How useful is genomic data for predicting maladaptation to future climate?

Brandon M. Lind[1,*], Rafael Candido-Ribeiro[1], Pooja Singh[2], Mengmeng Lu[2], Dragana Obreht Vidakovic[1], Tom R. Booker[3], Michael C. Whitlock[3], Sam Yeaman[2], Nathalie Isabel[4,5], Sally N. Aitken[1]

[1]Centre for Forest Conservation Genetics and Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada

[2]Department of Biological Sciences, University of Calgary, Calgary, AB, Canada

[3]Department of Zoology, University of British Columbia, Vancouver, BC, Canada

[4]Canada Research Chair in Forest Genomics, Centre for Forest Research and Institute for Systems and Integrative Biology, Université Laval, Québec, QC, Canada

[5]Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Québec, QC, Canada

10 February 2023

Running Title: *Validating genetic offsets*

Keywords: climate maladaptation, climate change, machine learning, genetic offset

[*]**Corresponding Author**

Brandon M. Lind – ORCID 0000-0002-8560-5417

430 Nahant Rd

Northeastern University Marine & Environmental Sciences

Nahant, MA 01908

Email: lind.brandon.m@gmail.com

Abstract

Methods using genomic information to forecast population maladaptation to climate change are becoming increasingly common, yet the lack of model validation poses serious hurdles toward their incorporation into management and policy. Here, we compare the validation of two methods – Gradient Forests (GF) and the Risk Of Non-Adaptedness – using exome capture pool-seq data from 35 to 39 populations across three conifer taxa: two Douglas-fir varieties and jack pine. We evaluate sensitivity of these algorithms to the number of input populations as well as the source and number of input loci (markers selected from genotype-environment associations [GEA] or those selected at random). We validate these methods against two-year and 52-year growth and mortality measured in independent transplant experiments. Overall, we find that both genetic offset methods often better predict transplant performance than climatic or geographic distances. We also find that while GF models are surprisingly not improved using GEA outliers, they are sensitive to the populations included in analysis. Even with promising validation results, ambiguity of model projections to future climates makes it difficult to identify the most maladapted populations using either method. Our work advances understanding of the sensitivity and applicability of these approaches, and we discuss recommendations for their future use.

1 | Introduction

Climate change poses unprecedented risk to global biodiversity loss[1,2]. Historically, the impacts of climate change on species abundance have been projected through species distribution modelling[3]. However, these methods often fail to account for the environmental drivers of local adaptation and the intrinsic mechanisms by which populations could respond to environmental change[4,5]. Recently, methods incorporating genomic information to forecast climate maladaptation have gained traction. Predominant among these, Gradient Forests (GF; *sensu* Ref[6]) and the Risk Of Non-Adaptedness[7] (RONA) use current relationships between genotype and climate to estimate offset (i.e., the mismatch between populations and their climate optima) under future climate scenarios (see Methods). If such offset methods were shown to be robust, they could circumvent the need for long-term field experiments and could be used to rapidly inform management priorities, or provide an option for species where experimentation is not feasible.

Despite their current popularity, genetic offset methods remain largely unvalidated with a few exceptions. For instance, Láruson et al.[8] used simulated data to train GF. They found that when 1) climate and genotypes are known without error, 2) all populations across the simulated landscape are locally adapted, and 3) validation is carried out within the climate space used in training, the predicted offset had a strong negative rank correlation with simulated fitness, and GF models trained using all markers

performed no better than GF models trained using causal markers. Further, they found that environmental distances calculated using environmental variables driving local adaptation also had a strong negative relationship with simulated fitness, though this was not the case when non-causal environments were included in distance calculations. Attempts to validate GF using empirical data where error is inherent have found weaker relationships between GF offset and common garden phenotypes[9].

In empirical settings, GF models are often used to project offset to areas of the species' range where no populations have been sampled[6,10–12] and have used disparate sets and sample sizes of both populations and loci to project future offset to changing climates. Projection to unsampled areas can lead to inaccuracies when models are poor, and the choice and number of loci may also impact model performance. Moreover, it is unclear whether more accessible forms of data (e.g., climate or geographic distance) perform as well as these genetically based methods in empirical settings. Validating these methods' predictions of maladaptation to future climate is challenging due to the temporal nature of such projections. However, transplant experiments (i.e., common gardens or provenance trials) can be used to quantify performance by correlating measurements of fitness-related phenotypes with the offset projected to the contemporary climate of the growing site[11].

Tree species are ideally suited to empirically validate predictions from offset methods because there is abundant evidence to suggest that many tree species are locally adapted to climate[13–16], a key underlying assumption of offset models[8,17,18]. Trees are also relevant systems for understanding maladaptation to future climate because of their ecological role

in terrestrial systems, as well as their capacity to sequester carbon. Historically, many forest tree species have experienced large geographic range shifts in response to changes in climate[19,20]. Yet, rates of projected climate change are likely to outpace maximum rates of historical migration for many of these species (e.g., Refs[19,21]) and therefore leave future outcomes largely unknown[22–25].

Here, we train offset methods using exome capture pool-seq data from three conifer taxa (Fig. 1) and validate results with phenotypes from independent transplant experiments at seedling (two-year) and adult (52-year) life stages. Using fitness-related phenotypes from juvenile life stages enables validation of projections for species where no longer-term phenotypic data exist (i.e., most species), while validation using phenotypes from adult life stages enables comparison of offset to more direct measures of total lifetime fitness. We use these datasets to address four main questions: Q1) How is performance of the offset method affected by the number and source of loci? Q2) How do genetically based offsets compare with non-genetic offset measures of climate and geographic distance? Q3) Will a greater number of training populations improve finer-scale predictions of offset? Q4) How well can GF predict offset in areas of the range not included in training?

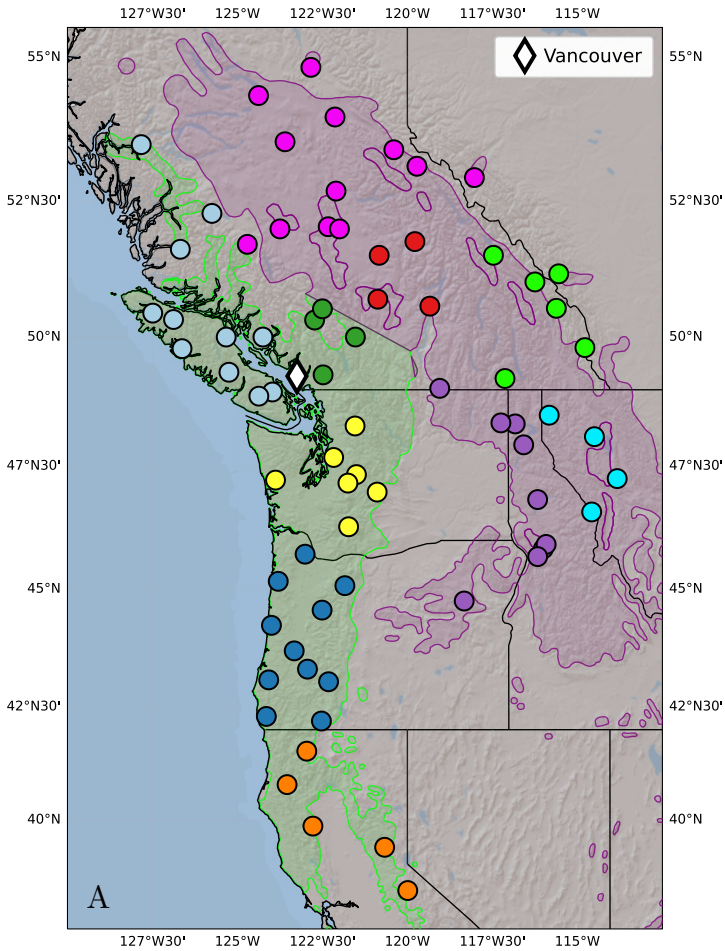
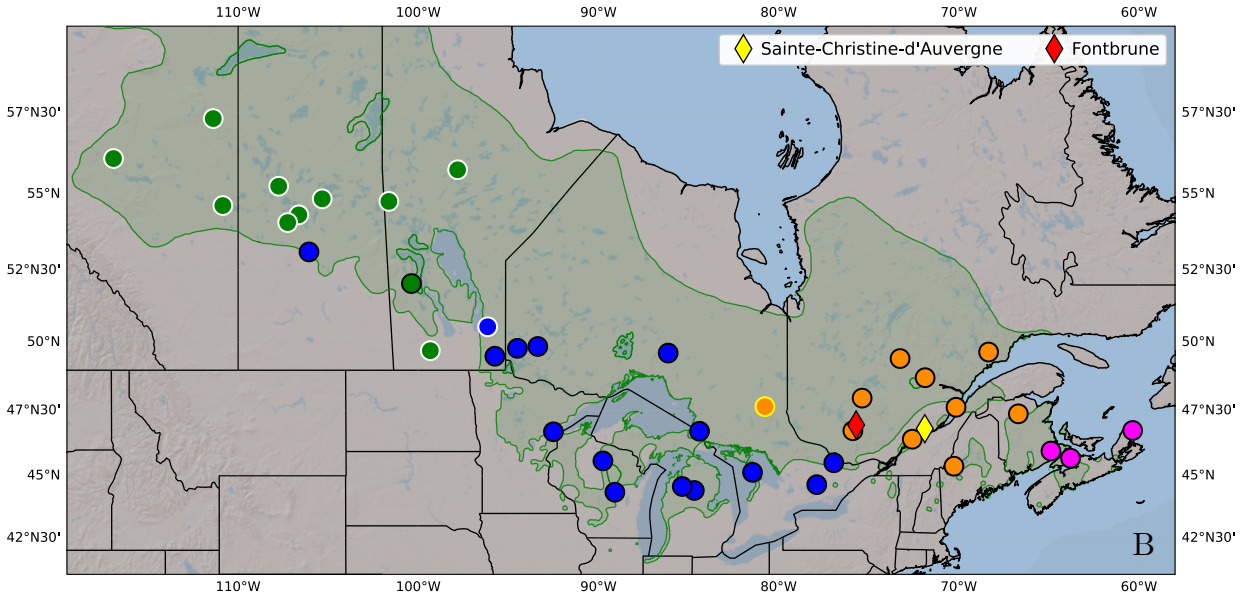


Fig. 1 North American source populations (circles) used for genomic and phenotypic data to train and validate both Gradient Forests (GF) and the Risk of Non-Adaptedness (RONA), and the common gardens used for offset validation: (A) Douglas-fir, and (B) jack pine. Shaded polygons are range maps for the full range of coastal Douglas-fir (lime, A), the northern and central range of interior Douglas-fir (purple, A), and the southern range of jack pine (green, B). All Douglas-fir populations were grown in the Vancouver common garden (blue diamond, A) and used for validation. Jack pine populations outlined in black were used for validation in both the Sainte-Christine (yellow diamond, B) and Fontbrune (red diamond, B) common gardens, while those outlined in yellow were only used for validation in Sainte-Christine, and those outlined in white were only used in model training but not validation. Color of population indicates the genetic groups over which stratified *k*-fold cross-validation was carried out in permutations of Gradient Forests. Code to generate these figures can be found in SN 15.99.



99

2 | Methods

Throughout this manuscript we will be referencing our code used to carry out specific analyses in-line with the text, most often in unstripped jupyter notebooks[26]. We refer to these notebooks as Supplemental Notebooks (SN) using a directory numbering system (e.g., SN 15.01). More information about the numbering system and archiving can be found in the Data Availability section.

2.1 Focal Species, Population Sampling, and Genetic Data

Three taxa of conifers across two species (Fig. 1) were used to assess the accuracy of genetic offset methods: 1) 38 range-wide populations of coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii* [Mirb.] Franco, *Pinaceae*), 2) 35 populations of interior Douglas-fir (*P. menziesii* var. *glauca*) from the variety's northern and central range, and 3) 39 jack pine (*Pinus banksiana* Lamb., *Pinaceae*) populations from across the species' southern range. We chose these species for their large and environmentally heterogeneous distributions, economic importance, and ecological relevance.

We used exome capture pool-seq data from these sampled populations. Briefly, we targeted exomic regions in DNA extracted from 33 to 40 individuals per population for Douglas-fir (39 Mb capture probe size), and 17 to 20 individuals per population for jack pine (41 Mb capture probe size) using exome capture probes described in Ref[27], where individuals within populations were pooled in equimolar quantities before sequencing. The sequencing depths used here exceed those used in Ref[27], as it was found that pool-seq

depth was one of the best predictors of the agreement of allele frequencies between sequence data for individuals and those from pool-seq data, despite generally strong agreement overall (Pearson's $r > 0.948$)[27].

Pool-seq libraries were sequenced in a 150bp paired-end format on an Illumina HiSeq4000 instrument at the Centre d'expertise et de Services Génome Québec, Montréal, Canada. We mapped reads from both varieties of Douglas-fir to the current reference genome of coastal Douglas-fir v1.01[28]. We mapped reads from jack pine to an amended version of its congener, loblolly pine[29–31] (*P. taeda* L., Pinaceae v2.01). In short, we used transcriptomic data from jack pine and amended non-mapping transcripts to the loblolly reference before mapping pool-seq data. Single nucleotide polymorphisms (SNPs) were called independently for Douglas-fir and jack pine as in Ref[27] using a VarScan pipeline[32] and filtered for missing data across populations $\leq 25\%$, minimum population allele depth ≥ 8, and global minor allele frequency ≥ 0.05. Here, we also created a fourth 'cross-variety' SNP set in which the same filtering process was carried out across both Douglas-fir varieties simultaneously (SN 02.01.01).

To address Q1, we use two methods for identifying genotype-environment association (GEA) outliers to ensure that genetic offset performance was not solely the outcome of the chosen method, as well as random-chosen sets of loci with numbers matching those of outlier sets to ensure that the number of loci was also not affecting the outcome. BayPass[33] is a single-locus genotype-environment association that evaluates support for each SNP independently for each environmental variable. We also use GEA results from Booker,

141 Singh et al.[34] that employed the Weighted Z Analysis[35] (WZA) using the same pool-seq
142 data. The WZA uses information across closely linked loci within genomic windows[34] (here
143 genic regions) to assess GEA support at the window level for a given environmental
144 variable. We performed GEA analyses using BayPass at the variety level for Douglas-fir
145 (see SN subfolder 02.02) and the species level for jack pine (see SN subfolder 07.02). For
146 BayPass, we identified all SNPs across all 19 climatic variables (Table S2) with Bayes
147 Factor (BF) in deciban units (dB) $\geq$ 15 following Jeffrey's rule indicating, at minimum,
148 very strong support (Extended Data Table 1). From the WZA output[34], we identified the
149 top 100 genes associated with each variable using p-values, and kept only those SNPs
150 from within these windows that had a Kendall's $\tau \geq 0.5$, calculated by correlating the
151 population-level allele frequencies with environmental values for each locus (Extended
152 Data Table 1). When using the 'cross-variety' SNPs filtered jointly across both Douglas-
153 fir varieties, we used the intersection of loci between 1) those that passed cross-variety
154 filtering and 2) those that were also GEA hits at the variety level (i.e., we did not perform
155 GEA across varieties). Hereafter, the BayPass and WZA marker sets are also referred to
156 more generally as 'outlier' marker sets.

157 Because we were interested in knowing how the input loci would affect genetic offset
158 methods (Q1), we also created a 'random' set of loci of equal sample size as each of the
159 two outlier sets by randomly choosing loci across our full datasets (SN 15.04). In total,
160 we generated four sets of SNPs for each of the three conifer taxa to use in training
161 (Extended Data Table 1). The comparison of models using random and outlier sets

addresses questions related to criteria of input loci and its impact on model performance, and comparison among models using the random sets of loci address questions related to the impact of the number of input loci to model performance. For the main text we present results using marker sets from BayPass, WZA, and the set of random loci with same sample size as WZA, and present all sets together within Supplemental Information. Note that for estimating RONA[7], we used a subset of each of these marker sets so that only loci with significant linear models were included in RONA calculations (see Section 2.3, Extended Data Table 1).

2.2 Training and Fitting Gradient Forests

Gradient Forests is a machine learning algorithm that incorporates Random Forest ensemble learning to predict genetic offset by using climate to split nodes of allele frequencies (from many loci) in a forest of decision trees, and uses this splitting information to construct monotonic turnover functions which are in turn aggregated and used to predict offset to future climate[6]. We used outlier and random marker sets (Section 2.1) to train GF (SN 15.04). For each marker set, we first created training sets that included all available populations, as is often done in the literature (e.g., Refs[11,12,36,37]). To address Q4, we also employed *k*-fold cross-validation to explore model sensitivity to input populations and the extent to which offset projected to areas of the range with no training populations can remain reliable (Supplemental Text S1.1-S1.2; Figs. S1-S2).

The climate data used in training included climate normals from 19 climatic environmental variables between the years 1961-1990, downloaded from AdaptWest.com on Februrary 5, 2021[38]; AdaptWest data is generated using ClimateNA[39]. These climate variables include those related to annual temperature (MAT, MWMT, MCMT, TD), 30-year minimum (EMT) and maximum (EXT) temperature extremes, annual precipitation (MAP, MSP, AHM, SHM, Eref, CMD), and the seasonality of both temperature (DD0, DD5, NFFD, FFP, bFFP, eFFP) and precipitation (PAS; see Table S2 for climatic abbreviations and units). We also included elevation. These variables were selected *a priori* based upon relevance to the species' biology and environmental variation across the species' ranges.

After clipping AdaptWest climate data (SN 15.03) to our species ranges (SN 15.02) using customized shapefiles from the United States Geological Survey[40], training sets ($N = 76$) were fed into customized training scripts (SN 15.05) to facilitate parallelization on Compute Canada's high-performance computing clusters (CCHPC; SN 15.04 section 5).

Each trained model from GF was fit to the climate of one (Douglas-fir) or two (jack pine) common gardens (SN 15.07) using the fitting script created in SN 15.05. The climate data used for fitting (SN 15.06) was the average climate (obtained from ClimateNA GUI between July 2-9 2021, Wang et al., 2016; Table S2), and elevation, of the common garden over the years in which the individuals were grown (see Section 2.5), and was treated as the 'future' climate of each population in offset projections.

201 An added utility of GF is that it can identify climatic variables driving patterns in
202 genetic data, without the need to project offset to future climates, particularly if offset is
203 not the primary goal. Gradient Forests outputs ranked environmental importance after
204 being trained. This has shown promise in identifying environmental drivers underlying
205 selection when using simulated data[8], even when there are multiple correlated
206 environmental variables. Using the outlier and random marker sets, we explore the
207 consistency of environmental importance ranks and how this differs among the training
208 runs that used all populations as well as those runs that used *k*-fold sampling of
209 populations. We also explore the consistency of environmental importance ranks between
210 outlier and random marker sets that used all populations in training (SN 15.13). We found
211 that GF was relatively insensitive to marker and population input with regard to
212 environmental importance (Supplemental Text S1.3; Figs. S3-S7).

213 **2.3 Estimating the Risk of Non-Adaptedness (RONA)**

214 In addition to GF, we also used RONA[7] to estimate genetic offset. The offset estimated
215 by RONA relies on linear relationships between allele frequencies for outlier loci and
216 climatic variables. This estimation is carried out in four steps: 1) identifying outlier loci
217 putatively underlying adaptation to the environment (e.g., from GEA), 2) subsetting this
218 list for loci that also have significant linear models relating allele frequency with
219 environmental variables, 3) using the current model of the linear relationship between
220 population allele frequencies and environment to estimate the allele frequency for a single

population in a new environmental (e.g., a value from projected climate change or a common garden), and 4) averaging the absolute difference between current and estimated future allele frequencies across loci for a given population for a given climatic variable (see equation and Fig. 2 on p. 5913 of Ref[7]).

Using the four marker sets described in Section 2.1 (two outlier sets and two random sets), we isolated loci with significant linear models relating current allele frequencies to climate variables (the same climate data used in GF training in Section 2.2; Table S2), then calculated RONA for each population for each variable (SN 15.09) using average environmental values for the years individuals were grown in the gardens (Section 2.5). Because RONA is calculated for a specific population and environmental variable, there is a range of RONA estimates for any given population, and thus the choice of environmental variables to consider for offset estimation could impact inferences regarding population performance. To address this, Rellstab et al.[7] used paired t-tests to determine which future environments were most different from their current state ($n = 5$), taking the top three variables after ranking p-values to use in estimating the range of RONA. For our validation, the vector containing future environments (i.e., common gardens) would be constant for a given variable across populations. In the context of a paired t-test, this is somewhat intractable with the test's null hypotheses that each vector in the pair is sampled from the same distribution, which could lead to biologically meaningless (yet statistically significant) inference. We explored groups of environmental variables (see next section) related to 'expert choice' or those used in guiding seed sourcing in

British Columbia. However, the top 5 environments from the canonical paired t-test as described above produced more accurate results than any of the other groups of environments (not shown, except in SN 15.09 section 8), and so we present the range of RONA using these top five environments from ranked t-test p-values. Because the top environments isolated in this way are often highly correlated, these top environments thus allow for the effective estimation of one RONA offset value.

2.4 Non-genetic Offset Measures

Could environmental data alone be used instead of genetic data for management decisions[41]? To address this question related to Q2, we also estimated population offset by calculating geographic and climatic distances from the source populations to the common garden (SN 15.08). To calculate geographic distance, we use the latitude and longitude of each population and garden to calculate distance via Vincenty's geodesic. To calculate climatic distance, we use the Mahalanobis distance for each population centered on the common garden using the same climate data in training and fitting of GF and RONA (Sections 2.2 and 2.3; Table S2). We explored three sets of environmental variables to estimate climate distance: 1) all geoclimatic variables (Table S2), 2) those climate variables used in climate-based seed transfer (CBST) guidelines for British Columbia[42] – mean annual temperature (MAT), mean coldest month temperature (MCMT), continentality (TD), mean annual precipitation (MAP), degree-days above 5°C (DD5), extreme minimum temperature (EMT), and 3) climate variables identified in previous and

independent reciprocal transplants not used here. For jack pine, we used the two climate variables from the transfer function used to best predict height of a sister species with which it readily hybridizes, lodgepole pine (*Pinus contorta* subsp. *latifolia* Douglas, *Pinaceae*)[43]: MAT (>64% variance explained) and annual heat-moisture index (AHM; where ln(AHM) explained >6% variance). For Douglas-fir, we used three variables found to be significant predictors in universal response functions of height and basal diameter for a large multiple common garden trial of North American populations planted in Central Europe[44]: MAT, summer heat-moisture index (SHM), and TD.

2.5 Common Garden Data

Measurements of fitness-related phenotypes from common gardens (diamonds, Fig. 1) used to validate genetic offset predictions were obtained by phenotyping individuals from the same populations sampled to obtain genetic data. For jack pine, we measured 52-year adult phenotypes – height, diameter at breast height (DBH), mortality – grown in a field provenance trial at two independent sites, Fontbrune (LAT 46.959, LONG -75.698) and Sainte-Christine-d'Auvergne (LAT 46.819, LONG -71.888), between 1966 and 2018. For Douglas-fir, we measured two-year seedling phenotypes – shoot biomass and height increment – grown in a Vancouver common garden (LAT 49.257, LONG -123.250) between 2018-2019[45]. For each common garden, we used the population mean phenotype to validate genetic offset (Section 2.6). For more information about phenotypic measurements, see Supplemental Text S1.4.

282 **2.6 Validating Offset Measures**

283 Population mean phenotypes (Section 2.5) were used as a proxy for fitness by which
284 to validate the genetic offsets predicted from GF (SN 15.11) and RONA (SN 15.09), by
285 correlating population mean phenotype with population offset, using Spearman's ρ as a
286 validation score (Supplemental Text S1.5). We expect a negative relationship between
287 offset and growth, and a positive relationship between offset and mortality. For GF models
288 trained using all populations, we used all available offset and phenotypes to calculate the
289 validation score. For GF models trained using a subset of populations from k-fold stratified
290 sampling, we used the offset and phenotypes from only those out-of-bag populations that
291 were excluded from training. We evaluate sensitivity of GF models to population input
292 using these out-of-bag validation scores and by plotting the relationship between the union
293 of out-of-bag offset scores with offset scores from the model trained using all populations.
294 We validated RONA for each environmental variable that ranked within the top five
295 environments that differed significantly (via t-test p-values) between the common garden
296 and climates used in training, calculating a validation score for each climate variable.

297 To determine if models were improved at fine-spatial scales by including a greater
298 number of populations in training (Q3), we leveraged genetic structure within and across
299 the two varieties of Douglas-fir. These two varieties (coastal and interior; lime and purple
300 ranges, respectively, in Fig. 1A) diverged ~2.11 Ma[46] and differ substantially both
301 morphologically and ecologically. While the coastal variety shows little genetic grouping
302 in PCA and instead differentiates along a latitudinal cline, populations in the northern

range of interior Douglas-fir populations form two distinct genetic groups (Fig. S1). This allowed us address Q3 by calculating our validation score across various levels of genetic hierarchy: 1) across both varieties (i.e., the cross-variety model), 2) across all interior variety populations, and 3) across each of the northwestern and southeastern interior Douglas-fir genetic subgroups (see Fig. S1). We evaluate these hierarchical scenarios using the GF models trained across both varieties as well as that trained using solely the interior variety (SN 15.11). We did not carry out an analogous assessment with RONA because of the effects of using fewer populations in the steps necessary to isolate loci with significant linear models.

2.7 Projecting genetic offset to future climates

As in Section 2.2, we downloaded future climate scenarios from AdaptWest.com[38,39]. Using the same environmental variables using in training, fit models trained using WZA loci from GF (SN 15.07) or RONA (SN 15.16) to project future genetic offsets. For future climates we used Representative Concentration Pathway (RCP) greenhouse concentration trajectories projected to the 2050s and 2080s: RCP4.5 2050s, RCP4.5 2080s, RCP8.5 2050s, and RCP8.5 2080s. RCP4.5 and RCP8.5 each represent radiative forcing units ($^{W}/_{m^2}$) and are, respectively, an intermediate scenario where emissions peak in the 2040s and then decline, or continue to rise throughout the 21st Century. As with estimating RONA using common gardens (Section 2.3), we identified the five environments most differentiated between current and future climates using paired *t*-tests and present these

estimates. We report patterns from RCP8.5 2050s in the main text, including Spearman's ρ between RONA estimates (SN 15.16), GF estimates (SN 15.18), and between RONA and GF (SN 15.16).

3 | Results

3.1 | Validation of offset with fitness-related phenotypes

3.1.1 / Jack pine

The performance of both GF and RONA differed between the two jack pine provenance trials validated using 52-year phenotypes of mean DBH, mean height, and mortality (Fig. 2). Mortality was better predicted than DBH and height. Genetic offset predictions of 52-year mortality were not demonstrably better than the best non-genetic offset measures at either location (Q2). Importantly, using loci chosen by GEA analysis did not improve predictive ability over randomly chosen loci for GF (Q1; Fig. 2). For RONA the validation scores from GEA sets tended to have similar scores as random loci when estimating DBH and height, but scores from the two sets became more differentiated when estimating mortality (Fig. 2, Fig. S8).

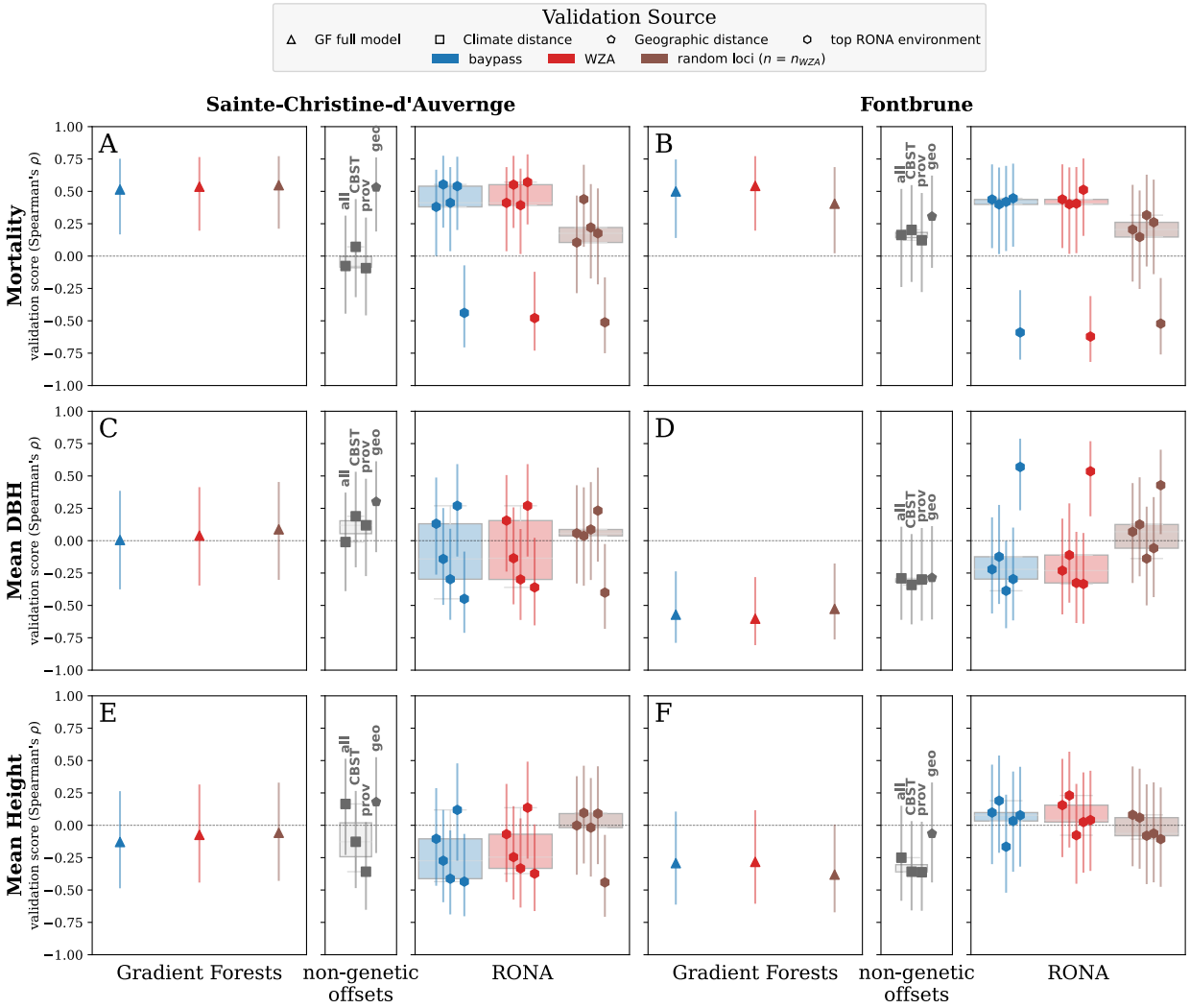


Fig. 2 Offset validation from 52-year jack pine phenotypes at Sainte-Christine-d'Auvergne (A, C, E) and Fontbrune (B, D, F) provenance trials using Gradient Forests, the Risk Of Non-Adaptedness (RONA), and climate and geographic distances. Triangles indicate performance of GF models trained and validated using all available populations. RONA background boxplots are indicative of the range of RONA estimates given for the top five environmental variables (hexagons) that differed significantly between source and common garden variables (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. Vertical bars indicate standard error estimated using a Fisher transformation (see Supplemental Text S1.3). Loci used in RONA calculations are a subset of those used in Gradient Forests that had significant linear models with the environment, see Extended Data Table 1 for loci counts. See Fig. S8 for all locus groups. Boxplot whiskers extend up to 1.5x the interquartile range. Code to create these figures can be found in SN 15.14.

The best non-genetic offset measure varied by phenotype and site, where validation scores among these metrics were not substantially differentiated (Fig. 2). In contrast to estimates for mortality (Fig. 2A-B), climate distance tended to perform better than geographic distance for DBH and height (Fig. 2C-F), but the set of climate variables used to calculate the best distance varied, and only once exceeded the scores from the full GF models (Fig. 2E).

3.1.2 | Douglas-fir

Both GF and RONA models substantially outperformed climate and geographic distance metrics for Douglas-fir (Q2, Fig. 3, Extended Data Fig. 1). As with jack pine, genetic offsets estimated using random loci performed equally well as GEA sets (Q1). However, fine-scale validation of GF and RONA was not substantially improved by models including more populations (Q3, Figs. S9-S10). The GF and RONA models that were trained and validated across both varieties of Douglas-fir had the greatest validation scores across all comparisons (Fig. 3A, Extended Data Fig. 1A), achieving much higher performance than in jack pine (Fig. 2). However, when models were trained and validated for each variety separately the relative performance decreased (Fig. 3B-C, Extended Data Fig. 1B-C). The stronger validation score from the cross-variety model validated using both varieties (e.g., Fig. 3A) compared to the scores validated within varieties is likely driven by the substantial genetic structure of the two varieties, as varieties are distinct when plotting cross-variety offset vs. phenotype (Extended Data Fig. 2).

Because the cross-variety validation statistic was inflated due to structure in the data, and because the scale of management is often applied on finer spatial scales, we were interested in how groups of Douglas-fir populations (i.e., varieties or genetic groups) would validate, and if there was any information gained from training across a greater number of populations than are used for validation (Q3). Comparing models, the cross-variety model validated using only variety-specific populations was not substantially different

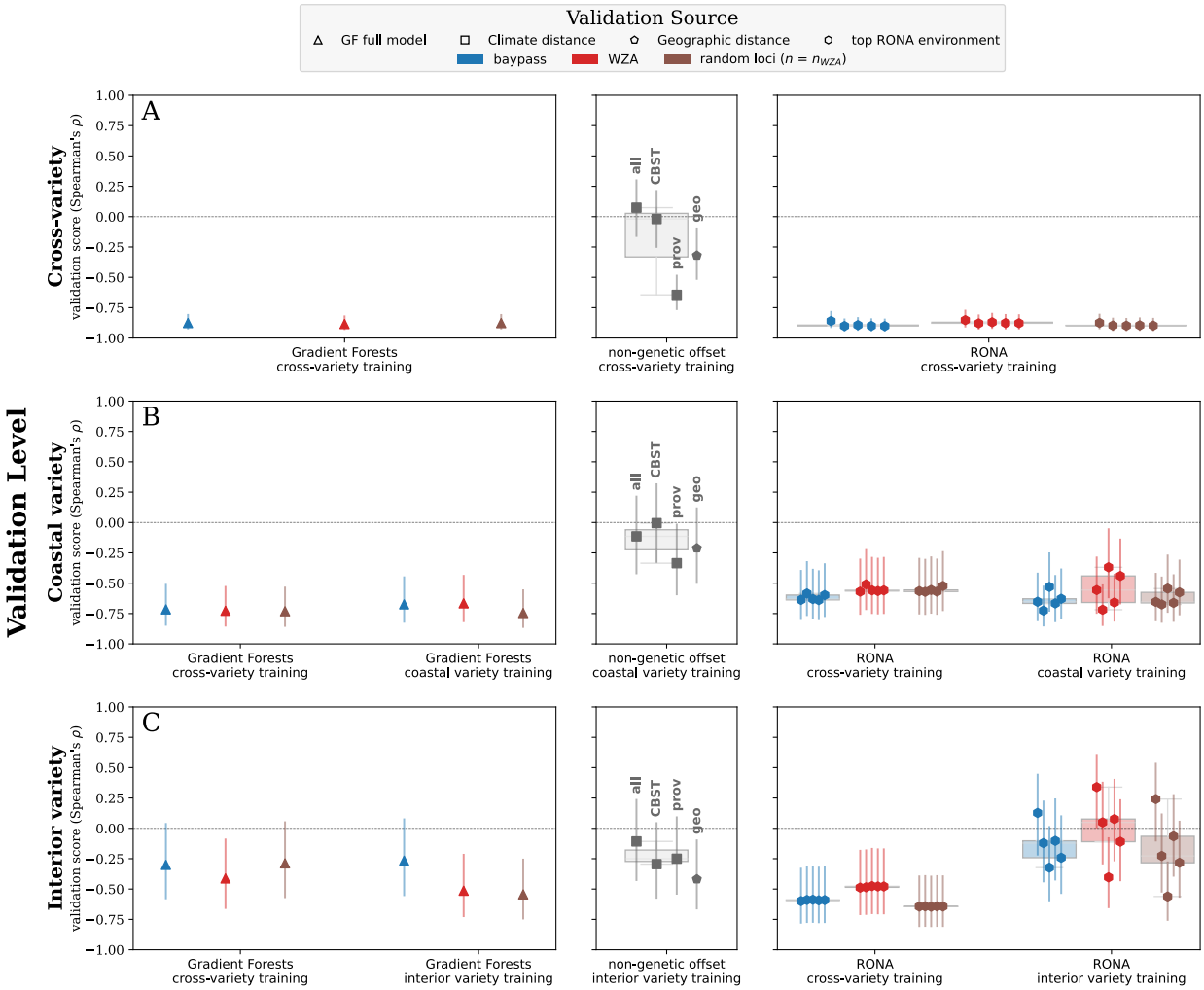


Fig. 3 Offset validation from two-year Douglas-fir height increment phenotypes at the Vancouver common garden (see Fig. 1A) using Gradient Forests (GF), the Risk of Non-Adaptedness (RONA), and climate and geographic distances. We assessed accuracy inference from trained models (x-axis groups) using populations (rows) across both varieties of Douglas-fir (A), at the variety level for the coastal (B) and interior varieties of Douglas-fir (C) to determine if greater numbers of training populations improve finer-scale predictions of offset. Genetic offset boxplots and shapes are shaded with respect to marker set source. Triangles indicate performance of GF models trained and validated using all available populations. RONA background boxplots are indicative of the range of RONA estimates given for the top five climatic variables (hexagons) that differed significantly between source population and the common garden (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. Vertical bars indicate standard error estimated using a Fisher transformation (see Supplemental Text S1.3). Loci used in RONA calculations are a subset of those used in Gradient Forests that had significant linear models with the environment, see Extended Data Table 1 for locus counts. See Extended Data Fig. 1 for similar validation using shoot biomass. See Fig. S9 for all locus groups. Boxplot whiskers extend up to 1.5x the interquartile range. Code to create these figures can be found in SN 15.14.

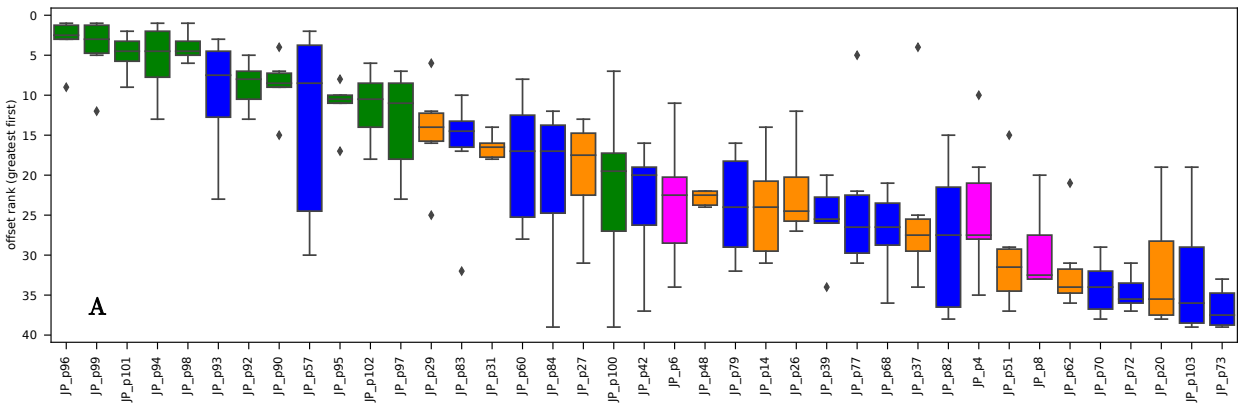
from models that were both trained and validated at the variety level (Fig. 3B-C, Extended Data Fig. 1B-C). Comparing the two varieties, the coastal variety models had greater validation scores than models for the interior variety (Fig. 3B-C, Extended Data Fig. 1B-C). Coastal variety genetic offsets often performed better than non-genetic offset measures, but genetic and non-genetic offsets performed similarly for the interior variety (Q2, center panels Fig. 3, Extended Data Fig. 1). To further explore impacts on the accuracy of fine-scale offset related to Q3, we subset populations from the interior variety into two distinct genetic groups to validate predictions from the GF cross-variety and interior-only models. We found similar patterns of accuracy between fine-scale validation of the cross-variety and interior-only genetic offset models, though fine-scale validation indicated stronger relationships between offset and performance within these genetic groups than at the variety level (Supplemental Text S1.6).

Validation scores from climate distance using variables inferred as important from independent provenance trials were often stronger than the other climate distance measures (Fig. 3, Extended Data Fig. 1), while validation scores from geographic distance were stronger than climate distance only in the validation of the interior Douglas-fir populations (Figs. S9-S10).

3.2 | Predicted genetic offset to future climates

3.2.1 / Jack pine

Maladaptation of jack pine populations to future climate (RCP8.5 2050s) inferred from GF and RONA models trained using WZA loci and all populations indicate that the western-most group (green populations, Fig. 1B) relative to all other populations are likely to experience the greatest maladaptive effects from changing climates (Fig. S11B-C). Four populations (JP_p96, JP_p99, JP_p98, and JP_p101) have consistently high



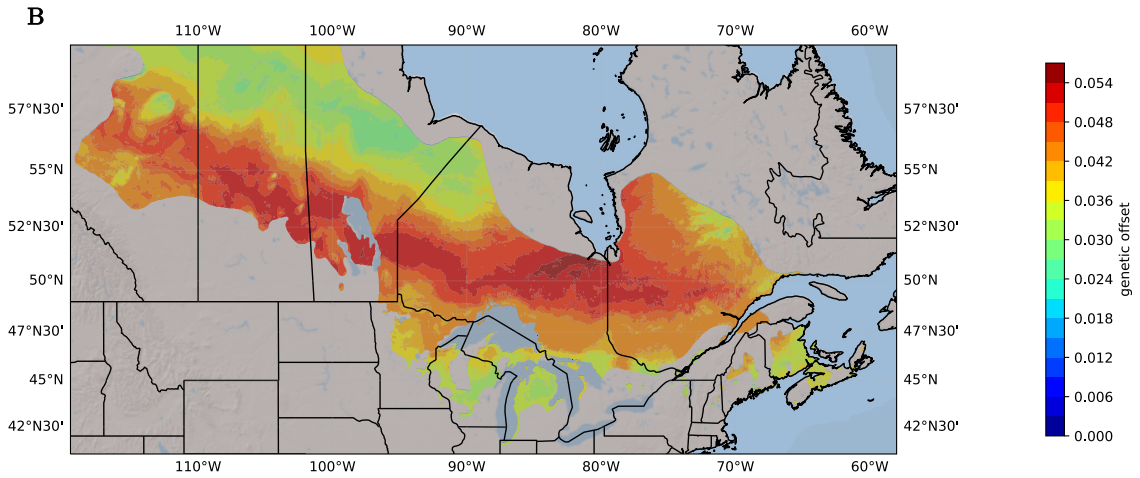


Fig. 4 Maladaptation of jack pine populations to future climate (RCP8.5 2050s) inferred from RONA and Gradient Forests (GF). A) distribution of offset rank for each jack pine population from RONA and GF (note all but one of the values that create each boxplot are from RONA) and are ordered from left to right by decreasing median offset rank. B) predicted offset from GF interpolated across the species' range. Environments used to estimate RONA were chosen based on ranking p-values from paired *t*-tests between current and future climate. Code to create these figures can be found in SN 15.17 and SN 15.18. Analogous figures created using climate models RCP4.5 2080s, RCP4.5 2050s, and RCP8.5 2080s show similar patterns and are not shown except within SN 15.17 and SN 15.18. To see populations overlayed onto (B), see Fig. S12.

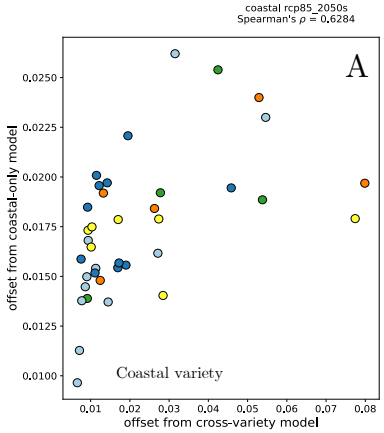
maladaptive rank across both GF and RONA with all ranks in the top 10 (Fig. 4A). From the projection of GF to areas of the jack pine range with no training data, it would seem that the central portion of the range will be similarly maladapted to future climate (red contours, Fig. 4B). Across the five environmental variables used to estimate RONA for this climate scenario (which were highly correlated, Fig. S11E-F), the predicted maladaptive rank from RONA was positively correlated with offset estimated using GF (Fig. S11C-D).

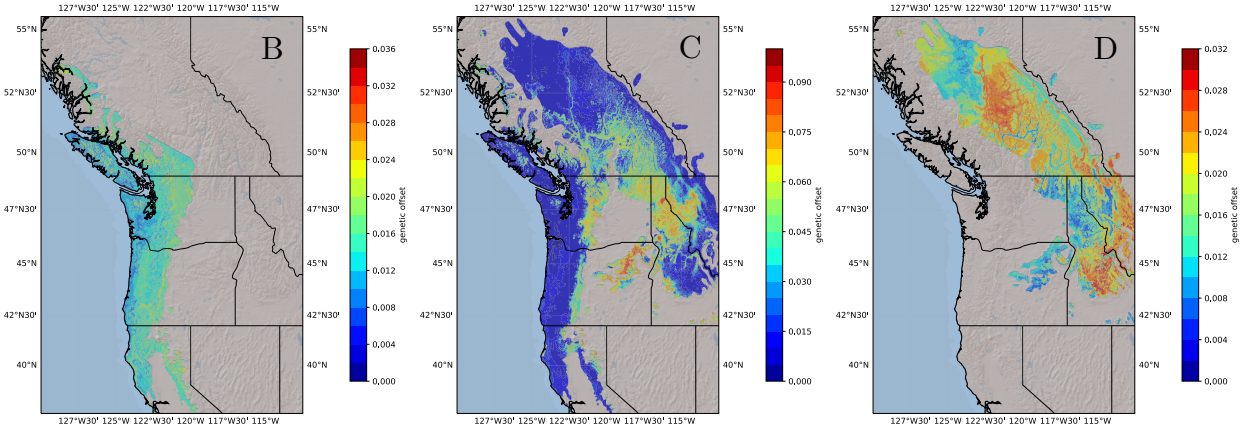
3.2.2 / Douglas-fir

Gradient Forest models fit to future climates (RCP8.5 2050s) using WZA loci gave inconsistent results as to which set of Douglas-fir populations were projected to be most

maladapted to new climates (Fig. 5). For the coastal variety, the cross-variety model and the coastal-only model of GF each identified the same two populations from coastal BC to be the least maladapted, but rank changed considerably among the remaining populations (Fig. 5A). For the interior variety, the cross-variety and the interior-only models were in conflict as to whether the northwestern genetic group (Fig. S1) would be more maladapted than the southeastern genetic group (compare Fig. 5C and Fig. 5D) whereas this was not the case when projecting offset to the common garden (Supplemental Text S1.7; Figs. S13-S15). For the northwestern interior genetic group, results from the cross-variety and interior-only models were generally similar, except that the population identified as the least maladapted with the cross-variety model was the most maladapted from the interior-only model (Fig. 5E). For the southeastern interior genetic group, there was a negative relationship between offset predicted by the two models (Fig. 5F).

The most maladapted interior Douglas-fir genetic group predicted from RONA was also inconsistent between the cross-variety and interior-specific models (Fig. S13B). However, RONA predictions were generally positively correlated for the interior variety and cross-variety models for the two interior genetic groups (Fig. S13C-D). Predictions from the cross-variety and coastal-only RONA models generally had positive, albeit relative weak, relationships (Fig. S13A).


coastal rcp85_2050s
Spearman's ρ = 0.6284
A
Coastal variety
offset from coastal-only model
offset from cross-variety model


B
C
D
genetic offset

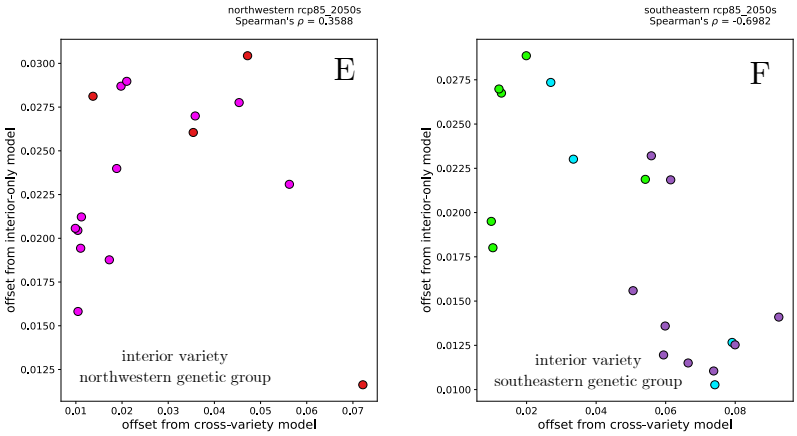

northwestern rcp85_2050s
Spearman's ρ = 0.3588
E
interior variety
northwestern genetic group
southeastern rcp85_2050s
Spearman's ρ = -0.6982
F
interior variety
southeastern genetic group
offset from interior-only model
offset from cross-variety model

Fig. 5 Maladaptation (offset) rank of Douglas-fir populations to future climate (RCP8.5 2050s) inferred from Gradient Forests is inconsistent between models trained using both varieties with those trained on a variety-specific basis. Shown are projected offsets to the range of Douglas-fir trained using WZA outliers and all populations from B) the coastal variety, C) both varieties, and D) the interior variety. For coastal Douglas-fir (A) and the two subvariety genetic groups of interior Douglas-fir (E, F), the relationship between the magnitude and rank of projected offset using the cross-variety model (pentagons, y-axes) is contrasted to those from the variety-specific model (squares, x-axes). Of note, the cross-variety model (C) and the interior-only model (D) indicate different interior variety genetic groups (populations in E or F) to be most maladapted to projected climate. Red lines within slope graphs are indicative of negative changes of rank. Populations are colored with respect to Fig. 1. Color legend is not standardized across B, C, and D to accentuate patterns in the data (offset values are meaningless outside of the current model). Code used to create these figures can be found in SN 15.18. Analogous figures created using climate models RCP4.5 2080s, RCP4.5 2050s, and RCP8.5 2080s show similar patterns and are not shown except within SN 15.18. To see populations overlayed onto B-D, see Fig. S12.

To select models for projecting offsets to future climate for Douglas-fir, we used four criteria when comparing cross-variety and variety-specific models: 1) validation scores, 2) validation model sensitivity, 3) agreement between future offsets from GF and RONA, and 4) agreement among RONA future offsets (Supplemental Text S1.8; Figs. S16-S21). Based on these criteria we use the cross-variety models to project maladaptation to future climate (RCP8.5 2050s; Fig. 6). For coastal Douglas-fir, many populations found along the Pacific Coast of California (orange) and Oregon (blue) had the greatest projected maladaptation (Fig. 6A). Populations from northwestern interior Douglas-fir near the Fraser River had consistently high offset ranks (Fig. 6B), whereas the remaining populations had a wide range of projected risks, and it is unclear which would be most affected by future climate. Finally, populations of southeastern interior Douglas-fir found in the Idaho, Montana, and eastern Washington and Oregon had consistently greater

predicted maladaptation to future climate than those found in Southeastern British Columbia (Fig. 6B).

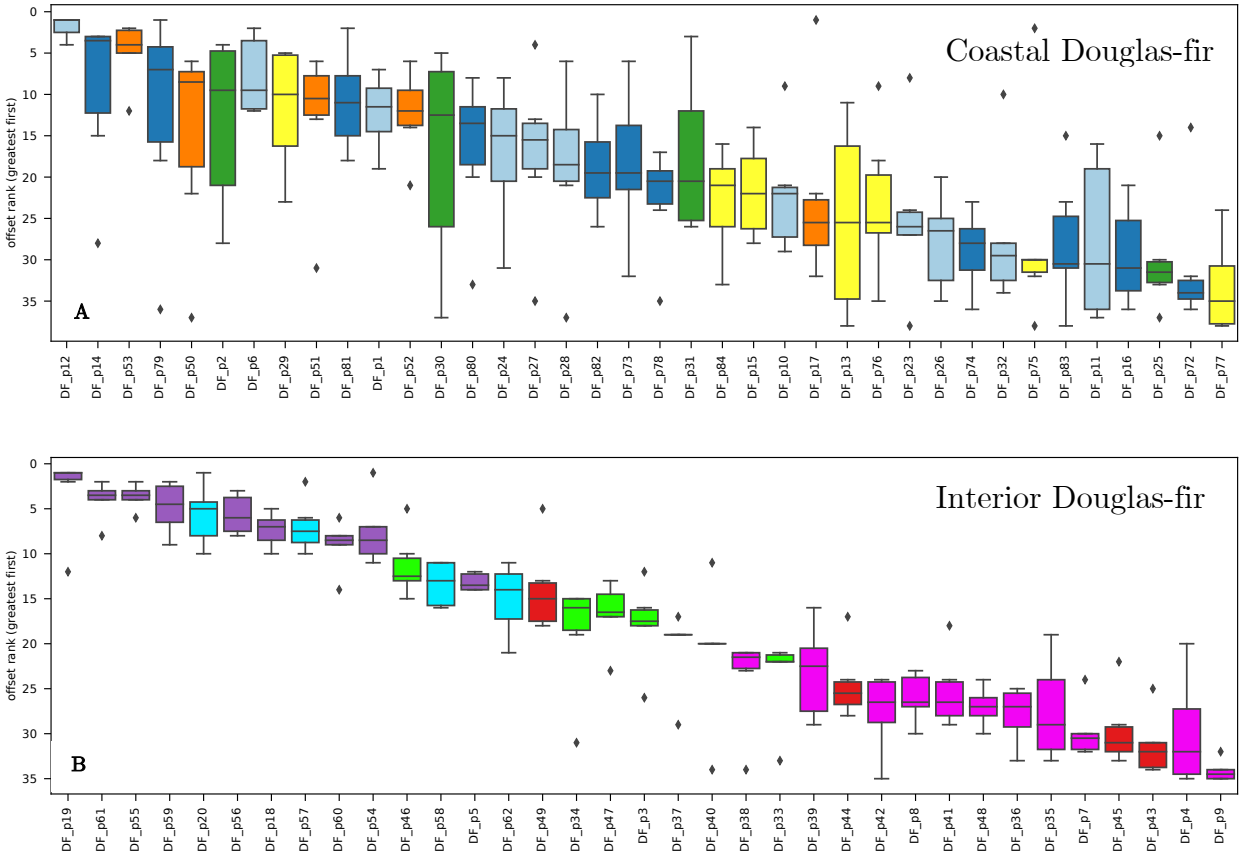


Fig. 6 Maladaptation of interior Douglas-fir to future climate (RCP8.5 2050s) inferred from RONA and Gradient Forests (GF) cross-variety models. Shown are the distributions of offset rank for each coastal (A) and interior Douglas-fir (B) population from RONA and GF and are ordered from left to right by decreasing median offset rank (note all but one of the values that create each boxplot are from RONA). Environments used to estimate RONA were chosen based on ranking p-values from paired *t*-tests between current and future climate. Populations are colored as in Fig. 1. For boxplots similar to A and B for all populations in a model see Figure S22. Code to create these figures can be found in SN 15.17. Analogous figures created using climate models RCP4.5 2080s, RCP4.5 2050s, and RCP8.5 2080s are not shown except within SN 15.17.

446 # 4 | Discussion

447 Projections of maladaptation of populations to climate change, i.e., genetic offset
448 estimates, have remained largely unvalidated despite the recent increase in their use. Here
449 use three taxa of conifers, four genomic marker sets, and common garden phenotypes from
450 independent two-year and 52-year individuals to demonstrate that genetic offset methods
451 perform as well or better than the best climate or geographic distance metrics when
452 predicting fitness-related phenotypes in transplant experiments (Q2, Fig. 2-3, Extended
453 Data Fig. 1). We further demonstrate that outlier sets provide little advantage over
454 random sets of loci, and that the number of loci used here have little impact on genetic
455 offset performance (Q1). We also show that fine-scale maladaptation is not better
456 predicted with models using a greater number of populations (Q3, Fig. S9-S10). GF models
457 are sometimes sensitive to input populations (Figs. S9-S10, S16-S20) and should not be
458 extrapolated to geographic areas where no training data exists (Q4). Lastly, we reveal
459 that when using future climate to fit these models, the Douglas-fir populations inferred to
460 be most maladapted depends on the model used (compare within and across Fig. 5, Fig.
461 6, and Figs. S13, S21). However, RONA and GF results largely agree when projecting
462 jack pine offset to future climates (Fig. S11). This highlights the fact that in the absence
463 of validation data, and without further knowledge of the behavior and sensitivity of these
464 genetic offset methods under a wider range of scenarios, it may be difficult to determine
465 whether a given set of populations can lead to reliable inference about maladaptation.

466 Together, these results suggest that acting on projections of maladaptation from genetic
467 offset methods through changes to policy and management practices should be considered
468 only after careful scrutiny of model performance. These findings also highlight the large
469 knowledge gap with respect to the ideal population and dataset parameters needed to
470 produce reliable genetic offset models.

471 Our results have several important implications for the use of offset models for
472 populations of any plant or animal species. First, our results suggest that genetic data
473 often contains more information regarding climate adaptation than can be expressed with
474 more readily accessible forms of data, providing evidence that climate differentiation alone
475 will not fully reveal the extent of maladaptation of populations to future climate change.

476 Second, the models trained using random markers performed no worse than those from
477 models using GEA loci (though there are minor exceptions for random sets used for
478 RONA, Figs. 2-3 and Extended Data Fig. 1). This suggests it may be unnecessary to
479 expend resources to identify adaptive regions when genome- or exome-wide data exist, a
480 finding that closely mirrors simulation results for GF[8]. The similar performance among
481 marker sets is perhaps due to the nature of our exome-targeted sequence data which
482 targeted functionally relevant coding regions. It remains to be seen if relatively inexpensive
483 sequencing techniques such as RAD-seq, which more often tags intergenic regions of large
484 genomes[47], would perform as well as the random marker sets used here. Even so, for
485 species with strong local adaptation where isolation-by-environment drives spatial genetic
486 structure, signals from genotyping-by-sequencing markers may contain sufficient

information for accurate offset projection and may therefore be a cost-effective alternative to the exome capture data used here.

Third, the contrast in the performance of these offset measures across the two jack pine provenance trial sites highlights the value of using multiple sources of validation in future work. Future studies will require validation to provide any degree of confidence in informing population- or site-specific management decisions. They will need to consider the extent to which the phenotypes and life stage used in validation are associated with total lifetime fitness, as well as how the common garden environment interacts with these phenotypes. For instance, while jack pine 52-year DBH may capture elements of fitness related to growth, it may miss aspects of fitness more directly related to survival and reproduction. It may also be more indicative of competitive ability in the planted common garden environment than fitness in the wild. Further, the location within climate space of the transplant experiment used in validation relative to that of training populations may also affect validation inference. In our study for both coastal Douglas-fir and jack pine, the transplant sites were at the periphery or within the climate space encompassed by the training populations, whereas the Vancouver common garden was well outside the climate space of the interior Douglas-fir populations (Fig. S23). Importantly, this observation may indicate that projections of maladaptation to future climates that differ greatly from climate data used in training may produce less robust estimates.

Finally, while model exploration will be necessary for future work[17], our results highlight the sensitivity of GF models to input populations, and the utility of stratified

508 *k*-fold cross-validation. Cross-validation will not only ensure that training and test sets
509 are similar but will also help reveal the extent to which projections to geographic areas
510 with no training data may produce reliable inferences. In many cases, the cross-validation
511 scores revealed relationships between offset and fitness that were of the opposite sign
512 expected (Figs. S8, S9), and further revealed sensitivity of population input to GF models
513 (Figs. S19, S20, S18, S16, S17). Our implementation of cross-validation represents a best-
514 case scenario because of the close geographic proximity of training and test populations.
515 Future validation of offset methods could also leave out entire genetic groups from
516 training, or instead group populations by source climate, to test yet another avenue for
517 understanding model sensitivity. Additionally, such exploration could be used alongside
518 other information to justify the choice among models that, for example, have similar
519 validation but differ in sensitivity, as we did for Douglas-fir.

520 Even with some promising results here, genetic offset estimates should be used with
521 caution to guide management decisions, as there are several circumstances under which
522 these estimates may be misleading with respect to true population maladaptation even
523 under otherwise ideal circumstances (e.g., in the presence of local adaptation).
524 Importantly, the most important factors for training and validating genetic offset models
525 are: (1) accurately estimating both climate data and allele frequencies for an adequate
526 number of populations; and (2) our ability to measure population phenotypes relevant to
527 total lifetime fitness. However, climate data is often interpolated across heterogenous
528 environments where weather stations are scarce. Further, the accurate, long-term

529 phenotyping of experiments from a sufficiently large number of individuals per population
530 for validation is resource- and time-intensive. With this first point in mind, we trained
531 additional models using climate normals from different periods and applied stricter
532 filtering of our genetic data but came to very similar conclusions (B. Lind unpublished
533 data).

534 In addition to data requirements necessary for accurate genetic offset predictions, not
535 all species (or groups of populations) are ideally suited for these models. These models
536 assume all populations are locally adapted to recent climates, that current genotype-
537 climate relationships are due solely to local adaptation and will remain optimal in the
538 future, and that deviations from these relationships will result in decreased fitness[17,18].
539 These models also ignore other dynamics that could alleviate or exacerbate maladaptation
540 experienced by future populations, such as gene flow (and perhaps subsequent swamping)
541 of adaptive alleles, release or onset of competition or disease, or the redundancy in the
542 genetic architecture underlying fitness[17,18,48]. Further, secondary contact or ongoing
543 hybridization may violate model assumptions (as is the case between Douglas-fir varieties
544 and for jack pine introgression with *Pinus contorta* in the westernmost part of its range).

545 Despite these concerns, and the long list of potential factors affecting survival
546 throughout the lifetime of long-lived individuals, genetic offset results are often used with
547 future climates projected many decades into the future. Our results for 52-year-old jack
548 pine across two provenance trial sites offers cautious optimism, as we found that genetic
549 offset predictions for mortality by both GF and RONA were often better than non-genetic

offset measures despite the long timespan. Yet, the disparity of performance across the two species studied here highlight the utility of alternative forms of information to guide management such as long-term reciprocal transplant trials when resources and budgets allow.

There is still considerable uncertainty in the performance of genetic offset methods applied to natural populations[8,17,18]. Even with ideal data, offset inferences will be impacted by both evolutionary factors (e.g., drift, pleiotropy, strength of selection) and experimental parameters (e.g., the sampling locations of populations)[8]. It is less clear, however, the circumstances under which we should expect multiple offset methods to come to the same conclusion, as they are likely to be affected to different degrees for any given set of parameters. For instance, Láruson et al.[8] highlight how genetic drift can mislead GF's offset magnitude and rank estimates, which may be driving some patterns observed here, for example, the extent to which the western-most group of jack pine is inferred to be the most maladapted (Figs. 4A, S11) or the extent to which the cross-variety model of Douglas-fir infers the southeastern groups of the interior variety to be most maladapted (Figs. 5C, 6B, S21.9). Because of this, we hesitate to recommend either GF or RONA over the other, particularly given their similar performance. Instead, we recommend further exploration of their performance under a wide variety of scenarios. A more detailed understanding of how genetic offset methods interact with complex multivariate selection, admixture, and lesser degrees of (or variation in) local adaptation also warrant further attention.

571 Empirical and simulation work estimating genetic offsets, at least in the short term,
572 will need to thoroughly explore the results by varying input loci, climate data, and
573 populations used in training to understand how sensitive the offset estimations are to the
574 data at hand. Most importantly, multiple sources of validation should be used to
575 corroborate offset estimates, which could include data beyond phenotypic measurements
576 from common garden experiments (e.g., time-series satellite data to estimate population
577 growth rates). Doing so will lead to a more complete understanding of the performance of
578 these models, and the circumstances under which they will fail.

579 While our validation results show promise, our future projections for Douglas-fir show
580 ambiguous results (Figs. 5C-5D, S13B). Because of this, we do not recommend using offset
581 estimates to strongly influence prescriptions to guide climate-adaptive management
582 practices such as assisted gene flow for individual populations until these approaches are
583 better understood and validated. It therefore may be more prudent to work under the
584 assumption that all populations are at some risk of maladaptation due to climate change.
585 Even so, offset methods could guide *ex situ* conservation collections to capture genetic
586 diversity from populations predicted to be most at risk of climate-related extirpation, e.g.,
587 for seed banks or living collections. Further, while our offset projections for Douglas-fir
588 show ambiguity, monitoring of populations could lead to evidence that support one
589 projection over the other and could serve as another source of validation. In general
590 practice, there may be situations where the risks of inaction may outweigh risks associated
591 with model precision, and these could be weighed accordingly. Finally, the value of

common garden experiments for evaluating risk of maladaptation should not be underestimated.

5 | Acknowledgements

The CoAdapTree project is funded by Genome Canada (241REF; Co-Project Leaders SNA, SY and Richard Hamelin), with co-funding from Genome BC and 16 other sponsors (http://coadaptree.forestry.ubc.ca/sponsors/), including Genome Alberta, Génome Québec, the BC Ministry of Forests, Natural Resources Operations and Rural Development, Alberta Innovates Bio Solutions, Vernon Seed Orchard Company, University of Alberta, University of British Columbia, Compute Canada, Mosaic Forest Management, and Western Forest Products. We thank Centre d'expertise et de services Génome Québec for sequencing service, University of Calgary Information Technologies for system support, as well as Dr. Pia Smets and Christine Chourmouzis for technical assistance. We also thank CoAdapTree Scientific Advisory Board members Drs. John Davis, Matias Kirst, and Graham Coop for their guidance. Dr. Sam Yeaman is also funded by the Natural Sciences and Engineering Research Council of Canada (NSERC Discovery Grant RGPIN/03950-2017) and Alberta Innovates (20150252). Dr. Sally Aitken is also funded by NSERC Discovery Grant (RGPIN-2020-05136). The funding bodies did not have any role in the design of the study, collection, analysis, or interpretation of data in writing the manuscript.

6 | References

1. Urban, M. C. Accelerating extinction risk from climate change. *Science* 348, 571–573 (2015).

2. Nadeau, C. P., Urban, M. C. & Bridle, J. R. Climates Past, Present, and Yet-to-Come Shape Climate Change Vulnerabilities. *Trends Ecol Evol* 32, 786–800 (2017).

3. Thuiller, W. *et al.* Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives Plant Ecol Evol Syst* 9, 137–152 (2008).

4. Waldvogel, A.-M. *et al.* Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters* 4, (2020).

5. ONeill, G. A., Hamann, A. & Wang, T. Accounting for population variation improves estimates of the impact of climate change on species growth and distribution. *Journal of Applied Ecology* 45, 1040–1049 (2008).

6. Fitzpatrick, M. C. & Keller, S. R. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters* 18, 1–16 (2015).

7. Rellstab, C. *et al.* Signatures of local adaptation in candidate genes of oaks (Quercusspp.) with respect to present and future climatic conditions. *Molecular Ecology* 25, 5907–5924 (2016).

8. Láruson, Á. J., Fitzpatrick, M. C., Keller, S. R., Haller, B. C. & Lotterhos, K. E. Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. *Evol Appl* 15, 403–416 (2022).

9. Fitzpatrick, M. C., Chhatre, V. E., Soolanayakanahally, R. Y. & Keller, S. R. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Mol Ecol Resour* (2021) doi:10.1111/1755-0998.13374.

10. Bay, R. A. *et al.* Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science* 359, 83–86 (2018).

11. Gougherty, A. V., Keller, S. R. & Fitzpatrick, M. C. Maladaptation, migration and extirpation fuel climate change risk in a forest tree species. *Nature Climate Change* 1–15 (2021) doi:10.1038/s41558-020-00968-6.

12. Vanhove, M. *et al.* Using gradient Forest to predict climate response and adaptation in Cork oak. *J Evolution Biol* (2021) doi:10.1111/jeb.13765.

13. Savolainen, O., Pyhäjärvi, T. & Knürr, T. Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics* 38, 595–619 (2007).

14. Boshier, D. *et al.* Is local best? Examining the evidence for local adaptation in trees and its scale. *Environmental Evidence* 4, 1–10 (2015).

15. Sork, V. L. *et al.* Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes* 9, 901–911 (2013).

16. Lind, B. M., Menon, M., Bolte, C. E., Faske, T. M. & Eckert, A. J. The genomics of local adaptation in trees: are we out of the woods yet? *Tree Genetics & Genomes* 14, 29 (2018).

17. Rellstab, C., Dauphin, B. & Exposito-Alonso, M. Prospects and limitations of genomic offset in conservation management. *Evol Appl* 14, 1202–1212 (2021).

18. Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M. & Keller, S. R. Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes. *Annu Rev Ecol Evol Syst* 51, 245–269 (2020).

19. Davis, M. B. & Shaw, R. G. Range Shifts and Adaptive Responses to Quaternary Climate Change. *Science* 292, 673–679 (2001).

20. Hamrick, J. L. Response of forest trees to global environmental changes. *Forest Ecology and Management* 197, 323–335 (2004).

21. McLachlan, J. S., Clark, J. S. & Manos, P. S. Molecular indicators of tree migration capacity under changing climate. *Ecology* 86, 2088–2098 (2005).

22. Mckenney, D. W., Pedlar, J. H., Lawrence, K., Campbell, K. & Hutchinson, M. F. Potential impacts of climate change on the distribution of North American trees. *BioScience* 57, 939–948 (2007).

23. Alberto, F. J., Aitken, S. N. & Alía, R. Potential for evolutionary responses to climate change–evidence from tree populations. *Global Change Biology* 19, 1645–1661 (2013).

24. Allen, C. D. *et al.* A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management* 259, 660–684 (2010).

25. Millar, C. I., Stephenson, N. L. & Stephens, S. L. Climate change and forests of the future: managing in the face of uncertainty. *Ecological applications : a publication of the Ecological Society of America* 17, 2145–2151 (2007).

26. Kluyver, T. *et al.* Jupyter Notebooks -- a publishing format for reproducible computational workflows. in *Positioning and Power in academic publishing: Players, agents and agendas* (eds. Loizides, F. & Schmidt, B.) 87–90 (IOS Press, 2016).

27. Lind, B. M. *et al.* Haploid, diploid, and pooled exome capture recapitulate features of biology and paralogy in two non-model tree species. *Mol Ecol Resour* 22, 225–238 (2022).

28. Neale, D. B. *et al.* The Douglas-Fir Genome Sequence Reveals Specialization of the Photosynthetic Apparatus in Pinaceae. *G3 Genes Genomes Genetics* 7, 3157–3167 (2017).

29. Wegrzyn, J. L. *et al.* Unique Features of the Loblolly Pine (Pinus taeda L.) Megagenome Revealed Through Sequence Annotation. *Genetics* 196, 891–909 (2014).

30. Neale, D. B. *et al.* Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology* 15, 1–13 (2014).

31. Zimin, A. *et al.* Sequencing and Assembly of the 22-Gb Loblolly Pine Genome. *Genetics* 196, 875–890 (2014).

32. Lind, B. M. GitHub.com/CoAdapTree/varscan_pipeline: Publication release (Version 1.0.0). *Zenodo* (2021) doi:https://doi.org/10.5281/zenodo.5083302.

33. Gautier, M. Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics* 201, 1555–1579 (2015).

34. Booker, T. R. *et al.* The molecular basis of convergent local adaptation among distantly related conifers. *manuscript in prep* (2023).

35. Booker, T. R., Yeaman, S. & Whitlock, M. C. The WZA: A window-based method for characterizing genotype-environment association. *Biorxiv* 2021.06.25.449972 (2021) doi:10.1101/2021.06.25.449972.

36. Gugger, P. F., Fitz-Gibbon, S. T., Albarrán-Lara, A., Wright, J. W. & Sork, V. L. Landscape genomics of Quercus lobata reveals genes involved in local climate adaptation at multiple spatial scales. *Mol Ecol* 30, 406–423 (2021).

37. Borrell, J. S., Zohren, J., Nichols, R. A. & Buggs, R. J. A. Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. *Evol Appl* 13, 161–175 (2019).

38. AdaptWest-Project. Gridded current and projected climate data for North America at 1km resolution, generated using the ClimateNA software. http://adaptwest.databasin.org (2021).

39. 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* 11, e0156720 (2016).

40. Little, E. L. Jr. Atlas of United States trees, volume 1, conifers and important hardwoods: U.S. Department of Agriculture Miscellaneous Publication 1146, 9 p., 200 maps. https://web.archive.org/web/20170127093428/https://gec.cr.usgs.gov/data/little/ (1971).

41. Mahony, C. R. *et al.* Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. *Evol Appl* 13, 116–131 (2020).

42. O'Neill, G. *et al.* A proposed climate-based seed transfer system for British Columbia. *Prov. B.C., Victoria, B.C. Tech. Rep. 099* (2009).

43. Wang, T., Hanann, A., Yanchuk, A., O'Neill, G. A. & Aitken, S. N. Use of response functions in selecting lodgepole pine populations for future climates. *Global Change Biol* 12, 2404–2416 (2006).

44. Chakraborty, D. *et al.* Selecting Populations for Non-Analogous Climate Conditions Using Universal Response Functions: The Case of Douglas-Fir in Central Europe. *Plos One* 10, e0136357 (2015).

45. Candido-Ribeiro, R. *et al.* Distribution of cold hardiness across Douglas-fir natural range and its implications for the assisted gene flow between populations. *manuscript in prep* (2022).

46. Gugger, P. F., Sugita, S. & Cavender-Bares, J. Phylogeography of Douglas-fir based on mitochondrial and chloroplast DNA sequences: testing hypotheses from the fossil record. *Mol Ecol* 19, 1877–1897 (2010).

47. Parchman, T. L., Jahner, J. P., Uckele, K. A., Galland, L. M. & Eckert, A. J. RADseq approaches and applications for forest tree genetics. *Tree Genet Genomes* 14, 39 (2018).

48. Láruson, Á. J., Yeaman, S. & Lotterhos, K. E. The Importance of Genetic Redundancy in Evolution. *Trends Ecol Evol* 35, 809–822 (2020).

7 | Data Availability

We reference the analysis code in the text of our documents by designating Supplemental Notebooks (SN) using a directory numbering system from our servers (as opposed to the order listed in the manuscript). For example, for Notebook 15 in Directory 3, we will refer to SN 03.15; for Notebook 1 in Subfolder 2 of Directory 5, we will refer to SN 05.02.01.

These notebooks not only contain the analysis code, but also contain code output and display attributes of the data objects being analyzed. Each of these directories will be archived on Zenodo.org and include a citation below, which will also link to the GitHub repository. Notebooks are best viewed within a local jupyter or jupyter lab session, but can also be viewed at nbviewer.jupyter.org using the web link to the notebook archived on GitHub. Analyses were carried out primarily using python v3.8.5 and R v3.5.1. Exact package and code versions are available at the top of each notebook.

At the time of submission, all coding records will be archived on Zenodo.

Supplemental Notebooks:

SN 15 – GitHub.com/brandonlind/offset_validation

Raw sequence data will be deposited on the Sequence Read Archive of the National Center for Biotechnology Information (NCBI SRA). All remaining data necessary for the replication of our work will be archived on Dryad.

8 | Author Contributions

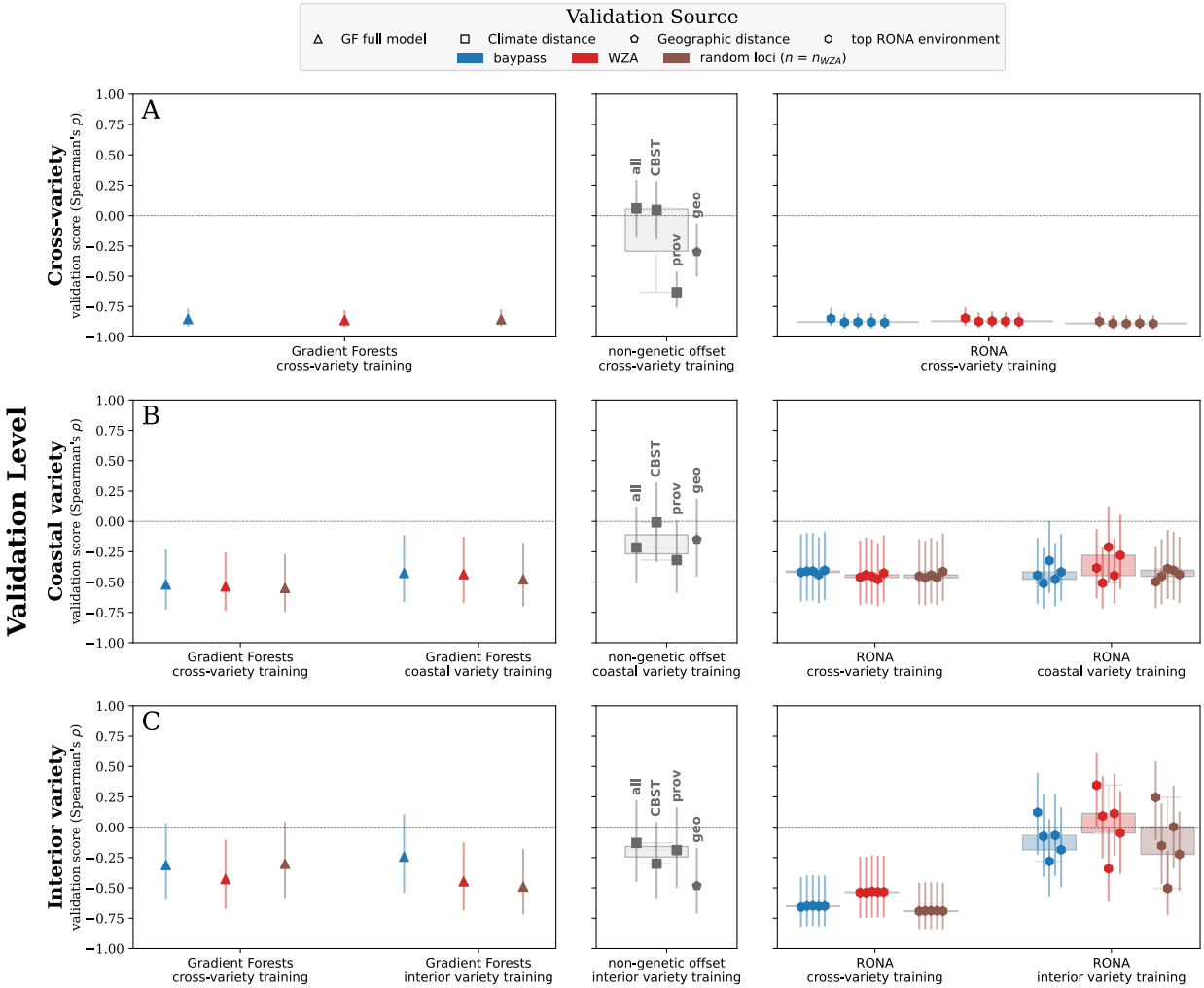
Sally Aitken and Sam Yeaman obtained funding. Brandon Lind and Sally Aitken conceived the offset validation study. Rafael Candido-Ribeiro and Nathalie Isabel

provided phenotypes used in validation for Douglas-fir and jack pine, respectively. Pooja Singh carried out gene annotation and implemented WZA, which was designed by Tom Booker, Michael Whitlock, and Sam Yeaman. Brandon Lind processed raw genetic data, called SNPs, and carried out single-locus GEA for each species. Brandon Lind designed stratified sampling of populations, and carried out implementation of genetic offset training, validation, and projection. Brandon Lind wrote the manuscript, created figures, and curated coding records for archiving. All authors contributed to improvements of the study design and the conceptualization of results.

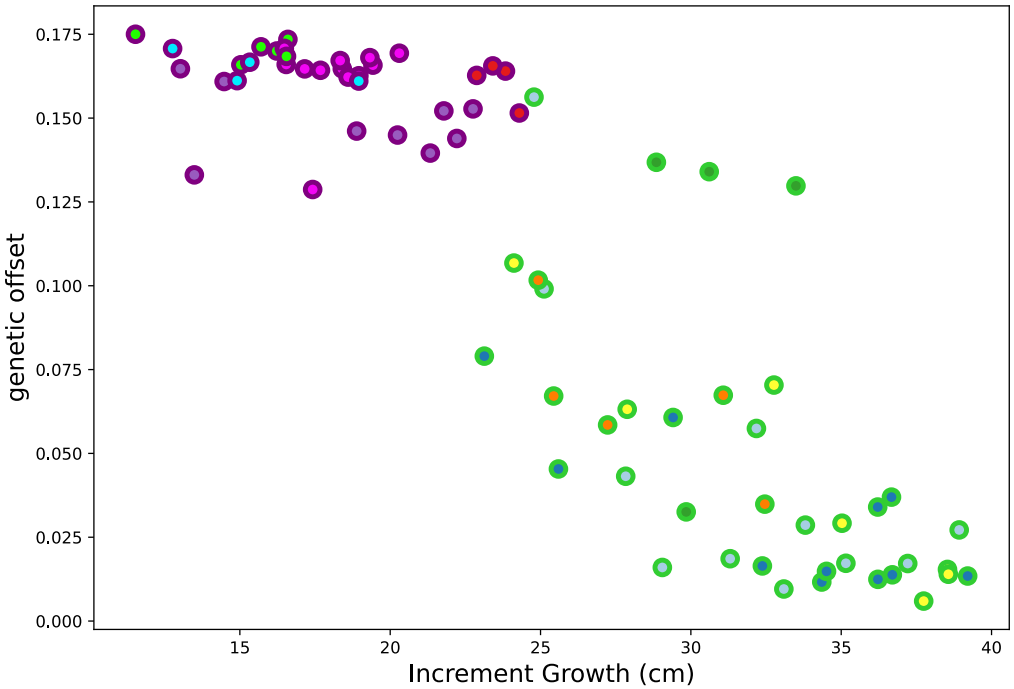
9 | Extended Data

Extended Data Table 1 Locus counts used in training of Gradient Forests (GF) and the Risk Of Non-Adaptedness (RONA) for each set of populations used: coastal Douglas-fir, interior Douglas-fir, jack pine, and across both varieties of Douglas-fir (cross-variety). Not shown are redundant counts of random marker sets with the same sample size as the BayPass and WZA sets used in GF. Marker sets used for RONA are subsets of those used in GF that had significant linear models with at least one environment. Subscript letters b and w refer to the original outlier sets (BayPass, and WZA, respectively) used to determine sample sizes for random marker sets used in GF which were then subset to form the counts shown for RONA. Code used to create this table can be found in SN 15.15.

	Gradient Forests		RONA			
	BayPass	WZA	BayPass	random_b	WZA	random_w
Jack pine	22635	4788	22571	11899	4788	2538
Cross-variety	25219	14760	24844	23147	14629	13537
Coastal	17516	4886	17436	10034	4881	2761
Interior	12938	11434	12276	7060	11305	6583



Extended Data Fig. 1 Offset validation from two-year Douglas-fir shoot biomass phenotypes at the Vancouver common garden (see Fig. 1A) using Gradient Forests (GF), the Risk of Non-Adaptedness (RONA), and climate and geographic distances. We used genetic hierarchy (rows) to assess accuracy inference from trained models (x-axis groups) using populations (rows) across both varieties of Douglas-fir (A), at the variety level for the coastal (B) and interior varieties of Douglas-fir (C) to determine if greater numbers of training populations improve finer-scale predictions of offset. Triangles indicate performance of GF models trained and validated using all available populations. RONA background boxplots are indicative of the range of RONA estimates given for the top five climatic variables (hexagon) that differed significantly between source population and the common garden (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. Vertical bars indicate standard error estimated using a Fisher transformation (see Supplemental Text S1.3). Loci used in RONA calculations are a subset of those used in Gradient Forests that had significant linear models with the environment, see Table 1 for locus counts. See Fig. 3 for similar validation using shoot biomass. See Fig. S9 for all locus groups. Boxplot whiskers extend up to 1.5x the interquartile range. Code to create these figures can be found in SN 15.14. Code to create these figures can be found in SN 15.14.



Extended Data Fig. 2 Genetic structure drives high validation scores in the Douglas-fir cross-variety models of Gradient Forest. Shown is the relationship between increment growth of coastal and interior varieties of Douglas-fir and offset values from the cross-variety model of Gradient Forests (trained using WZA outliers and all populations). Edges are colored using the shade of each variety's geographic range in Fig. 1 (lime: coastal Douglas-fir, and purple: interior Douglas-fir), and the interior of each point is colored with respect to the genetic groups from Fig. 1. Code used to create this figure can be found in SN 15.22.

Supplemental Information

How useful is genomic data for predicting maladaptation to future climate?

Brandon M. Lind[1,*], Rafael Candido-Ribeiro[1], Pooja Singh[2], Mengmeng Lu[2], Dragana Obreht Vidakovic[1], Tom R. Booker[3], Michael C. Whitlock[3], Sam Yeaman[2], Nathalie Isabel[4,5], Sally N. Aitken[1]

[1]Centre for Forest Conservation Genetics and Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada

[2]Department of Biological Sciences, University of Calgary, Calgary, AB, Canada

[3]Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada

[4]Canada Research Chair in Forest Genomics, Centre for Forest Research and Institute for Systems and Integrative Biology, Université Laval, Québec, QC, Canada

[5]Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Québec, QC, Canada

10 February 2023

Running Title: *Validating genetic offsets*

Keywords: climate vulnerability, climate change, machine learning, genetic offset

***Corresponding Author**

Brandon Lind

2424 Main Mall

3027 Forest Science Centre

University of British Columbia

Vancouver, BC, Canada

Email: lind.brandon.m@gmail.com

1 | Supplemental Text

1.1 | Implementation of *k*-fold cross validation for Gradient Forests

For each marker set used to train GF, we created *k* groups of training populations via stratified *k*-fold sampling (SN 15.01), where the overall proportion of populations from a given genetic group (circle colors in Fig. 1; Supplemental Text S1.2) used in any of the *k*-fold sets for training is roughly equivalent to the proportion of that genetic group in the full data set (see Table S4 for *k*-fold proportions). This ensured that 1) we were not randomly leaving out an excess of populations from any given portion of the range, 2) that validation sets were similar to those used in training, and 3) that each available population was used only once for cross-validation of a given model. Doing so will reveal the sensitivity of these GF models with respect to both marker and population input (Q4). For instance, if 1) the relationship between offsets from a model using all populations is not strongly correlated to the union of out-of-bag offsets from *k*-fold sets, or 2) if cross-validation values vary widely about the estimates from models using all populations, this would indicate that GF is sensitive to population input. If so, any projected offset in areas of the range that lack training populations should not necessarily be taken at face value. Our implementation represents a best-case scenario for modelling offset in geographic areas where training data does not exist because of the proximity of the populations left out of training with those that were used in training.

Within the text and code, we refer to these *k*-fold groups using numerical IDs – e.g., k1, k2, k3, k4; each *k*-fold group refers to a distinct set of test populations (SN 15.01). If there are four *k*-fold cross-validation sets (Douglas-fir), this indicates that 25% of the populations were left out of training for each fold, while when there are three *k*-fold cross-validation sets (jack pine), this indicates that 33% of the data was left out of training for each fold. The number of *k*-fold sets was limited by the number of populations within color groups of Fig. 1 (i.e., minimum group count was either n=3 or n=4). We used these population assignments consistently across GF analyses so that we could more easily compare results between marker sets.

1.2 | Genetic grouping

We used genetic groups as informed by PCA *k*-means clustering (Fig. S1-S2). To choose the optimal number of genetic groups we used a combination of the mean silhouette coefficient scores from first three axes of PCA fit using loci with no missing data and the second derivative of the line fit (hereafter, Line 1) using the Number of Clusters and the Within-Cluster Sum of Squares (SN 07.01, SN 02.12.01, SN 02.12.03). The within-cluster sum of squares used here is the unweighted sum of squared distances of samples to their closest cluster center for a given value of *k*. The silhouette coefficient is calculated on a per-sample basis using the mean intra-cluster distance and the mean nearest-cluster distance; the best value is 1 and the worst value is -1. We calculated mean silhouette coefficient scores for *k*=2-20, each time recording the within-cluster sum of squares, and

chose the optimal number of k when the mean silhouette coefficient first reached a local maximum (for any $k > 2$) and Line 1 appeared to have a second derivative close to zero. For Douglas-fir varieties, this generally resulted in less than three optimal clusters; for the interior variety this was the northwestern and southeastern genetic groups, and for the coastal variety this was the central range and range-edge populations. So we used information from the PCA, k-means clustering, and geography (see subfolder 02.03) to increase the number of groups over which to carry out cross-validation, in line with groupings found when using larger values of k during k-means clustering of PCA data. This ensured a more even sampling from throughout the range when we implemented k-fold (Supplemental Text S1.1). PCAs colored with according to the k-means clustering (i.e., colors in Fig. 1) can be seen in Fig. S1 for Douglas-fir and Fig. S2 for jack pine.

1.3 | Consistency of Environmental Variable Importance Rank

The environmental variables driving local adaption are often unknown *a priori*, and landscape ecologists must often choose among highly correlated environmental variables when analyzing data. Random forest models, like those implemented within Gradient Forests, are well suited to identify features important for predictive models, even when presented with correlated features. In the context of GF, these features are environmental variables. We assessed the feature importances output by GF using both accuracy importance and weighted accuracy importance and visualized the consistency of these importance ranks from each dataset using slope graphs. If the various marker sets and

populations used to train GF models have little impact with respect to climate feature importance, ranks should be stable across these outcomes and therefore these slope graphs should contain relatively little noise (i.e., no large or numerous changes in rank), at least for the top-ranked climates which may contain considerably larger importance values than lower ranked climate features. If these figures are noisy throughout all ranks, this tells us that the input data impacts model outcomes with respect to inferred climate importance.

Among marker sets for a given set of populations used in training, the environmental ranks using either accuracy importance or weighted importance from models with outlier or random loci were relatively stable across runs (i.e., the *k*-fold iterations as well as the model using all populations; Supplemental Figs. S3-S4).

Across the Douglas-fir models trained using BayPass or WZA outliers, top environmental ranks from weighted importance values were most similar across runs between the cross-variety and interior-only models, where TD, EMT, Elevation, and MCMT were consistently ranked in the top five environments across most runs (except for Elevation which ranked sixth in *k*-fold set k3 of the cross-variety model; Fig. S7A-C). For the coastal variety models trained using outlier sets, Eref and CMD were consistently ranked in the top three variables across runs, MSP was consistently ranked in the top six variables across runs (though generally greater than rank 5), and SHM was generally ranked above the sixth rank except for WZA *k*-fold sets *k1* and *k2* (Fig. S7).

For all Douglas-fir models trained using outliers and all populations (blue circles and diamonds, Fig. S7), the top six environments were consistently ranked between

BayPass and WZA outlier sets. Together with the environments consistently ranked across runs, these data suggest that the interaction of summer precipitation and temperature play an important role in structuring sampled genetic variation in the coastal variety, while cold temperatures predominate as important factors structuring sampled genetic variation within the interior variety. Similar to Douglas-fir, models trained using all jack pine populations and loci from BayPass and WZA outliers produced consistent rank in the top five environmental variables: MCMT, MAP, DD_0, MAT, and MSP (Fig. S7D; Table S2). These results indicate both cold temperatures and precipitation drive patterns within our genetic data for jack pine (Fig. S7). For both species, the ranks from weighted importance were in general agreement with those from accuracy importance (Fig. S5).

Except for the GF model trained using the coastal variety, GF models generally ranked the majority of CBST environmental variables within the top 50% of ranks (blue and red environmental labels in Fig. S7, Figs. S3-S5). However, environmental variables explaining significant variation in Douglas-fir provenance trials were generally in the bottom 50% of ranks, except for MAT in the interior-only models (brown and red environmental labels in Fig. S7A-C, panels A-D in Figs. S3-5). For jack pine, the provenance trial environments (except for SHM) generally ranked in the top 50%, where environments related to both CBST and provenance trials generally predominated top ranks (brown, red, and blue labels in Fig. S7D, panels E-F in Figs. S3-5). In general, the magnitudes of weighted importance indicated that between 1 to 5 climate variables could

be inferred to have the greatest importance in these models (Fig. S6; analogous histograms to Fig. S6 for accuracy importance show similar patterns and are not shown, except in SN 15.13). Given the relatively consistent rank across these top-ranked climate features, our analyses indicate that both marker and population input has relatively less impact on climate importance inference than it does when predicting genetic offset.

Our analyses suggest that inference related to climate importance is relatively less sensitive than offset predicted from models of GF, though whether the most important climates ranks indicate the climates most relative to adaptation, as opposed to just patterns from the data, remains to be seen and warrants further attention.

1.4 | Common garden phenotypic information for Douglas-fir and jack pine

1.4.1 | jack pine

The *Pinus banksiana* genetics program was initiated by Petawawa Experimental Forest Research in 1950. Between 1955-1961, seeds were collected by Canadian Forest Service (CFS) and United States (US) collaborators from 99 *P. banksiana* provenances throughout the species' natural range (Yeatman 1974). In 1962, seeds were sown and seedlings were then grown and evaluated at the Petewawa nursery. Three to four years later, twelve field tests, with varying numbers of provenances, were established in Canada (Ontario, Quebec, and New Brunswick) and the US to assess the pattern of adaptive genetic variation (Yeatman 1974).

In September 2018, a census was conducted in two Quebec field tests. Tree status (e.g. alive or dead) was noted along with tree height (m) and diameter at breast height (cm) for each of the trees still present. Provenance numbers 73 and 74 were assigned to Marl Lake with the same LAT and LONG coordinates, had the same mortality levels, and similar means for height and diameter at breast heigh (DBH) – we therefore averaged these populations for height and DBH and assigned these values in the data to Provenance 73 and removed Provenance 74.

Since we were using a trial planted long ago, we matched provenance ID (LAT and LONG) used for phenotyping with that from genomic populations used for offset training by ensuring each provenance ID was < 1km from only one other population ID from our pool-seq data (SN 15.06; yellow- and black-edged circles, Fig. 1B); the population-provenance pairs that were kept all had a distance of 0.0 km, except for one pair that had a distance of 0.048 km.

1.4.2 / Douglas-fir

The Douglas-fir common garden experiment was established in outdoor nursery raised beds at the University of British Columbia campus in March 2018. One-year-old seedlings of 73 natural populations (38 var. *menziesii* and 35 var. *glauca*) spanning most of the natural range of the species (see Fig. 1A) were randomized into 11 blocks with an unbalanced block design. Each block contained 240 seedlings surrounded by a row of unmeasured edge seedlings. We used a spacing of 8 x 8 cm between seedlings and between one and 13 seedlings per provenance in each block. The total number of seedlings per

166 provenance in the experiment varied between 11 and 95. All individuals in the experiment
167 were assessed for initial height in 2018 (before bud flush of the second growing season)
168 and final height in 2019 (after bud set of the third growing season), and final total shoot
169 dry biomass.

170 Best linear unbiased estimates (BLUEs) of the two-year height increments were
171 obtained for provenances before testing for associations with the genetic offsets. Height
172 increments were first log transformed to meet the assumptions of normality of residuals
173 and homoscedasticity of variances in the models. The following mixed effect model
174 implemented in ASReml-R 4.0 (Butler et al. 2007) was used:

175

176 $$Y_{ij} = \mu + \beta x_1 + \alpha x_2 + \epsilon_{ij}$$

177

178 where Y_{ij} is the phenotype height increment corresponding to individuals from provenance
179 *i* and block *j*; μ is the phenotype global mean across all individuals within the experiment
180 (fixed intercept), β is the coefficient for the fixed effect of provenance (x_1) and α is the
181 random effect of blocks (x_2) and ϵ_{ij} is the error term for Y_{ij}.

182 **1.5 | Estimating error in Spearman's ρ estimates**

183 We used Fishers Transformation to estimate the lower and upper confidence
184 intervals of Spearman's ρ estimates used in validation via the following equations:

185 $$F(\rho) = artanh\,(\rho)$$

186 $$SE = {}^{1}/_{\sqrt{N-3}}$$

$$lower = \tanh\left(\mathrm{F}(\rho) - (1.96 * SE)\right)$$

$$upper = \tanh(F(\rho) + (1.96 * SE))$$

Where $F(\rho)$ is Fisher's transformation of Spearman's ρ, and SE is the standard error standardized using the number of populations used in the correlation estimate (N). Some cross-validation scores for Douglas-fir were equal to $\rho = -1.0$ (Supplemental Figs. S9D-E and S10D-E) and in these cases we set $\rho = -0.9999999999999999$ to avoid domain errors when calculating $F(\rho)$.

1.6 | Exploring fine-scale validation of models trained across larger spatial scales

We continued exploring the effects of fine-scale validation by subsetting the interior variety into its two distinct genetic groups (the northwestern group and the southeastern group, see Supplemental Fig. S1) to use for validation from models trained across both varieties or those trained using all interior populations. While the GF models using all interior populations in validation (triangles, Fig. 3C) were in some cases outperformed by a non-genetic offset measure, validation scores using either of the two interior genetic groups often exceeded these same climate and geographic distance measures (Fig. S9D-E). Furthermore, these sub-interior variety scores (Fig. S9D-E) were often stronger than scores obtained at the interior level (Fig. 3C), and even exceed those from the coastal variety (Fig. 3B). However, cross-variety and interior-only models showed similar performance (Fig. S9C-E). Boxplots in Fig. S9D-E span a wide range, due in part to the small sample sizes of cross-validation assignments from stratified sampling assigned from cross-variety and interior-only models (colors, Fig. 1A). Similar patterns were found for shoot biomass (Fig S10).

1.7 | Inconsistency between projected maladaptation between common garden and future climates for interior Douglas-fir

The relative rank of projected maladaptation between the two genetic groups within interior Douglas-fir was inconsistent between that predicted for the Vancouver common garden and for future climate. For projections of maladaptation to future climate, the cross-variety and interior-only models of both GF (Fig. 5C-D) and RONA

(Supplemental Fig. S13B) disagreed as to whether the northwestern genetic group was more maladapted to RCP 8.5 2050s than the southeastern genetic group. However, for projections to the common garden (i.e., models depicted in Fig. 3 and Extended Data Fig. 1) the cross-variety model and interior-only model of GF were in agreement (Fig. S14) and were in disagreement for RONA (Fig. S15). It is unclear why such a 'flip-flop' would occur between cross-variety and interior-only GF models for future climate but not for this model when GF projected to the Vancouver common garden. In each case, we used the same GF model object from R and then input either the Vancouver climate or the future climate (SN 15.07). The difference between these input climates was of course the values used, but in the case of the Vancouver climate input, all of the populations had a single value for a given climate variable. Other than this, it is unclear what could be causing these differences. It may be because of the differences between current and future (common garden or RCP) climates, where the future climates are more differentiated than that of the common garden from current values. Or perhaps ancestry violating model assumptions (the northwestern interior genetic group shares secondary contact with the coastal variety). Future simulation work could explore the effects of ancestry or climate dissimilarity from current values to determine if such patterns would impact projected offset rank, but this task is beyond the ability of the sparse pool-seq data available here.

233 **1.8 | Choosing future offset models to predict the most maladaptive populations**

234 Based on the similar level of validation performance among marker sets of either
235 outlier or random loci for both jack pine and Douglas-fir, we explored offset to future
236 climate using WZA outlier loci. Sensitivity of this model to population input was also
237 explored for both St. Christine (Fig. S16) and Fontbrune (Fig. S17) which was similar to
238 other locus sets (SN 15.21).

239 Next, for Douglas-fir, we needed to decide between the cross-variety or variety-
240 specific models (i.e., which input populations to use) to project offset to future climate
241 change. We used four criteria to choose among models by using insight from model
242 validation and projection: 1) validation scores, 2) model sensitivity to input populations
243 inferred from the strength of the relationship between the union of out-of-bag *k*-fold offset
244 predictions with the offset predictions from models trained using all populations, 3) the
245 strength of the relationship between offset predicted from GF with that of RONA, 4) the
246 strength of the relationships among RONA estimates from the top environments
247 differentiated between current and future scenarios.

248 1) While the cross-variety and variety-specific models of GF had similar
249 performance when validated at the variety level, the cross-variety model of RONA had
250 stronger validation scores for the interior variety populations than did the interior-only
251 model (Fig. 3, Extended Data Fig. 1).

252 2) Additionally, projections from interior-only GF models for the Vancouver
253 common garden had clear outliers and a much weaker relationship between the union of

out-of-bag offset predictions with the offset predicted using all populations (Figs. S18) relative to the cross-variety models (Fig. S19) or the coastal variety models (Supplemental Fig. S20). We therefore use cross-variety models to project maladaptation to future climate (RCP 8.5 2050s; Fig. 6).

3) For both the coastal and interior variety, and the genetic groups from within the interior variety, the offset predicted from the GF cross-variety models had a stronger relationship with that predicted from RONA cross-variety models than the comparison between GF and RONA variety-specific models (Fig. S21).

4) Furthermore, models for interior Douglas-fir had stronger relationships among RONA estimates from cross-variety models than from interior-only models (Supplemental Fig. S21 iii-viii).

It is worth noting that the cross-variety model ($n_{\text{populations}} = 73$) may be less sensitive than the interior-only model ($n_{\text{populations}} = 35$) due to the relatively greater number of input populations. However, given the less sensitive behavior of coastal-only models ($n_{\text{populations}} = 38$) relative to interior-only models, where input population numbers are more similar between varieties relative to either variety and the cross-variety models, the population numbers input to interior variety models may not be completely driving sensitivity patterns seen here.

2 | Supplemental Tables

Environment	p-value	Common Garden	Population set
EXT	1.084316e-27	Ste-Christine-d'Auvergne	jack pine
MSP	9.095524e-16	Ste-Christine-d'Auvergne	jack pine
MAP	5.559867e-12	Ste-Christine-d'Auvergne	jack pine
SHM	5.607879e-10	Ste-Christine-d'Auvergne	jack pine
Elevation	7.337828e-09	Ste-Christine-d'Auvergne	jack pine
EXT	1.229238e-25	Fontbrune	jack pine
MSP	1.640940e-07	Fontbrune	jack pine
SHM	3.166373e-06	Fontbrune	jack pine
AHM	2.229745e-05	Fontbrune	jack pine
MAP	2.495356e-05	Fontbrune	jack pine
bFFP	3.134143e-44	Vancouver	cross-variety
FFP	6.294516e-37	Vancouver	cross-variety
NFFD	2.498133e-27	Vancouver	cross-variety
eFFP	9.501555e-27	Vancouver	cross-variety
DD5	4.645407e-26	Vancouver	cross-variety
bFFP	1.312699e-21	Vancouver	coastal variety
FFP	1.996474e-18	Vancouver	coastal variety
eFFP	2.866360e-12	Vancouver	coastal variety
NFFD	3.582954e-12	Vancouver	coastal variety
EMT	6.254842e-12	Vancouver	coastal variety
bFFP	3.205920e-32	Vancouver	interior variety
FFP	1.292912e-31	Vancouver	interior variety
eFFP	6.564966e-30	Vancouver	interior variety
NFFD	8.936509e-30	Vancouver	interior variety
TD	1.051689e-29	Vancouver	interior variety

Table S1 Top five environments that differed between source populations and common gardens ranked by t-test p-values. These ranks were used to decide which locus-environment relationships to use when calculating the range of RONA (Figs. 2-4 of the main text). Data used to create this table is in SN 15.09 cell 64.

Annual variables:

MAT - mean annual temperature (°C)

MWMT – mean warmest month temperature (°C)

MCMT – mean coldest month temperature (°C)

TD – temperature difference between MWMT and MCMT, or continentality (°C)

MAP – mean annual precipitation (mm)

MSP – May to September precipitation (mm)

AHM – annual heat-moisture index (MAT+10)/(MAP/1000)

SHM – summer heat-moisture index MWMT/(MSP/1000)

Derived annual variables:

DD0 – degree-days below 0°C

DD5 – degree-days above 5°C

NFFD – number of frost-free days

FFP – frost-free period

bFFP -the day of the year on which FFP begins

eFFP – the day of the year on which FFP ends

PAS – precipitation as snow (mm) between August in previous year and July in current year

EMT – extreme minimum temperature over 30 years

EXT – extreme maximum temperature over 30 years

Eref – Hargreaves reference evaporation (mm)

CMD – Hargreaves climatic moisture deficit (mm)

Table S2. Climate variables used for offset calculations.

Jack pine

	group_count	group_perc	k1_count	k1_perc	k2_count	k2_perc	k3_count	k3_perc
blue	15	0.384615	10	0.384615	10	0.384615	10	0.384615
darkorange	10	0.256410	7	0.269231	7	0.269231	6	0.230769
green	11	0.282051	7	0.269231	7	0.269231	8	0.307692
magenta	3	0.076923	2	0.076923	2	0.076923	2	0.076923

Cross-variety (both varieties of Douglas-fir)

	group_count	group_perc	k1_count	k1_perc	k2_count	k2_perc	k3_count	k3_perc	k4_count	k4_perc
California	5	0.068493	3	0.055556	4	0.072727	4	0.072727	4	0.072727
Eastern BC	6	0.082192	4	0.074074	5	0.090909	5	0.090909	4	0.072727
Lower Mainland	4	0.054795	3	0.055556	3	0.054545	3	0.054545	3	0.054545
Montana	4	0.054795	3	0.055556	3	0.054545	3	0.054545	3	0.054545
North BC	12	0.164384	9	0.166667	9	0.163636	9	0.163636	9	0.163636
Oregon	11	0.150685	9	0.166667	8	0.145455	8	0.145455	8	0.145455
South Central BC	4	0.054795	3	0.055556	3	0.054545	3	0.054545	3	0.054545
Tri-state	9	0.123288	7	0.129630	6	0.109091	7	0.127273	7	0.127273
Washington	7	0.095890	5	0.092593	6	0.109091	5	0.090909	5	0.090909
Western BC	11	0.150685	8	0.148148	8	0.145455	8	0.145455	9	0.163636

Interior Douglas-fir

	group_count	group_perc	k1_count	k1_perc	k2_count	k2_perc	k3_count	k3_perc	k4_count	k4_perc
Eastern BC	6	0.171429	4	0.153846	4	0.153846	5	0.192308	5	0.185185
Montana	4	0.114286	3	0.115385	3	0.115385	3	0.115385	3	0.111111
North BC	12	0.342857	9	0.346154	9	0.346154	9	0.346154	9	0.333333
South Central BC	4	0.114286	3	0.115385	3	0.115385	3	0.115385	3	0.111111
Tri-state	9	0.257143	7	0.269231	7	0.269231	6	0.230769	7	0.259259

Coastal Douglas-fir

	group_count	group_perc	k1_count	k1_perc	k2_count	k2_perc	k3_count	k3_perc	k4_count	k4_perc
California	5	0.131579	4	0.142857	3	0.107143	4	0.137931	4	0.137931
Lower Mainland	4	0.105263	3	0.107143	3	0.107143	3	0.103448	3	0.103448
Oregon	11	0.289474	8	0.285714	9	0.321429	8	0.275862	8	0.275862
Washington	7	0.184211	5	0.178571	5	0.178571	6	0.206897	5	0.172414
Western BC	11	0.289474	8	0.285714	8	0.285714	8	0.275862	9	0.310345

Table S4. Implementation of stratified *k*-fold sampling. Groups of populations over which stratified *k*-fold sampling was carried out are shown as rows (see colors in Fig. 1 of the main text). Each group represents a given number of populations in the full dataset (`group_count`) and therefore a proportion of the overall dataset at hand (`group_perc`). Stratified *k*-fold sampling is implemented so that within each fold (`k1`, `k2`, `k3`, `k4`) each group of populations maintains a similar proportion in the training and test data as that group did in the full dataset. Shown for each *k*-fold are the number of populations representing a group (e.g., `k1_count`) and the proportion of that group's populations in the fold's training set (e.g., `k1_perc`). This contrasts with traditional cross-validation, where in each fold populations are chosen at random irrespective of their population, location, or genetics. These tables were produced in SN 15.01.

3 | Supplemental Figures

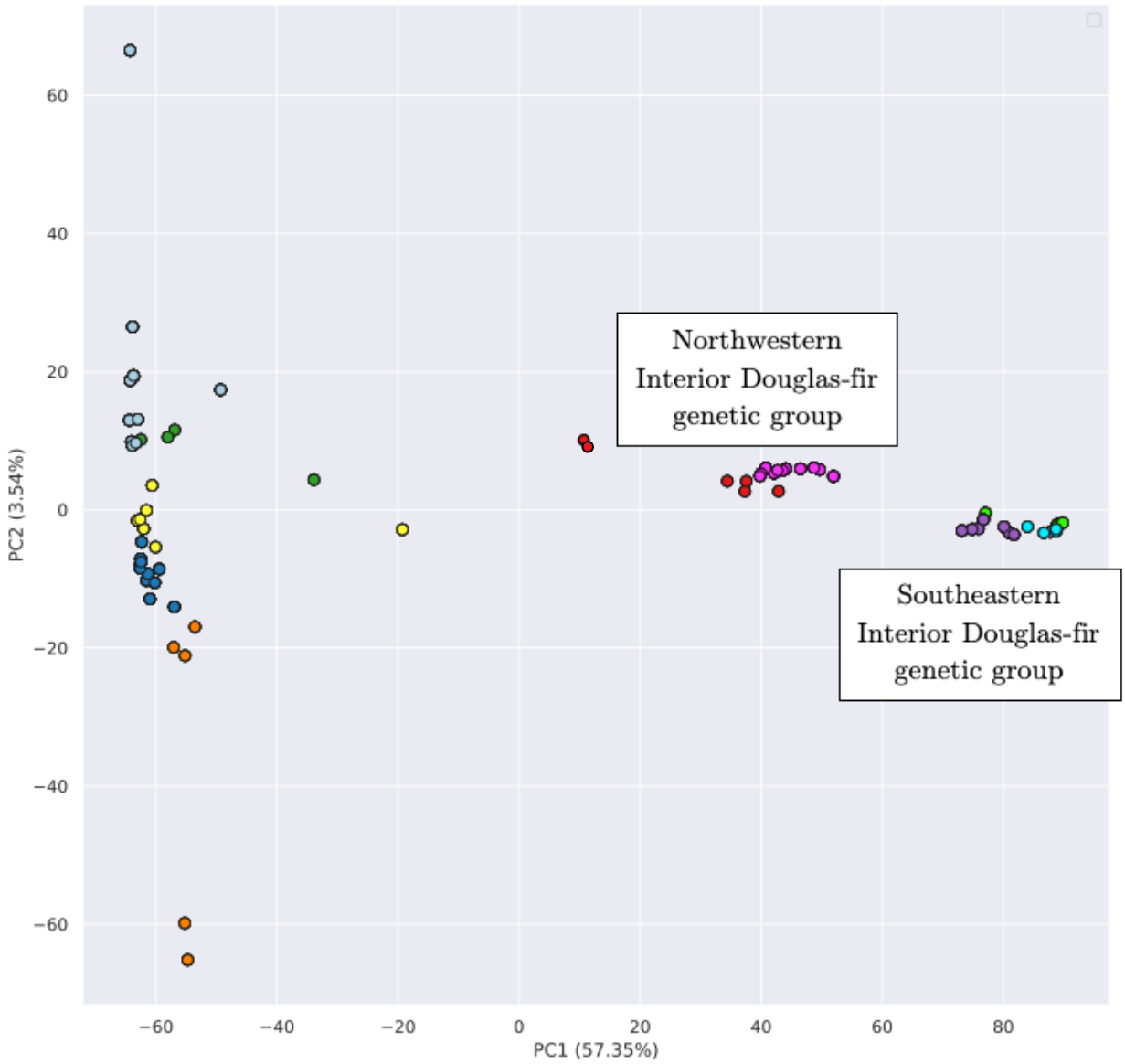


Fig. S1 Principle components analysis (PCA) of the two varieties of Douglas-fir using SNPs called across both varieties. Populations are colored as in Fig. 1 of the main text. Code used to create this figure can be found in SN 02.12.02.

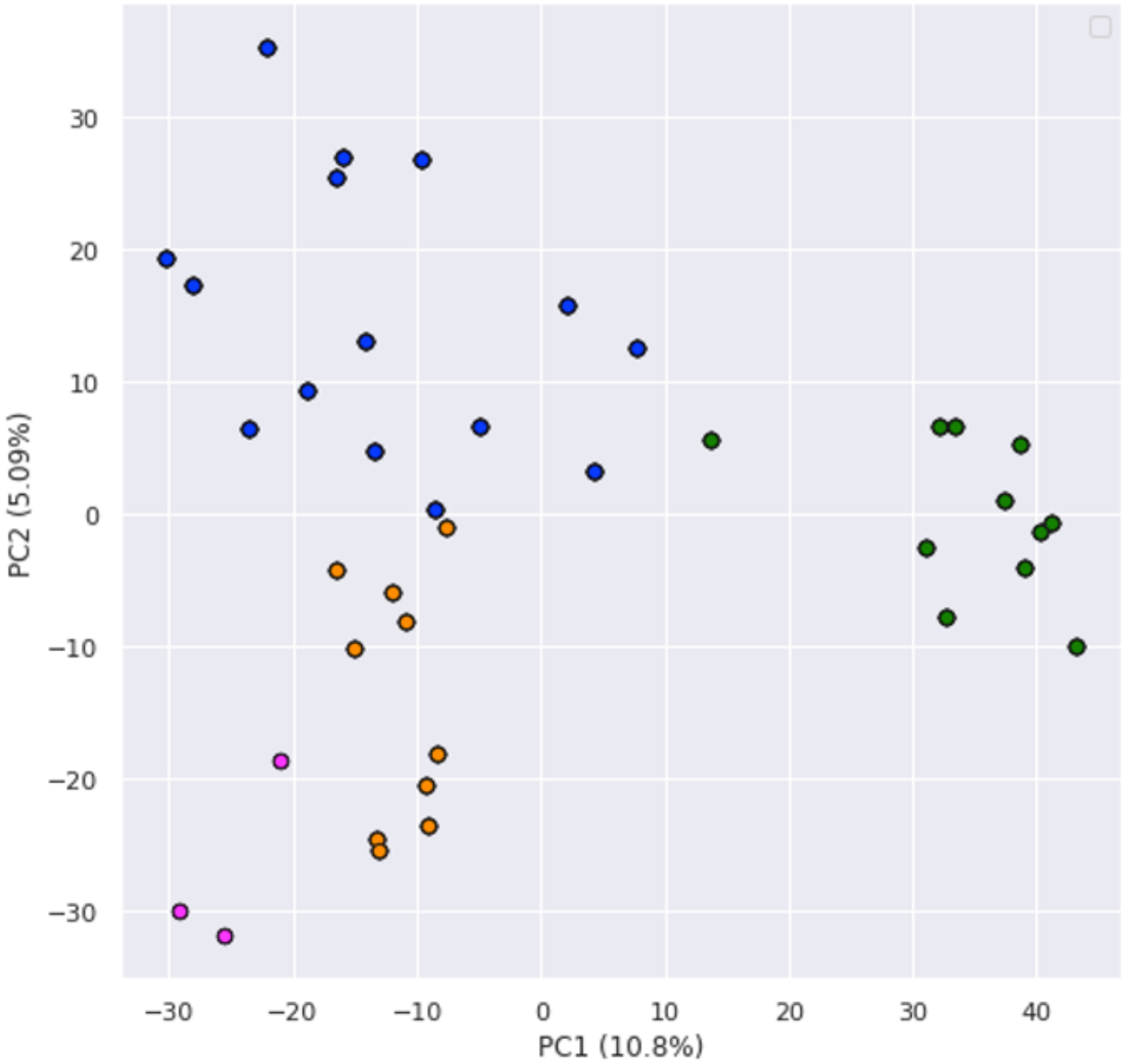
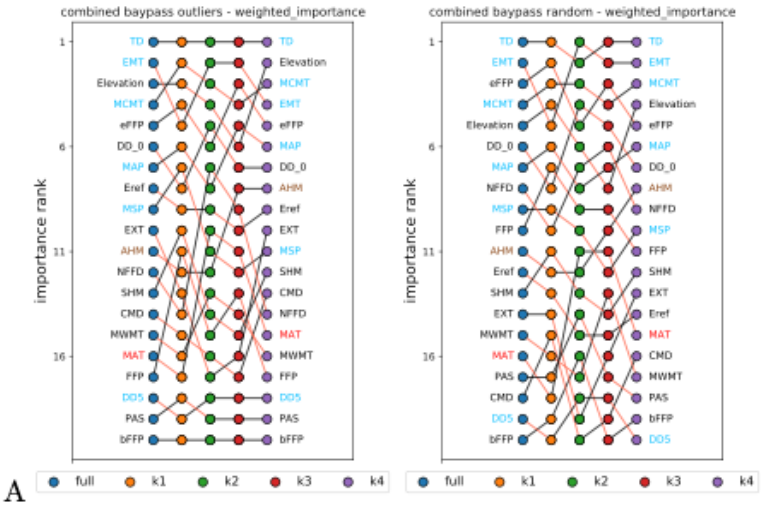


Fig. S2 Principle components analysis (PCA) of jack pine. Populations are colored as in Fig. 1 of the main text. Code used to create this figure can be found in SN 02.12.02.

Fig S3 – Consistency of weighted environmental importance from Gradient Forests.



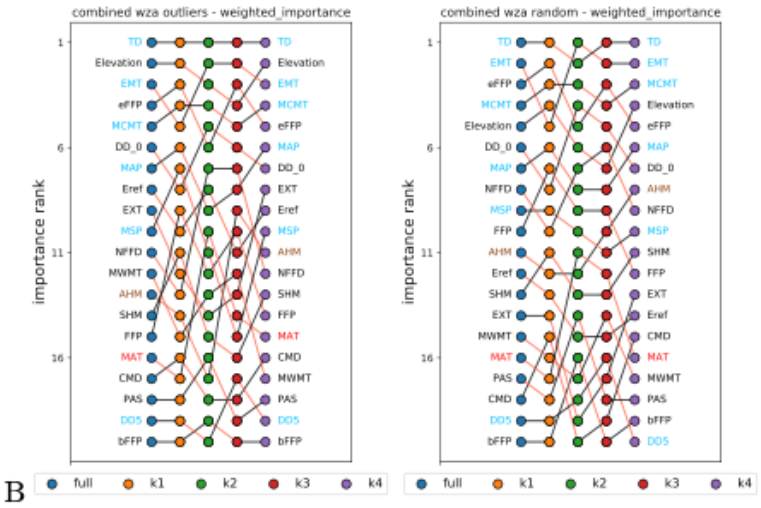


Figure S3 continued

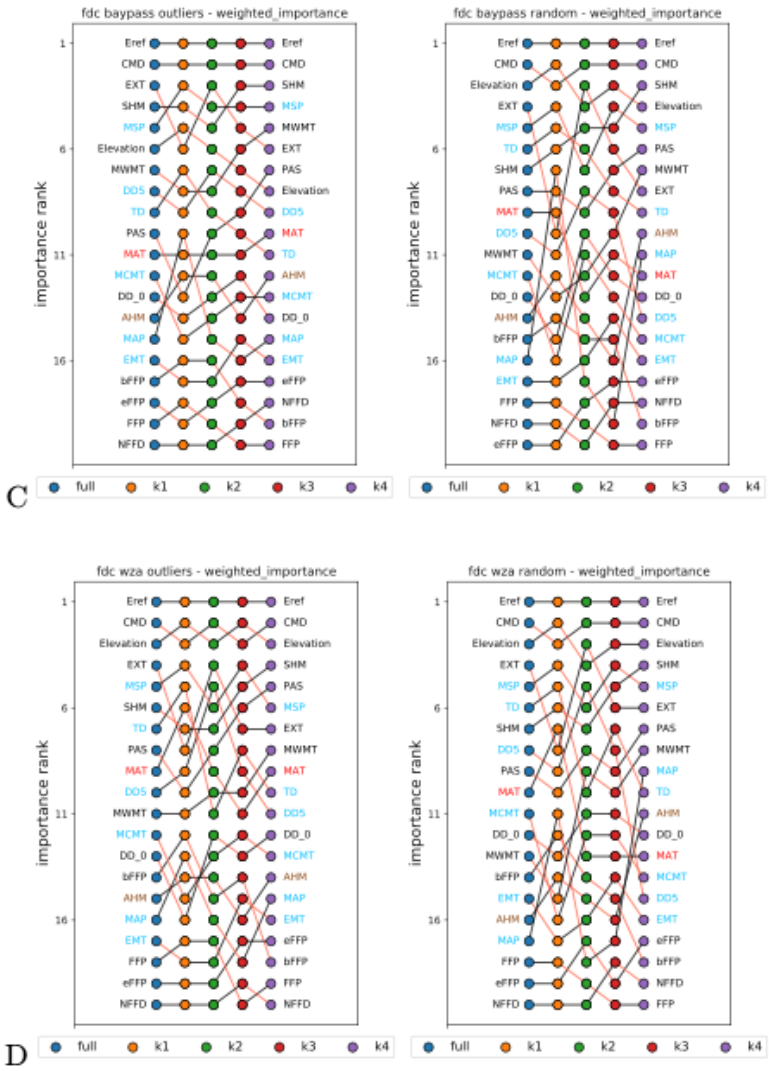
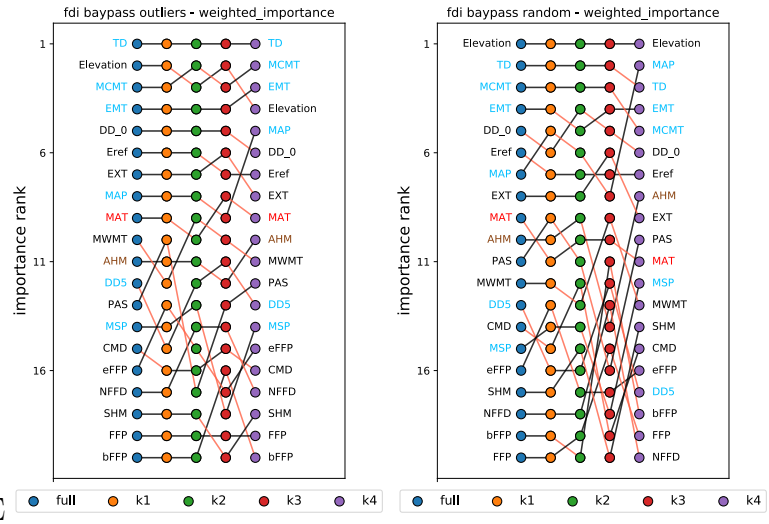
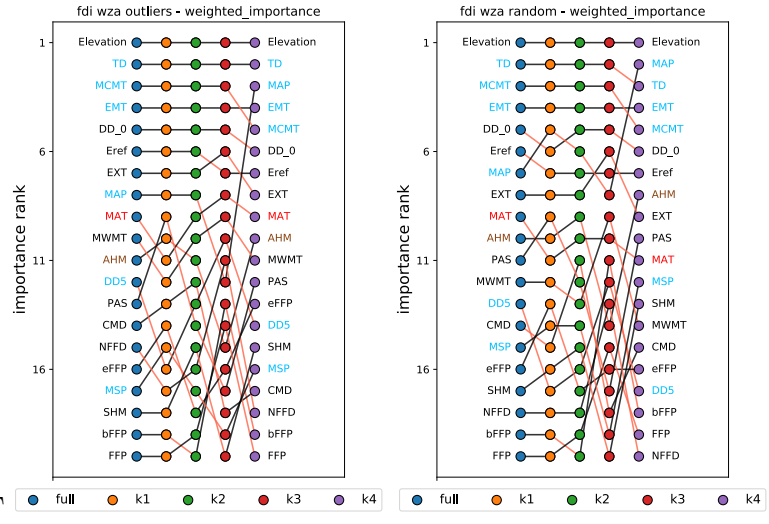


Figure S3 continued



E



F

Figure S3 continued

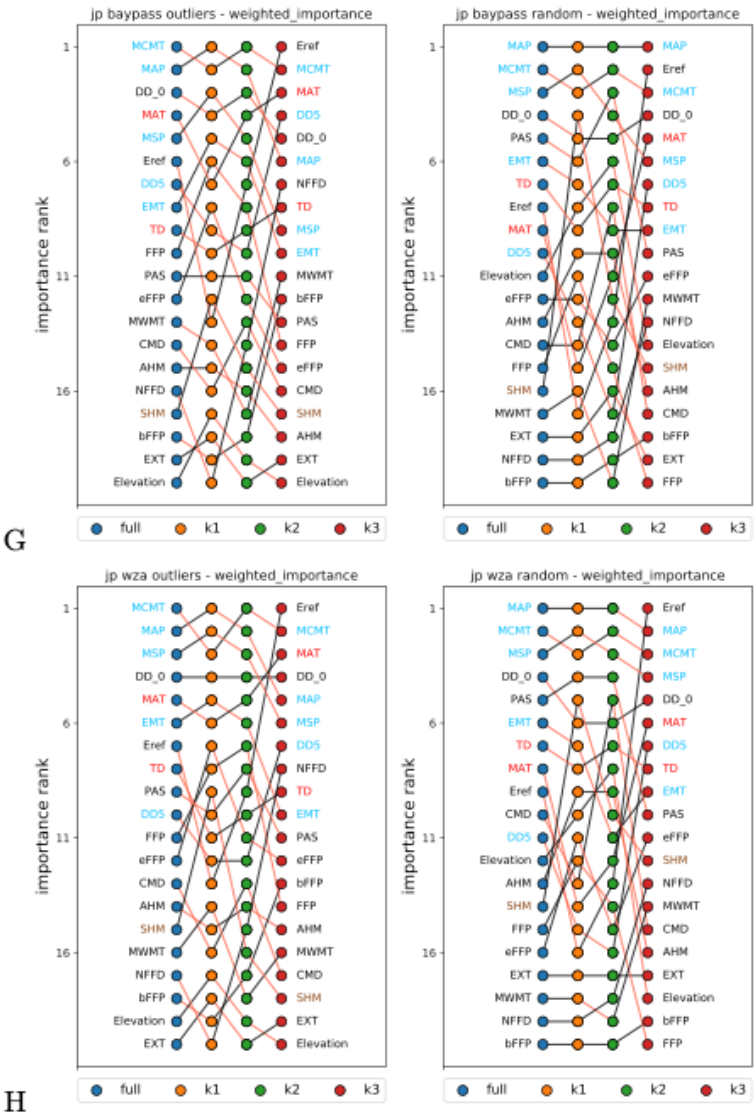
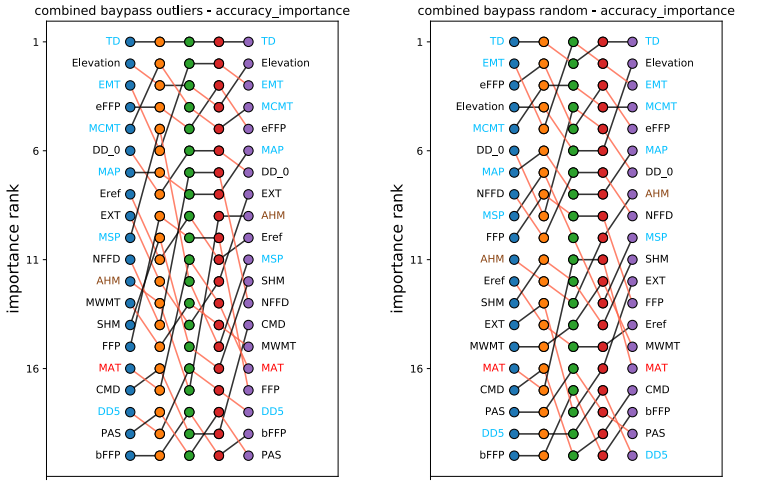
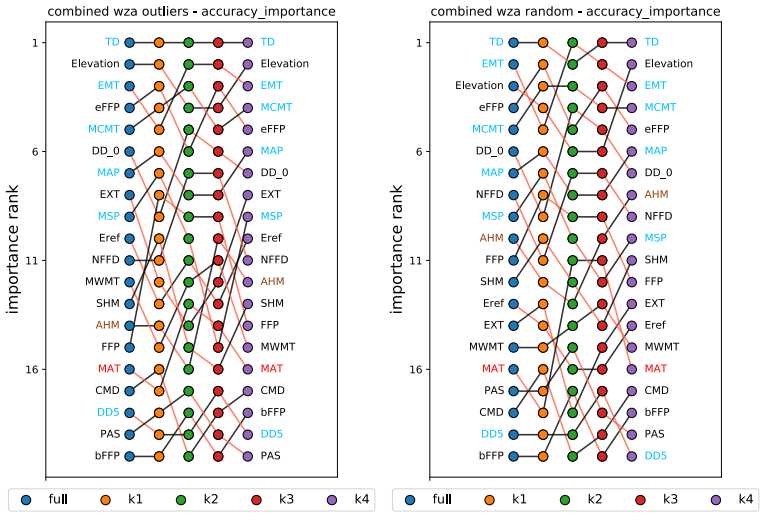


Fig. S3 Consistency of environmental weighted importance rank output from trained models of Gradient Forests *k*-fold stratified sampling (orange, green, red, and purple). Models (i.e., a set of populations) are paired every two rows, where the first row represents BayPass-related marker sets and the second row represents WZA-related marker sets (i.e., boxplot colors in Supplemental Figs. S8-S10): A-B) combined cross-variety using both varieties of Douglas-fir; C-D) FDC coastal Douglas-fir; E-F) FDI interior Douglas-fir; G-H) jp jack pine. Within each figure, red lines indicate negative changes in rank, while black lines indicate changes in rank ≥0. Environment labels are colored according to membership within 1) those used in climate-based seed transfer (CBST) guidelines in British Columbia (sky blue), 2) those explaining significant variation from provenance trials (brown), 3) those overlapping CBST and provenance trial environments (red), and 4) remaining environments (black). Code to create these figures can be found in SN 15.13.

Figure S4 - Consistency of environmental accuracy importance from Gradient Forests.

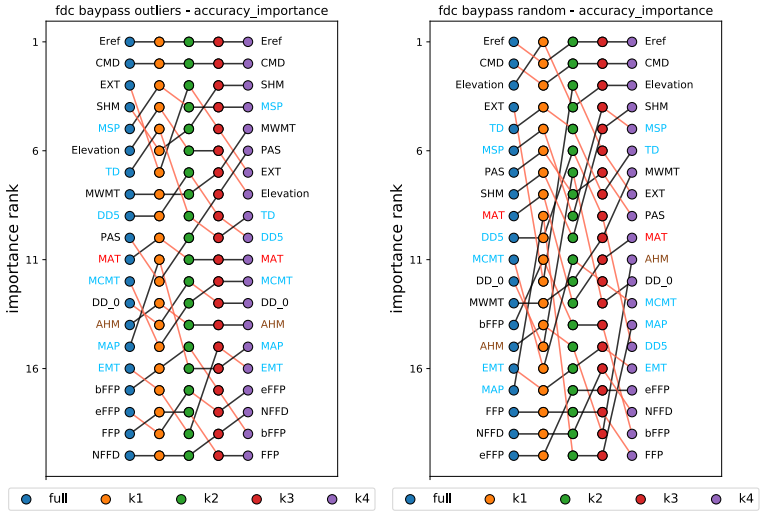


A

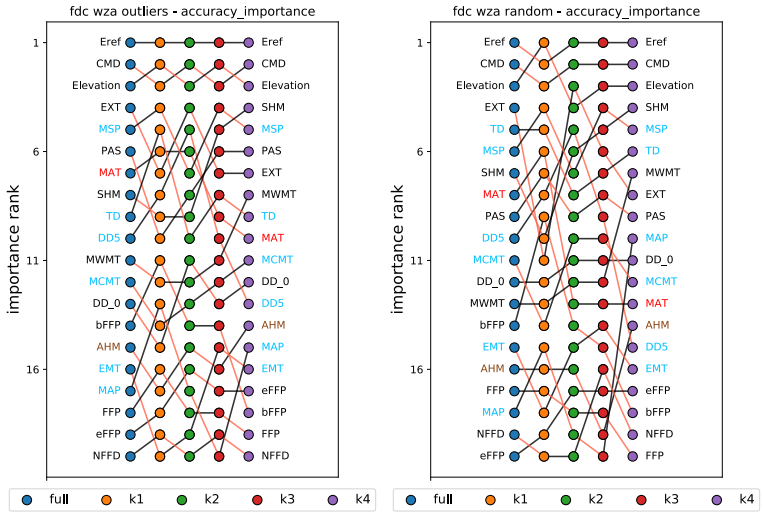


B

Figure S4 continued

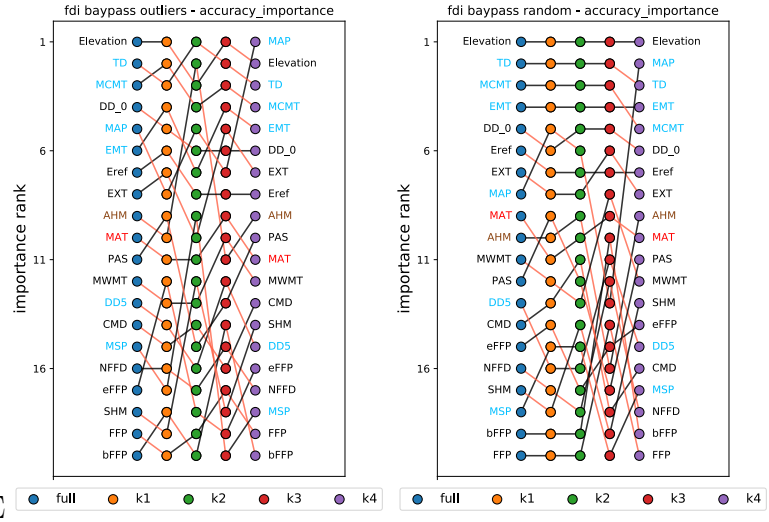


C

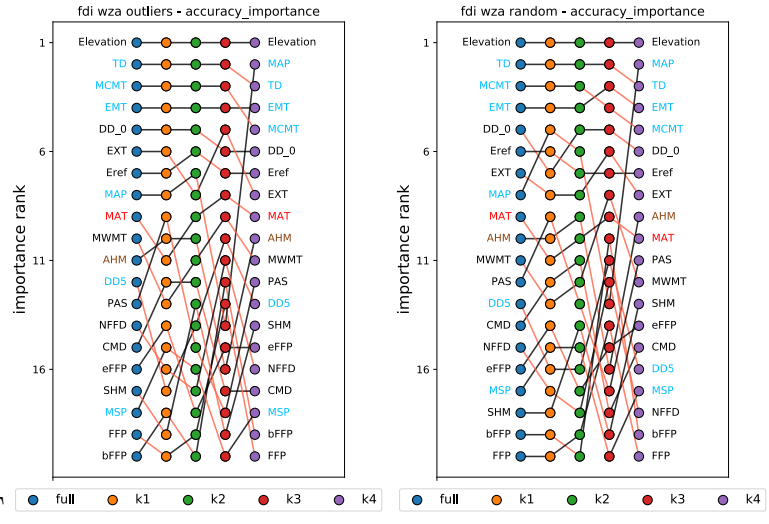


D

Figure S4 continued

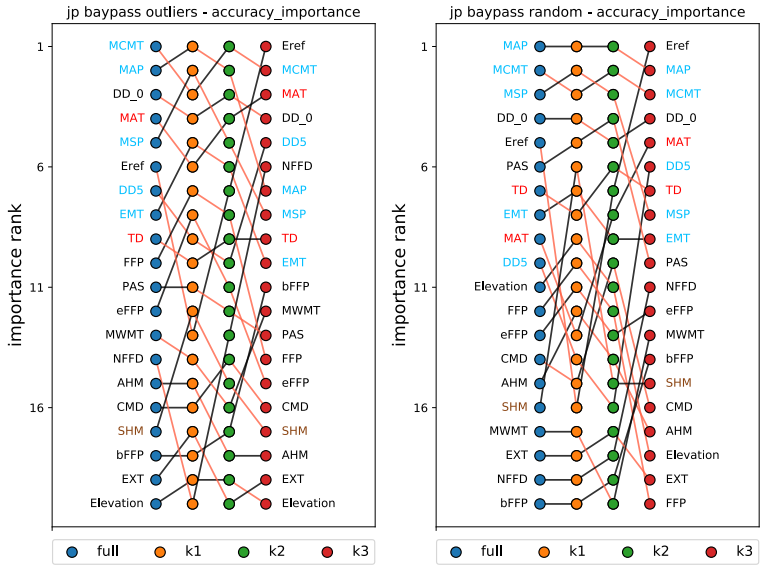


E

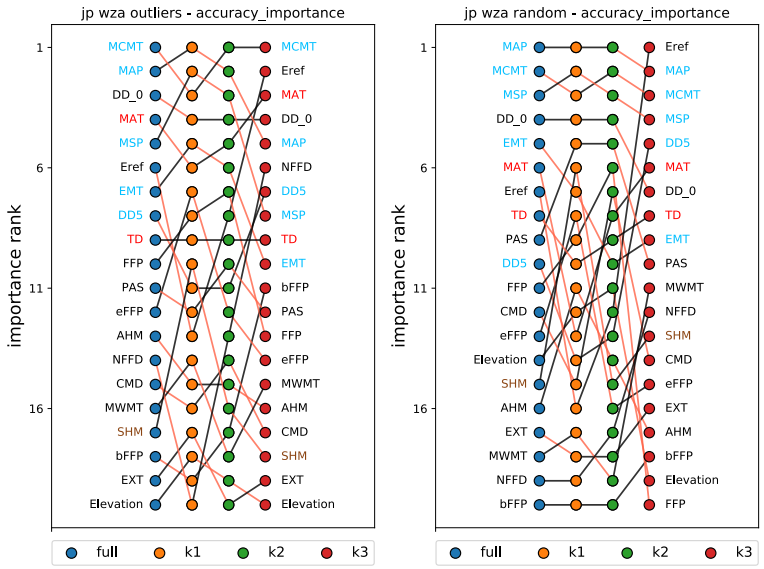


F

Figure S4 continued



G



H

Fig. S4 Consistency of environmental accuracy importance rank output from trained models of Gradient Forests using all populations (full, blue) or *k*-fold stratified sampling (orange, green, red, and purple). Models (i.e., a set of populations) are paired every two rows, where the first row represents BayPass-related locus sets and the second row represents WZA-related locus sets (i.e., boxplot colors in Supplemental Figs. S8-S10): A-B) combined cross-variety using both varieties of Douglas-fir, C-D) FDC coastal Douglas-fir, E-F) FDI interior Douglas-fir, G-H) jp jack pine. Within each figure, red lines indicate negative changes in rank, while black lines indicate changes in rank ≥0. Weighted importance is only shown once within each figure because it was unchanged across the *k*-fold iterations (including models using all populations) as well as across locus sets for a given model. Environments are colored according to membership within 1) those used in climate-based seed transfer (CBST) guidelines in British Columbia (sky blue), 2) those explaining significant variation from provenance trials (brown), 3) those overlapping CBST and provenance trial environments (red), and 4) remaining environments (black). Code to create these figures can be found in SN 15.13.

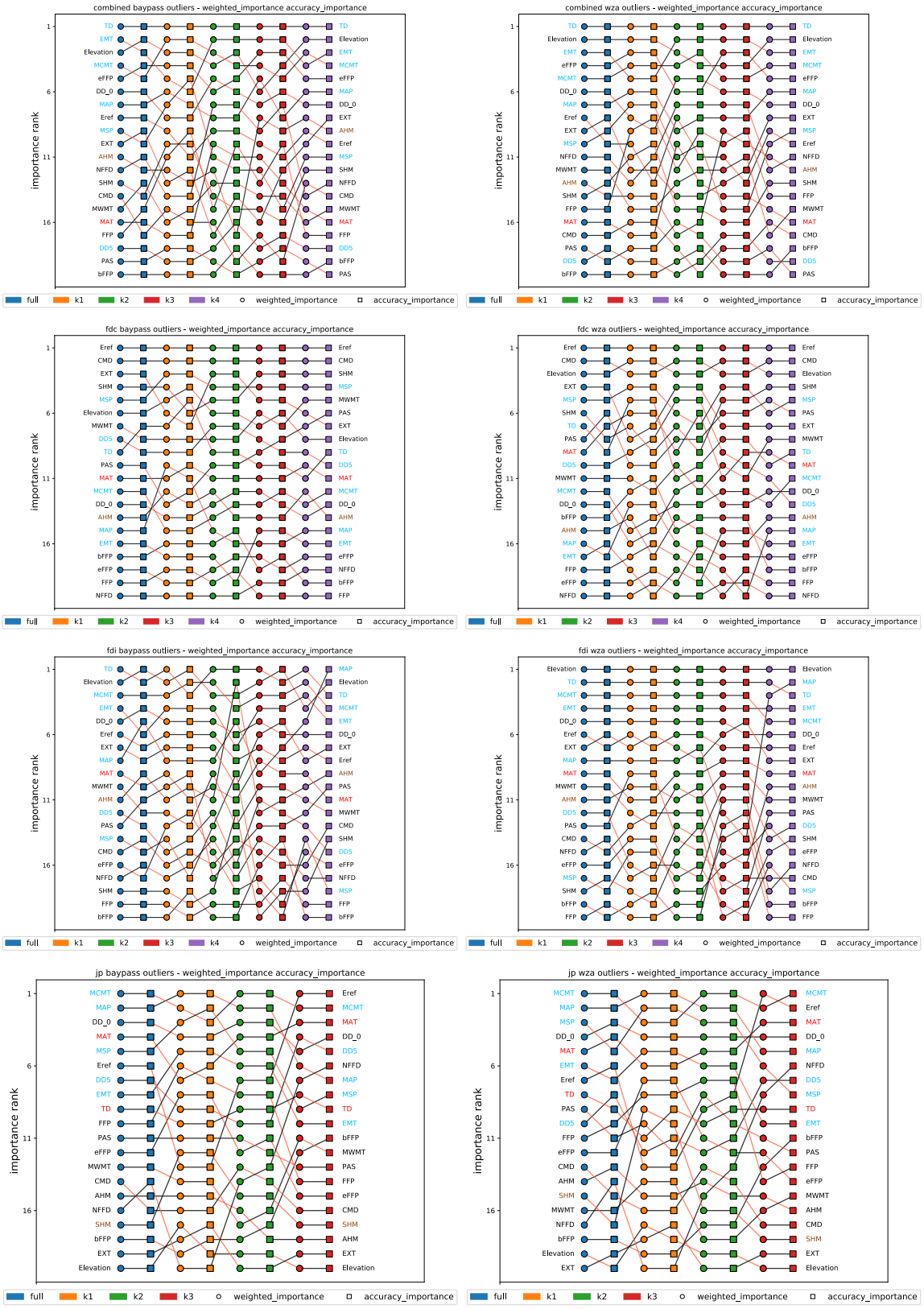


Fig. S5 Consistency of environmental accuracy importance rank output from trained models of Gradient Forests using outlier loci sets (BayPass and WZA) for A) the cross-variety model, B) FDC coastal

variety, C) FDI interior variety, D) JP jack pine. Legend (excluding importance type) and abbreviations as in Fig S4. Code to create these figures can be found in SN 15.13.

Fig S6.

Cross-variety model

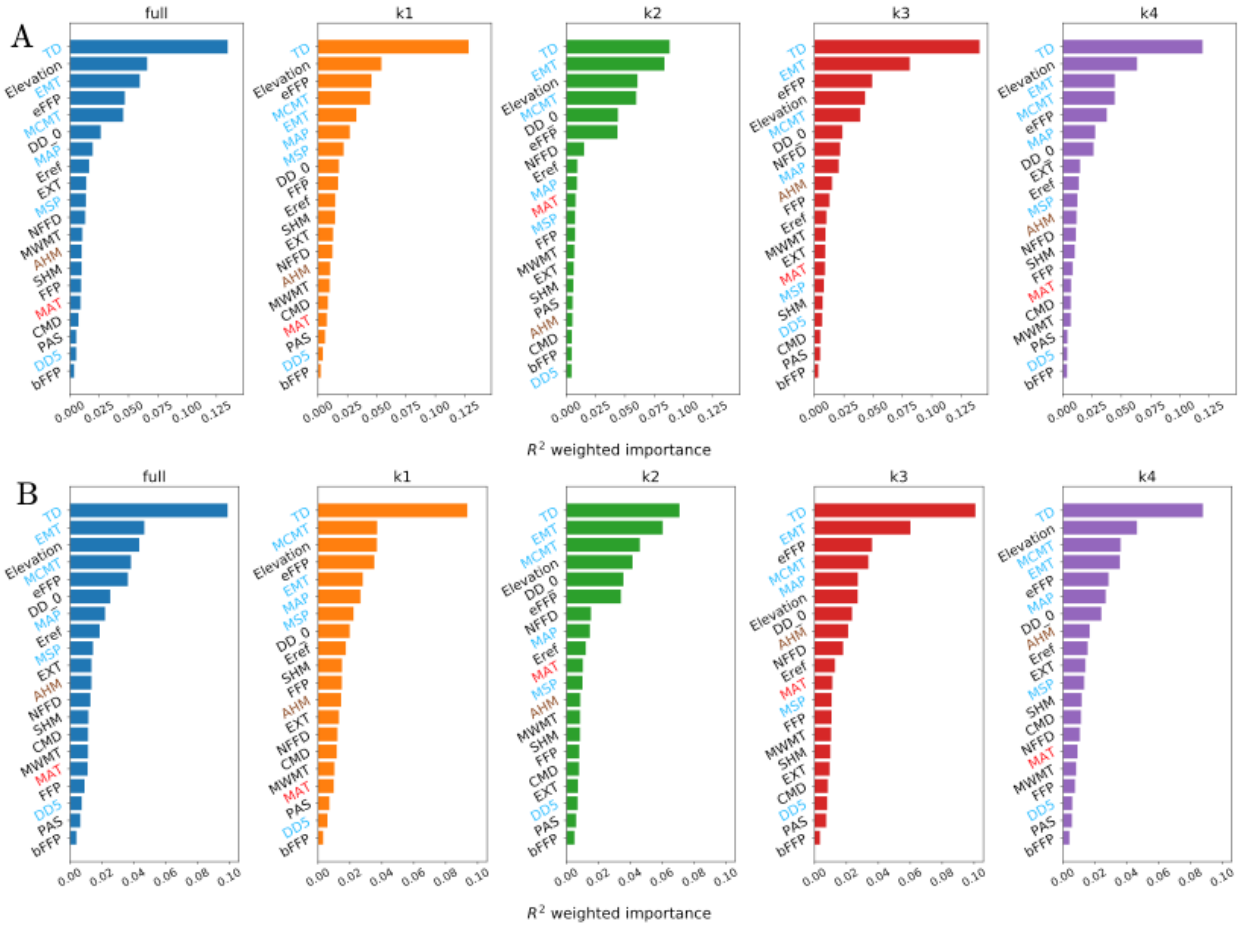


Fig S6 (cont'd)

Coastal variety model

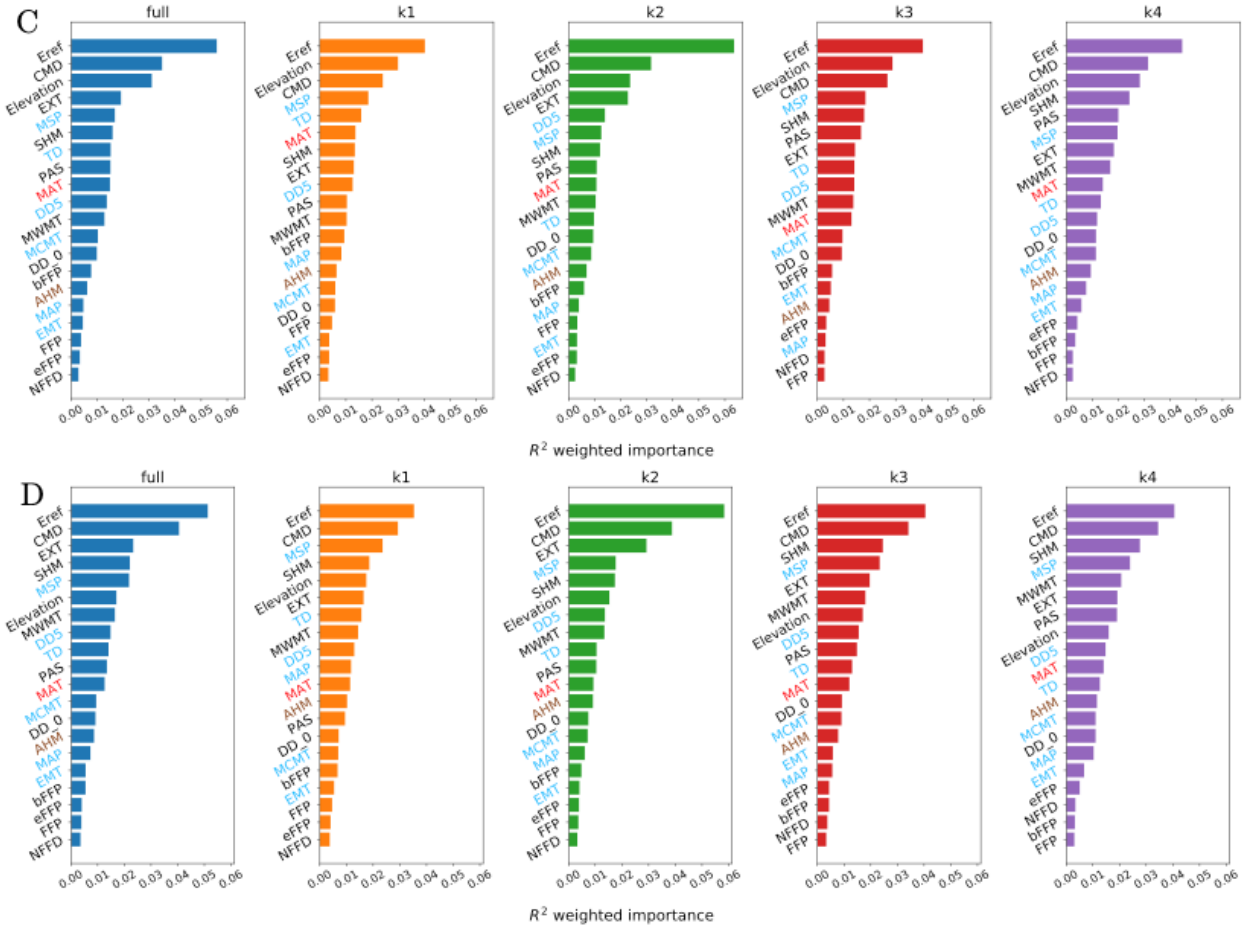


Fig S6 (cont'd)

Interior variety model

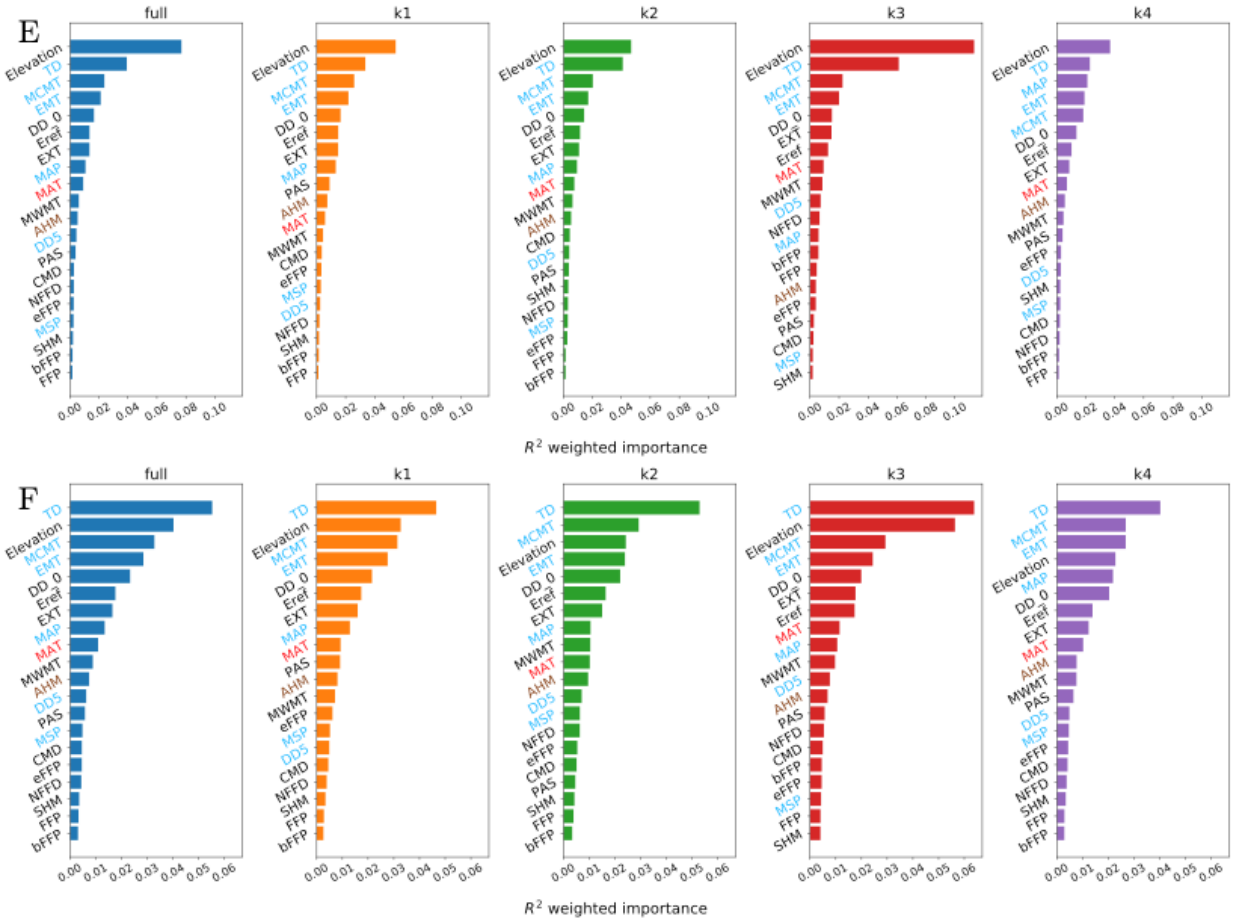


Fig S6 (cont'd)

Jack pine

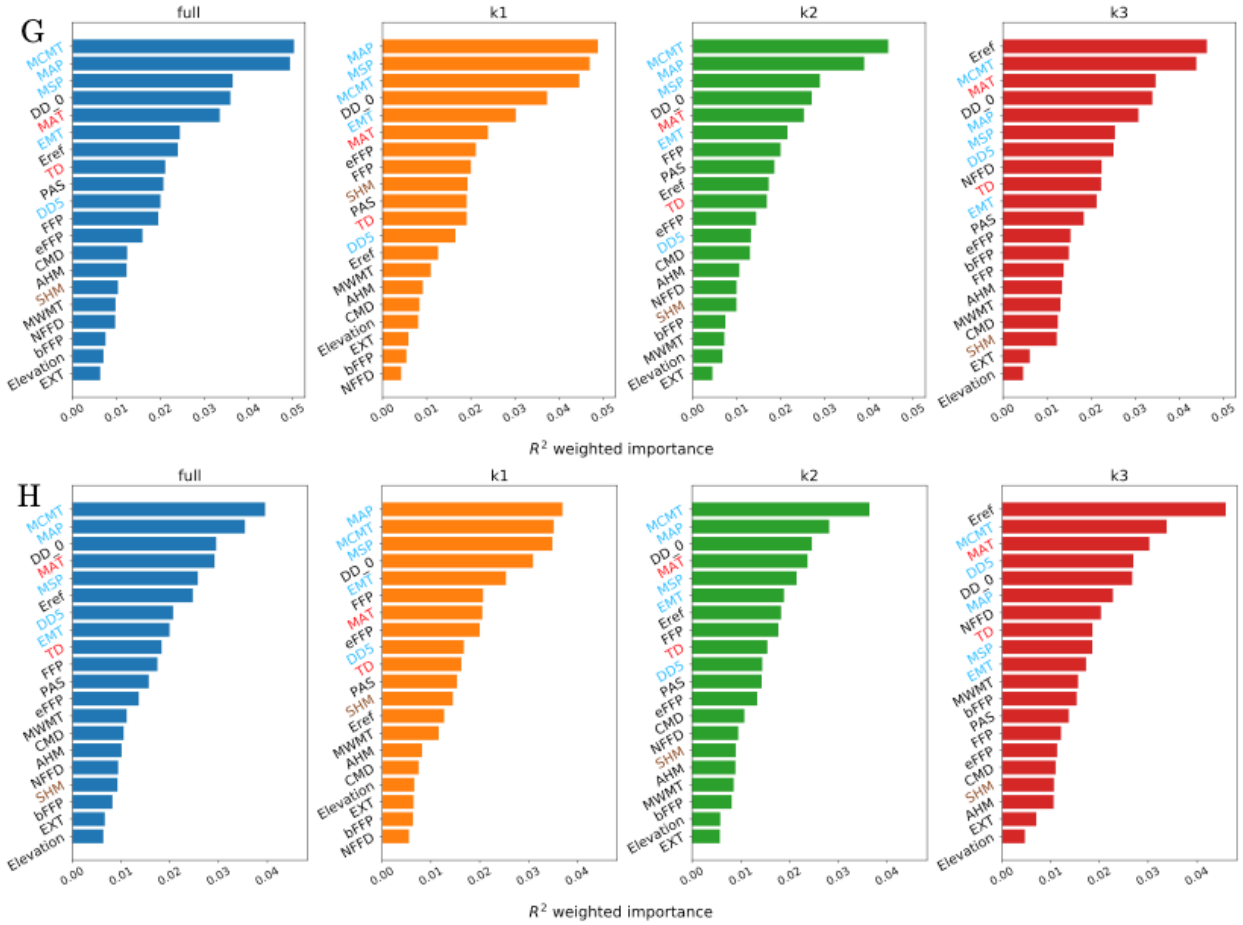


Fig S6. Weighted feature (climate) importance from Gradient Forest models trained using either WZA (A, C, E, G) or BayPass loci (B, D, F, H) using both varieties of Doulgas-fir (A, B), the coastal variety of Douglas-fir (C, D), the interior variety of Douglas-fir (E, F), or jack pine (G, H). Colors refer to the set of populations used: blue – all populations, orange, green, red, and purple – populations from stratified *k*-fold sampling. See Fig. S17 for rank visualization across loci and *k*-fold sets. Code to create this figure can be found in SN 15.13.

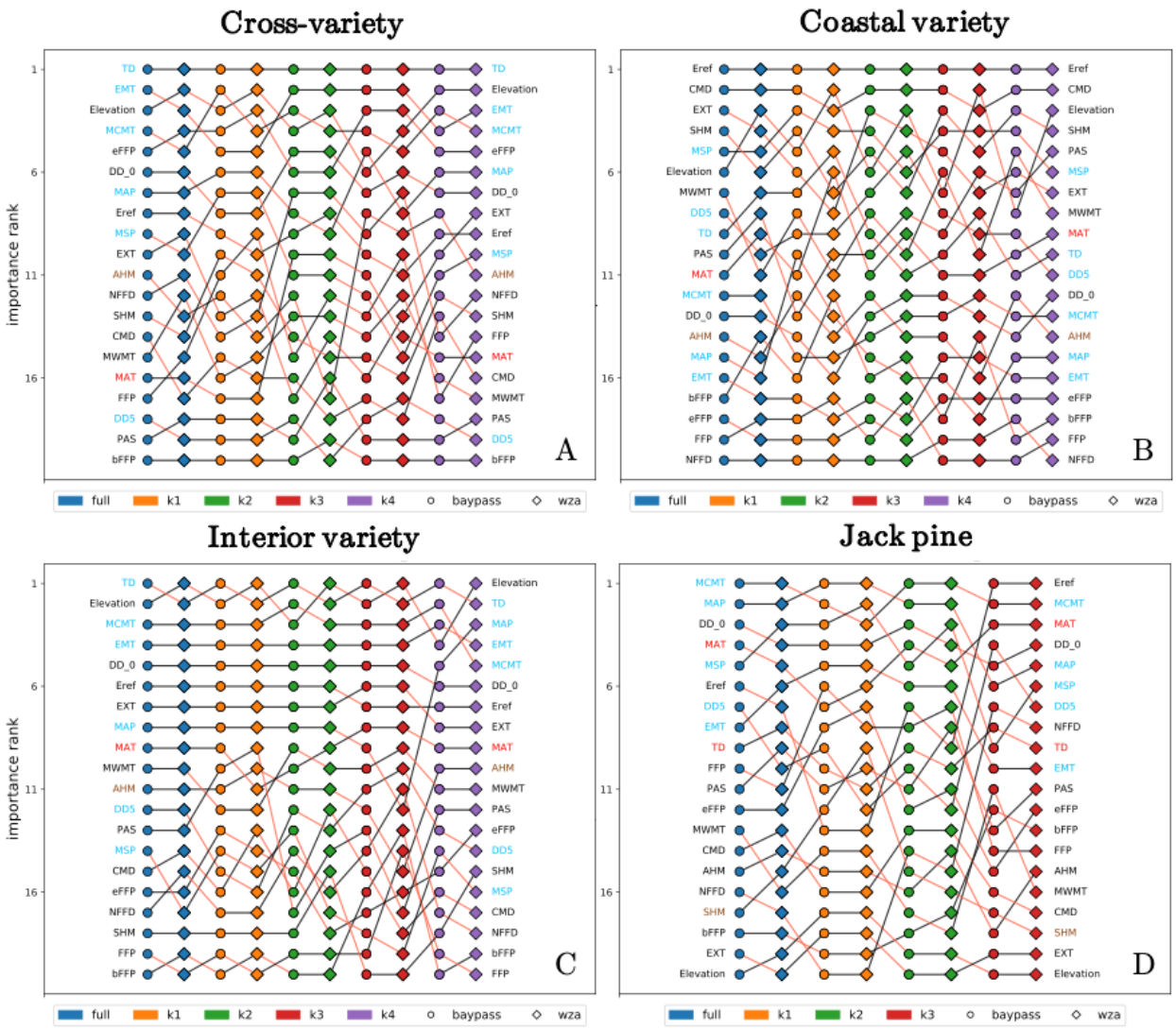


Fig. S7 Consistency of environmental weighted importance rank across models of Gradient Forests trained using outlier marker sets from BayPass (circles) or the WZA (diamonds) and either all populations (full, blue) or *k*-fold stratified sampling (orange, green, red, and purple) for A) both Douglas-fir varieties, B) coastal Douglas-fir, C) interior Douglas-fir, and D) jack pine. Within each figure, red lines indicate negative changes in rank, while black lines indicate changes in rank ≥0. Environmental labels are colored according to membership within 1) those used in climate-based seed transfer (CBST) guidelines in British Columbia (sky blue), 2) those explaining significant variation from provenance trials (saddle brown), 3) those overlapping CBST and provenance trial environments (red), and 4) remaining environments (black). Histograms of these values can be found in Supplemental Fig. S6. Figures showing rank consistency from both accuracy and weighted importances using only baypass or only WZA outliers (as well as the other individual marker sets) are found in Supplemental Fig. S3. Code used to create this figure can be found in SN 15.13.

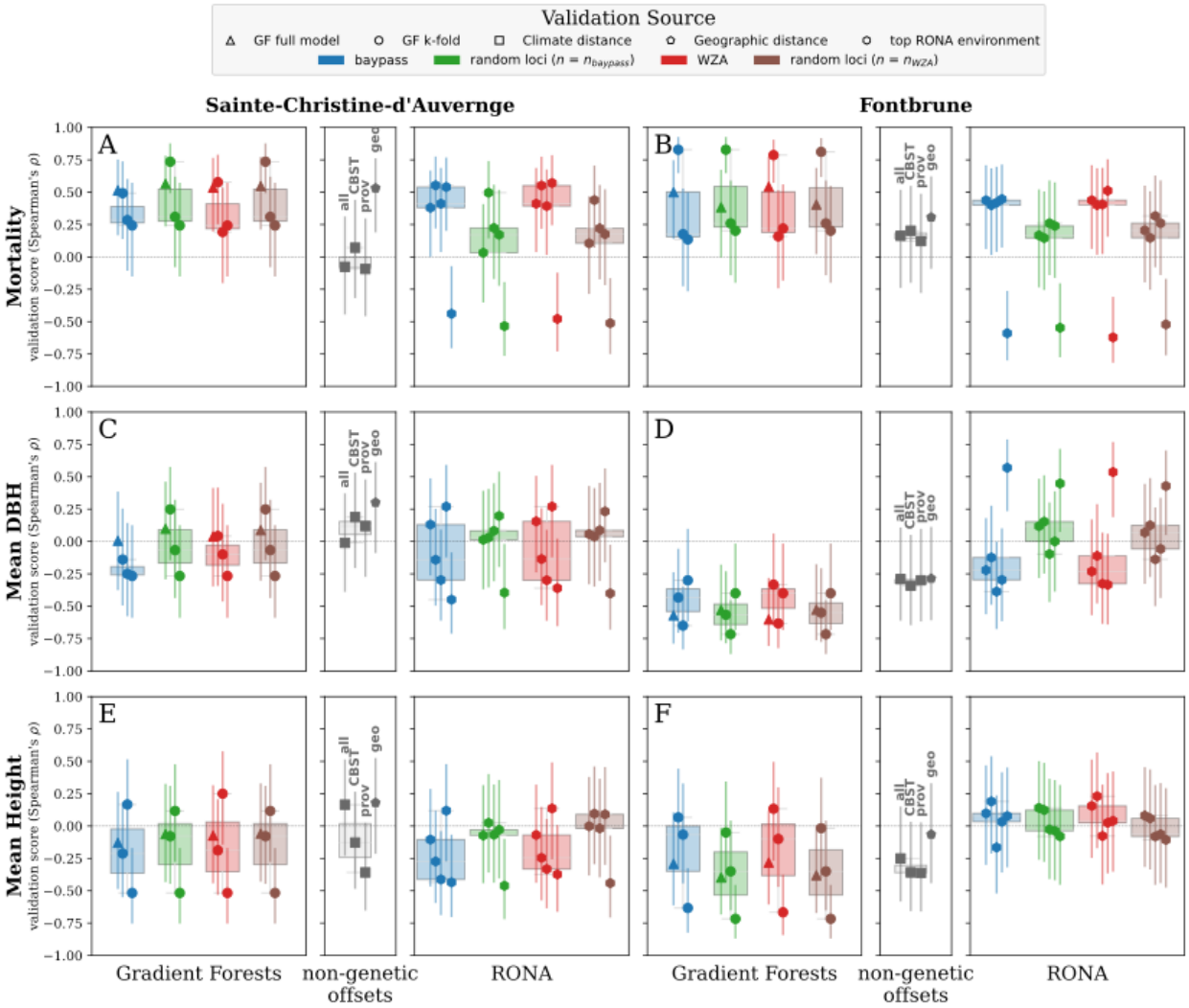


Fig. S8 Offset validation from 52-year jack pine phenotypes at Sainte-Christine-d'Auvergne (A, C, E) and Fontbrune (B, D, F) common gardens using Gradient Forests, the Risk Of Non-Adaptedness (RONA), and climate and geographic distances. Gradient Forests boxplots are indicative of cross-validation from stratified sampling using out-of-bag populations (circles) while triangles indicate performance of models trained and validated using all available populations (see Fig. 1B). RONA boxplots are indicative of the range of RONA estimates given for the top five climatic variables (hexagons) that differed significantly between source (i.e., populations used in validation) and common garden variables (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. Loci used in RONA calculations are a subset of those used in Gradient Forests that had significant linear models with the environment. Vertical bars indicate standard error estimated using a Fisher transformation (see Supplemental Text S1.5). See Extended Data Table 1 for loci counts. Whiskers extend up to 1.5x the interquartile range. Code to create these figures can be found in SN 15.14.

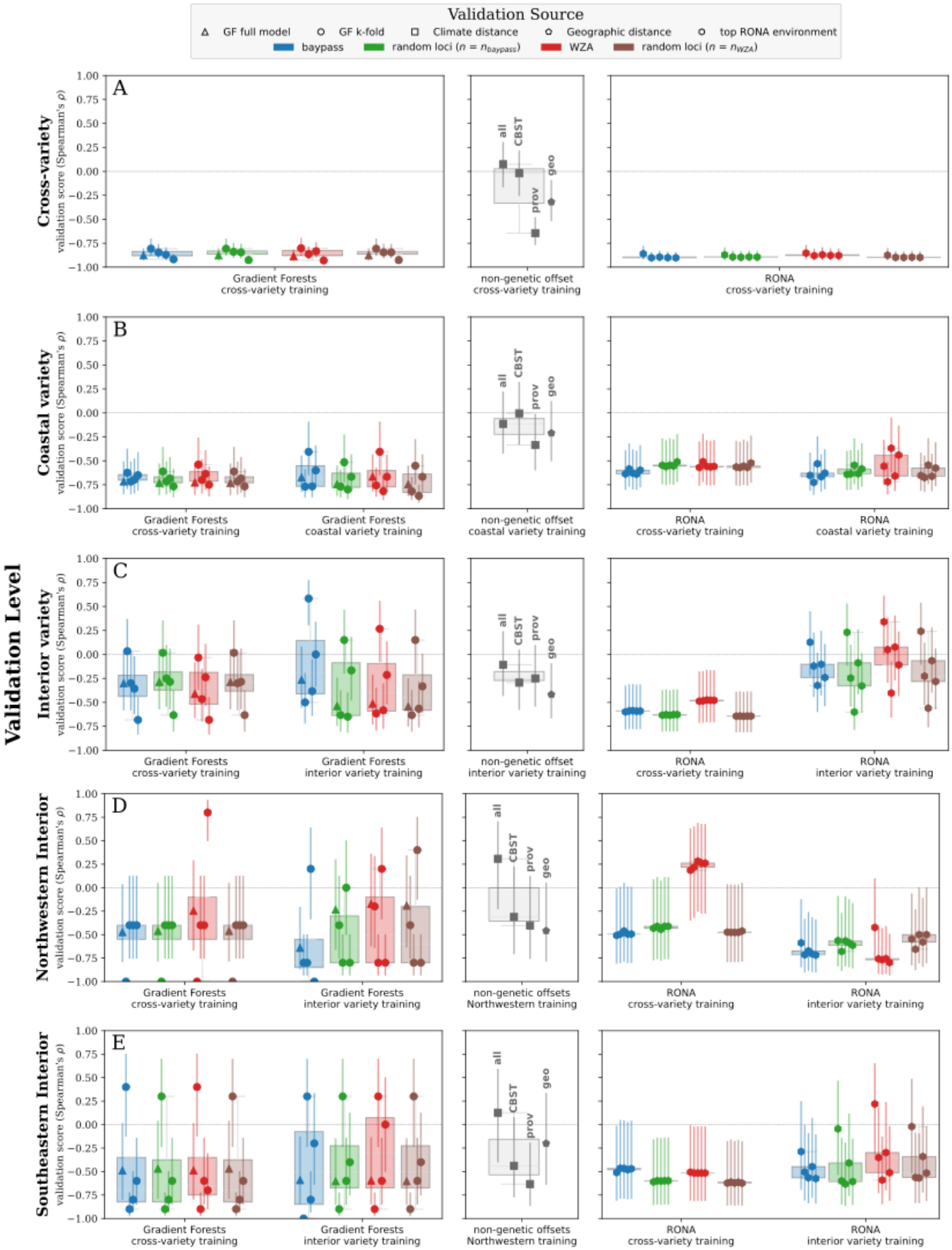


Fig. S9 Offset validation from two-year Douglas-fir height increment phenotypes at the Vancouver common garden (see Fig. 1A) using Gradient Forests (GF), the Risk of Non-Adaptedness (RONA), and

climate and geographic distances. Gradient Forests boxplots are indicative of cross-validation from stratified sampling using out-of-bag populations (circles) while triangles indicate performance of models trained and validated using all available populations. RONA boxplots are indicative of the range of RONA estimates given for the top five climatic variables (hexagons) that differed significantly between source (i.e., populations used in validation) and common garden variables (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. We used genetic hierarchy to assess accuracy inference using the validation statistic at the variety level for coastal and interior varieties, and at the subvariety level within the interior variety. X-axis labels indicate populations used in training sets, and the rows indicates populations used in validation sets. Vertical bars indicate standard error estimated using a Fisher transformation (see Supplemental Text S1.5). Locus counts in Extended Data Table 1. See Fig. S10 for similar validation using shoot biomass. See Supplemental Fig. S9 for all locus groups. Code to create these figures can be found in SN 15.14.

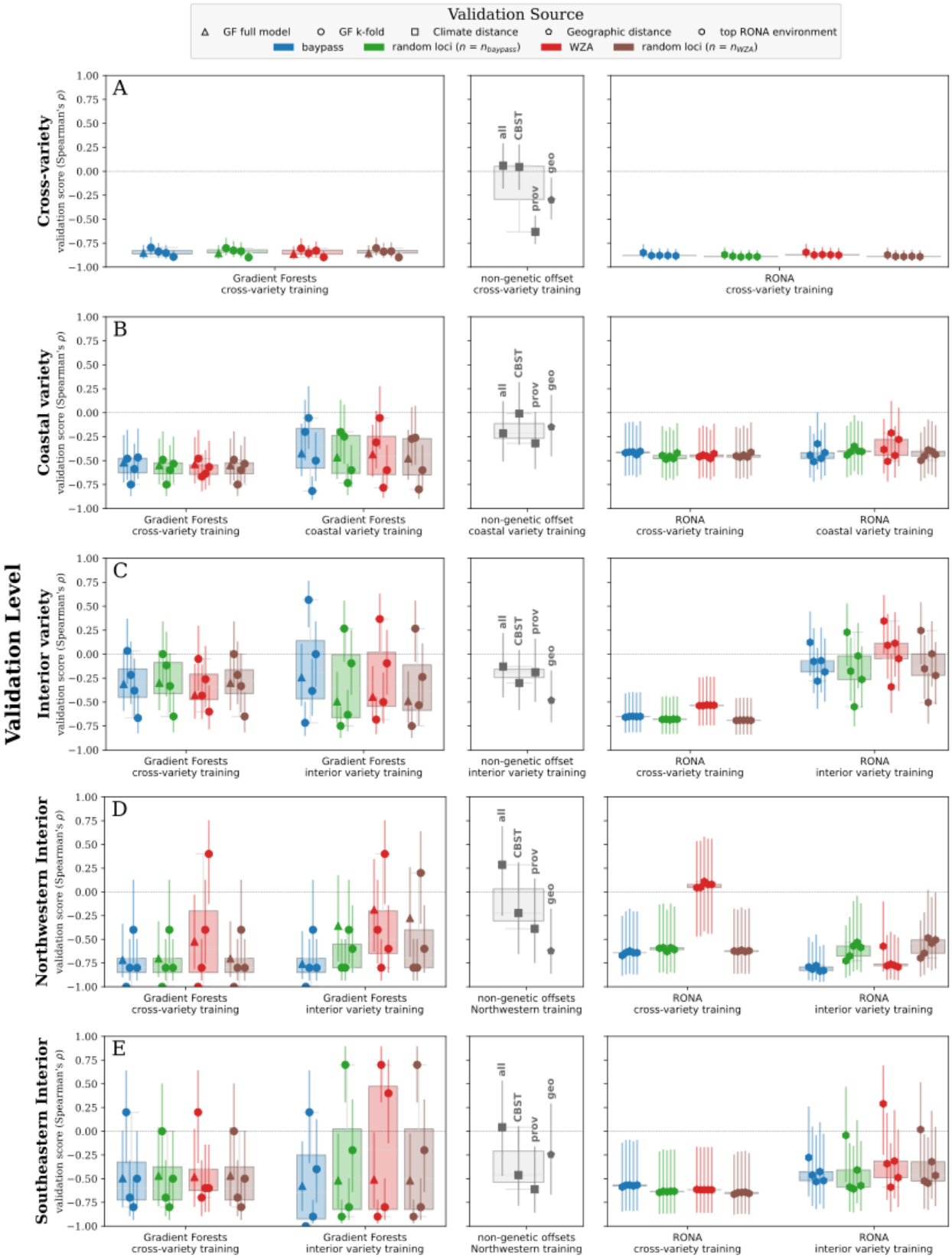

Validation Source
GF full model
GF k-fold
Climate distance
Geographic distance
top RONA environment
baypass
random loci ($n = n_{baypass}$)
WZA
random loci ($n = n_{WZA}$)
Validation Level
Cross-variety
Coastal variety
Interior variety
Northwestern Interior
Southeastern Interior
validation score (Spearman's ρ)
A
B
C
D
E
all
CBST
prov
geo
Gradient Forests cross-variety training
non-genetic offset cross-variety training
RONA cross-variety training
Gradient Forests coastal variety training
non-genetic offset coastal variety training
RONA coastal variety training
Gradient Forests interior variety training
non-genetic offset interior variety training
RONA interior variety training
non-genetic offsets Northwestern training

Fig. S10 Offset validation from two-year Douglas-fir shoot biomass phenotypes at the Vancouver common garden (see Fig. 1A) using Gradient Forests (GF), the Risk of Non-Adaptedness (RONA), and climate and geographic distances. Genetic offset boxplots and shapes are shaded with respect to marker set source. Gradient Forests boxplots are indicative of cross-validation from stratified sampling using out-of-bag populations (circles) while triangles indicate performance of models trained and validated using all available populations. RONA boxplots are indicative of the range of RONA estimates given for the top five climatic variables (hexagons) that differed significantly between source (i.e., populations used in validation) and common garden variables (see Table S1). Climate distances (squares) were calculated using 1) all climate variables, or 2) those variables used for climate-based seed transfer (CBST) in British Columbia, or 3) those explaining significant variation in provenance trials. We used genetic hierarchy to assess accuracy inference using the validation statistic at the variety level for coastal and interior varieties, and at the subvariety level within the interior variety. X-axis labels indicate populations used in training sets, and the rows indicates populations used in validation sets. Vertical bars indicate error estimated using a Fisher transformation (see Supplemental Text S1.5). Locus counts in Extended Data Table 1. See Fig. S10 for similar validation using shoot biomass. See Supplemental Fig. S9 for all locus groups. Code to create these figures can be found in SN 15.14.

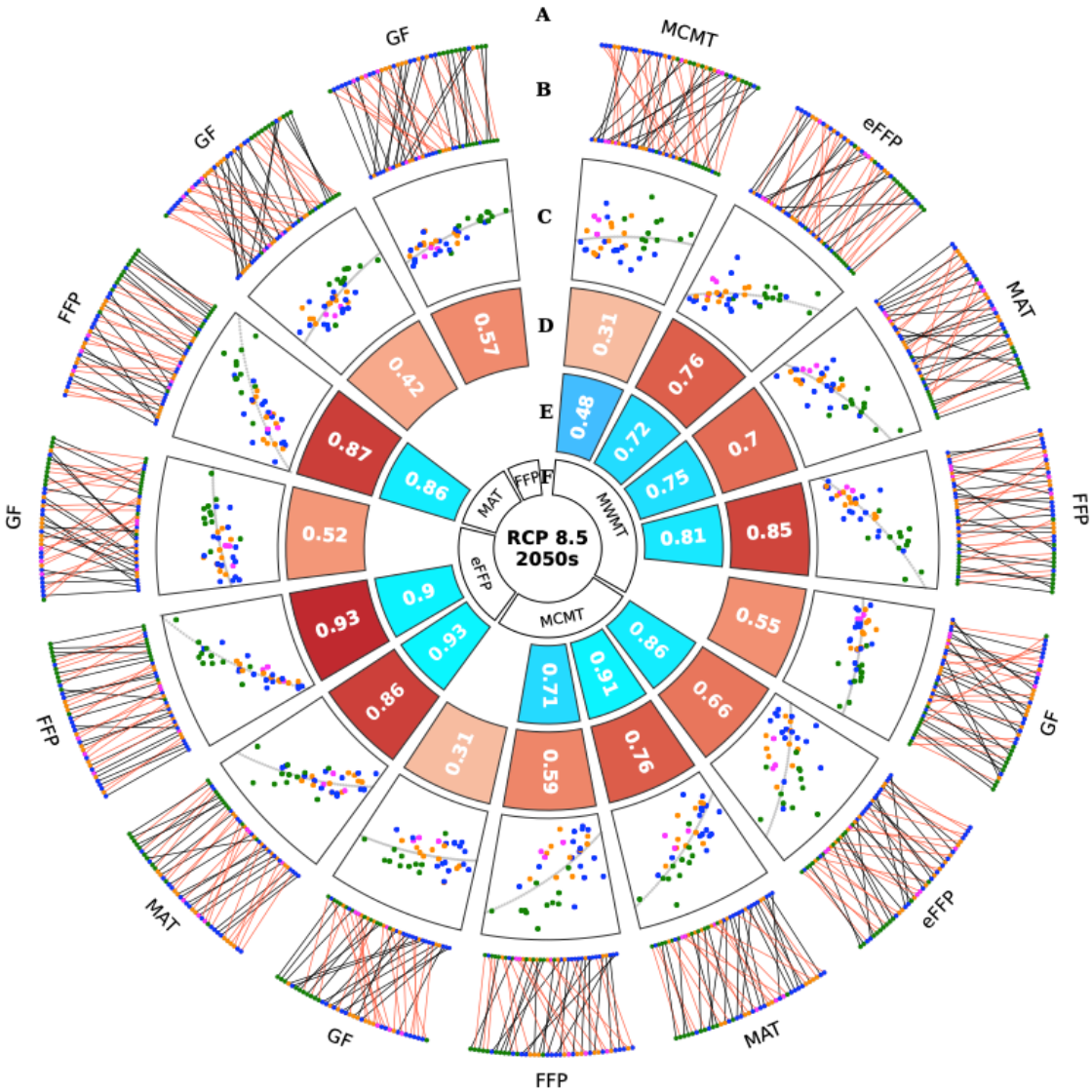


Fig S11 Comparisons of Gradient Forest (GF) and RONA top environments for projected maladaptation to RCP 8.5 2050s for jack pine. Circos plots show pairwise comparisons (A and F) between the offset predicted from the top five environments from RONA and GF (B-D) where B* shows the changes in rank between comparisons, C** shows the linear relationship (with gray line of best fit), and D shows the Spearman's rank correlation of offset predictions. E shows the correlation between future environmental variables compared in A and F. Environments used to estimate RONA were chosen based on ranking p-values from paired *t*-tests between current and future climate across all populations that went into a given model (this is why some environments differ between cross-variety and variety-specific models). Populations in B, C, and G are colored as in Fig. 1. Red lines in B are indicative of negative changes in rank between F and A. Code to create this figure can be found in SN 15.17. Analogous figures created using climate models RCP4.5 2080s, RCP4.5 2050s, and RCP8.5 2080s show similar patterns and are not shown except within SN 15.17.

* outer ranks are from A, inner ranks are from F

** y-axis is from A and x-axis is from F

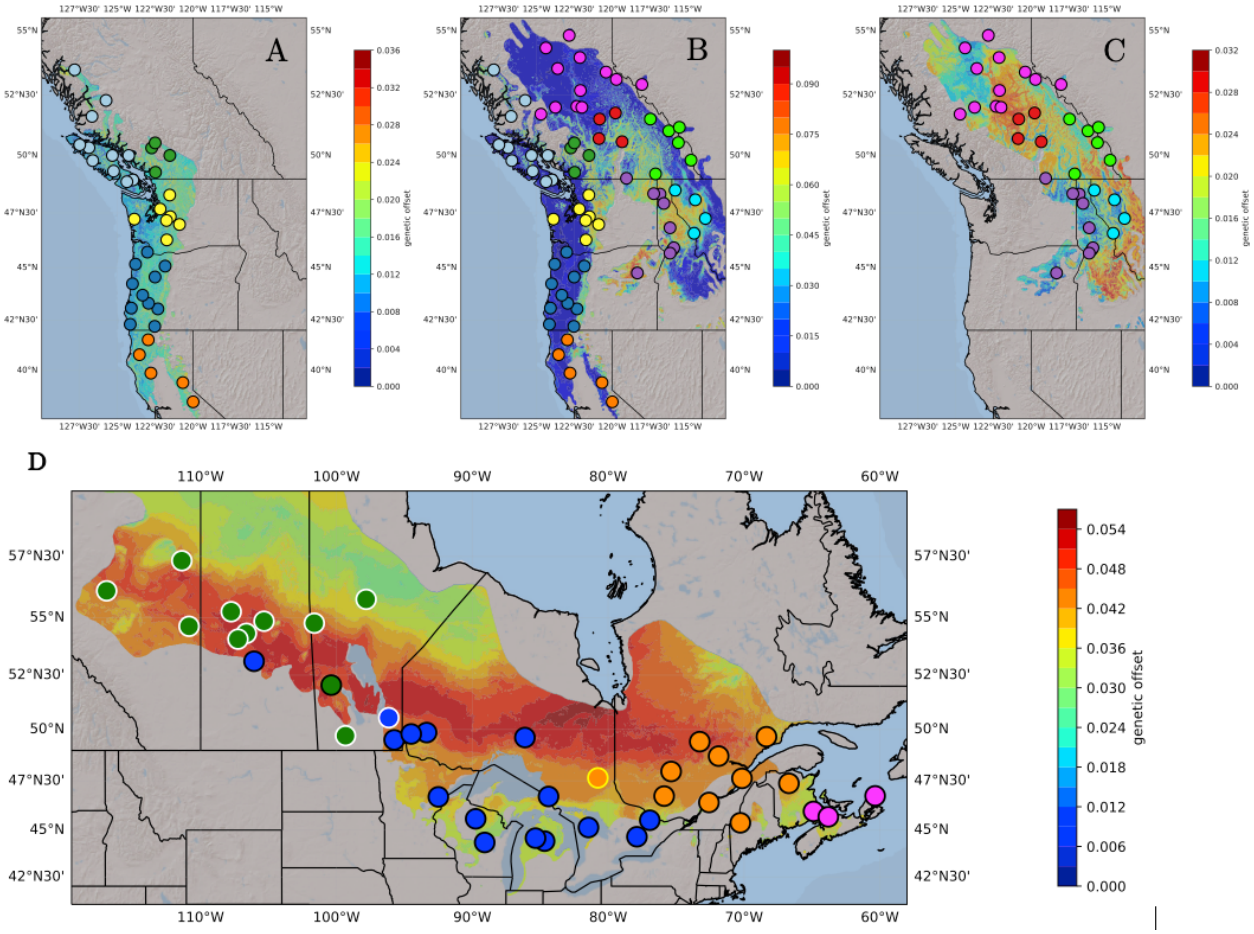


Fig S12 These figures are identical to maps shown in Figs. 4-5, except with the overlay of sample populations. Note color of population corresponds to Fig. 1 and does not reflect values of predicted offset. Code to create these figures can be found in SN 15.18.

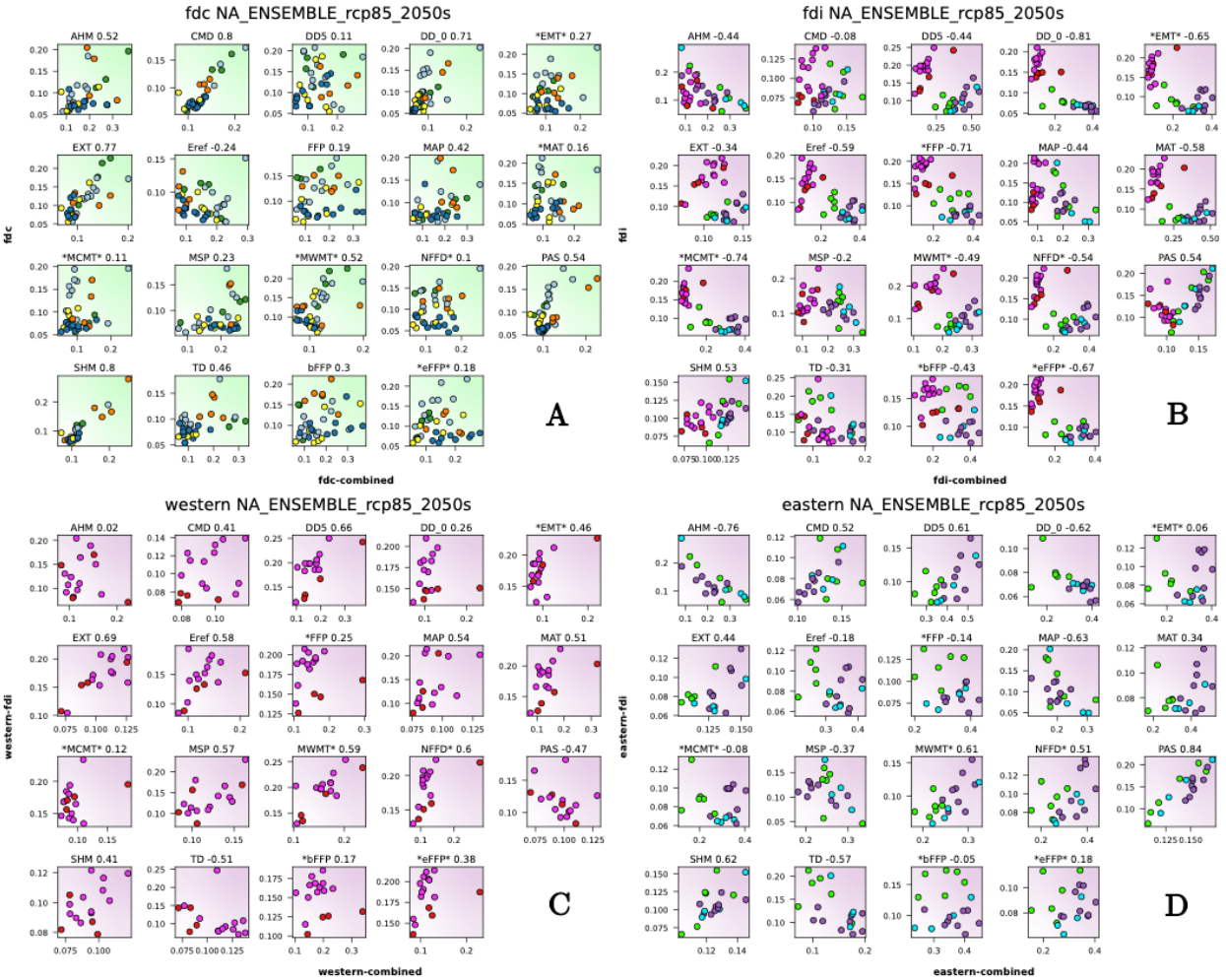


Fig S13 Relationship between Douglas-fir RONA offset predictions to future climates (RCP 8.5 2050s) from variety-specific models and cross-variety models (A, B) or for genetic groups in the interior variety-specific models with the same genetic groups taken from the cross-variety model (C, D). Environmental names are at the top of each plot – an asterisk on the left side of an environmental name indicates that the variable was significantly different between current and future climates at the variety level; similarly an asterisk on the right side of an environmental variable indicates a significant difference in climate between current and future conditions using populations from both varieties. Background color of each figure indicates population membership to either the coastal variety (lime green) or the interior variety (purple). FDI = interior Douglas-fir; FDC = coastal Douglas-fir; eastern and western refer to the genetic groups within interior Douglas-fir; combined refers to the cross-variety model. Compare with Fig. 5 in the main text comparing similar datasets output from Gradient Forests. Code to create these figures can be found in SN 15.17.

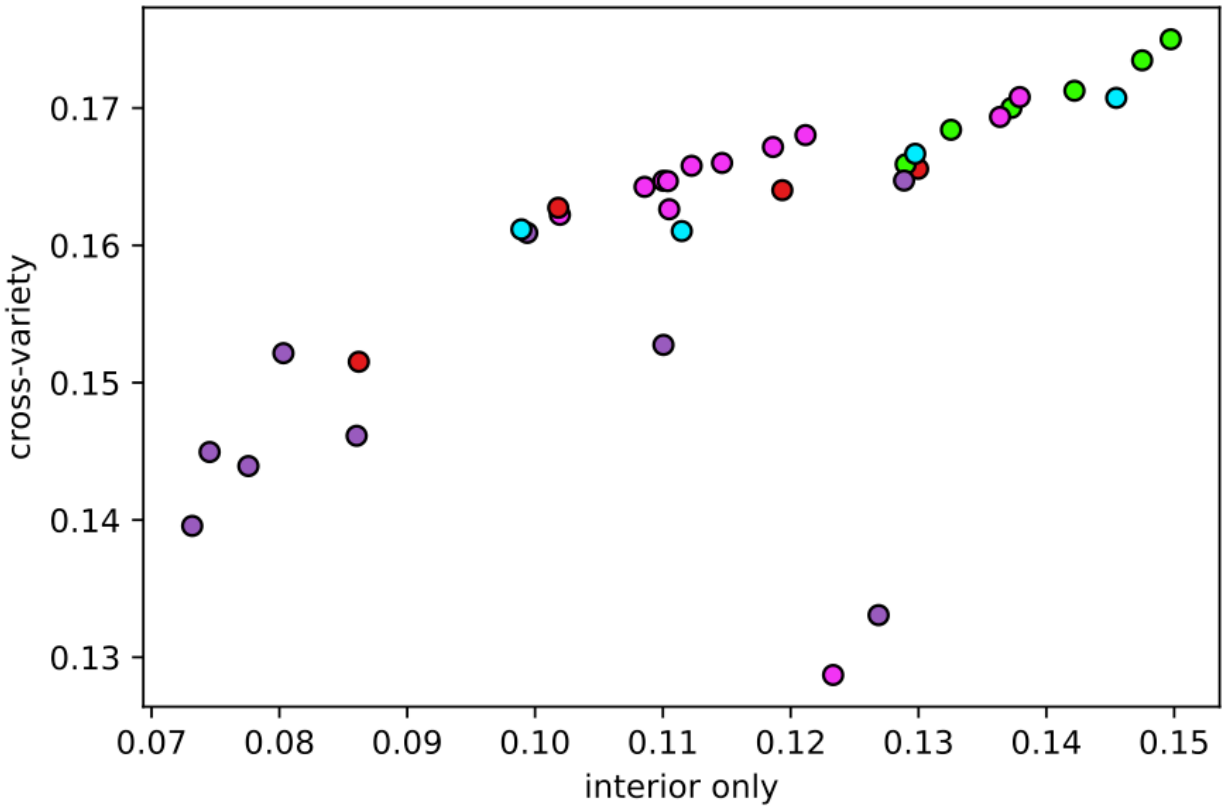


Fig. S14 Gradient Forest (GF) model agreement as to whether the northwestern or southeastern genetic group of interior Douglas-fir is most maladapted to the climate of the Vancouver common garden used for validation. Note that the same GF models were in disagreement as to whether either genetic group was most maladapted to RCP 8.5 2050s. Models shown are the cross-variety (y-axis) and interior-only (x-axis) models of GF trained using WZA loci and all populations. Population colors are as in Fig. 1 of the main text. Code used to create this figure can be found in SN 15.20.

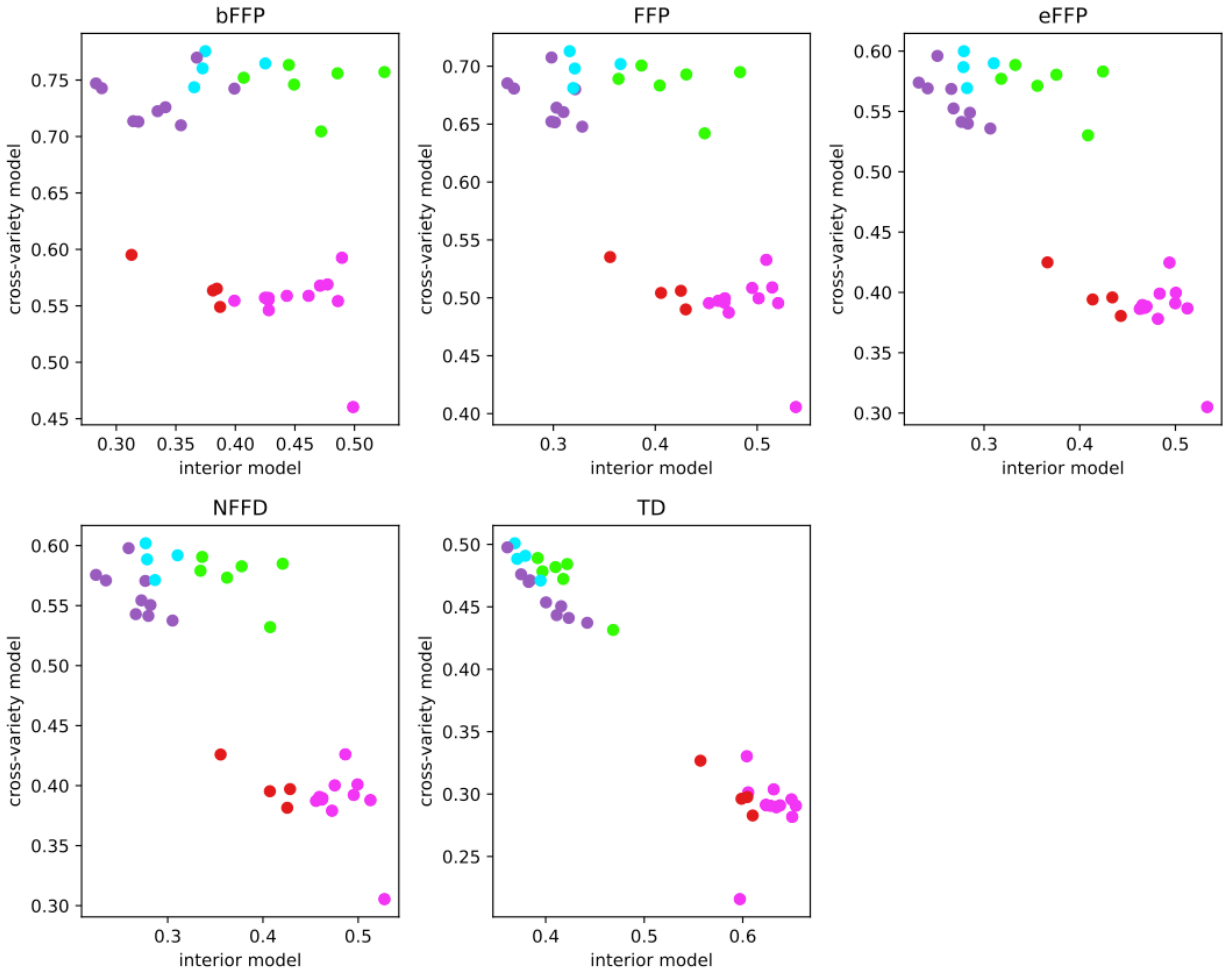


Fig. S15 RONA model comparison of cross-variety (y-axes) and interior-only (x-axes) offset projected to the Vancouver common garden. Population colors are as in Fig. 1 of the main text. Code used to create this figure can be found in SN 15.20.

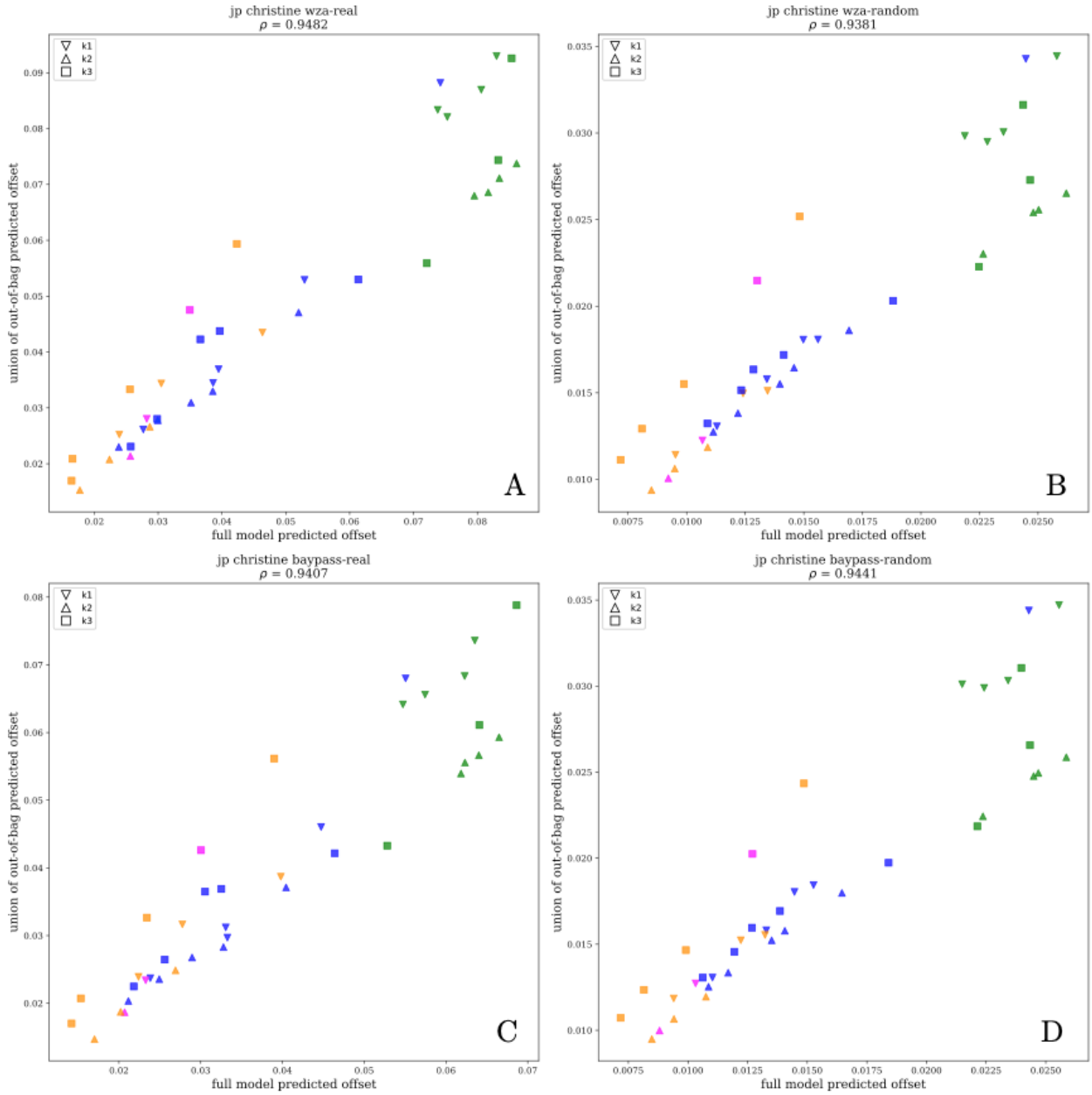


Fig S16 Relationship between estimated offset to the Sainte Christine-d'Auvernge provenance trial site predicted from full models (i.e., trained using all populations) of Gradient Forests (x-axes) and the union of out-of-bag offset estimations across *k*-fold models of Gradient Forests (y-axes) for jack pine. Models shown are those using either outlier (A, C) or random loci (B, D) with respect to marker sets from either the WZA (A, B) and baypass (C, D) genotype-environment association analyses. Points are color coded with respect to groups from Fig. 1 of the main text, and shapes corresponding to *k*-fold membership. Code to create these figures can be found in SN 15.21.

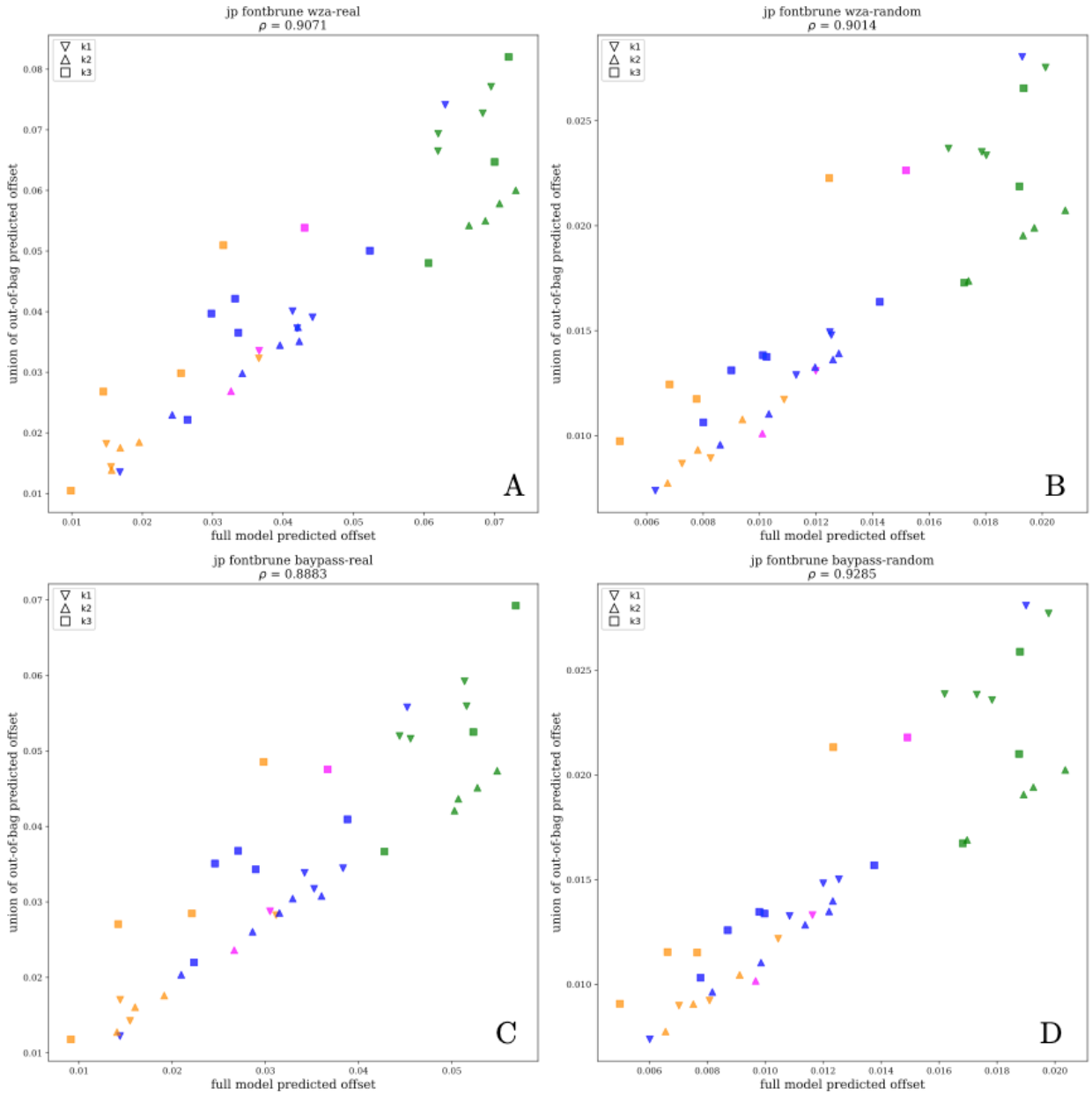


Fig S17 Relationship between estimated offset to the Fontbrune provenance trial site predicted from full models (i.e., trained using all populations) of Gradient Forests (x-axes) and the union of out-of-bag offset estimations across *k*-fold models of Gradient Forests (y-axes) for jack pine. Models shown are those using either outlier (A, C) or random loci (B, D) with respect to the WZA (A, B) and baypass (C, D) genotype-environment association analyses. Points are color coded with respect to groups from Fig. 1 of the main text, and shapes corresponding to *k*-fold membership. Code to create these figures can be found in SN 15.21.

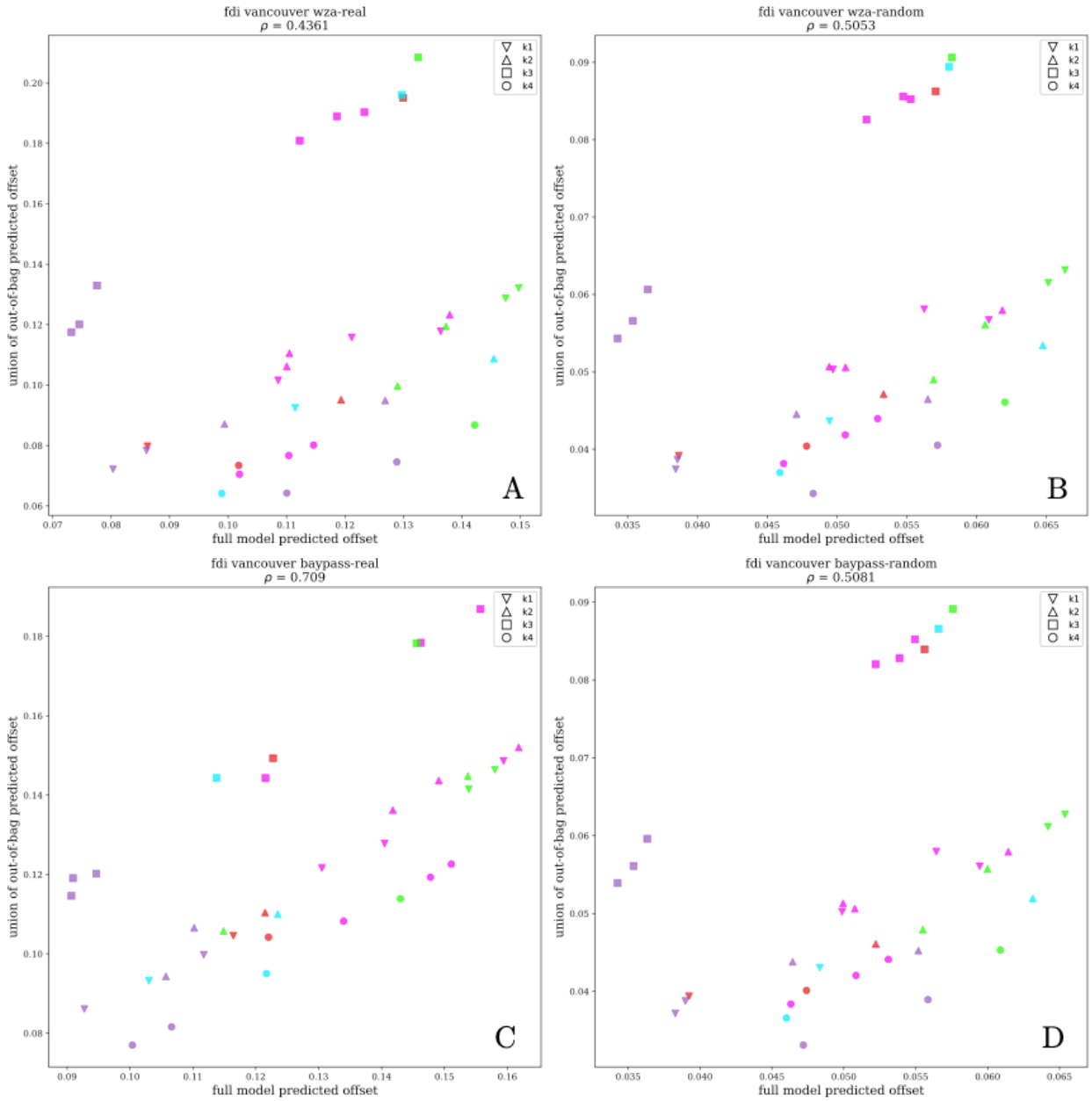


Fig S18 Relationship between estimated offset to the Vancouver common garden site predicted from full models (i.e., trained using all populations) of Gradient Forests (x-axes) and the union of out-of-bag offset estimations across *k*-fold models of Gradient Forests (y-axes) for the interior variety models of Douglas-fir. Models shown are those using either outlier (A, C) or random loci (B, D) with respect to marker sets from either the WZA (A, B) and baypass (C, D) genotype-environment association analyses. Points are color coded with respect to groups from Fig. 1 of the main text, and shapes corresponding to *k*-fold membership. Code to create these figures can be found in SN 15.21.

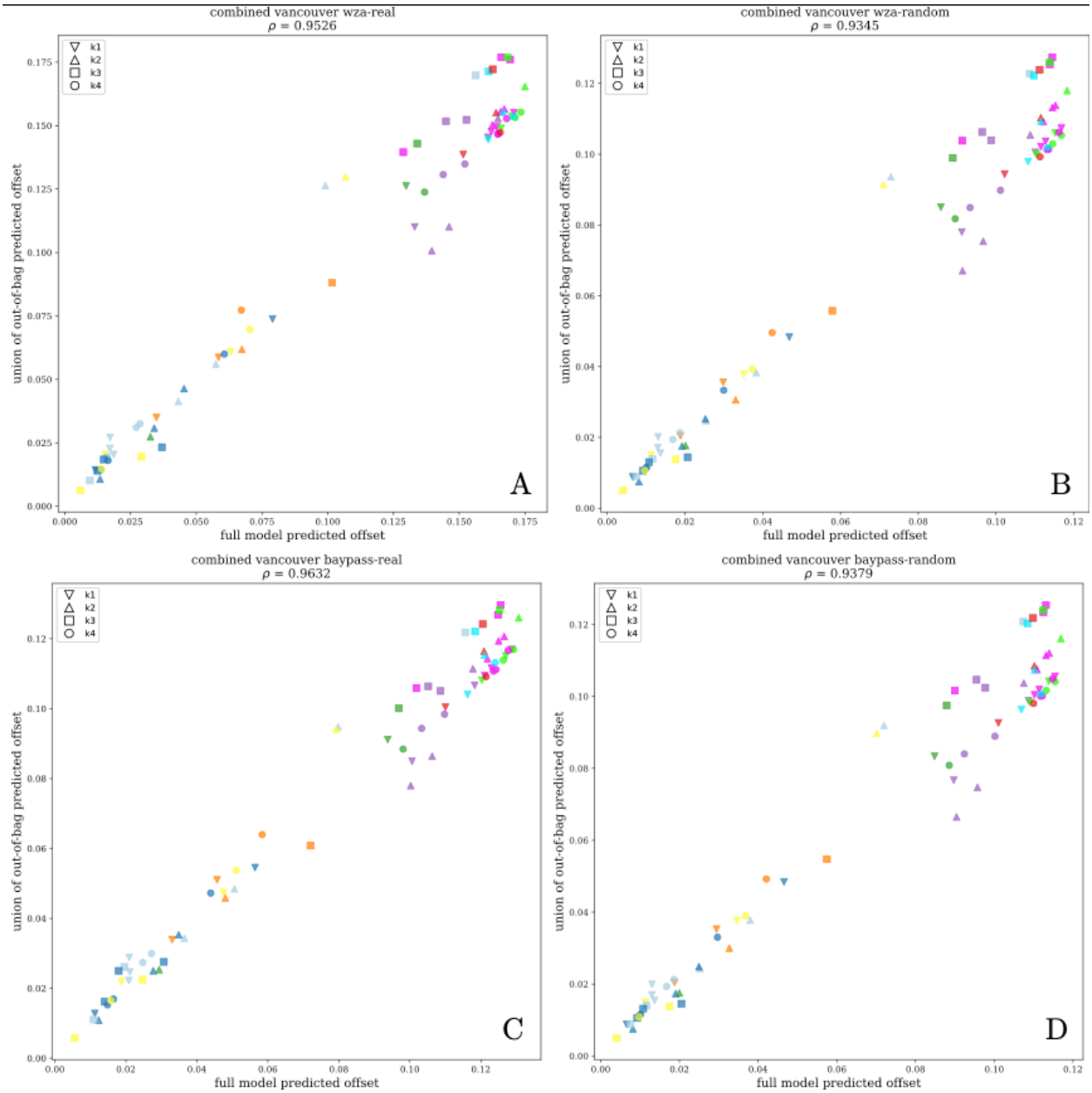


Fig S19 Relationship between estimated offset to the Vancouver common garden site predicted from full models (i.e., trained using all populations) of Gradient Forests (x-axes) and the union of out-of-bag offset estimations across *k*-fold models of Gradient Forests (y-axes) for the cross-variety models of Douglas-fir. Models shown are those using either outlier (A, C) or random loci (B, D) with respect to marker sets from either the WZA (A, B) and baypass (C, D) genotype-environment association analyses. Points are color coded with respect to groups from Fig. 1 of the main text, and shapes corresponding to *k*-fold membership. Code to create these figures can be found in SN 15.21.

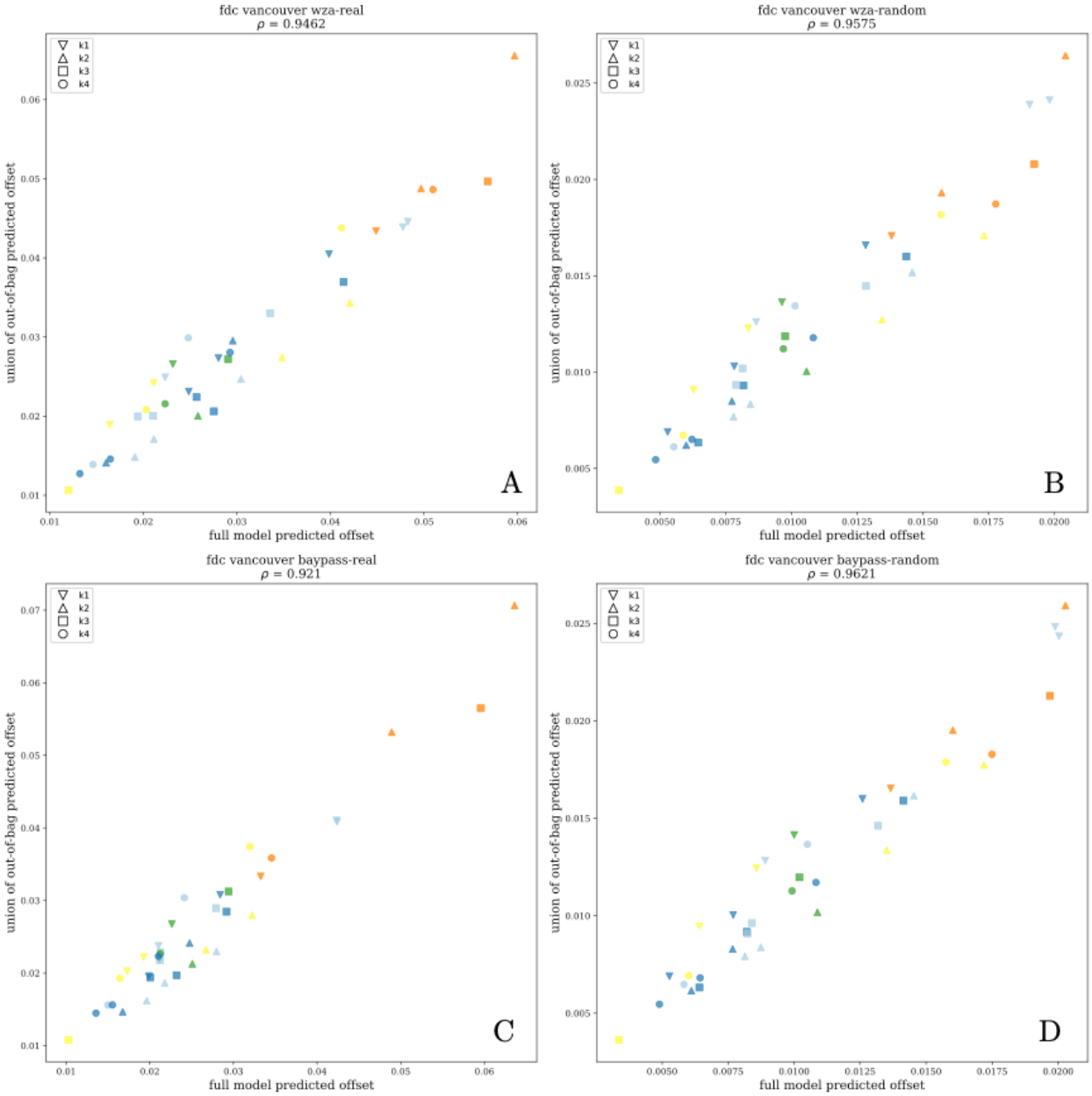
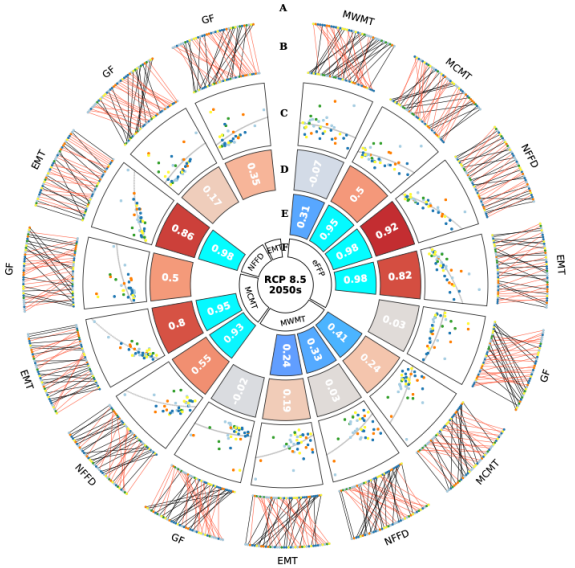


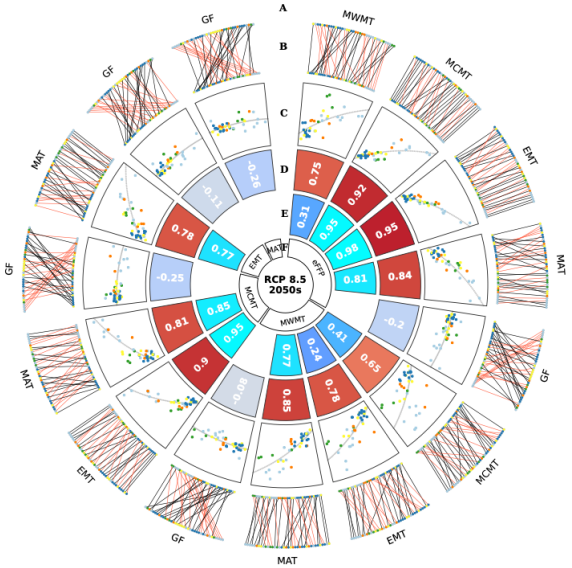
Fig S20 Relationship between estimated offset to the Vancouver common garden site predicted from full models (i.e., trained using all populations) of Gradient Forests (x-axes) and the union of out-of-bag offset estimations across *k*-fold models of Gradient Forests (y-axes) for the coastal variety models of Douglas-fir. Models shown are those using either outlier (A, C) or random loci (B, D) with respect to marker sets from either the WZA (A, B) and baypass (C, D) genotype-environment association analyses. Points are color coded with respect to groups from Fig. 1 of the main text, and shapes corresponding to *k*-fold membership. Code to create these figures can be found in SN 15.21.

Fig S21. Comparison of Douglas-fir offset projections

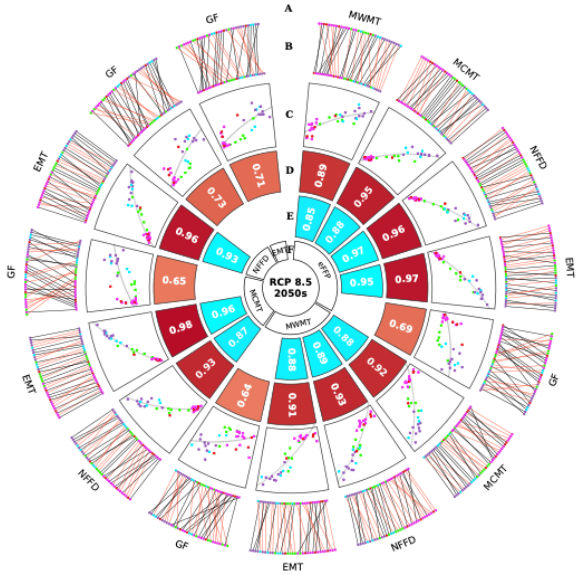
i) Coastal Douglas-fir
Cross-variety model



ii) Coastal Douglas-fir
Coastal variety model



iii) Interior Douglas-fir
Cross-variety model



iv) Interior Douglas-fir
Interior variety model

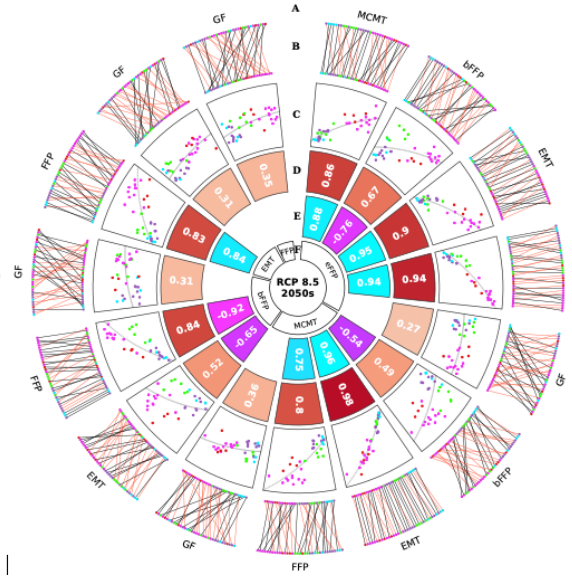
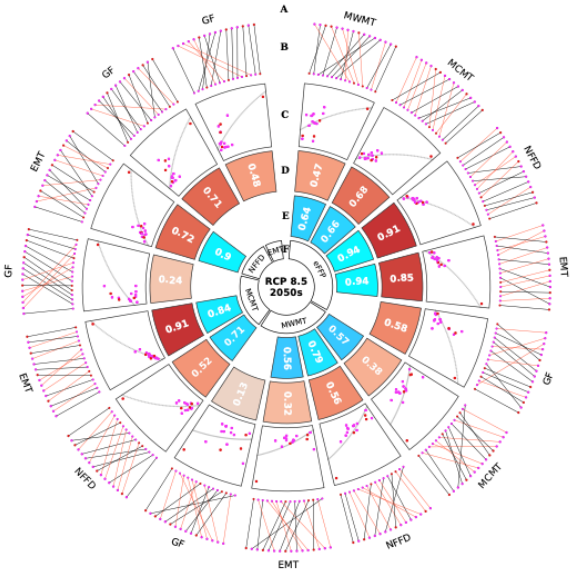


Fig S21. Comparison of Douglas-fir offset projections (cont'd)

v) Northwestern Interior Douglas-fir
Cross-variety model


RCP 8.5
2050s
A
B
C
D
E
MWMT
MCMT
NFFD
EMT
GF
EMT
GF
NFFD
EMT
GF
MCMT
GF

ix) Both Douglas-fir varieties
cross-variety model

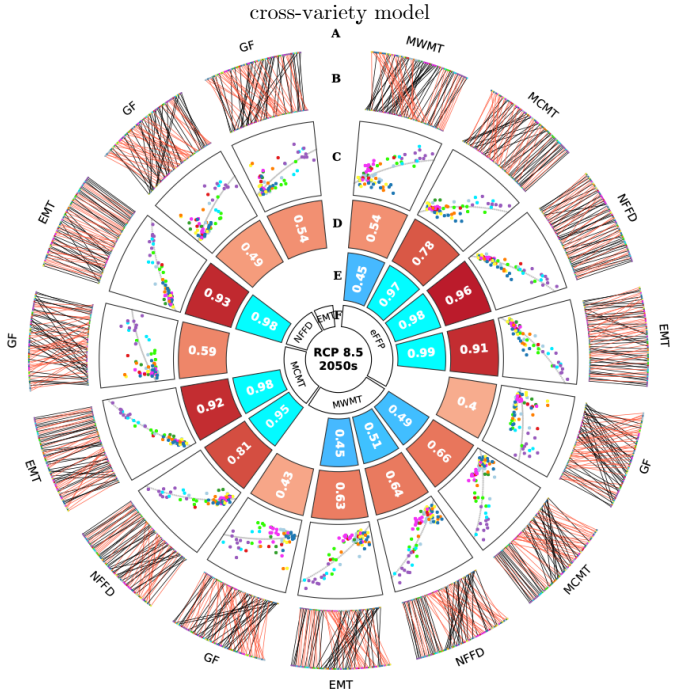


Fig S21 Comparisons of Gradient Forest (GF) and RONA top environments for projected maladaptation to RCP 8.5 2050s (models labeled within each panel). Circos plots show pairwise comparisons (A and F) between the offset predicted from the top five environments from RONA and GF (B-D) where B* shows the changes in rank between comparisons, C** shows the linear relationship (with gray line of best fit), and D shows the Spearman's rank correlation of offset predictions. E shows the correlation between future environmental variables compared in A and F. Environments used to estimate RONA were chosen based on ranking p-values from paired *t*-tests between current and future climate across all populations that went into a given model (this is why some environments differ between cross-variety and variety-specific models). Populations in B, C, and G are colored as in Fig. 1. Red lines in B are indicative of negative changes in rank between F and A. Boxplot ranks similar to Figs. 4A, 6A-B can be found in Supplemental Fig. S22. Code to create these figures can be found in SN 15.17. Analogous figures created using climate models RCP4.5 2080s, RCP4.5 2050s, and RCP8.5 2080s show similar patterns and are not shown except within SN 15.17.

* outer ranks are from A, inner ranks are from F

** y-axis is from A and x-axis is from F

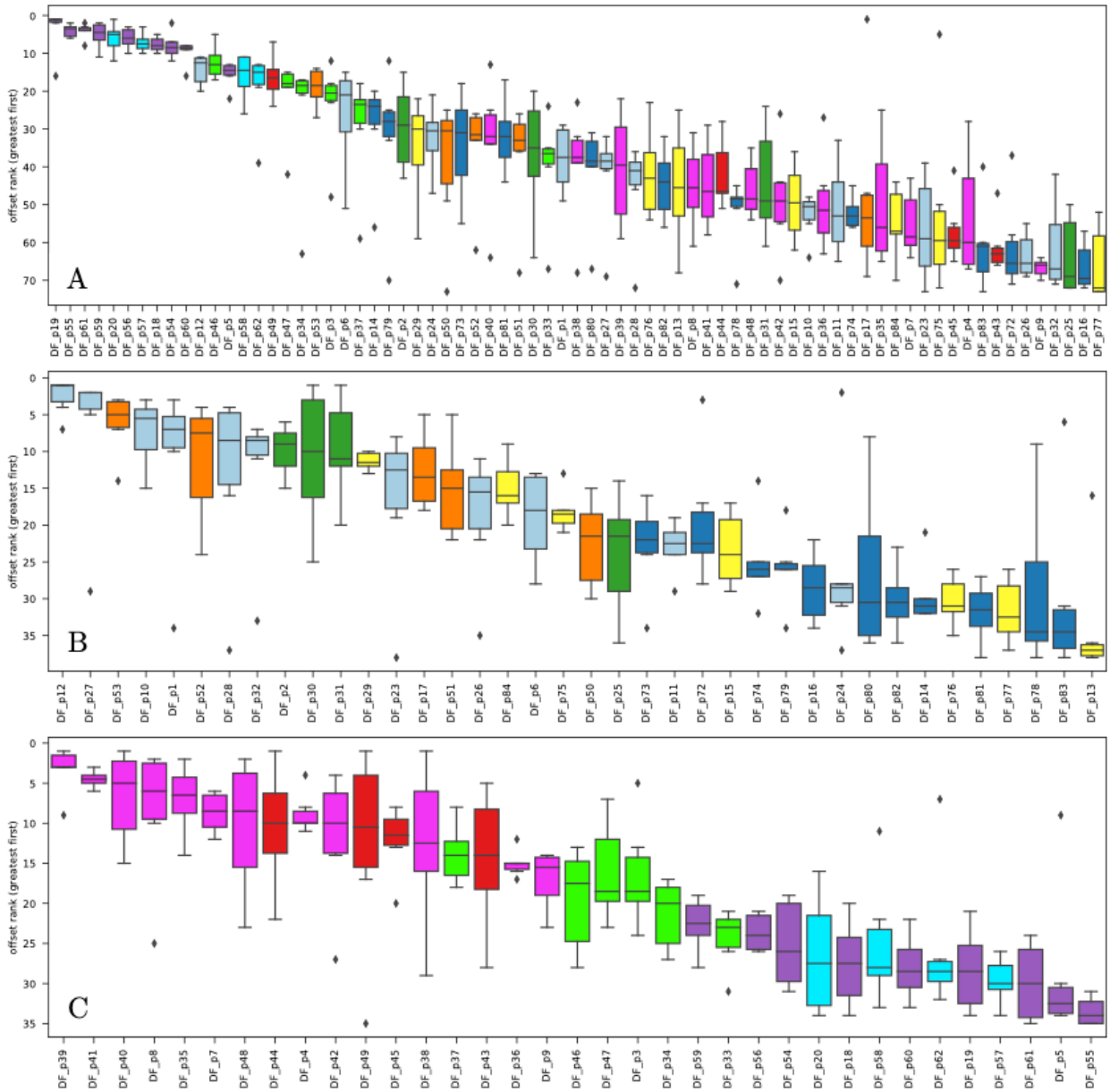


Fig. S22 Maladaptation of Douglas-fir populations to future climate (RCP 8.5 2050s) inferred from the A) cross-variety B) coastal-only, and C) interior-only models of RONA and Gradient Forests. This figure shows the distribution of offset rank for each Douglas-fir population from RONA and GF (note all but one of the values that create each boxplot are from RONA) and are ordered from left to right by increasing median offset rank. Information in A is the combined information from Fig. 6A-B. Code to create this figure can be found in SN 15.17. Analogous figures created using climate models RCP4.5 2080s, RCP8.5 2050s, and RCP8.5 2080s are not shown except within SN 15.17.

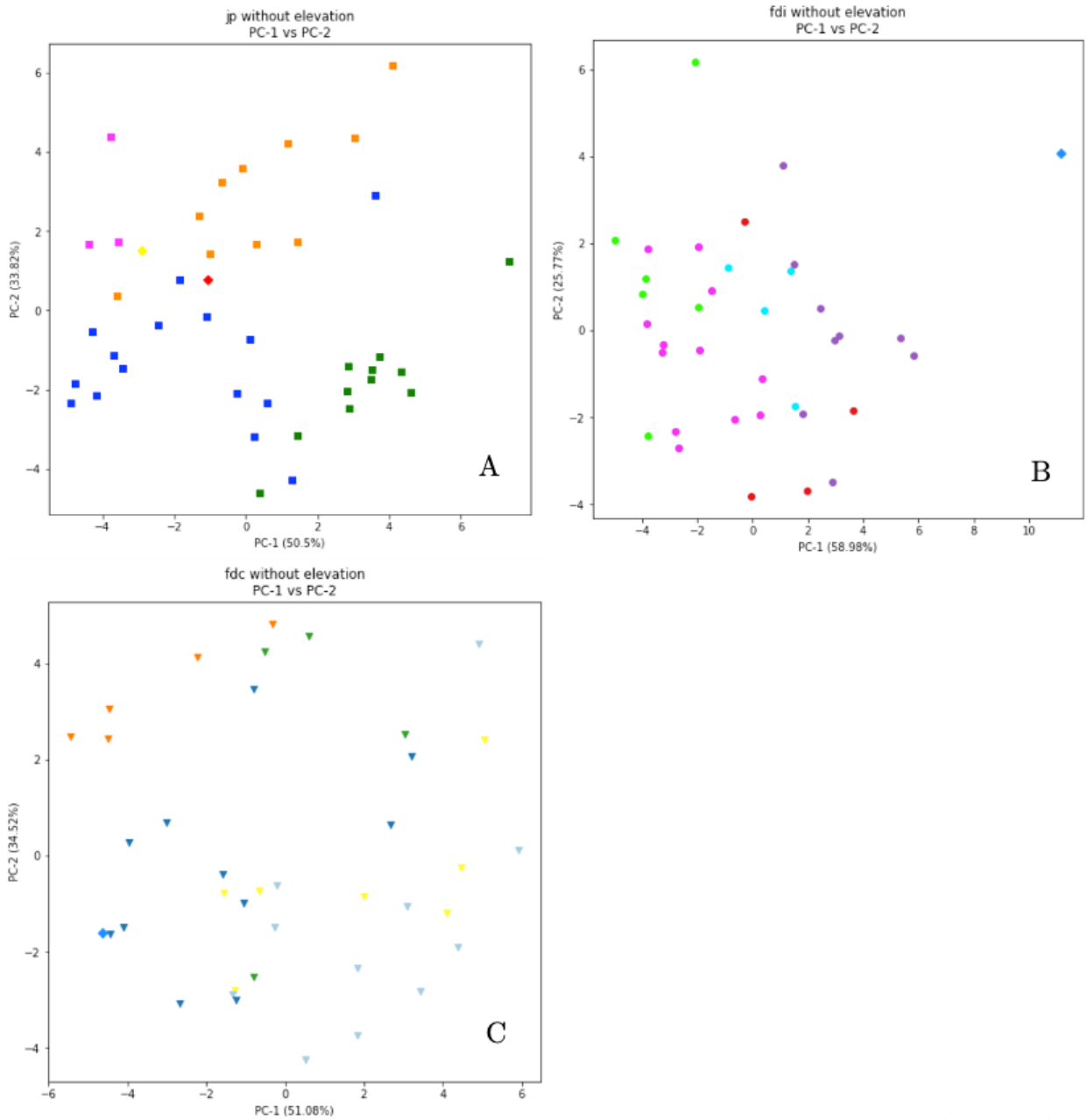


Fig S23 Climate space as inferred from Principle Component Analysis of climate normals (1961-1990) for populations of jack pine (squares, A); coastal Douglas-fir (circles, B), and interior Douglas-fir (triangles, C) relative to common gardens used for validation (diamonds). Colors in figures are the same as the genetic groups and common gardens used in Fig. 1. Code to create these figures can be found in SN 15.19.

Supplemental References

Yeatman, C. W. 1974. The jack pine genetics program at Potawa Forest Experiment Station 1950-1970. Pp. 1–33 *in* The jack pine genetics program at Potawa Forest Experiment Station 1950-1970.