



Genome-wide shifts in climate-related variation underpin responses to selective breeding in a widespread conifer

Ian R. MacLachlan^{a,b,1}, Tegan K. McDonald^c, Brandon M. Lind^a, Loren H. Rieseberg^d, Sam Yeaman^c, and Sally N. Aitken^a

^aDepartment of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; ^bDepartment of Ecology and Conservation Biology, Texas A&M University, College Station, TX 77843; ^cDepartment of Biological Sciences, University of Calgary, Calgary, AB T2N 1N4, Canada; and ^dDepartment of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Edited by Ronald R. Sederoff, North Carolina State University, Raleigh, NC, and approved January 21, 2021 (received for review August 10, 2020)

Locally adapted temperate tree populations exhibit genetic trade-offs among climate-related traits that can be exacerbated by selective breeding and are challenging to manage under climate change. To inform climatically adaptive forest management, we investigated the genetic architecture and impacts of selective breeding on four climate-related traits in 105 natural and 20 selectively bred lodgepole pine populations from western Canada. Growth, cold injury, growth initiation, and growth cessation phenotypes were tested for associations with 18,600 single-nucleotide polymorphisms (SNPs) in natural populations to identify “positive effect alleles” (PEAs). The effects of artificial selection for faster growth on the frequency of PEAs associated with each trait were quantified in breeding populations from different climates. Substantial shifts in PEA proportions and frequencies were observed across many loci after two generations of selective breeding for height, and responses of phenology-associated PEAs differed strongly among climatic regions. Extensive genetic overlap was evident among traits. Alleles most strongly associated with greater height were often associated with greater cold injury and delayed phenology, although it is unclear whether potential trade-offs arose directly from pleiotropy or indirectly via genetic linkage. Modest variation in multilocus PEA frequencies among populations was associated with large phenotypic differences and strong climatic gradients, providing support for assisted gene flow policies. Relationships among genotypes, phenotypes, and climate in natural populations were maintained or strengthened by selective breeding. However, future adaptive phenotypes and assisted gene flow may be compromised if selective breeding further increases the PEA frequencies of SNPs involved in adaptive trade-offs among climate-related traits.

climatic adaptation | selective breeding | lodgepole pine | positive effect alleles | polygenic traits

Local adaptation of climate-related traits in widespread temperate conifers has been demonstrated for centuries using extensive long-term common garden experiments (1, 2). As early as the 17th century, foresters were instructed to recognize variation in desirable traits and select seed from trees with favorable phenotypes (3). Modern tree improvement programs systematically select upon genetic variation, primarily to achieve growth gains and meet economic objectives. Estimates of genetic variation and gains from selection made using quantitative genetic models assume many anonymous loci of small effect underlie both variation in continuously distributed phenotypes and their responses to selective breeding. However, the type, quantity, effect size, distribution, and dynamics of genes underlying locally adaptive phenotypic variation and responses to selective breeding in forest trees are still poorly understood (4).

Directional selection over hundreds or thousands of generations has led to genomic features of domestication in agricultural crops including simplified genetic architectures underlying many traits, reduced genome-wide diversity, and numerous selective sweeps

(5–7). Beyond high-gain, short-rotation clonal forestry [e.g., *Eucalyptus* spp. (8)], we know little about the effects of artificial selection on adaptive genetic variation in forest trees, yet many tree species undergo some degree of selective breeding. Two or three generations of conifer breeding is not expected to have the same magnitude of genetic effects seen in domesticated crops, but if artificial selection for increased productivity is detectable in conifer genomes, it may expose genetic relationships and potential sources of trade-offs between growth and climatically adaptive phenotypes.

Climate-related adaptive traits are often intercorrelated due to pleiotropy, natural selection, or linkage disequilibrium (LD), so that strong directional selection on one trait can cause correlated responses in others. Pleiotropic allelic variants associated with phenotypes do not function in isolation. Antagonistic pleiotropic effects among traits can generate adaptive trade-offs among traits within populations, and limit gains from selection on a focal trait (9). Trait–trait correlations can also arise through strong selection acting in parallel on unlinked loci or from LD mediated by physical linkage of loci on chromosomes. Average genome-wide LD estimates in conifers appear to be low (10, 11)

Significance

Temperate trees originating from warmer localities usually grow faster and acclimate to winter later than trees of the same species from colder localities. However, when trees are selected for faster growth, are climatically adaptive genes and traits affected? Our research demonstrates a simple, sensitive, broadly applicable, and elusive approach to integrating complex polygenic variation into applied environmental management. We show that selective breeding increases allele frequencies of many trait-associated genes, and that alleles that increase most display strong genetic linkage and potential trade-offs among traits. Increasing tree growth while maintaining adaptation is essential to reforestation in a changing climate, but this may be challenging because trait-associated genetic variation that underpins climate adaptation is responsive to selection for tree growth.

Author contributions: I.R.M., L.H.R., and S.N.A. designed research; I.R.M., T.K.M., and S.Y. performed research; I.R.M. and T.K.M. contributed new reagents/analytic tools; I.R.M., T.K.M., B.M.L., and S.Y. analyzed data; and I.R.M., T.K.M., B.M.L., L.H.R., S.Y., and S.N.A. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: ian.maclachlan@alumni.ubc.ca.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2016900118/-DCSupplemental>.

Published March 1, 2021.

but may be greater (r^2 of 0.2–0.4) within genes under strong selection (12).

Conifer studies have identified putatively adaptive phenotype-associated alleles on a locus-by-locus basis using quantitative trait loci (QTL) mapping or genotype–phenotype associations (GPAs) (also known as genome-wide association studies [GWAS]) (13). Genotype–environment association analyses in conifers have identified putatively adaptive environmentally associated loci (e.g., 14, 15), but loci are usually anonymous relative to adaptive phenotypes. All of these approaches are biased toward detecting loci with large phenotypic effects, but expectations that genome scans will discover individual adaptive loci with large effects or frequency differences among populations may be biologically and statistically unrealistic (16, 17). Conifer GPA studies typically detect relatively few statistically significant loci, and locus-by-locus analyses are insufficient to characterize adaptive genome-wide variation associated with adaptive traits and signatures of selective breeding. Multilocus tests for adaptive polygenic signatures of selection have been developed (e.g., refs. 18 and 19), but significant limitations remain (reviewed in refs. 20 and 21).

Uncertainty about the effects of selective breeding on adaptive genetic variation is layered upon expectations that forest trees will become maladapted as climates shift (22). Efforts are being made to estimate maladaptation using genome-wide variation associated with adaptive traits and climate (23, 24), because conserving, managing, and efficiently redeploing genetic variation associated with adaptive phenotypes will be a necessary element of strategies to mitigate the effects of shifting climates on forest resources (25). Assisted gene flow strategies in temperate and subboreal forests generally aim to move trees to cooler climates in anticipation of future warming, but trees must then contend with the increased short- to medium-term risk of maladaptation to damaging frost. This means cold tolerance is, perhaps surprisingly, an important trait when planning for warming climates.

Genetic approaches have the potential to efficiently and accurately characterize local adaptation to climate. Understanding whether this potential can be realized in a technically robust and operationally feasible way, and whether selection for faster growth compromises genetic variation associated with phenotypic

adaptation to climate, has far-reaching implications for developing effective assisted gene flow strategies that mitigate negative climate change impacts on forest health and timber production (26, 27). In this context, our research objectives are to 1) identify the genetic architecture of climate-related adaptive traits in interior lodgepole pine (*Pinus contorta* Dougl. ex Loud. var. *latifolia* Engelm.); 2) identify genome-wide effects of artificial selection for increased productivity on climate-related traits; and 3) assess the implications of genetic responses to selection for assisted gene flow strategies.

Our study combines climatic data, genotype data from ~50,000 lodgepole pine single-nucleotide polymorphisms (SNPs), and seedling phenotypic data for height, cold injury, growth initiation, and growth cessation traits. These data were collected from a seedling common garden that sampled reforestation seed lots from 105 natural populations and 20 breeding populations from across the species' range in Alberta (AB) and British Columbia (BC), Canada (Fig. 1 and *SI Appendix, Table S1*). For each of the four traits, we identify range-wide GPAs using 929 seedlings from all 105 natural populations. Then for the 1% most strongly phenotype-associated SNPs, we examine how artificial selection within breeding populations has changed allele frequencies at individual SNP loci, within individual seedlings, populations (breeding zones), and three climatic regions (Fig. 1). Using elements from the approach of Turchin et al. (18), we study changes in frequency of the alleles that have a positive effect on adaptive traits (positive effect alleles [PEAs]). At each SNP locus, a PEA is the allele associated with increasing numeric values of the respective phenotype, determined in this case through GPA analyses in the natural seedling populations. PEAs reported here are associated with greater seedling height, greater cold injury, delayed growth initiation, and delayed growth cessation. To parse physical genetic linkage from allelic associations due to other causes, we compare LD estimated from our natural seedlings with estimates of recombination among haploid megagametophytes from a single maternal parent, where physical linkage is the only cause of LD. Integrating genetic, climatic, and phenotypic data gives us a robust basis to detect the effects of artificial selection on climate-related

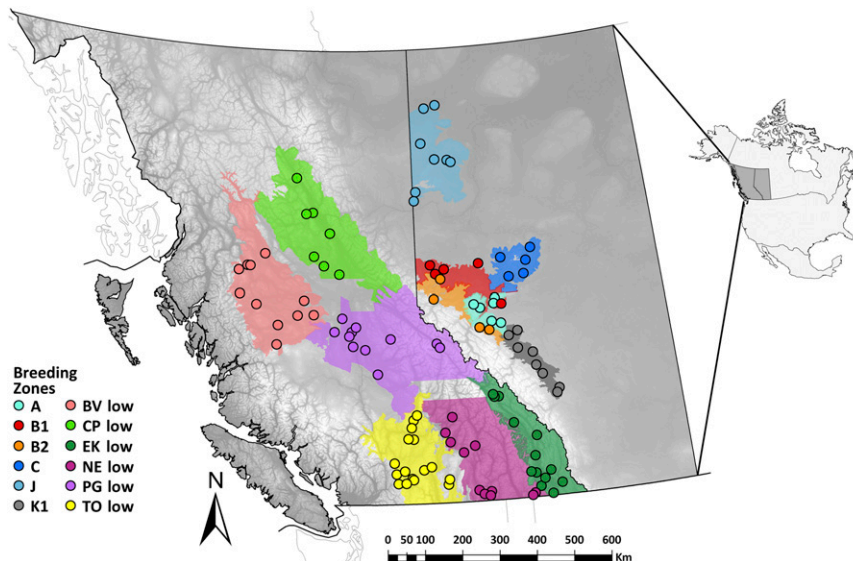


Fig. 1. Geographic origins of the natural and selected seedling populations sampled from across the range of lodgepole pine in Alberta (AB) and British Columbia (BC). Natural populations are represented by filled circles; selected seedling breeding zones are represented by filled polygons. The three climatic regions we used were AB, BC-Central, and BC-South. AB breeding zones are formally identified as A, B1, B2, C, J, and K1. BC-Central breeding zone abbreviations are as follows: BV, Bulkley Valley; CP, Central Plateau; and PG, Prince George. BC-South breeding zone abbreviations are as follows: EK, East Kootenay; NE, Nelson; and TO, Thompson–Okanagan. Reprinted from ref. 35, with permission from Elsevier.

genotypes that are relevant to breeding and assisted gene flow strategies.

Results

The lodgepole pine SNP array used in this study contained many previously detected candidates for associations with growth and phenology phenotypes, and climate (28). Following filtering for SNP quality, call rate, and minor allele frequency, and reduction to one SNP per genomic contig, our genetic dataset comprised 3,935 intergenic control SNPs and 18,563 candidate SNPs for GPA analyses from an initial array of 50,298 SNPs. Genotypes were successfully obtained for 2,084 seedlings (natural $n = 929$, selected $n = 1,155$). Multidimensional scaling (MDS) of the control SNPs indicated that population structure of the natural seedlings was continuous, forming a single large MDS cluster that was modestly structured but not discretely stratified based on breeding zone of origin (*SI Appendix, Fig. S1*). The same large cluster was present in MDS of control SNPs in selected seedlings, but 40 seedlings were outliers relative to the remaining selected or natural seedlings (*SI Appendix, Fig. S2*). To remove confounding effects of possible seed lot or DNA contamination, or genotyping errors, these individuals were excluded from further analyses, leaving 1,115 selected seedlings. GPAs for each trait using natural population seedlings had SNP P values that closely matched theoretical quantiles for a uniform distribution of 18,563 SNPs indicating little systematic bias in the strength of GPAs due to population structure (*SI Appendix, Fig. S3*). From these GPAs, we identified the top 1% ($n = 186$) SNPs with lowest GPA P values in the filtered panel of 18,563 SNPs and analyzed the effects of artificial selection on their PEA frequencies in breeding populations.

Seedling PEA Shifts in Response to Selective Breeding. The average number of PEAs per seedling (hereafter, the “PEA proportion”) was greater in selected seedlings than in natural seedlings for almost all traits and regions (Fig. 2 and *SI Appendix, Table S2 and Fig. S4*). Selection for increased height growth resulted in positive shifts in the distributions of PEA proportions for SNPs associated with increased height and greater cold injury in the regions BC-South and BC-Central (Fig. 2 *A–D*). These shifts corresponded to an average increase of PEA proportions in selected seedlings of 0.02 to 0.03, or a haploid change of 8 to 12 PEAs (*SI Appendix, Table S2*). Similar distributions of PEA proportions for height and cold injury were present in Alberta, but positive shifts in selected seedlings were much weaker (*SI Appendix, Fig. S4 A and B*).

The phenology traits had trimodal frequency distributions for the PEA proportions per seedling in all regions (Fig. 2 *E–H* and *SI Appendix, Fig. S4 C and D*). Positive shifts in average seedling PEA proportion for SNPs associated with delayed growth initiation and delayed growth cessation were large for selected seedlings from the warmest region, BC-South, where average PEA proportion increased by ~ 0.06 , or 20 to 25 additional PEAs. Increases in mean seedling PEA proportion were relatively small in BC-Central, and decreased slightly in Alberta, meaning that selection here slightly favors PEAs associated with earlier growth initiation and cessation (*SI Appendix, Table S2*).

To assess the significance of these results relative to background genetic variation, we drew 1,000 random samples of 186 SNPs from the panel of 18,563 candidate SNPs and tested differences in PEA proportions between natural and selected individuals in each random sample (e.g., *SI Appendix, Fig. S5*). PEA proportions differed between natural and selected seedlings for the top 186 SNPs significantly more than for random samples for all traits in BC-South, as well as height and cold injury in BC-Central, but not for phenology traits in BC-Central or any traits in Alberta (*SI Appendix, Fig. S6*). There was also substantially less variation in distributions of seedling PEA proportions calculated using random vs. phenotype-associated top 186 SNPs (*SI Appendix, Fig. S7*).

PEA Responses of SNPs to Selective Breeding. The responses of individual PEAs to selective breeding for height varied greatly among traits and regions, even though their constituent breeding populations were all selected for the same trait: height growth. For each of the four traits, the top 20 to 30 GPA SNPs by P value responded strongly to selection for height in BC-South and BC-Central, increasing in frequency by 0.08 to 0.14 (Fig. 3). In contrast, frequencies of the top 20 to 30 trait-associated SNPs in Alberta seedlings increased only slightly on average (<0.04), and inconsistently (*SI Appendix, Fig. S8*). Interestingly, most of the top ~ 30 SNPs were associated with all four traits (red bars, Fig. 3 and *SI Appendix, Fig. S8*), although this trend was weaker for the growth initiation SNPs.

For top 1% SNPs associated with either height or cold injury, the responses of PEAs to selection were very similar between BC-South and BC-Central (Fig. 3 *A–D*); however, the responses of phenology traits to selection for height contrast strongly between these regions. In BC-South, the warmest region, we observed a PEA frequency increase of ~ 0.1 for the majority of growth initiation and growth cessation-associated SNPs (Fig. 3 *E and G*). This response was largely absent in seedlings from the cooler BC-Central region (Fig. 3 *F and H*) and Alberta (*SI Appendix, Fig. S8 C and D*).

PEA–Phenotype Relationships. To quantify multilocus genotype–phenotype relationships, we regressed phenotypic means estimated for each of the 12 breeding zones upon mean breeding zone PEA frequencies of all top 1% ($n = 186$) phenotype-associated SNPs. PEA frequency–phenotype relationships for height-associated SNPs were moderate in natural seedlings, and slightly stronger in selected seedlings. An increase in the mean frequency of height PEAs by 0.1 corresponded to 12 cm (28%) greater height in selected seedlings (Fig. 4A). Of our four traits, cold injury had the strongest relationship with PEA frequencies among breeding zones, and a PEA frequency increase of 0.1 corresponded to 18% greater cold injury among breeding zones in both natural and selected seedlings (Fig. 4B). Phenology traits had a slightly greater range of mean PEA frequencies among breeding zones compared to height and cold injury. An increase in growth initiation PEA frequency by 0.1 corresponded to a 1.25-day delay in growth initiation in both seedling types (Fig. 4C). Relationships between breeding zone PEA frequency and growth cessation were moderately strong, with a 0.1 increase in PEA frequency delaying growth cessation by 4.5 days in both seedling types (Fig. 4D). To assess significance of PEA–phenotype associations, we used the same 1,000 random draws of 186 SNPs as in *Seedling PEA Shifts in Response to Selective Breeding* and calculated the phenotypic variance explained by their PEAs, relative to PEAs of the top 186 phenotype-associated SNPs for each trait (*SI Appendix, Fig. S9* and, e.g., *SI Appendix, Fig. S10*). In five of eight trait-by-seedling type comparisons, the strength of PEA–phenotype regressions for the top 186 trait-associated SNPs was greater than the 95th percentile of the random 1,000 samples and in seven of eight they were greater than the median value.

PEA–Climate Relationships. Strong positive clines in mean frequencies of trait-associated PEAs with temperature-related variables were observed for all four traits (Fig. 5 and *SI Appendix, Table S5*). Mean annual temperature (MAT), extreme minimum temperature (EMT), summer heat–moisture index (SHM), and climate PC1 explained moderate to large amounts of variation in PEA frequency for both seedling types ($0.58 \leq r^2 \leq 0.94$ for natural populations; $0.82 \leq r^2 \leq 0.94$ for selected populations) (Fig. 5). The strongest PEA frequency–climate clines were for height and cold injury PEAs with extreme minimum temperature and PC1 (Fig. 5 *C, D, G, and H*), but these clines were identical for natural and selected populations. Where differences in cline strength did occur between seedling types, they were always stronger in

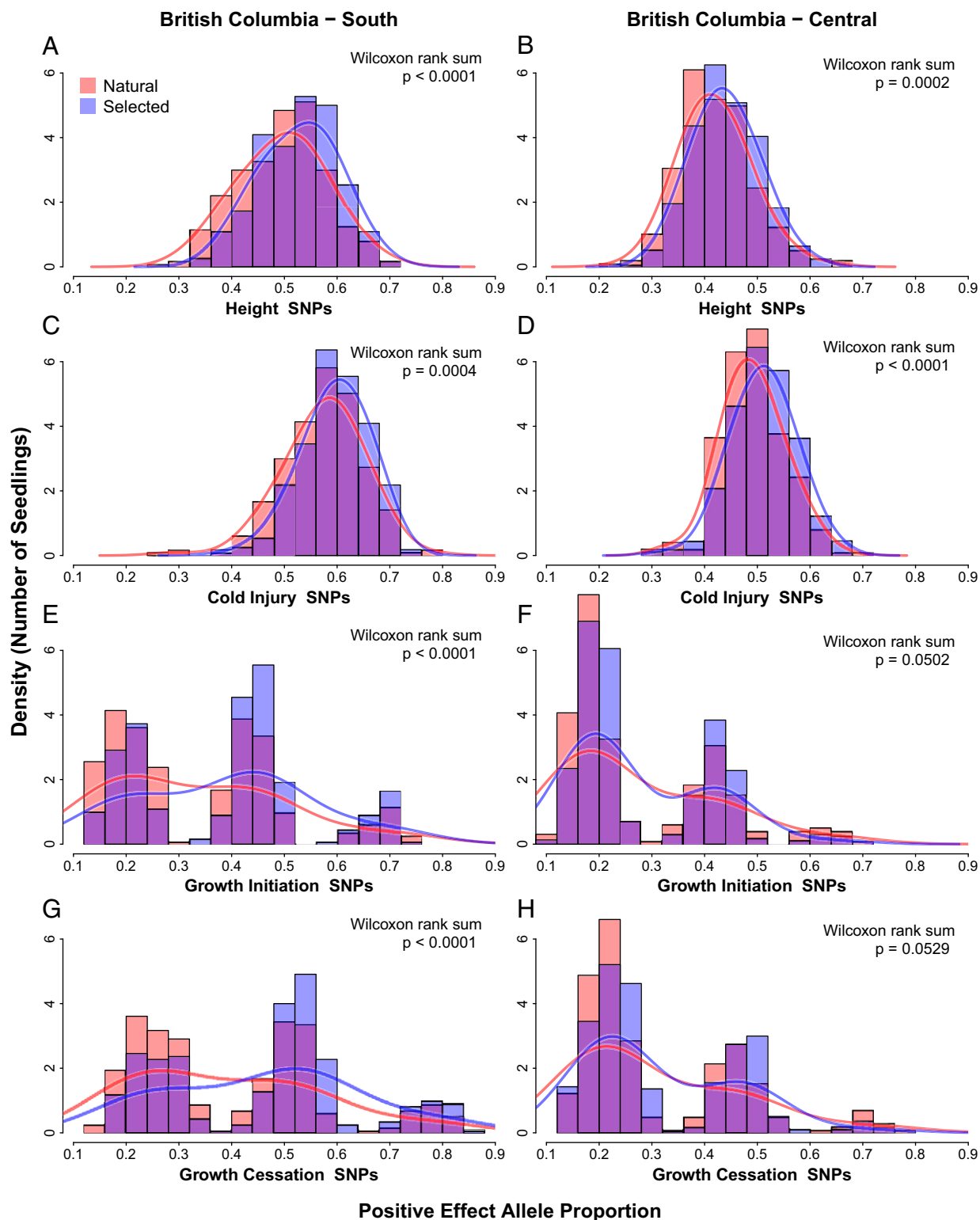


Fig. 2. Distributions of seedling PEA proportions for natural vs. selected seedlings from the milder British Columbia-South and cooler British Columbia-Central regions (columns) using the top 1% ($n = 186$) of SNPs associated with height (A and B), cold injury (C and D), growth initiation (E and F), and growth cessation (G and H). Wilcoxon rank sum tests for significant differences in the mean frequency of PEAs between seedling types are significant at a Bonferroni adjusted cutoff value of $\alpha = 0.0125$ for four comparisons within each region. Fig. 2 corresponds to *SI Appendix, Table S2*. For equivalent Alberta seedling plots, see *SI Appendix, Fig. S4*.

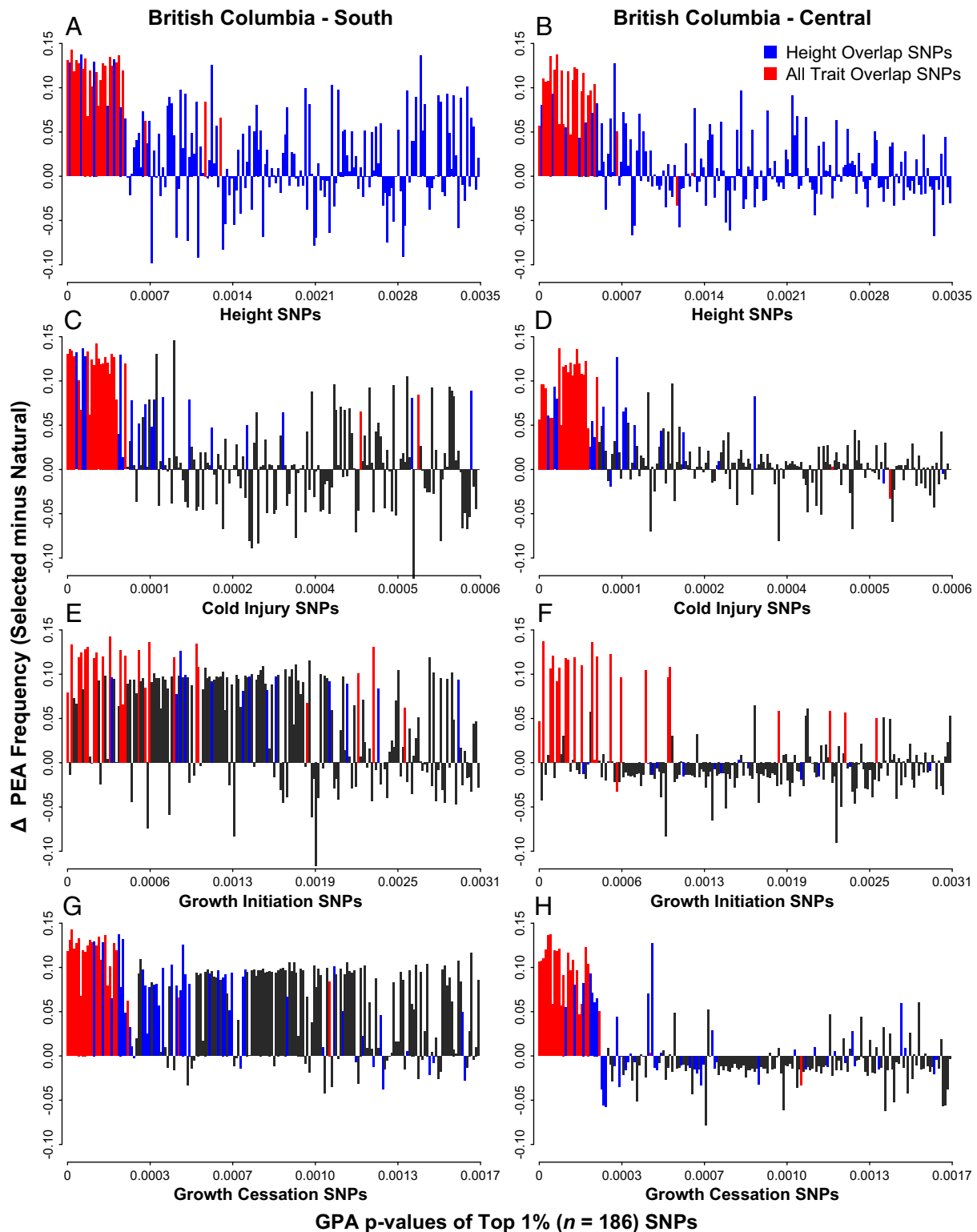


Fig. 3. Differences between natural and selected PEA frequencies from the milder British Columbia-South and cooler British Columbia-Central regions (columns) for each of the top 1% ($n = 186$) SNPs associated with height (A and B), cold injury (C and D), growth initiation (E and F), and growth cessation (G and H) ranked by increasing P value on the x axis. Differences were calculated as selected minus natural seedling PEA frequencies. SNPs that have overlapping identities with height-associated SNPs are colored in blue. SNPs colored red are those 23 SNPs with identities overlapping among the top 1% ($n = 186$) associated SNPs of all four traits. The remaining SNPs are colored black. Mean differences in PEA frequency between seedling types are equivalent to Δf values in *SI Appendix, Table S2*. For equivalent Albertan seedling plots, see *SI Appendix, Fig. S8*.

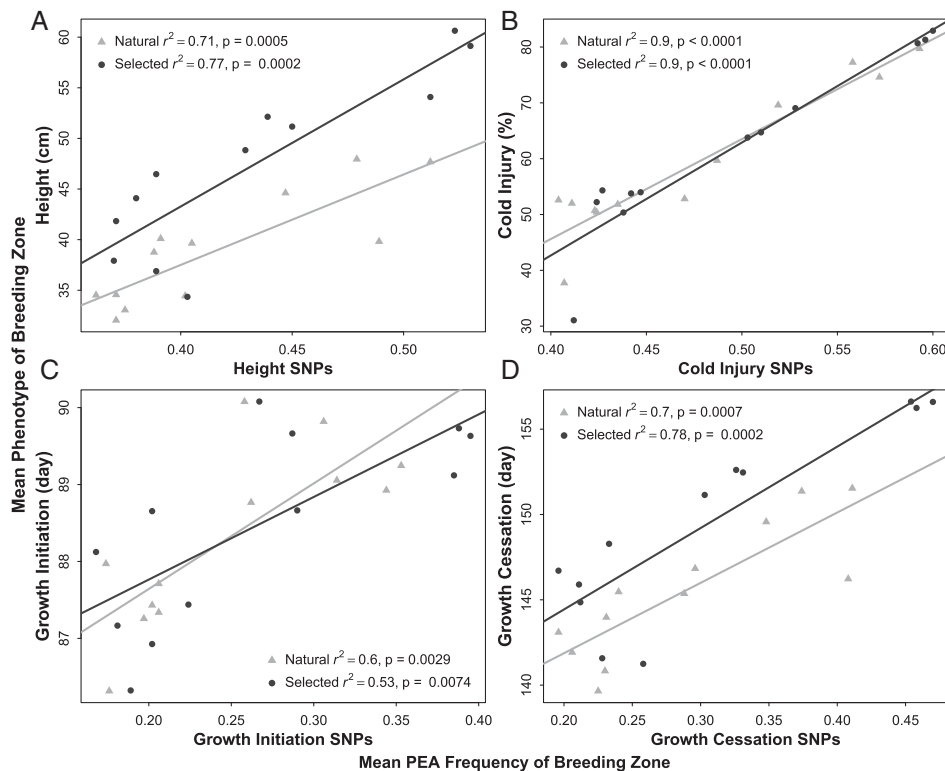


Fig. 4. Phenotypic variation in each trait for both natural (gray) and selected (black) seedling types explained by mean PEA frequency of the top 1% ($n = 186$) of phenotype-associated SNPs in each breeding zone (*SI Appendix, Table S3*). PEA–phenotype relationships are shown for height (A), cold injury (B), growth initiation (C), and growth cessation (D). Phenotypic values are best linear unbiased estimates of breeding zone means (*SI Appendix, Table S4*). P values are significant at a Bonferroni-adjusted cutoff value of $\alpha = 0.0125$ for four comparisons. No differences between natural and selected seedling slopes were statistically significant.

selected than in natural seedlings (Fig. 5 and *SI Appendix, Table S5*), but none was statistically significant.

Linkage Disequilibrium. Average pairwise LD for height and cold injury-associated SNPs was an order of magnitude greater than the background value and three orders of magnitude greater for phenology SNPs (*SI Appendix, Table S8*). One-third to one-half of the height and cold injury-associated SNPs had strong LD with other candidate SNPs in five to seven clusters (Fig. 6 A and B and *SI Appendix, Fig. S11 A and B*). By contrast, each phenology heat map is dominated by a single large cluster in strong LD containing nearly half of the top 1% ($n = 186$) associated SNPs (Fig. 6 C and D and *SI Appendix, Fig. S11 C and D*). Selective breeding increased the average strength of pairwise LD only slightly for all traits except cold injury (*SI Appendix, Table S8*) and did not change the patterns of pairwise LD among SNPs (*SI Appendix, Fig. S11*).

We parsed the relative importance of physical linkage and natural selection on LD using the mapping population. Most clusters of SNPs with high LD in natural seedlings also had high LD in the mapping population (Fig. 6, clusters above vs. below diagonal), indicating stronger physical linkage and lower rates of recombination within clusters than in the genomic background. However, in other cases, we found no pairwise LD in the mapping population between SNPs that showed high LD in natural populations (Fig. 6). This was particularly evident for the large cluster of SNPs in each phenology trait, where SNPs with high LD in natural populations were split among several smaller “stutter” clusters that showed no LD in the mapping population (Fig. 6 C and D). Similar but less prominent patterns also occurred for height and cold injury-associated SNPs, with a large cluster in each trait showing high natural population LD, but low mapping population LD for a small number of other SNPs (Fig. 6 A and B).

Gene Ontology and Annotation. Functional annotations were assigned to 33 SNPs from the large growth initiation and growth cessation LD clusters (Fig. 6 C and D), and 2 of the 23 SNPs that overlapped among all traits (*SI Appendix, Table S9*). It appears that at least 27 of the annotations were associated with primary or secondary metabolic processes or RNA replication, and cannot be related to specific phenotypes. However, the parent contig of one of the 23 SNPs that overlapped among all four traits is annotated to the Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Map 4712 associated with plant circadian rhythm.

We were unable to identify any Gene Ontology (GO) terms associated with the 23 SNPs that overlap among the top 186 SNPs for all traits. However, 16 GO terms were associated with the large clusters of SNPs in strong LD for both growth initiation and growth cessation, and were significantly overrepresented in our data (unadjusted Fisher’s P value, < 0.05) (*SI Appendix, Table S10*). Eight of these GO terms were associated with cellular or biological process involved in structural components of membranes and organelles. A further six GO terms are associated with biological or molecular processes that appear to have functions relating to release or transport of energy within primary metabolism.

Discussion

Two generations of selection for faster growth in lodgepole pine populations has produced substantial shifts in the proportions (Fig. 2) and frequencies (Fig. 3) of PEAs that were associated with increased growth, greater cold injury, and delayed phenology in natural populations. Twenty-three SNPs with the lowest GPA P values overlapped among traits and had substantially higher PEA frequencies in selected seedlings relative to natural seedlings, indicating these SNP markers are relevant to climate adaptation and highly responsive to phenotypic selection (Fig. 3).

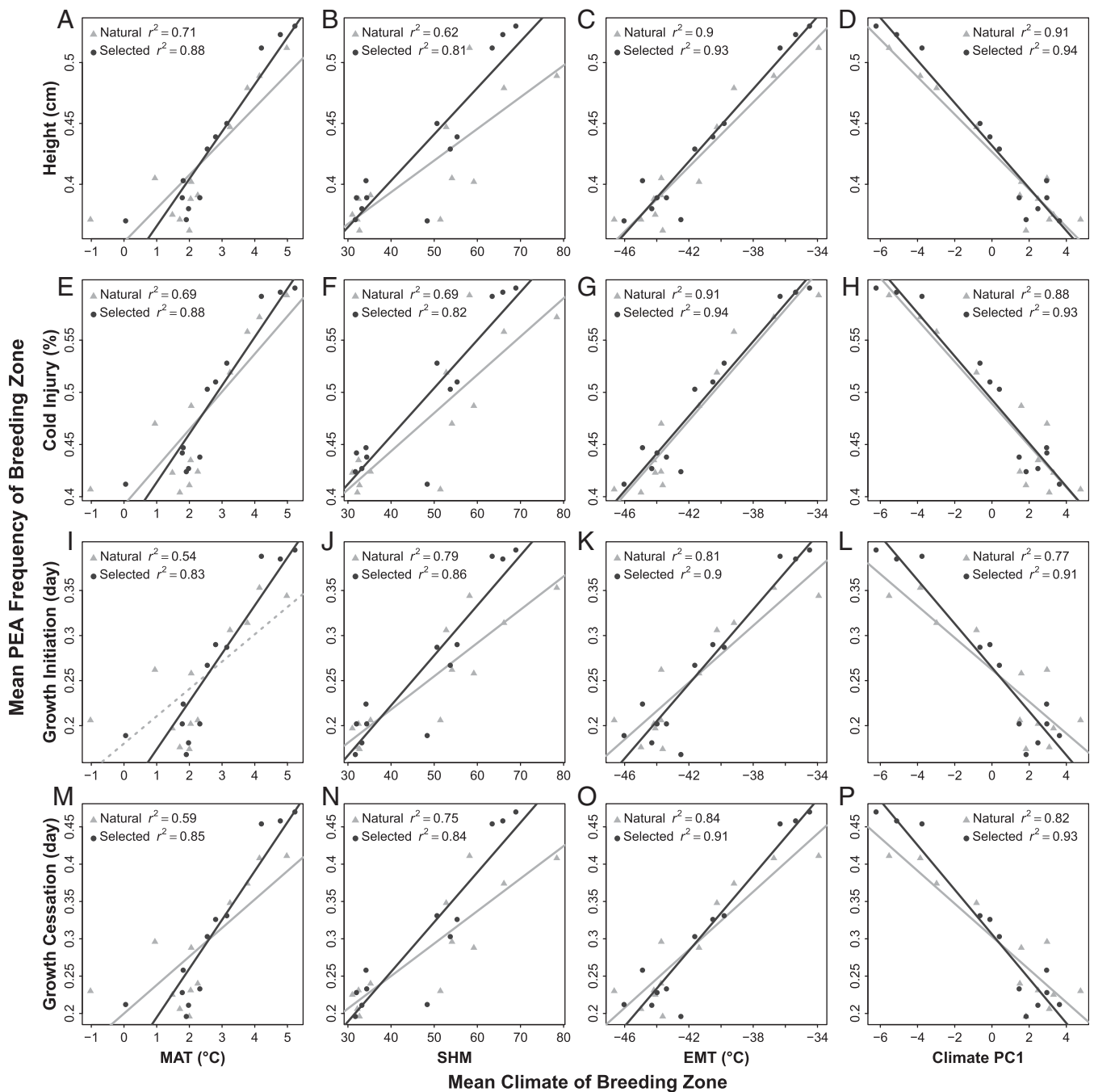


Fig. 5. Clines in the mean PEA frequency of each trait's top 1% ($n = 186$) phenotype-associated SNPs with four climatic variables (columns), MAT, SHM, EMT, and climate PC1. Clines in PEA frequencies with the four climatic variables are shown for height (A–D), cold injury (E–H), growth initiation (I–L), and growth cessation (M–P). None of the slopes within panels differs significantly between natural (gray) and selected (black) seedling types. r^2 and P values for all four traits and 11 climate variables are given in *SI Appendix, Table S5*. Points represent the mean climatic values (*SI Appendix, Table S7*) and PEA frequency values (*SI Appendix, Table S3*) for natural and selected populations in each breeding zone. Clines represented by a dashed line were not significant in *SI Appendix, Table S5* using a Bonferroni-adjusted cutoff value P value of $\alpha = 0.0045$ for 11 comparisons.

Tantalizingly, one of these 23 SNPs was annotated to a known circadian rhythm pathway (KEGG Pathway Map 4712). We found contrasting LD patterns for height and cold injury vs. phenology traits, indicating substantial differences in the number and genomic distribution of genes affecting these ecologically and economically important traits. Contrasting LD patterns appeared to be driven by pleiotropy or physical linkage, but we also found evidence for LD among physically unlinked genomic regions (Fig. 6). On average, we did not find evidence to date that

selective breeding for faster growth has compromised local adaptation to climate in the four traits we studied. However, the highly conserved LD seen within some genomic regions may constrain future responses to selection for growth, or exacerbate negative trade-offs among climate-related traits within breeding populations. While our PEA approach is simplistic, it is sensitive to and useful for characterizing both climatic adaptation and the effects of selective breeding among individuals and populations.

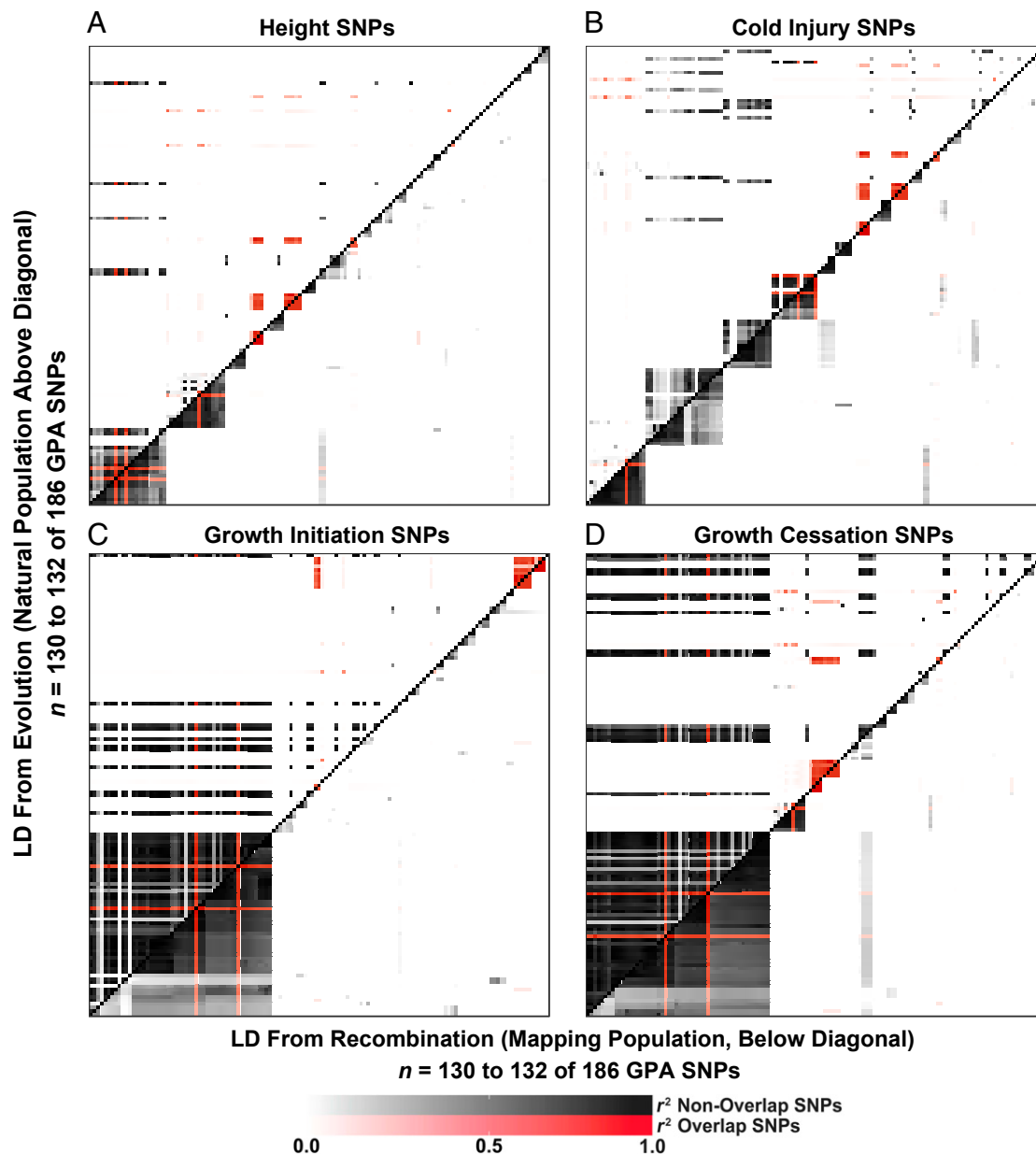


Fig. 6. Heat maps of pairwise LD r^2 values among the top 1% ($n = 186$) SNPs from natural seedlings (above diagonal) and equivalent r^2 values of their parent genomic contigs from the mapping population (below diagonal) associated with height (A), cold injury (B), growth initiation (C), and growth cessation (D). Approximately 50 SNPs associated with each trait in the natural population were on parent genomic contigs that could not be mapped to the *Pinus taeda* genome assembly, version 2.0, and were omitted here. Pairwise comparisons are colored red for the 23 SNPs that are common to the top 1% SNPs associated with all four traits. See *SI Appendix, Fig. S11* for a heat map of all 186 SNPs comparing natural and selected seedlings r^2 values.

Signatures of Selective Breeding on Genetic Architectures of Adaptive Traits. Positive shifts in the distribution of height-associated PEAs from natural to selected seedlings were substantial, and echoed by the selection responses of PEAs associated with cold injury and phenology traits (Fig. 2). These PEA proportion shifts after only two generations of selection and breeding quantify changes in the occurrence of genome-wide variants associated with climate-related traits in natural populations (*SI Appendix, Table S2*). They suggest either pleiotropy or linkage may cause some of the quantitative genetic trade-offs between growth and climate-related traits such as cold injury and phenology. Multimodal distributions in all regional subsets of our phenology-associated SNPs (Fig. 2 E–H and *SI Appendix, Fig. S4 C and D*) are caused by SNPs in the large LD clusters (*SI Appendix, Fig. S12 C and D*), contrasting with

unimodal distributions of seedling PEAs for height and cold injury (Fig. 2 A–D), and phenology SNPs outside the large LD clusters (*SI Appendix, Fig. S12 A and B*). Trimodal distributions of phenology SNPs are present in all three breeding regions, but vary in amplitude among regions (Fig. 2 E–H and *SI Appendix, Fig. S4 C and D*) and correspond little with population structure (*SI Appendix, Fig. S13*). These results indicate that phenology-associated variation is primarily partitioned among individuals and secondarily among geographic regions, and that responses to selective breeding for height in phenology-associated SNPs differ strongly among regions (Fig. 3 E–H and *SI Appendix, Fig. S8 C and D*). Loss or gain of one adaptive haplotype in each large phenology LD cluster affects ~60 associated SNPs, and therefore 60 different genomic contigs (Fig. 6 C and D), likely causing larger responses to selection

in the frequencies of most phenology-associated SNPs seen in BC-South vs. BC-Central (Fig. 3 *E–H*).

Climate-related traits of conifers are typically highly polygenic (2), and our results reflect these expectations. The 23 SNPs that overlap among all four traits (red bars, Fig. 3) appear to be pleiotropic and were those with the strongest statistical support (lowest *P* values) in independent GPA analyses of each trait, indicating that these SNPs (or linked causal variants nearby) are important to multiple adaptive traits and highly responsive to artificial selection. Despite a relatively consistent positive response to selection among the top 30 SNPs, PEA frequencies also responded positively for a large proportion of the remaining 156 phenotype-associated SNPs (Fig. 3). However, selected seedlings, on average, have fewer than 10 additional PEAs compared to the average natural seedling for most traits and regions; phenology traits in BC-South are the exception with up to 25 additional PEAs (*SI Appendix, Table S2*). This implies no single SNP, or small suite of SNPs, can reliably characterize genetic responses to selection for a trait, and that different suites of PEAs appear to exhibit some genotypic redundancy by producing similar positive responses to selection in a phenotype (29, 30). Such genotypic redundancy also indicates one reason why the potential of marker-assisted selection in forest trees has not been fully realized by operational tree breeding programs and is being outperformed by genomic selection (13, 31).

The 23 SNPs associated with all traits were highly responsive to selection, showed strong LD within five separate clusters (red clusters, Fig. 6), and based on Fig. 3 likely statistically capture a substantial portion of genome-wide responses to selection for height. LD clusters for height and cold injury were small to moderate in size and numerous, indicating that many loci associated with these two traits occur in discrete, unlinked clusters dispersed throughout the genome, or are unlinked with any other trait-associated loci in our analyses (Fig. 6 and *SI Appendix, Fig. S11*). By contrast, LD clusters for phenology traits were large and few. Interestingly, in some cases, we also observed extensive LD among individual SNPs or small clusters of SNPs that were unlinked in the mapping population. This stutter cluster pattern accounts for 11 growth initiation and 15 growth cessation SNPs, approximately one-fifth of SNPs in natural seedling LD clusters. It is most likely driven by strong divergent selection among natural populations. This LD pattern is most obvious for the phenology traits (*y* axis, Fig. 6 *C* and *D*); similar patterns occurred for height and cold injury to a lesser degree (*y* axis, Fig. 6 *A* and *B*) but were absent in the mapping population.

PEA frequency changes of SNPs in response to selective breeding give some insight into the dynamics and genetic architecture of LD clustering observed in natural populations. Frequency changes of SNPs in the LD clusters of Fig. 6 and their respective stutter clusters generally respond in similar positive or negative directions to selection (*SI Appendix, Fig. S14*). Small numbers of SNP loci responding in relative synchrony to selection appear to tag single phenotype-associated genomic regions. Even so, allele frequency responses to selection can vary substantially from zero to +0.12 in the largest LD clusters, and some are even slightly reversed in the BC-Central region (*SI Appendix, Fig. S14 E–H*). There are also several SNPs in the large natural population LD clusters that have weaker LD than in the mapping population (Fig. 6 *C* and *D*). These observations imply that SNPs in large LD clusters may represent multiple physically distant regions of the genome linked by strong selection favoring particular combinations of alleles, or that within a single large genomic region tagged by several tens of SNPs in a large LD cluster, sufficient recombination exists to allow independent segregation of some loci and variable responses of individual PEAs to artificial selection. This latter possibility is consistent with the rapid decay of LD reported in conifers over distances of a few hundred base pairs (11, 12, 32), which are far shorter than the length of genomic contigs used in our study. Our analyses cannot determine which SNPs are causal, and many SNPs in large LD

clusters of both the mapping and natural populations could be nearly neutral in relation to our traits (Fig. 6 and *SI Appendix, Fig. S14*). However, regardless of whether the LD clusters we observe arise due to pleiotropy, physical linkage, or natural selection, it appears that multiple causal genetic variants that are underlying these ecologically and economically relevant traits have concentrated and highly conserved genetic architectures, resulting in strong genetic trade-offs that will complicate or confound the future efforts of breeding programs.

Despite uncertainty over causation vs. correlation among the top PEAs, functional genetic analyses of phenology-associated loci in the large LD clusters provide intriguing insights. Most annotations for parent contigs of phenology-associated SNPs could not be readily linked with any phenotype. However, one notable biological pathway annotation, KEGG Pathway Map 4712 (*SI Appendix, Table S9*), and the genes it encompasses, are part of the plant circadian clock that controls biological activity in response to day length (33), a primary cue for seasonal phenological synchronization in temperate conifers. KEGG Pathway Map 4712 is annotated to one parent contig of the 23 SNPs associated with all four traits, indicating the likely importance of pleiotropic phenology-associated SNPs in the responses of cold injury and phenology to selection on height. Eight significantly overrepresented GO terms were identified (*SI Appendix, Table S10*), and six were associated with structural components of intracellular membranes or organelle formation. These functions could be part of cell expansion during growth initiation, intracellular remodeling during growth cessation in preparation for autumn cold hardening and senescence (34), or resistance of intracellular membranes to physical rupturing by ice crystals during freezing events.

Geographic Variation in Genetic Responses to Selection. Achieving gains in wood production through selection for greater height is the primary objective of lodgepole pine breeding. Genome-wide responses to selection for increased height differed substantially among traits between the BC-South and BC-Central breeding populations (Figs. 2 and 3), despite similar program histories, genetic gains for growth, and breeding, selection, and testing strategies implemented by the same breeder. Including data from Alberta meant we could capture a much greater range of phenotype and climate-associated genetic variation than from British Columbia alone, but in Alberta the genetic responses to selective breeding were weak for all traits (*SI Appendix, Figs. S4 and S8*). This is most likely due to younger tree improvement programs that have applied lower intensities of selection but could also relate to environmental constraints of colder boreal climates. It is possible that more extreme cold temperatures in Alberta than British Columbia exert greater selective pressure on genetic variants associated with both height growth and adaptation to low temperatures that leads to their fixation in natural populations, or preserves cold hardiness at the expense of responses to artificial selection for increased height.

Selection produced similar PEA frequency changes for height and cold injury-associated SNPs in both BC regions, but the changes in frequencies of phenology-associated PEAs were substantially greater in the warmer BC-South region than in cooler BC-Central (Fig. 3). Increased phenology PEA frequencies appear to be an important component of the large genetic gains for height seen in breeding zones from BC-South, compared to colder BC-Central or Alberta breeding zones (35), where the shorter growing season duration and extreme winter cold are more limiting to increases in growth. The increase of most phenology PEA frequencies in BC-South (Fig. 3) reflects later growth initiation and later growth cessation in faster-growing selected seedlings than natural seedlings (Fig. 4). It is most likely caused by strong pairwise LD among many phenology-associated SNPs and their contigs, and subsets of those SNPs associated with height, or all four traits that respond strongly to selection (blue and red bars respectively,

Fig. 3). These regional differences suggest that responses to directional selection on height are phenotypically similar among discrete breeding populations, but their underlying genome-wide responses are not necessarily parallel and appear to be constrained by trade-offs between growth and climate adaptation that vary among regional environments and associated field test climates.

The 23 SNPs associated with greater growth and autumn cold injury, and delayed growth phenology indicate loci that underlie trade-offs among these traits due to pleiotropy or physical linkage in natural populations. Selection increased PEA frequencies of these 23 SNPs, suggesting these trade-offs among traits are expected to be more prevalent within seedlings and breeding populations selected for faster growth. However, under current selection intensities and testing regimes in lodgepole pine breeding programs, we do not yet see much phenotypic evidence for this effect, on average, among breeding zones (35). Frequencies of cold injury-associated PEAs identified from GPAs of natural populations had relationships with population phenotypes and population temperature variables that were highly conserved, on average, between natural and selected seedlings (Figs. 4B and 5 G and H). Cold hardiness has repeatedly been identified as the phenotypic trait showing the strongest climate-associated population differentiation in temperate and boreal trees (2, 36). Our results show this has a substantial genome-wide basis and we cannot not discount the possibility that negative growth cessation–cold injury trade-offs may be more prevalent in individual selectively bred populations, or with further generations of selection, particularly within the BC-South region. Cold injury testing in seasons other than autumn, or using cold injury-associated loci identified from GPA analyses of selectively bred seedlings, may also reveal additional trade-offs between faster growth and cold injury.

Modest genetic differences in mean PEA frequencies among breeding populations were associated with relatively strong climatic differentiation, and reflect the genetic clines in lodgepole pine found by Mahony et al. (37) for 281 natural populations (including the 105 in this study). Mean height and cold injury PEA frequencies in our analyses vary by 0.15 across breeding zones, corresponding to 6 °C variation in mean annual temperature, and 12 °C in extreme minimum temperature (Fig. 5). Mean annual temperature (MAT) across our study area has increased 1.6 °C since the early 1900s; similar to the mean MAT difference between breeding zones EK low and BV low for the 1961–1990 reference period (Fig. 1 and *SI Appendix, Table S7*). The overall range in mean MAT values for natural breeding zones during the 1961–1990 reference period was 6.0 °C (*SI Appendix, Table S7*), but mean MAT across the range of our study is predicted to increase a further 4.5 °C by the 2070s under the RCP 4.5 scenario (38). Therefore, predicted warming is equivalent in magnitude to climatic differences among breeding zones that are associated with high levels of genetic divergence for both natural and selected seedlings.

Our results provide further evidence that assisted gene flow may be necessary to match genotypes of planted trees with future climates. Selection, testing, and breeding within zones have produced slightly stronger adaptive genetic clines along climatic gradients of individual temperature-related variables yet maintained nearly identical genetic clines in height and cold injury with climatic variables summarized as PC1 (Fig. 5). The synchrony of growth, phenology, dormancy, and cold hardiness appears to be largely conserved in selectively bred populations relative to climate, despite substantial genome-wide shifts in phenology and cold injury-associated PEA frequencies. Our findings provide genetic evidence that at the current stage in lodgepole pine breeding programs, similar assisted gene flow prescriptions can be applied equally to natural and selectively bred seed lots.

Using PEAs to Identify and Manage Adaptive Genetic Variation. PEA frequencies estimated climatically adaptive multilocus clines for complex polygenic traits (Fig. 5) far more effectively than previous

single-locus approaches (e.g., refs. 14, 39, and 40). Within populations, PEA frequencies were sensitive to small genome-wide shifts in adaptive alleles in response to selection and appeared to capture functional redundancy among different suites of loci responding to selection.

Accurate PEA effect size estimates could be expected to produce more effective polygenic scores and phenotypic predictions than a simple count of PEAs. However, in preliminary analyses, we found GPA effect sizes did not consistently identify loci with the largest responses to selection. Similarly, randomly selected samples of SNPs captured little genomic variation among seedlings (*SI Appendix, Figs. S5 and S7*), or among breeding populations where the range of PEA frequencies was ~ 0.02 (x axes, *SI Appendix, Fig. S10*). This produced a wide range of PEA–phenotype r^2 values because the chance sampling, or not, of any large-effect SNPs would substantially affect the strength of random PEA–phenotype associations (*SI Appendix, Fig. S9*). By contrast, top 1% phenotype-associated SNPs capture much greater genetic variation than random samples in all of our analyses and are more robust to small stochastic SNP sampling effects, even if PEA–phenotype associations are not always the strongest (*SI Appendix, Fig. S9*).

Our approach of selecting a single SNP per contig for analysis of growth initiation, growth cessation, and cold injury SNPs based on P values for height had no appreciable effect on our results (*SI Appendix, Figs. S15–S17*). We arbitrarily chose the top 1% ($n = 186$) SNPs by P value, but our results suggest this captured sufficient genetic variation to detect fine-scale effects of selective breeding within and among breeding zones. By repeating the phenotype–PEA and PEA–climate regressions with increasing numbers of SNPs ($n = 1–400$) (*SI Appendix, Figs. S18–S22*), we found that the number of SNPs required to obtain the maximum r^2 value cannot be reliably generalized, although it was usually fewer than $n = 186$. Therefore, most of our analyses conservatively estimate the power of PEAs to describe adaptive genetic variation relative to the number of PEAs used. Optimizing the number of PEAs used for each trait could make characterization of adaptive genetic variation more accurate and increase the average PEA frequency response to selection, but may exclude some loci that explain phenotypic variation and limit detection of fine-scale selection responses. Nonetheless, the PEA approach is compatible with classical clinal analyses of local adaptation in genecology, and has promise for simplifying the genetic characterization of polygenic traits that are relevant to the management of adaptive genetic variation and artificial selection in a changing climate.

The utility of combining PEAs into simple polygenic scores for managing genetic resources like forest tree populations lies in their potential to rapidly and accurately characterize genome-wide variation in climate-related traits. PEA analyses can exploit the increasing simplicity and decreasing costs of genotyping using SNP arrays or sequence capture approaches for nonmodel species with large genomes, and can characterize adaptive variation independently from long-term provenance trials. They could be an efficient means to monitor ongoing climatic maladaptation and associate populations with future climates for assisted gene flow prescriptions. As a genetic estimate of phenotypic variation, PEAs are also unaffected by nongenetic sources of error inherent in quantitative genetic parameters. Given the apparent sensitivity of PEA frequencies to selection, they may be a suitable measure to routinely characterize the phenotypic and climatically adaptive effects of selection, and estimate phenotypes of families within tree breeding programs. Validating these possibilities would require further analyses of PEA data from large-scale multifamily, multisite provenance trials.

Materials and Methods

Experimental Sampling, Design, and Data Collection. As described previously (35), 20 open-pollinated, selectively bred reforestation seed lots with the highest available genetic worth for growth were obtained from seed orchards

for 12 lodgepole pine breeding zones across Alberta (AB) and British Columbia (BC) (Fig. 1). Breeding zones represent discrete selective breeding management units across AB and BC, with zone-specific seed orchards. The seed lots we sampled represented breeding programs at varying stages of advancement from 1.0 to 2.5 generations. All breeding programs select for height growth as the primary trait and a proxy for timber volume. Selections are typically made based on family performance in progeny trials between 15 and 20 y of age across a range of representative climatic environments within each breeding zone. From each breeding zone, 4 to 16 open-pollinated natural population seed collections were sampled for a total of 105 natural seed lots. An outdoor seedling common garden was established from these seed lots in May 2012 in Vancouver, BC, Canada. Each breeding zone was represented by 60 to 112 natural seedlings, and each selectively bred seed lot was represented by 60 seedlings (total $n = 2,176$; natural $n = 976$; selected $n = 1,200$) (SI Appendix, Table S1).

Throughout the third growing season (2014), phenotypic data were collected for seedling height (in centimeters). Growth initiation day and growth cessation day were interpolated from growth curves fitted to the height time series data. Autumn cold injury was measured in October 2013 using Flint et al.'s (41) index of cold injury from fresh needle tissue using the artificial freeze testing and electrolyte leakage (42).

DNA was extracted from fresh needle tissue using the Macherey-Nagel Nucleospin 96 Plant II Core kit. Samples from all seedlings were genotyped using the AdapTree lodgepole pine Affymetrix Axiom 50,298 SNP array, which included candidate SNPs identified from exome capture using genotype-environment and genotype-phenotype associations (28, 43). After seedling mortality, and filtering for SNP quality and call rate, our genetic dataset contained genotypes for 2,084 individuals at 40,163 SNP loci (SI Appendix, section C.1).

Genotype-Phenotype Associations. Phenotype-associated SNPs were identified for seedlings from the range-wide sample of natural populations only, to avoid biases from allele frequency differences between comparative natural and selected seedling populations (20). Natural seedling genotypes ($n = 929$) in the SNP table were filtered further to retain 36,384 SNPs with minor allele frequencies ≥ 0.01 . GPA (also known as GWAS) analyses of 32,449 SNPs in 929 natural seedlings were implemented using the *mlma* function in GCTA (44) and the seedling phenotypic residual values of each trait (SI Appendix, section C.2). Population structure was controlled using a relatedness matrix reconstructed from a set of 3,935 filtered intergenic control SNPs included in the SNP array for this purpose and excluded from GPA analyses (SI Appendix, section C.2). Population structure in the natural seedlings used for these GPAs was assessed using coordinate plots from MDS of the control SNP relatedness matrix, implemented using the *cdmscale* function in R. This was repeated using the control SNPs in the selected seedlings as an additional population structure check, although selected seedlings were not used for GPAs.

To reduce biases arising from SNPs in high LD, only the SNP from each genomic contig with the lowest GPA P value for height was retained, leaving 18,563 SNPs. GPA results of the remaining three traits were reduced to match this suite of height-associated SNPs so that we could compare the effects of selection for height on the same SNPs among traits. Finally, SNPs in the bottom 0.01 quantile of GPA P values for each trait were identified as the top 1% associated SNPs ($n = 186$ SNPs per trait) and used for our remaining analyses.

PEA Responses to Selective Breeding. From the top 1% SNPs associated with each trait, we summarized the effects of selection on PEA proportions for individual seedlings and PEA frequencies individual SNPs. The total number PEAs carried by an individual seedling was referred to as the PEA proportion, calculated across all 186 SNPs for each trait. Distributions and means of seedling PEA proportions were compared between natural and selected seedlings. We validated these analyses by conducting 1,000 random draws of 186 SNPs without replacement and using the Wilcoxon rank sum test (*wilcox.test* in R) to compare distributions of PEA proportion at these loci between natural vs. selected individuals. Significance was assessed by comparing the P values for the top 1% SNPs with the P values from these random draws. The 2σ values of PEA proportion distributions from each iteration and region were used to demonstrate the difference in variation captured by random vs. phenotype-associated PEAs.

Individual SNP PEA frequencies were calculated separately across natural and selected seedlings, and frequency differences between seedling types were used to quantify responses to selection of each SNP. We also identified SNPs that overlapped among traits. To assess the effects of regional climate on genome-wide responses to selection, both the individual seedling and individual SNP analyses were run for regional subsets of the data that had different phenotypic responses to selection for growth and different MATs (35). The climatic regions were as follows: southern British Columbia (BC-South,

MAT 4.52 °C, breeding zones EK low, NE low, and TO low); central British Columbia (BC-Central, MAT 2.46 °C, breeding zones BV low, CP low, and PG low); and Alberta (MAT 1.53 °C, all breeding zones) (Fig. 1) (SI Appendix, sections C.3 and C.4).

Genotype-Phenotype and Genotype-Climate Relationships. Within every combination of breeding zone-by-seedling type (natural or selected), the number of PEAs was counted across all seedlings for each associated SNP. These counts were then summed across all 186 SNPs and divided by the maximum possible number of PEAs [n individuals \times (186 \times 2)] to give the breeding zone mean PEA frequency for each trait (SI Appendix, Table S3). Genotype-phenotype relationships were then summarized by regressing breeding zone-by-seedling type phenotypic means from MacLachlan et al. (35) (SI Appendix, Table S4) on breeding zone-by-seedling type PEA frequencies (SI Appendix, section C.5). The repeated random sampling approach described in *PEA Responses to Selective Breeding* was also used to validate these relationships by calculating mean breeding zone PEA frequencies and the genotype-phenotype r^2 values for 1,000 iterations of 186 randomly sampled SNPs.

Data for 19 climatic variables from natural seed lot and seed orchard parent tree provenances were estimated for the 1961–1990 climate normal period using *ClimateNA*, version 5.21 (38) (SI Appendix, Tables S6 and S7). Clines in breeding zone PEA frequency with breeding zone means for eight of the climatic and three geographic variables (SI Appendix, Table S5) were used to assess the proportion of genotypic variation explained by climate, and identify differences in the slopes of genetic clines between natural and selectively bred seedlings (SI Appendix, section C.6).

Linkage Disequilibrium. Pairwise LD was calculated separately for natural and selected seedlings as the squared correlation (r^2) among all pairs of associated SNPs ($n = 17,112$ pairwise comparisons for 186 SNPs) for each trait. Background pairwise LD among all 18,563 SNPs was estimated as the mean from repeated samples (SI Appendix, section C.7). Linkage blocks observed in hierarchically clustered heat maps of pairwise LD r^2 values were validated using comparisons with physical linkage from a lodgepole pine genetic mapping population consisting of 106 haploid megagametophytes from a single maternal parent (SI Appendix, section C.7).

GO and Annotation. To probe the functional basis of large phenology-associated SNP clusters and clusters of SNPs that overlapped among all traits in the pairwise LD analysis, we used the Eukaryotic nonmodel Transcriptome Annotation Pipeline (EnTAP, v.0.9.0) (45) to provide GO term analyses and functional annotations for coding regions of our lodgepole transcriptome represented in the SNP array. The GO and KEGG databases, as well as three protein databases: UniProt (46), National Center for Biotechnology Information (NCBI) RefSeq complete (47), and the TAIR 10 database (48) were all retrieved on June 5, 2019. Homology searches were run on the UniProt, NCBI RefSeq, and TAIR 10 databases. Annotated transcripts were mapped to our reference, and we transferred transcript annotations to individual SNPs using positions from the mapped transcripts (SI Appendix, section C.8).

Data Availability. The data and custom analysis scripts used in this paper have been deposited in the Dryad Digital Repository (DOI: [10.5061/dryad.ncjsxkstp](https://doi.org/10.5061/dryad.ncjsxkstp)).

ACKNOWLEDGMENTS. This research was part of the AdapTree project led by S.N.A. It was funded by Genome Canada, Genome BC, Genome Alberta, Alberta Innovates BioSolutions, the Forest Genetics Council of British Columbia, the British Columbia Ministry of Forests, Lands and Natural Resource Operations (BCMFLNRO), Virginia Polytechnic University, and the University of British Columbia. Seeds were kindly donated by 63 forest companies and agencies in Alberta and British Columbia (listed at <https://adapttree.forestry.ubc.ca/seed-contributors/>). Seed donation was facilitated by the Alberta Tree Improvement and Seed Centre, and the BCMFLNRO Tree Seed Centre. Our research would not have been possible without extensive technical assistance from the S.N.A. laboratory at University of British Columbia (UBC), and especially Pia Smets and Joanne Tuytel, at all stages of experimental establishment and data collection. Kristin Nurkowski was the AdapTree laboratory technician whose dynamic combination of exuberance, care, determination, and molecular laboratory expertise coaxed genetic data from many thousands of unsuspecting conifer seedlings as the basis for this manuscript. Kay Hodgins and Katie Lotterhos contributed extensively to development of AdapTree lodgepole pine SNP array. Jeremy B. Yoder contributed advice on the SNP table processing and GPA analyses. Seane Tehearne was extremely helpful at the UBC Totem Field site. Yousry El-Kassaby (UBC) and Greg O'Neill (BCMFLNRO) provided helpful comments and suggestions on the PhD thesis chapter from which this work originated. Last, we thank two anonymous reviewers whose feedback led to the inclusion of insightful additional analyses and valuable manuscript improvements.

1. O. Savolainen, T. Pyhäjärvi, T. Knürr, Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Evol. Syst.* **38**, 595–619 (2007).
2. F. J. Alberto *et al.*, Potential for evolutionary responses to climate change—evidence from tree populations. *Glob. Change Biol.* **19**, 1645–1661 (2013).
3. J. Evelyn, *Sylva, or, A Discourse of Forest-Trees, and the Propagation of Timber in His Majesties Dominions* (Royal Society of London, London, 1664).
4. B. M. Lind, M. Menon, C. E. Bolte, T. M. Faske, A. J. Eckert, The genomics of local adaptation in trees: Are we out of the woods yet? *Tree Genet. Genomes* **14**, (2018).
5. S. I. Wright *et al.*, The effects of artificial selection on the maize genome. *Science* **308**, 1310–1314 (2005).
6. J. F. Doebley, B. S. Gaut, B. D. Smith, The molecular genetics of crop domestication. *Cell* **127**, 1309–1321 (2006).
7. B. L. Gross, K. M. Olsen, Genetic perspectives on crop domestication. *Trends Plant Sci.* **15**, 529–537 (2010).
8. M. D. V. Resende *et al.*, Genomic selection for growth and wood quality in *Eucalyptus*: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol.* **194**, 116–128 (2012).
9. A. B. Paaby, M. V. Rockman, The many faces of pleiotropy. *Trends Genet.* **29**, 66–73 (2013).
10. D. B. Neale, O. Savolainen, Association genetics of complex traits in conifers. *Trends Plant Sci.* **9**, 325–330 (2004).
11. V. E. Chhatre, T. D. Byram, D. B. Neale, J. L. Wegrzyn, K. V. Krutovsky, Genetic structure and association mapping of adaptive and selective traits in the east Texas loblolly pine (*Pinus taeda* L.) breeding populations. *Tree Genet. Genomes* **9**, 1161–1178 (2013).
12. N. Pavy, M.-C. Namroud, F. Gagnon, N. Isabel, J. Bousquet, The heterogeneous levels of linkage disequilibrium in white spruce genes and comparative analysis with other conifers. *Heredity* **108**, 273–284 (2012).
13. D. B. Neale, A. Kremer, Forest tree genomics: Growing resources and applications. *Nat. Rev. Genet.* **12**, 111–122 (2011).
14. A. J. Eckert *et al.*, Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* **185**, 969–982 (2010).
15. B. M. Lind *et al.*, Water availability drives signatures of local adaptation in whitebark pine (*Pinus albicaulis* Engelm.) across fine spatial scales of the Lake Tahoe Basin, USA. *Mol. Ecol.* **26**, 3168–3185 (2017).
16. A. Kremer, V. Le Corre, Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity* **108**, 375–385 (2012).
17. S. Yeaman, Local adaptation by alleles of small effect. *Am. Nat.* **186** (suppl. 1), S74–S89 (2015).
18. M. C. Turchin *et al.*, Genetic Investigation of Anthropometric Traits (GIANT) Consortium, Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat. Genet.* **44**, 1015–1019 (2012).
19. J. J. Berg, G. Coop, A population genetic signal of polygenic adaptation. *PLoS Genet.* **10**, e1004412 (2014).
20. G. Sella, N. H. Barton, Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annu. Rev. Genomics Hum. Genet.* **20**, 461–493 (2019).
21. J. J. Berg *et al.*, Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* **8**, e39725 (2019).
22. S. N. Aitken, S. Yeaman, J. A. Holliday, T. Wang, S. Curtis-McLane, Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evol. Appl.* **1**, 95–111 (2008).
23. J.-P. Jaramillo-Correa *et al.*, Molecular proxies for climate maladaptation in a long-lived tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics* **199**, 793–807 (2015).
24. L. Browne, J. W. Wright, S. Fitz-Gibbon, P. F. Gugger, V. L. Sork, Adaptational lag to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted gene flow. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25179–25185 (2019).
25. B. Fady *et al.*, Forests and global change: What can genetics contribute to the major forest management and policy challenges of the twenty-first century? *Reg. Environ. Change* **16**, 927–939 (2016).
26. S. N. Aitken, J. B. Bemmels, Time to get moving: Assisted gene flow of forest trees. *Evol. Appl.* **9**, 271–290 (2016).
27. S. N. Aitken, M. C. Whitlock, Assisted gene flow to facilitate local adaptation to climate change. *Annu. Rev. Ecol. Evol. Syst.* **44**, 367–388 (2013).
28. S. Yeaman *et al.*, Convergent local adaptation to climate in distantly related conifers. *Science* **353**, 1431–1433 (2016).
29. Á. J. Láruson, S. Yeaman, K. E. Lotterhos, The importance of genetic redundancy in evolution. *Trends Ecol. Evol.* **35**, 809–822 (2020).
30. M. A. Nowak, M. C. Boerlijst, J. Cooke, J. M. Smith, Evolution of genetic redundancy. *Nature* **388**, 167–171 (1997).
31. D. Grattapaglia, “Status and perspectives of genomic selection in forest tree breeding” in *Genomic Selection for Crop Improvement*, R. K. Varshney, M. Rookival, M. E. Sorrells, Eds. (Springer, 2017), pp. 199–257.
32. M. Lu *et al.*, Exome genotyping, linkage disequilibrium and population structure in loblolly pine (*Pinus taeda* L.). *BMC Genomics* **17**, (2016).
33. M. Kanehisa, S. Goto, KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
34. G. R. Strimbeck, P. G. Schaberg, C. G. Fossdal, W. P. Schröder, T. D. Kjellsen, Extreme low temperature tolerance in woody plants. *Front. Plant. Sci.* **6**, 884 (2015).
35. I. R. MacLachlan, T. Wang, A. Hamann, P. Smets, S. N. Aitken, Selective breeding of lodgepole pine increases growth and maintains climatic adaptation. *For. Ecol. Manage.* **391**, 404–416 (2017).
36. G. T. Howe *et al.*, From genotype to phenotype: Unraveling the complexities of cold adaptation in forest trees. *Can. J. Bot.* **81**, 1247–1266 (2003).
37. C. R. Mahony *et al.*, Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. *Evol. Appl.* **13**, 116–131 (2019).
38. T. Wang, A. Hamann, D. Spittlehouse, C. Carroll, Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One* **11**, e0156720 (2016).
39. J. A. Holliday, M. Yuen, K. Rittland, S. N. Aitken, Postglacial history of a widespread conifer produces inverse clines in selective neutrality tests. *Mol. Ecol.* **19**, 3857–3864 (2010).
40. J. Chen *et al.*, Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in Norway spruce (*Picea abies*). *Genetics* **191**, 865–881 (2012).
41. H. L. Flint, B. R. Boyce, D. J. Beattie, Index of cold injury—a useful expression of freezing injury to plant tissues as determined by the electrolytic method. *Can. J. Plant Sci.* **47**, 229–230 (1967).
42. M. Hannerz, S. N. Aitken, J. N. King, S. Budge, Effects of genetic selection for growth on frost hardiness in western hemlock. *Can. J. For. Res.* **29**, 509–516 (1999).
43. H. Suren *et al.*, Exome capture from the spruce and pine giga-genomes. *Mol. Ecol. Resour.* **16**, 1136–1146 (2016).
44. J. Yang, S. H. Lee, M. E. Goddard, P. M. Visscher, GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
45. A. J. Hart *et al.*, EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Mol. Ecol. Resour.* **20**, 591–604 (2020).
46. UniProt Consortium, UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).
47. K. D. Pruitt, T. Tatusova, D. R. Maglott, NCBI reference sequences (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61–D65 (2007).
48. E. Huala *et al.*, The *Arabidopsis* Information Resource (TAIR): A comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. *Nucleic Acids Res.* **29**, 102–105 (2001).