Variation in gene expression patterns across a conifer hybrid zone highlights the architecture of adaptive evolution under novel selective pressures

Mitra Menon[1], Jared Swenson[2], Ehren Moler[3], Kristen Waring[4], Amy Whipple[2], and Andrew Eckert[5]

[1]University of California Davis
[2]Northern Arizona University
[3]California State Polytechnic University Pomona
[4]Northern Arizona University College of Engineering Forestry and Natural Sciences
[5]Virginia Commonwealth University

October 30, 2023

Abstract

Natural plant populations often exhibit marked differences in gene expression patterns that can reflect heterogeneity in selective pressures. Analyzing gene expression as a quantitative trait provides a unique opportunity to evaluate the underlying genomic basis of a plethora of traits and their interactions in driving adaptive evolution. We investigated patterns and processes driving expression differentiation under conditions mimicking future climates by combining common garden experiments with transcriptome-wide datasets obtained from hybrid populations of Pinus strobiformis and P. flexilis. We found strong signals of genotype-environment interactions (GEI) at the individual transcript and the co-expression module levels suggesting a marked influence of drought related variables on adaptive evolution. Overall, survival was positively associated with P. flexilis ancestry, but it exhibited an environment-specific pattern. Co-expression modules exhibiting strong associations with survival and genomic ancestry were representative of similar functional categories across both gardens. Using network topology measures, putatively adaptive garden-specific expression traits were pleiotropic and belonged to modules exhibiting high population differentiation yet low preservation across gardens. Overall, our study suggests the presence of substantial genetic variation underlying univariate and multivariate traits in novel climates that may enable populations of long-lived forest trees to respond to rapid shifts in climatic conditions in early seedling stages when mortality tends to be the highest. Our finding of pleiotropic trait architectures underlying adaptive traits, however, implies rapid adaptive responses to changing selection pressures depend on whether trait covariances align with the direction of change in selection pressures.

Variation in gene expression patterns across a conifer hybrid zone highlights the architecture of adaptive evolution under novel selective pressures

Mitra Menon[1,2#], Jared Swenson[3], Ehren Moler[3], Amy V. Whipple[3], Kristen M. Waring[4], Andrew J. Eckert[2]

[1]Department of Evolution & Ecology and Center for Population Biology, University of California, Davis, CA 95616, USA.[2]Department of Biology, Virginia Commonwealth University, Richmond, VA 23284, USA. [3]Department of Biological Sciences and Center for Adaptive Western Landscapes, Northern Arizona University, Flagstaff, AZ 86011, USA. [4]School of Forestry, Northern Arizona University, Flagstaff, AZ 86011, USA.

[#]Corresponding author:*mbmenon@ucdavis.edu*

Current word count = 7760

Figures = 4, Tables = 2

Abstract

Natural plant populations often exhibit marked differences in gene expression patterns that can reflect heterogeneity in selective pressures. Analyzing gene expression as a quantitative trait provides a unique opportunity to evaluate the underlying genomic basis of a plethora of traits and their interactions in driving adaptive evolution. We investigated patterns and processes driving expression differentiation under conditions mimicking future climates by combining common garden experiments with transcriptome-wide datasets obtained from hybrid populations of *Pinus strobiformis* and *P. flexilis* . We found strong signals of genotype-environment interactions (GEI) at the individual transcript and the co-expression module levels suggesting a marked influence of drought related variables on adaptive evolution. Overall, survival was positively associated with *P. flexilis*ancestry, but it exhibited an environment-specific pattern. Co-expression modules exhibiting strong associations with survival and genomic ancestry were representative of similar functional categories across both gardens. Using network topology measures, putatively adaptive garden-specific expression traits were pleiotropic and belonged to modules exhibiting high population differentiation yet low preservation across gardens. Overall, our study suggests the presence of substantial genetic variation underlying univariate and multivariate traits in novel climates that may enable populations of long-lived forest trees to respond to rapid shifts in climatic conditions in early seedling stages when mortality tends to be the highest. Our finding of pleiotropic trait architectures underlying adaptive traits, however, implies rapid adaptive responses to changing selection pressures depend on whether trait covariances align with the direction of change in selection pressures.

Keywords*: Adaptive evolution, conifers, climate change, gene expression, hybridization, quantitative genetics*

INTRODUCTION

Spatial variation in selection pressures leads to ecological specialisation and genetic differentiation among populations. Such ecological specialisation, often referred to as local adaptation, contributes to biodiversity. Common garden experiments and provenance trials of numerous species have demonstrated local adaptation via a fitness reduction when populations are translocated to abiotic and biotic conditions that diverge from those of their home environment (Clausen *et al.* , 1948; Savolainen *et al.* , 2013; Pickles*et al.* , 2015). Common garden studies have also revealed among-population variation for phenotypic plasticity (i.e., environment-dependent expression of a trait value; Dayan *et al* ., 2015), which itself is a phenotypic trait that can be adaptive (Reed et al. 2011). As climate change brings with it drastic mismatches and interannual fluctuations, variation in phenotypic plasticity among genotypes – referred to here as genotype-environment-interactions (GEI) – will be a major determinant of survival in the long term (Chevin, Collins & Lefevre, 2013; Franks, Weber & Aitken, 2014). Despite the abundance of studies investigating local adaptation and demonstrating GEI, few empirical studies have evaluated the architecture underlying GEI under novel selection pressures (Etterson & Shaw, 2001). Experimentally, novel selection pressures can be imposed through space-for-time substitution designs (Pickett, 1989) conducted by using common gardens located beyond or at the climate margin of a species range (Geber & Eckhart, 2005). Contrary to the larger contribution of small effect loci towards adaptive evolution in polygenic selection models (Fisher, 1935; Barghi *et al* ., 2020), novel selective pressures may disproportionately favor architectures of large effect loci. Some of these larger effect loci might have pleiotropic consequences, impacting several correlated traits which may not directly be the target of selection (Orr, 1998). The fitness benefits of pleiotropic architecture are likely to be transient and dominant only in the early phases of adapting to a new optima, specifically in species with high migration rates (Battlay *et al* ., 2023; Hamala *et al* ., 2021).

Populations of forest trees are often locally adapted (Lind *et al* ., 2018) with both intra- and inter-specific variants as key sources contributing towards the architecture of adaptive evolution, specifically under novel selective pressures (Aitken *et al* ., 2008; Taylor & Larson, 2019; Bolte & Eckert, 2020). This is likely

because inter-specific gene flow via hybridization can increase additive genetic variation, which influences trait responses to selection (Falconer & McKay, 1996). Studies in spruce, pine and poplar assessing fitness-related traits such as height, volume, bud set and disease resistance have demonstrated higher heritability and phenotypic variance in hybrid populations as compared to non-hybrid populations, as well as increased hybrid performance in novel environmental conditions (Dungey, 2001; De La Torre *et al* ., 2014; Suarez-Gonzalez *et al* ., 2016, 2018). Investigations of the architecture underlying GEI in novel environments and the contribution of hybrid ancestry to the evolution of this architecture, however, have lagged behind traditional investigations of adaptive evolution in forest trees. This is partially because of the difficulty in evaluating total lifetime fitness due to the longevity of trees.

Combining gene expression with survival – a key fitness component in trees – can help overcome the challenge posed by the longevity of tree lifecycles. Regulatory elements affecting gene expression disproportionately drive signals of adaptive evolution, especially for polygenic traits (Mei *et al* ., 2018). As such, it is expected that gene expression, which often also shows high heritability and responses to selection (Whitehead & Crawford, 2006; Eckert *et al* ., 2013), should be informative about architectures of adaptive evolution in natural populations. Even before the availability of genome-wide transcriptomic datasets, studies in systems biology demonstrated that metabolites, proteins, and gene expressions operate in the context of functional modules and are related to each other through a complex network of interactions (Hartwell *et al* ., 1999; Tohge *et al* ., 2005; Civelek & Lusis, 2014). The modular nature of biological networks permits environmental cues to target specific functional modules, limiting impact on other modules. Leveraging the modular nature of gene expression patterns (Hartwell *et al* ., 1999) and treating gene expression itself as a quantitative trait (Roberge *et al* ., 2007) can aid a better understanding of the multivariate architecture underlying adaptive evolution (Fagny & Austerlitz, 2021). This is enabled by estimation of co-expression networks (Barabási & Oltvai, 2004) using genetic values of expression levels that are treated as quantitative traits (i.e., one trait per locus). Genetic value is widely used in quantitative genetics as it represents the combined effect of all the alleles underlying a trait that an individual carries (Falconer & Mackay, 1996). Variation in genetic value is likely reflective of heterogeneity in selection pressures across the studied genotypes and is key for facilitating heritable responses to selection. The patterns and strength of connections among traits in co-expression networks, moreover, is often reflective of differing selection pressures. For example, strongly connected expression traits located at the core of networks often experience strong selective constraints, while those with lower connectivity are located at the periphery and often involved in GEI (Cork & Purugganan, 2004; Josephs *et al* ., 2017). When co-expression networks are constructed using genetic values rather than raw expression values, the connectivity patterns can be indicative of genetic covariances which are important components of the genetic architecture (Lande, 1980). We can thus use co-expression networks connectivity as a surrogate for understanding the relative role of two components of the genetic architecture - pleiotropy and linkage disequilibrium (LD). Genetic covariances between traits is key in determining the directionality of response to selection but has received limited attention in climate change studies, partly due to the extensive and rigorous experimental designs needed (Shaw & Etterson, 2012). Due to the rapid decay of LD in forest trees (Neale & Savolainen, 2004), high connectivity observed in networks is more likely to be indicative of strong pleiotropy as opposed to LD.

Our study uses the natural hybrid zone formed between two ecologically divergent species of pine, *Pinus strobiformis* Engelm. and*P. flexilis* E. James, to evaluate signatures of GEI and its contribution towards adaptive evolution. Both species have broad geographic distributions across western North America, with hybrid populations inhabiting sky-island ecosystems of New Mexico, Arizona, Texas and southern Colorado (Bisbee, 2014; Menon *et al* ., 2018) where they experience ongoing gene flow from *P. flexilis*(Critchfield, 1975; Menon *et al* ., 2018). Species distributed in fragmented populations are often vulnerable to extreme environmental fluctuations due to limited standing genetic diversity preventing adaptive responses (Aguilar *et al* ., 2008; Willi *et al* ., 2006). Previous work in this system has demonstrated no reduction in genetic diversity despite the high degree of fragmentation (Menon *et al* ., 2020) which may be the result of adaptive introgression (Menon *et al* ., 2021). These findings set the stage for evaluating the effect of interactions between interspecific gene flow and novel selective pressures as experienced under changing climatic conditions

on the genomic architecture of GEI in long-lived species.

We were specifically interested in evaluating the performance of the hybrid seedlings to climatic conditions that diverge from the historical norms of temperature and moisture availability and are similar to those expected under climate change models. To simulate these novel selective pressures, we utilized a space-for-time substitution design wherein we planted hybrid seedlings across two common gardens that represented warm to cool mean annual temperatures on an elevational gradient. Using this design, we tested the following four hypotheses about the role of GEI towards adaptive evolution employing transcriptome-wide expression traits at the per-transcript and co-expression module levels:

H1: Sampled populations will demonstrate strong signals of local adaptation at both the per-transcript and the module level. These will be reflective of heterogeneity in source populations' environmental conditions as well as novel selective pressures to which seedlings were exposed at the common gardens.

H2: Environmental differences (see Fig. S1) between common gardens will result in garden-specific patterns of trait differentiation (i.e., GEI) at both the per-transcript and the module level.

H3: Based on previous work on adaptive introgression in this system (Menon *et al* ., 2021) hybrid genomic ancestry will impact GEI at both the per-transcript and the module level.

H4: Traits with low connectivity within the co-expression network will dominate the architecture of adaptive evolution because such traits likely experience weak selective constraints and are more amenable to physiological fine tuning.

Overall, we demonstrate the prevalence of GEI across the transcriptome of our focal species and the key role it plays in driving adaptive evolution towards novel climatic conditions. By leveraging the connectivity patterns of gene expression traits within a quantitative genetic framework, we suggest the initial steps toward tracking novel climate optima disproportionately involve pleiotropic genetic architectures.

MATERIALS & METHODS

Seed sources, seedling culture, and common garden design

Seeds were collected from 30 maternal trees representing 10 populations (3 maternal trees/population; Fig. 1b) occurring across the *Pinus strobiformis* - *P. flexilis* hybrid zone in Arizona, Colorado, New Mexico, and Texas (Menon *et al* ., 2018). Cones were dried and processed prior to seed sowing and the first growth season in the Research Greenhouse at Northern Arizona University as described in Bucholz *et al* . (2020). Replicates of each maternal family in two gardens were planted into 107 cm × 117 cm × 91 cm raised bed garden boxes constructed at two Southwest Experimental Garden Array (SEGA) common gardens in the Kaibab National Forest, AZ (Fig. 1) also as described in Bucholz *et al* . (2020). Sets of seedlings from one tree are mixed-sib families (maternal families), and it is unknown whether seedlings within a maternal family are full or half siblings. Further details about seed sources and common garden design are detailed in Bucholz *et al* . (2020).

Seedling sampling

In August 2019, we sampled 90 seedlings per garden (3 siblings/maternal tree) representing 30 maternal trees originating from 10 populations (3 maternal trees/population) (Fig. 1) across the *P. strobiformis-P. flexilis* hybrid zone (Menon *et al* ., 2018). To evaluate the selective pressures experienced by the seedlings in the gardens relative to that in their native range we used climate variables at 1 km resolution from ClimateWNA v6.1 (Wang *et al* ., 2016). Population-level estimates of climate variables were obtained by averaging values across the three maternal trees selected per population. During 2019, the high elevation cold garden experienced drier and colder conditions relative to the values noted across a 60-year period in the native range of the sampled hybrid populations. The low elevation warm garden, experienced warmer and drier conditions relative to the sampled populations and relative to the cold garden as well (Fig. S1). The seedlings in each garden were further exposed to a 50% experimental reduction in growing season precipitation for three growing seasons prior to sampling in 2019 (see Bucholz *et al* . (2020) for details).

Transcriptomic dataset

For each seedling, up to three fascicles from the current year's growth were sampled and flash frozen in liquid nitrogen for RNA extraction. Tissue sampling was performed across two days with consistent weather conditions between the hours of 12:00 and 17:00. A maximum of 100 mg of needle tissue per seedling was ground into a fine powder using liquid nitrogen with a mortar and pestle with 40 mg of polyvinylpyrrolidone (PVP-40) added. Total RNA was extracted using Spectrum Plant Total RNA Kits (Sigma-Aldrich). Following polyA tail selection to enrich for mRNA, libraries were prepared following standard protocols for the NEBNext Ultra II RNA Library Prep (Illumina) and sequenced on the Novaseq6000 platform at Novogene Corporation (Sacramento, CA).

Raw data were assessed for quality using fastqc and trimmed to remove adapters and low-quality reads using TrimGalore v.0.6.4 (Krueger, 2015). Due to lack of a well-annotated genome from a closely related species, we built a *de novo* transcriptome assembly using Trinity v.2.8.5 (Grabherr *et al* ., 2011; Haas *et al* ., 2013) with ak -mer length of 24 and a minimum contig size of 300bp. Further details about the assembly pipeline, filtering steps, quality assessment and annotations using EnTAP (Hart *et al* ., 2020) are detailed in the Supporting Information (Methods S1).

Estimation of hybrid ancestry and population structure

We obtained genomic DNA from 270 trees using DNeasy 96 Plant Kits (Qiagen). These trees included the 30 maternal trees described above along with 154 parental *P. strobiformis* and 86 parental *P. flexili* s trees (Menon *et al* ., 2018; Fig. S2a). Following the ddRADseq library preparation detailed in Parchman *et al* . (2012), we size selected 300-400 bp regions and performed single-end sequencing (1 × 150 bp) on the Illumina HiSeq 4000 platform (Novogene Corporation) for multiplexed sets of 96 trees. The resulting data were processed using dDocent v.1.0 (Puritz *et al* ., 2014) as in Menon*et al* . (2018, 2021) to obtain genotype calls. Following post-filtering steps implemented in VCFTOOLS v.0.1.153 (Danecek *et al* ., 2011) and custom functions (Methods S2), we obtained 23,261 SNPs. We used this full dataset of 270 individuals and 23,261 SNPs to estimate hybrid ancestry through the genotype likelihood-based clustering algorithm implemented in NGSAdmix (Skotte *et al* ., 2013), with the number of clusters (*K)* set to the two hybridizing species (K = 2). We estimated multilocus F_{ST} from 11,432 of the 23,261 SNPs that were polymorphic across the 30 maternal trees using the method of Weir and Cockerham (1984) as implemented in the hierfstat v.0.04-22 package (Goudet, 2005) in R v.4.0.2 (R Core Team, 2020).

Estimation of maternal expression trait values

Prior to conducting further analyses using the initial 47,651 transcripts obtained after a series of standard filtering steps (Methods S1), we normalised read count data for varying library sizes using the trimmed mean of M values (TMM) approach as implemented with the*calcNormFactors* function from the edgeR v.3.14.0 package (Robinson *et al* ., 2010) in R. We further filtered transcripts that were not expressed across all three siblings nested within the same maternal tree in either garden. We implemented a linear mixed effects model for each of the remaining 27,441 transcripts (Eq. 1):

$$Y_{\text{ijklmn}} = \mu + \text{Batch}_j + \text{SY}_k + \text{Garden}_n | \left(\frac{\text{Pop}_l}{\text{Fam}_m}\right) + \epsilon_{\text{ijklnm}} (1)$$

Here, Y_{ijklmn} represents the normalised expression value of a transcript for the i^{th} seedling in thej^{th} sequencing batch sowed in thek^{th} year from the m^{th}maternal tree nested within the l^{th} population at n^{th} garden. The global mean across all seedings for a transcript is given by μ. The terms *Batch*and *SY* are fixed effects and represent the date of sequencing and year of seed sowing, respectively. The term *Fam* represents the maternal tree that is nested within *Pop* , which represents the source population of the seeds for a maternal tree. Both *Fam* and*Pop* are treated as random effects that are allowed to vary by the fixed effect of *Garden* . This linear mixed effect model was fitted using the *fitVarPartModel* function with log-transformed normalised counts and precision weights, which are ideal for RNAseq datasets (Law *et al* ., 2014), obtained from the*voomWithDreamWeights* function in the variancePartition v.1.16.1 package (Hoffman & Schadt,

2016) in R. Using the fitted model in equation (1), we obtained garden-specific genetic values for maternal trees and populations for each transcript.

Assessing per-transcript level signatures of adaptive evolution

We evaluated the among population component of additive genetic variance (Q_{ST}) for each transcript given our quantitative genetic study design (Spitze, 1993). Under neutrality, Q_{ST} should be similar to its genome-wide equivalent (F_{ST}) (Leinonen *et al* ., 2008). Under spatially divergent selection pressures driving local adaptation, however, Q_{ST} should statistically exceed F_{ST}. Using the per transcript variance components obtained from equation (1), we estimated Q_{ST} as:

$$Q_{\mathrm{ST}} = \frac{\sigma_a^2}{6\sigma_w^2 + \sigma_a^2}, \quad (2)$$

where σ_a^2 represents the among population variance component and σ_w^2 represents the within population variance component. The constant 6 originates from our assumption that the mixed-sib design is a 50:50 mixture of half and full siblings (Gilbert & Whitlock, 2015). Confidence intervals (CI, 95%) per transcript per garden were obtained using 1000 parametric bootstrap estimates generated using the *bootMer* function implemented in the lme4 v.1.1-28 package in R (Bates *et al* ., 2015). To address our hypotheses concerning garden specific patterns of adaptive differentiation at the per-transcript level (H1 & H2), we compared the 0.025 quantile of Q_{ST} for each transcript with the 0.95 quantile of F_{ST} and identified transcripts exhibiting signatures of local adaptation. Since we are primarily interested in the architecture underlying GEI, subsequent analyses and our main results focus on three Q_{ST}categories: (a) conditionally adaptive in cold garden (Cold-condA), transcripts with $Q_{ST} > F_{ST}$ only in high elevation garden; (b) conditionally adaptive in the warm garden (Warm-condA), transcripts with $Q_{ST} > F_{ST}$ only in low elevation garden; and (c) adaptive plasticity (Ad-Pl), transcripts with $Q_{ST} > F_{ST}$ across both gardens and a significant Q_{ST} reaction norm ($p < 0.05$) assessed using a t -test (Fig. 2). For each Q_{ST} category, we performed gene ontology (GO) enrichment analyses using a hypergeometric test with 1000 permutations to compute the family-wise error rate as implemented in the GOfuncR v1.12 package (Grote, 2020) in R. The background for GO enrichment analyses was the full set of GO terms across all annotated and non-contaminant transcripts from EnTAP (Table S1; Methods S1).

Assessing co-expression module level signatures of adaptive evolution

Local adaptation is often influenced by trait covariation (O'Brien *et al* ., 2019), thus modules formed by sets of strongly correlated transcripts effectively capture the architecture of local adaptation. We used the estimated genetic trait values for maternal trees from eqn. 1 to construct garden-specific co-expression networks using WGCNA v1.70 (Langdelfer & Horvath, 2008) in R. This is different from the standard practice of constructing co-expression networks using raw expression values across tissues, individuals or environments. Since we are using genetic values and not raw expression levels the co-expression network can be treated as a proxy for variance-covariance matrix (G -matrix), which has a long history of use in understanding the genetic architecture of local adaptation (summarized by MacPherson *et al* ., 2015). Each co-expression network consists of modules formed by sets of strongly correlated transcripts. This modularity permits selection to act on only certain functional categories. Modules formed within each network were visualised using the prefuse force-directed edge-weighted layout in Cytoscape v.3.8.2 (Shannon *et al* ., 2003). We filtered out weakly connected transcripts and edges with weights below the 75^{th} quantile for the module for clarity (Fig. S3).

Transcripts classified into each module were used to perform GO enrichment analyses using the same approach as described above for the three Q_{ST} categories of interest. To address our hypotheses concerning adaptive evolution (H1 & H2) at the module level, we evaluated among-population differences in the co-expression module using a linear model with the first eigengene of each module as the response variable and population of origin as the predictor variable. We also assessed the relative representation (RR) of individual adaptive transcripts (identified using the Q_{ST}-F_{ST} approach above, Fig. 2) in each module. For modules detected at the cold garden, we used the Cold-condA as the Q_{ST} category, while for modules detected at

the warm garden we used the Warm-condA category. In all cases, RR was estimated as:

$$RR = \frac{\frac{N_{Qst-module}}{N_{\mathrm{Qst}}}}{\frac{N_{\mathrm{module}}}{N_{\mathrm{network}}}}, \quad (3)$$

where $N_{Q\ \mathrm{st}\text{-module}}$ is the number of transcripts in a module that were also classified under the respective Q_{ST} category, $N_{Q\ \mathrm{st}}$ is the number of transcripts classified under the Q_{ST} category (Fig. 2), N_{module} is the total number of transcripts in a module and N_{network} is the total number of transcripts used in WGCNA. Statistically significant over- or under-represented modules (two tailed test, $p < 0.05$) were identified by permuting the transcript labels and calculating 10,000 null RR estimates for each module. For significant modules, we evaluated whether RR was associated with module size using a Kruskal-Wallis test as implemented with the stats v. 4.0.1 package in R.

Estimating the effect of environment on among population expression trait differentiation

To further address H1 & H2 at the per-transcript level, we identified climatic drivers of adaptive differentiation for each Q_{ST} category (Fig. 2) using redundancy analysis (RDA) as implemented in the vegan v.2.5.6 package in R (Oksanen *et al* ., 2013). Following Menon *et al* . (2021), we classified the environmental variables obtained from ClimateNA into drought- and freeze-associated variables, as these two axes are likely the primary selective agents across the hybrid zone populations. To reduce the dimensionality of the environmental dataset, we performed principal component analysis (PCA) separately for the drought- and freeze-associated variables and retained PC axes that explained at least 90% of the variance. We used the population-level estimates of gene expression as the response matrix with the drought-associated PCs, freeze-associated PCs, and geography as the predictor matrices. Geography was represented by scaled and centered estimates of latitude and longitude. For each Q_{ST} category, we evaluated the effect of the full model that included all three predictors, as well as the marginal effects of terms when the full model was significant ($p < 0.05$).

Influence of P. flexilis ancestry on expression patterns and survival

Survival for seedlings was recorded in August 2019 as a binary trait. To obtain maternal tree and population-level estimates of survival, we fitted a similar linear mixed effect model as used for the expression data, but with a binomial error distribution (Swenson 2021), as implemented with the glmmTMB v. 1.1.2.3 package (Brooks *et al* ., 2017) in R.

To assess the impact of genomic ancestry on adaptive trait differentiation (H3) at the per-transcript level, we correlated mean population-level estimates of *P. flexilis* ancestry obtained from NGSAdmix (Fig. S2b) with population-level expression values for each Q_{ST} category. For Cold-condA and Warm-condA, we used expression values estimated at cold and warm gardens, respectively. For Ad-Pl we used the absolute difference in expression values across gardens. For each Q_{ST} category, the observed distribution of Pearson's correlation coefficients (r) was compared against an empirically determined background distribution based on 10,000 permutations that were matched to the expression levels of transcripts in the category of interest. Significant deviation in the observed dataset was evaluated using the Kolmogorov-Smirnov test implemented in R ($p < 0.05$). We report results only from a simple permutation test that does not involve expression level binning since our results were not sensitive to binning and expression levels did not vary much by category (results not shown). Using a similar approach, we evaluated whether the three Q_{ST} categories deviated significantly from the background set of transcripts in their association with the estimated population-level survival measured at the respective gardens and the mean across gardens for the Ad-Pl category. Similarly, we tested the third hypothesis at the module level by evaluating whether the eigengene expression for a module was associated with both *P. flexilis* ancestry and survival using a false discovery rate-based multiple testing correction as implemented in the stats package in R ($q < 0.05$).

Connectivity patterns of Q_{ST} transcripts and a multivariate perspective on GEI

The positioning of traits within a co-expression network relies on their patterns of connectivity with other traits. Identifying where adaptive traits fall within the co-expression network can help us understand whether adaptive evolution proceeds through large pleiotropic effects or through smaller fine tuning of weakly

pleiotropic traits (Jordan*et al* ., 2004; Des Marais *et al* ., 2017; Josephs *et al* ., 2017). Additionally, comparing co-expression networks across environments allows identification of modules targeted by selection and those that are un-altered by changing environmental conditions, which we investigated by calculating three measures of connectivity for each expression trait. These reflected intramodular connectivity (kWithin), connectivity to all transcripts disregarding module membership in the network (kTotal) and the difference between inter- and intra-modular connectivity (kDiff). Specifically, to evaluate garden-specific patterns of GEI (H2) from a multivariate perspective, we performed comparative analyses using 10,000 permutations with the WGCNA package (Langfelder*et al* ., 2011). We used two aggregate summary statistics to declare a module as being preserved across gardens. $Z_{summary}$ represents a normalised value of various connectivity and density-based measures following a permutation test procedure, while *medianRank* simply scores each module based on the observed preservation statistic. Following Langfelder *et al* . (2011), we used both $Z_{summary}$ and*medianRank* for inference of module preservation. We declared modules as preserved across gardens if their $Z_{summary}$ scores were higher than 10 and*medianRank* scores were below 5.

To assess whether weakly pleiotropic traits dominate the architecture of adaptive evolution (H4), we implemented the network characterization approach developed by Mähler et al. (2017) to define the core of a module based on the top 10% of transcripts with the highest kTotal. We determined if the observed relationship betweenQ_{ST} categories and connectivity was significantly different than expected by chance by performing 10,000 permutations. Our permuted sets were matched on bins of expression levels to be representative of the core transcripts since expression levels are often associated with connectivity in a co-expression network (Fig. S4).

RESULTS

Transcript and module level signatures of adaptive evolution

At the per-transcript level, estimates of Q_{ST} across both gardens ranged from 0 to 1 (Fig. 3), with a mean ($\pm$ 1 *sd*) of 0.31 $\pm$ 0.37 for the cold garden and 0.37$\pm$ 0.38 for the warm garden. Approximately 50% of the transcripts exhibited differentiation below 0.10 in the cold garden, while 40% were below 0.10 in the warm garden. Only 20% of the transcripts exhibited strong differentiation ($>$ 0.90) in either garden. The distribution of F_{ST} using genome-wide ddRADseq markers was clustered around zero, with a multilocusF_{ST} of 0.015 (95% CI [0.06, 0.08]) (Fig. S2.c). Using Q_{ST}-F_{ST} comparisons, we classified 234 transcripts as Cold-condA, 344 transcripts as Warm-condA and 26 transcripts as Ad-Pl (Fig. 2) thus supporting H1 and H2 at the per-transcript level.

At the cold garden, the co-expression network consisted of 22,757 transcripts, which were grouped into 31 modules (Methods S3, Fig S5), with the number of transcripts per module ranging from 75 to 5,668. At the warm garden, the co-expression network consisted of 23,468 transcripts, which were grouped into 26 modules, with the number of transcripts per module ranging from 96 to 9,410 (Methods S3, Fig S5). Using the eigengene values of each module, we identified 20 modules at the warm garden and 21 at the cold garden that were strongly differentiated across populations (Table S3). Of the 20 modules identified at the warm garden, five were significantly enriched for transcripts classified as Warm-condA ($p < 0.05$). Five modules were significantly depleted for Warm-condA transcripts and did not overlap with the 20 strongly differentiated modules (Fig. 4; Table S3). Similarly, at the cold garden four of the 21 strongly differentiated modules were significantly enriched ($p < 0.05$) for transcripts categorised under Cold-condA. Three modules were significantly depleted ($p < 0.05$) for Cold-condA transcripts and did not overlap with the 21 strongly differentiated modules (Fig. 4; Table S3). At the warm garden, the number of transcripts in enriched modules ranged from 125 to 3,439, while at the cold garden they ranged from 139 to 5668. Q_{ST}enrichment or depletion was not associated with module size across either garden (Cold garden: $H = 4.78$, $df = 2$, $p = 0.09$; Warm garden: $H = 0.68$, $df = 2$, $p = 0.71$). These patterns of strong module level population differentiation noted across both gardens render support for H1.

Drivers of adaptive evolution at the per-transcript level

Drought, freeze and geography associated variables jointly explained 50% of the expression variance for the

Cold-condA category, 17% for the Warm-condA category and 19% for the Ad-Pl category (Table 1). The full model with all three predictors was significant only for the Cold-condA category and this was driven by geography and drought on the first RDA axis ($p_{\text{geo}} = 0.004$,$p_{\text{drought}} = 0.011$).

Estimates of genomic ancestry for the 30 maternal trees ranged from 0.18 to 1.00 (Fig. S2b), with 100% genomic ancestry from *P. flexilis*set at a value of 1.0. Survival estimates ranged from 0.37 to 1.00 at the Cold garden and from 0.23 to 0.90 at the Warm garden. Across both gardens, estimated survival of maternal trees was positively correlated with their genomic ancestries (Pearson's r ; Warm garden: r= 0.24, p = 0.18; Cold garden: $r = 0.40$, $p = 0.02$). On average (± 1 sd), population-level expression values were weakly correlated with population-level estimates of ancestry (-0.02 ± 0.31 in Cold garden and 0.02 ± 0.31 in Warm garden). Correlation coefficients for individual transcripts, however, ranged from -0.6 to 0.8 for Cold-condA, from -0.8 to 0.8 for Warm-condA and from -0.6 to 0.4 for Ad-Pl. The cumulative distribution of correlation coefficients between ancestry and transcript abundance for all categories was not significantly different from their respective matched background set of transcripts (Cold-condA: $D = 0.05$, $p = 0.702$; Warm-condA:$D = 0.03$, $p = 0.84$; Ad-Pl: $D = 0.13$, $p = 0.74$). Several transcripts in the Warm-condA and Cold-condA categories displayed significant associations with ancestry ($p < 0.05$), but none were significant after multiple test corrections. No transcript in the Ad-Pl category was significantly associated with ancestry.

Expression values on average (± 1 sd) were weakly correlated with survival estimates (Cold garden: r = -0.008 ± 0.32; Warm garden:$r = 0.029 \pm 0.33$). Correlation coefficients for individual transcripts, however, ranged from r = -0.68 to 0.76 for Cold-condA, from r = -0.89 to 0.73 for Warm-condA and from$r$ = -0.42 to 0.55 for Ad-Pl category. The cumulative distributions of correlation coefficients for all categories were not significantly different from their respective backgrounds (Cold-condA:$D = 0.073$, $p = 0.817$; Warm-condA: D = 0.065,$p = 0.11$; Ad-Pl: $D = 0.17$, $p = 0.39$). Similar to the case of ancestry, we identified several transcripts for the Warm-condA and Cold-condA categories as significantly associated with survival ($p < 0.05$), although none passed the multiple testing correction. No transcript in the Ad-Pl category was significantly associated with survival.

While no GO terms were enriched in either Q_{ST}categories, significant depletion was noted for GO terms related to signal transduction and cell communication in Cold-condA, and for hydrolase activity in Warm-condA (Table S2). Several transcripts related to freeze and drought tolerance, as well as cell wall modulation were classified under the Cold-condA category. These included *VIN3* ,*PPR* , *XTH* and *Aquaporins* such as TIP1 (Sung & Amasino, 2005; Raimund, 2015; Tucker *et al* ., 2018). Similarly, the Warm-condA category contained transcripts related to drought stress response and photosynthesis, including numerous *ERF* family genes,*SWEET* , *WRKY, psbE* and multiple auxin responsive elements, such as *ARF2* (Lata *et al* ., 2015; Zhang *et al* ., 2020). Transcripts classified in the Ad-Pl category, included auxin responsive elements, such as *SAUR50* (Sun *et al* ., 2016), which is critical for light signaling, *α-εξπανσιν* and ABA signaling pathway family genes, such as *LPPD* , which have known roles in numerous plant developmental and stress response pathways (Marowa *et al* ., 2016). Thus, even without clear relationships between population-level transcript abundances, survival, and genomic ancestry by Q_{ST} category, there were several notable examples of functionally sensible genes in eachQ_{ST} category consistent with adaptive responses to drought and freeze gradients. Together with the analyses presented in the previous section these results provide support for hypothesis H1 and H2 at the per-transcript level but render minimal support for H3.

Drivers of adaptive evolution at the co-expression module level

At the module level across both gardens, we noted a wide range of eigengene correlations with ancestry, ranging from r = -0.56 to 0.35 at the warm garden and from r = -0.32 to 0.41 at the cold garden (Table S3). Compared to the association with ancestry, eigengene expression was generally more strongly related to survival and ranged from -0.50 to 0.72 at both gardens (Table S3). At the warm garden, for example, the ME11 module exhibited a significant ($p < 0.05$) negative association with ancestry, as well as with survival (Table 2), however the association with ancestry was not significant after multiple test corrections. Similarly at the cold garden, the ME13 module exhibited a significant ($p < 0.05$) but positive association with ancestry and survival, however the association with ancestry was again not significant after multiple test corrections

(Table 2). Similar to the case at the per-transcript level, the results at the module level do not support H3.

GO enrichment analysis for modules exhibiting at least one of the following - Q_{ST} category enrichment, Q_{ST} category depletion, or strong correlation with ancestry and/or survival -identified a wide range of functional groups across both gardens (Table 2, Table S2). Across both gardens, Q_{ST} category enrichment modules were predominantly associated with GO terms and genes related to stress response, cell wall modulation and response to light stimuli. In contrast, Q_{ST} category depleted modules across gardens were predominantly represented by developmental and metabolic processes. Functional terms associated with metabolic processes were generally under-represented, and terms associated with response to abiotic stress were generally over-represented in the modules strongly associated with survival and ancestry at the warm garden (e.g., ME11 module), as well as at the cold garden (e.g., ME13 module). Across both gardens, modules exhibiting the strongest association with survival (i.e., the ME26 module at the cold garden and ME19 at the warm garden) were primarily associated with GO terms related to glutathione transferase, detoxification and response to abiotic stimuli.

Relationship between network connectivity and adaptive evolution

Across both gardens, Q_{ST} was positively associated with module connectivity (kWithin; Cold garden: $r = 0.36$, $p < 0.0001$; Warm garden: $r = 0.46$,$p < 0.0001$) and with total network connectivity (kTotal; Cold garden $r = 0.53$, $p < 0.0001$; Warm garden: $r = 0.56$, $p < 0.0001$). Additionally, transcripts in the Warm-condA and Cold-condA categories were significantly over-represented at the core of their respective modules (Warm garden enrichment = 1.7, $p < 0.0001$; Cold garden enrichment = 2.5, $p < 0.0001$), as well as at the core of the garden-specific network (Warm garden enrichment = 2.0, $p< 0.0001$; Cold garden enrichment = 3.9, $p < 0.0001$) (Fig. 4; Fig. S3). This pattern persisted even after accounting for differences in expression levels between the core and the periphery transcripts. For most modules exhibiting strong population differentiation, the transcript with highest Q_{ST}also had the highest module membership and hence were located at the module core (Fig. S3). The average Q_{ST} for core transcripts, however, was higher than that of the periphery for only 38% of the strong differentiated modules at the cold garden and for only 20% of the differentiated modules at the warm garden (Table S3). Overall, results from our module and network level relative representation analyses do not support H4.

Only five modules out of the 31 identified at the cold garden and four out of the 26 modules identified at the warm garden were strongly preserved across gardens using our joint threshold (Table S3; Fig. 4). Generally, modules enriched for Q_{ST} categories were weakly preserved across gardens, while Q_{ST}depleted modules had consistently strong preservation across gardens (Fig. 4a,c) providing further support for H2 at the module level.

Discussion

Hybrid zone populations exhibit adaptive potential to novel climatic conditions

Current patterns of genetic diversity in populations of forest trees are outcomes of various evolutionary processes structured by past events (Hampe & Petit, 2005; Mayol *et al* ., 2015), which often results in extant populations exhibiting adaptational lags to their current climate conditions. Although this has been documented for several tree taxa (Seliger *et al* ., 2021), it does not necessarily mean these populations are devoid of adaptive potential. Theory suggests that GEI should be prevalent in populations exhibiting adaptational lags and those occurring in heterogenous environments (Via & Lande, 1985; Ghalambor *et al* ., 2007). Sequential founder events, high levels of landscape fragmentation and the associated loss of genetic diversity, however, can restrict the evolution of GEI (Schmid *et al* ., 2019).

The populations sampled in this study naturally occur across the fragmented landscapes of the southwestern United States, where they experience markedly different seasonal and annual means and fluctuations in environmental conditions, as well as gene flow from a northern congener, *P. flexilis* (Menon *et al* ., 2018). By treating gene expression patterns and resulting co-expression modules as quantitative traits, we revealed strong signals of garden-specific adaptive trait differentiation (i.e., GEI) thus supporting our first and second

hypotheses. In general, our estimates ofQ_{ST} agree with previously published estimates from studies of forest trees (Lind *et al* ., 2018), and more specifically with the estimates obtained using expression datasets (Roberge *et al* ., 2007; Leder *et al* ., 2015). Although the presence of neutral population structure in the transcriptomic dataset used to construct the co-expression modules could generate false positives, eight of the strongly differentiated modules were also enriched for various Q_{ST} categories (Fig. 4). We suggest these are likely true candidates for garden-specific signatures of adaptive evolution, given the overall weak population structure in this hybrid zone (Menon *et al* ., 2018).

The signals of local adaptation noted under the novel environmental conditions of our study suggest that populations of long-lived tree species, such as conifers, might not be limited in their ability to adapt to rapidly changing climatic conditions. Our results contrast with predictions of extensive future maladaptation suggested for other long-lived tree species such as oak (Browne *et al* ., 2019), poplar (Fitzpatrick *et al* ., 2020; Gougherty *et al* ., 2021) and spruce (Frank *et al* ., 2017). This contrast may relate to methodological differences among past studies and ours. Specifically, by using a space-for-time substitution study we allow for the architecture of adaptive evolution to reflect response to novel conditions. However, similarities between the space-for-time substitution design used in the present study and by Fitzpatrick *et al* . (2020) and Browne*et al* . (2019) suggest that our contrasting results might follow from the inclusion of hybrid populations, which matters because hybrid populations frequently show increased additive genetic variance (Reif *et al* ., 2007; Kulmuni, Wiley & Otto, 2023) (Table 2). Nevertheless, we advise caution in the interpretation that populations of long-lived tree species may be more adaptable to novel climates than expected because the presence of additive genetic variation underlying climatically relevant traits is only one of the important conditions needed for an adaptive response. Correlations between traits can be modulated by environmental conditions (Wood & Brodie, 2015) and correlations antagonistic to the direction of selection can impede adaptive responses (Walsh & Blows, 2009). While this does not seem to be the case in our study, as is evident throughQ_{ST} enrichment at the core of the networks and strong co-expression module differentiation, we cannot conclude that evolution is occurring in the direction of the novel selective pressure. Such a conclusion would require a larger sample size and more thorough quantitative evaluation of the multivariate trait space.

Drought and hybrid ancestry influence adaptive evolution and generate GEI

Results from our transcript categorisation (Fig. 2) and co-expression module enrichment analyses (Fig 4) support the second hypothesis of garden-specific trait differentiation. Despite the more arid conditions at the warm garden, the marginal effect of drought on population-level transcript differentiation was significant only at the cold garden. The strong impact of drought on population differentiation found here conforms with previous studies conducted in *P. strobiformis* (Goodrich *et al* ., 2016; Bucholz *et al* ., 2020; Menon *et al* ., 2021). While drought was an important selective pressure in both gardens, its effect may have been exaggerated at the cold garden because of freeze-related rupturing of cell wells during the previous winter season (Bachofen *et al* . 2015) leading to elevated susceptibility to drought stress. Given the overlap in molecular pathways leading to stress tolerance (Blödner et al. 2005) such carry-over effects or cross-susceptibility may be common across long-lived plants, although the opposite indicating higher tolerance to the second stressor has also been noted (Kong & Henry, 2019).

The absence of association with freezing stress could be due to global rather than local adaptation (Booker *et al* ., 2020) aided by expression of variants introgressed from *P. flexilis* , strong confounding between geography and freeze-related variables or compounded effects of drought and freeze stress in the cold garden. Ongoing work demonstrating higher post-winter survival in the cold garden when compared to the warm garden (Moler ERV. 2020) as well as a significant association between *P. flexilis* ancestry and survival in the cold garden provide some evidence for garden-specific effect of introgressed variants. At the network level, we noted very limited preservation across gardens using our joint threshold (Table S3; Fig. 4). The modules that were strongly preserved consisted of transcripts with low garden specific population differentiation and were generally related to metabolic processes that are likely essential for basic organismal functioning. Overall, this highlights strong and consistent patterns of GEI.

At the multivariate level, modules correlated with ancestry were often also strongly correlated with survival (Table S3; Table 2). Specifically, two Q_{ST} enriched and weakly preserved modules – ME24 (warm garden) and ME13 (cold garden) – were strongly associated with *P. flexilis* ancestry, as well as with survival, making the traits and the loci underlying them a key candidate for further detailed studies of GEI (Fig. 4; Table 2). Furthermore, none of the transcripts encompassing these modules were shared across the two gardens. Since we sampled the transcriptomes of juvenile hybrid trees, it is possible that some of the strongly differentiated traits documented here are involved in post-zygotic isolating barriers (Lindtke*et al* ., 2014; Zhao *et al* ., 2014). Overall associations between ancestry and expression levels of the three Q_{ST} categories, however, were not significant. There are at least two potential explanations for these results. First, ancestry estimates were obtained from a ddRADseq dataset, which although is representative of overall genomic ancestry, provides less coverage of genic regions in species with large genomes such as pines (Parchman*et al* ., 2018). Second, ancestry was obtained for maternal trees and correlated with the maternal tree's genetic values that were estimated using only the surviving seedlings. Given that postzygotic barriers are often expressed in seedling stage (Ogasawara & Okubo, 2009) resulting in high mortality, it is likely that our design could not fully capture the impact of ancestry on maternal tree's genetic expression value as we only used the surviving seedlings without strong incompatibilities.

Interestingly, the proportion of *P. flexilis* ancestry was positively associated with survival across both gardens, but significant only at the cold garden. Survival and ancestry remained uncorrelated to population-level transcript abundances. These results thus provide only partial support for the third hypothesis of hybrid ancestry interacting with GEI and highlight the need for an in-depth evaluation mapping expression difference between parentals and hybrid trees in novel environments. Nevertheless, drawing from our previous work (Menon*et al* ., 2021) and the likely wider climatic tolerance of *P. flexilis* (Windmuller-Campione & Long, 2016), we speculate that the significant correlation between ancestry and survival only at the cold garden is indicative of *P. flexilis* -like variants enhancing survival in cooler environments, which aligns with the higher latitudinal distribution of the species range of *P. flexilis* than*P. strobiformis* .

Transcriptome wide profiling illustrates the prevalence of pleiotropic architectures

Strongly correlated traits within biological networks often experience selective constraints, as variants occurring in the genes underlying these traits could cause deleterious pleiotropic effects (Jordan*et al* ., 2004). On the other hand, traits with lower connectivity will often be involved in adaptive evolution and in GEI, as selection can fine tune population responses to local environments without disrupting key functional components (Cork & Purugganan, 2004; Josephs*et al* ., 2017). Despite this, our work demonstrates an overrepresentation of GEI transcripts at the core of networks and of the modules, thereby leading us to reject our fourth hypothesis of adaptive evolution being dominated by weakly connected peripheral traits. Following Fisher's geometric model (Fisher, 1930), if populations are further from their fitness optima, variants that impart pleiotropic influences are advantageous during the initial stages of the adaptive walk (Orr, 1998). We suggest that this is likely the case for the expanding hybrid zone populations (Menon *et al* ., 2020) that are exposed to novel selective pressures. Under this scenario of demographic and climatic mismatch, pleiotropy speeds up evolution to fitness optima. However, our study uses a narrow definition of pleiotropy restricted to the number of associated expression traits which could have a multitude of different functions. While gene expression is widely treated as a quantitative trait, its association with other genes within a co-expression network may not always reflect pleiotropy and could be confounded by the effect size of unaccounted for cis-variants (Josephs*et al* ., 2017. However, considering the large genome size of conifers, and that most variants in our study are located in non-genic regions, this would be a challenging task to accomplish without further improvement in genomic resources.

Conclusions

Studies focusing on seedling adaptive potential to multivariate stressors mimicking changing climatic conditions are timely given the focus on assisted migration efforts. A multitude of common garden experiments and genome-wide association studies provide remarkable evidence for local adaptation and GEI (Langlet, 1971; Savolainen*et al* ., 2013). Here, we demonstrate strong evidence of adaptive evolution and GEI at both

the per-transcript and co-expression module levels, with partial evidence that these patterns interact with interspecific gene flow through the survival component of fitness. We further highlight that populations of long-lived tree species may be more able than expected to respond to rapidly changing climatic conditions through ample genetic variation underlying expression traits and a modulation of genetic architecture depending on the degree and direction of environmental change. This response, moreover, appears to be provided by pleiotropic loci underlying strongly connected adaptive expression traits. The generality of our primary finding of rapid adaptive potential aided by pleiotropic loci additionally awaits further work across phylogenetically diverse tree species distributed across various climatic gradients with differing demographic histories.

Data availability statement:

Raw RNAseq reads are available at the NCBI SRA Archive (SRAxxxx). Matrix of gene expression and all analysis scripts will be available through the lead author's GitHub account (https://github.com/mitramenon/Ecology_expression_Pine).

Acknowledgments:

This work is supported by the National Science Foundation EF-1442597 to KW, EF-1442486 to AE and Rice Rivers research scholarship to MM. We also acknowledge NSF Grant No.1126840, which funded the Southwest Experimental Garden Array used in this research. We thank the field crew for help setting up the common gardens and specifically Dr. Ethan Bucholz for his continued support and guidance through sampling design and common garden establishment. We also thank the computational resources available through the VCU HPC grid engine system.

Author contributions:

EM, JS, KW and AW designed and planted the gardens. MM and AE conceived and designed the study. MM and JS performed the analysis. MM wrote the manuscript with edits from all co-authors. The final version has been approved for submission by all co-authors.

Table 1 : Redundancy analyses (RDA) based modelR^2 and significance of the multivariate models along with representative candidate genes for the threeQ_{ST} categories. Candidates were determined manually based on the annotations from EnTAP (see Table S1 & Methods S1) by focusing on transcripts with highest Q_{ST}values in each category.

Category	**Model**	R^2	R^2_{adj}	p**-value**	**Candidates**
Cold-condA	Drought + Freeze + Geography	0.88	0.53	0.003	WRKY4, CYC4, XTH9, LRR, VIN3, CO-like,
Warm-condA	Drought + Freeze + Geography	0.81	0.17	0.20	CyP450, ARF2, ERF2, SWEET, Ycf3, WRKY
Adap-Pl	Drought + Freeze + Geography	0.82	0.19	0.39	SAUR50, α-expansin, ERG3, LPPD

Table 2 : Candidate modules associated with survival, ancestry and/or exhibiting Q_{ST} enrichment along with the annotation of the transcript with the highest or the next highest module membership (hMEM) in cases where the topmost transcript was lacking a clear annotation. Q-values are given for p-values < 0.05.

Garden	**Module ID**	**Survival (*r*)**	**Ancestry (*r*)**	***E*nrichment value (RR)**	**hMEM annotation**
COLD	ME13	0.62 (p = 2e-04, q = 0.003)	0.41 (p = 0.02, q = 0.23)	2.87	Golgi intermediate protein 3
	ME26	0.71 (p = 1e-05, q = 3e-04)	0.29 (p = 0.01, q = 0.23)	0.64	WRKY14

	ME16	-0.19 ($p = 0.32$)	-0.17 ($p = 0.36$)	1.37	Non-race specific disease resistance-1 (NDR1)
	ME12	-0.18 ($p = 0.34$)	-0.11 ($p = 0.55$)	2.12	Snakin-2
	ME14	0.29 ($p = 0.12$)	0.049 ($p = 0.79$)	2.12	E3 ubiquitin-protein ligase (KEG)
WARM	ME24	0.53 ($p = 3e\text{-}03$, $q = 0.02$)	0.23 ($p = 0.22$)	3.96	β-carotene hydrolase (CHY-β)
	ME19	0.69 ($p = 3e\text{-}05$, $q = 3e\text{-}04$)	0.35 ($p = 6e\text{-}02$, $q = 0.40$)	0.99	Light regulated WD1
	ME8	-0.27 ($p = 0.20$)	-0.56 ($p = 1e\text{-}03$, $q = 0.03$)	0	Trehalose-6-phosphate synthase 1
	ME11	-0.43 ($p = 0.02$, $q = 0.04$)	-0.47 ($p = 9e\text{-}03$, $q = 0.12$)	1.62	E3-ubiquitin protein ligase (UPL5-like)
	ME2	0.43 ($p = 0.02$, $q = 0.04$)	0.17 ($p = 0.36$)	1.41	Phospholipase C Gene Family-like (PI-PLC)
	ME1	-0.24 ($p = 0.20$)	-0.017 ($p = 0.90$)	2.10	RNA binding protein
	ME3	-0.22 ($p = 0.23$)	0.11 ($p = 0.60$)	2.02	Filament like protein (FPP7)
	ME15	0.3 ($p = 0.08$)	0.21 ($p = 0.3$)	1.72	ubiquitin-like modifier (ULP)

Figures

Figure 1: Location of sampled populations, common gardens and the geographic range of the hybridizing species. Inset shows the study area on an elevation map.

Figure 2: Schematic representation of Q_{ST} categories.

Figure 3: Distribution of transcript expression Q_{ST} values across gardens.

Figure 4: Relationship between Q_{ST} transcript enrichment, module size and module preservation measures across gardens (WARM: Low elevation garden – panels **a & b**); (COLD: High elevation garden – panels **c & d**). The vertical dotted line indicates cutoff for declaring a module as preserved using *medianRank* and $Z_{summary}$ score. Only modules significantly over-enriched or under-enriched for Q_{ST} outlier transcripts have been labelled.

Supporting information

Methods S1: Details of the pipeline and summaries used for *de novo* transcriptome assembly.

Methods S2: Pipeline details for SNP calling and filtering using ddRAD-seq genotyping.

Methods S3 : Details of steps and summary statistics for construction of the co-expression networks.

Figure S1: a) PCA of 76 environmental variables from climateWNA for the normal 1981-2010 for the 10 sampled populations and for the study year 2019 for the common gardens. **(b)** Heatmap representing the proportion of times drought or freeze related environmental variables measured at the gardens in 2019 were

outside the range experienced by the assayed population across a 60-year period. Garden environment was declared as an outlier if it fell outside the 95 percent confidence interval of the climatic condition experienced by the populations for each year. This evaluation was done on a year-by-year basis. Warmer (red) colours indicate higher drought intensity and colder (blue) colours indicate higher freeze events experienced by the gardens relative to the assayed populations.

Figure S2: a) Geographical range map of *P. strobiformis*and *P. flexilis* along with the populations sampled for ddRAD-seq genotyping. **b)** STRUCTURE plot of genomic ancestry for the 30 maternal trees represented by the mixed-sicold sampled for RNA-seq and**c)** Distribution of F_{ST} values for the 30 maternal trees.

Fig S3: Visualization of Q_{ST} enriched modules at the cold and warm gardens with circles representing the nodes (transcripts) and lines representing the edges (correlation between transcripts). Transcripts classified as Cold-condA or Warm-condA are highlighted in color while the rest are represented by black circles. The location of the transcripts in the module are representative of their edge weights (i.e intra-modular connectivity).

Fig S4: Relationship between mean transcript expression and co-expression connectivity levels for the **a)** Warm and**b)** Wold garden. Transcripts were divided into bins of mean expression by their position in the co-expression network (core vs. periphery) to highlight their level of connectivity.

Figure S5: A) Evaluation of scale free topology at various soft-thresholding powers and **B)** clustering dendrogram of transcripts using a soft-thresholding of 8 for the Cold garden and 9 for the Warm garden.

Table S1 : Annotations for all 47,651 transcripts generated through EnTAP.

Table S2: Summary of GO terms and the enrichment analyses for three Q_{ST} categories (Cold-condA, Warm-condA, Ad-Pl) and for all modules across the two gardens.

Table S3: Summary statistics for modules identified across the gardens. Columns represent eigengene association with survival and ancestry, Q_{ST} enrichment score, annotation of the transcript with the highest module membership, preservation scores and difference in mean Q_{ST} between transcripts classified in the core and periphery of a module (Q_{ST} periphery - Q_{ST}core).

REFERENCES

Adams HD, Kolb TE. 2004 . Drought responses of conifers in ecotone forests of northern Arizona: tree ring growth and leaf δ 13 C.*Oecologia.* **140** : 217– 225.

Aitken SA, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications* . **1:**95-111.

Barabási AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet.***5(2):** 101–13.

Barghi N, Hermission J, Schlotterer C. 2020 . Polygenic adaptation: a unifying framework to understand positive selection.*Nature Review Genetics.* **21** (12):769-781.

Bates D, Mächler M, Bolker B, Walker S. 2015 . Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* ,**67** (1), 1–48.

Battlay P, Yeaman S, Hodgins KA . **2023** . Impacts of pleiotropy and migration on repeated genetic adaptation. BioRxiv. https://doi.org/10.1101/2021.09.13.459985

Bisbee J. 2014. *Cone morphology of the Pinus ayacahuite-flexilis complex of the southwestern United States and Mexico* . Bulletin of the Cupressus conservation project, 3, 3–33.

Blödner, C., T. Skroppa, O. Johnson and A. Polle. 2005. Freezing tolerance in two Norway spruce (Picea abies [L.] Karst.) progenies is physiologically correlated with drought tolerance. *J. Plant Physiol* . **162** :549–558.

Browne L, Wright JW, Fitz-Gibbon S, Gugger PF, Sork VL. 2019. Adaptational lag to temperature in valley oak (Quercus lobata) can be mitigated by genome-informed assisted gene flow. *Proc Natl Acad Sci USA.* **116** (50):25179-25185.

Bucholz ER, Waring KM, Kolb TE, Swenson JK, and Whipple AV. 2020. Water relations and drought response of *Pinus strobiformis* . *Canadian Journal of Forest Research* .**50(9)** : 905-916.

Chevin L, Collins S, Lefèvre F. 2013 . Phenotypic plasticity and evolutionary demographic responses to climate change: taking theory out to the field. *Funct. Ecol* . **27** : 967–979.

Clausen Jens, Keck David D, Hiesey William M. 1948 .*Experimental studies on the nature of species. III: Environmental responses of climatic races of Achillea* . 581; Washington, D.C.: Carnegie Institution of Washington.

Cork JM, Purugganan MD. 2004. The evolution of molecular genetic pathways and networks. *Bioessays.* **26(5):** 479-84.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E *et al.* 2011. The Variant Call Format and VCFtools, *Bioinformatics* .**27(15)** :2156-8

Dayan DI, Crawford DL, Oleksiak MF. 2015. Phenotypic plasticity in gene expression contributes to divergence of locally adapted populations of Fundulus heteroclitus. *Molecular Ecology* .**24(13)** :3345-59.

Des Marais DL, Guerrero RF, Lasky JR, Scarpino SV. 2017. Topological features of a gene co-expression network predict patterns of natural diversity in environmental response. *Proceedings of Royal Society of Biological Science* s. **284(1856)** :20170914

Etterson JR, Shaw RG. 2001 . Constraint to adaptive evolution in response to global warming. *Science* **294** : 151–154.

Falconer DS & Mackay T. 1996 . *Introduction to Quantitative Genetics* . Longman, New York.

Fagny M, Austerlitz F. 2021. Polygenic Adaptation: Integrating Population Genetics and Gene Regulatory Networks. *Trends in Genetics.* **37** (7):631-638.

Fisher RA. 1930. *The genetic theory of natural selection* . Oxford: The Carendon Press.

Fitzpatrick MC, Chhatre VE, Soolanayakanahally RY, Keller SR. 2021. Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests.*Molecular Ecology Resources* . **21** (8):2749-2765.

Franks SJ, Weber JJ, Aitken SN. 2014 . Evolutionary and plastic responses to climate change in terrestrial plant populations.*Evol. Appl. 7* : 123–139.

Geber MA, Eckhart VM. 2005. Experimental studies of adaptation in Clarkia xantiana. II. Fitness variation across a sucoldpecies border.*Evolution* **59(3)** :521–31.

Ghalambor CK, McKay JK, Carroll SP, Reznick DN. 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional ecology.*2007. **21(3**):394-407.

Goodrich BA, Waring KM & Kolb TE. 2016. Genetic variation in*Pinus strobiformis* growth and drought tolerance from southwestern US populations. *Tree Physiology* . **36:** 1219-1235.

Goudet J. 2005 . hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes* , **5** , 184–186.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA *et al.* 2013. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nature Biotechnology.* **29(7)** : 644–652.

Grote S. 2020. *GOfuncR: Gene ontology enrichment using FUNC* . R package version 1.10.0.

Haas, B., Papanicolaou, A., Yassour, M. *et al.* 2014.*De novo* transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. *Nature Protocols* .**8:** 1-43.

Hämälä T, Gorton AJ, Moeller DA, Tiffin P. 2021. Pleiotropy facilitates local adaptation to distant optima in common ragweed (*Ambrosia artemisiifolia*). *PLOS Genetics* **16(3)** : e1008707.

Hampe A. & Petit RJ. 2005 . Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters* .**8:** 461–467.

Hart AJ, Ginzburg S, Xu M, *et al.* 2020. EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Molecular Ecology Resources* . **20** : 591-604.

Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999). From molecular to modular cell biology. *Nature* . **402** : C47-52.

Hoffman GE and Schadt EE. 2016. variancePartition: interpreting drivers of variation in complex gene expression studies. *BMC Bioinformatics* **17** : 483.

Holliday JA, Ralph SG, White R, Bohlmann J, Aitken SN. 2008. Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (Picea sitchensis).*New Phytologist* . **178** :103–22.

Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH 2003. From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany***81** : 1247–1266.

Jordan IK, Marino-Ramirez L, Wolf YI & Koonin EV. 2004. Conservation and Coevolution in the Scale-Free Human Gene Coexpression Network. *Molecular Biology and Evolution* ,**21(11)** :2058–2070.

Josephs EB, Wright SI, Stinchcombe JR, Schoen DJ. 2017. The relationship between selection, network connectivity, and regulatory variation within a population of Capsella grandiflora. *Genome Biology Evolution* . **9(4):** 1099–1109.

Kong, R, Henry, H. 2019. Cross acclimation effects of spring freezing and summer drought on plant functional groups and ecosystem properties. *Environmental and Experimental Botany* . **164** : 52-57.

Kulmuni J, Wiley B & Otto S 2023. On the fast track: hybrids adapt more rapidly than parental populations in a novel environment, *Evolution Letters* . 1-9

Krueger F. Trim Galore!: 2015. *A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files.http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/*

Lande R. 1980. The Genetic Covariance between Characters Maintained by Pleiotropic Mutations. *Genetics* . **94** (1), 203–215.

Langlet O. 1971. Two hundred years of genecology. *Taxon***20** :653–721

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* . **9** :559.

Langfelder P, Luo R, Oldham MC, Horvath S. 2011. Is My Network Module Preserved and Reproducible? *PLoS Computational Biology* .**7(1):** e1001057.

Langmead B & Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods.* **9** :357-359.

Lata C., Muthamilarasan M., Prasad M. 2015. Drought Stress Responses and Signal Transduction in Plants. In: Pandey G. (eds)*Elucidation of Abiotic Stress Signaling in Plants.* Springer, New York, NY.

Law CW, Chen Y, Shi W, Smyth GK 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology.* **15:** R29.

Leder EH, McCairns RJ, Leinonen T, Cano JM, Viitaniemi HM, Nikinmaa M, Primmer CR, Merilä J. 2015. The Evolution and Adaptive Potential of Transcriptional Variation in Sticklebacks—Signatures of Selection and Widespread Heritability. *Molecular Biology and Evolution* . **32(3)** :674–689.

Leinonen T, O'Hara RB, Cano JM and Merilä J. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biolog* y.**21** :1–17.

Lind BM, Menon M, Bolte CE, Faske TM & Eckert AJ. 2018 .The genomics of local adaptation in trees: Are we out of the woods yet?. *Tree genetics & genomes.* **14(2):** 29.

Lindtke D, Gompert Z, Lexer C, Buerkle CA. 2014 . Unexpected ancestry of Populus seedlings from a hybrid zone implies a large role of postzygotic selection in the maintenance of species. *Molecular Ecology.* **23** : 4316-4330.

MacPherson, A., P. A. Hohenlohe, and S. L. Nuismer. (2015). Trait dimensionality explains widespread variation in local adaptation.*Proc. R. Soc. B* **282** : 20141570.

Mähler N, Wang J, Terebieniec BK, Ingvarsson PK, Street NR, Hvidsten TR. 2017. Gene co-expression network connectivity is an important determinant of selective constraint. *PLoS Genetics* .**13(4)** :e1006402.

Marowa P, Ding A, Kong Y. 2016. Expansins: roles in plant growth and potential applications in crop improvement. *Plant Cell Reports.* **5(5)** :949-965.

Mayol M, Riba M, Gonzalez-Martinez SC et al. 2015 . Adapting through glacial cycles: insights from a long-lived tree (Taxus baccata).*New Phytologist* . **208** : 973–986.

Mei W, Stetter MG, Gates D, Stitzer MC, Ross-Ibarra J. 2018. Adaptation in plant genomes: bigger is different. *American journal of botany* . **104** (1): 16-19.

Menon M, Landguth E, Leal-Saenz A *et al.* 2019. Tracing the footprints of a moving hybrid zone under a demographic history of speciation with gene flow. *Evolutionary Applications* .**13(1):** 195-209.

Menon M, Bagley JC, Page GFM. *et al.* 2021. Adaptive evolution in a conifer hybrid zone is driven by a mosaic of recently introgressed and background genetic variants. *Communications Biol* ogy. **4:** 160

Menon M, Bagley JC, Friedline CJ, Whipple AV, *et al.*2018 . The role of hybridization during ecological divergence of southwestern white pine (*Pinus strobiformis*) and limber pine (*P. flexilis*). *Molecular Ecol* ogy. **27** : 1245–1260.

Moler, E.R.V., 2020. Climate change studies of Southwestern White Pine: Phenotypic and life history trait responses to temperature manipulations during seed development and seedling growth.

Doctoral dissertation, Northern Arizona University.

Mollie E. Brooks, Kasper Kristensen, Koen J. van Benthem, Arni Magnusson, Casper W. Berg, Anders Nielsen, Hans J. Skaug, Martin Mächler, Benjamin M. Bolker 2017. Modeling zero-inflated count data with glmmTMB. *BiorXiv.https://doi.org/10.1101/132753.*

Pickles, B.J., Twieg, B.D., O'Neill, G.A., Mohn, W.W. and Simard, S.W., 2015. Local adaptation in migrated interior Douglas-fir seedlings is mediated by ectomycorrhizas and other soil factors.*New Phytologist* , **207(3)** : 858-871.

Pickett STA. 1989. *Space-for-time sucoldtitution as an alternative to long-term studies* . In: Long-term studies in ecology: approaches and alternatives. Springer-Verlag, New York, New York, USA. 110–135.

Reed TE., Schindler DE., & Waples RS. (2011) . Interacting effects of phenotypic plasticity and evolution on population persistence in a changing climate. *Conservation Biology.* **25(1):**56-63.

Reif JC, Gumpert FM, Fischer S, Melchinger AE. 2007. Impact of interpopulation divergence on additive and dominance variance in hybrid populations. *Genetics* . **176(3):** 1931-1934.

Roberge C, Guderley H and L. Bernatchez. 2007. Genome wide Identification of Genes Under Directional Selection: Gene Transcription Q_{ST} Scan in Diverging Atlantic Salmon Subpopulations. *Genetics* **177** : 1011–1022.

Savolainen O, Lascoux M, Merila J. 2013. Ecological genomics of local adaptation. *Nature Reviews Genetics* . **14:** 807–820.

Schmid M, Dallo R & Guillaume F. 2019. Species' range dynamics affect the evolution of spatial variation in plasticity under environmental change. *The American Naturalist* . **193** : 798-813.

Seliger, B. J., McGill, B. J., Svenning, J. C., & Gill, J. L. 2021. Widespread underfilling of the potential ranges of North*American trees. Journal of Biogeography* . **48(2):** 359-371.

Shaw RG, Etterson JR. 2012 Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytol* . **195** : 752–765.

Skotte L Korneliussen TS & Albrechtsen A. 2013. Estimating individual admixture proportions from next generation sequencing data.*Genetics* **195:** 693–702.

Spitze, K. 1993. Population-structure in Daphnia-obtusa - quantitative genetic and allozymic variation. *Genetics***135** : 367–374.

Sung S. & Amasino R.M. 2004. Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature***427** : 159–163.

Squillace AE. Average genetic correlations among offspring from open-pollinated forest trees. *Silvae Genetica.* 23: 149- 156 (1974).

Swenson, Jared Karl. 2021. *Environment specific alleles related to adaptive seasonal traits and the fitness consequences of phenology for Pinus strobiformis.* Masters thesis, Northern Arizona University.

Tucker MR, Lou H, Aubert MK, Wilkinson LG, Little A, Houston K, Pinto SC, Shirley NJ. 2018. Exploring the Role of Cell Wall-Related Genes and Polysaccharides during Plant Development. *Plants (Basel)* . **7(2)** :42.

Ogasawara O & Okubo K. 2009. On theoretical models of gene expression evolution with random genetic drift and natural selection.*PLoS One* **4** :e7943

Oksanen, O., Blanchet, F.G., Kindt, R. *et al.* 2013.*vegan: Community Ecology Package*** . R package version 2.5-2.

Orr HA. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution.*Evolution* . **52** : 935–949.

Puritz JB, Hollenbeck CM & Gold JR. 2014. dDocent: a RADseq, variant -calling pipeline designed for population genomics of non -model organisms. *PeerJ,* **2** : e431.

R Core Team. R v.3.6.3: A Language and Environment for Statistical Computing. Vienna, Austria : *The R Foundation for Statistical Computing. R Foundation for Statistical Computing* , Vienna, Austria (2020)

Robinson MD, McCarthy DJ, Smyth GK. 2010 . edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* , **26(1**), 139-140.

Via S, Lande R. 1985. Genotype-environment Interaction and The Evolution Of Phenotypic Plasticity. *Evolution.***39(3)** :505-522.

Wang T, Hamann A, Spittlehouse DL, & Carroll C. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. *PLoS One* **11** : e0156720.

Whitehead A and DL Crawford. 2006. Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology* . **15** : 1197–1211.

Willi, Y., van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the Adaptive Potential of Small Populations. *Annual Review of Ecology, Evolution, and Systematics* . **37** : 433–458.

Windmuller-Campione MA & Long JN, 2016. Limber Pine (Pinus flexilis James), a Flexible Generalist of Forest Communities in the Intermountain West. *PloS One.* **11(8)** , 1–16.

Zhang Y, Zhang A, Li X, Lu C. 2020. The Role of Chloroplast Gene Expression in Plant Responses to Environmental Stress.*International Journal of Molecular Sciences* .**21(17)** :6082.

Zhao W, Meng J, Wang B, Zhang L, Xu Y, Zeng QY, Li Y, Mao JF, Wang XR 2014. Weak Crossability Barrier but Strong Juvenile Selection Supports Ecological Speciation of the Hybrid Pine Pinus Densata on the Tibetan Plateau. *Evolution* . **68(11)** : 3120-3133.