Conservation Genetics
https://doi.org/10.1007/s10592-021-01378-7

RESEARCH ARTICLE

Genomic drivers of early-life fitness in *Picea rubens*

Thibaut Capblancq[1] · Helena Munson[1] · John R. Butnor[2] · Stephen R. Keller[1]

Received: 2 June 2020 / Accepted: 12 May 2021

Abstract
Red spruce (*Picea rubens* Sarg.) is a coniferous tree with a highly fragmented range in eastern North American montane forests. It serves as a foundational species for many locally rare and threatened taxa and has therefore been the focus of large-scale reforestation efforts aimed at restoring these montane ecosystems, yet genetic input guiding these efforts has been lacking. To tackle this issue, we took advantage of a common garden experiment and a whole exome sequencing dataset to investigate the impact of different population genetic parameters on early-life seedling fitness in red spruce. The level of inbreeding, genetic diversity and genetic load were assessed for 340 mother trees sampled from 65 localities across the species range and compared to different fitness traits measured on 5100 of their seedlings grown in a controlled environment. We identified an overall positive influence of genetic diversity and negative impact of genetic load and population-level inbreeding on early-life fitness. Those associations were most apparent for the highly fragmented populations in the Central and Southern Appalachians, where lower genetic diversity and higher inbreeding were associated with lower germination rate, shorter height and reduced early-life fitness of the seedlings. These results provide unprecedented information that could be used by field managers aiming to restore red spruce forests and to maximize the success of future plantations.

Keywords Conifers · Fitness · Inbreeding depression · Genzetic load · Genetic diversity · Whole exome sequencing

Introduction

Identifying factors that induce variation in the fitness of natural populations is a central question in conservation biology. Besides the large influence of environmental stressors (abiotic and biotic) or the direct impact of human activities (Anderson 2015; Liu et al. 2013), the intrinsic genetic characteristics of populations, arising from their demographic and evolutionary histories, are known to play important roles in driving fitness variation (Allendorf et al. 2010; Allendorf and Luikart 2006; Frankham et al. 2010; Lande, 1995). Among those factors, the level of genetic diversity, the accumulation of a genetic load of deleterious mutations, and the degree to which genetic load is expressed through inbreeding are especially relevant for small or declining populations, especially in the context of conservation and management (Kardos et al. 2016; Latta 2008; Willi et al. 2018).

Inbreeding, either through mating among genetically related individuals (consanguineous mating) or through self-reproduction (selfing), increases the rate of population genetic drift and may incur individual and/or population-level fitness effects (Angeloni et al. 2011; Mosseler et al. 2000; Reed and Frankham 2003). It is well established that inbreeding increases the frequency of homozygous genotypes in an inbred population (Wang et al. 1998). As a result, there is an increased probability of homozygosity for lethal alleles as well as increased expression of mildly deleterious (i.e., non-lethal) alleles across the genome that are otherwise masked when in the heterozygous state (Charlesworth and Willis 2009). When the trend is strong enough it causes a reduction in fitness called inbreeding depression (Keller and Waller 2002). The accumulation of deleterious mutations can also occur independently from inbreeding when the effective population size (N_e) is very

Thibaut Capblancq and Helena Munson have contributed equally to this work.

✉ Thibaut Capblancq
thibaut.capblancq@gmail.com

✉ Stephen R. Keller
Stephen.Keller@uvm.edu

[1] Department of Plant Biology, University of Vermont, Burlington, VT 05405, USA

[2] USDA Forest Service, Southern Research Station, University of Vermont, Burlington, VT 05405, USA

Published online: 18 May 2021

small, for example due to population isolation, decline or a bottleneck (Lynch et al. 1995). In this context, genetic drift can be strong enough to impede the effect of purifying selection and allow an accumulation of deleterious mutations over time (Peischl et al. 2013; 2015). This genetic load of deleterious mutations can negatively impact fitness if they drift to high frequencies in populations (Agrawal and Whitlock 2012). In addition to inbreeding depression and genetic load, low levels of genetic diversity have also been identified as one of the genetic threats to small population persistence (Lande, 1995). Erosion of genetic diversity can not only increase the likelihood of inbreeding depression, but also reflects a lowered level of adaptive potential that could prevent the population from responding evolutionarily to future changes in the environment (Markert et al. 2010).

Inbreeding, genetic load, and genetic diversity are strongly dependent on current patterns of gene flow and population connectivity, as well as historical processes such as range expansions or contractions in response to past climate changes. Evidence from fossil pollen suggests that migration of many tree species occurred towards higher elevations and more northerly latitudes that supported favorable environmental conditions as the climate changed at the end of the last ice age (Boisvert-Marsh et al. 2014; Koo et al. 2014; Schauffler and Jacobson 2002). These migration patterns often resulted in the establishment of "leading" and "trailing (rear) edge" populations within a species range that can possess distinct yet complex demographic and genetic histories (Hampe and Petit 2005). The rear range edge of many species became more fragmented as ranges shifted northward, which may have led to higher levels of inbreeding and potentially lower fitness than populations in the leading edge found in more favorable environments (Levin 2011; Willi et al. 2018). Conversely, rear edge populations may retain more ancestral genetic diversity, which can be reduced at the leading edge through founder effects and genetic drift due to successive migration events (Mosseler et al. 2003; Pluess 2011). For example, a study on holm oak (*Quercus ilex* L.) provided evidence that successive long-distance dispersal events have produced reductions in genetic diversity at the (northern) leading edge due to founder effects (Hampe et al. 2013). The succession of founder events can also lead to the accumulation of deleterious mutations in leading-edge populations (i.e., expansion load), by increasing genetic drift and lowering the effect of purifying selection (Peischl et al. 2013). Thus, both current population connectivity and historic demographic events and range expansion are likely to leave different patterns of genetic diversity in different parts of the range. In the context of conservation efforts, the main goal then becomes understanding the impact of such patterns on individual and population-level fitness across the species range.

Red spruce (*Picea rubens* Sarg.) is a conifer species endemic to the eastern United States and southeastern Canada. Historically, there is evidence to suggest that the range of red spruce expanded northward over the Holocene period in response to climatic changes (Schauffler and Jacobson 2002). Fossil *Picea* pollen preserved in sediments of lakes and bogs show that climate change at the end of the last ice age resulted in dramatic latitudinal and elevational shifts in abundance, degree of population fragmentation, and genetic diversity (Davis and Shaw 2001). Probably due to those historical events, the current spatial distribution of red spruce shows a north–south gradient of range fragmentation (Koo et al. 2014; Major et al. 2015) associated with three main genetic clusters (Capblancq et al. 2020). These three genetic clusters differentiate a population in the main contiguous portion of the range in the northeastern US and southeastern Canada (this part of the range is hereafter called the 'Core' region) from a population in Pennsylvania where the range becomes fragmented (hereafter called the 'Margin' region) and finally a low latitude, highly fragmented trailing edge population in the southern Appalachians (hereafter called the 'Edge' region), where populations occur as mountaintop "sky islands". Recently, red spruce populations have also experienced significant human-caused decline from logging and fire in the late 1800 to early 1900s, and from atmospheric pollution leading to acid rain (Mathias and Thomas 2018; Rentch et al. 2016). These events likely impacted genetic diversity, potentially removing many rare genetic variants when entire forests disappeared but also accentuating population fragmentation, which is known to reduce gene flow and enhance inbreeding (Leimu et al. 2010; Ratnam et al. 2014). The decline of red spruce, by as much as 95% of its original areal extent in the highly fragmented southern portion of its range (Rentch and Schuler 2009), has become a major conservation focus among resource managers and restoration ecologists, which has led to the formation of multi-partner cooperatives (e.g., the Central and Southern Appalachian Spruce Restoration Initiatives; CASRI and SASRI, respectively) aimed at restoring functional red spruce ecosystems via large scale restoration plantings and relevant silvicultural practices. A better assessment of the associations between red spruce fitness, especially at the sensitive seedling stage, and genetic diversity or inbreeding levels is then critical to optimize the success of such restoration effort, for example by helping to select the best mix of seed sources for a particular planting site.

The purpose of this study is to investigate associations between early-life fitness traits and estimates of inbreeding, genetic load, and genetic diversity across the range of red spruce, and to understand these associations with regard to the evolutionary and demographic history of the species. We sampled tissue and seeds from 340 mother trees, germinated and grew 15 seeds per mother tree to measure different

early-life fitness traits. We focused on early-life traits as the transition from seed to seedling stage is of significant importance for the establishment of trees and their long-term reproductive success in forest ecosystems (Le et al. 2012). By comparing early-life fitness traits with genetic parameters estimated using data from a recently published exome-capture sequencing on the same mother trees (Capblancq et al. 2020), we tested for the influence of inbreeding, genetic load, and genetic diversity in shaping the early-life performance of red spruce seedlings. Specifically, we test the predictions that early-life fitness traits should be positively associated with estimates of genetic diversity and negatively associated with estimates of inbreeding and genetic load. We interpret the results in light of the genetic history of range expansion, population structure, and demographic change in red spruce to better understand the relationship between genetic variation and early-life fitness in this species of conservation and restoration importance.

Materials and methods

Sampling design

We sampled open-pollinated seeds and needle tissue from 340 individual mother trees of red spruce during the late summer and fall of 2017. Those 340 mother trees were sampled at 65 sampling locations (also called localities hereafter) distributed across the entirety of the species range in eastern North America (Fig. 1 and Supplementary Material S1). The sample sizes range from 2 to 11 individuals per locality, but most (N = 44 locality) included 5 or 6 individuals (Supplementary Material S1). Considering that red spruce is a wind-pollinated species with a principally outcrossing mating system, the sampled seeds represented 340 maternal families with a mix of full and potentially half siblings of different paternal ancestry within each family. Most seed were collected from naturally occurring red

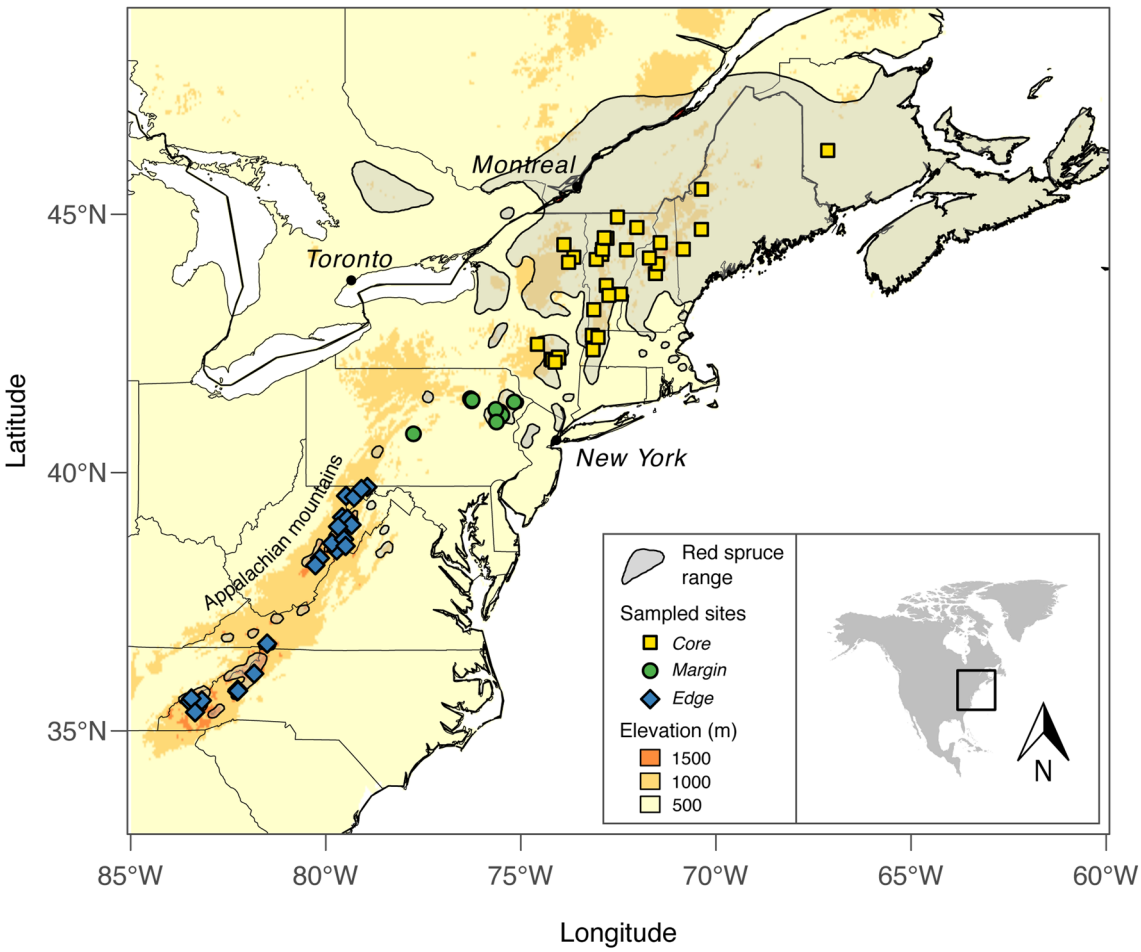

Fig. 1 Map showing the geographic range of red spruce in the eastern United States. Sampling locations are indicated by shape and color based on three genetically defined "regions" (Capblancq et al. 2020). Altitude across the range is indicated by color, with southern edge localities being predominantly at higher altitudes

spruce stands with the exception of 'NBTIC' (Supplementary Material S1). This latter locality consisted of open-pollinated seed collected from six mother trees growing in a red spruce orchard established by the New Brunswick Tree Improvement Council. Each NBTIC mother tree was originally collected from a different natural locality in New Brunswick, Canada. We pooled those trees into a single locality to be able to estimate the different genetic parameters used for the analyses, and because we wanted to include samples from this northernmost part of the species range. However, it is important to recognize that families from NBTIC do not constitute a local mating population in the same way our other sampled localities do. We consider this point again in the Discussion.

Phenotypic seedling trait collection

In order to determine variation in early-life fitness traits, we germinated 50 seeds and grew 15 seedlings from each of the 340 maternal families and measured four fitness-related traits. First, 50 filled seeds per family (N = 17,000 seeds total) were weighed in bulk to the nearest milligram and used to calculate a per-seed average (hereafter 'Seed Weight'). The weighed seeds were put into petri dishes containing wetted filter paper and coarse sand and placed in a germination chamber set to a daily temperature and photoperiod regime of 20 °C for 16 h of dark and 30 °C for 8 h of light. Germination was scored once the radicle of a seedling had visibly penetrated the seed coat, and the overall germination proportion (hereafter 'Germination') out of the initial 50 seeds was determined for each family. From the germinating seeds for each family, we transplanted 15 seedlings individually into 164 ml Ray Leach cone-tainers filled with ProMix BX potting media. Cone-tainer racks were transferred to the University of Vermont greenhouse to grow in the spring of 2018, maintaining temperatures of 21–24 °C during 16-h light and 15–18 °C during 8-h dark periods. After 12 weeks of growth in the greenhouse, the height of each seedling was measured as the stem length containing live foliage (sensu Butnor et al. 2019) to the nearest 0.1 mm using calipers (hereafter 'Height'). Survival for each seedling was also recorded after 12 weeks (hereafter 'Survival'). To calculate an overall multiplicative estimate of early-life fitness for each family (hereafter 'Early-Life Fitness'), we multiplied *Germination* by *Height,* weighting the average height per family by its *Survival* (i.e., dead seedlings contributing height values of 0).

Genetic parameter estimation

We based our estimation of genetic parameters on recently published exome capture sequencing of each of the 340 maternal trees (Capblancq et al. 2020). Briefly, this consisted of designing 80,000 probes based on almost 100 Mbp of genomic and exomic regions spanning an estimated 38,570 unigenes of the *P. glauca* genome. Exome-capture libraries were sequenced with Illumina to an average depth of 2.3X, mapped to the WS711 *P. glauca* reference genome (Birol et al. 2013). One individual that showed an abundance of PCR duplicates in the raw sequence data was removed (YRB_01), leaving 339 families for further analysis. Linkage Disequilibrium (LD) decreases very rapidly along the genome for red spruce (Capblancq et al. 2020) and exome capture is targeting many small fragments distributed all along the genome, limiting potential biases due to strong LD among genetic sites when estimating genomic parameters.

We used the program ANGSD (Analysis of Next Generation Sequencing Data) (Korneliussen et al. 2014) to identify single nucleotide polymorphisms (SNPs). This program is particularly suited to analyze low coverage sequencing because it takes into account genotypic uncertainty by calculating genotype likelihoods and/or probabilities for each SNP based on the depth of aligned reads and the associated mapping and sequencing quality scores (Nielsen et al. 2011). Genotype likelihoods were estimated using the SAMtools genotype likelihood model, using only reads having unique best hits ("-uniqueOnly 1"), setting a minimum MapQ score to keep a read to 20 ("-minMapQ 20"), a min nucleotide Q score to consider a site to 20 ("-minQ 20"), a minimum number of 2 individuals with coverage to keep a site ("-minInd 2"), a maximum of 17 reads to estimate genotype likelihood for one individual ("-setMaxDepthInd 17"), a minimum number of 15 reads across the complete sampling to estimate genotype likelihoods for a site ("-setMinDepth 15"), keeping only biallelic sites ("-skipTriallelic 1"), performing the base alignment quality (BAQ: Phred-scaled probability of a read base being misaligned)(Li 2011) as in SAMtools ("-baq 1"). The resulting genotype probabilities were then used to call genotypes with a probability threshold of 0.8 ("-doGeno 2", "-postCutoff 0.8").

We annotated the resulting SNPs based on best BLAST hits to the Norway spruce (*P. abies*) genome annotation available from congenie.org (Nystedt et al. 2013). Using the mapped positions of the red spruce SNPs within the Norway spruce reference, we then used SNPeff (Cingolani et al. 2012) to annotate variants to functional class (up- and downstream, synonymous, nonsynonymous, intronic or intergenic sites). ANGSD analyses and the annotations produced by SnpEff were used to estimate the following parameters of genetic diversity, inbreeding and genetic load at the sampling locations (Table 1):

Genetic diversity

First, in order to account for a putative positive impact of genetic diversity on individual survival and growth, we

Table 1 Summary of the genetic factors calculated and used for analysis in this study with the method applied for calculating values for each factor and a justification for its use in this study

Variable	Method	Purpose
Genetic diversity	Expected heterozygosity in a population (2pq)	Use as an estimator of genetic diversity within a population
Population homozygosity	Proportion of loci showing an excess in homozygosity compared to the expected value in the population (1-2pq)	Use as a proxy of non-random mating within a population (Population level inbreeding)
Family homozygosity	Proportion of individual's homozygote genotypes across all the polymorphic sites	Use as a proxy of alleles identity by descent levels for a Family (Family level inbreeding)
Genetic load	Mean number of non-synonymous mutation weighted by their frequencies in the population	Estimate the amount of deleterious alleles within a population
Genetic structure	Individuals scores along the first two axes of a genetic PCA (PC1 & PC2)	Integrate the species genetic structure at the Family level

The parameters p and q mentioned represent the frequency of the major and the minor allele in the population, respectively

calculated the expected heterozygosity within each sampling location, as a descriptor of its genetic diversity. For each polymorphic site, we estimated allele frequencies and calculate the expected heterozygosity under Hardy–Weinberg equilibrium (i.e., 2pq). To avoid bias due to sites under strong selection in the coding regions of genes, we only used SNPs in introns or up- and down-stream of genes to estimate the expected heterozygosity within sampling locations.

Population-level excess of homozygosity

To characterize inbreeding at the population level, we tested for loci that deviated significantly from Hardy–Weinberg equilibrium based on comparing the expected heterozygosity estimated above to the mean observed heterozygosity. We then identified the proportion out of all SNPs that showed an excess of homozygosity and used this proportion as a metric for the magnitude of inbreeding at the population level. Here again, only SNPs in introns and up- or down-stream of genes were used.

Family-level homozygosity

We estimated the individual-level inbreeding coefficient by calculating for each family (i.e., per maternal tree) the proportion of homozygous genotypes within the intronic and up- and down-stream genes sites.

Genetic load

Effective population size and demographic expansion history can strongly influence the accumulation of deleterious alleles in populations (Peischl et al. 2015), negatively impacting fitness. We estimated genetic load for each sampled locality according to Willi et al. (2018) to approximate the local abundance and relative frequency of deleterious alleles. Based on the annotations made with SNPEff, we estimated genetic load as $P_n f_n / P_s f_s$, where P_n and P_s refers to the number of polymorphic nonsynonymous and synonymous sites, respectively, within a focal locality, and f_n and f_s similarly refer to the corresponding mean relative frequencies of nonsynonymous and synonymous sites within the same locality. Assuming that the majority of the non-synonymous polymorphisms are deleterious, this ratio provides a proxy of genetic load at each sampling location that accounts for both the abundance and mean frequencies of deleterious variants (Willi et al. 2018).

Genetic structure

To include in our model a potential role for different genetic backgrounds in explaining fitness traits as a result of variation in demographic histories, we used the three genetically-based groups (Core, Margin, and Edge) identified in Capblancq et al. (2020) as a categorical co-variable named "Region" (Fig. 1). These groups showed evidence of unique demographic histories, including separating into distinct clusters in a principal component analysis (PCA) based on > 100,000 SNPs, and estimated divergence times of > 8000 years ago with minimal subsequence gene flow, based on demographic modeling (Capblancq et al. 2020).

Statistical Analyses

Fitness trait variation

To visualize trait variation across the species range, we first generated boxplot distributions for the four traits within each locality. Values of *Height* and *Early-Life Fitness* were corrected for the rack effect before plotting using random effects ANOVA. Then, we tested the correlation among traits by running pairwise linear regressions, using the *lm*

function of the R package *stats* (R core team 2018). We also performed, for each trait, a one-way ANOVA with *Region* as the explanatory variable, to test the influence of genetic background on trait values.

Geographic variation in genetic parameters

To visualize the variation of the different genetic parameters across the species range, we performed ordinary kriging using the function *krige.conv* of the *geoR* R-package (Ribeiro and Diggle 2016) and fitting a Gaussian function model. The predictions obtained were superimposed on species range maps.

Association between genetic parameters and fitness traits

We used multivariate linear regression models to investigate the relative roles of genetic diversity, population- and individual-level homozygosity, genetic load, and genetic structure in driving family-level variation in fitness traits. We ran separate models for *Germination*, *Survival*, *Height*, and *Early-Life Fitness*. To account for potential (non-genetic) maternal environmental effects due to variation in resource availability or environment during seed production, we included *Seed Weight* as a co-variable in the multivariate regression for each trait. Interaction terms between the continuous genetic predictors and *Region* (as a discrete factor) were also included to test if the influence of genetic variables on fitness traits was region-specific. To summarize, the explanatory variables included in the model for each trait were: the four genetic parameters, *Seed Weight*, *Region* and the four interactions between *Region* and the genetic parameters.

Results

Variation in seedling early-life fitness traits

The distribution of early-life fitness traits (*Seed Weight, Germination, Survival, Height and Early-Life Fitness*) exhibited abundant variation, both among and within sampled localities (Fig. 2). On average, localities from the Margin region tended to achieve the highest values for

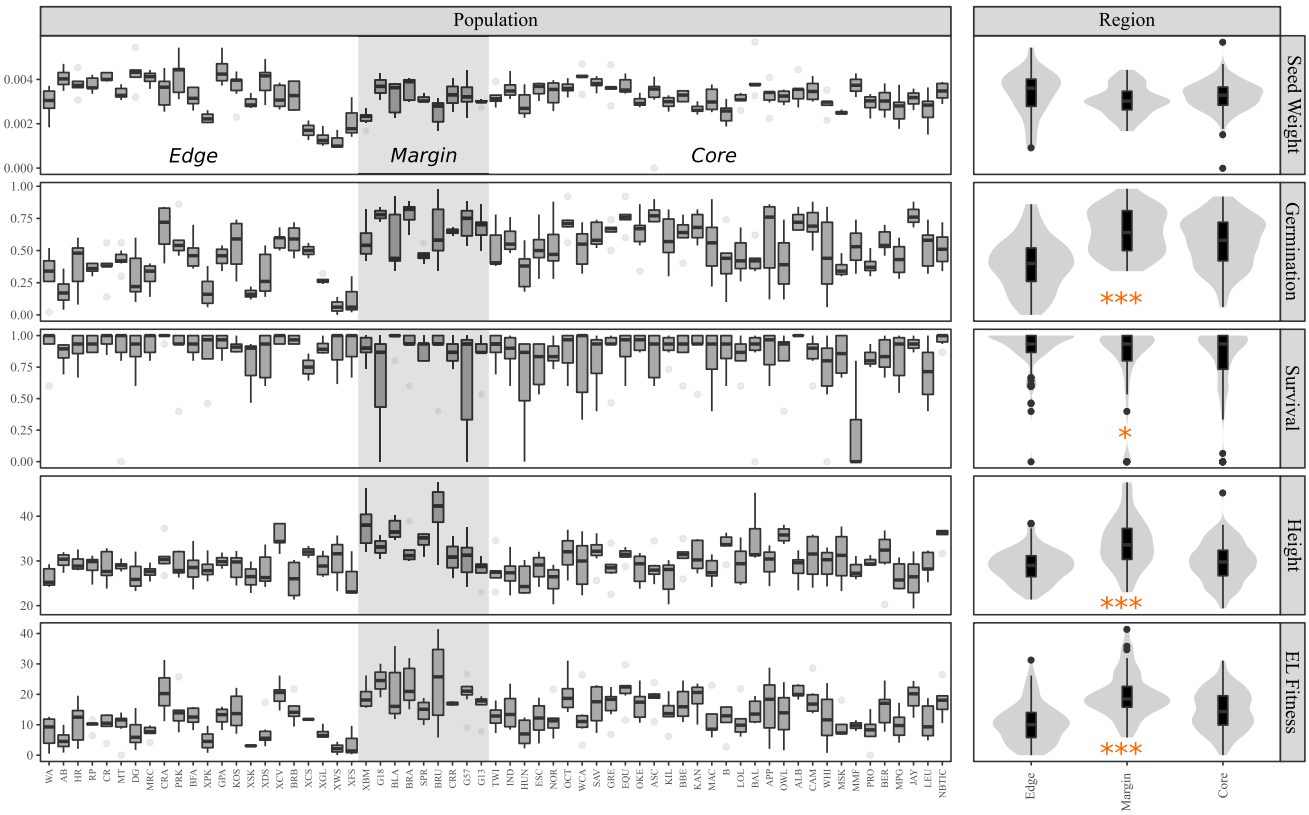

Fig. 2 Boxplot distribution of early-life fitness traits with respect to sampling location. Localities are ordered depending on their latitude, with the southernmost locality (WA) on the extreme left and the northernmost locality (NBTIC) on the extreme right of the x axis. The violin plots in the right panels show the variation in traits among the three genetic regions. Orange stars indicate the significance level in one-way ANOVA (***$P < 0.001$, *$0.01 \geq P \leq 0.05$)

most traits. Conversely, the Edge region tended to have lower trait values as well as more variation among localities than Core or Margin regions, the latter especially for *Seed Weight* and *Germination*.

Correlations among the individual early-life fitness traits were generally weak (R^2 0.01–0.06; Fig. 3). The exception was for correlations with *Early-Life Fitness,* which was logically strongly correlated with each of its component variables, but most strongly reflected variation in *Germination* ($R^2 = 0.75$) compared to the other components *Survival* ($R^2 = 0.24$) and *Height* ($R^2 = 0.19$). Interestingly, *Seed Weight* was significantly associated with *Germination* ($P < 0.001$, $R^2 = 0.06$) and *Fitness* ($P < 0.001$, $R^2 = 0.05$) but not with *Survival* ($P = 0.14$) and *Height* ($P = 0.28$).

Geographic structure in landscape genetic variation

The four different genetic parameters (family homozygosity, population homozygosity, genetic diversity and genetic load) showed geographic variation across the range but the patterns were more complex than simple latitudinal clines (Fig. 4). For example, family homozygosity was maximal in the southern Edge region, but was also high in the western part of the Core region in New York and Vermont. In contrast, population homozygosity was maximal in the northeast and moderately high elsewhere, except for Edge sampling locations in WV and MD and the parts of the Core region in northern VT. Genetic diversity and genetic load showed a surprisingly similar distribution, with both parameters highest in the northeastern Core and in the Margin regions.

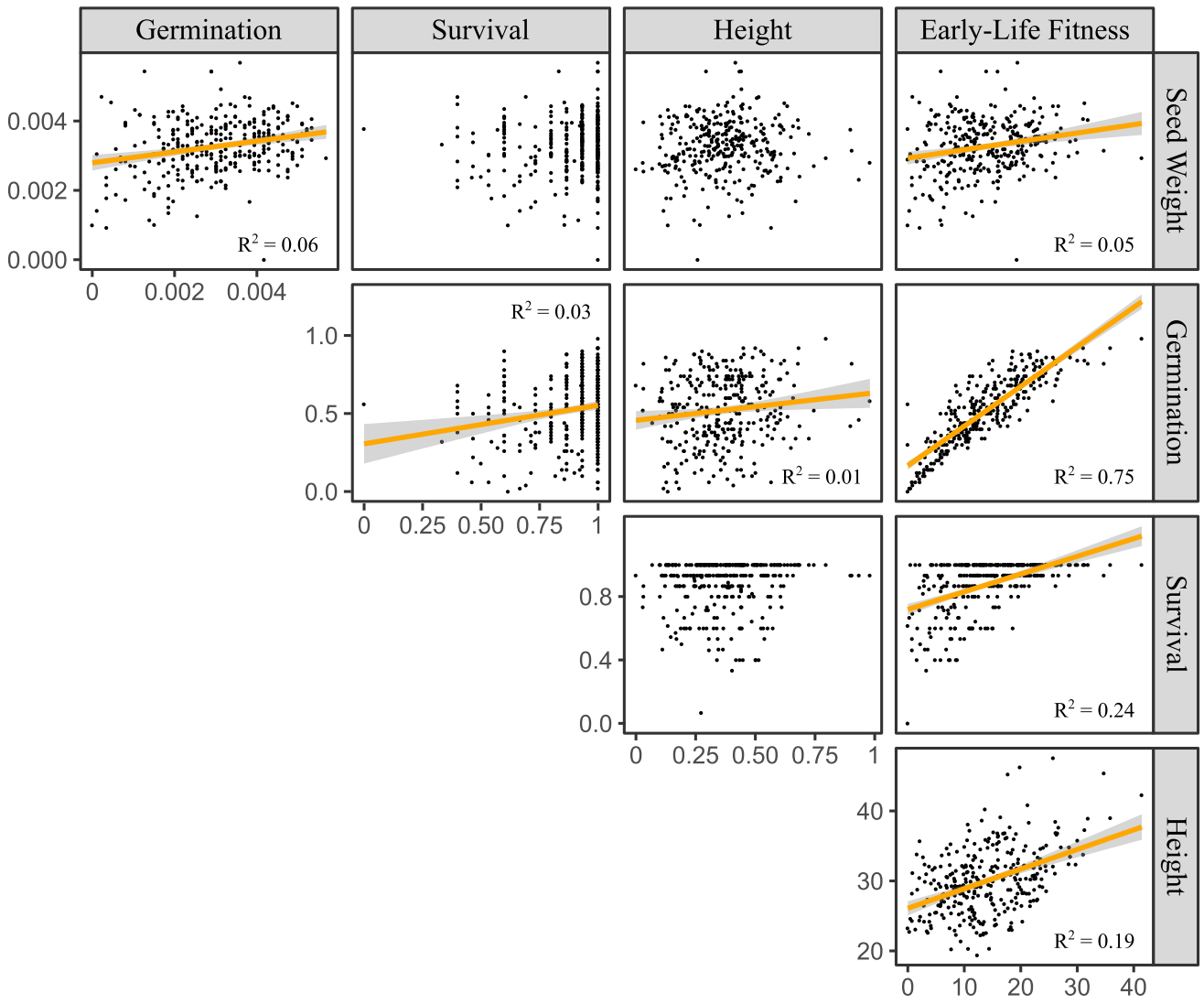

Fig. 3 Pairs plot showing pairwise simple correlations between all traits. Linear regression lines and confidence intervals were plotted when the correlation test was significant ($P < 0.05$)

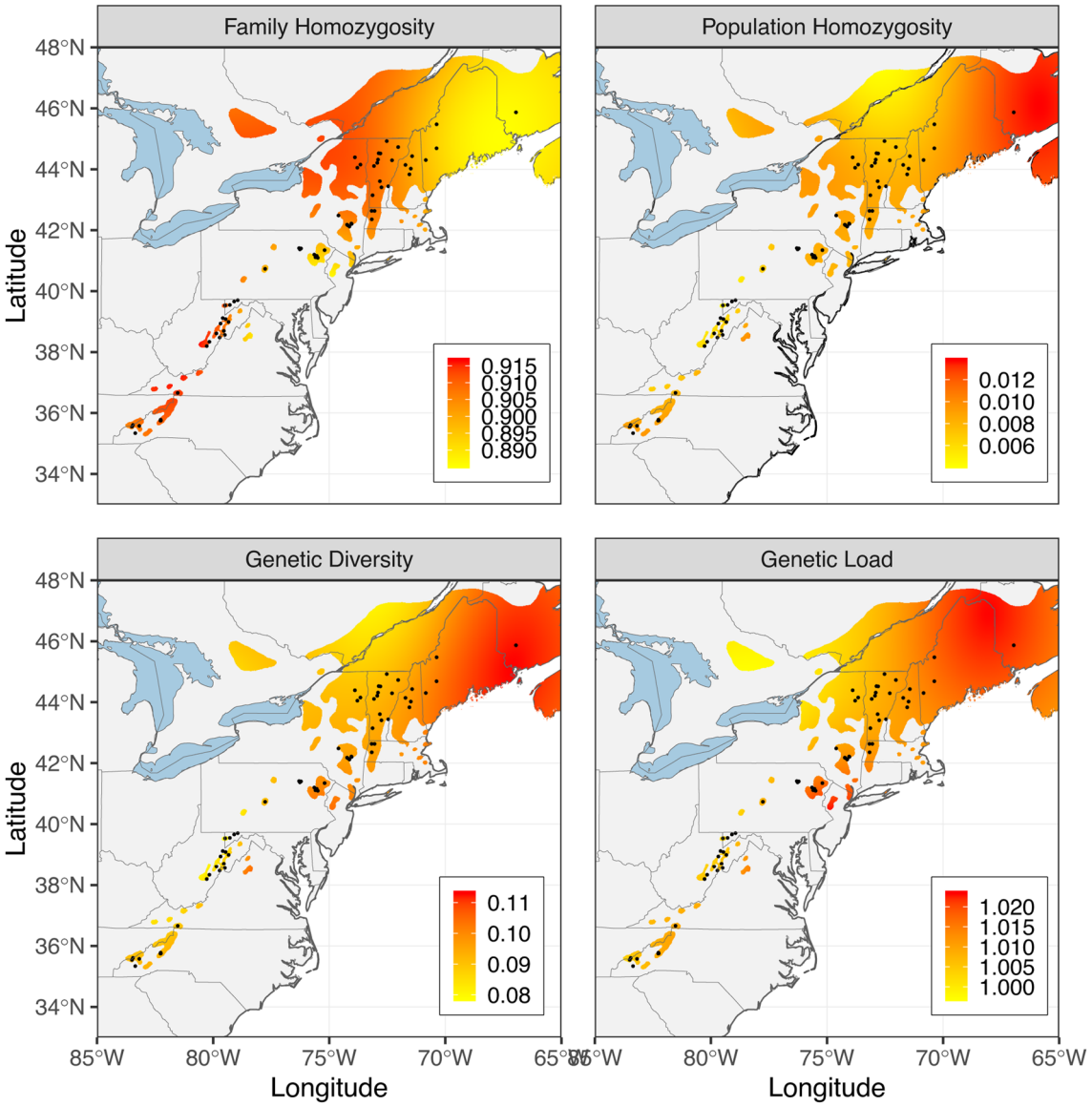

Fig. 4 Red spruce geographic range maps superimposed with the results of kriging of each of the genetic parameters using a Gaussian covariance function. Gradient scales range from low (yellow) to high (red) for each of the different parameters. Sampling locations are indicated by the black dots shown on each map

Genetic drivers of variation in earl- life fitness traits

Multivariate linear models revealed extensive variation among the different early-life fitness traits in how they responded to genetic parameters (Fig. 5). The overall model was significant for *Germination* ($P < 0.001$) and explained almost 28% of the variation in germination proportion among our range-wide sample of seed families ($R^2 = 0.279$). *Germination* was highly influenced by *Seed Weight*, but also showed significantly higher values in the Margin region. In contrast, *Survival* did not show any significant association with the genetic parameters (Overall model $P = 0.57$). The overall model for mean seedling *Height* was highly significant ($P < 0.001$, $R^2 = 0.23$) driven by a strong positive association with *Seed Weight* (estimate = 955.35, $P = 0.004$) but also by positive associations with the Edge region (estimate = 188.79, $P = 0.04$), genetic load (estimate = 102.5, $P = 0.03$) and several interaction effects between genetic parameters and *Region*. When looking at these interactions, we observed that in comparison with the Core region (i.e., the intercept), *Height* of families from the Margin region was more negatively affected by population homozygosity ($P = 0.004$) and more positively affected by genetic diversity ($P = 0.004$). Additionally, *Height* of families from the Edge region was more negatively affected by genetic load ($P = 0.004$). Lastly, 25% of variation in *Early-Life Fitness*

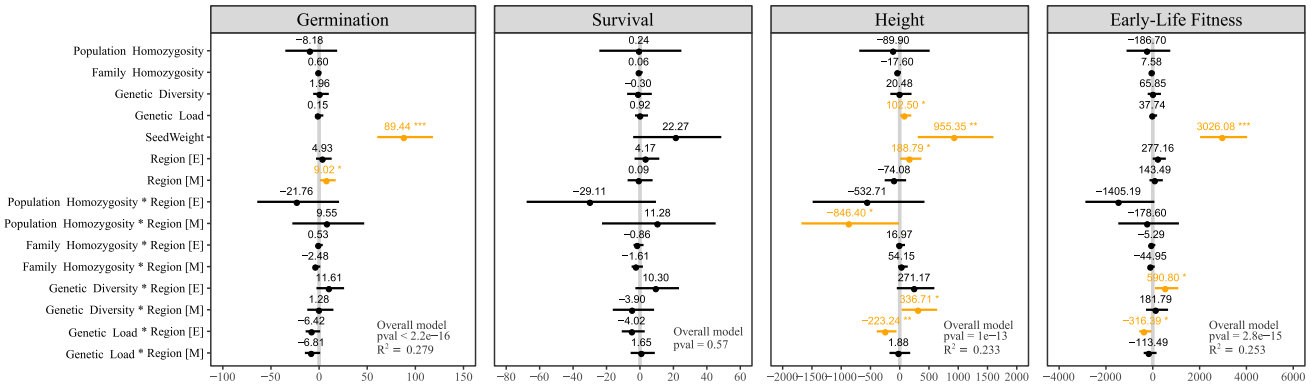

Fig. 5 Summary statistics of multivariate linear models performed for each phenotypic trait. The overall model *P*-values, and adjusted R^2 values are shown together with the coefficient estimates for each variable included in the models. Values are colored in orange when significant and include asterisks to note level of significance (***$P < 0.001$, **$0.001 \leq P \leq 0.01$, and *$0.01 \leq P \leq 0.05$)

was explained by the full model ($P < 0.001$), which was strongly influenced by *Seed Weight* (estimate = 3026.08, $P = 8.5e{-}09$), but also showed significant interactions effects with *Region.* Specifically, *Early-Life Fitness* in the Edge region exhibited stronger positive associations with genetic diversity ($P = 0.02$) and stronger negative associations with genetic load ($P = 0.018$) (Fig. 6). Population homozygosity also showed a trend towards a more negative relationship with *Early-Life Fitness* in the Edge region (Fig. 6), although this was on the boundary of statistical significance (Fig. 5).

DIscussion

Variation in fitness traits and genetic parameters across the range

By growing thousands of seedlings from 65 different sampling locations in a controlled environment, and pairing measurements of early-life traits with genomic sequences from their maternal trees, we uncovered key genetic drivers of variation in early-life fitness traits across red spruce's range. This variation, associated to some extent with geographic patterns, gives important information on red spruce's capacity to maintain and recover from logging, fire and acid rains that greatly affected the species in the last two centuries (Mathias and Thomas 2018; Rentch et al. 2016).

We originally posed two main hypotheses that could explain variation in early-life fitness traits in association with genetic parameters across a species range. First, the edges of a species' geographic distribution are expected to present environmental conditions that are less optimal than the center of the range for the survival and growth of individuals (a.k.a., the central-marginal hypothesis) (Sagarin and Gaines 2002). This gradient in suitability is predicted to lead to smaller population sizes at the range periphery, where selection is weakened and genetic drift strengthened (Kimura, 1955, 1957; Wright, 1931). As a result, genetic diversity is expected to be lower in those populations (Eckert et al. 2008) and deleterious alleles would accumulate more rapidly, increasing genetic load and decreasing fitness in comparison with the center of the range (Peischl et al.

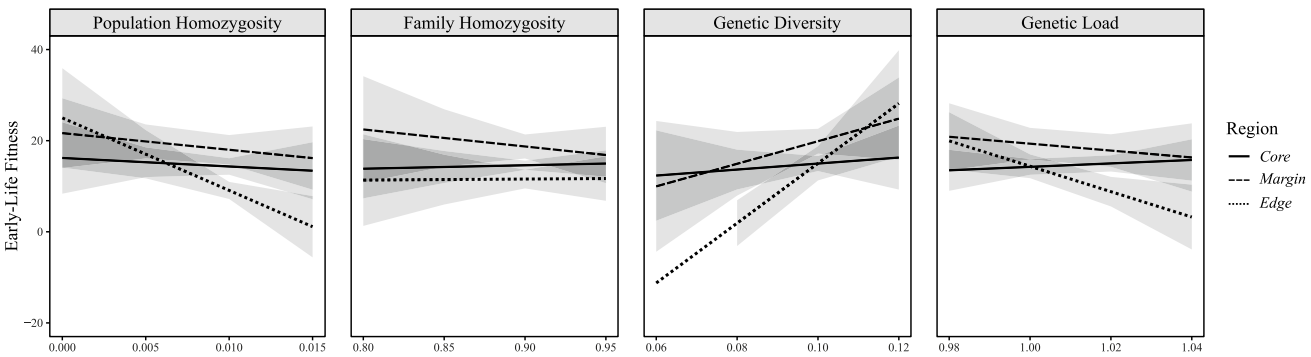

Fig. 6 Marginal effects estimated from the multivariate linear model for *Early-Life Fitness*, plotted for each genetic parameter and including interactions between genetic parameters and *Regions*

2015; Willi et al. 2018). Second, past range shifts may have strongly influenced the distribution of genetic diversity and genetic load across the species' range. The post-glacial recolonization most species have undergone since the LGM often resulted in the establishment of "leading" and "trailing (rear) edge" populations (Hampe and Petit 2005; Hewitt 2000). Due to successive bottlenecks during expansion, leading edge populations are expected to show low genetic diversity and high levels of genetic load, whereas trailing edge populations are expected to exhibit low genetic diversity due to fragmentation and population decline, but should not present high levels of genetic load (González-Martínez et al. 2017; Hampe and Petit 2005).

Based on the observed geographic structure of genetic variation, none of those patterns are clearly visible across red spruce range (Fig. 4). In accordance with the central-marginal hypothesis, the highly fragmented trailing edge in the south showed a decrease in genetic diversity, while sampling locations in the northern core of the range showed some of the highest level of diversity, even though they probably are the most recently colonized locations. Interestingly, the degree of genetic load, which should follow an opposite trend (highest in demographically more marginal localities), seemed in fact strongly associated with genetic diversity (Fig. 4). In fact, we observed some of the highest levels of genetic load in the central part of the range (Margin region), which neither the leading/trailing edge hypothesis nor the central-marginal hypothesis would easily explain. Some of these complexities likely reflect multiple processes at work over different temporal or spatial scales. For example, past hybridization in parts of red spruce's range may contribute to the genetic patterns we observed. Red spruce is known to hybridize with its close congener black spruce (*P. mariana*), and hybrids have been reported in the far northeast of its range but also in Pennsylvania (de Lafontaine and Bousquet 2017; De Lafontaine et al. 2015), which coincides with the Margin region in our exome data. Thus, one hypothesis for the high diversity observed in the Margin is that it reflects past introgression with black spruce, which might also influence our estimate of genetic load by introducing nonsynonymous polymorphisms from black spruce as a consequence of introgression.

Another source of complexity in the genetic patterns we observed may involve more recent demographic changes. It is well known that genetic diversity in a population at demographic equilibrium should closely approximate N_e (effective population size)(Ellegren and Galtier 2016). However, recent strong bottlenecks over time scales of a few generations can actually temporarily increase genetic diversity, when estimated as expected heterozygosity at polymorphic sites, because loss of rare alleles leads to allele frequencies that are more even (i.e., less skewed; Luikart et al. 1998). Lastly, we note that the high level of population homozygosity in the northeast primarily reflects the influence of the NBTIC locality (Fig. 4), which may be caused by a Wahlund effect of pooling diversity from several different genetically differentiated sources into a single sample. Altogether, our results suggest a complex demographic-genetic history for red spruce, whereby populations have probably undergone a south to north post-glacial expansion but have also found refuge at higher elevations in the Southern and Central Appalachians, creating islands of isolated forest that later faced a dramatic decline (~90%) due to intensive logging, fires and acid rains in the last two centuries (Mathias and Thomas 2018; Rentch et al. 2016; Rentch and Schuler 2009). Hybridization with black spruce likely contributes as well, either historically, more recently, or both (Bashalkhanov et al. 2013). Thus, the patterns of genetic and phenotypic variation in current populations are intrinsically associated with this complex demographic history, obviating a simple linking of observations to the straightforward theories described above.

Factors influencing early-life fitness

Our results confirm the existence of significant variation among the three genetically determined regions in shaping the early-life fitness of red spruce seedlings. These regional groups are derived from the main genetic clusters structuring the genetic background of red spruce populations (Capblancq et al. 2020), and highlight the importance of evolutionary history in shaping phenotypic variation and fitness in natural populations (Miller et al. 2019). For most of our measured fitness-related traits, sampled localities in the Margin returned the highest scores, localities in the Core showed intermediate values, and localities in the Edge the lowest. Those results suggest that the specific historical, geographical, and perhaps environmental features associated with these three different regions have important impacts on the early-life fitness of spruce seedlings.

We observed a very low germination rate for some of the Edge sampling locations (e.g., 'XWS' and 'XFS' < 10%), which could have a negative impact on forest regeneration in this fragmented part of the range. Eight of the 22 Edge sampling locations had a germination rate less than 25%, even when the seeds were germinated in a suitable controlled environment. It could be explained, to some extent, by particularly poor local conditions for seed production in the southern Appalachians the year the seeds were collected (2017), but some southern sampling locations show germination rates superior to 60%, similar to most sampling locations in the Margin or the Core regions. While it would be necessary to confirm this trend by extending the experiment to multiple years and seed cohorts (Zhang et al. 2017), it nonetheless suggests a poor recruitment capacity in those southern forests, which are already the most threatened ones

(Nowacki et al. 2010). On a more positive note, once the seeds have germinated, the survival rate was very high for most families, with more than 80% of the seedlings surviving the first 12 weeks. When looking at seedling height after 12 weeks of growth, sampling locations from the Edge region again exhibited lower growth efficiency than locations in the Core or the Margin regions. The combination of poor germination rates and smaller seedling heights led to a pattern of lower early-life fitness in the Edge region compared to the Margin and Core regions. Interestingly, it seemed that the lowest fitness values within the Edge were found in the Central Appalachians (WV and western MD) and not further south in NC and TN. It is not currently known what historic processes (ancient or more recent) led to Central Appalachian red spruce exhibiting lower seedling fitness, but it may reflect the dramatic losses of spruce forests in this region to logging and fire in the late 19th and early 20th centuries, and later on to atmospheric pollution as well (Adams and Stephenson 1989; Mathias and Thomas 2018; Rentch et al. 2016).

We also found that average seed weight had a significant effect on germination, as well as on height and early-life fitness (Fig. 4). Considering the wide range of source environments represented in our seed collection (Table S1), we used average seed weight as a covariate when analyzing each trait under the assumption that seed weight is a good proxy of maternal resource allocation and thus reflects environmental conditions during the growing season (Roach and Wulff, 1987). In plants, early-life traits are the most susceptible to environmental maternal effects, and seed mass in particular has been shown to vary depending on environmental conditions due to the maternal inheritance of seed tissue that is distinct from parental genetic factors (Singh et al. 2017). For example, the climatic harshness (e.g., drought or cold) of the environment can directly affect the quantity and/or quality of tissue produced per seed in plants (Baker 1972), which can in turn directly influence germination proportion and growth of the offspring (Wahid and Bounoua 2013). Not taking this effect into account could thus lead to erroneous conclusions when analyzing the relationship between seedling phenotypes and their genetic characteristics (Wolf and Wade 2009). Our results confirmed that average seed weight explained a significant part of the variation in red spruce germination, height and early-life fitness traits.

Even after accounting for regional genetic backgrounds and maternal effects using region and average seed weight as co-variables in the models, we identified a significant positive influence of genetic diversity on seedling height and overall fitness, a significant negative association between those traits and genetic load, and a negative association between population homozygosity and seedling height. If not always significant, the trends of co-variation between genetic parameters and early-life fitness were consistent with expectations (Fig. 5). Population homozygosity and genetic load were negatively associated with early-life fitness in all regions except for the Core, where genetic load returned a slightly positive association. In contrast, genetic diversity was positively associated with early-life fitness in all three regions even if more strongly in the Edge and Margin regions. This confirms for red spruce the pattern largely found for other species in the conservation literature—namely, that genetic diversity is usually positively correlated with fitness traits (Reed and Frankham 2003). Such a trend was also identified in eastern Canadian red spruce localities for which genetic diversity had been identified as a positive factor for growth measured in natural stands (Mosseler et al. 2003). Interestingly, the same study suggested that the presence of rare deleterious alleles was negatively impacting growth rate, which is confirmed by the negative influence of genetic load on height and fitness in our study (Figs. 5 and 6). Finally, population homozygosity was negatively associated with seedling fitness in our experiment, suggesting a deleterious effect of inbreeding (Angeloni et al. 2011; Mosseler et al. 2000; Reed and Frankham 2003). Inbreeding depression has been frequently identified as a major factor affecting survival in the wild (Keller & Waller 2002), especially for small and isolated populations (Naish et al. 2013). Some of our sampled red spruce localities indeed represent very small and fragmented forest stands where the species would likely suffer from inbreeding depression, which in return could also accelerate genetic diversity erosion and deleterious mutation accumulation in the future (Lynch et al. 1995).

Early-life fitness, genomics and conservation

There is growing attention given to genetic diversity and its role in adaptation ("evolvability") in conservation or management plans (Hendricks et al. 2018; Khan et al. 2016; Supple and Shapiro 2018). Maximizing genetic diversity is often proposed to avoid the deleterious effects of inbreeding and provides the population the raw genetic variability required to adapt to changing environmental conditions (Keller and Waller 2002). The results of this study confirm the importance of genetic diversity for red spruce seedling fitness. These findings are then directly relevant to restoration ecologists and resource managers aimed at restoring red spruce forests. Lowering the number of deleterious alleles (genetic load) and maximizing genetic diversity is especially important in the Central and Southern Appalachians (Edge region) where the highly fragmented populations appear more sensitive to variation in these parameters. We see a great opportunity for integration of current knowledge on genetic diversity and genetic load into a seed selection model that identifies optimal combinations of potential seed sources so as to

maximize levels of genetic diversity at planting sites while minimizing the number of deleterious alleles. We believe that integrating genomic tools to inform restoration programs could enhance the survival rate of red spruce individuals in the reforested areas and help decrease levels of inbreeding and the erosion of genetic diversity in the highly fragmented southern parts of the species' range.

Finally, effective population size and degree of gene flow among populations are known to greatly influence fitness by mediating levels of inbreeding and genetic diversity (Naish et al. 2013), but some of the variation in seedlings traits could also reflect local adaptation to divergent environments associated with the variety of environments populations experience across the species range (Butnor et al. 2019; Marks 2007). Depending on the spatial scale of local adaptation that may occur in red spruce, selecting seed sources with the appropriate adaptive alleles may also be an important consideration. This will require a better understanding of adaptive genetic variation across the species range and the spatial scale at which those differences manifest, which would allow for more refined predictions of the optimal genetic composition under future climates for different reforestation sites (Capblancq et al. 2020). Future studies aimed at combining our present results on population genetic diversity, inbreeding, and genetic load, with knowledge about the influence of local adaptation will help produce an unprecedented approach to inform management plans for the conservation of red spruce, with high value in the current context of climate change.

Supplementary Information The online version of this article (https://doi.org/10.1007/s10592-021-01378-7).

Acknowledgments We appreciate the help of partners who shared access to their seed collections: Barbara Crane, Matt Fitzpatrick, Robert Jetton, John Major, John Malcom, Dave Nelson, and Brittany Verrico. We also thank the members of the Keller lab and three anonymous reviewers for helpful comments on the manuscript. This project was supported by awards from the National Science Foundation (1656099) and USDA-HATCH (1006810) to SK, and a UVM SURF award to HM.

Author contributions Stephen Keller conceived the study. SK and John Butnor collected the samples. Helena Munson germinated and grew the seedlings with help from SK and JB. HM measured fitness traits and Thibaut Capblancq analyzed genomic data. TC performed the statistical analyses with the help of HM. TC and HM wrote a first draft of the manuscript and all authors provided critical feedbacks.

Funding This project was supported by awards from the National Science Foundation (1656099) and USDA-HATCH (1006810) to Stephen Keller, and a UVM SURF award to Helena Munson.

Data availability A data table including the different genomic and trait values for each individual and family is available on Github (https://github.com/stephenrkeller/Prubens_Capblancq_etal_2020_ConsGen). together with the script used to perform the different analyses

Code availability The script used to perform the different analyses is available on Github (https://github.com/stephenrkeller/Prubens_Capblancq_etal_2020_ConsGen).

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

Adams HS, Stephenson SL (1989) Old growth red spruce communities in the mid-Appalachians. Vegetatio 85:45–56

Agrawal AF, Whitlock MC (2012) Mutation load: the fitness of individuals in populations where deleterious alleles are abundant. Annu Rev Ecol Evol Syst 43(1):115–135. https://doi.org/10.1146/annurev-ecolsys-110411-160257

Allendorf FW, Luikart GH (2006) Conservation and the genetics of populations (Blackwell). Blackwell, London

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. Nat Rev Genet 11(10):697–709. https://doi.org/10.1038/nrg2844

Anderson JT (2015) Plant fitness in a rapidly changing world. New Phytol 210(1):81–87. https://doi.org/10.1111/nph.13693

Angeloni F, Ouborg NJ, Leimu R (2011) Meta-analysis on the association of population size and life history with inbreeding depression in plants. Biol Cons 144(1):35–43. https://doi.org/10.1016/j.biocon.2010.08.016

Baker HG (1972) Seed weight in relation to environmental conditions in California. Ecology 53(6):997–1010

Bashalkhanov S, Eckert AJ, Rajora OP (2013) Genetic signatures of natural selection in response to air pollution in red spruce (*Picea rubens*, Pinaceae). Mol Ecol 22(23):5877–5889. https://doi.org/10.1111/mec.12546

Birol I et al. (2013) Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics*, *29*(12), 1492–1497. https://doi.org/10.1093/bioinformatics/btt178

Boisvert-Marsh L, Périé C, De Blois S (2014) Shifting with climate? Evidence for recent changes in tree species distribution at high latitudes. Ecosphere 5(7):1–33. https://doi.org/10.1890/ES14-00111.1

Butnor JR, Verrico BM, Johnsen KH, Maier CA, Vankus V, Keller SR (2019) Phenotypic variation in climate-associated traits of red spruce (*Picea rubens* Sarg) along elevation gradients in the Southern Appalachian Mountains. Castanea 84(2):128. https://doi.org/10.2179/0008-7475.84.2.128

Capblancq T et al. (2020) Whole-exome sequencing reveals a long-term decline in effective population size of red spruce (*Picea rubens*). Evolut Appl (2), 1–16. https://doi.org/10.1111/eva.12985

Capblancq T, Fitzpatrick MC, Bay RA, Exposito-Alonso M, Keller SR (2020b) Genomic prediction of (Mal)adaptation across current and future climatic landscapes. Annu Rev Ecol Evol Syst 51(1):245–271. https://doi.org/10.1146/annurev-ecolsys-020720-042553

Charlesworth D, Willis JH (2009) The genetics of inbreeding depression. Nat Rev Genet 10(11):783–796. https://doi.org/10.1038/nrg2664

Cingolani P et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly, 6(2), 80–92. https://doi.org/10.4161/fly.19695

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. Science 292(5517):673–679. https://doi.org/10.1126/science.292.5517.673

de Lafontaine G, Bousquet J (2017) Asymmetry matters: a genomic assessment of directional biases in gene flow between hybridizing spruces. Ecol Evol 7(11):3883–3893. https://doi.org/10.1002/ece3.2682

De Lafontaine G, Prunier J, Gérardi S, Bousquet J (2015) Tracking the progression of speciation: Variable patterns of introgression across the genome provide insights on the species delimitation between progenitor-derivative spruces (*Picea mariana* × *P. rubens*). Mol Ecol 24(20):5229–5247. https://doi.org/10.1111/mec.13377

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. Mol Ecol 17(5):1170–1188. https://doi.org/10.1111/j.1365-294X.2007.03659.x

Ellegren H, Galtier N (2016) Determinants of genetic diversity. Nat Publish Group 17(7):422–433. https://doi.org/10.1038/nrg.2016.58

Frankham R, Ballou JD, Briscoe DA (2010) Introduction to conservation genetics. Cambridge University Press, Cambridge

González-Martínez SC, Ridout K, Pannell JR (2017) Range expansion compromises adaptive evolution in an outcrossing plant. Curr Biol 27(16):2544–2551.e4. https://doi.org/10.1016/J.CUB.2017.07.007

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. Ecol Lett 8(5):461–467. https://doi.org/10.1111/j.1461-0248.2005.00739.x

Hampe A, Pemonge MH, Petit RJ (2013) Efficient mitigation of founder effects during the establishment of a leadingedge oak population. Proc R Soc B. https://doi.org/10.1098/rspb.2013.1070

Hendricks S et al. (2018) Recent advances in conservation and population genomics data analysis. Evol Appl 11(8), 1197–1211. https://doi.org/10.1111/eva.12659

Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405(6789):907–913. https://doi.org/10.1038/35016000

Kardos M, Taylor HR, Ellegren H, Luikart G, Allendorf FW (2016) Genomics advances the study of inbreeding depression in the wild. Evol Appl 9(10):1205–1218. https://doi.org/10.1111/eva.12414

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17(5):230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Khan S, Nabi G, Ullah MW, Yousaf M, Manan S, Siddique R, Hou H (2016) Overview on the role of advance genomics in conservation biology of endangered species. Int J Genomics. https://doi.org/10.1155/2016/3460416

Kimura M (1955) Solution of a process of random genetic drift with a continuous model. Proc Natl Acad Sci 41(2):144–150

Kimura M (1957) Some problems of stochastic processes in genetics. Ann Math Stat 28(4):882–901. https://doi.org/10.1214/aoms/1177706791

Koo KA, Madden M, Patten BC (2014a) Projection of red spruce (*Picea rubens* Sargent) habitat suitability and distribution in the Southern Appalachian Mountains, USA. Ecol Model 293:91–101. https://doi.org/10.1016/j.ecolmodel.2014.06.005

Koo KA, Patten BC, Teskey RO, Creed IF (2014b) Climate change effects on red spruce decline mitigated by reduction in air pollution within its shrinking habitat range. Ecol Model 293:81–90. https://doi.org/10.1016/j.ecolmodel.2014.07.017

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. BMC Bioinform 15:356–368. https://doi.org/10.1007/978-1-4614-1174-1_8

Lande R (1995) Mutation and conservation. Conservation Biology. Conservation Biology 9(4):782–791

Latta RG (2008) Conservation genetics as applied evolution: from genetic pattern to evolutionary process. Evol Appl 1(1):84–94. https://doi.org/10.1111/j.1752-4571.2007.00008.x

Le HD, Smith C, Herbohn J, Harrison S (2012) More than just trees: assessing reforestation success in tropical developing countries. J Rural Stud. https://doi.org/10.1016/j.jrurstud.2011.07.006

Leimu R, Vergeer P, Angeloni F, Ouborg NJ (2010) Habitat fragmentation, climate change, and inbreeding in plants. Ann N Y Acad Sci 1195:84–98. https://doi.org/10.1111/j.1749-6632.2010.05450.x

Levin DA (2011) Mating system shifts on the trailing edge. Ann Bot 109:613–620

Li H (2011) Improving SNP discovery by base alignment quality. Bioinformatics 27(8):1157–1158. https://doi.org/10.1093/bioinformatics/btr076

Liu Y et al. (2013) Environmental factors influence both abundance and genetic diversity in a widespread bird species. Ecol Evol, 3(14), 4683–4695. https://doi.org/10.1002/ece3.856

Luikart G, Allendorf FW, Cornuet JM, Sherwin WB (1998) Distortion of allele frequency distributions provides a test for recent population bottlenecks. J Hered 89(3):238–247. https://doi.org/10.1093/jhered/89.3.238

Lynch M, Conery IJ, Burger R (1995) Mutation accumulation and the extinction of small populations. Am Nat 146(4):489–518

Major JE, Mosseler A, Johnsen KH, Campbell M, Malcolm J (2015) Growth and allocation of *Picea rubens*, *Picea mariana*, and their hybrids under ambient and elevated CO 2. Can J For Res 45(7):877–887. https://doi.org/10.1139/cjfr-2014-0525

Markert JA et al. (2010) Population genetic diversity and fitness in multiple environments. BMC Evol Biol, 10(1), 5–9. https://doi.org/10.1186/1471-2148-10-205

Marks CO (2007) The causes of variation in tree seedling traits: the roles of environmental selection versus chance. Evolution 61(2):455–469. https://doi.org/10.1111/j.1742-4658.2007.00021.x

Mathias JM, Thomas RB (2018) Disentangling the effects of acidic air pollution, atmospheric CO2, and climate change on recent growth of red spruce trees in the Central Appalachian Mountains. Glob Change Biol 24(9):3938–3953. https://doi.org/10.1111/gcb.14273

Miller ML, Kronenberger JA, Fitzpatrick SW (2019) Recent evolutionary history predicts population but not ecosystem-level patterns. Ecol Evolut 9(24):14442–14452. https://doi.org/10.1002/ece3.5879

Mosseler A et al. (2000) Indicators of population viability in red spruce, *Picea rubens*. I. Reproductive traits and fecundity. Can J Bot *78*(7), 928–940. https://doi.org/10.1139/b00-065

Mosseler A, Major JE, Rajora OP (2003) Old-growth red spruce forests as reservoirs of genetic diversity and reproductive fitness. Theoret Appl Genet 106(5):931–937. https://doi.org/10.1007/s00122-002-1156-1

Naish KA, Seamons TR, Dauer MB, Hauser L, Quinn TP (2013) Relationship between effective population size, inbreeding and adult fitness-related traits in a steelhead (Oncorhynchus mykiss) population released in the wild. Mol Ecol 22(5):1295–1309. https://doi.org/10.1111/mec.12185

Nielsen R, Paul JS, Albrechtsen A, Song YS (2011) Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet 12(6):443–451. https://doi.org/10.1038/nrg2986

Nowacki G, Carr R, Van Dyck M (2010) The current status of Red Spruce in the Eastern United States : distribution, population trends, and environmental drivers. Proceedings from the Conference on the Ecology and Management of High-Elevation Forests in the Central and Southern Appalachian Mountains, pp 140–162

Nystedt B et al. (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, *497*(7451), 579–584. https://doi.org/10.1038/nature12211

Peischl S, Dupanloup I, Kirkpatrick M, Excoffier L (2013) On the accumulation of deleterious mutations during range expansions. Mol Ecol 22(24):5972–5982. https://doi.org/10.1111/mec.12524

Peischl S, Kirkpatrick M, Excoffier L (2015) Expansion load and the evolutionary dynamics of a species range. Am Nat 185(4):E81–E93. https://doi.org/10.1086/680220

Pluess AR (2011) Pursuing glacier retreat: genetic structure of a rapidly expanding Larix decidua population. Mol Ecol 20(3):473–485. https://doi.org/10.1111/j.1365-294X.2010.04972.x

Ratnam W et al. (2014) Genetic effects of forest management practices: global synthesis and perspectives. For Ecol Manage 333, 52–65. https://doi.org/10.1016/j.foreco.2014.06.008

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17(1):230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Rentch JS, Schuler TS (2009) Pennsylvania boreal conifer forests and their bird communities: past, present, and potential. Proceedings from the Conference on the Ecology and Management of High-Elevation Forests in the Central and Southern Appalachian Mountains, 48–73

Rentch JS, Ford WM, Schuler TS, Palmer J (2016) Release of suppressed red spruce using canopy gap creation—ecological restoration in the central appalachians. Nat Areas J 36(1):29. https://doi.org/10.3375/043.036.0108

Ribeiro PJ, Diggle PJ (2016) geoR: A package for geostatistical data analysis using the R software. *R News*, 1(6), 15–18. Retrieved from http://www.leg.ufpr.br/geoR/

Roach DA, Wulff RD (1987) Maternal effects in plants. Ann Rev Ecol System. 18(1987):209–235. https://doi.org/10.1146/annurev.es.18.110187.001233

Sagarin RD, Gaines SD (2002) The "abundant centre" distribution: to what extent is it a biogeographical rule? Ecol Lett 5(1):137–147. https://doi.org/10.1046/j.1461-0248.2002.00297.x

Schauffler M, Jacobson GL (2002) Persistence of coastal spruce refugia during the Holocene in northern New England, USA, detected by stand-scale pollen stratigraphies. J Ecol 90(2):235–250. https://doi.org/10.1046/j.1365-2745.2001.00656.x

Singh J, Michelangeli JAC, Gezan SA, Lee H, Vallejos CE (2017) Maternal effects on seed and seedling phenotypes in reciprocal F1hybrids of the common bean (*Phaseolus vulgaris* L.). Front Plant Sci 8(1):1–13. https://doi.org/10.3389/fpls.2017.00042

Supple MA, Shapiro B (2018) Conservation of biodiversity in the genomics era. Genome Biol 19(1):1–12. https://doi.org/10.1186/s13059-018-1520-3

Wahid N, Bounoua L (2013) The relationship between seed weight, germination and biochemical reserves of maritime pine (Pinus pinaster Ait.) in Morocco. New Forests 44(3):385–397. https://doi.org/10.1007/s11056-012-9348-2

Wang J, Caballero A, Hill WG (1998) The effect of linkage disequilibrium and deviation from Hardy-Weinberg proportions on the changes in genetic variance with bottlenecking. Heredity 81(2):174–186. https://doi.org/10.1038/sj.hdy.6883900

Willi Y, Fracassetti M, Zoller S, Van Buskirk J (2018) Accumulation of mutational load at the edges of a species range. Mol Biol Evol 35(4):781–791. https://doi.org/10.1093/molbev/msy003

Wolf JB, Wade MJ (2009) What are maternal effects (and what are they not)? Philos Trans R Soc B 364(1520):1107–1115. https://doi.org/10.1098/rstb.2008.0238

Wright S (1931) Evolution in mendelian populations. Genetics 16:97–159. https://doi.org/10.4161/hv.21408

Zhang R, Baskin JM, Baskin CC, Mo Q, Chen L, Hu X, Wang Y (2017) Effect of population, collection year, after-ripening and incubation condition on seed germination of Stipa bungeana. Sci Rep 7(1):1–11. https://doi.org/10.1038/s41598-017-14267-2