

RESEARCH

Open Access



Spatially heterogeneous selection and inter-varietal differentiation maintain population structure and local adaptation in a widespread conifer

Pablo Peláez¹, Gustavo P. Lorenzana¹, Kailey Baesen¹, Jose Ruben Montes² and Amanda R. De La Torre^{1*}

Abstract

Background Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) plays a critical role in the ecology and economy of Western North America. This conifer species comprises two distinct varieties: the coastal variety (var. *menziesii*) along the Pacific coast, and the interior variety (var. *glauca*) spanning the Rocky Mountains into Mexico, with instances of inter-varietal hybridization in Washington and British Columbia. Recent investigations have focused on assessing environmental pressures shaping Douglas-fir's genomic variation for a better understanding of its evolutionary and adaptive responses. Here, we characterize range-wide population structure, estimate inter-varietal hybridization levels, identify candidate loci for climate adaptation, and forecast shifts in species and variety distribution under future climates.

Results Using a custom SNP-array, we genotyped 540 trees revealing four distinct clusters with asymmetric admixture patterns in the hybridization zone. Higher genetic diversity observed in coastal and hybrid populations contrasts with lower diversity in inland populations of the southern Rockies and Mexico, exhibiting a significant isolation by distance pattern, with less marked but still significant isolation by environment. For both varieties, we identified candidate loci associated with local adaptation, with hundreds of genes linked to processes such as stimulus response, reactions to chemical compounds, and metabolic functions. Ecological niche modeling revealed contrasting potential distribution shifts among the varieties in the coming decades, with interior populations projected to lose habitat and become more vulnerable, while coastal populations are expected to gain suitable areas.

Conclusions Overall, our findings provide crucial insights into the population structure and adaptive potential of Douglas-fir, with the coastal variety being the most likely to preserve its evolutionary path throughout the present century, which carry implications for the conservation and management of this species across their range.

Keywords Douglas-fir, Hybridization, Single nucleotide polymorphisms, Population structure, Genetic diversity, Local adaptation

*Correspondence:

Amanda R. De La Torre
Amanda.de-la-torre@nau.edu

¹School of Forestry, Northern Arizona University, Flagstaff, AZ, USA

²Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, México



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Intraspecific gene flow can generate novel genetic pools and promote adaptive evolution in differentiated populations. Intraspecific variation and hybridization in natural populations with dissimilar evolutionary histories are widely acknowledged as crucial factors for different patterns of local adaptation, and understanding these evolutionary processes has become more relevant due to climate change [1]. Forest trees have been proposed to be susceptible to rapid climate change due to their long generation times. However, trees also present high levels of standing genetic diversity due to large population sizes, extensive levels of gene flow and wide distributions across different environments, which makes them ideal systems to detect adaptive signals and to study the genetic basis of adaptation [2]. In the past, efforts to dissect the genetic basis of adaptation relied on population genetic analyses that focused mainly on the genetic differentiation of the populations without considering environmental heterogeneity among them [3]. Recent approaches in landscape genomics integrate multiple genetic, spatial, and environmental data to shed light on the genetic variants underlying adaptation to different climate scenarios.

Douglas-fir constitutes one of the world's most important timber trees. It grows under a wide range of climatic conditions and has been part of the landscape of western North America since the Pleistocene. With warming environmental conditions, research on Douglas-fir has focused on testing the ability of the species to grow while withstanding changes in temperature (heat stress) and low soil water content (drought stress) [4–8]. Two varieties with adaptive variation are formally recognized: the coastal variety along the Pacific coast (*P. menziesii* var. *menziesii*), and the interior variety across the Rocky Mountains (*P. menziesii* var. *glauca*). Populations of Mexican Douglas-fir are small, extremely fragmented, and non-continuous in distribution, except in a larger area in the state of Chihuahua [9]. Compared to populations from northern Mexico, populations from central Mexico are genetically, morphologically, and phenologically more distinct from populations from the United States [10–13].

Fossil records exist for Douglas-fir from the early Miocene to the late Holocene [13, 14]. Fossils from the Miocene and Pliocene were located along the west coast (from British Columbia to California) and in the Columbia Plateau and Great Basin. Pleistocene fossils were detected in the Rocky Mountains and in the west coast; however, no fossil records were found in the Columbia Plateau and the western Great Basin [13, 14]. This change in the distribution of fossil records correlates with the Pliocene orogeny of the Sierra Nevada and Cascade Mountain ranges, which is thought to be the cause of the vicariant separation of the species into interior and

coastal varieties [13, 14]. In addition, fossil records suggest that subsequent differentiation within varieties may have occurred due to climate change refugia during the Pleistocene glaciation [13, 15]. Posterior events of secondary contact in British Columbia and the Washington Cascades led to the formation of two inter-varietal hybrid zones in these geographic locations.

Demographic, phylogenetic and population molecular studies have shown differentiation between the varieties and populations of Douglas-fir [13, 16–19]. Genetic differentiation between the two varieties was reported using allozymes more than three decades ago [17]. Chloroplast and mitochondrial markers revealed Pleistocene divergence of the Mexican populations, which resulted in the classification of the interior variety into two different lineages: the Rocky Mountain and Mexican lineages [13, 16]. Population structure analyses using nuclear microsatellite markers identified refugial populations with intra-varietal and inter-varietal gene flow caused by hybridization and introgression [15, 19].

Studies that assess the relation between genetic diversity and environmental responses in Douglas-fir have been scarce. Recently, based on genome-wide sequencing approaches, genes associated with local adaptation have been identified [7, 20–23]. A comparative transcriptome analysis of Douglas-fir trees from two provenances, from a coastal and interior habitat, with contrasting natural environments was carried out to evaluate the effect of abiotic environmental factors on gene expression responses [21]. Targeted sequence capture and mixed effect models were used to detect high differentiation of drought tolerance genes between interior and coastal trees grown in experimental conditions [22]. In coastal populations, single nucleotide polymorphisms have been associated with cold-hardiness and phenology related traits [7, 20]. In this study, we genotyped trees of the two varieties to characterize the population structure across the natural, species-range distribution of Douglas-fir in North America; identify the evolutionary processes maintaining genetic variation and population structure; assess the levels of inter-varietal hybridization; identify candidate loci for climate adaptation; and predict the impacts of climate change in the future distribution of the species.

Materials and methods

Plant material and DNA extraction

Seeds and needle tissue were collected from 577 open-pollinated trees across the Douglas-fir natural distribution range, from Mexico to British Columbia, Canada (Table 1). Trees were collected at a minimum distance of 20 m to avoid relatedness among individuals. Prior to extraction, seeds were soaked in a solution of 70% water and 30% of 3% hydrogen peroxide for 12 h. Ten haploid

Table 1 Geographic location and mean environmental variables of individuals included in this study

Group	Latitude	Longitude	MAT (°C)	MAP (mm)	Sample size
Coastal	45.8	-123.0	8.93	1861.2	325
Hybrid	51.4	-120.2	3.11	752.2	27
Interior north	47.9	-111.87	4.02	793.5	61
Interior south	35.5	-108.3	6.88	584.1	127

MAT, mean annual temperature; MAP, mean annual precipitation

megagametophytes for each family were pooled to infer the maternal genotype. DNA from megagametophytes was extracted using the Qiagen DNeasy mini-prep Plant kit and an Eppendorf automated pipetting workstation. When needle tissue was available, DNA was extracted using a modified CTAB method [24] or the MPBio (MPBiomedicals LLC, Ohio, USA) RapidPure DNA Plant kit. The CTAB method was modified to include a wash step of homogenized tissue with CTAB buffer prior to 65 °C incubation. DNA quality and concentration was evaluated with a NanoDrop Spectrophotometer, a Qubit 4.0 Fluorometer and agarose gels using Invitrogen's E-Gel Power Snap system.

SNP genotyping and filtering

Samples were genotyped using a custom-designed gene-based Illumina Infinium SNP array containing 16,146 Douglas-fir single nucleotide polymorphisms (SNPs) at the University of California-Davis Genome Center. This array was designed to represent genome-wide variation in the species by using 10X whole-genome re-sequencing data from individuals across the species' geographic range as input for array construction [7, 8]. Genotypes were called using Illumina's Genome Studio Genotyping Module v 2.0.5 (Illumina, 2016). Filtering criteria included a SNP call frequency > 0.85, individual call rate > 0.85, non-monomorphic and a minor allele frequency > 0.01.

Population structure and hybridization analyses

Principal component analysis (PCA) was carried out to infer population structure using the gdsfmt v.1.34 and SNPRelate v.1.32 packages [25]. Population structure was also inferred by a Discriminant Analysis of Principal Components using the find.clusters function of the adegenet (v.2.1.10) package [26]. The best number of genetic clusters was determined using the Bayesian Information Criterion (BIC). To estimate ancestral differences between varieties and populations, ADMIXTURE software v.1.3.0 [27] was employed using 10 independent runs for k-values ranging from 2 to 10. Admixture proportions obtained for K=2 and the two K-values with the lowest averaged cross-validation error were chosen. CLUMPAK v.1.1 was used to produce bar plots [28]. Pie

chart maps of individual admixture proportions were constructed with QGIS and the option around the point for better visualization. Posterior probabilities of hybrid categories were calculated with NewHybrids v.2.0 [29]. All individuals classified as coastal, hybrid, and interior south from Additional file 1 were used. The individuals from the interior north were not included because they presented higher variation in the admixture proportions (Q values) among the estimated ancestral populations. One hundred randomly selected SNPs were used to estimate posterior probabilities with 50,000 sweeps for 3 replicates.

Genetic diversity analyses

Population inbreeding coefficients were obtained from a kinship matrix with the popkin package v.1.3.23 [30]. Trees delimited in a region with similar topography and climate conditions as well as similar latitude or longitude were considered an independent population. Geographic positions of populations on the map were estimated from the centroid of all individuals in each population. The natural distribution map of the two varieties was based on the Atlas of United States trees [31]. The VCFtools program v.0.1.16 was used to determine the heterozygosity and pairwise fixation index (Fst) values [32]. Heterozygosity values were calculated per individual with the observed number of homozygotes [O(HOM)] and the number of non-missing genotypes [N(NM)] using the formula: $[N(NM) - O(HOM)]/N(NM)$. Violin plots were created to visualize the results using the ggplot2 v. 3.4.2 R package [33].

Isolation by distance and environment analyses

To evaluate the association between genetic and geographic distances, a Mantel test was performed using the gl.ibd function implemented in the dartR R package v. 1.0.2 with 999 permutations [34]. Isolation by environment and association between geographic and environmental distances analyses were performed with vegan (v. 2.6-4) and geosphere (v. 1.5–18) R packages. The Spearman correlation method, 9999 permutations, and SNPs without missing data were used for Mantel tests. Climatic data for the environmental variables (BIO1 and BIO12) were obtained from the WorldClim database (www.worldclim.org) with a resolution of 2.5 arcsec using the raster and sp R packages and the coordinates of the trees collected.

Detection of loci under selection

Identification of SNPs under selection was carried out with BayeScan (v. 2.0) and pcadapt (v. 4.3.3) packages [35, 36]. The parameters used for outliers' detection with BayeScan were the following: 20 pilot runs of 5000 iterations, burn-in length of 50,000, thinning interval size of

100 and prior odds ratio of 100. SNPs were considered outliers based on a false discovery rate q -value threshold of 0.05. Pcadapt was used to detect signatures of selection using two principal components. A q -value threshold of 0.1 was used to choose outliers. SNPs were functionally annotated using KOBAS-i and the annotation file of the Douglas-fir genome v. 1.0 [37]. KOBAS-i was used with default parameters for Gene Ontology (GO) and KEGG pathways enrichment analyses. Climate data of 25 environmental variables were extracted from the ClimateNA application [38]. All climate data are based on annual averages for the years 1962–1990. The variables used in the analysis are related to monthly, seasonal, and annual temperature and precipitation measurements (Additional file 2). To identify SNPs involved in local adaptation to climate, the Bayenv2 program was used [39]. A covariance matrix was obtained from the average of 5 independent covariance matrices generated with different seed numbers and 100,000 iterations. Five independent runs were performed for the populations of Douglas-fir and 25 climate variables with different seed sizes and 500,000 iterations. Bayes factors (BF) were averaged over the independent runs. A BF value of 10 was considered as threshold for outlier loci.

Ecological niche modeling (ENM)

The ecological niche of Douglas-fir was modeled with the Wallace2 interactive web app [40], which incorporates the Maxent algorithm for estimating potential distribution using presence-only data [41, 42]. Additionally, ecological niche shifts were forecasted (i.e., potential distribution gains and losses in response to climate change) throughout the present century (years 2050 and 2070). Douglas-fir's presence data was derived from the 540 geo-referenced trees retained after the SNP filtering procedure. From those, 236 duplicated coordinates were further removed. The geographic database was complemented with Douglas-fir site records ($n=44$) reported by Wei et al. (2011), which allowed to fill some geographic gaps, especially for Mexican populations, increasing the overall presence data across the species range. The resulting 348 non-duplicated locations were then subdivided into three major regions (ancestral populations) according to PCA and admixture results: coastal ($n=250$), interior-north ($n=41$), interior-south ($n=48$), and a hybrid zone ($n=9$). Coastal region included northern California to British Columbia; interior-north encompassed Idaho and Wyoming up north to British Columbia, and interior-south comprised high elevations from Utah and Colorado to southern Mexico. Hybrid zone is located in British Columbia and Washington.

Nineteen bioclimatic variables derived from present-day averaged temperature and precipitation data available at WorldClim [43] were used as raster layers with a

resolution of 2.5 arc minutes (≈ 5 km). Occurrence data was further processed to reduce sampling bias using a spatial thinning technique, where the minimum distance between occurrence locations (i.e., nearest neighbor distance) was 10 km, resulting in a thinned dataset of 263 localities. From those points, the spatial extent for niche model building and evaluation was determined, using a bounding box with a buffer distance of 1 geographic degree and 10,000 sample background points. The environmental space occupied by each of the regional Douglas-fir populations was characterized as an approximation to their Hutchinsonian niche, defined as “the n -dimensional hypervolume where a species can persist and reproduce in a mathematical space defined by non-depletable environmental gradients” [40, 44]. Wallace2 reduces the dimensionality of the predicted niche using three modules: (1) “environmental ordination”, which is basically a Principal Component Analysis, (2) “occurrence density grid”, depicting the portion of the environmental space that is more densely occupied by the species, given the availability of environmental conditions present within the background extent, and (3) “niche overlap”, quantified as an overall index (overlap D) [45], ranging from 0 to 1, where 0 represents a null overlap and 1 is a complete ecological overlap between populations. The environmental ordination module conducts a Principal Component Analysis to maximize the variation contained in the predictor variables, allowing the identification of collinear variables, and retaining a reduced set of highly explicative variables.

A spatial partition of the occurrence data for training and testing the predictive models, was performed using the “checkerboard 1” option, with $k=2$, and aggregation factor 2. Then, niche models were built using the Maxent module. Maxent is a machine learning algorithm that models the potential distribution of a species in response to a set of environmental conditions (e.g., climate), which are constrained to be relatively uniform across the geographic space encompassed by the input locations [41, 46]. Increasingly complex models were obtained using linear, quadratic, hinge, and product feature classes, and their predictive performance was evaluated using a combination of metrics included in Wallace2 such as the Area Under the Curve (AUC), omission rate, Continuous Boyce Index (CBI), and Akaike Information Criterion (AIC). The best models per species and per variety were transferred to geographic space within the range of Douglas-fir in North America, and as well to future climate scenarios over the next 50 years. The threshold we used for deciding predicted presence (potential distribution) was the 10th percentile of training presence, which roughly corresponds to a 10% omission rate. Potential distribution gains and losses were calculated using the

estimated area of present-day ecological niche extent as a baseline.

Results

Data collection and genotyping

Assessments of genetic diversity, population structure and natural hybridization of Douglas-fir were based on the genotypes of 577 trees (37 populations) collected throughout the distribution of the two varieties including the hybrid zone located in British Columbia and the Washington Cascades (Fig. 1; Additional file 1; Additional file 3). After filtering steps (SNP and sample call frequency rates and minor allele frequency), 540 trees and 11,320 SNPs representing 70.1% of the SNPs in the array were kept for further analyses. The total genotype rate was 96%.

Population structure and hybridization

Due to the broad geographic distribution of the two recognized varieties and the presence of contact zones, this study began by determining the population structure of the dataset. Principal component analysis with

the SNPRelate package [25] distinguished individuals from coastal, hybrids, interior-north, and interior-south groups/varieties according to the samples' geographic location (Fig. 2A). Principal component 1, which accounted for 16.7% of the variation in the dataset, separated interior populations from coastal populations and hybrids, suggesting coastal and interior-south are the most genetically differentiated. Genetic variation between coastal and interior-south groups was higher than variation within each of the groups. Individuals from the Mexican populations (TLX and PUE) clustered with interior-south individuals. Principal component 2, which accounted for 4.15% of the variation in the dataset, separated interior-north and hybrids from most interior-south and coastal individuals (Fig. 2A, B). Discriminant analysis of principal components (DAPC) grouped individuals into four optimal clusters (Fig. 2C). Individuals from interior-north and hybrids were clustered together, while interior-south individuals were separated into two clusters that distinguished trees from Arizona (ARD, SPR, OAC, CAN), New Mexico (SVC, SAM, RIT), Puebla

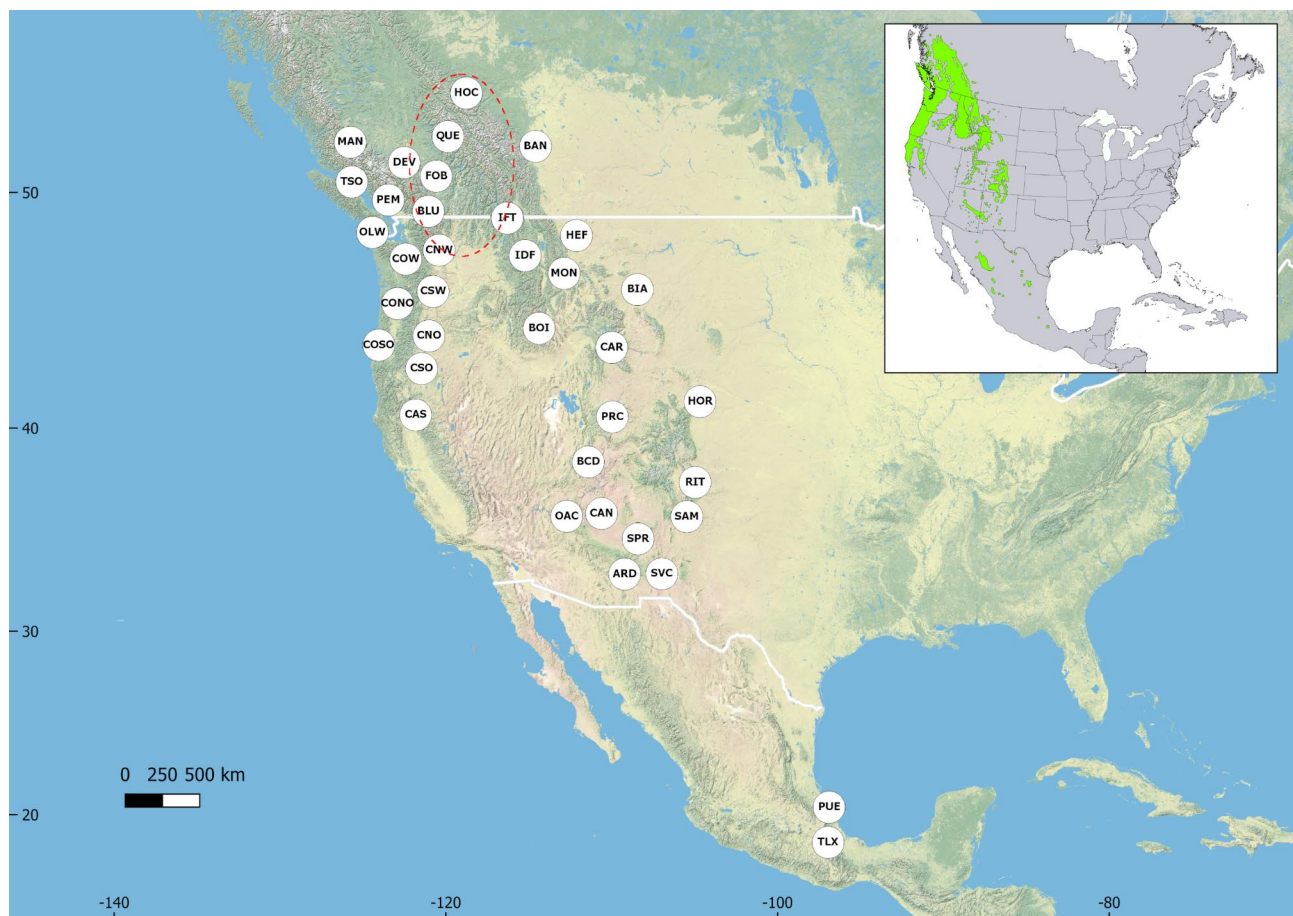


Fig. 1 Douglas-fir sampling localities. Circles represent each of the 37 sampled localities. Natural distribution range of the species (top right inset). The red dotted circle indicates the hybrid zone

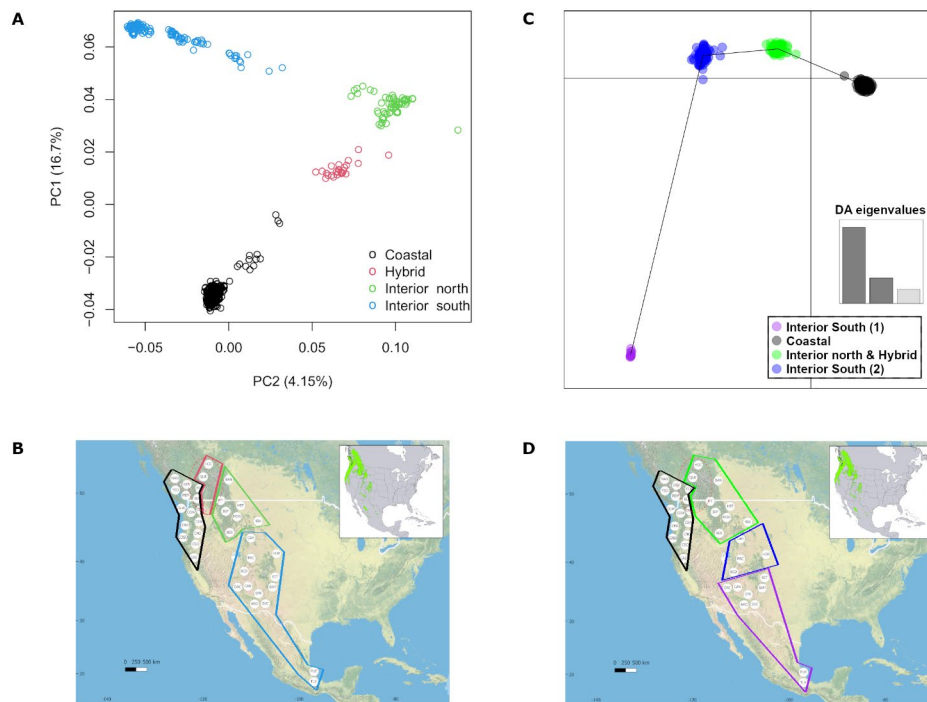


Fig. 2 Population genetic structure of Douglas-fir. **(A)** Principal component analysis of coastal ($N=325$) and interior ($n=188$) varieties, including hybrids ($n=27$), using 11,320 SNPs. **(B)** Distribution map of populations within PCA clusters. **(C)** DAPC for Douglas-fir populations. Black lines denote genomic distance between clusters. **(D)** Distribution map of clusters and populations determined by the DAPC analysis. Natural distribution range of the species (top right inset). The red dotted circle indicates the hybrid zone

(PUE), and Tlaxcala (TLX), from those in Utah (BCD) and Colorado (HOR; Fig. 2D).

Further analysis of the structure of the population with ADMIXTURE resulted in four ancestral populations (best K) and different levels of hybridization across the geographic distribution of the species (Fig. 3; Additional file 4; [27]). When considering two ancestral populations, the two varieties are mostly separated, and higher levels of admixture are observed in individuals of the interior variety from the north and in individuals from the hybrid zone (Fig. 3A). Although DAPC analyses clustered interior north and hybrids into one cluster, admixture proportions (Q values) revealed differences between the two groups (Fig. 3). For $K=3$, the interior variety is separated into north and south, with an admixture zone appearing between them (Fig. 3B). On the other hand, a greater contribution in the levels of admixture by the ancestral population corresponding to the individuals of the interior variety from the north is observed in the hybrids (Fig. 3B). For $K=4$, each variety was separated into two north to south ancestral populations, with regions of hybridization between them (Fig. 3C). Coastal Douglas-fir from British Columbia (BLU, TSO, PEM, MAN, and DEV) were differentiated from coastal individuals from the United States (COW, CNO, CAS, CONO, CNW, CSW, and COSO). Interior individuals from Alberta (BAN), Idaho (IDE, BOI), and Montana (MON, HEE,

BIA) clustered together, whereas a different genetic cluster was composed by interior individuals in the South (BCD, PRC, HOR, ARD, SPR, OAC, CAN, SVC, SAM, RIT, PUE, and TLX). The most diverse events of admixture occurred among individuals located in the contact zones in British Columbia and the Washington Cascades (HOC, QUE, and CNW; Fig. 3C). The prevailing admixture component in these individuals was from the interior north variety, suggesting asymmetric introgression. The second most important admixture contribution in the hybrids corresponded to the northern ancestral population of the coastal variety in British Columbia, which was coincident with the geographic position of the contact zone. Coastal and interior-north individuals located near the contact zones also presented low levels of admixture among their populations (Fig. 3C). Genetic structuring of trees varied among methods. Coastal, hybrids, interior-north, and interior-south groups showed clustering variation in the PCA analysis, while in the DAPC analysis the hybrid and interior-north groups are clustered together and certain populations of the interior-south group appear largely differentiated from the rest of the populations. Nonetheless, in the admixture analysis the coastal, hybrid, interior-north, and interior-south groups presented particular admixture proportions for each of the K values evaluated (Fig. 3).

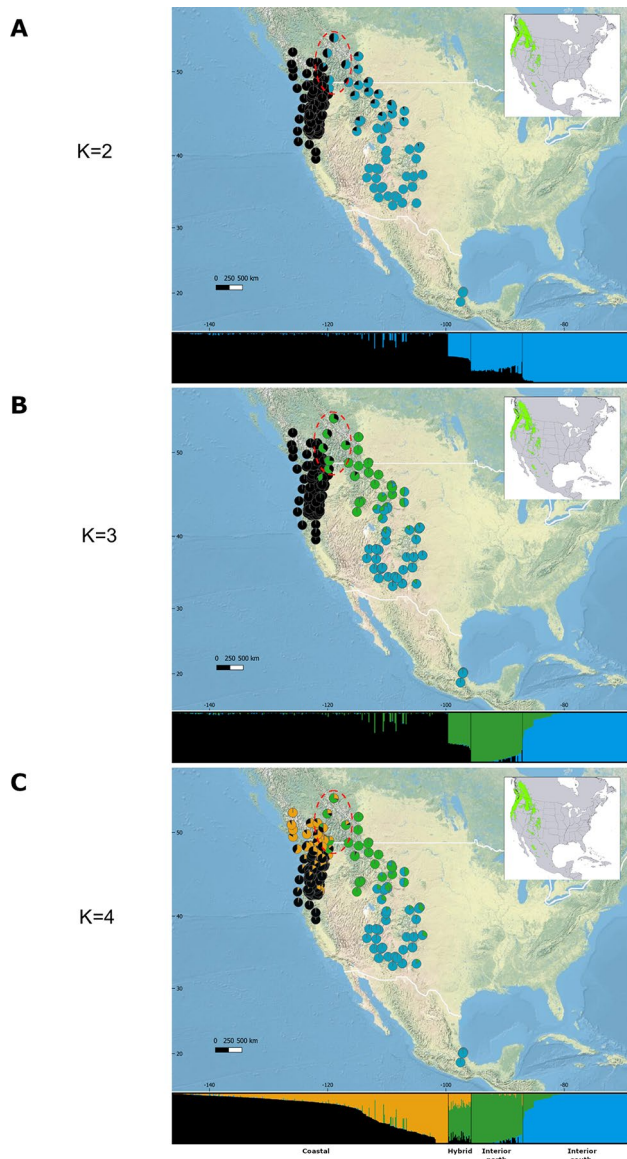


Fig. 3 Admixture analysis of Douglas-fir varieties. Clustering of individuals and geographic distribution of admixture proportions obtained with the Admixture software for **(A)** K2, **(B)** K3, and **(C)** K4. Natural distribution range of the species (top right inset). The red dotted circle indicates the hybrid zone

Hybrid classes

To get further insights into the attributes of the populations and especially the hybrids of the species, the NewHybrids program was used [29]. This program estimates the posterior probability of each individual of the population falling into different hybrid classes. All 325 coastal individuals, all 127 individuals from the interior-south group, and all 27 hybrids were used as input. The six genotype groups chosen were: pure interior parent, pure coastal parent, F1 hybrid, F2 hybrid, backcross with pure interior parent, and backcross with pure coastal parent. Almost all coastal and interior-south individuals

fell into the pure coastal parent and pure interior parent categories, respectively (Fig. 4). The majority (87%) of hybrids were classified as F2s, and no individual was predicted to be a first-generation (F1) hybrid.

Genetic diversity, population differentiation, and isolation by distance and environment

Inbreeding, heterozygosity, population differentiation (pairwise F_{st}) and isolation by distance were estimated from the dataset in this study. The population inbreeding coefficients ranged from 0.08 to 0.92. The highest inbreeding coefficients were found in Mexican populations and southern populations of the interior variety (Fig. 5); while the lowest inbreeding coefficients corresponded to coastal and hybrid populations, most of them located in British Columbia (Fig. 5). Coincident with inbreeding coefficient values, the lowest heterozygosity values corresponded to the Mexican and interior-south populations ($\mu=0.09$) and the highest values to the coastal and hybrids populations ($\mu=0.24$ and $\mu=0.23$, respectively; Fig. 6A). The northern interior populations presented an intermediate mean of 0.17. The two populations with the lowest heterozygosity values corresponded to the populations of the interior variety from Mexico (Puebla and Tlaxcala), while the two populations with the highest values corresponded to coastal populations located in British Columbia (Additional file 5). Higher genetic population differentiation based on pairwise F_{st} values was observed between populations of the different genetic groups than within their own populations (Fig. 6B; Additional file 6). The widest range in F_{st} values was found when comparing populations of the interior-south genetic group. The coastal and the interior-south populations were among the most genetically differentiated ones. Comparisons involving hybrids showed that these populations were more genetically similar to the interior-north populations, and more genetically distant from the interior-south populations (Fig. 6B). Aiming to examine if nearby populations were more genetically similar, a Mantel test was conducted to explore the correlation between geographic and genetic distances. The Mantel test showed a significant strong positive correlation between genetic and geographic distance among all the populations of the species (Fig. 6C; $r=0.74$ and $p=0.001$). Isolation by environment analysis also showed positive correlations between genetic distance and annual mean temperature (Fig. 6D; $r=0.35$ and $p=0.0001$) and precipitation variables (Fig. 6E; $r=0.11$ and $p=0.0001$). A significant positive relationship between geographic distance and temperature ($r=0.219$ and $p=0.0001$) and precipitation ($r=0.213$ and $p=0.0001$) variables was observed as well (Additional file 7).



Fig. 4 NewHybrids analysis of Douglas-fir trees. Bar plot of posterior probabilities of category membership for each Douglas-fir individual. Categories are divided as: pure interior parent (Pure 1), pure coastal parent (Pure 0), F1 hybrid (F1), F2 hybrid (F2), backcross with pure interior parent (0 Bx) and backcross with pure coastal parent (1 Bx)

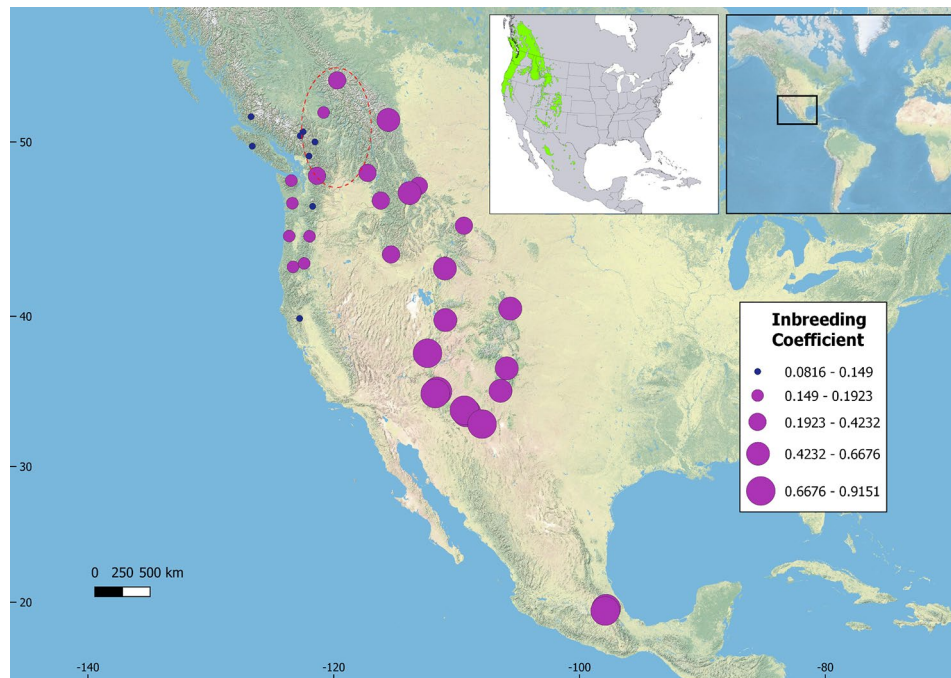


Fig. 5 Inbreeding coefficient analysis of Douglas-fir populations. Each circle represents one of the 37 collected populations. Circle size is proportional to the average inbreeding coefficient. Natural distribution range of the species (top right inset). The red dotted circle indicates the hybrid zone

Signatures of selection

Douglas-fir varieties differ in adaptive traits such as drought tolerance, frost hardiness and bud phenology, which allow populations to differentially respond to contrasting environmental conditions [47–49]. Signatures of selection in loci can reveal evolutionary processes of adaptive differentiation. To detect loci under selection in the 37 Douglas-fir populations, we used BayeScan and pcadapt tools [35, 36]. In total, 390 SNPs were identified as outliers by both tools, which matched 300 predicted Douglas-fir genes (Additional file 8). Enrichment analysis based on GO and KEGG pathway terms identified metabolic process, cellular process, and primary metabolism

among the most highly represented biological categories. DNA binding, catalytic activity and transcription regulation activity were the top three represented categories in the molecular function classification. For KEGG categories, the most highly represented pathways were phenylalanine metabolism, methane metabolism, and phenylpropanoid biosynthesis.

Signatures of local adaptation can be elucidated by analyzing markers that correlate with climatic data [39]. To evaluate (putative) signatures of local adaptation in Douglas-fir populations, 25 climatic variables related to temperature and precipitation were extracted from ClimateNA database [38]. Bayes factors for all pure and

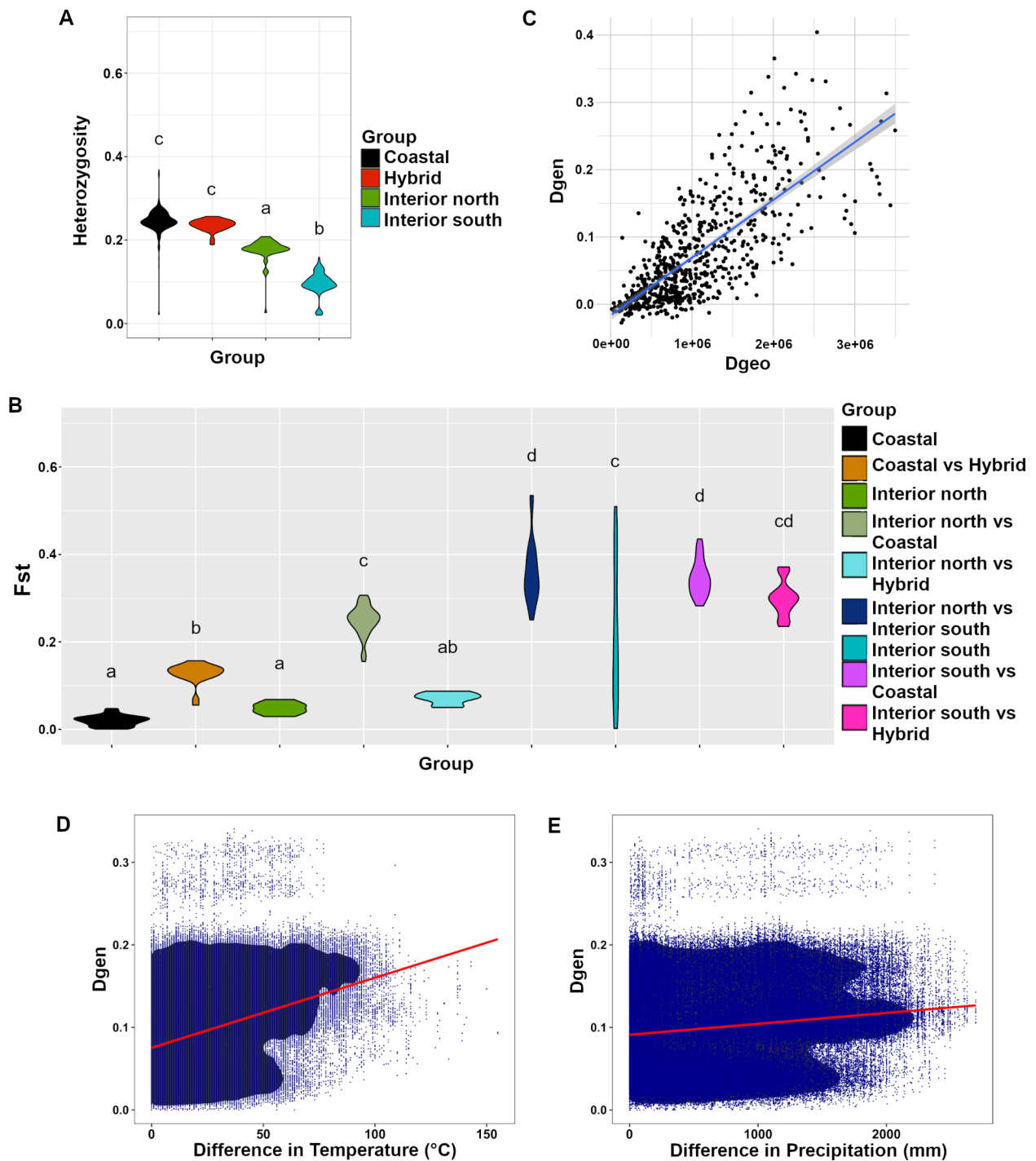


Fig. 6 Heterozygosity, Fst and Mantel analyses Violin plots of heterozygosity (**A**) and pairwise Fst (**B**) between coastal, interior-south, interior-north and hybrid groups. (**C, D, E**) Mantel test analysis of the correlations between geographic, genetic and environmental distances of Douglas-fir populations. Statistical analysis for group comparisons was performed using the SPSS software (version 29.0) through one-way ANOVA followed by Tukey's test ($p \leq 0.05$). Mean values sharing the same letter were not significantly different

hybrid populations identified 971 unique SNPs associated with the 25 climatic variables. These 971 SNPs were associated with 754 unique predicted genes (Additional file 8). The three most represented KEGG categories

for these genes were metabolic metabolism, phenylpropanoid biosynthesis and phenylalanine metabolism. The most represented GO categories for molecular function were catalytic activity, binding, and protein binding.

Response to stimulus, response to chemical compounds and metabolic processes were the top three biological processes. The number of SNPs that were associated with environmental variables and under selection were 64 (Fig. 7A). These SNPs were assigned to 36 annotated genes (Table 2). In addition, we evaluated associations between genotypes and environments in each variety separately using only the coastal or interior populations (Additional files 9–10). We detected 976 unique SNPs associated with environmental variables in either the interior (596 SNPs) or the coastal populations (380 SNPs). Only 18 SNPs were common between the two varieties (Fig. 7B). The top represented GO categories of the genes corresponding to coastal populations for biological processes, cellular component and molecular function were response to stimulus, cell-part, and

catalytic activity, respectively. In the case of the interior populations, the top represented GO categories of the genes associated to environmental variables for the biological processes, cellular component and molecular function categories were response to stimulus, cell, and catalytic activity, respectively.

Environmental niche modeling

The characterization of the environmental space occupied by Douglas-fir populations was based on a pairwise comparison that allowed us to identify the climate variables driving a niche differentiation between clusters. We identified 11 highly correlated variables and 8 variables that showed a high predictive power as indicated by the explained variance on the PCA's first two components. Accordingly, the most explicative set of variables were slightly different for each of the pairwise comparisons (listed in Additional file 11). The environmental ordination analysis identified that precipitation of the coldest quarter (BIO19), temperature annual range (BIO7) and maximum temperature of warmest month (BIO5) segregated coastal from interior north populations. Similarly, minimum temperature of the coldest month (BIO6), mean temperature of the coldest quarter (BIO11), and precipitation of the coldest quarter separated coastal from hybrid populations. The rest of the relationships are shown in Additional file 12.

The highest bioclimatic niche overlap occurred between interior north–hybrid populations ($D=0.26$, $p=0.05$), followed by interior north–interior south ($D=0.14$, $p=0.06$) and coastal–interior south ($D=0.1$, $p=0.08$). Interestingly, the smallest overlap was found between coastal and both interior north and hybrid populations ($D=0.02$ and 0.06 , respectively), despite coastal being geographically contiguous to both (Additional file 13, Additional file 14A–F). Occurrence density grids, depicting a pairwise comparison of the portion of the environmental space that is more densely occupied by each population, further clarify niche overlap patterns (Additional file 15A–F).

Projections of potential distribution

Spatially explicit projections of niche models allow us to discern the habitat suitability across the Douglas-fir range (Fig. 8A). For the coastal variety, a large tract of continuous suitable habitat occurs along the Pacific Northwest, especially in Washington, extreme southern British Columbia and Vancouver Island, and Northern California east of the Great Valley. As for interior north populations highest suitability is predicted for most of Idaho (except for the Snake River plain), southeastern British Columbia, central Montana, and eastern portions of Washington and Oregon. Interior south suitability is scattered across the mountain ranges of Colorado, Utah,

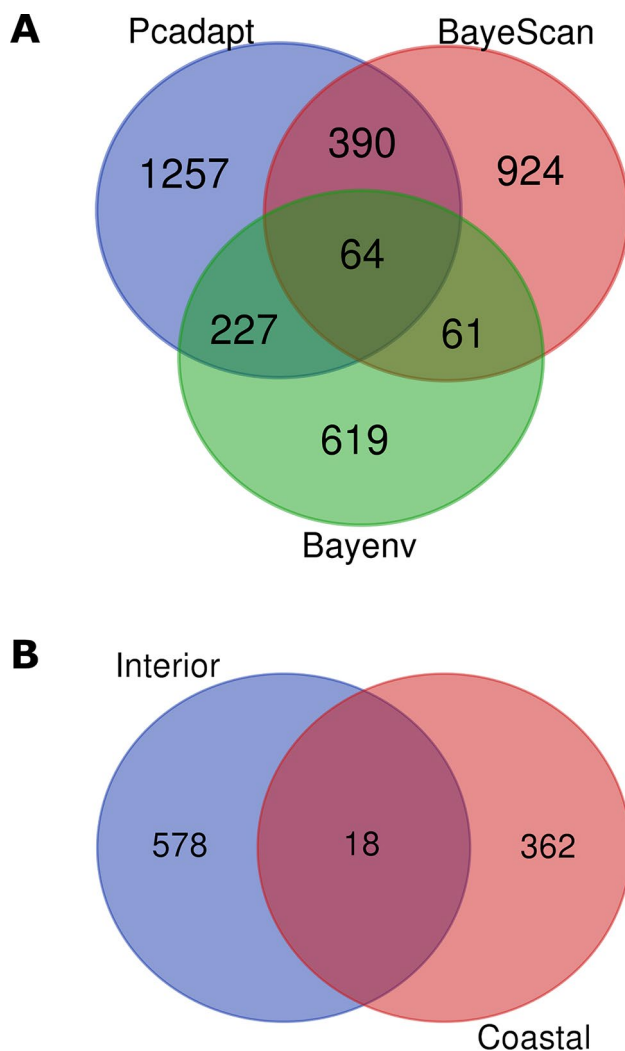


Fig. 7 Venn diagrams of the number of outlier SNPs and SNPs-associated with environmental variables. **(A)** Venn diagram showing overlap between the SNPs detected by each method using all populations of Douglas-fir. **(B)** Venn diagram showing overlap between the SNPs detected by Bayenv among coastal and interior populations

Table 2 Genes associated with environmental variables and detected as outliers

Gene ID	Locus (AT)	Gene description	SNP ID	Environmental variables
PSME_03744	AT1G72470	EXOCYST SUBUNIT EXO70 FAMILY PROTEIN D1 (EXO70D1)	seq-rs8794-DF	TD, EMT
PSME_09672	AT5G26600	Pyridoxal phosphate (PLP)-dependent transferase protein	seq-rs7221-DF	MSP, CMD
PSME_11002	AT4G14950	KILLING ME SLOWLY 1 (KMS1)	seq-rs5525-DF	TD, bFFP, EMT, MAR
PSME_13942	AT5G58610	PHD finger transcription factor	seq-rs2638-DF	TD, MAR
PSME_15259	AT4G15020	hAT transposon superfamily	seq-rs11624-DF	TD, EMT, MAR
PSME_16148	AT1G69040	ACT DOMAIN REPEAT 4 (ACR4)	seq-rs2413-DF	MCMT, TD, bFFP, EMT, MAR
PSME_17634	AT1G65010	Microtubule-associated protein	seq-rs7694-DF	DD18, DD1040
PSME_18531	AT5G52390	PAR1 protein	seq-rs14799-DF	MAT
PSME_19641	AT2G13600	SLOW GROWTH 2 (SLO2)	seq-rs20280-DF	MAT, NFFD, FFP, PAS
PSME_19644	AT2G28080	TARGETS UNDER ETTIN CONTROL3 (TEC3)	seq-rs12377-DF, seq-rs12378-DF	MCMT, TD, EMT, MAT, DD_0, DD5, DD_18, bFFP, FFP
PSME_21214	AT3G07350	Phosphorus (P) stress-inducible DUF506 gene	seq-rs404-DF	EMT
PSME_23343	AT3G17450	hAT dimerization domain-containing protein	seq-rs5518-DF	MCMT, TD, DD_18, bFFP, FFP, EMT, MAR
PSME_23814	AT3G48060	BAH and TFIS domain-containing protein	seq-rs3329-DF, seq-rs3331-DF	MAR
PSME_25837	AT1G71050	HEAVY METAL ASSOCIATED ISOPRENYLATED PLANT PROTEIN 20 (HIPP20)	seq-rs15203-DF	TD, FFP, MAR
PSME_26257	AT3G13460	EVOLUTIONARILY CONSERVED C-TERMINAL REGION 2 (ECT2)	seq-rs12858-DF, seq-rs12859-DF	MCMT, TD, DD_0, DD_18, bFFP, FFP, EMT, MAR
PSME_26527	AT3G17450	hAT dimerization domain-containing protein	seq-rs6469-DF	DD_0
PSME_26913	AT3G17450	hAT dimerization domain-containing protein	seq-rs24300-DF	DD5, bFFP, FFP, MAR
PSME_27258	AT4G21960	PRXR1	seq-rs5609-DF	TD, MAR
PSME_27924	AT3G17450	hAT dimerization domain-containing protein	seq-rs16255-DF	MAR
PSME_28111	AT4G25610	C2H2-like zinc finger protein	seq-rs12212-DF	MAT, DD_0
PSME_30358	AT4G33240	FORMS APLOID AND BINUCLEATE CELLS 1 A (FAB1A)	seq-rs2130-DF	MAR
PSME_31523	AT1G68150	WRKY DNA-BINDING PROTEIN 9 (WRKY9)	seq-rs5088-DF	TD, MAP, MSP
PSME_32133	AT3G12250	TGACG MOTIF-BINDING FACTOR 6 (TGA6)	seq-rs6065-DF	MAT, DD5, CMD, DD1040
PSME_36236	AT3G45630	RNA binding (RRM/RBD/RNP motifs) protein	seq-rs3467-DF	MAT, MCMT, TD, DD_0, DD_18, bFFP, FFP, EMT, MAR
PSME_37507	AT2G27900	Coiled-coil protein	seq-rs14589-DF	DD_0
PSME_38327	AT2G07690	MINICHROMOSOME MAINTENANCE 5 (MCM5)	seq-rs15648-DF	DD_0
PSME_40364	AT4G39420	Spatacsin carboxy-terminus protein	seq-rs4908-DF, seq-rs4906-DF, seq-rs4909-DF, seq-rs4901-DF, seq-rs4902-DF, seq-rs4905-DF	MAR
PSME_42442	AT3G18110	EMBRYO DEFECTIVE 1270 (EMB1270)	seq-rs15950-DF, seq-rs15954-DF	DD_0, MCMT, MAT
PSME_46484	AT5G20480	EF-TU RECEPTOR (EFR)	seq-rs6514-DF	TD, DD_0, EMT
PSME_46939	AT4G15020	hAT transposon superfamily	seq-rs10937-DF	DD_0
PSME_48792	AT4G26640	WRKY Transcription Factor 20	seq-rs3486-DF, seq-rs3487-DF	MAT, MCMT, TD, DD_0, DD_18, bFFP, FFP, EMT, MAR
PSME_50000	AT1G13580	LAG1 LONGEVITY ASSURANCE HOMOLOG 3 (LAG13)	seq-rs6971-DF	TD, bFFP, EMT, MAR
PSME_50606	AT3G17450	hAT dimerization domain-containing protein	seq-rs6498-DF	MAT, DD5, FFP, DD1040
PSME_50684	AT4G03400	DWARF IN LIGHT 2 (DFL2)	seq-rs15961-DF	MCMT, TD, EMT, MAR
PSME_50787	AT4G15020	hAT transposon superfamily	seq-rs14879-DF	MCMT
PSME_51247	AT4G39420	Spatacsin carboxy-terminus protein	seq-rs2918-DF	

Functional annotation of genes (ID of Douglas-fir genome version Psme v1.0) shared by all three methods for all populations of Douglas-fir

New Mexico, and Arizona, as well as high elevations in Mexico (Fig. 8A).

Projections of Douglas-fir potential distribution over the next 50 years showed a leading- and trailing-edge pattern. On one hand, northernmost populations (i.e., coastal, and interior north) are expected to experience a poleward shift with a concomitant overall gain

in potential distribution. On the other hand, interior south, and especially Mexican populations are expected to experience severe losses of suitable habitat, particularly at lower elevations, despite some upslope advances on the highest peaks. Depending on any given Representative Concentration Pathway (RCP) scenario between 2050 and 2070, the overall distribution of Douglas-fir

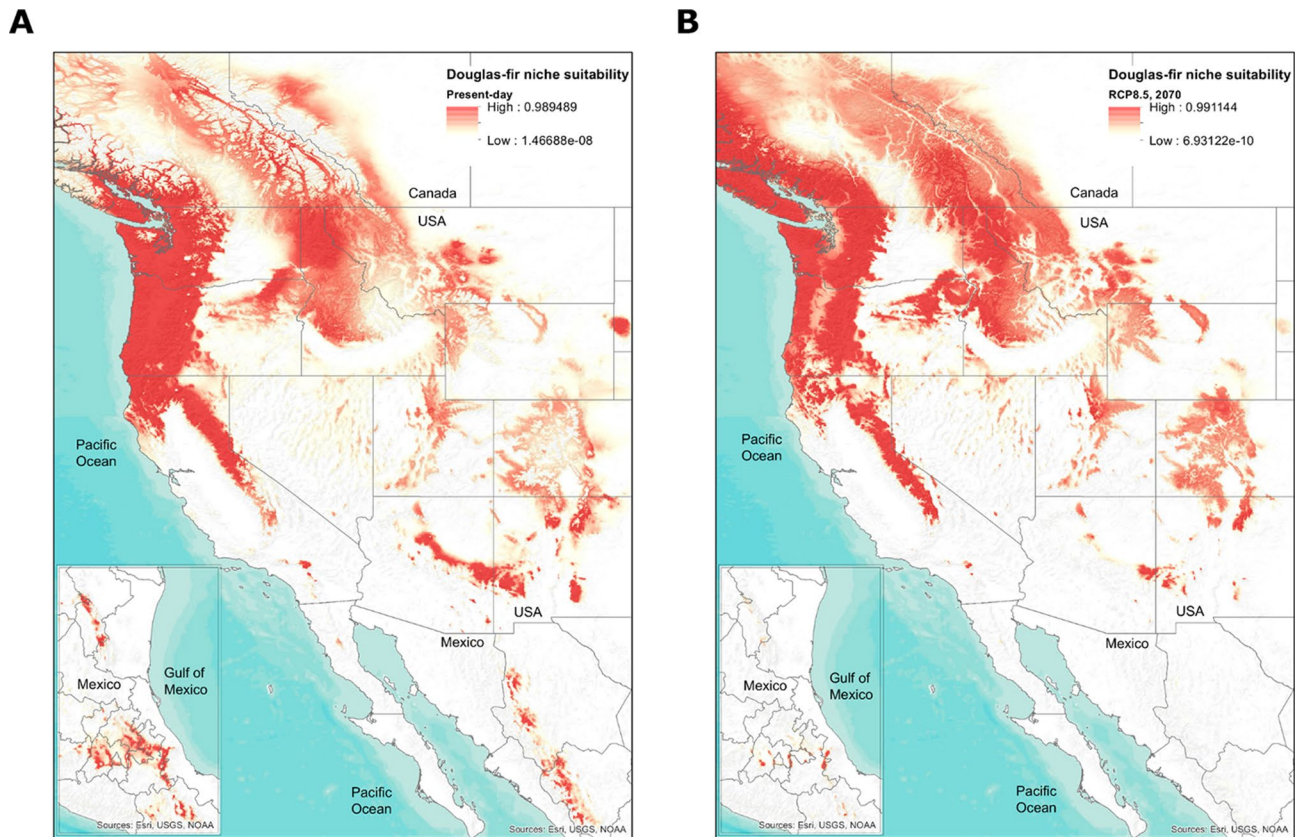


Fig. 8 Douglas-fir ecological niche suitability across western North America, with the inset showing eastern Mexico. **(A)** Present-day potential niche. **(B)** Potential niche at RCP8.5 scenario at year 2070

Table 3 Douglas-fir potential distribution predicted shifts over the next 50 years in North America in relation to different representative concentration pathways (RCP)

Population cluster	Contemporary niche extent (km ²)	Future niche gains (+) or losses (-) in relation to contemporary niche extent (%)							
		RCP2.6		RCP4.5		RCP6.0		RCP8.5	
Rangewide	~600k	2050	2070	2050	2070	2050	2070	2050	2070
		804k	798k	816k	852k	804k	750k	810k	792k
		(+34)	(+33)	(+36)	(+42)	(+34)	(+25)	(+35)	(+32)
Coastal	~80k	114k	104k	134k	143k	143k	136k	142k	168k
		(+43)	(+30)	(+68)	(+79)	(+79)	(+70)	(+78)	(+110)
		640k	623k	661k	691k	691k	704k	696k	760k
Interior north	~427k	(+50)	(+46)	(+55)	(+62)	(+62)	(+65)	(+63)	(+78)
		53k	55k	48k	40k	54k	35k	32k	20k
		(-43)	(-41)	(-48)	(-57)	(-42)	(-62)	(-65)	(-78)

in North America could increase by 25–42% (Table 3). Nonetheless, on a per-population basis, gains and losses are exacerbated. Under the RCP2.6 scenario in 2070, the coastal population is projected to experience at least a one-third increase in suitable conditions, while in the event of an RCP8.5 scenario during that decade, the increase may exceed a doubling (Table 3). On the contrary, interior south populations are expected to endure a drastic reduction of the contemporary suitable habitat, with losses reaching 65–78% in the worst-case scenarios

(Table 3). Changes in suitability for the interior north populations are predicted to be mostly gains although not as pronounced as for those in the Pacific Northwest (Fig. 8B). Finally, we consider that small sample size for the hybrid population precluded a robust estimation of their niche shifts across time.

Discussion

Role of selection and gene flow maintaining population structure in the species

The environmental adaptation of long-lived trees is crucial for forest survival. The distribution of the species from the genus *Pseudotsuga* is discontinuous, but wide. They are present in North America, Mexico, and Asia. Douglas-fir, the only species from the Pinaceae family with 26 chromosomes, has one of the broadest ranges of any conifer from North America. It has been introduced into temperate regions since the mid-19 century. There are natural populations of both varieties of Douglas-fir that have high levels of genetic diversity with potential to confer resilience to varying climates. In this study, we used a custom-designed SNP array to study the nuclear genome variation, its relation to space and the local adaptation of several natural populations of Douglas-fir.

The SNPs found in the present study revealed clear genetic structure between the two varieties as previously reported [13, 15–19]. The best K-value for the number of ancestral populations in our admixture analysis was four. The number of ancestral populations previously reported with different genetic markers and different methods is variable; however different ancestral populations have been reported within each variety that distinguishes north and south individuals, which indicates that the structure is geographically shaped. Indeed, we observed that F_{st} values in pairwise comparisons were higher between distant populations of the two varieties, suggesting reduced gene flow in these populations. Strong correlation of patterns of population genetic variation that derive from spatially limited gene flow were also detected through the isolation by distance analysis. Geographic distance was significantly correlated with genetic distance, a stronger correlation was even detected than with environmental variables; however, we also found significant correlation between geographic distance and environmental distances, highlighting spatial and environmental heterogeneous selection in this species.

Inter-variety hybridization and introgression

Natural hybridization between the interior and coastal varieties was found in the contact zones in British Columbia and the Washington Cascades. Hybrids in both locations present a higher ancestry from the interior north variety, suggesting asymmetric introgression from the interior to the coastal variety [13, 15]. Mountain ranges are often a physical barrier to gene flow in widely dispersed species. Conifers, being wind pollinated species, may be susceptible to elevational differences between populations [50]. Douglas-fir grows from sea level to over 3000 m of elevation with northern interior populations growing at the highest altitude [8, 51]. Hybrids grow at intermediate elevations between northern coastal and

northern interior populations which may explain the higher ancestry of the interior north variety we observed as gene flow from the coastal populations could be limited due to the elevational gradient [52]. Previous studies have identified asymmetric introgression in hybrids between *Pinus strobiformis* and *Pinus flexilis* where cold-resilient genes from *P. flexilis* were favored and maintained in hybrid populations [53]. A similar pattern may exist in our study system where asymmetric introgression from interior Douglas-fir confers a fitness advantage in natural populations of hybrids [13, 15]. Previous work in Douglas-fir suggests that adaptive introgression from the interior variety may have resulted in natural populations of hybrids with increased water-use efficiency (WUE) and heat tolerance [8].

Our results also suggest that certain regions of the hybrid zone may not be of recent origin (as inferred by the lack of F1 hybrids). All hybrids found in this study were advanced-generation hybrids, probably resulting from the cross of two hybrid parents. This suggests that the hybrid zones could be maintained by some level of hybrid fitness, in which hybrids are able to mate, and survive at higher rates than pure varieties in novel environments. Hybrid individuals may be able to occupy unique novel environments that differ from the environment that the parents experienced. This may occur because of hybridization mechanisms such as transgressive segregation which generates unique packages of alleles and can generate extreme phenotypes that exceed that of the pure varieties [54]. Future research to evaluate transgressive segregation in hybrids of Douglas-fir may shed light on the cause of the establishment of these populations in harsh conditions such as the Washington Cascades and mountains in British Columbia.

Candidate loci for local adaptation

Understanding the genomic basis of local adaptation in long-lived trees is crucial to predict their responses to the oncoming effects of climate change. Local adaptation emerges across heterogeneous environments, which as previously shown is the case with Douglas-fir due to its wide distribution. With the use of the Bayenv mixed model we found inter- and intra-variety differences in the frequency of gene loci that indicate local adaptation in Douglas-fir. The functions of several of the genes associated with environmental variables and detected as outliers are associated with abiotic stress responses in *Arabidopsis thaliana* and are therefore probably relevant to local adaptation and with similar function in Douglas-fir. For example, we found that some genes (*ECT2*, *SLO2*, and *FABIA*) that were associated with climatic variables respond or are related to abscisic acid (ABA), which is crucial for abiotic stress responses (Table 2). The *ECT2* (EVOLUTIONARILY CONSERVED C-TERMINAL

REGION 2) gene encodes a YTH domain-containing reader protein that regulates transcriptional and post-transcriptional gene expression through recognition of m6A modifications. ECT2, ECT3 and ECT4 have genetically redundant functions in ABA response regulation and their disruption destabilizes mRNAs of ABA signaling related genes resulting in ABA hypersensitivity [55]. The SLO2 gene encodes a pentatricopeptide repeat protein that functions as a mitochondrial RNA editing factor. Disruption of SLO2 function results in ABA hypersensitivity, insensitivity to ethylene and increased drought and salt tolerance [56]. The FAB1A gene encodes a 1-phosphatidylinositol-3-phosphate 5-kinase. Null mutants of phosphatidylinositol 3-phosphate 5-kinases in *Arabidopsis* presented delayed stomatal closure during ABA treatment and increased water loss [57]. Other genes detected as outliers respond to light stimuli (TEC3, BAH and TFIIS domain-containing protein, and DFL2), oxidative stress (PRXR1), and temperature (C2H2-like zinc finger protein). The transcription factor WRKY9 regulates salt tolerance and LAG13 is involved in hypoxia tolerance [58, 59].

As probably expected, due to its wide distribution, we also detected signals of local adaptation within populations of each variety in Douglas-fir. Most of the SNPs associated to environmental variables within each variety were not shared, which denotes differentiation and genetic diversity among them; however, the top represented GO term categories in each variety were similar, suggesting that genes associated with environmental variables in both varieties play related roles. Only 18 SNPs were shared between the two varieties. Within the genes that contain shared SNPs, we found genes relevant to the stress responses in plants. For example, one of these genes is the previously mentioned SLO2 gene that encodes a pentatricopeptide repeat protein [56]. Also, within this group of genes, we found the gamma tonoplast intrinsic protein 2 (*TIP2*) gene, which is involved in ABA and salinity stress responses [60]. The D6 protein kinase like 2 and RAD1, involved in phototropism and resistance to UV radiation, respectively were detected as well [61, 62]. An AP2/ERF transcription factor, member of the ethylene signaling pathway, that confers resistance to heat, and hydrogen peroxide stresses was also detected with shared SNPs [63].

Future distribution of the species under climate change

A few studies have assessed the ecological niche of Douglas-fir, among other conifers, and the potential shifts in distribution in response to future scenarios of climate change [64–66]. In general, the evidence we found in our study supports the general findings of those investigations, which can be summarized in two major aspects. First, the ample separation regarding the environmental

space occupied by the two main Douglas-fir varieties (i.e., coastal, inland), which in turn is reflected in a geographic clustering on a regional scale, coherent with their genetic structuring; and second, the contrasting impacts of climate change over the next decades on the distribution of the distinct populations. As global climate gets increasingly warmer, leading-edge populations of temperate tree species such as Douglas-fir are expected to migrate poleward [67–69], either to track suitable conditions or to colonize new available areas, while trailing-edge populations will be forced to engage in upslope migrations, which nonetheless will be greatly constrained by local geomorphological features. As a result, over the next 50 years, the Douglas-fir suitable habitat is expected to increase at least 25% in relation to their current extent. The potential gains are even larger when considering the extreme warming that could take place in an RCP8.5 scenario. Nevertheless, the considered timespan (5 decades) is certainly small considering the longevity and generation time of the Douglas-fir. Beyond that period, erratic oscillations in global climate patterns could disrupt any temporary equilibrium in ecological communities, including temperate coniferous forests.

Conclusions

Our range-wide population structure and genetic diversity analyses of Douglas-fir varieties indicate high genetic variation in their populations with clear structured populations and differential adaptive potential to different climate scenarios as result of spatially heterogeneous selection and dissimilar evolutionary histories. The genetic variation associated with climate data found here helps our understanding of the molecular mechanisms controlling long-living trees adaptation. The future distributions predicted in this study revealed different evolutionary paths for Douglas-fir varieties, which constitutes valuable information regarding the conservation and management of the species.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-024-02304-4>.

Additional file 1. Summary of sample collection

Additional file 2. Environmental variables used for SNP-climate associations

Additional file 3. Summary of population statistics

Additional file 4. Bar plots of admixture proportions of Douglas-fir populations from K2 to K4

Additional file 5. Violin plots of heterozygosity values per population

Additional file 6. Pairwise F_{st} values per population

Additional file 7. Mantel test correlations between geographic and environmental distances

Additional file 8. SNPs associated with environmental variables and detected as outliers in Douglas-fir populations

Additional file 9. SNPs associated with environmental variables of coastal populations

Additional file 10. SNPs associated with environmental variables of interior populations

Additional file 11. Rangewide evaluation of the Ecological Niche models

Additional file 12. Environmental ordination of Douglas-fir populations based on a Principal Component Analysis characterization of their Hutchinsonian niche

Additional file 13. Pairwise niche overlap (D) between Douglas-fir population clusters

Additional file 14. Niche overlap (D) between Douglas-fir populations

Additional file 15. Pairwise comparison of occurrence density grids

Acknowledgements

The authors would like to thank Benjamin Wilhite and Matthew Weiss for collecting some of the samples used in this study.

Author contributions

A.R.D.L.T. designed and supervised the project and obtained research funding; K.B. and J.R.M. collected samples; K.B. performed all lab analyses; P.P. carried out all population genomic analyses; G.P.L. performed the climate niche modeling; P.P., A.R.D.L.T., K.B. and G.P.L. wrote the manuscript; all authors approved the final version of the manuscript.

Funding

This project was supported by the National Science Foundation [CAREER project 2145834] to A.R.D.L.T.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 12 December 2023 / Accepted: 28 August 2024

Published online: 03 September 2024

References

- Sork VL, Aitken SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genet Genomes*. 2013;9:901–11.
- Savolainen O, Pyhäjärvi T, Knürr T. Gene Flow and local adaptation in Trees. *Annu Rev Ecol Syst*. 2007;38:595–619.
- Feng L, Du FK. Landscape Genomics in Tree Conservation under a changing environment. *Front Plant Sci*. 2022;13.
- Bansal S, Harrington CA, Gould PJ, St.Clair JB. Climate-related genetic variation in drought-resistance of Douglas-fir (*Pseudotsuga menziesii*). *Glob Chang Biol*. 2015;21:947–58.
- Montwé D, Spiecker H, Hamann A. Five decades of growth in a genetic field trial of Douglas-fir reveal trade-offs between productivity and drought tolerance. *Tree Genet Genomes*. 2015;11:29.
- Marias DE, Meinzer FC, Woodruff DR, McCulloh KA. Thermotolerance and heat stress responses of Douglas-fir and ponderosa pine seedling populations from contrasting climates. *Tree Physiol*. 2016. <https://doi.org/10.1093/treephys/tpw117>.
- De La Torre AR, Wilhite B, Puiu D, St. Clair JB, Crepeau MW, Salzberg SL, et al. Dissecting the Polygenic Basis of Cold Adaptation Using Genome-Wide Association of Traits and Environmental Data in Douglas-fir. *Genes (Basel)*. 2021;12:110.
- Compton S, Stackpole C, Dixit A, Sekhwal MK, Kolb T, De la Torre AR. Differences in heat tolerance, water use efficiency and growth among Douglas-fir families and varieties evidenced by GWAS and common garden studies. *AOB Plants*. 2023;15.
- Dominguez-Álvarez FA. Análisis histórico-ecológico De Los bosques de Pseudotsuga en México. *Folleto Técnico INIFAP*. 1994:43.
- Reyes-Hernández V, Vargas-Hernández J, López-Upton J, Vaquera-Huerta H. Phenotypic similarity among Mexican populations of Pseudotsuga Carr. *Agrociencia*. 2006;40:545–56.
- Acevedo-Rodríguez R, Vargas-Hernandez JJ, Lopez-Upton J, Mendoza JV. Effect of geographic origin and nutrition on shoot pre-nology of Mexican Douglas-fir (*Pseudotsuga* sp.) seedlings. *Agrociencia*. 2006;40:125–37.
- Cruz-Nicolás J, Vargas-Hernández JJ, Ramírez-Vallejo P, López-Upton J. Genetic diversity and differentiation of *Pseudotsuga menziesii* (Mirb.) Franco populations in Mexico. *Revista Fitotecnia Mexicana*. 2011;34:233–40.
- Gugger PF, Sugita S, Cavender-Bares J. Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Mol Ecol*. 2010;19:1877–97.
- Hermann RK. The genus *Pseudotsuga*: ancestral history and past distribution. *Forest Research Laboratory Special Publication 2b*. Edited by: Corvallis, OR. Oregon State University, College of Forestry; 1985.
- van Loo M, Hintsteiner W, Pötzelsberger E, Schüller S, Hasenauer H. Intervarietal and intravarietal genetic structure in Douglas-fir: nuclear SSRs bring novel insights into past population demographic processes, phylogeography, and intervarietal hybridization. *Ecol Evol*. 2015;5:1802–17.
- Wei X-X, Beaulieu J, Khasa DP, Vargas-Hernández J, López-Upton J, Jaquish B, et al. Range-wide chloroplast and mitochondrial DNA imprints reveal multiple lineages and complex biogeographic history for Douglas-fir. *Tree Genet Genomes*. 2011;7:1025–40.
- Li P, Adams WT. Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). *Can J for Res*. 1989;19:149–61.
- Neophytou C, Weisser A-M, Landwehr D, Šeho M, Kohnle U, Ensminger I, et al. Assessing the relationship between height growth and molecular genetic variation in Douglas-fir (*Pseudotsuga menziesii*) provenances. *Eur J Res*. 2016;135:465–81.
- Hintsteiner WJ, van Loo M, Neophytou C, Schueler S, Hasenauer H. The geographic origin of old Douglas-fir stands growing in Central Europe. *Eur J Res*. 2018;137:447–61.
- Eckert AJ, Wegryzn JL, Pande B, Jermstad KD, Lee JM, Liechty JD, et al. Multilocus Patterns of Nucleotide Diversity and divergence reveal positive selection at candidate genes related to Cold Hardiness in Coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). *Genetics*. 2009;183:289–98.
- Müller T, Freund F, Wildhagen H, Schmid KJ. Targeted re-sequencing of five Douglas-fir provenances reveals population structure and putative target genes of positive selection. *Tree Genet Genomes*. 2015;11:816.
- Hess M, Wildhagen H, Junker LV, Ensminger I. Transcriptome responses to temperature, water availability and photoperiod are conserved among mature trees of two divergent Douglas-fir provenances from a coastal and an interior habitat. *BMC Genomics*. 2016;17:682.
- George J, Schueler S, Grabner M, Karanitsch-Ackerl S, Mayer K, Stierschneider M, et al. Looking for the needle in a downsized haystack: whole-exome sequencing unravels genomic signals of climatic adaptation in Douglas-fir (*Pseudotsuga menziesii*). *Ecol Evol*. 2021;11(12):8238–53.
- Nelson TC, Stathos AM, Vanderpool DD, Finseth FR, Yuan Y, Fishman L. Ancient and recent introgression shape the evolutionary history of pollinator adaptation and speciation in a model monkeyflower radiation (*Mimulus* section *Erythranthe*). *PLoS Genet*. 2021;17:e1009095.
- Zheng X, Levine D, Shen J, Gogarten SM, Laurie C, Weir BS. A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. 2012;28:3326–8.
- Jombart T. Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008;24:1403–5.
- Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009;19:1655–64.

28. Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol Ecol Resour.* 2015;15:1179–91.
29. Anderson EC, Thompson EA. A model-based method for identifying species hybrids using Multilocus Genetic Data. *Genetics.* 2002;160:1217–29.
30. Ochoa A, Storey JD. Estimating FST and kinship for arbitrary population structures. *PLoS Genet.* 2021;17:e1009241.
31. Little EL. Atlas of United States trees. Washington, D.C.: U.S. Dept. of Agriculture, Forest Service; 1971.
32. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27:2156–8.
33. Wickham H. ggplot2. New York, NY: Springer New York; 2009.
34. Gruber B, Unmack PJ, Berry OF, Georges A. Darta: an R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour.* 2018;18:691–9.
35. Foll M, Gaggiotti O. A genome-scan Method to identify selected loci appropriate for both Dominant and Codominant markers: a bayesian perspective. *Genetics.* 2008;180:977–93.
36. 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.
37. Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, et al. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. *Nucleic Acids Res.* 2021;49:W317–25.
38. Wang T, Hamann A, Spittlehouse D, Carroll C. Locally downscaled and spatially customizable Climate Data for historical and future periods for North America. *PLoS ONE.* 2016;11:e0156720.
39. Günther T, Coop G. Robust identification of local adaptation from Allele Frequencies. *Genetics.* 2013;195:205–20.
40. Kass JM, Pinilla-Buitrago GE, Paz A, Johnson BA, Grisales Betancur V, Meenan SI et al. wallace 2: a shiny app for modeling species niches and distributions redesigned to facilitate expansion via module contributions. *Ecography.* 2023;2023.
41. Phillips SJ, Anderson RP, Schapire RE. Maximum entropy modeling of species geographic distributions. *Ecol Modell.* 2006;190:231–59.
42. Phillips SJ, Anderson RP, Dudík M, Schapire RE, Blair ME. Opening the black box: an open source release of Maxent. *Ecography.* 2017;40:887–93.
43. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol.* 2005;25:1965–78.
44. Hutchinson GE. Concluding remarks. *Cold Spring Harb Symp Quant Biol.* 1957;22:415–27.
45. Schoener TW. The Anolis lizards of Bimini: resource partitioning in a Complex Fauna. *Ecology.* 1968;49:704–26.
46. Elith J, Phillips SJ, Hastie T, Dudík M, Chee YE, Yates CJ. A statistical explanation of MaxEnt for ecologists. *Divers Distrib.* 2011;17:43–57.
47. Candido-Ribeiro R, Aitken SN. Weak local adaptation to drought in seedlings of a widespread conifer. *New Phytol.* 2024;241:2395–409.
48. Jermstad KD, Bassoni DL, Jech KS, Ritchie GA, Wheeler NC, Neale DB. Mapping of quantitative trait loci Controlling adaptive traits in Coastal Douglas Fir. III. Quantitative trait loci-by-environment interactions. *Genetics.* 2003;165:1489–506.
49. Rehfeldt GE. Growth and cold hardiness of intervarietal hybrids of douglas-fir. *Theor Appl Genet.* 1977;50:3–15.
50. Li Y-S, Shih K-M, Chang C-T, Chung J-D, Hwang S-Y. Testing the Effect of Mountain ranges as a physical barrier to current Gene Flow and environmentally dependent adaptive divergence in *Cunninghamia Konishii* (Cupressaceae). *Front Genet.* 2019;10.
51. Gould PJ, Harrington CA, St. Clair JB. Incorporating genetic variation into a model of budburst phenology of coast Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). *Can J for Res.* 2011;41:139–50.
52. Silen RR. Genetics of Douglas-fir. United States Forest Service Research Paper WO-35 1978.
53. Menon M, Bagley JC, Page GFM, Whipple AV, Schoettle AW, Still CJ, et al. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. *Commun Biol.* 2021;4:160.
54. Rieseberg LH, Widmer A, Arntz AM, Burke B. The genetic architecture necessary for transgressive segregation is common in both natural and domesticated populations. *Philos Trans R Soc Lond B Biol Sci.* 2003;358:1141–7.
55. Song P, Wei L, Chen Z, Cai Z, Lu Q, Wang C, et al. m6A readers ECT2/ECT3/ECT4 enhance mRNA stability through direct recruitment of the poly(A) binding proteins in Arabidopsis. *Genome Biol.* 2023;24:103.
56. Zhu Q, Dugardeyn J, Zhang C, Mühlenbock P, Eastmond PJ, Valcke R, et al. The Arabidopsis thaliana RNA editing factor SLO2, which affects the Mitochondrial Electron Transport Chain, participates in multiple stress and hormone responses. *Mol Plant.* 2014;7:290–310.
57. Bak G, Lee E-J, Lee Y, Kato M, Segami S, Sze H, et al. Rapid Structural Changes and Acidification of Guard Cell vacuoles during Stomatal Closure Require Phosphatidylinositol 3,5-Bisphosphate. *Plant Cell.* 2013;25:2202–16.
58. Krishnamurthy P, Vishal B, Bhal A, Kumar PP. WRKY9 transcription factor regulates cytochrome P450 genes CYP94B3 and CYP86B1, leading to increased root suberin and salt tolerance in Arabidopsis. *Physiol Plant.* 2021;172:1673–87.
59. Xie L-J, Chen Q-F, Chen M-X, Yu L-J, Huang L, Chen L, et al. Unsaturation of very-long-chain Ceramides protects Plant from Hypoxia-Induced damages by modulating Ethylene Signaling in Arabidopsis. *PLoS Genet.* 2015;11:e1005143.
60. Pih KT, Kabilan V, Lim JH, Kang SG, Piao HL, Jin JB, et al. Characterization of two new channel protein genes in Arabidopsis. *Mol Cells.* 1999;9:84–90.
61. Willige BC, Ahlers S, Zourelidou M, Barbosa ICR, Demarsy E, Trevisan M, et al. D6PK AGCVIII kinases are required for Auxin Transport and Phototropic Hypocotyl bending in Arabidopsis. *Plant Cell.* 2013;25:1674–88.
62. Jiang C-Z, Yen C-N, Cronin K, Mitchell D, Britt AB. UV- and Gamma-Radiation sensitive mutants of Arabidopsis thaliana. *Genetics.* 1997;147:1401–9.
63. Ogawa T, Pan L, Kawai-Yamada M, Yu L-H, Yamamura S, Koyama T, et al. Functional analysis of Arabidopsis Ethylene-Responsive element binding protein conferring resistance to Bax and abiotic stress-Induced Plant Cell Death. *Plant Physiol.* 2005;138:1436–45.
64. Rehfeldt GE, Jaquish BC, López-Upton J, Sáenz-Romero C, St Clair JB, Leites LP, et al. Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotsuga menziesii*: realized climate niches. *Ecol Manage.* 2014;324:126–37.
65. Campbell JL, Shinneman DJ. Potential influence of wildfire in modulating climate-induced forest redistribution in a central Rocky Mountain landscape. *Ecol Process.* 2017;6:7.
66. Zhao Y, O'Neill GA, Wang T. Predicting fundamental climate niches of forest trees based on species occurrence data. *Ecol Indic.* 2023;148:110072.
67. Corlett RT, Westcott DA. Will plant movements keep up with climate change? *Trends Ecol Evol.* 2013;28:482–8.
68. Périé C, de Blois S. Dominant forest tree species are potentially vulnerable to climate change over large portions of their range even at high latitudes. *PeerJ.* 2016;4:e2218.
69. Elsen PR, Saxon EC, Simmons BA, Ward M, Williams BA, Grantham HS, et al. Accelerated shifts in terrestrial life zones under rapid climate change. *Glob Chang Biol.* 2022;28:918–35.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.