**Genetic differentiation across a steep and narrow environmental gradient: Quantitative genetic and genomic insights into Lake Superior populations of *Quercus rubra***

María José Gómez Quijano1, 2, Briana L. Gross2, Julie R. Etterson2

1Queen’s University, Department of Biology, Biosciences Complex 4425, 116 Barrie St, Kingston, ON K7L 3N6

2University of Minnesota Duluth, Department of Biology, Swenson Science Building 207, 1035 Kirby Drive, Duluth, MN 55812

Corresponding author: 21mjgq@queensu.ca / gomez312@d.umn.edu

**ABSTRACT**

Adaptive differentiation of traits and underlying loci can occur at a small geographic scale if natural selection is stronger than countervailing gene flow and drift. We investigated this hypothesis using coupled quantitative genetic and genomic approaches for a wind-pollinated tree species, *Quercus rubra*, along the steep, narrow gradient of the Lake Superior coast that encompasses four USDA Hardiness Zones within 100 km. For the quantitative genetic component of this study, we examined phenotypic differentiation among eight populations in a common garden, measuring seed mass, germination, height, stem diameter, leaf number, SLA, and survival. For the genomic component, we quantified genetic differentiation for 26 populations from the same region using RAD-seq. Because hybridization with *Quercus ellipsoidalis* occurs in other parts of the species’ range, we included two populations of this congener for comparison. In the common garden study, we found a strong signal of population differentiation that was significantly associated with at least one climate factor for nine of ten measured traits. In contrast, we found no evidence of genomic differentiation among populations based on *FST* or any other measures. However, both distance-based and genotype-environment association analyses identified loci showing the signature of selection, with one locus in common across five analyses. This locus was associated with the minimum temperature of the coldest month, a factor that defines the climate zones and was also significant in the common garden analyses. In addition, we documented introgression from *Q. ellipsoidalis* into *Q. rubra*, with rates of introgression correlated with the climate gradient. In sum, this study reveals signatures of selection at the quantitative trait and genomic level consistent with climate adaptation, a pattern that is more often documented at a much broader geographic scale, especially in long-lived wind-pollinated species.

**KEY WORDS:** Population divergence, hybridization, geographic variation, *Quercus*, selection

**1 INTRODUCTION**

At both large and small spatial scales, population differentiation can evolve along environmental gradients as long as strength of natural selection exceeds the countervailing effects of gene flow and drift. Adaptive divergence across large geographic scales, such as when large species’ ranges encompass broad latitudinal gradients, is well-documented (e.g., Rehfeldt et al., 1999, Olsson and Ångren, 2002, Halbritter et al., 2015, Rauschendorfer et al., 2021, Lindback et al., 2023). While this large-scale differentiation should not be assumed in the absence of evidence, it is a reasonable expectation for species inhabiting large ranges, especially for plants, where individuals are literally rooted into place. This immobility means that gene flow among populations will decrease with increasing geographical distance (e.g., Twyford et al., 2020), even for species with long-distance dispersal of pollen and seeds, allowing populations to evolve in response to local selective pressures (Savolainen et al., 2007). At small geographical scales, the homogenizing effect of gene flow is expected to be a dominant force, and divergence requires either strong selection, reduced movement of gametes and propagules, or some combination of these factors. The restrictive conditions necessary for population differentiation at a small scale mean that this outcome cannot be assumed in the same way as differentiation at a large scale.

Approaches used to evaluate population differentiation fall into two main categories: those utilizing a quantitative genetic framework and those using population genetic tools, or both. In essence, quantitative genetics starts with phenotypic data from related individuals and infers genetic differentiation for traits (Lynch and Walsh, 1998).Quantitative genetic approaches at the population level require common gardens to examine genetic differentiation and reciprocal transplant studies to confirm the adaptive nature of the differentiation. This makes quantitative genetic studies challenging in terms of space and time, especially for long-lived species like trees. These studies are nonetheless fundamental to identifying genetically based trait differences among populations in a range of environments, assessing local adaptation, quantifying natural selection on traits, and predicting evolutionary rates (Falconer and MacKay, 1983). In contrast, genomics starts with sequence and aims to link differentiation at this level to phenotypic traits or environmental variables (Dalziel et al., 2009). Population genomic approaches permit dissection of evolutionary forces (i.e., drift and selection), characterization of genetic differences at the sequence level, genomic selection, and identification genomic-environment associations in natural populations (de Villemereuil et al., 2016; Sork et al., 2013). Genomic approaches are especially valuable for long-lived species because they provide information that corroborates and complement quantitative genetic data that is more laborious to obtain. For example, in one of the first studies that combined these techniques in *Alnus glutinosa* across Europe*,* multiple outlier loci were associated with leaf sizewhich also varied according to a temperature gradient (De Kort et al., 2014). While these approaches are useful in isolation, pairing quantitative genetic and population genomic techniques with environmental data increases inferential power by exposing associations between the environment and the observed genomic and quantitative genetic differences (De Kort et al., 2014; Lepais and Bacles, 2014).

Both quantitative and population genomic approaches have shown population divergence at small scales, despite the challenges posed by gene flow. Most of this work has utilized short-lived herbaceous species (Sato and Kudoh, 2014; Ren et al., 2017; Frei et al., 2012). The question of whether this also occurs in long-lived tree species, especially wind-pollinated taxa with high rates of gene flow, has not been asked with the same frequency. Indeed, the vast majority of studies in trees investigate genetic differentiation across the species range and generally report positive findings in alignment with expectation (e.g., Rauschendorfer et al., 2021, Filipe et al., 2022; Lindback et al., 2023). When differentiation across smaller scales has been investigated in long-lived trees, the results are mixed. Some studies have shown that populations are differentiated along steep, compact elevational gradients but that the differences are due to environmental rather than genetic effects and attribute this negative finding to the presence of strong gene flow (e.g., King et al., 2013). A few studies in trees report genetic differentiation at a small scale. These studies have primarily relied entirely on either quantitative genetics (Lobo et al., 2017) or population genomics (Gugger et al., 2021; von Takach et al., 2021; Scotti et al., 2023). Studies combining these approaches are rare, with the lone examples coming from two elegant investigations of fine scale adaptation in *Pinus* species from the Lake Tahoe Basin (USA) (Eckert et al., 2015; Lind et al., 2017). Thus, integrative investigations of divergence are still needed to strengthen our understanding of population dynamics at a small scale.

The presence of genetic variation is essential to fuel the evolutionary process. In wind-pollinated trees, where gene flow among populations is high, the variation within a population generally reflects intraspecific allelic diversity, which tends to be substantial in these long-lived organisms. This diversity can provide the basis for populations to respond to natural selection across the landscape, as has been demonstrated many studies (McKown et al., 2014; Howe et al., 2003). Alternatively, variation can originate from interspecific gene flow (hybridization and introgression), which is common across many groups of trees (Kremer et al., 2014; Kremer, 2010; Torres et al., 2019). Neutral introgression between sympatric species will result whenever reproductive isolating barriers are weak, but in some cases the variation introduced in this process lends a selective advantage. This has been demonstrated at both a large geographic scale, in the case of the post-glacial expansion of *Quercus petraea* into northern Europe (Petit et al., 2002; Leroy et al., 2020), and at a microgeographical scale, in the case of soil type characteristics in *Quercus rubra* and *Quercus ellipsoidalis* (Khodwekar and Gailing, 2017). Differences among populations due to introgression can be just as relevant to the evolutionary trajectory of a species as those due to the sorting of intraspecific variation, but most analytical frameworks assume only the latter. Thus, regardless of whether introgression is adaptive or neutral in nature, it is critical to account for this source of variation to avoid spurious conclusions.

Regardless of the scale of the study, an understanding of the degree of population differentiation is inherently interesting and has practical importance in the age of restoration in a changing climate (Benito‐Garzón et al., 2013, Lindback et al., 2023). Foresters and restoration ecologists are increasingly considering the magnitude and direction of climate change when deciding on source material for restoration and reforestation (Twardek et al., 2023; Etterson et al., 2020). Given the rapid rate of climate change, it is possible translocated southern populations may outperform local populations if they are adapted to climate. However, the scale of adaptation is critical to inform the optimal scale of translocation; this is often not known and may differ across the species range (Etterson, 2008). For example, at a broad scale, populations of *Q. rubra* sampled from southern part of the species range and planted into Michigan showed a phenological mismatch with climate that resulted in frost damage and, presumably, negatively impacted plant fitness (Lindback et al., 2023). However, at a smaller scale, central Minnesota populations of *Q. rubra* translocated into northern Minnesota had higher survival, growth, and better phenological matching with the extended growing season (Etterson et al., 2020). The study reported herein examines differentiation at an even smaller scale, from the coast of Lake Superior into the interior of Minnesota. We anticipated that there might be population differentiation along this steep and narrow gradient because forest managers will not purchase seed from Lake Superior Coastal populations based on their experience that the seedlings do not thrive in inland climates (MN DNR, 2024).

Northern red oak *(Quercus rubra* L.) has an expansive range in North America, and one small part of this range occurs across a steep environmental gradient from the cool, moist coastal shores of Lake Superior to the drier, warmer interior forest habitats. Within this narrow gradient (< 50 km), there are four USDA hardiness zones (3a–4b) that are generated by proximity to the deep, cold waters of Lake Superior, the largest freshwater lake in the world by surface area. For comparison, this range of hardiness zones is also found across latitudes from the north shore of Lake Superior to the southern border of Minnesota. The climate impacts of this massive northern lake include warmer falls and winters, reduced frost in the spring, and cooler summers by the shore, all leading to an extended growing season (Moen, 2018). *Quercus rubra* also hybridizes with at least seven other species in the subgenus *Quercus* section *Lobatae* (Moran et al., 2012; Sander, 1990), one of them being the closely related species Northern pin oak (*Quercus ellipsoidalis* E.J. Hill) (Sander, 1990). These two species’ ranges overlap in the Great Lakes region (including Wisconsin, Michigan, and Minnesota in the USA), and several studies have identified hybridization and adaptive introgression from *Q. ellipsoidalis* into *Q. rubra* based on morphometric traits and genetic markers (Gailing et al., 2012; Gailing and Zhang, 2018; Jensen et al., 1993; Lind and Gailing, 2013). However, the potential for introgression between the two species has not yet been studied in areas of sympatry west of Lake Superior*.*

Our goal was to determine whether the steep and narrow environment gradient adjacent to Lake Superior has driven adaptive differentiation using population genomic and quantitative genetic approaches. Our multi-faceted approach included an assessment of neutral population structure to control for the influence of geographic distance coupled with a common-garden assessment of genetic divergence in quantitative traits. We complemented these finding by using multiple methods to detect loci showing distance-based signatures of selection and loci showing associations with environmental factors. Finally, because hybridization between *Q. rubra* and *Q. ellipsoidalis* is well-documented, albeit not in in this study area, we also assessed the level of introgression to ensure that our results were not biased by the inclusion of hybrid zones in regions of sympatry. Here, we address the following questions:1) Is there genetic differentiation among populations as a function of distance from a large climate-moderating Great Lake? 2) Are there loci under differential selection in coastal vs. noncoastal environments? 3) Is there evidence of introgression between *Q. rubra* and *Q. ellipsoidalis* in our study area?

**2 MATERIALS AND METHODS**

**2.1 Study system**

*Quercus* (Fagaceae) is a widespread genus that has radiated across the northern hemisphere since its estimated origin 56 million years ago (Kremer and Hipp, 2019). The genus ranges from the equator to boreal regions, with extremes that include a latitude of 60ºN in Europe and 4,000m above sea level in China (Kremer and Hipp, 2019). Northern red oak (*Quercus rubra*) is a deciduous tree native to North America and widely distributed through eastern United States and southeastern Canada (Dey et al., 2007), where it inhabits a diverse range of hardiness zones (3–8) (USDA, 2019) and soils (Tirmenstein, 1991). As a masting tree, energetically costly large acorn crops occur every 3–5 years, and these fruits are a major source of food for wildlife (Sander, 1990).

**2.2 Leaf tissue and acorn sampling**

We used the Minnesota Department of Natural Resources Relevé data (MNDNR, 2013) to locate 30 *Q. rubra* sampling locations (referred to as “populations,” hereafter) that differed with respect to distance from Lake Superior: “Coastal” (0–16 km), “Inland” (17–80 km), and “Interior” (81–160 km) (Fig. 1a, Supplementary Table 1). For the population genomic study, we collected leaf tissue from 15–18 individuals from 30 populations: ten each from coastal, inland, and interior regions (Supplementary Table 1). To verify the genetic identity of the common garden populations (see below), we also collected leaf tissue from 73 seedlings, distributed across maternal lines, of the eight *Q. rubra* populationsfrom the common garden experiment (Supplementary Table 1). Finally, given the documented hybridization in eastern Great Lakes populations with sister species *Q. ellipsoidalis* (Lind and Gailing et al., 2013), estimated to have diverged from *Q. rubra* between 20-25 million years ago (Rauschendorfer et al., 2022), we were compelled to check for it even though it has not previously been recorded in the study area. To this end, we collected tissue from two *Q. ellipsoidalis* populations from locations near Minneapolis, MN (Fig. 1a, Supplementary Table 1). Only eight of the 30 populations from which we had collected leaf tissue also produced acorns in 2018 (three coastal, four inland, one interior). For these populations, we collected 20–30 acorns from each of 10–15 trees (i.e., maternal lines; Supplementary Table 1). Given the lack of balance with respect to population number between the three distance-from-the-lake categories (Supplementary Table 1) for the common garden experiment, we pooled the inland and interior populations and refer to them collectively as “Noncoastal” populations (17–160 km) and analyzed the quantitative trait data to detect differences between these populations and “Coastal” populations (0–16 km).

**2.3 Common garden experimental design and measurements**

To determine seed viability, we subjected the acorns to a float test (Gribko and Jones, 1995) and stratified them for three months at 4°C before planting. Seed mass and radicle emergence was recorded for each acorn. This is important because differences in the maternal plant’s environment influences offspring traits (e.g., provisioning to the embryo, gene regulation) and often manifests itself as differences in seed weight and other early traits (Roach and Wulff, 1987). Seed weight data can be used as a covariate in the model to account for at least some of the environmental variance in offspring that traces to the parental environment (e.g., Platenkamp and Shaw, 1993; Zas and Sampedro, 2015). In total, 1,438 acorns from 102 maternal lines from eight populations were planted in a randomized block design with four blocks, with populations and maternal lines distributed as evenly as possible across the blocks. Acorns were planted in 2.7 x 10” Deepot cells (D40L – Stuewe & Sons, Inc.) in low water retention soil (PRO-MIX BRK) under a maximum temperature setting of 20°C, a minimum temperature setting of 11.6°C and watered every 48hrs at the University of Minnesota Duluth greenhouse (46°49’00.7” N 92°05’12.1” W).

We recorded seedling emergence as a measure of germination every other day and indicators of juvenile growth weekly (stem height, stem diameter, leaf number and survival). Oak seedlings had a rapid growth period during the first four weeks of the experiment, followed by a period of relative stasis; therefore, we calculated growth rate as the total increase in growth measured by stem height and diameter during the early season (Weeks 1–4) and the late season (Weeks 5–11). Lastly, we collected the uppermost fully expanded leaf at week ten from each available seedling to calculate specific leaf area (SLA cm2/g).

**2.4 Common garden data analysis**

We leveraged two approaches to analyze our quantitative genetic data, using: 1) distance categories from Lake Superior (coastal and noncoastal regions; hereafter referred to as the “regional model”), and 2) climate data from the seedlings’ home environments as predictors (hereafter referred to as the “climate model”). The first analysis tests for differentiation between regions and among populations within each region, whereas the second analysis provides a more detailed examination of the climate factors that might be associated with differentiation. For the regional analyses, we used an ANCOVA for continuous response variables (seed mass, germination phenology, growth, SLA, and week 4 measurements of height, diameter, and leaf number). Seed mass and the presence of an emergent radicle were used as covariates to account for the environmental carryover effects from the maternal environment (Platenkamp and Shaw, 1993; Zas and Sampedro, 2015). Block was considered a random factor, and region and population nested within region were fixed factors. Binomial response variables (presence of emergent radicle, germination, and survival) were analyzed with logistic regression. All statistical models were constructed in JMP Pro 16.0.0 (SAS Institute Inc., 2021).

For the climate analyses, we obtained 19 site-specific bioclimatic variables (http://www.worldclim.org) and performed a forward stepwise model-selection algorithm implemented with ‘adespatial’ R package (Dray et al., 2022). The factors that were retained were: (1) mean annual temperature, (2) precipitation of the driest month and, (3) minimum temperature of the coldest month. These three continuous climate variables replaced distance categories in the model.

**2.5 DNA isolation, genotyping, and SNP calling**

DNA was isolated using a modified CTAB extraction (OpenWetWare, 2018) and quantified using a Qubit 3.0 fluorometer. Genotyping was performed using standard restriction associated DNA sequencing (RAD-seq) methods at the University of Minnesota Genomic Center (UMGC). RAD libraries were prepared using *BamHI* + *NsiI* restriction enzymes. DNA isolation and RAD-seq methods are detailed in the Supplementary Methods.

Variant calling was conducted using the standard ipyRAD v. 9.93 pipeline (Eaton and Overcast, 2020), using the *Quercus rubra* v.2.1 reference genome (Kapoor et al., 2023). We constructed independent ipyRAD assemblies for the following combinations of samples (Supplementary Table 2): (1) field and common garden *Q. rubra* and *Q. ellipsoidalis,* hereafter referred as dataset 1; and (2) field and common garden collected *Q. rubra* (coastal, inland, and interior regions), hereafter referred as dataset 2*.* We removed any SNPs that had not been characterized as part of a chromosomal region in the *Q. rubra* v.2.1 nuclear refence genome as we identified that some of these SNPs were the same as an inserted mis-assembly region of a mitochondrial sequence in the *Q. lobata* nuclear reference genome (Sork et al., 2022). We then performed further SNP filtering for both datasets using VCFtools v.1.16 (Danecek et al., 2011). Details of SNP calling and filtering are available in the Supplementary Methods.

**2.6 Diversity, population structure, and individual assignment**

To characterize genetic differentiation among populations, we calculated pairwise genetic differentiation (*FST*) among populations of both species in dataset 1 using ARLEQUIN v3.5.2.1 (Excoffier and Lischer, 2010). We used STRUCTURE v2.3.4 (Pritchard et al., 2000) to infer population clustering and individual assignment; details of the STRUCTURE runs, evaluating levels of K, and visualization are available in the Supplementary Methods. We evaluated population clustering within *Q. rubra* using dataset 2 and admixture between *Q. rubra* and *Q. ellipsoidalis* using dataset 1. To assess significance of the *Q. ellipsoidalis* assignment proportions within each region for dataset 1, we performed an ANOVA on the *K* = 2 Q-values using region and population nested within regions as predictors in R v.4.2.2 (R core team, 2021). We analyzed genetic groupings for both datasets using Identity-By-Descent measures and constructed principal component analyses (PCA) with the R package ‘SNPrelate v.1.32.2’ (Zheng et al., 2012). We generated estimates of genetic diversity within the wild *Q. rubra* populations in two ways. First, we generated the number of heterozygous sites per individual and averaged this within each population. Second, we calculated p in 100 bp sliding windows for each population, retained only the windows that contained variable sites across all populations, and then averaged across these common windows within each population. Both heterozygosity and p were calculated in VCFtools. We then calculated the pairwise Nei *FST* to estimate genetic distance and pairwise Euclidean distances of the latitude and longitude coordinates of each population to estimate geographic distances using ‘adegenet v.2.1.10’ (Jombart, 2008) and evaluated the linear relationship between the two distances. Lastly, we used the pairwise genetic and geographic distances to perform a mantel test with a Pearson correlation method in ‘vegan v.2.6-4’ (Oksanen et al., 2017) to investigate patterns of Isolation-by-Distance (IBD). All packages were implemented in in R v.4.3.2.

**2.7 Genomic signatures of selection**

To assess signatures of selection, we constructed a dataset of unlinked SNPs using only field collected *Q. rubra* from dataset 2. To reduce the physical linkage between SNPs that could bias the selection results, we calculated pairwise linkage disequilibrium between the remaining SNPs using plink v1.9 (Purcell et al., 2007), implementing windows of 50 markers with ten marker sliding windows, and removed loci with an *r2* above 0.25.

We performed a distance-based outlier loci analysis on the resulting SNP set by implementing a per-locus pairwise *FST* approach using two methods: OutFLANK v.0.2 (Whitlock and Lotterhos, 2015) and BayeScan 2.1 (Foll and Gaggiotti, 2008). OutFLANK was implemented in R v.4.2.2, allowing a minimum heterozygosity of 4%, and trimming 35% of the left-hand side of the *FST* distribution. Loci identified in OutFLANK were considered significant outliers if their q-value < 0.05. BayeScan 2.1 was conducted with prior odds of 100, a burn-in of 50,000, a thinning interval of 10, a sample size of 5,000, and 20 pilot runs with 5,000 iterations each. We identified outliers from BayeScan using the R script provided by the program authors with a 0.05 FDR correction threshold (Benjamini and Hochberg, 1995). We searched for functional annotation of outlier loci using the Jbrowse v.1.16.9 tool from Phytozome v.13 (Goodstein et al., 2012) and using the *Q. rubra* genome as a reference and calculated allele frequencies for all outlier loci in each region using VCFtools.

Lastly, we performed a genotype-environment association analysis (GEA) to identify loci associated with environmental variables in the three *Q. rubra* regions. We obtained climate data for each population from the 19 bioclimatic variables in the Worldclim dataset (http://www.worldclim.org). For all GEA analyses, we performed a correlation matrix and a variance inflation factor analysis (VIF). We removed variables with an *R2* > 0.7 and a VIF > 10. To ensure that the three variables used in the original common garden analysis were still significant for our GEA, we included them in the calculation for the correlation matrix and VIF analysis. The bioclimatic variables retained for each GEA analysis can be found in Figure 6b.

We performed the GEA using three methods: BayeScEnv v.1.1 (de Villemereuil and Gaggiotti, 2015), a Latent Factor Mixed Model (LFMM2) analysis as implemented in ‘LEA v.3.10.2’ (Frichot and François, 2015), and a genomic redundancy analysis (RDA) as implemented in ‘vegan v.2.6-4’. BayeScEnv uses an *FST*-based approach that incorporates standardized environmental variables. This method considers two locus-specific effects, the first one being divergent selection and the second one being other processes different from local adaptation (de Villemereuil and Gaggiotti, 2015). The LFMM2 approach implemented in R v.4.2.2 identifies loci associated with environmental factors while including the effect of population structure as a random factor (latent factor) (Frichot et al., 2013). The genomic RDA is a supervised multivariate ordination analysis that determines how groups of loci covary in response to multiple environmental factors (Rellstab et al., 2015).

For BayeScEnv, we standardized each population environmental variable by subtracting the mean and dividing it by the standard deviation. We set up the model using the standard parameters and performed a 0.05 FDR correction. For the LFMM2, we performed a PCA on the environmental factors to reduce the number of variables added to the model. We retained the first three PC axes as synthetic predictors as they explained 93% of the variation. Consequently, we performed three independent runs for *K* = 2, 4, 6 for each of the principal components with a burn-in period of 5,000 and a total of 10,000 Monte Carlo Markov Chains. The association results for each value of *K* were very similar; thus, we chose *K* = 2 to avoid overparameterizing the model. We adjusted the *P*-values for the correlations between each locus and PC1 using a 0.05 FDR correction. Lastly, we performed a genomic RDA using the ‘vegan v.2.6-4’’ package. We imputed the genotype missing data using the most common genotype at each SNP across individuals and performed the RDA between the genotype and the environmental factors using the same environmental variable reduction process as in LFMM2. Lastly, loci that loaded more than three standard deviations from the mean the first three loading distributions were identified as outliers (Kamvar et al., 2016).

**3 RESULTS**

**3.1 Regional Model – Quantitative genetic differentiation between coastal and noncoastal populations**

Overall, the regional model detected fewer differences between coastal and noncoastal regions relative to population differences within regions (36% and 91% of traits, respectively) (Table 1a). The two covariates in the analyses, seed mass and radicle emergence at the time of planting, significantly influenced quantitative genetic traits throughout the measurement period (82% and 55% of traits, respectively), as did the spatial blocks in the common garden experiment (100% of traits).

**Seed traits:** Seed mass did not differ between coastal and noncoastal regions, but did differ significantly between populations within region (Table 1a, Supplementary Fig. 1a).Heavier acorns were more likely to have an emergent radicle at the time of planting. Radicle presence differed significantly between both regions and among populations within regions. Coastal populations were 37% more likely to have emergent radicles compared to noncoastal populations (Fig. 2a). Like seed mass, germination percentage and germination phenology did not differ between regions but both traits differed significantly among populations within regions. Populations within the coastal region germinated within five days of each other, whereas populations in the noncoastal region germinated more synchronously within 1.5 days of each other (Supplementary Fig. 1b-c). Germination rate and germination phenology were positively associated with presence of emergent radicles and seed mass. Germination phenology was not influenced by other factors in the model.

**Growth Traits:** Oak seedlings had a rapid growth period during the first four weeks of the experiment, followed by a period of relative stasis in above ground growth. The change in stem height during Weeks 1-4, “Early Growth Rate”, was significantly different between regions and between populations within each region (Table 1a). Coastal populations grew 10% (0.2 cm/week) faster than noncoastal populations on average (Fig. 2b). Seedlings that grew the fastest had a high seed mass but were less likely to have an emergent radicle at the time of planting. The later stage of seedling growth during Weeks 5-11, “Late Growth Rate”, was only significantly influenced by seed mass with larger seeds growing faster.

At Week 4, noncoastal seedlings were significantly taller (12%) than coastal seedlings, although populations also differed significantly within regions (Table 1a). At the end of the common garden experiment, noncoastal seedlings remained 6% taller than coastal seedlings although this difference was not significant (not shown). Seedling height was positively associated with seed mass. There were no significant regional differences for stem diameter, leaf number, or SLA (Supplementary Fig. 1d-f), although populations within regions differed significantly for each of these traits. Acorns that germinated earlier and had an emergent radicle at the time of planting produced seedlings with wider stem diameters and had lower SLA (i.e., thicker leaves).Leaf number was also positively associated with seed mass.

**Survival:** Noncoastal populations had small but significant survival advantage (6%) compared to coastal populations (Fig. 2d). Survival was positively associated with radicle presence prior to planting.

**3.2 Climate Model – Geographic differentiation with respect to climate variables**

Overall, the climate model detected significant relationships between climate variables and all measured traits except for late growth rate, including annual mean temperature (82% of traits), minimum temperature of the coldest month (55%), and precipitation in the driest month (55%) (Table 1b).Like the regional model, the covariates, seed mass and radicle presence significantly influenced most quantitative genetic traits (73% and 64% of traits, respectively), as did the spatial blocks in the common garden (100% of traits).

**Seed traits:** Acorns sampled from populations with higher mean annual temperature tended to have a lower seed mass (Supplementary Fig. 2a), were less likely to have an emergent radicle (Supplementary Fig. 2b), were more likely to germinate (Fig. 3a) but at a later time (Supplementary Fig. 2c) compared to acorns sampled from populations with lower mean annual temperature (Table 1b).Extreme cold, as indicated by the coldest minimum winter temperature, resulted in the production of acorns with lower seed mass (Supplementary Fig. 2a) that were less likely to germinate (Fig. 3a). Acorns from populations with lower precipitation in the driest month were less likely to have an emergent radicle at the time of planting (Supplementary Fig. 2b) but were more likely to germinate (Fig. 3a). Germination phenology was not significantly affected by the annual temperature or minimum temperature of the coldest month. As with the regional model, heavier acorns were more likely to have an emergent radicle and were more likely to germinate (50%) and earlier (7 days).

**Growth Traits:** Seedlings from populations with higher mean annual temperature had faster early growth rates (Supplementary Fig. 2d) and, at week 4, were taller (Fig. 3b) but had fewer leaves (Supplementary Fig. 2f) with lower SLA (thicker, Fig. 3c) compared to seedlings from populations with lower mean annual temperature (Table 1b). Seedlings from populations that are typically exposed to extreme cold, as measured by the minimum temperature in the coldest month, were shorter (Fig. 3b), had more narrow stem diameters (Supplementary Fig. 2e), and produced leaves with higher SLA (thinner, Fig. 3c).Populations that receive more precipitation in the driest month of the year produced seedlings that had faster early growth rates (Supplementary Fig. 2d), were taller (Fig. 3b), and had lower SLA (Fig. 3c) relative to populations that had low precipitation in the driest month. Late growth rate was not significantly influenced by any of the climatic factors in the model but was positively associated with seed mass. Heavier acorns also produced taller seedlings with wider stem diameters, and more leaves with lower SLA. Acorns with an emergent radicle produced seedlings with 14.8% slower early growth rates, 7.5% smaller stem diameters, and leaves with 1% lower SLA compared to seedlings from acorns with no emergent radicle.

**Survival:** Seedling mortality was significantly affected by all factors in the climate models except for acorn mass (Table 1b, Fig 3d). In general, seedlings from sites with higher mean annual temperature had higher survival while seedlings from sites that experience the lowest minimum temperature in the coldest month had lower survival. Seedlings from sites with the highest precipitation in the driest month were also more likely to survive. Radicle presence at the time of planting was associated with higher survival.

**3.3 Diversity, population structure, and individual assignment**

For our two independent ipyRAD assemblies, we initially identified 259,016 SNPs across 494 individuals for dataset 1 and 254,042 SNPs across 468 individuals for dataset 2. After filtering, there were a total of 15,834 SNPs across 453 individuals for dataset 1, and a total of 16,112 SNPs across 429 individuals for dataset 2 (Supplementary Table 2). Interspecific *FST* was 0.084 and pairwise *FST* among all *Q. rubra* and *Q. ellipsoidalis* populations ranged between 0.0101 and 0.2622. Most intraspecific *Q. rubra* pairwise *FST* values were much lower than values between *Q. rubra* and *Q. ellipsoidalis* populations. However, four field populations (SP, BL, II, GW; Supplementary Table 1, Supplementary Fig. 3) that had been identified as *Q. rubra* by the MN DNR showed elevated pairwise *FST* values ranging between 0.1157 and 0.2622 with other *Q. rubra* populations and low interspecific pairwise *FST* values ranging between 0.0151 and 0.0381 with *Q. ellipsoidalis* populations (Fig. 1b). Principal component analysis also showed that these four *Q. rubra* populations clustered more closely with *Q. ellipsoidalis* populations than with other *Q. rubra* populations (Supplementary Fig. 4). These populations were removed from both datasets for all subsequent analyses. This resulted in 404 individuals in dataset 1 and 380 individuals in dataset 2 (Supplementary Table 2). After removing the misidentified populations from dataset 2, the global *FST* for *Q. rubra* alonewas 0.028 and the pairwise *FST* for *Q. rubra* populations ranged between 0.0125 and 0.0761.

Given the finding of misidentified *Q. rubra* populations, it was incumbent upon us to examine the overall extent of introgression between *Q. rubra* and *Q. ellipsoidalis* in our samples. To this end, admixture between field-collected *Q. rubra* and *Q. ellipsoidalis* individuals in dataset 1 (excluding the misidentified populations) was evaluated based on STRUCTURE assignments at *K =* 2. For *Q. rubra* populations, the proportion assignment to *Q. ellipsoidalis* increased as distance from Lake Superior increased (Fig. 4a-b; Supplementary Table 3; Supplementary Fig. 5). Interior populations had the highest average assignment to *Q. ellipsoidalis* (9.8%), and three interior individuals had an assignment proportion above 45%. The inland populations had lower average assignment proportions (5.2%) than interior populations, and two inland individuals had more than 45% assignment to *Q. ellipsoidalis*. Coastal populations had the lowest average assignment *Q. ellipsoidalis* (1.7%). The two *Q. ellipsoidalis* populations showed less than 3.5% average assignment to *Q. rubra* (WO = 3.3%, HA = 1.5%)*.*

Population structure within *Q. rubra* was performed using the field and common garden *Q. rubra* populations in dataset 2 (excluding the misidentified populations). The optimal value of K for the STRUCTURE analysis was K = 5, but each of the five groups were present in each population and there was no evidence of differentiation between regions; analyses performed at K = 2 through 4 show the gradual partitioning of each individual into these five groups (Supplementary Fig. 6). The most prominent pattern was that the five individuals that had >45% assignment to *Q. ellipsoidalis* in the multi-species STRUCTURE analysis were assigned to a single group at all levels of K. This suggests that some of the slight structure identified within *Q. rubra* is likely to be an artifact of admixture with *Q. ellipsoidalis*. At K = 5, individuals from the AA (coastal) and MG (inland) populations were assigned primarily (although not entirely) to two different clusters. Neither of these populations showed high degrees of admixture with *Q. ellipsoidalis*. At K = 2 there was no difference between the common garden and field-collected individuals (collection site(region): df = 3, 152, F = 1.330, *p* = 0.27), indicating that the common garden samples were good representatives of populations in the field. The PCA of field-collected populations explained a small amount of the variance between regions (2.56% across the first two PCs) and clustered all regions into one group, though some individuals from inland (including the MG population) and interior regions diverged from the main cluster along PC1 and PC2 (Fig. 4c). Overall, there was no evidence of strong differentiation between *Q. rubra* regions or populations, which is consistent with the low pairwise *FST* values (Fig. 1b).

Values of p ranged from 0.0045 to 0.0058. These values are lower than those estimated based on whole-genome sequencing (0.0077-0.0079; Kapoor et al., 2023) because the RAD-seq approach used here does not capture all variable sites, leading to a systematic under-estimation of variation when calculated across the genome. Nonetheless, the values are internally consistent and useful for comparison among populations within this study. The average number of heterozygous sites per population ranged from 817 to 1013, with the lowest population showing ~80% of the heterozygosity of the most heterozygous population. Heterozygosity and p were lowest for coastal populations and highest for interior populations, and both of these values were positively correlated with distance from the lake (Supplementary Fig. 5). This increase in diversity with distance from the lake mirrors the increase in average assignment to *Q. ellipsoidalis* (Fig. 4a-b; Supplementary Fig. 5). A Mantel test for all populations in dataset 2 showed no evidence of IBD among populations (*r2*=0.155, *p* = 0.133; Supplemental Fig. 7).

**3.4 Genomic signatures of selection**

To identify outliers representative of selection in our wild populations and to ensure that only characterized chromosomal regions from the *Q. rubra* reference genome were included, we removed all common garden individuals from dataset 2 and only retained loci from the characterized chromosomal regions, retaining 312 field-collected *Q. rubra* individuals and 15,659 SNPs across 3,330 RAD loci. We identified moderate levels of linkage between SNPs on different RAD loci; of the initial 15,659 SNPs, 30.3% exhibited high levels of linkage (*r2* > 0.25). SNPs with high linkage levels were removed yielding a final set of 4,741 SNPs in linkage equilibrium across 2,577 RAD loci for use in the selection analyses (Supplementary Table 2).

The distance-based outlier analysis using BayeScan (Fig. 5a) and OutFLANK (Fig. 5b) identified 64 and six loci, respectively, with elevated levels of divergence at a 5% significance level. Four of the six outlier loci identified by OutFLANK were also identified by BayeScan (Chr4:33901924, Chr5:53324407, Chr8:14061577, and Chr9:33364097; all positions are based on the *Q. rubra* reference genome) and exhibited *FST* levels ranging between 0.36 and 0.47 based on the OutFLANK analysis (Supplementary Table 4). We calculated allele frequencies for the four outlier loci in each *Q. rubra* geographical region; outlier loci in chromosomes 4,5, and 8 were polymorphic within each region, although Chr4:33901924 was nearly fixed in the inland region and Chr8:14061577 was nearly fixed the interior region. Chr4:33901924 and Chr8:14061577 had the lowest and higher *FST* values (036 and 0.47) respectively based on the OutFLANK analysis (Supplementary Table 4). The outlier locus in chromosome 9 was fixed for the coastal region but remained polymorphic in the inland and interior region. Chr9:33364097 had the second highest (0.44) and highest (0.34) *FST* values based on the OutFLANK and BayeScananalysesrespectively.

The three GEA methods identified variable numbers of loci with significant associations to environmental factors. BayeScEnv identified 14 loci across nine different chromosomes (2-4, 6, 8-12) with a significant association to at least one environmental variable (Fig. 6a). No loci were associated with mean annual temperature (Fig. 6b). Only locus Chr9:33364097, was associated with more than one environmental variable (minimum temperature of the coldest month, precipitation of the driest month and precipitation of the warmest quarter) and had an average *FST* value (0.33) across the three associated variables in the BayeScEnv analysis. The LFMM2 analysis revealed a total of four unique loci associated with PC1 and PC3 environmental variables at 5% significance (Fig. 6a). PCA loadings showed that PC1 of the LFMM2 was driven by the minimum temperature of the coldest month, while PC3 was driven by isothermality and precipitation of the driest month (data not shown). Neither mean annual temperature and precipitation of the warmest quarter were significant in the LFMM2 analysis (Fig. 6b). Lastly, the genomic RDA detected 42 candidate loci (Fig. 6a) that were associated with the five environmental variables (Fig. 6b). RDA loadings showed that RDA1 was driven by the precipitation of the driest month, precipitation of the warmest quarter, and minimum temperature of the coldest month, RDA2 by mean annual temperature and minimum temperature of the coldest month, and RDA3 by isothermality (Supplementary Fig. 8). Only one locus, Chr9:33364097, was significant across all three GEA methods and the two distance-based outlier analyses (Fig. 6a). This locus was associated with minimum temperature of the coldest month, precipitation of the driest month and precipitation of the warmest quarter for BayeScenv and the genomic RDA, and minimum temperature of the coldest month for LFMM2.

**4 DISCUSSION**

Understanding the drivers and consequences of divergence at a small spatial scale has implications for our understanding of basic questions in evolutionary biology and for conservation actions. There is no question that strong selection can drive divergence among populations of short-lived, herbaceous species that are physically proximate, and there is growing evidence indicating that the same patterns can be found in some tree species. The communities’ exploration of this question has been greatly facilitated by population genomic approaches that allow an assessment of wild populations that enriches the information obtained from labor intensive common garden studies. While this has encouraged the investigation of population differentiation at a small scale (Gugger et al., 2021; von Takach et al., 2021; Scotti et al., 2023), there are few studies that investigated this question using a quantitative genetic framework (Lobo et al., 2017), and fewer still that have combined these analytical techniques in a single study (Eckert et al., 2015; Lind et al., 2017). Here, we present the first study combining quantitative genetics and population genomics to investigate fine-scale population divergence in *Quercus* and show that the narrow but steep environmental gradient produced by the cold water of Lake Superior is sufficiently strong to promote diversification in traits and loci, despite the lack of neutral population structure in this region. In addition, we documented hybridization between *Quercus rubra* and its congener *Quercus ellipsoidalis* for the first time west of Lake Superior.

**4.1 Is there genetic differentiation among populations as a function of distance from a large climate-moderating Great Lake?**

We did not detect strong genomic differentiation among *Q. rubra* populations based on *FST* values or clustering analyses nor did we detect IBD across this small geographic area (Fig. 1, 4c; Supplementary Fig. 7). Although there was a slight trend towards differential assignment to the arbitrary groups identified in the STRUCTUREanalysis within *Q. rubra*, a close examination suggests that the apparent structure at K = 2 is likely to be an artifact of increased levels of admixture with *Q. ellipsoidalis* at greater distances from the lake (Fig. 4a; Supplementary Fig. 6). The overall lack of population structure within *Q. rubra* based on genomic variation is consistent with high rates of gene flow, likely due to the long generation time and the extended pollen dispersal range (recorded up to 100 km) that characterize oaks (Schueler and Schlünzen, 2006). Both factors contribute to the dispersal and maintenance of allelic diversity across populations. Indeed, our global *FST* value (0.028) is similar to those found in *Quercus* and other tree species (e.g., *Q. oleoides*, *FST* = 0.09, Ramírez-Valiente et al., 2018; *Quercus robur* L., *FST* = 0.07, Vakkari et al., 2006; *Eucalyptus albens* Benth., *FST* = 0.018, Murray et al., 2019; *Pinus taeda* L., *FST* = 0.04, Bassar et al., 2010).

Interestingly, despite this lack of structure and the inference of high gene flow, the within-population diversity was not equivalent across the region. Instead, diversity increased with distance from the lake. As noted above, distance from the lake is also positively correlated with increased admixture between *Q. rubra* and *Q. ellipsoidalis*. Thus, there are two equally plausible explanations for this pattern of diversity on the landscape. One explanation is that the higher diversity of interior populations is due to introgression from *Q. ellipsoidalis* introducing new alleles into *Q. rubra* populations. Alternatively, the low diversity of the coastal populations may be due to the fact that they are essentially at a range edge, where gene flow is geographically constrained (i.e., propagules and pollen are unlikely to flow inwards from the lake). Differentiating between these two possibilities is an intriguing area of future study and will require comparative sampling from other locations in *Q. rubra*’s range.

Despite the overall lack of genomic differentiation, we found strong evidence of quantitative trait differentiation across the small spatial scale of the study (Table 1a; Fig. 2). In the climate model, all traits except one (late season growth) exhibited significant clinal variation associated with climate (Table 1b; Fig. 3, Supplementary Fig. 2). Four traits, including the closest fitness proxy (survival), showed significant covariation with all three bioclimatic variables (mean annual temperature, precipitation of the driest month and, minimum temperature of the coldest month). This suggests that climate has molded phenotypes along environmental gradients either through genetic differentiation or environmental carryover effects from the maternal environment. However, we did attempt to account for differences in seed provisioning that traces to the quality of the parental environment (i.e., maternal effects) by including traits typically associated with maternal provisioning of the seeds in the model (i.e., seed mass and radicle presence after stratification, Wulff, 1995). These traits are appropriate proxies for maternal effects because environmental carryover is most pronounced at the seed stage (Platenkamp and Shaw, 1993; Zas and Sampedro, 2015) and has diminishing impacts on later life history stages (Roach and Wulff, 1987). In our study, we found that seed mass accounted for a significant amount of variation for all traits except for survival. In addition, early radicle presence was significant in two-thirds of our analyses. This suggests that these early life-history traits, that are known to be sensitive to the maternal environment, valuably accounted for variance transmitted across generations (Table 1a). Although maternal effects can also have genetic component (Galloway et al., 2009; Platenkamp and Shaw, 1993; Thiede, 1998) and affect fitness in nature (Galloway and Etterson, 2007), the environmental component of maternal effects is typically larger than the genetic component and is especially impactful early in establishment (Roach and Wulff, 1987) which in *Q. rubra* can take up to five years. Recognition of the need to account for maternal environmental carryover effects in common garden experiments is also growing in the forestry literature (Vivas et al., 2020). Despite this, it is not commonly considered in common garden studies in trees (Baquedano et al., 2008; Gao et al., 2021; Ramírez-Valiente et al., 2017). In fact, in some cases, seed mass itself is considered a trait that has undergone adaptive differentiation even though it is likely to be confounded with maternal environmental effects (Ramírez-Valiente et al., 2009). The extent to which seed mass accurately accounts for environmental maternal effects or is a genetically-based trait that has itself undergone adaptive differentiation can only be disentangled with experiments that are difficult to conduct on long-lived species, such as trees (Dickerson, 1947).

Our hypothesis of genetic differentiation as a function of proximity to one of the climate-moderating Great Lakes was supported by congruent evidence about functional trait adaptation. Specifically, we observed parallel phenotypic (Table 1a; Fig 2; Supplementary Fig. 1) and environmental clines (Table 1b; Fig 3; Supplementary Fig. 2) that are illustrated by three examples. First, the radicle, a structure that has been shown to be drought-sensitive in other species (Boureima et al., 2011), was less likely to have emerged during stratification among acorns sampled from warmer and drier inland climates. Earlier germination as indicated by radical emergence may be more favorable in moist coastal environments where spring temperatures are moderated, and risk of freezing is low. Second, plants from warmer environments grew faster and were taller compared to those from sites exposed to extreme low winter temperatures (Table 1, Fig. 2, 3), a trend that has also been widely observed across other temperature gradients (e.g., Akalusi and Bourque, 2018; Mimura and Aitken, 2010; Rehfeldt et al., 2002, 2014).Third, plants from warmer and drier climates also produced significantly fewer and thicker leaves (lower SLA) (Table 1, Fig. 2, 3), patterns that are commonly associated with drought tolerance (Dwyer et al., 2014). Collectively, our results suggest that *Q. rubra* populations have undergone genetic differentiation in a narrow geographical area with steep environmental gradients, as evidenced by both the significant categorical comparisons of coastal and noncoastal populations in the regional model, and the strong covariation between trait values and environmental variables in the climate model.

**4.2 Are there loci under differential selection in coastal vs. noncoastal environments?**

Despite the low population differentiation that is consistent with high rates of gene flow, we found patterns of genomic variation: 1) consistent with selection based on outlier analyses and 2) significantly associated with environmental variables. These findings suggest a scenario of population differentiation due to strong natural selection acting against the homogenizing effect of gene flow, a well-established pattern in oaks (e.g., Kremer and Hipp, 2019, Ramírez-Valiente et al., 2022), and the genomic results are largely consistent with population divergence identified in the quantitative genetic analysis. However, the key environmental factors identified by these two approaches, while overlapping, are not identical (Fig. 6b). Precipitation of the driest month and minimum temperature of the coldest moth were consistently significant across all analyses, but two out of five factors were only significant in the genomic analyses. Meanwhile, mean annual temperature was significant for the quantitative genetic analysis but only significant in one out of three GEA analyses. While some of these differences are no doubt due to the assumptions underlying the quantitative genetic and genomic approaches employed, there are two other critical factors that likely drive the different results from these complementary approaches. First, the datasets include different sets of populations and individuals. The quantitative genetic analysis included only eight *Q. rubra* populations due to space constraints, so it is possible that the genomic analysis of 26 populations captured more variation and led to the identification of additional significant bioclimatic associations. Second, the life stage at which the environmental associations were being identified differed. The quantitative genetic analysis focused on juvenile traits by necessity, which describes only one life history stage in this long-lived perennial species. In contrast, genotype-environmental associations were based on adult trees in natural populations, potentially revealing the impact of selection acting at any point across the species’ lifespan. Thus, the incomplete overlap between the significant bioclimatic associations, and the identification of more significant variables associated in the genomic than the quantitative genetic analysis, is not surprising.

In this study, we identified four loci that showed signatures consistent with selection across both distance-based outlier approaches (Fig. 5, Supplementary Table 4). One of these loci was also significant across all three environmental association analyses (Fig. 6a). This locus (Chromosome 9, position 33364097) was associated with minimum temperature of the coldest month and had an *FST* that was much higher than the global *FST* across *Q. rubra* populations (BayeScan *FST* for Chr9:33364097 = 0.43, OutFLANK *FST* for Chr9:33364097 = 0.45 Supplementary Table 4; global *FST* = 0.028). It is interesting to note that minimum winter temperature is a key parameter defining the four USDA hardiness zones in this region (Fig. 1a; Daly et al., 2012). This locus was polymorphic in the inland and interior populations but fixed for a single allele in the coastal populations (Supplementary Table 4), a pattern that is consistent with a scenario of a beneficial allele being driven towards fixation in the coastal populations. The association of Chr9:33364097 with minimum temperature of the coldest month in the environmental association analyses is consistent with the fact that the moderating effect of Lake Superior results in relatively warmer winter conditions for the coastal populations, while populations further inland experience intense cold in the winter months. The outlier locus Chr9:33364097 does not have a functional annotation in the *Q. rubra* v. 2.1 reference genome but is located 7,986 bp upstream of the NADH-ubiquinone oxidoreductase chain II gene in the *Q. rubra* genome (Qurub.09G162300). The NADH-ubiquinone oxidoreductase chain II gene is responsible for the production of ubiquinone reductase which functions as an electron transporter in aerobic respiration (Liu and Lu, 2016). The locus is also 2,464 bp downstream of a homolog of the AT1G43760 gene in *Arabidopsis thaliana* that encodes a DNAse I-like superfamily protein. Neither of these annotations is obviously related to tolerance of winter temperatures, so it is possible that the SNP is located in or linked to a regulatory region or an as-yet unannotated RNA- or protein-coding gene.

Our identification of loci under selection in *Q. rubra* is part of a growing body of work in the genus, where genes showing the signature of selection have been identified based on outlier or distance-based methods in wild populations of *Q. rubra, Q. ellipsoidalis*, and *Q. lobata* (Lind-Riehl et al., 2014; Pettenkofer et al., 2020). More recently, the abundance of genome-scale data has facilitated the use of environmental association analyses to more directly link selection to climatic variables. To our knowledge, only three studies have used environmental association analyses in *Quercus*.Two of these studies were performed in *Q. lobata* and identified 34 loci using SNPs and 43 sites using single methylation variants (SMVs) (Gugger et al., 2016, 2021). More recently, the publication of the *Q. rubra* genome was accompanied by a GEA based on a genomic RDA similar to our study, although conducted at a much larger geographic scale (Kapoor et al., 2023). The results of that study show striking concordance with our investigation, in that minimum temperature of the coldest month was identified as a critical environmental factor, and associated loci associated were on chromosome 9, although the location was approximately 20 Mbp from the SNP identified here.

Environmental association analyses using SNP data in other plant species have found between 54 and 1,011 loci associated with the environment (e.g., Christmas et al., 2016; Pais et al., 2017; Rellstab et al., 2016). The small number of significant loci in this study may seem anomalous by comparison, but it is a logical outcome of the fact that our study focuses only on the overlapping loci between two distance-based and three GEA approaches and that previous studies are conducted on broader geographical scales than our study, ranging from 200–800 km. In comparison, the populations in this study span a smaller geographical scale (160 km), where only very strong selection will overcome the pervasive effects of gene flow, leading to fewer loci showing the signature of selection.

**4.3 Is there evidence of introgression between *Q. rubra* and *Q. ellipsoidalis*?**

Interspecific hybridization among oak lineages is the rule, rather than the exception (Kremer and Hipp, 2019; Rauschendorfer et al., 2022), and there is substantial evidence of hybridization between *Q. rubra* and *Q. ellipsoidalis* in eastern areas of the Great Lakes region (Gailing et al., 2012; Jensen et al., 1993; Lind and Gailing, 2013). Given that these two species are also sympatric in our study area (Supplementary Fig. 3), it was important to include *Q. ellipsoidalis* populations as references to confirm *Q. rubra* genomic identity and exclude populations that might otherwise bias our results. This also allowed a preliminary investigation of the degree of introgression between these two species west of Lake Superior. As a result, we eliminated multiple populations based on their genomic identity and also documented admixture between these species in a pattern across the landscape that opens doors for future studies of gene flow and potentially adaptive introgression*.*

The inclusion of the *Q. ellipsoidalis* samples revealed that four “*Q. rubra*”populations in our initial collections were genetically more similar to *Q. ellipsoidalis* than *Q. rubra* (SP, BL, GW, II; Fig. 1b, Supplementary Fig. 4). These results are surprising since the MN-DNR vegetation surveys that guided our population search did not indicate the presence of *Q. ellipsoidalis* in these sites. It is possible that the trees were misidentified in the original surveys, given the close morphological resemblance between species (Gailing et al., 2012, Jensen et al., 1993), or that our samples were from a stand of *Q. ellipsoidalis* close to, but not included in, the original MN-DNR survey site. Regardless of the nature of these populations, these results confirm the importance of screening *Quercus* populations based on genetic identity before testing for selection; inclusion of these populations would certainly impact conclusions from the distance-based and GEA methods employed herein. The assessment of admixture across the remaining populations showed widespread, but moderate, levels of introgression from *Q. ellipsoidalis* into *Q. rubra*. Our results differed from previous surveys of introgression that have documented up to 24% introgressive forms in *Q. rubra* populations (individuals assigned >10% to *Q. ellipsoidalis*; Lind-Riehl and Gailing, 2017). This frequency was only approached in the interior populations, where 21% of individuals showed >10% assignment to *Q. ellipsoidalis* (Supplementary Table 3). This difference is not surprising, given that we specifically excluded populations that showed signs of being hybrid zones and retained only “pure” populations, rather than focusing on sympatric or parapatric populations of the two species. The moderate levels of admixture detected here are consistent with occasional, historic hybridization events followed by either neutral or selective elimination of *Q. ellipsoidalis* genomic content.

While the proportion of the *Q. rubra* genome assigned to *Q. ellipsoidalis* was generally low, it increased significantly with distance from the lake (Fig. 4b, Supplementary Fig. 5). This result is intriguing because gene flow from the drought-tolerant *Q. ellipsoidalis* (Abrams, 1990; Hipp, 2010) into the mesic-adapted *Q. rubra* would potentially be advantageous in inland/interior regions, where annual precipitation is low and summer temperatures are high. Our study was not specifically designed to identify adaptive introgression, so we lack the power to thoroughly explore this scenario, but we did collect genotypic data for 66 seedlings from the quantitative genetic study to ensure that they were representative of field-collected samples. When the proportion of *Q. ellipsoidalis* assignment for these individuals was added to the climate model, it explained a significant amount of additional variance for two traits both of which were related to germination. Individuals with a higher assignment to *Q. ellipsoidalis* were less likely to have an emergent radical at the time of planting ( 2 = 4.20; *P* = 0.04) and were slower to germinate (F1, 57 = 5.46; *P* = 0.02). As discussed earlier, delayed radical emergence may be advantageous in drought-prone regions away from Lake Superior. This pattern is consistent with previous studies that have demonstrated the adaptive potential of introgression from *Q. ellipsoidalis* into *Q. rubra* in sympatric populations in the eastern Great Lakes region based on patterns of variation at a CONSTANS-like gene associated with drought and soil quality (Khodwekar and Gailing, 2017; Lind-Riehl and Gailing, 2016; Lind-Riehl et al., 2014). More broadly, it echoes work in other *Quercus* species where introgression has been shown to be environmentally dependent (Khodwekar and Gailing, 2017; Nagamitsu et al., 2020) and in the well-studied *Q. petraea/Q. robur* species pair, where the northern expansion of *Q. petraea* is attributed partially to the adaptive introgression of *Q. robur* alleles and partially to climate change (Truffaut et al., 2017; Kremer and Hipp, 2019; Leroy et al., 2020).

We note that while potential for adaptive introgression in the region is intriguing, an equally plausible explanation for the different levels of admixture across regions is a neutral process attributable to the increasing overlap between the species in inland and interior areas (Supplementary Fig. 3). We also acknowledge that our study included only two populations of *Q. ellipsoidalis*, so much deeper sampling of *Q. ellipsoidalis* is necessary to understand which phenomenon is driving the introgression. In either case, these discoveries have created a new testing ground for a more in-depth investigation of oak hybridization and the potential adaptive significance of gene flow across the landscape in an area that is experiencing some of the most extreme climate change in the continental USA, including an increased frequency of mid-summer drought.

**4.4 Conclusion**

The steep environmental gradient generated along the shore of Lake Superior is well documented and easily observed, but it represents only a tiny fraction of the considerable range of *Q. rubra*. Our investigation of whether populations of a this long-lived, wind-pollinated species could respond to selective pressures across such a small scale could have revealed population homogeneity in the region. Instead, we found convincing evidence of population divergence, identified loci under differential selection, and demonstrated that these results are tied to the key environmental variables that define the tight climate zone contours adjacent to the lake and experienced by populations in the wild. This enhances our basic understanding of the evolutionary dynamics of forest trees and reveals the need to consider fine-scale environmental variation. This information is especially valuable for state and non-profit agencies who source seeds from wild populations for reforestation efforts, indicating that unique source populations of *Q. rubra* should be recognized, even across a small geographic scale. Finally, we have laid the groundwork for further studies of potential adaptive introgression from *Q. ellipsoidalis* into *Q. rubra* in the region, a phenomenon that may well accelerate as the climate change drives more frequent droughts.

**REFERENCES**

Abrams, M. D. (1990). Adaptations and responses to drought in *Quercus* species of North America. *Tree Physiology*, *7*, 227–238. https://doi.org/10.1093/treephys/7.1-2-3-4.227

Akalusi, M. E., And Bourque, C. P.-A. (2018). Effect of climatic variation on the morphological characteristics of 37-year-old balsam fir provenances planted in a common garden in New Brunswick, Canada. *Ecology and Evolution*, *8*(6), 3208–3218. https://doi.org/10.1002/ece3.3852

Baquedano, F. J., Valladares, F., And Castillo, F. J. (2008). Phenotypic plasticity blurs ecotypic divergence in the response of *Quercus coccifera* and *Pinus halepensis* to water stress. *European Journal of Forest Research*, 127(6), 495–506. https://doi.org/10.1007/s10342-008-0232-8

Bassar, R. D., Marshall, M. C., López-Sepulcre, A., Zandonà, E., Auer, S. K., Travis, J., Reznick, D. N. (2010). Local adaptation in Trinidadian guppies alters ecosystem processes. *Proceedings of the National Academy of Sciences*, 107(8), 3616–3621. doi: 10.1073/pnas.0908023107

Benito‐Garzón, M., Ha‐Duong, M., Frascaria‐Lacoste, N. and Fernández‐Manjarrés, J., (2013). Habitat restoration and climate change: dealing with climate variability, incomplete data, and management decisions with tree translocations. *Restoration Ecology*, 21(5), pp.530-536.

Benjamini, Y., And Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B* (Methodological), 57(1), 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Boureima, Seyni, Murielle Eyletters, M. Diouf, T. A. Diop, and PJRJoES Van Damme (2011). Sensitivity of Seed Germination and Seedling Radicle Growth to Drought Stress in Sesame *Sesamum indicum L*. *Research Journal of Environmental Sciences* 5, no. 6: 557.

Christmas, M. J., Biffin, E., Breed, M. F., And Lowe, A. J. (2016). Finding needles in a genomic haystack: Targeted capture identifies clear signatures of selection in a nonmodel plant species. *Molecular Ecology*, 25(17), 4216–4233. doi: 10.1111/mec.13750

Daly, C., Widrlechner, M.P., Halbleib, M.D., Smith, J.I. and Gibson, W.P., (2012). Development of a new USDA plant hardiness zone map for the United States. *Journal of Applied Meteorology and Climatology*, 51(2), pp.242-264.

Dalziel, A.C., Rogers, S.M. and Schulte, P.M., (2009). Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology. *Molecular ecology*, 18(24), pp.4997-5017.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. doi: 10.1093/bioinformatics/btr330

De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O., And Mergeay, J. (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology*, 23(19), 4709–4721. doi: 10.1111/mec.12813

de Villemereuil, P., And Gaggiotti, O. E. (2015). A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, 6(11), 1248–1258. https://doi.org/10.1111/2041-210X.12418

de Villemereuil, P., Gaggiotti, O.E., Mouterde, M. and Till-Bottraud, I., (2016). Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity*, 116(3), pp.249-254.

Dey, D. C., Miller, G. W., And Kabrick, J. M. (2007). Sustaining northern red oak forests: managing oak from regeneration to canopy dominance in mature stands

Dickerson, G.E., (1947). Composition of hog carcasses as influenced by heritable differences in rate and economy of gain. *Iowa Agric. Home Econ. Exp*. Stn. Res. Bull., 28, p.1

Dray, S., Bauman, D., And Blanchet, G. (2022). “adespatial” Multivariate multiscale spatial analysis. Retrieved from https://github.com/sdray/adespatial

Dwyer, J. M., Hobbs, R. J., And Mayfield, M. M. (2014). Specific leaf area responses to environmental gradients through space and time*. Ecology*, 95(2), 399–410.

Earl, D. A., And vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361. doi: 10.1007/s12686-011-9548-7

Eaton, D. A. R., And Overcast, I. (2020). ipyRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, 36(8), 2592–2594. doi: 10.1093/bioinformatics/btz966

Eckert, A. J., Bower, A. D., González-Martínez, S. C., Wegrzyn, J. L., Coop, G., And Neale, D. B. (2010). Back to nature: Ecological genomics of loblolly pine (*Pinus taeda*, *Pinaceae*). *Molecular Ecology*, 19(17), 3789–3805. doi: 10.1111/j.1365-294X.2010.04698.x

Etterson, J.R., Cornett, M.W., White, M.A. and Kavajecz, L.C., (2020). Assisted migration across fixed seed zones detects adaptation lags in two major North American tree species. *Ecological Applications*, 30(5), p.e02092.

Etterson, J.R. (2008). Evolution in response to climate change. Chuck Fox and Scott Carroll (eds). *Conservation Biology: Evolution in Action*. Oxford University Press. New York, NY pp. 145-163

Excoffier, L., And Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3), 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Falconer, D.S., and Mackay, T.F., (1983). Quantitative genetics. London: Longman.

Filipe, J. C., Rymer, P. D., Byrne, M., Hardy, G., Mazanec, R., and Ahrens, C. W. (2022). Signatures of natural selection in a foundation tree along Mediterranean climatic gradients. *Molecular Ecology*, 31(6), 1735-1752. doi: 10.1111/mec.16351

Foll, M., And Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominantmarkers: A Bayesian perspective. *Genetics*, 180(2), 977–993. doi: 10.1534/genetics.108.092221

Francis, R. M. (2017). pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27–32. doi: 10.1111/1755-0998.12509

Frei, E. S., Scheepens, J. F., and Stöcklin, J. (2012). High genetic differentiation in populations of the rare alpine plant species *Campanula thyrsoides* on a small mountain. *Alpine Botany*, 122(1), 23-34. doi:10.1007/s00035-012-0103-2

Frichot, E., And François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925–929. doi: 10.1111/2041-210X.12382

Frichot, E., Schoville, S. D., Bouchard, G., And François, O. (2013). Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution*, 30(7), 1687–1699. doi: 10.1093/molbev/mst063

Gailing, O., And Zhang, R. (2018). Experimental evidence for selection against hybrids between two interfertile red oak species. *Silvae Genetica*, 67(1), 106–110. doi: 10.2478/sg-2018-0015

Gailing, O., Lind, J., And Lilleskov, E. (2012). Leaf morphological and genetic differentiation between *Quercus rubra L*. and *Q. ellipsoidalis* E.J. Hill populations in contrasting environments. *Plant Systematics and Evolution*, 298(8), 1533–1545. doi: 10.1007/s00606-012-0656-y

Galloway, L. F., And Etterson, J. R. (2007). Transgenerational plasticity is adaptive in the wild. *Science*, 318(5853), 1134–1136. doi: 10.1126/science.1148766

Galloway, L. F., Etterson, J. R., And McGlothlin, J. W. (2009). Contribution of direct and maternal genetic effects to life-history evolution. *The New Phytologist*, 183(3), 826–838.

Gao, J., Liu, Z.-L., Zhao, W., Tomlinson, K. W., Xia, S.-W., Zeng, Q.-Y., Wang, X.-R., And Chen, J. (2021). Combined genotype and phenotype analyses reveal patterns of genomic adaptation to local environments in the subtropical oak *Quercus acutissima*.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, 40(Database issue), D1178–D1186. https://doi.org/10.1093/nar/gkr944

Gribko, L. S., And Jones, W. E. (1995). Test of the float method of assessing northern red oak acorn condition. 5.

Gugger, P. F., Fitz-Gibbon, S., PellEgrini, M., And Sork, V. L. (2016). Species-wide patterns of DNA methylation variation in *Quercus lobata* and their association with climate gradients. *Molecular Ecology*, 25(8), 1665–1680. doi: 10.1111/mec.13563

Gugger, P. F., Fitz‐Gibbon, S. T., Albarrán‐Lara, A., Wright, J. W., And Sork, V. L. (2021). Landscape genomics of *Quercus lobata* reveals genes involved in local climate adaptation at multiple spatial scales. *Molecular Ecology*, 30(2), 406–423. doi: 10.1111/mec.15731

Halbritter, A. H., Billeter, R., Edwards, P. J., and Alexander, J. M. (2015). Local adaptation at range edges: comparing elevation and latitudinal gradients. *Journal of Evolutionary Biology*, 28(10), 1849-1860. https://doi.org/10.1111/jeb.12701

Hipp, A.L., (2010) Hill’s oak: the taxonomy and dynamics of a Western Great Lake endemic. *Arnoldia*, 67(4):2–14

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*. 3;81:1247–1266.

Jakobsson, M., And Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806. doi: 10.1093/bioinformatics/btm233

Jensen, R. J., Hokanson, S. C., Isebrands, J. G., And Hancock, J. F. (1993). Morphometric variation in oaks of the Apostle Islands in Wisconsin: Evidence of hybridization between *Quercus rubra* and *Q. ellipsoidalis* (Fagaceae). *American Journal of Botany*, 80(11), 1358–1366. doi: 10.1002/j.1537-2197.1993.tb15375.x

Jombart T. (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 24 (11):1403-5. doi: 10.1093/bioinformatics/btn129. Epub 2008 Apr 8. PMID: 18397895.

Kamvar, Z. N., López-Uribe, M. M., Coughlan, S., Grünwald, N. J., Lapp, H., & Manel, S. (2016). Developing educational resources for population genetics in R: An open and collaborative approach. *Molecular Ecology Resources*. https://doi.org/10.1111/1755-0998.12558

Kapoor, B., Jenkins, J., Schmutz, J., Zhebentyayeva, T., Kuelheim, C., Coggeshall, M., Heim, C., Lasky, J. R., Leites, L., Islam-Faridi, N., Romero-Severson, J., DeLeo, V. L., Lucas, S. M., Lazic, D., Gailing, O., Carlson, J., & Staton, M. (2023). A haplotype-resolved chromosome-scale genome for *Quercus rubra L*. provides insights into the genetics of adaptive traits for red oak species. *G3* (Bethesda, Md.), 13(11), jkad209. https://doi.org/10.1093/g3journal/jkad209

Khodwekar, S., And Gailing, O. (2017). Evidence for environment-dependent introgression of adaptive genes between two red oak species with different drought adaptations. *American Journal of Botany*, 104(7), 1088–1098. doi: 10.3732/ajb.1700060

King, G. M., Gugerli, F., Fonti, P., & Frank, D. C. (2013). Tree growth response along an elevational gradient: climate or genetics? *Oecologia*, 173(4), 1587–1600. https://doi.org/10.1007/s00442-013-2696-6

Kremer, A., And Hipp, A. L. (2019). Oaks: An evolutionary success story. *New Phytologist*, 226(4), 987–1011. doi: 10.1111/nph.16274

Kremer, A., Potts, B.M. and Delzon, S. (2014). Genetic divergence in forest trees: understanding the consequences of climate change. *Functional Ecology*, pp.22-36.

Kremer, A. (2010). Evolutionary responses of European oaks to climate change. *Irish Forestry*.

Lepais, O., And Bacles, C. F. E. (2014). Two are better than one: Combining landscape genomics and common gardens for detecting local adaptation in forest trees. *Molecular Ecology*, 23(19), 4671–4673. doi: 10.1111/mec.12906

Leroy, T., Louvet, J. M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J. M., and Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist*, 226(4), 1171-1182. doi: 10.1111/nph.16095

Lind, J.F. and Gailing, O., (2013). Genetic structure of *Quercus rubra L*. and *Quercus ellipsoidalis* EJ Hill populations at gene-based EST-SSR and nuclear SSR markers. *Tree genetics & genomes*, 9, pp.707-722.

Lind, B., Menon, M., Bolte, C., Faske, T., and Eckert, A. (2017). The genomics of local adaptation in trees: Are we out of the woods yet? *Tree Genetics And Genomes*, 14. doi: 10.1007/s11295-017-1224-y

Lind-Riehl, J., And Gailing, O. (2016). Adaptive variation and introgression of a CONSTANS-like gene in North American red oaks. *Forests*, 8(1), 3. doi: 10.3390/f8010003

Lind-Riehl, J. F., Sullivan, A. R., And Gailing, O. (2014). Evidence for selection on a CONSTANS-like gene between two red oak species. *Annals of Botany*, 113(6), 967–975. doi: 10.1093/aob/mcu019

Lindback, E.C., Rauschendorfer, J.K., Burton, A.J., Külheim, C. and Cavaleri, M.A., (2023). Common garden study reveals frost-tolerant northern seed sources are best suited to expand range of *Quercus rubra*. *Forest Ecology and Management*, 539, p.120985.

Liu, M., & Lu, S. (2016). Plastoquinone and Ubiquinone in Plants: Biosynthesis, Physiological Function and Metabolic Engineering. *Frontiers in plant science*, 7, 1898. https://doi.org/10.3389/fpls.2016.01898

Lobo, A., Hansen, J. K., Hansen, L. N., & Dahl, K. E. (2018). Differences among six woody perennials native to northern Europe in their level of genetic differentiation and adaptive potential at fine local scale. *Ecology and Evolution*, 8, 2231–2239. https://doi.org/10.1002/ ece3.3824

Lynch, M., & Walsh, B. (1998). Genetics and analysis of quantitative traits (Vol. 1, pp. 535-557). Sunderland, MA: Sinaue

McKown, A.D., Guy, R.D., Klápště, J., Geraldes, A., Friedmann, M., Cronk, Q.C., El‐Kassaby, Y.A., Mansfield, S.D. and Douglas, C.J., (2014). Geographical and environmental gradients shape phenotypic trait variation and genetic structure in P *opulus trichocarpa*. *New Phytologist*, 201(4), pp.1263-1276.

Mimura, M., And Aitken, S. N. (2010). Local adaptation at the range peripheries of Sitka spruce. *Journal of Evolutionary Biology*, 23(2), 249–258. https://doi.org/10.1111/j.1420-9101.2009.01910.x

MNDNR. (2013). The Relevé method 2nd edition. MNDNR. Retrieved from MNDNR website: https://files.dnr.state.mn.us/eco/mcbs/releve/releve\_booklet.pdf

MNDNR. (2024).Seed zone collection map. MNDNR. Retrieved from MNDNR website: https://arcgis.dnr.state.mn.us/portal/home/item.html?id=94c92579e94941c0a9035d8b40290436

Moen, S. (2018). Lake Superior’s Natural process. Minnesota Sea Grant. Retrieved from Minnesota Sea Grant website: http://www.seagrant.umn.edu/superior/processes.

Moran, E. V., Willis, J., And Clark, J. S. (2012). Genetic evidence for hybridization in red oaks (*Quercus* sect. Lobatae , Fagaceae). *American Journal of Botany*, 99(1), 92–100. doi: 10.3732/ajb.1100023

Murray, K. D., Janes, J. K., Jones, A., Bothwell, H. M., Andrew, R. L., And Borevitz, J. O. (2019). Landscape drivers of genomic diversity and divergence in woodland Eucalyptus*. Molecular Ecology*, 28(24), 5232–5247. doi: 10.1111/mec.15287

Nagamitsu, T., Uchiyama, K., Izuno, A., Shimizu, H., And Nakanishi, A. (2020). Environment-dependent introgression from *Quercus dentata* to a coastal ecotype of *Quercus mongolica var. Crispula* in northern Japan. *New Phytologist*, 226(4), 1018–1028. https://doi.org/10.1111/nph.16131

Oksanen, F.J., et al. (2017) Vegan: Community Ecology Package. R package Version 2.4-3. https://CRAN.R-project.org/package=vegan

Olsson, K., and Ågren, J. (2002). Latitudinal population differentiation in phenology, life history and flower morphology in the perennial herb *Lythrum salicaria*. *Journal of Evolutionary Biology*, 15(6), 983-996. can't find DOI - might not exist

OpenWetWare. (2018). Sork Lab: Protocols. OpenWetWare. Retrieved from https://openwetware.org/mediawiki/index.php?title=Sork\_Lab:ProtocolsAndoldid=1037048

Pais, A. L., Whetten, R. W., And Xiang, Q. (Jenny). (2017). Ecological genomics of local adaptation in *Cornus florida L*. by genotyping by sequencing. *Ecology and Evolution*, 7(1), 441–465. doi: 10.1002/ece3.2623

Petit, R. J., Latouche-Hallé, C., Pemonge, M. H., and Kremer, A. (2002). Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. *Forest Ecology and Management*, 156(1-3), 115-129. doi:10.1016/S0378-1127(01)00638-7

Pettenkofer, T., Finkeldey, R., Müller, M., Krutovsky, K. V., Vornam, B., Leinemann, L., And Gailing, O. (2020). Genetic variation of introduced red oak (*Quercus rubra*) stands in Germany compared to North American populations. *European Journal of Forest Research*, 139(2), 321–331. doi: 10.1007/s10342-019-01256-5

Platenkamp, G. A. J., And Shaw, R. G. (1993). Environmental and genetic maternal effects on seed characters in *Nemophila menziesii. Evolution*, 47(2), 540–555. https://doi.org/10.2307/2410070

Pritchard, J. K., Stephens, M., And Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. doi: 10.1093/genetics/155.2.945

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559–575. doi: 10.1086/519795

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

Ramírez-Valiente, J. A., Valladares, F., Gil, L., And Aranda, I. (2009). Population differences in juvenile survival under increasing drought are mediated by seed size in cork oak (*Quercus suber L*.). *Forest Ecology and Management*, 257(8), 1676–1683. https://doi.org/10.1016/j.foreco.2009.01.024

Ramírez-Valiente, J. A., Center, A., Sparks, J. P., Sparks, K. L., Etterson, J. R., Longwell, T., Pilz, G., And Cavender-Bares, J. (2017). Population-level differentiation in growth rates and leaf traits in seedlings of the neotropical live oak *Quercus oleoides* grown under natural and manipulated precipitation regimes. *Frontiers in Plant Science*, 8. https://www.frontiersin.org/articles/10.3389/fpls.2017.00585

Ramírez-Valiente, Deacon, N. J., Etterson, J., Center, A., Sparks, J. P., Sparks, K. L., Cavender-Bares, J. (2018). Natural selection and neutral evolutionary processes contribute to genetic divergence in leaf traits across a precipitation gradient in the tropical oak *Quercus oleoides. Molecular Ecology*, 27(9), 2176–2192. doi: 10.1111/mec.14566

Ramírez‐Valiente, Jose Alberto, Luis Santos del Blanco, Ricardo Alía, Juan J. Robledo‐Arnuncio, and Jose Climent. (2022) Adaptation of Mediterranean forest species to climate: Lessons from common garden experiments. *Journal of Ecology* 110 (5) 1022-1042. doi:10.1111/1365-2745.13730

Rauschendorfer, J.K., Lindback, E.C., Rooney, R.A., Frantti, S., Peck, V., Cavaleri, M.A. and Külheim, C., (2021). *Quercus rubra* seedling biomass response related to mean annual temperature conditions of associated provenance. *In Northern Hardwood Conference 2021: Bridging Science and Management for the Future* (p. 280).

Rauschendorfer, J.K., Rooney, R.A., and Külheim, C., (2022). Strategies to mitigate shifts in red oak (*Quercus sect. Lobatae*) distribution under a changing climate. *Tree Physiology*. 2022; 42:2383–2400, https://doi.org/10.1093/treephys/tpac090

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., and Hamilton, D. A. (1999). Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. *Ecological Monographs*, 69(3), 375–407. https://doi.org/10.2307/2657162

Rehfeldt, G. E., Tchebakova, N. M., Parfenova, Y. I., Wykoff, W. R., Kuzmina, N. A., And Milyutin, L. I. (2002). Intraspecific responses to climate in *Pinus sylvestris*: responses to climate in Pinus sylvestris. *Global Change Biology*, 8(9), 912–929. https://doi.org/10.1046/j.1365-2486.2002.00516.x

Rehfeldt, G. E., Leites, L. P., Clair, J. B. S., Jaquish, B. C., Saenz-Romero, C., Lopez-Upton, J., And Joyce, D. G. (2014). Comparative genetic responses to climate in the varieties of Pinus ponderosa and *Pseudotsuga menziesii*: Clines in growth potential. Forest *Ecology and Management*. 324: 138-146., 138–146. https://doi.org/10.1016/j.foreco.2014.02.041

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., And Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17), 4348–4370. doi: 10.1111/mec.13322

Rellstab, C., Zoller, S., Walthert, L., Lesur, I., Pluess, A. R., Graf, R., Gugerli, F. (2016). Signatures of local adaptation in candidate genes of oaks (*Quercus* spp.) with respect to present and future climatic conditions. *Molecular Ecology*, 25(23), 5907–5924. doi: 10.1111/mec.13889

Ren, G. P., Abbott, R. J., Zhou, Y. F., Zhang, L. R., Peng, Y. L., and Liu, J. Q. (2012). Genetic divergence, range expansion and possible homoploid hybrid speciation among pine species in Northeast China. *Heredity*, 108(5), 552-562. doi: 10.1038/hdy.2011.123

Roach, D. A., And Wulff, R. D. (1987). Maternal effects in plants. Annual Review of *Ecology and Systematics*, 18(1), 209–235. https://doi.org/10.1146/annurev.es.18.110187.001233

Sander, I. L. (1990). *Quercus rubra*. In Silvics of North America: Volume 2. Hardwoods (Vol. 2, pp. 727–733). United States Department of Agriculture (USDA), Forest Service, Agriculture Handbook 654

SAS Institute Inc. (2021). Carry, NC: JMP Version 6.0.0.

Sato, Y., and Kudoh, H. (2013). Relative strength of phenotypic selection on the height and number of flowering-stalks in the rosette annual *Cardamine hirsuta* (Brassicaceae). *Journal of Ecology and Environment*, 36(3), 151-158. http://dx.doi.org/10.5141/ecoenv.2013.151

Savolainen, O., Pyhäjärvi, T. and Knürr, T. (2007). Gene flow and local adaptation in trees. *Annu. Rev. Ecol. Evol. Syst*., 38, pp.595-619.

Schueler, S., And Schlünzen, K. H. (2006). Modeling of oak pollen dispersal on the landscape level with a mesoscale atmospheric model. *Environmental Modeling And Assessment,* 11(3), 179–194. doi: 10.1007/s10666-006-9044-8

Scotti, I., Lalagüe, H., Oddou‐Muratorio, S., Scotti‐Saintagne, C., Ruiz Daniels, R., Grivet, D., and Vendramin, G. G. (2023). Common microgeographical selection patterns revealed in four European conifers. *Molecular Ecology*, 32(2), 393-411. https: //doi.org/10.1111/mec.16750Citations

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., And Neale, D. B. (2013). Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics And Genomes*, 9(4), 901–911. doi: 10.1007/s11295-013-0596-x

Sork, V. L., Cokus, S. J., Fitz-Gibbon, S. T., Zimin, A. V., Puiu, D., Garcia, J. A., Gugger, P. F., Henriquez, C. L., Zhen, Y., Lohmueller, K. E., Pellegrini, M., And Salzberg, S. L. (2022). High-quality genome and methylomes illustrate features underlying evolutionary success of oaks. *Nature Communications*, 13, 2047. https://doi.org/10.1038/s41467-022-29584-y

Thiede, D. A. (1998). Maternal inheritance and its effect on adaptive evolution: A quantitative genetic analysis of maternal effects in a natural plant population. *Evolution*, 52(4), 998–1015. https://doi.org/10.2307/2411232

Tirmenstein, D. A. (1991). *Quercus rubra*. In: Fire Effects Information System. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. Retrieved from U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory website: https://www.fs.fed.us/database/feis/plants/tree/querub/all.html

Torres-Ruiz, J.M., Kremer, A., Carins Murphy, M.R., Brodribb, T., Lamarque, L.J., Truffaut, L., Bonne, F., Ducousso, A. and Delzon, S. (2019). Genetic differentiation in functional traits among European sessile oak populations. *Tree physiology*, 39(10), pp.1736-1749.

Truffaut, L., Chancerel, E., Ducousso, A., Dupouey, J. L., Badeau, V., Ehrenmann, F., And Kremer, A. (2017). Fine-scale species distribution changes in a mixed oak stand over two successive generations. *New Phytologist*, 215(1), 126–139. https://doi.org/10.1111/nph.14561

Twardek, W.M., Taylor, J.J., Rytwinski, T., Aitken, S.N., MacDonald, A.L., Van Bogaert, R., and Cooke, S.J. (2023). The application of assisted migration as a climate change adaptation tactic: An evidence map and synthesis. *Biological Conservation*, 280, p.109932.

Twyford, A.D., Wong, E.L. and Friedman, J. (2020). Multi-level patterns of genetic structure and isolation by distance in the widespread plant *Mimulus guttatus*. *Heredity*, 125(4), pp.227-239.

USDA. (2019). USDA Hardiness Zones. Retrieved from https://planthardiness.ars.usda.gov/PHZMWeb/

Vakkari, P., Blom, A., Rusanen, M., Raisio, J., And Toivonen, H. (2006). Genetic variability of fragmented stands of pedunculate oak (*Quercus robur*) in Finland. *Genetica*, 127(1–3), 231–241. doi: 10.1007/s10709-005-4014-7

Vivas, M., Wingfield, M.J. and Slippers, B., 2020. Maternal effects should be considered in the establishment of forestry plantations. *Forest Ecology and Management*, 460, p.117909.

von Takach, B., Penton, C. E., Murphy, B. P., Radford, I. J., Davies, H. F., Hill, B. M., and Banks, S. C. (2021). Population genomics and conservation management of a declining tropical rodent. *Heredity*, 126(5), 763-775. doi: 10.1038/s41437-021-00418-9

Wang, J. (2017), The computer program structure for assigning individuals to populations: easy to use but easier to misuse. *Mol Ecol Resour*, 17: 981-990. https://doi.org/10.1111/1755-0998.12650

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of F(ST). *The American naturalist*, 186 Suppl 1, S24–S36. https://doi.org/10.1086/682949

Wulff, R. D. (1995). Environmental maternal effects on seed quality and germination. In Seed Development and Germination. Routledge

Zas, R., And Sampedro, L. (2015). Heritability of seed weight in Maritime pine, a relevant trait in the transmission of environmental maternal effects. *Heredity*, 114(1), Art.1. https://doi.org/10.1038/hdy.2014.76

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., And Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, 28(24), 3326–3328. doi: 10.1093/bioinformatics/bts606

**Data Accessibility Statement:** FASTA and VCF files for dataset 1 and 2, quantitative genetic data and metadata for the genomic and common garden samples (including tissue collected, latitude and longitude of collections and year of collection) and are freely available at the University of Minnesota Data Repository (<https://conservancy.umn.edu/handle/11299/212844>).

**Benefits-Sharing Statement:** Benefits from this research include the sharing of both the results of the study and the data the results are based on, as described above. These results contribute to the conservation and management of Lake Superior coastal forests.

**Acknowledgements:** This study was funded in part by the Coastal Zone Management Act of 1972, as amended, administered by the Office for Coastal Management, National Oceanic and Atmospheric Administration under Award NA17NOS4190062 provided to the Minnesota Department of Natural Resources for Minnesota’s Lake Superior Coastal Program. The authors acknowledge the Minnesota Supercomputing Institute (MSI) at the University of Minnesota for providing resources that contributed to the research results reported within this paper. URL:<http://www.msi.umn.edu>. This work was also supported by the resources and staff at the University of Minnesota Genomics Center (<https://genomics.umn.edu>). We would also like to acknowledge Sophie Leaf, Michael Laidlaw, Joe Rose, Daniela Gamba, Taylor Koehnen and Samantha Ykema for assistance with data collection, and fieldwork. We also thank greenhouse manager Matt Jahnke and the University of Minnesota Duluth Greenhouse for providing us with support and space to perform this study. Lastly, We would like to help everyone who provided feedback for this manuscript especially the Gross and Etterson Labs.

**Author Contributions:** All authors contributed equally to the conception and design of this study and the writing of the manuscript. All data collection and analyses were completed by María José Gómez Quijano, with assistance from Julie R. Etterson and Briana L. Gross