerging interdisciplinary field that integrates dendrology, genetics and genomics to study the biological processes underlying wood formation and tree adaptation. While often associated with the analysis of mature tree rings, this field inherently encompasses fundamental tree-growth traits, which provide the phenotypic basis connecting environmental variation with genomic information [3]. Recent advances in genotyping large numbers of trees have made it increasingly possible to link genetic variation to specific growth-related phenotypes, even at early developmental stages [4].

The study of early tree growth is particularly important in forest genomics, as juvenile traits often serve as early indicators of long-term productivity and climate resilience [5]. Genome-wide association studies (GWAS) have identified key loci associated with seedling height and early biomass accumulation in conifers such as Douglas-fir [6], *Pinus ponderosa* [6], *Taxus baccata* [7] and *Picea crassifolia* [8]. Furthermore, genomic markers linked to early growth and physiological traits have been demonstrated in *Picea koraiensis* [9]. These findings underscore the potential of evaluating genomic variation at early developmental stages to accelerate conservation and selection efforts.

Dendrogenomic approaches have also revealed relationships between genetic variation and drought-induced dieback in *Nothofagus dombeyi* [10], adaptive capacity in *Abies marocana* forests [11] and responses to climatic stress in *Larix sibirica* [12] and *Pinus sibirica* [13]. Recently, Segura-Sanchez et al. [14] demonstrated genomic mechanisms of resilience in *Picea martinezii*. Notably, Chen et al. [15] conducted the largest (to date) genome-wide association study for a tree species (*Picea abies*), emphasizing the importance of evaluating early growth for genomic selection and designing conservation strategies.

The genus *Picea* comprises 37 species distributed across the Northern Hemisphere, with major diversity centers in eastern Asia, North America and Europe [16]. Genomic analyses have been conducted on numerous *Picea* species, driven by the ecological and economic importance of each, including *P. abies*, *P. glauca*, *P. sitchensis*, *P. engelmannii*, *P. crassifolia*, *P. mariana*, *P. martinezii* and *P. mexicana* [17, 8, 18, 14]. Mexico harbors three endemic spruce species representing the southernmost distribution of the genus [19]. Among these, *P. martinezii* and *P. mexicana* represent genetically distinct evolutionary lineages [16] whose limited distributions have led to some conservation concern. The species occur in contrasting elevational zones: *P. martinezii* occurs at mid-elevations (1,820–2,515 m; four populations, ~ 89,266 individuals, 157 ha in total) within the Sierra Madre Oriental, and *P. mexicana* occurs at high elevations (3,096–3,528 m; three populations, ~ 39,059 individuals, 173 ha in total) within the Sierra Madre Occidental and Oriental [19]. Both species are listed as endangered [20] and face ongoing pressures from fire, logging, habitat degradation and climate change [19, 14]. Mendoza-Maya et al. [21] analyzed the relationship between genetic diversity, population size and reproductive traits, reporting a positive correlation between population size and seed development and indicating that individual genome-wide heterozygosity was linked to seed traits in *P. martinezii*.

The present study aimed to elucidate the genomic growth mechanisms in *P. martinezii* and *P. mexicana* for the first time as follows: (i) by analyzing family- and population-related variations in seedling basal diameter (*D*) and height (*H*) after 12 months of growth under common garden conditions, representing early phenotypic growth traits, as well as seed weight (*SW*), a phenotypic maternal provisioning trait known to influence early tree growth [22]; and (ii) by identifying genomic loci (SNPs) associated with variations in *D*, *H* and *SW* through genotyping-by-sequencing (GBS). We hypothesized that the combined analysis of tree growth and genomic data will facilitate identifying biologically plausible candidate genes involved in tree-growth processes [3].

Materials and methods

Seed collection

Seeds and foliar tissue of *Picea martinezii* and *Picea mexicana* were collected from wild populations by the working group of Christian Wehenkel, Instituto de Silvicultura e Industria de la Madera (ISIMA), Universidad Juárez del Estado de Durango (UJED), Mexico under the corresponding permits issued by the Secretariat of Environment and Natural Resources, Mexico (SEMARNAT permit SGPA/DGVS/008797/18) and complied with Mexican environmental legislation and UJED institutional guidelines. Formal taxonomic identification was performed by Dr. Christian Wehenkel [20] using standard keys for Mexican

spruces. Voucher specimens were deposited in the Herbarium of the Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional (CIIDIR), Instituto Politécnico Nacional (IPN), Durango, Mexico. The deposition numbers for *P. martinezii* 67981–67984 and for *P. mexicana* 66539–66558 identified by Maria Socorro Gonzalez Elizondo.

Seeds were collected from all four documented populations of *Picea martinezii Patterson*, encompassing a broad range of effective population sizes (*Ne*), from very small populations (N_e = 55∼99) to a large population (N_e = 2,455), as reported by Mendoza-Maya [21], these populations are located in the Mexican state of Nuevo León, growing at 1,820–2,515 m elevation, with a mean annual temperature of 13.5°C and annual precipitation of 788 mm. On the other hand, we sampled the three reported populations of *Picea mexicana Martínez* with N_e ranging from 225 to 1,773 [21]. These populations are distributed across the states of Chihuahua (one population) and Coahuila (two populations) at elevations between 3,096 and 3,528 m, with a mean annual temperature of 9.3°C and annual precipitation of 858 mm (Fig. 1). Within each population, 12–30 dominant, phenotypically healthy trees were randomly selected, and 50–250 cones per tree were harvested during autumn 2018. This sampling yielded adequate seed quantities for genotyping and common garden trials, resulting in 92 *P. martinezii* and 87 *P. mexicana* paternal families [23].

Seed processing and phenotypic measurements

The phenotypic seedling dataset analyzed here is identical to that used in [20], where variance components and heritability were estimated using classical quantitative genetic approaches without molecular markers: a common garden experiment was established in May 2019 at a nursery in El Salto, Pueblo Nuevo, Durango, Mexico (23°47'11.72" N, 105°21'55.18" W; elevation of 2,580 m). The experimental site has a mean annual temperature of 11.2°C and a mean annual precipitation of 1,050 mm. This site is geographically and climatically intermediate between the native ranges of both species. These neutral common garden conditions allowed assessment of genetic effects on growth and survival independent of local environmental adaptation. The experiment continued until May 2020 (12 months).

A total of 5,548 seeds were sown for phenotypic evaluation from 92 *P. martinezii* and 87 *P. mexicana* parental trees (named hereafter “families”). Seed weight was determined as the weight of 1000 seeds (SW) per parental tree using an analytical balance (± g; Velab VE-5000H, Mexico), following [20]. Prior to sowing, seeds were surface-sterilized in 2% (v/v) hydrogen peroxide for 24 h and then sown individually in 170 cm³ plastic tubes containing a substrate mixture of 60% peat moss, 20% perlite, and 20% vermiculite, supplemented with 3 kg·m⁻³ controlled-release fertilizer (Multicote 15-7-15 + 2MgO + microelements). Tubes were arranged in 56 plastic trays (98 cavities per tray, 7 × 14 configuration) following a randomized complete block design, with family identity maintained throughout the experiment using the coding system "species—population— parental tree—seedling number".

Seedlings received standardized cultural treatments to ensure uniform growing conditions across families. Soluble fertilizer application followed a three-stage protocol: early growth (0–2 months: N-P-K 08-25-17 at 50 mg L^{-1}, weekly), rapid growth (2–6 months: N-P-K 20-09-20 at 20% v/v, weekly), and hardening (6–12 months: N-P-K 04-25-40 at 5% v/v, weekly). Automated irrigation was applied three times weekly. Environmental conditions were modified throughout the growing period: germination phase (0–6 months) under 720-gauge polyethylene cover for UV protection without shading; preconditioning phase (6–8 months) under 50% shade cloth; hardening phase (8–12 months) under full ambient conditions. Seedling basal diameter (D, mm) was measured monthly at the root collar using digital calipers (± mm precision; AVEDISTANT LCD6), and total height (H, mm) was measured from substrate surface to apical bud using a flexible tape measure (± 1 mm precision; Uline Accu-Lock H-1766). Survival was recorded monthly, with 4,753 seedlings (85.7%) surviving to 12 months. The family-level mean *D* and *H* (averaged across surviving offspring per parental tree) after 12 months of growth were used as phenotypic variables in genome-wide association analyses, along with mean seed weight per family.

DNA extraction, sequencing, genotyping, and SNP filtering

Although growth traits were measured on 4,753 individual seedlings, genomic data were obtained at the parental level, and associations were therefore explored at the family level rather than through individual-based genome-wide association analyses. Of the 92 *P. martinezii* and 87 *P. mexicana* parental trees initially sampled, 77 *P. martinezii* and 74 *P. mexicana* trees (families) were successfully genotyped and used for association mapping after quality control filtering (Table 1): genomic DNA was extracted from 15 mg of fresh needle tissue using a modified CTAB protocol [24]. Modifications included substitution of polyethylene glycol with 1.0% polyvinylpyrrolidone (PVP40), replacement of 5 M NaCl with LiCl, and addition of µL of 2-mercaptoethanol and 50 µL of Proteinase K (20 mg mL^{-1}) to improve DNA yield and purity [24]. DNA concentration and quality were assessed using a NanoDrop spectrophotometer (Thermo Fisher Scientific), with samples exhibiting A_{260}/A_{280} ratios of 1.8–and A_{260}/A_{230} ratios >selected for library preparation.

Table 1

Descriptive statistics for mean seedling base diameter, height and seed weight in Picea martinezii and P. mexicana parental families after growth for 12 months under common garden conditions.

Species	Trait	n	Mean ± SD	Range	CV (%)	Skewness	Kurtosis
P. martinezii	Diameter (mm)	290	1.53 ± 0.30	0.90–2.54	19.6	0.34	0.21
	Height (cm)	290	10.76 ± 3.29	4.37–21.36	30.6	0.47	-0.12
	Seed weight (g)*	290	0.64 ± 0.12	0.31–0.95	19.2	0.18	-0.31
P. mexicana	Diameter (mm)	416	1.33 ± 0.14	0.99–1.91	10.5	0.42	0.67
	Height (cm)	416	9.10 ± 2.19	5.22–15.21	24.1	0.38	-0.09
	Seed weight (g)*	416	0.63 ± 0.08	0.43–0.94	13.9	0.23	-0.18

Note: * = per 1000 seeds; CV = Coefficient of variation; Skew = Skewness; Kurt = Excess kurtosis. Welch's t-tests detected significant differences between *P. martinezii* and *P. mexicana* for all three traits: basal diameter ($t_{118.7} = 5.23$, $p = 7.5\times10^{-7}$), height ($t_{148.4} = 19.84$, $p < 2.2\times10^{-16}$) and seed weight ($t_{72.1} = 31.69$, $p < 2.2\times10^{-16}$). Welch's two-sample t-tests were performed to test for significant differences in trait means between *P. martinezii* and *P. mexicana*, accounting for unequal variances between species.

Genotyping-by-sequencing (GBS) libraries were prepared following [25]. Briefly, 200 ng of high-quality genomic DNA per sample was double-digested with restriction enzymes PstI (R0106L) and MspI (R3140L; New England Biolabs, USA). Sample-specific barcodes were ligated using T4 DNA ligase (New England Biolabs), and libraries were pooled, purified using AMPure XP beads (Beckman Coulter), and amplified via PCR at the *Plateforme d'analyses génomiques, Institut de biologie intégrative et des systèmes* (IBIS), Université Laval, Canada. Paired-end sequencing (2 × 150 bp) was performed on an Illumina NovaSeq 6000 S4 platform (one lane, 330 sequencing cycles including index reads) at the *Centre d'expertise et de services Génome Québec, Canada*.

Raw sequencing reads were quality-filtered, adapter-trimmed, and *de novo* assembled into locus-specific catalogs using TrimGalore [26] with default parameters optimized for outbreeding species (minimum stack depth = 3, maximum nucleotide mismatches = 2, maximum indels = 3). Separate *de novo* assemblages were generated for *P. martinezii* and *P. mexicana* to account for potential species-specific genomic variation. Variant calling was performed within each species, and resulting VCF files were filtered using VCFtools v0.1.17 [27] with the following parameters: minimum minor allele count (--mac) = 2, minimum mean depth (--min-meanDP) = 10, maximum mean depth (--max-meanDP) = 120, and minimum genotyping rate per SNP = 80% (--max-missing 0.8). After filtering, 19,057 high-quality SNPs were retained for *P. martinezii* and 31,678 for *P. mexicana* (for complete bioinformatic pipeline details, see [22] and [14]).

Statistical strategy and methodological rationale

To optimize the detection of relevant genomic signals with a limited sample size (Table 1), a two-step procedure was chosen. First, we identified outlier SNPs (*i.e.*, instead of genotype-first, an Outlier-first approach) (see details in Section 2.5). Subsequently, the covariation of these candidates with the phenotypes was evaluated. A classical genome-wide scan of all SNPs would have led to a prohibitive burden of multiple testing with Bonferroni correction (i.e., p < 0.0000003). This threshold massively increases the risk of type II errors (false negatives) and would mask polygenic signals of small effect size [28, 29]. The targeted preselection of outlier SNPs acts as dimension reduction, which lowers the multiple-testing hurdle and increases statistical power within this rare study system [30]. The use of the Spearman rank correlation (r_s) also allows the detection of robust, non-linear gene effects without a normal distribution assumption [31] (see in Section 2.6).

However, outlier preselection can artificially inflate *p*-values through circular inference, increasing the risk of false positives. Moreover, correlations may be driven by population structure or genetic drift rather than functional causality [32]. The chosen approach therefore does not allow precise estimation of direct gene effect sizes. To assess the robustness of the pre-selected outlier SNPs, a genome-wide association analysis using a Mixed Linear Model (MLM) was applied to the complete GBS-derived SNP dataset, which controls for population structure and kinship [33] (see Section 2.6).

A post-hoc functional annotation (see in section 2.7) is used to distinguish between stochastic noise and true signals. The identification of evolutionary conserved regulators of plant development is considered as independent biological evidence. A high concordance between statistical loci and plausible gene functions validates that the pipeline has captured functional genomic regions and no statistical artifacts.

Identification of outlier SNPs

Four methodologies were employed to detect outlier SNPs in *P. martinezii* and *P. mexicana*: i) PCAdapt v.4.3.3 package [34, 35]; ii) genotype-environment association (GEA) approach using latent factor mixed models (LFMM2) from the R package LEA [36, 37, 38]; iii) the Local Outlier Factor (LOF) algorithm from the Python package scikit-allel Python library (https://scikit-allel.readthedocs.io) [39, 40]; and iv) OutFLANK v.0.2 R package [41, 38].

PCAdapt offers benefits over alternative genome-scanning tools such as BayeScan, hapflk, OutFLANK, and sNMF, particularly regarding computational efficiency, capacity to manage missing data, and individual-based approach that obviates population grouping requirements [34]. This package implements principal component analysis (PCA)-based outlier SNP detection while controlling for population structure [34]. The methodology relies on Mahalanobis distance (Dm) statistics, generating a Z-score vector through regression of each SNP against *K* (*K* = 2) principal components [42]. Principal component number *K* was determined applying the Cattell's criterion [43]. *P*-values were derived by transforming Dm via chi-square distribution [43]. Bonferroni adjustment was implemented, designating SNPs with corrected *p*-values < 0.05 as selection candidates.

LFMM2 facilitated GEA assessment. Non-collinear bioclimatic and edaphic parameters were selected using absolute Spearman rank correlation (|rs|) < 0.7, yielding six bioclimatic and thirteen edaphic variables. Analysis incorporated four latent factors determined by variance explained per principal component for climatic and edaphic parameters independently. Bonferroni correction was applied to minimize false discovery rates, with significance threshold set at adjusted *p*-value < 0.05.

The LOF algorithm utilized 20 nearest neighbors (K = 20) to optimize deviation detection while maintaining precision. The kNN-LOF algorithm employs distance-based weightings to modulate neighbor influence, enabling robust performance with irregularly distributed datasets [44], incorporating a contamination threshold of 1%, Z threshold of 2, and Manhattan distance metric [45].

OutFLANK identifies outlier SNPs through *F*-statistics, applying standard parameters: neutral loci mean *F*-statistic of 0.05, neutral loci proportion of 0.95, and minimum degrees of freedom (*dfmin*) of 10. The algorithm automatically excludes extreme *Fst*-value loci to infer neutral distribution, trimming 5% from each *Fst* distribution tail by default, assuming central 90% represents neutral variation. The 95% neutral loci proportion reflects an implicit 5% contamination rate (1–0.95), representing the maximum non-neutral loci fraction tolerable without biasing neutral distribution estimation [46].

Association analyses between growth parameters and SNP genotypes

Association analyses were conducted to examine the relationship between single nucleotide polymorphism (SNP) genotypes and early phenotypic tree-growth traits following the protocols established by [14], that tested SNP – tree ring resilience indices. For this purpose, annual mean values of seedling measurements were correlated with the GBS-derived SNP markers obtained from parent trees.

The use of family-level mean phenotypes as the response variable is a well-established strategy in forest genomics. This approach integrates maternal genomic data with the additive genetic performance of the offspring while effectively reducing individual environmental micro-variation [47, 30]. To distinguish the genetic signal of the seedling basal diameter (*D*) and height (*H*) after 12 months growing under common garden conditions, from maternal effects (maternal provisioning), seed weight (*SW*) was recorded for each parental tree and incorporated into subsequent models [23].

Each outlier SNP was numerically classified into genotypic classes "0", "1", and "2", corresponding respectively to homozygous genotypes for the major allele (AA), heterozygous genotypes (Aa), and homozygous genotypes for the minor or alternate allele (aa). Missing genotypic data were imputed using the median value derived from the respective SNP dataset.

Statistical correlations between individual SNPs and growth metrics were evaluated using Spearman's rank correlation coefficient (r_s), a non-parametric measure appropriate for data that may not follow normal distributions. The correlation analyses were performed separately for the variables *D* and *H*, and *SW* of both *P. martinezii* and *P. mexicana* using the Hmisc package in R v.4.[31]. Spearman's r_s values and their corresponding *p*-values were calculated for each SNP-variable combination [48, 49, 50, 51]. A Bonferroni-adjusted significance level was calculated to account for the increased possibility of type-I error (false positives).

Additionally, a Principal Component Analysis (PCA) was performed using the prcomp function in *R* [38]. The PC1 was subsequently used to determine r_s and their *p*-value between PC1 scores and each of the seedlings *D*, *H* or *SW* variable using only those outlier SNPs that demonstrated statistically significant r_s ($p < 0.05$) with any of these variables.

To further validate the exploratory findings, a genome-wide association analysis was implemented using the GBS-derived SNP dataset via MLM calculated by the GAPIT3 package [33]. MLM was employed that simultaneously accounted for *SW* as a fixed covariate and corrected for individual kinship and population structure via Principal Components (PCs) [33]. While r_s captures robust, monotonic relationships, the MLM provides a stringent framework to neutralize complex genomic backgrounds and maternal provisioning.

A Random Forest (*RF*) regression [52] was implemented to evaluate the predictive power of the identified candidate SNPs and their epistatic interactions affecting both seedling growth (*D*, *H*) and *SW* using the randomForest [53] and caret packages [54] in *R*. This non-parametric framework accounts for non-linear relationships that might be overlooked by standard linear models. Multivariate associations were quantified using the pseudo coefficient of determination (R^2_{rf}) and 99.9% confidence intervals ($CI_{99.9\%}$) derived from 1,000 bootstrap iterations. Furthermore, to disentangle the relative contributions of genetic factors and the other two of three variables *D*, *H* and *SW*, block-wise conditional variance partitioning ($\Delta R^2_{rf,c}$) was employed [55]. This approach allowed us to quantify the unique explanatory power of putatively growth-related SNPs while controlling for shared effects with two of three variables.

Results were compared across all methods. The identification of top-ranked SNPs demonstrating significant signals consistently across MLM (controlled for *SW* and kinship), Spearman's r_s, and *RF* analysis was established as the primary criterion for biological validity. This methodological triangulation ensures that the detected candidate SNPs represent functional signals rather than statistical artifacts or demographic effects [33, 3].

Post-hoc functional annotation of putatively growth-associated SNPs

Post-hoc functional annotation was performed to validate the biological relevance of the identified associations and to distinguish between potential statistical artifacts and genuine genomic signals. Genomic sequences flanking growth-associated SNPs (120–410 bp fragments) were extracted from the assembled loci for both species and compared against the *P. abies* reference genome [17] using BLASTn [56] via the PlantGenIE.org platform [57]. Homologous matches were identified using stringent criteria: *E*-value < 10^{-5}, query coverage > 70%, and sequence identity > 80%. For each significant hit, we recorded the top *P. abies* gene model ID, gene function annotation, sequence similarity percentage, and *E*-value. Multiple sequence alignments of *P. martinezii, P. mexicana*, and *P. abies* homologs were generated using AliView [58] to verify orthology and visualize SNP positions relative to conserved domains. SNPs that could not be annotated (*E*-value > 10^{-5} or no significant BLAST hit) were classified as non-annotated and retained for analysis as they may represent species-specific or rapidly evolving genomic regions.

Population-level validation of genotype-phenotype associations

Finally, significant differences in the frequency of homozygous minor-allele genotypes (aa) at putatively growth-related (i.e., *D*, *H*, *SW*) SNPs in both spruce species were assessed at the population level using Dunn's post-hoc pairwise comparisons with Bonferroni correction for multiple testing by FSA [59] and dunn.test packages in *R* [60].

Results

Seedling diameter, height and seed weight measurements

Seedling growth after 12 months under common garden conditions revealed significant variation across parental families and pronounced differences between species. *Picea martinezii* seedlings exhibited significantly larger mean basal diameter (*D*) (1.53 ± 0.30 mm vs. 1.33 ± 0.14 mm; $t_{118.7} = 5.23$, $p = 7.5\times10^{-7}$), greater mean height (*H*) (10.76 ± 3.29 cm vs. 9.10 ± 2.19 cm; $t_{148.4} = 19.84$, $p < 2.2\times10^{-16}$) and greater mean seed weight (*SW*) (0.64 ± 0.12 g vs. 0.63 ± 0.88 g per 1000 seedss; $t_{72.1} = 31.69$, $p < 2.2\times10^{-16}$) than *P. mexicana* (Table 1; Table S1). Notably, *P. martinezii* displayed greater phenotypic variability across all traits (CV = 19.2–30.6%) than *P. mexicana* (CV = 10.5–24.1%), suggesting broader adaptive potential or greater environmental sensitivity in the former species.

Seedling basal diameter (*D*) and height (*H*) were strongly correlated within both *P. martinezii* ($r^2 = 0.73$, $p = 7.35\times10^{-14}$) and *P. mexicana* ($r^2 = 0.84$, $p = 1.19\times10^{-20}$). Seed weight (*SW*) was moderately and positively correlated with both *D* and *H* in *P. martinezii* ($r^2 = 0.39$ for both traits, $p < 0.001$) but was not significantly correlated with either *D* or *H* variable in *P. mexicana* ($r^2 = -0.05$ for *D* and 0.04 for *H*, $p > 0.70$), indicating that maternal provisioning effects on early growth are species specific.

Candidate SNP identification for diameter, height and seed weight

LFMM2 filtering confirmed 19 of the 32 PCAdapt-filtered SNPs (59%) in *P. martinezii* and six of ten PCAdapt-SNPs (60%) in *P. mexicana*. Notably, neither OutFLANK nor LOF detected outlier SNPs in either species. In *P. martinezii*, PC1 of the principal component analysis including the 32 outlier SNPs was significantly associated with *D* (*r*s, PC1 = − 0.52, $p < 0.00001$), *H* (*r*s, PC1 = − 0.40, $p = 0.0003$) and *SW* (*r*s, PC1 = − 0.42, $p = 0.0001$) after Bonferroni correction ($p < 0.001$). Remarkably, all 32 PCAdapt-SNPs exhibited uniformly negative correlations with both *D* and *H*. After Bonferroni correction, all 32 SNPs retained statistical significance for at least one trait, with 29 significantly associated with both *D* and *H*. The strongest associations (*r*s < − 0.45) and SNPs verified by MLM ($p < 0.05$) are highlighted in Table 2. The 99.9% confidence intervals (CI 99.9%) of the pseudo-coefficient of determination (R^2_{rf}) for the multivariate associations of *D*, *H*, and *SW* with the 32 outlier SNPs (0.37–0.66, 0.32–0.60, and 0.10–0.56, respectively) provide strong evidence for robust associations for *D* and *H*, whereas the broader and lower CI 99.9% for *SW* suggests a weaker model.

Table 2
Genomic loci associated with growth traits and their functional annotation in two endangered Mexican spruce specie

Species		Seedling diameter		Seedling height		Seed weight						
	SNPs	r_s	*p*-value	r_s	*p*-value	r_s	*p*-value	Q.S. (b.p)	Gene	G.R.	Similarity *Picea abies*	*E*-va
P. martinezii	PC1ѧ	-0.517	0.0000014	-0.404	0.00027	-0.423	0.0001					
	50362:161ѧ	-0.511	$3.19E^{-11}$	-0.373	0.000003	-0.269	0.0009	377	MA_10292150	NA	98.57%	6.5 62
	33345:136*ѧ	-0.487	$3.36E^{-10}$	-0.493	$1.88E^{-10}$	-0.218	0.007	220	NA	NA	NA	NA
	22169:33*ѧ	-0.475	$1.03E^{-09}$	-0.484	$4.39E^{-10}$	-0.267	0.001	253	MA_10435027g0010	2607	98.38%	2.3 117
	28920:117ѧ	-0.469	$1.81E^{-09}$	-0.470	$1.65E^{-09}$	-0.237	0.003	377	MA_10204399g0010	1887	94.75%	1.1 58
	8237:220ѧ	-0.466	$2.27E^{-09}$	-0.456	$5.50E^{-09}$	-0.275	0.0006	250	MA_34782g0010	807	89.07%	2.1 112
	7331:244ѧ	-0.435	$3.09E^{-08}$	-0.455	$6.05E^{-09}$	-0.326	0.0001	255	MA_34146	NA	88.06%	4.9 114
	8789.65ѧ	-0.435	$3.16E^{-08}$	-0.454	$6.29E^{-09}$	-0.275	0.0006	201	MA_10435550g0020	1002	98.31%	3.2 95
	39836:85*ѧ	-0.425	$7.26E^{-08}$	-0.438	$2.47E^{-08}$	-0.164	0.045	188	MA_84859g0010	288	91.67%	1.2 60
	2645:276ѧ	-0.425	$7.28E^{-08}$	-0.438	$2.51E^{-08}$	-0.153	0.062	290	MA_6446g0010	5997	82.78%	3.5 103
	38134:111*ѧ	-0.422	$8.64E^{-08}$	-0.434	$3.49E^{-08}$	-0.163	0.047	248	NA	NA	NA	NA
	21038:50*ѧ	-0.418	$1.21E^{-07}$	-0.434	$3.37E^{-08}$	-0.246	0.002	410	MA_101621g0020	3924	87.74%	1.2 121
	5866:17ѧ	-0.412	$1.87E^{-07}$	-0.431	$4.35E^{-08}$	-0.153	0.0626	284	MA_7154225	NA	80.89%	3.2 72
	44879:209ѧ	-0.406	$2.94E^{-07}$	-0.409	$2.31E^{-07}$	-0.158	0.054	357	MA_10150655g0010	1116	97.93%	1.7 112
	55:46*ѧ	-0.391	$8.74E^{-07}$	-0.414	$1.68E^{-07}$	-0.194	0.017	342	MA_444615g0010	993	96.97%	2.2 104
	16330.110ѧ	-0.388	0.000001	-0.371	0.000003	-0.286	0.0004	360	NA	NA	NA	NA
	34326.285ѧ⋃	-0.38	0.000002	-0.359	0.000007	-0.307	0.0001	392	MA_184747g0010	4128	87.10%	4.3 71
	24370:25*ѧ	-0.374	0.000003	-0.388	0.000001	-0.185	0.023	178	MA_15508g0010	3369	93.82%	3.2 69
	41521:196*ѧ	-0.369	0.000004	-0.216	0.008	-0.049	0.551	324	MA_31722g0010	924	89.14%	3.8 135
	22682:63*ѧ	-0.368	0.000004	-0.382	0.000002	-0.077	0.35	291	MA_10429973g0010	2978	87.13%	5.3 101
	29628.135ѧ⋃	-0.368	0.000004	-0.289	0.0003	-0.326	0.00005	282	MA_10221218g0010	1656	98.75%	2.1 157

Note: (r_s) = Spearman correlation coefficient; Q.S. = Query sequence length (bp); Gene = *P. abies* gene model ID; G.R. = Genomic reference; *E*-value = BLAST e principal component from PCA of outlier SNPs. *SNPs validated by both PCAdapt and LFMM2; ѧ = SNPs significant after Bonferroni correction ($\alpha = 0.001$ ar significant BLAST hit. ⋃ = verified by MLM ($p < 0.05$); ⋃=verified by MLM ($p = 0.06$)

Species		Seedling diameter		Seedling height		Seed weight						
	SNPs	r_s	*p*-value	r_s	*p*-value	r_s	*p*-value	Q.S. (b.p)	Gene	G.R.	Similarity *Picea abies*	*E*-va
	29628.135ሐ⋃	-0.368	0.000004	-0.289	0.0003	-0.326	0.00005	282	MA_766447g0010	3000	82.35%	2.1 81
	7241:238*ሐ	-0.364	0.000005	-0.380	0.000002	-0.176	0.032	290	MA_47536g0010	645	84.34%	9.(111
	5566:269*ሐ	-0.355	0.000009	-0.369	0.000004	-0.267	0.001	289	NA	NA	NA	NA
	473:152*ሐ	-0.348	0.000014	-0.380	0.000002	-0.187	0.022	205	NA	NA	NA	NA
	47948:247*ሐ	-0.326	0.00005	-0.327	0.000048	-0.157	0.056	256	MA_338690g0010	888	92.13%	1.9 99
	35763:31*ሐ⋃	-0.310	0.0001	-0.325	0.00005	0.033	0.689	344	NA	NA	NA	NA
	35337:53*ሐ	-0.301	0.0002	-0.314	0.0001	-0.085	0.301	300	MA_126211g0010	2880	99.60%	8.(123
	51806.107ሐ	-0.270	0.0009	-0.282	0.0005	-0.269	0.0009	258	MA_10429567g0010	995	97.44%	1.3 133
	55487:25*ሐ	-0.266	0.0011	-0.283	0.0004	-0.038	0.638	228	NA	NA	NA	NA
	38814:35*ሐ	-0.265	0.0011	-0.292	0.0003	-0.062	0.447	212	NA	NA	NA	NA
	90892.226ሐ	-0.252	0.0019	-0.266	0.001	-0.275	0.0006	280	NA	NA	NA	NA
	57886:61*ሐ	-0.164	0.045	-0.231	0.004	-0.270	0.0008	368	NA	NA	NA	NA
	34589:168*ሐ	-0.154	0.061	-0.169	0.039	-0.119	0.147	272	NA	NA	NA	NA
P. mexicana	PC1ሐ	0.441	0.00008	0.422	0.00018	NA	NA					
	51510:145ሐ	0.377	0.0009	0.248	0.033	-0.120	0.306	180	NA	NA	NA	NA
	5754:95ሐ	0.373	0.001	0.299	0.009	-0.088	0.455	360	NA	NA	NA	NA
	9021:109ሐ	0.371	0.001	0.300	0.009	-0.063	0.592	416	MA_3967g0010	2000	93.91%	2.(45
	32877:235	0.322	0.005	0.243	0.036	-0.044	0.704	312	MA_19894g0010	720	96.83%	1.4 113
	4124:36*⋃	0.316	0.006	0.297	0.01	-0.116	0.324	335	MA_10429901g0010	3000	88.66%	7.(24
	50422:144*	0.292	0.011	0.250	0.031	-0.201	0.084	228	MA_140203g0010	2674	97.58%	4.(73
	44765:107	-0.172	0.14	– 0.235	0.043	0.012	0.913	300	NA	NA	NA	NA
	65739:66 ⋃	-0.202	0.083	– 0.302	0.008	-0.029	0.802	180	MA_377091g0010	893	90.71%	6.7 33
	34584:94	-0.229	0.049	– 0.257	0.026	0.139	0.235	240	NA	NA	NA	NA

Note: (r_s) = Spearman correlation coefficient; Q.S. = Query sequence length (bp); Gene = *P. abies* gene model ID; G.R. = Genomic reference; *E*-value = BLAST e principal component from PCA of outlier SNPs. *SNPs validated by both PCAdapt and LFMM2; ሐ = SNPs significant after Bonferroni correction (α = 0.001 ar significant BLAST hit. ⋃ = verified by MLM ($p < 0.05$); ⋃=verified by MLM ($p = 0.06$)

Species		Seedling diameter		Seedling height		Seed weight						
	SNPs	r_s	*p*-value	r_s	*p*-value	r_s	*p*-value	Q.S. (b.p)	Gene	G.R.	Similarity *Picea abies*	*E*-va
	53265:41*	-0.246	0.034	–0.275	0.017	0.088	0.452	240	MA_112260g0010	2280	93.46%	1.7 10

Note: (r_s) = Spearman correlation coefficient; Q.S. = Query sequence length (bp); Gene = *P. abies* gene model ID; G.R. = Genomic reference; *E*-value = BLAST e principal component from PCA of outlier SNPs. *SNPs validated by both PCAdapt and LFMM2; ⋔ = SNPs significant after Bonferroni correction (α = 0.001 ar significant BLAST hit. ⋃ = verified by MLM ($p < 0.05$); ⋃=verified by MLM ($p = 0.06$)

Conditional *RF*–based variance partitioning ($\Delta R^2_{rf,c}$) revealed substantial shared explanatory contributions among phenotypic variables *D*, *H* and *SW* beyond the variance explained by outlier SNPs alone (i.e. controlling joint effects with two of these three variables). When *D* was modeled as the response variable, the combined inclusion of *SW* and *H* increased the model explanatory power by $\Delta R^2_{rf,c}$ = 0.23 relative to a SNP-only model, with a wide but consistently positive bootstrap confidence interval ($CI_{99.9\%}$ 0.08–0.44). Similarly, modeling *H* as the response showed that *D* and *SW* together explained an additional $\Delta R^2_{rf,c}$ = 0.29 of variance beyond SNP effects ($CI_{99.9\%}$ 0.08–0.52). Conversely, early growth traits (*D* and *H*) accounted for an additional $\Delta R^2_{rf,c}$ = 0.22 of variation in *SW* after controlling for outlier SNP associations ($CI_{99.9\%}$ 0.04–0.55). However, permutation tests indicated that these conditional trait effects were not statistically separable from SNP-associated variation (p = 0.92–0.99).

In *P. mexicana*, the PC1 of the principal component analysis including the ten outlier SNPs revealed significant positive associations with *D* (r_s = 0.44, p = 0.00008) and *H* (r_s = 0.42, p = 0.0002), contrasting sharply with the negative associations observed in *P. martinezii*. Unlike *P. martinezii*, only the SNPs 51510:145, 5754:95 and 9021:109 showed moderate significant positive associations with *D* (r_s = + 0.37–0.38) after Bonferroni correction ($p < 0.004$) (Table 2). MLM confirmed the preselected SNPs 65739:66 (p = 0.04) and 4124:36 (p = 0.06). The $CI_{99.9\%}$ of R^2_{rf} for the largest multivariate association (i.e. for *D*, *H*, and *SW* with the ten outlier SNPs) were [0.08–0.37], [0.16–0.42] and [0.0004–0.17], respectively, demonstrating that *D* and *H* show meaningful single-SNP associations. For *SW*, however, the $CI_{99.9\%}$ extended almost to zero, indicating that the data do not provide strong support for a genuine association. In contrast to *P. martinezii*, $\Delta R^2_{rf,c}$ revealed markedly stronger phenotypic integration among *D*, *H* and *SW*. The phenotypic variable blocks explained between 47% and 64% additional variance beyond single-SNP models, with consistently positive and relatively narrow bootstrap confidence intervals. Permutation tests in *P. mexicana* yielded lower *p*-values than in *P. martinezii*, particularly for the effect of early growth traits on seed weight (p = 0.26).

Annotation of candidate genes in Picea martinezii and P. mexicana

In the absence of reference genomes for *P. martinezii* and *P. mexicana*, the annotated *P. abies* genome was used to perform BLAST (Basic Local Alignment Search Tool) searches and provide functional annotation of sequences containing outlier SNPs. With genetic similarities ranging from 80.89% to 99.60% (*E*-values 10^{-58} to 10^{-157}), the 160–377 bp (base pairs) DNA sequences including growth-associated SNPs ($p < 0.05$) in *P. martinezii* were located in sequences of *P. abies* genes encoding the following: histone-lysine N-methyltransferase *Arabidopsis trithorax* (ATX4)-like (MA_10433638g0010), tobacco mosaic virus (TMV) resistance N-like protein (MA_11646g0010), AP2 (APETALA2)-like ethylene-responsive transcription factor ANT (AINTEGUMENTA) (MA_10434193g0010), gibberellin-regulated family protein (MA_10269445g0010), cytokinin hydroxylase-like (MA_130655g0010), water-stress inducible protein (MA_19119g0010), lipase-like PAD4 (Phytoalexin Deficient 4) (MA_9176g0010), beta-catenin 1 (MA_132337g0010), glutamine synthetase (MA_10429533g0010), cytochrome c oxidase subunit 2 (MA_10435788g0010), cytochrome P450 CYP76AA22 (MA_8449754g0010), kinesin KP1 (MA_10268978g0010) and ATP synthase F1 assembly factor 2 (MA_93821g0010). Moreover, four of the significantly growth-associated SNPs ($p < 0.001$) were found in *P. abies* genes with no documented functionality. Finally, 12 of these SNPs did not correspond to *P. abies* genes (Table 2).

In *P. mexicana*, with genetic similarities ranging from 88.66% to 97.58% (*E*-values 10^{-24} to 10^{-113}), growth-associated SNPs ($p < 0.05$) were located in *P. abies* genes encoding: mediator of RNA polymerase II transcription subunit 15a-like (MA_10430289g0010), suppressor of auxin resistance 1 SAR1/NUP160 (Nucleoporin 160) (MA_10427978g0010), 2-oxoglutarate and Fe(II)-dependent oxygenase superfamily protein (MA_10427396g0010), Class I glutamine amidotransferase-like superfamily protein (MA_19064g0010) and ER membrane protein DUF962 (Domain of Unknown Function 962) (MA_116398g0010). Finally, six of the growth-associated SNPs (60%; $p < 0.004$) lacked reliable matches to *P. abies* genes (Table 2).

Population-level validation of genotype-phenotype associations

Population-level genomic analyses revealed significant differences in the frequency of homozygous minor-allele genotypes (aa) at growth (i.e. seedling *D*, *H* and *SW*)- related SNPs in both spruce species. In *P. martinezii*, the La Encantada (LE) population showed the significantly highest median aa frequency (41.5%) together with the almost always significantly lowest median seedling diameter (1.2 mm), height (8.9 cm), and seed weight (0.37 g per 1000 seeds). By contrast, the Agua Fría (AF) population exhibited an aa frequency close to zero (0.0%) and in many cases the significantly highest median diameter (1.4 mm), height (11.9 cm) and seed weight (0.52g per 1000 seeds). In *P. mexicana*, the El Mohinora (EM) population displayed the significantly highest, but not significantly different, aa frequency (43.0%) and the significantly highest median seedling diameter (1.3 mm) and height (8.3 cm). Conversely, the El Coahuilón (EC) population showed the significantly lowest aa frequency (9.6%) along with the significantly lowest diameter (1.1 mm) and height (6.9 cm). The La Marta (LM) population exhibited intermediate values for both genotype frequencies (25.2%) and growth traits (Tables 1, S2, S3, S4, S5).

Discussion

The study findings support our hypothesis that integrating tree growth-related traits and genomic data will enable the identification of key genes involved in dendrogenomic growth mechanisms [3]. In *P. martinezii*, significant associations were detected between 32 outlier SNPs and variations in seedling basal diameter (D) and height (H) after 12 months of growth under common-garden conditions and also in seed weight (SW). This indicates that both early growth traits and the maternal provisioning trait (influencing early growth) (up to r²,PC1 = -0.52, $p < 0.00001$ in *P. martinezii*) [23] may contribute to the genomic signatures of growth. Similarly, in *P. mexicana*, three outlier SNPs and PC1 (up to r²,PC1 = + 0.44, $p < 0.0001$) were significantly and positively correlated with early growth traits, further substantiating the hypothesized link between the growth phenotype and the underlying genomic variation (Table 2). Conditional variance partitioning (ΔR^2rf,c) indicated that these conditional trait effects were not statistically separable from SNP-associated variation, reflecting extensive shared explanatory structure between growth traits, maternal provisioning and SNP-based predictors, rather than independent additive contributions.

Despite the observation of putative SNP–trait associations (Tables 1 and 2), evidence for adaptive divergence remains limited. Thus, only 19 of the 32 SNPs in *P. martinezii* (but none of the three SNP loci in *P. mexicana*) (Table 2) were jointly identified by PCAdapt and LFMM2, and these signals were not supported by additional outlier tests. Together with the restricted genomic coverage of GBS [61], this substantially weakens confidence in strong selection. Nonetheless, the overall pattern is more compatible with locally varying selective pressures (local adaptation) acting across heterogeneous environments than with balancing selection, which typically maintains variation within more uniform populations [62, 63]. Population-based results support local adaptation by selection at these putatively growth-related SNP loci. Both the LE (*P. martinezii*) and EM (*P. mexicana*) populations showed a strongly increased proportion of aa homozygotes. The LE population, characterized by low growth (i.e. small D and H) and small SW, was the *P. martinezii* population growing at the highest elevation. The EM population, with above average growth and very large SW, was the *P. mexicana* population growing at the lowest elevation [20] (Table S4, Fig. 1).

Beyond these methodological constraints, false positive associations in GWAS can also arise from circular inference when the same data are used for both SNP discovery and validation, as well as from confounding effects of population structure and genetic drift [32]. However, the biological plausibility of the candidate genes mitigates these concerns: the 17 genes identified in *P. martinezii* and six identified in *P. mexicana* form clusters in functional categories with well-established roles in growth regulation, including defense-growth trade-offs, hormonal signaling (gibberellins, cytokinins, ethylene), epigenetic control and energy metabolism. Although not structural growth genes per se, these genetic categories are linked to key regulatory and metabolic bottlenecks that modulate growth in long-lived forest trees. Considered together, these modules suggest that variation in growth is driven by integrated gene regulatory networks rather than isolated effects of individual growth genes [62].

In *P. martinezii*, growth-associated SNPs clustered into four major functional modules: (1) defense and stress perception, primarily located at the plasma membrane and cytoplasm, which likely modulates growth through trade-offs between protection and biomass accumulation (Calcium-dependent kinase [64], disease resistance protein TIR-NBS-LRR class [65], TMV resistance N-like [66], lipase-like PAD4 [67], water-stress inducible protein [68], GPI-anchored adhesin PGA55 related to cell adhesion [69], cytochrome P450 CYP76AA22 [70]; (2) hormonal signaling pathways controlling cell division and elongation (gibberellin-regulated family protein [71]; cytokinin hydroxylase-like [74]; AP2-like ethylene-responsive transcription factor ANT [72]; beta-catenin 1 [73]); (3) epigenetic regulators shaping the stability and responsiveness of growth programs (histone-lysine N-methyltransferase ATX4-like [75]; and (4) mitochondrial and chloroplast energy metabolism (cytochrome c oxidase subunit 2 [76]; ATP synthase F1 assembly factor 2 [77]; glutamine synthetase [78]; kinesin KP1 [79]), which constrains carbon and nitrogen allocation to growth. Additional genes are involved in protein turnover in nucleus and cytoplasm (E3 ubiquitin-ligase Praja-2) [80]). Among the candidates identified, the AP2-like ethylene-responsive transcription factor ANT stands out as the most prominent growth-related gene, given its well-established role in controlling cell proliferation, organ size and overall plant growth, with additional support from gibberellin- and cytokinin-associated genes that directly regulate cell division and elongation [71, 74, 72]. Both a SNP in the Calcium-dependent kinase gene and in the disease resistance protein TIR-NBS-LRR gene were verified by the MLM approach, reinforcing the robustness of these associations ($p = 0.01$). Notably, one SNP showed homology to a sequence annotated as Internalin A in PlantGenIE (Table 2). Although Internalin A was originally characterized as an adhesion protein in the bacterium *Listeria monocytogenes* [81], its detection here likely reflects a plant homolog containing conserved leucine-rich repeat (LRR) domains. Such LRR-containing proteins are widespread in plants and function in cell-cell adhesion and receptor-mediated signaling pathways [82, 83, 84].

By contrast, the six annotated genes are linked to growth-associated SNPs in *P. mexicana* mapped to transcriptional regulation in the nucleus and endoplasmic reticulum (mediator of RNA polymerase II transcription subunit 15a-like protein (MED15a) [85]; suppressor of auxin resistance 1 (SAR1) [86]), signal transduction at the plasma membrane (receptor kinase [87]), metabolic flexibility in cytoplasm and chloroplasts (2-oxoglutarate oxygenase [88]; glutamine amidotransferase [89]), and ER membrane function (DUF962 [90]). Therefore, the growth-associated SNPs detected in *P. mexicana* could primarily affect regulatory and metabolic pathways that modulate growth indirectly through plastic responses and resource allocation (plasticity-driven strategy) rather than a growth-defense trade-off strategy. Notably, both a SNP in the receptor kinase and 2-oxoglutarate (2OG) and in the Fe(II)-dependent oxygenase superfamily protein gene were also verified by the MLM approach, supporting the biological relevance of these associations.

From an adaptive perspective, growth-associated SNPs related to defense, hormonal signaling, epigenetic regulation, transcriptional control and energy metabolism likely reflect selection on resource-allocation strategies rather than growth per se. In long-lived conifers, growth variation results from how carbon, nitrogen and energy are allocated between growth, stress defense and maintenance in a process tightly regulated by hormonal networks and epigenetic mechanisms [28]. Such allocation-mediated control of growth is key to local adaptation in heterogeneous and stressful environments [91].

Comparison of the findings of this study and the large-scale GWAS conducted in *Picea abies* [15] highlights the contrasting dynamics of continental versus relict population genomics. While the high genetic heterogeneity across the European range of *P. abies* typically dilutes effect sizes, the unique demographic history of relict species may facilitate the detection of adaptive signals. Genetic drift and strong directional selection in marginal habitats could concentrate adaptive alleles, making them detectable even with moderate sample sizes [3]. Specifically, the identification of key regulatory genes, such as the AP2-like transcription factor ANT (Table 2), suggests genuine genomic mechanisms of growth regulation. These signals likely reflect the long-term demographic

purging of deleterious alleles following historical population collapses, dated to ~ 30 ka for *P. martinezii* and ~ 800 ka for *P. mexicana* [22]. The finding that three of four MLM-verified SNPs lack annotation was expected, given the immense complexity and incomplete characterization of the 20-Gb conifer genome [17]. By identifying a high-confidence genomic signal linked to key growth regulators, this study provides a robust foundation for understanding the genomic mechanisms of early development in these endangered spruces.

Although separated by oceanic barriers for 20–25 million years (late Oligocene/early Miocene), the great similarity of 80.9–99.6% (93% on average, Table 2) between genomic sequences containing growth-related outlier SNPs from the two Mexican spruces and *P. abies* could arise from evolutionarily conserved genes, parallel evolution and shared selective pressure [92, 16]. This aligns with patterns in other widespread conifer genera (*Pinus*, *Abies*), in which stress-responsive genes exhibit greater conservation than lineage-specific traits like reproductive timing and terpene profiles [28].

Beyond these direct functional interpretations, several alternative explanations should be considered for the observed associations. First, linkage disequilibrium may play a role, where the observed SNPs could be linked to neighboring genes that influence seedling growth, rather than being the causal variants themselves [91]. Furthermore, the observed signals may reflect pleiotropic effects, where some loci influence multiple traits through shared regulatory networks, a known constraint in the evolution of adaptive phenotypes [93]. The influence of epigenetic mechanisms on gene expression related to growth traits should also not be overlooked [94]. Finally, recurrent mutation could also explain the maintenance of certain variants, whereby specific loci or genomic regions may have a higher probability of change due to their inherent sequence characteristics [95].

While the pre-selection of outlier SNPs can introduce statistical bias [32], this risk was mitigated through a rigorous validation framework. The implementation of a conservative Mixed Linear Model (MLM) alongside the block-wise conditional variance partitioning for multivariate association modeling enabled disentanglement of genetic signals from maternal effects (SW) and kinship structures. The stringency of this approach is evidenced by the fact that only four core SNPs originally identified through outlier pre-selection maintained nominal significance ($p < 0.05$) under the MLM framework (Table 2).

The robustness of these multivariate associations was further confirmed by CI99.9% of the pseudo-RF coefficient of determination, which remained significantly higher than zero, ensuring the precision of the genomic predictors. Factors such as the highly polygenic architecture of growth [96], limited GBS coverage and relatively small sample sizes [29] predispose the analysis to the Beavis effect [97] and inherently constrain the detection of small-effect loci. However, the consistency across non-parametric, linear and multivariate models confirmed the validity of the findings. Notably, the absence of any detectable heterozygote advantage in these candidate SNPs suggests that the observed growth variations are likely driven by additive genetic effects, in a pattern consistent with the previously described demographic purging in these fragmented populations [22]. Consequently, the convergence between statistical validation, variance partitioning and functional plausibility indicates that these loci represent robust candidates for local adaptation in Mexican spruces, rather than mere statistical artifacts.

Conclusion

This study provides the first analysis of associations between growth traits in the endangered endemic spruces *P. martinezii* and *P. mexicana* using genotyping-by-sequencing, extending dendrogenomic approaches to relict conifer populations. Despite small sample sizes, of 77 and 74 families, we identified 32 and 10 outlier SNPs yielding respectively 17 and six annotated candidate genes with contrasting multivariate associations (Table 2). This finding suggests species-specific growth regulation: defense-oriented in *P. martinezii* and plasticity-driven in *P. mexicana*. Notably, several of the candidate detected genes encode key components of growth hormone pathways (including gibberellin-regulated protein, cytokinin hydroxylase and the AP2-like transcription factor ANT). This finding provides direct mechanistic links between genetic variation and hormonal control of cell proliferation and organ size. While no single approach establishes causality, the convergence of population-genomic signals, parent-genotype–dependent effects, machine-learning patterns and functional annotation increases confidence that the identified SNPs and genes represent biologically meaningful candidates influencing growth.

Integrating these findings with the contrasting demographic histories of the two spruce species reveals how population bottlenecks enhance the detection of growth-associated variants by reducing background genetic variation. These genomic resources provide information for prioritizing conservation measures and assisted gene flow to maintain adaptive potential under climate change conditions and also for designing future breeding programs. With 80.9–99.6% sequence identity to conserved *P. abies* homologs, the applicability of these genomic markers may extend across the genus, facilitating comparative conservation genomics approaches in other threatened spruce species.

Abbreviations

D
seedling basal diameter
H
seedling height
SW
seed weight
SNP
single nucleotide polymorphism
GBS
genotyping-by-sequencing
GWAS

genome-wide association study
PCA
principal component analysis
MLM
mixed linear model
RF
Random Forest
BLAST
Basic Local Alignment Search Tool
LFMM
latent factor mixed model
LOF
local outlier factor
PC1
first principal component.

Declarations

Ethics approval and consent to participate

Seed collection was conducted under permits issued in accordance with Mexican environmental regulations (NOM-059-SEMARNAT-2010). The Nagoya Protocol on access and benefit-sharing was observed. No human subjects or animals were involved in this study.

Consent for publication:

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was funded by CONACYT-CONAFOR (Project 2017-4-292615) and SECIHTI (scholarship 776540 to C.A.S.-S.).

Author Contribution

C.A.S.-S. conceived the study, performed the analyses, and wrote the manuscript. J.H.-V. contributed to data analysis and manuscript revision. V.C. contributed to genomic data analysis. E.M.-M. conducted seed collection and common garden experiments. J.V.-D. contributed to dendrochronological analysis. P.L.-S. contributed to spatial analysis. C.W. supervised the study, provided resources, and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgement

We thank the communities and ejidos that granted access to the natural populations of both spruce species.

Data Availability

The datasets generated during the current study are available in the NCBI BioProject under accession PRJNA1146750 repository (http://www.ncbi.nlm.nih.gov/bioproject/1146750). The Stacks catalogue files (FASTA format, consensus loci) generated by the assembly pipeline applied in this study are deposited at Zenodo (https://doi.org/10.5281/zenodo.20057775). The Picea abies reference genome used for putative functional annotation by homology (BLAST) is publicly accessible through PlantGenIE (https://plantgenie.org/). All other datasets supporting the conclusions of this article are included within the article and its Supplementary Information files.

References

1. Taiz L, Møller IM, Murphy A, et al. Plant Physiology and Development. 7th ed. Oxford, UK: Oxford University Press; 2023.
2. Knighton J, Sanchez-Martinez P, Anderegg L. A global dataset of tree hydraulic and structural traits imputed from phylogenetic relationships. Sci Data. 2024;11:1336. https://doi.org/10.1038/s41597-024-04254-4.

3. Krutovsky KV. Dendrogenomics is a new interdisciplinary field of research of the adaptive genetic potential of forest tree populations integrating dendrochronology, dendroecology, dendroclimatology, and genomics. Russ J Genet. 2022;58:1273–86. https://doi.org/10.1134/s1022795422110059.
4. Heer K, Behringer D, Piermattei A, et al. Linking dendroecology and association genetics in natural populations: stress responses archived in tree rings associate with SNP genotypes in silver fir (Abies alba Mill). Mol Ecol. 2018;27:1428–38. https://doi.org/10.1111/mec.14538.
5. Jurcic EJ, Villalba PV, Dutour J, et al. Breeding value predictive accuracy for scarcely recorded traits in a Eucalyptus grandis breeding population using genomic selection and data on predictor traits. Tree Genet Genomes. 2023;19:31. https://doi.org/10.1007/s11295-023-01611-z.
6. De La Torre AR, Sekhwal MK, Puiu D, et al. Genome-wide association identifies candidate genes for drought tolerance in coast redwood and giant sequoia. Plant J. 2022;109:7–22. https://doi.org/10.1111/tpj.15547.
7. Chybicki IJ, Suszka J, Meyza K, et al. Heritability of early seedling growth in Taxus baccata, a declining conifer. Sci Rep. 2025;15:35214. https://doi.org/10.1038/s41598-025-21939-x.
8. Feng S, Xi E, Wan W, et al. Genomic signals of local adaptation in Picea crassifolia. BMC Plant Biol. 2023;23:534. https://doi.org/10.1186/s12870-023-04539-7.
9. Wang Y, Jiang Z, Qin A, et al. Population structure, genetic diversity and candidate genes for the adaptation to environmental stress in Picea koraiensis. Plants. 2023;12:1266. https://doi.org/10.3390/plants12061266.
10. Fasanella M, Suarez ML, Hasbún R, et al. Individual-based dendrogenomic analysis of forest dieback driven by extreme droughts. Can J Res. 2021;51:420–32. https://doi.org/10.1139/cjfr-2020-0221.
11. Méndez-Cea B, García-García I, Sánchez-Salguero R, et al. Tree-level growth patterns and genetic associations depict drought legacies in the relict forests of Abies marocana. Plants. 2023;12:873. https://doi.org/10.3390/plants12040873.
12. Novikova SV, Oreshkova NV, Sharov VV, et al. Study of the genetic adaptation mechanisms of Siberian larch (Larix sibirica Ledeb.) regarding climatic stresses based on dendrogenomic analysis. Forests. 2023;14:2358. https://doi.org/10.3390/f14122358.
13. Novikova SV, Oreshkova NV, Sharov VV, et al. Study of the genetic mechanisms of Siberian stone pine (Pinus sibirica Du Tour) adaptation to the climatic and pest outbreak stresses using dendrogenomic approach. Int J Mol Sci. 2024;25:11767. https://doi.org/10.3390/ijms252111767.
14. Segura-Sanchez CA, Hernández-Velasco J, Villanueva-Díaz J, et al. Dendrogenomic resilience mechanisms of two endangered Mexican spruces. Plant Divers. 2026;48:151–9. https://doi.org/10.1016/j.pld.2025.05.011.
15. Chen ZQ, Zan Y, Milesi P, et al. Leveraging breeding programs and genomic data in Norway spruce (Picea abies L. Karst) for GWAS analysis. Genome Biol. 2021;22:179. https://doi.org/10.1186/s13059-021-02392-1.
16. Lockwood JD, Aleksić JM, Zou J, et al. A new phylogeny for the genus Picea from plastid, mitochondrial, and nuclear sequences. Mol Phylogenet Evol. 2013;69:717–27. https://doi.org/10.1016/j.ympev.2013.03.017.
17. Nystedt B, Street NR, Wetterbom A, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013;497:579–84. https://doi.org/10.1038/nature12211.
18. Lo T, Coombe L, Gagalova KK et al. Assembly and annotation of the black spruce genome provide insights on spruce phylogeny and evolution of stress response. G3 (Bethesda). 2024;14:jkad247. https://doi.org/10.1093/g3journal/jkad247
19. Mendoza-Maya E, Gómez-Pineda E, Sáenz-Romero C, et al. Assisted migration and the rare endemic plant species: the case of two endangered Mexican spruces. PeerJ. 2022;10:e13812. https://doi.org/10.7717/peerj.13812.
20. NOM-059-SEMARNAT-2010. Protección ambiental–Especies nativas de México de flora y fauna silvestres–Categorías de riesgo y especificaciones para su inclusión, exclusión o cambio–Lista de especies en riesgo. México: Diario Oficial de la Federación, Secretaría de Medio Ambiente y Recursos Naturales; 2019.
21. Mendoza-Maya E, Giles-Pérez GI, Vargas-Hernández JJ, et al. Evolutionary drivers of reproductive fitness in two endangered forest trees. New Phytol. 2024;244:1086–100. https://doi.org/10.1111/nph.20073.
22. Castro J. Seed mass versus seedling performance in Scots pine: a maternally dependent trait. New Phytol. 1999;144:153–61. https://doi.org/10.1046/j.1469-8137.1999.00495.x.
23. Wehenkel C, Torres-Valverde JM, Hernández-Díaz JC, et al. Adaptive trait variation in seedlings of rare endemic Mexican spruce provenances under nursery conditions. Forests. 2023;14:790. https://doi.org/10.3390/f14040790.
24. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
25. Abed A, Légaré G, Pomerleau S, et al. Genotyping-by-sequencing on the ion torrent platform in barley. Methods Mol Biol. 2019;1900:233–52. https://doi.org/10.1007/978-1-4939-8944-7_15.
26. Krueger F, James F, Ewels P et al. FelixKrueger/TrimGalore: v0.6.7. Zenodo. 2021. https://doi.org/10.5281/zenodo.5127899
27. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. Bioinformatics. 2011;27:2156–8. https://doi.org/10.1093/bioinformatics/btr330.
28. Neale DB, Kremer A. Forest tree genomics: growing resources and applications. Nat Rev Genet. 2011;12:111–22. https://doi.org/10.1038/nrg2931.
29. Wang M, Xu S. Statistical power in genome-wide association studies and quantitative trait locus mapping. Heredity. 2019;123:287–306. https://doi.org/10.1038/s41437-019-0205-3.
30. Isik F, Holland J, Maltecca C. Genetic Data Analysis for Plant and Animal Breeding. Cham, Switzerland: Springer; 2017.
31. Harrell FE Jr. Regression Modeling Strategies. 2nd ed. Cham, Switzerland: Springer; 2015. https://doi.org/10.1007/978-3-319-19425-7.

32. Forester BR, Lasky JR, Wagner HH, et al. Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. Mol Ecol. 2018;27:2215–33. https://doi.org/10.1111/mec.14584.
33. Wang J, Zhang Z. GAPIT version 3: Boosting power and accuracy for genomic association and prediction. Genom Proteom Bioinform. 2021;19:629–40. https://doi.org/10.1016/j.gpb.2021.08.005.
34. Luu K, Bazin E, Blum MGB. pcadapt: an R package to perform genome scans for selection based on principal component analysis. Mol Ecol Resour. 2017;17:67–77. https://doi.org/10.1111/1755-0998.12592.
35. Privé F, Luu K, Vilhjálmsson BJ, et al. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. Mol Biol Evol. 2020;37:2153–4. https://doi.org/10.1093/molbev/msaa053.
36. Caye K, Jumentier B, Lepeule J, et al. LFMM 2: Fast and accurate inference of gene-environment associations in genome-wide studies. Mol Biol Evol. 2019;36:852–60. https://doi.org/10.1093/molbev/msz008.
37. Gain C, François O. LEA 3: Factor models in population genetics and ecological genomics with R. Mol Ecol Resour. 2021;21:2738–48. https://doi.org/10.1111/1755-0998.13366.
38. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria 2022. https://www.R-project.org/
39. Breunig MM, Kriegel HP, Ng RT, et al. LOF: Identifying density-based local outliers. SIGMOD Rec. 2000;29:93–104. https://doi.org/10.1145/335191.335388.
40. Buitinck L, Louppe G, Blondel M et al. API design for machine learning software: experiences from the scikit-learn project. arXiv:1309.0238. 2013.
41. Li WQ, Zhang MJ, Gan PF, et al. CLD1/SRL1 modulates leaf rolling by affecting cell wall formation, epidermis integrity and water homeostasis in rice. Plant J. 2017;92:904–23. https://doi.org/10.1111/tpj.13728.
42. Cattell RB. The scree test for the number of factors. Multivar Behav Res. 1966;1:245–76. https://doi.org/10.1207/s15327906mbr0102_10.
43. Xu H, Zhang L, Li P, et al. Outlier detection algorithm based on k-nearest neighbors-local outlier factor. J Algorithms Comput Technol. 2022;16:174830262210781. https://doi.org/10.1177/17483026221078111.
44. Liu X, Liu X, Zhang R, et al. Securely computing the Manhattan distance under the malicious model and its applications. Appl Sci. 2022;12:11705. https://doi.org/10.3390/app122211705.
45. Whitlock MC, Lotterhos KE. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. Am Nat. 2015;186(Suppl 1):S24–36. https://doi.org/10.1086/682949.
46. Lotterhos KE, Whitlock MC. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. Mol Ecol. 2015;24:1031–46. https://doi.org/10.1111/mec.13100.
47. Grattapaglia D, Resende MDV. Genomic selection in forest tree breeding. Tree Genet Genomes. 2011;7:241–55. https://doi.org/10.1007/s11295-010-0328-4.
48. Whitacre JM. Biological robustness: paradigms, mechanisms, and systems principles. Front Genet. 2012;3:67. https://doi.org/10.3389/fgene.2012.00067.
49. Smeeth D, Beck S, Karam EG, et al. The role of epigenetics in psychological resilience. Lancet Psychiatry. 2021;8:620–9. https://doi.org/10.1016/S2215-0366(20)30515-0.
50. Flotildes MJ, Garcia G, Piol AM, et al. Lived experiences and resilience of hospital pharmacists during the COVID-19 pandemic. Explor Res Clin Soc Pharm. 2023;11:100299. https://doi.org/10.1016/j.rcsop.2023.100299.
51. Thorogood R, Mustonen V, Aleixo A, et al. Understanding and applying biological resilience, from genes to ecosystems. NPJ Biodivers. 2023;2:16. https://doi.org/10.1038/s44185-023-00022-6.
52. Breiman L. Random forests. Mach Learn. 2001;45:5–32. https://doi.org/10.1023/a:1010933404324.
53. Liaw A, Wiener M. Classification and regression by randomForest. R J. 2007;2:18–22.
54. Kuhn M. Building predictive models in R using the caret package. J Stat Softw. 2008;28:1–26. https://doi.org/10.18637/jss.v028.i05.
55. Strobl C, Boulesteix AL, Kneib T, et al. Conditional variable importance for random forests. BMC Bioinformatics. 2008;9:307. https://doi.org/10.1186/1471-2105-9-307.
56. Altschul SF, Gish W, Miller W, et al. Basic local alignment search tool. J Mol Biol. 1990;215:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.
57. Sundell D, Mannapperuma C, Netotea S, et al. The plant genome integrative explorer resource: PlantGenIE.org. New Phytol. 2015;208:1149–56. https://doi.org/10.1111/nph.13557.
58. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30:3276–8. https://doi.org/10.1093/bioinformatics/btu531.
59. Ogle DH, Wheeler P, Dinno A. FSA: fisheries stock analysis. R package version 0.8.17. 2017.
60. Dinno A, Dinno MA. Package 'dunn.test'. CRAN Repos. 2017;10:1–7.
61. Visscher PM, Wray NR, Zhang Q, et al. 10 years of GWAS discovery: biology, function, and translation. Am J Hum Genet. 2017;101:5–22. https://doi.org/10.1016/j.ajhg.2017.06.005.
62. Alberto FJ, Aitken SN, Alía R, et al. Potential for evolutionary responses to climate change – evidence from tree populations. Glob Change Biol. 2013;19:1645–61. https://doi.org/10.1111/gcb.12181.

63. Vieira LD, Silva-Junior OB, Novaes E, et al. Comparative population genomics in Tabebuia alliance shows evidence of adaptation in Neotropical tree species. Heredity. 2022;128:141–53. https://doi.org/10.1038/s41437-021-00491-0.
64. Valmonte GR, Arthur K, Higgins CM, et al. Calcium-dependent protein kinases in plants: evolution, expression and function. Plant Cell Physiol. 2014;55:551–69. https://doi.org/10.1093/pcp/pct200.
65. Dubey N, Singh K. Role of NBS-LRR proteins in plant defense. Molecular Aspects of Plant-Pathogen Interaction. Singapore: Springer; 2018. pp. 115–38.
66. Zhu M, Feng M, Tao X. NLR-mediated antiviral immunity in plants. J Integr Plant Biol. 2025;67:786–800. https://doi.org/10.1111/jipb.13821.
67. Wang H, Song S, Gao S, et al. The NLR immune receptor ADR1 and lipase-like proteins EDS1 and PAD4 mediate stomatal immunity in Nicotiana benthamiana and Arabidopsis. Plant Cell. 2024;36:427–46. https://doi.org/10.1093/plcell/koad270.
68. Ying S, Scheible WR, Lundquist PK. A stress-inducible protein regulates drought tolerance and flowering time in Brachypodium and Arabidopsis. Plant Physiol. 2023;191:643–59. https://doi.org/10.1093/plphys/kiac486.
69. Beihammer G, Maresch D, Altmann F, et al. Glycosylphosphatidylinositol-anchor synthesis in plants: A glycobiology perspective. Front Plant Sci. 2020;11:611188. https://doi.org/10.3389/fpls.2020.611188.
70. Hamberger B, Bohlmann J. Cytochrome P450 mono-oxygenases in conifer genomes: discovery of members of the terpenoid oxygenase superfamily in spruce and pine. Biochem Soc Trans. 2006;34:1209–14. https://doi.org/10.1042/BST0341209.
71. Herzog M, Dorne AM, Grellet F. GASA, a gibberellin-regulated gene family from Arabidopsis thaliana related to the tomato GAST1 gene. Plant Mol Biol. 1995;27:743–52. https://doi.org/10.1007/bf00020227.
72. Kiba T, Takei K, Kojima M, et al. Side-chain modification of cytokinins controls shoot growth in Arabidopsis. Dev Cell. 2013;27:452–61. https://doi.org/10.1016/j.devcel.2013.10.004.
73. Krizek BA, Bantle AT, Heflin JM, et al. AINTEGUMENTA and AINTEGUMENTA-LIKE6 directly regulate floral homeotic, growth, and vascular development genes in young Arabidopsis flowers. J Exp Bot. 2021;72:5478–93. https://doi.org/10.1093/jxb/erab223.
74. Kulich I, Vogler F, Bleckmann A, et al. ARMADILLO REPEAT ONLY proteins confine Rho GTPase signalling to polar growth sites. Nat Plants. 2020;6:1275–88. https://doi.org/10.1038/s41477-020-00781-1.
75. Shang JY, Cai XW, Su YN, et al. Arabidopsis Trithorax histone methyltransferases are redundant in regulating development and DNA methylation. J Integr Plant Biol. 2022;64:2438–54. https://doi.org/10.1111/jipb.13406.
76. Dahan J, Tcherkez G, Macherel D, et al. Disruption of the CYTOCHROME C OXIDASE DEFICIENT1 gene leads to cytochrome c oxidase depletion and reorchestrated respiratory metabolism in Arabidopsis. Plant Physiol. 2014;166:1788–802. https://doi.org/10.1104/pp.114.248526.
77. Duan Z, Li K, Zhang L, et al. F-type ATP synthase assembly factors Atp11 and Atp12 in Arabidopsis. Front Plant Sci. 2020;11:522753. https://doi.org/10.3389/fpls.2020.522753.
78. Fortunato S, Nigro D, Lasorella C, et al. The role of glutamine synthetase (GS) and glutamate synthase (GOGAT) in the improvement of nitrogen use efficiency in cereals. Biomolecules. 2023;13:1771. https://doi.org/10.3390/biom13121771.
79. Yang XY, Chen ZW, Xu T, et al. Arabidopsis kinesin KP1 specifically interacts with VDAC3, a mitochondrial protein, and regulates respiration during seed germination at low temperature. Plant Cell. 2011;23:1093–106. https://doi.org/10.1105/tpc.110.082420.
80. Senatore E, Chiuso F, Rinaldi L, et al. The TBC1D31/praja2 complex controls primary ciliogenesis through PKA-directed OFD1 ubiquitylation. EMBO J. 2021;40:e106503. https://doi.org/10.15252/embj.2020106503.
81. Gaillard JL, Berche P, Frehel C, et al. Entry of L. monocytogenes into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci. Cell. 1991;65:1127–41. https://doi.org/10.1016/0092-8674(91)90009-n.
82. Shiu SH, Bleecker AB, Sci STKE. 2001;2001:re22. https://doi.org/10.1126/stke.2001.113.re22
83. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. Curr Opin Struct Biol. 2001;11:725–32. https://doi.org/10.1016/s0959-440x(01)00266-4.
84. Dolan J, Walshe K, Alsbury S, et al. The extracellular leucine-rich repeat superfamily; a comparative survey and analysis of evolutionary relationships and expression patterns. BMC Genomics. 2007;8:320. https://doi.org/10.1186/1471-2164-8-320.
85. Hernández-García J, Serrano-Mislata A, Lozano-Quiles M, et al. DELLA proteins recruit the Mediator complex subunit MED15 to coactivate transcription in land plants. Proc Natl Acad Sci USA. 2024;121:e2319163121. https://doi.org/10.1073/pnas.2319163121.
86. Fernández-Jiménez N, Martinez-Garcia M, Varas J, et al. The scaffold nucleoporins SAR1 and SAR3 are essential for proper meiotic progression in Arabidopsis thaliana. Front Cell Dev Biol. 2023;11:1285695. https://doi.org/10.3389/fcell.2023.1285695.
87. Goff KE, Ramonell KM. The role and regulation of receptor-like kinases in plant defense. Gene Regul Syst Biol. 2007;1:167–75. https://doi.org/10.1177/117762500700100015.
88. Sun F, Chen Q, Chen Q, et al. Genome-wide identification and expression analysis of the 2OG-Fe(II) oxygenase gene family in upland cotton (Gossypium hirsutum L). Physiol Mol Biol Plants. 2021;27:1969–77. https://doi.org/10.1007/s12298-021-01065-4.
89. Kambhampati S, Pajak A, Marsolais F. Evidence that class I glutamine amidotransferase, GAT1_2.1, acts as a glutaminase in roots of Arabidopsis thaliana. Plant Sci. 2021;311:111033. https://doi.org/10.1016/j.plantsci.2021.111033.
90. Ushio M, Ishikawa T, Matsuura T, et al. MHP1 and MHL generate odd-chain fatty acids from 2-hydroxy fatty acids in sphingolipids and are related to immunity in Arabidopsis thaliana. Plant Sci. 2023;336:111840. https://doi.org/10.1016/j.plantsci.2023.111840.
91. Grattapaglia D, Silva-Junior OB, Resende RT, et al. Quantitative genetics and genomics converge to accelerate forest tree breeding. Front Plant Sci. 2018;9:1693. https://doi.org/10.3389/fpls.2018.01693.

92. Ran JH, Wei XX, Wang XQ. Molecular phylogeny and biogeography of Picea (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. Mol Phylogenet Evol. 2006;41:405–19. https://doi.org/10.1016/j.ympev.2006.05.039.
93. Bertolini F et al. [Reference pending verification]. 2025.
94. Yakovlev I, Fossdal CG, Skrøppa T, et al. An adaptive epigenetic memory in conifers with important implications for seed production. Seed Sci Res. 2012;22:63–76. https://doi.org/10.1017/s0960258511000535.
95. Gillespie JH. Population Genetics: A Concise Guide. Baltimore, MD: JHU; 2004.
96. Hall D, Hallingbäck HR, Wu HX. Estimation of number and size of QTL effects in forest tree traits. Tree Genet Genomes. 2016;12:107. https://doi.org/10.1007/s11295-016-1073-0.
97. Xu S. Theoretical basis of the Beavis effect. Genetics. 2003;165:2259–68. https://doi.org/10.1093/genetics/165.4.2259.

Figures

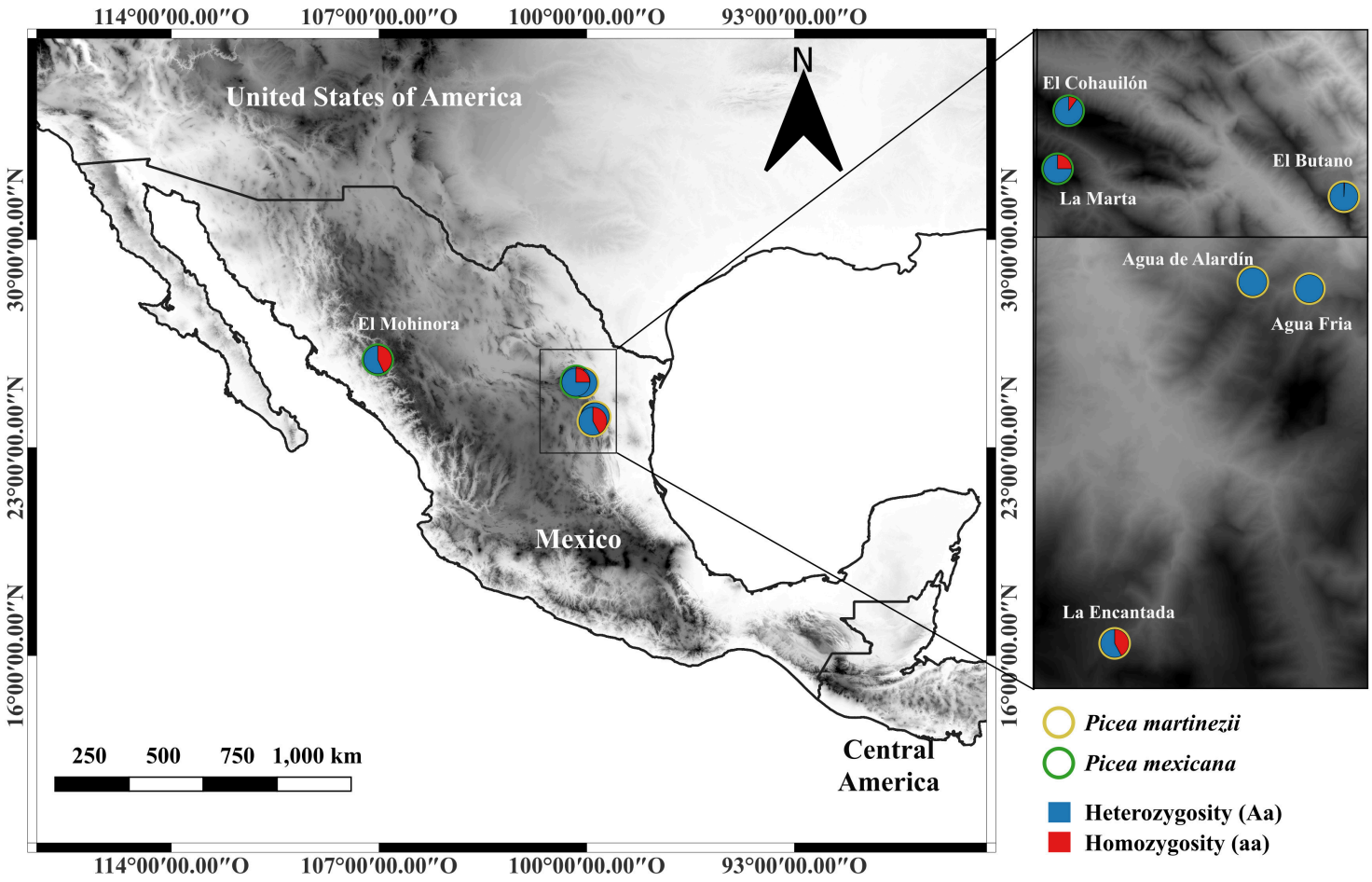

Figure 1

Location of *P. martinezii* populations (yellow circles) in the Sierra Madre Oriental (Nuevo León) and of *P. mexicana* populations (green circles) in the Sierra Madre Oriental (Coahuila) and Sierra Madre Occidental (Chihuahua). The pie charts indicate the proportion of observed heterozygosity (yellow) and homozygosity (green) at growth-associated SNP loci. The inset shows a detailed view of the eastern populations.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupportingInformationBMC.docx