bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Heritable and climatic sources of variation in juvenile tree growth in an Austrian common garden experiment of Central European Norway spruce populations

Laura Morales[1,2] and Kelly Swarts[1,3]*

[1]Gregor Mendel Institute, Austrian Academy of Sciences, Vienna, Austria

[2]Institute of Biotechnology in Plant Production, Department of Agrobiotechnology, University of Natural Resources and Life Sciences Vienna, Tulln an der Donau, Austria

[3]Max Perutz Labs, University of Vienna, Vienna, Austria

*Kelly Swarts: corresponding, kelly.swarts@gmi.oeaw.ac.at, kelly.swarts@univie.ac.at

Laura Morales: laura.morales@gmi.oeaw.ac.at

Abstract

We leveraged publicly available data on juvenile tree height of 299 Central European Norway spruce populations grown in a common garden experiment across 24 diverse trial locations in Austria and weather data from the trial locations and population provenances to parse the heritable and climatic components of juvenile tree height variation. Principal component analysis of geospatial and weather variables demonstrated high interannual variation among trial environments, largely driven by differences in precipitation, and separation of population provenances based on altitude, temperature, and snowfall. Tree height was highly heritable and modeling the covariance between populations and trial environments based on climatic data led to more stable estimation of heritability and population × environment variance. Climatic similarity among population provenances was highly predictive of population × environment estimates for tree height.

Key words

Norway spruce; tree height; prediction; climate; genotype-by-environment

Introduction

Under rapid anthropogenic climate change, ecosystem dominant conifer species, such as Norway spruce (*Picea abies*), are facing environmental conditions to which they have not been adapted (Lindner et al. 2010; Seidl et al. 2017). Although Norway spruce has broad genetic diversity and geographic adaptation, long generation times and the confounded effects of population structure and local adaptation make breeding populations for new environments challenging (Leslie et al. 2012; Yeaman et al. 2016; Wang et al. 2020). Common garden experiments can be used to evaluate a select number of diverse individuals or populations in different environments (Oleksyn et al. 1998; Kapeller et al. 2012; Aitken and Bemmels 2016). Assisted migration can then be applied, where well-adapted individuals are introduced into

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

locations with compatible (current or forecasted) climates (Williams and Dumroese 2013; Aitken and Bemmels 2016).

Precise estimation of genetic and environmental effects on tree growth phenotypes is critical for assisted migration and breeding program selections (Williams and Dumroese 2013; Aitken and Bemmels 2016). In addition to proper experimental design, germplasm choice, and phenotyping methods, statistical approaches can also increase the genetic signal and prediction accuracy of phenotypes (Heslot et al. 2015; Crossa et al. 2017; Isidro y Sánchez and Akdemir 2021). For example, leveraging genetic marker, geospatial, and/or climatic data to model relationships between individuals, trial locations, and genotype × environment interactions is a common approach in agricultural and forest genetics and breeding (Heslot et al. 2014; Crossa et al. 2016; Rodríguez-Álvarez et al. 2018; Avanzi et al. 2019; Bustos-Korts et al. 2019).

In this study, we combined publicly available datasets on (a) juvenile tree height from a common garden experiment of Norway spruce, in which 299 Central European populations were evaluated across five years in 24 locations in Austria and (b) historical monthly precipitation and temperature measurements from the common garden trial locations and population provenances (Efthymiadis et al. 2006; Hiebl et al. 2009; Chimani et al. 2011, 2013; Kapeller et al. 2012, 2017). Previous studies have used this dataset to understand varying selection and intraspecific variation in climate response from an ecological perspective (Kapeller et al. 2012, 2017). Here, we employed agricultural statistical methods to parse and to predict the heritable (genetic) and environmental components of phenotypic variance in tree height.

Materials and methods

Common garden experiment

We used a publicly available dataset to investigate population and environmental effects on juvenile tree growth in Norway spruce (*Picea abies*) (Nather and Holzer 1979; Schulze 1985; Kapeller et al. 2012, 2016, 2017). The plant material was derived from 299 Norway spruce populations collected across Central Europe during commercial seed harvests in 1971. Seeds from each population were germinated and seedlings were then transplanted into a nursery in Austria in 1973. In 1978, a total of 65,534 five-year-old trees were transplanted into 24 trial locations across Austria in a balanced incomplete block design. On average, each population was grown at two locations, 27 populations were grown in each location, and 102 trees (replications) from each population were split into three randomized blocks within each location. The height (cm) of each tree was recorded at the ages of 7-10 and 15 years in 1980-1983 and 1988, respectively, resulting in a total of 300,310 observations. The previously published dataset also included the altitude, latitude, and longitude for all trial locations and for 278 population provenances.

Weather data

We accessed monthly precipitation (total, liquid, solid) and temperature data from the publicly available Historical Instrumental Climatological Surface Time Series of the Greater Alpine

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Region (HISTALP) resource (Efthymiadis et al. 2006; Hiebl et al. 2009; Chimani et al. 2011, 2013). The HISTALP dataset includes monthly temperature and precipitation grids from 1780-2009 and 1801-2003, respectively, at 5 min x 5 min resolution ranging from 4-19°E and 43-46°N. We extracted data from the original Network Common Data Form (NetCDF) formatted HISTALP files using the "netcdf4" package in R (R Core Team 2020; Pierce 2021).

Climatic analysis

All altitude, latitude, longitude, and HISTALP weather variables used in the analyses described in this subsection were centered and scaled using the "scale" function in R (R Core Team 2020). We conducted principal component analysis (PCA) using the "FactoMineR" package in R (Lê et al. 2008; R Core Team 2020). To investigate spatiotemporal variation across the common garden experiment, we extracted monthly temperature and precipitation data from the grids closest to the trial locations in 1980-1983 and 1988 from the HISTALP dataset. We then conducted PCA using the altitude, latitude, longitude and monthly weather data from each trial location in each year, henceforth referred to as trial environments (PCA_{Env}). To explore climatic differences among population provenances, we extracted monthly temperature and precipitation data from the grids closest to the population provenances from 1801 (earliest year with both temperature and precipitation data in HISTALP) to 1970 (year before seed harvests). Weather data from this time period should reflect the climates experienced by the parental populations of the trees evaluated in the common garden experiment. We then conducted a PCA using the altitude, latitude, longitude, and monthly weather data from each population provenance (PCA_{Prov}).

We used the altitude, latitude, longitude, and HISTALP weather data described above to model the relationships among population provenances and environments (trial location-years). We first calculated Euclidian distance matrices using said geospatial and weather variables for all pairs of population provenances ($\boldsymbol{D}_{Prov}$) and trial environments ($\boldsymbol{D}_{Env}$) with the "dist" function in R (R Core Team 2020). We then transformed the distance matrices into similarity matrices as $\boldsymbol{S}_{Prov} = \frac{\max(\boldsymbol{D}_{Prov}) - \boldsymbol{D}_{Prov}}{\max(\boldsymbol{D}_{Prov})}$ and $\boldsymbol{S}_{Env} = \frac{\max(\boldsymbol{D}_{Env}) - \boldsymbol{D}_{Env}}{\max(\boldsymbol{D}_{Env})}$ for further use in phenotypic modeling.

Phenotypic analysis using climatic data

Due to the large number of observations (N=300,310) in the common garden experiment, it was not computationally feasible to specify complex variance-covariance matrices for the entire dataset in one model. As such, we used a two-stage approach common in multi-environment plant breeding and genetics studies (Piepho et al. 2012, 2020), with mixed linear models fit with the "breedR" package in R (Muñoz and Sanchez 2020; R Core Team 2020). In the first stage, we fit the following mixed model within each environment (location-year):

$$Height_{ij} = \mu + Pop_i + Block_j + \varepsilon_{ij} \tag{1}$$

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

where $Height_{ij}$ is the tree height response, μ is the overall mean, Pop_i is the fixed effect of population i, $Block_j$ is the random effect of block j, and ε_{ij} is the random error. The random effects were independent and identically normally distributed as $Block_j \sim N(0, \boldsymbol{I}\sigma_b^2)$ and $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$. We then extracted the fitted values ($\hat{y}_{ij}$) for each population i within each block j from each within-environment model (1) as

$$\hat{y}_{ij} = \mu + popBLUE_i + blockBLUP_j \tag{2}$$

Where μ is the intercept, $popBLUE_i$ is the best linear unbiased estimate (BLUE) of population i, and $blockBLUP_j$ is the best linear unbiased predictor (BLUP) of block j. We combined the fitted values across environments, resulting in a reduced and refined set of phenotypes for further analysis (N = 7,314). Only fitted values for populations with climatic information were included in the second stage of the analysis (278/299 populations).

In the second stage, we fit a naïve across-year mixed model to understand the environmental and genetic (population-level) effects on juvenile tree growth as:

$$\hat{y}_{ij} = \mu + Year + Pop_i + Env_j + PopEnv_{ij} + \varepsilon_{ij} \tag{3}$$

where $\hat{y}_{ij}$ are the fitted values from each within-environment model (2), μ is the overall mean, $Year$ is the continuous fixed effect of year, Pop_i is the random effect of population i, Env_j is the random effect of trial environment j, $PopEnv_{ij}$ is the random effect of the interaction between population i and trial environment j, and ε_{ij} is the random error. The random effects were assumed to be independent and identically normally distributed, where $Pop_i \sim N(0, \boldsymbol{I}\sigma_p^2)$, $Env_j \sim N(0, \boldsymbol{I}\sigma_e^2)$, $PopEnv_{ij} \sim N(0, \boldsymbol{I}\sigma_{pe}^2)$, and $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$.

We also fit an across-year mixed model with form similar to model (3) using historical weather data to model similarity among trial environments and populations provenances as:

$$\hat{y}_{ij} = \mu + Year + Prov_i + Env_j + ProvEnv_{ij} + \varepsilon_{ij} \tag{4}$$

where $\hat{y}_{ij}$ are the fitted values from each within-environment model (2), μ is the overall mean, $Year$ is the continuous fixed effect of year, $Prov_i$ is the random effect of population provenance i, Env_j is the random effect of trial environment j, $ProvEnv_{ij}$ is the random effect of the interaction between population provenance i and trial environment j, and ε_{ij} is the random error. The random effects were distributed as $Prov_i \sim N(0, \boldsymbol{S}_{Prov}\sigma_p^2)$, $Env_j \sim N(0, \boldsymbol{S}_{Env}\sigma_e^2)$, $ProvEnv_{ij} \sim N(0, \boldsymbol{I}\sigma_{pe}^2)$, and $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$.

We fit naïve within-year mixed models with the form in order to understand the environmental and genetic effects on juvenile tree height within developmental stage (tree age) as:

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

$$\hat{y}_{ij} = \mu + Pop_i + Env_j + PopEnv_{ij} + \varepsilon_{ij} \tag{5}$$

where $\hat{y}_{ij}$ are the fitted values from each within-environment model (2) from a given year (1980, 1981, 1982, or 1983), μ is the overall mean, Pop_i is the random effect of population i, Env_j is the random effect of trial environment j, $PopEnv_{ij}$ is the random effect of the interaction between population i and trial environment j, and ε_{ij} is the random error. The random effects were assumed to be independent and identically normally distributed, where $Pop_i \sim N(0, \boldsymbol{I}\sigma_p^2)$, $Env_j \sim N(0, \boldsymbol{I}\sigma_e^2)$, $PopEnv_{ij} \sim N(0, \boldsymbol{I}\sigma_{pe}^2)$, and $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$.

We also fit a within-year mixed model with form similar to model (5) using environmental and population provenance weather data as:

$$\hat{y}_{ij} = \mu + Prov_i + Env_j + ProvEnv_{ij} + \varepsilon_{ij} \tag{6}$$

where $\hat{y}_{ij}$ are the fitted values from each within-environment model (2) from a given year (1980, 1981, 1982, or 1983), μ is the overall mean, $Prov_i$ is the random effect of population provenance i, Env_j is the random effect of trial environment j, $ProvEnv_{ij}$ is the random effect of the interaction between population provenance i and trial environment j, and ε_{ij} is the random error. The random effects were distributed as $Prov_i \sim N(0, \boldsymbol{S}_{Prov}\sigma_p^2)$, $Env_j \sim N(0, \boldsymbol{S}_{Env}\sigma_e^2)$, $ProvEnv_{ij} \sim N(0, \boldsymbol{I}\sigma_{pe}^2)$, and $\varepsilon_{ij} \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$, where $\boldsymbol{S}_{Prov}$ and $\boldsymbol{S}_{Env}$ were subset for the population provenances and environments included in each within-year model, respectively.

In models (4) and (6), we used the term “population provenance” rather than “population” for two reasons: (i) to distinguish models (4) and (6) from models (3) and (5) and (ii) the population variance was modeled as the similarity matrix estimated from climatic data from the provenances (seed source locations) of the populations. Because of insufficient population × environment replication in 1988, we did not fit within-year models for 1988.

For each second stage model (3-6), we estimated broad-sense heritability (H^2) as

$$H^2 = \frac{\sigma_P^2}{\sigma_p^2 + (\sigma_{pe}^2/n_e) + (\sigma_\varepsilon^2/n_e n_r)} \tag{7}$$

where n_e and n_r are the number of trial environments (n_e = 10 across years; n_e = 2 within years) and replications per population/provenance per trial environment (n_r = 3), respectively, and σ_p^2, σ_{pe}^2, and σ_ε^2 are the variance components for population provenance, population/provenance × environment, and error, respectively (Holland et al. 2003; Piepho and Moehring 2007).

Predictions using climatic covariance

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Genomic prediction, in which phenotypes are predicted based on genetic similarity (kinship), is a widely used approach in plant and animal breeding for quantitative traits (controlled by many genes with small individual effects) (Hickey et al. 2017). In addition, genotype × environment interactions can be modeled with both genetic and weather-based covariance structures to improve genomic prediction accuracy (Jarquín et al. 2014; Chen et al. 2017; Gillberg et al. 2019; Costa-Neto et al. 2021). We extracted the population, environment, and population × environment BLUPs from the naïve across-year phenotypic model (3) and then fit the following mixed models:

$$popBLUP_i = \mu + Prov_i + \varepsilon_i \tag{8}$$

$$envBLUP_j = \mu + Env_j + \varepsilon_i \tag{9}$$

$$popEnvBLUP_i = \mu + Prov_i + \varepsilon_i \tag{10}$$

$$popEnvBLUP_j = \mu + Env_j + \varepsilon_i \tag{11}$$

$$popEnvBLUP_{ij} = \mu + Prov_i + Env_j + \varepsilon_i \tag{12}$$

$$(popBLUP + popEnvBLUP)_i = \mu + Prov_i + \varepsilon_i \tag{13}$$

$$(popBLUP + popEnvBLUP)_j = \mu + Env_j + \varepsilon_i \tag{14}$$

$$(popBLUP + popEnvBLUP)_{ij} = \mu + Prov_i + Env_j + \varepsilon_i \tag{15}$$

where $popBLUP_i$ is the population BLUPs from model (3) as the response for model (8), $envBLUP_j$ is the environment BLUPs from model (3) as the response for model (9), $popEnvBLUP_i$, $popEnvBLUP_j$, and $popEnvBLUP_{ij}$ are the population × environment BLUPs from model (3) as the response for models (10-12), respectively, $(popBLUP + popEnvBLUP)_i$, $(popBLUP + popEnvBLUP)_j$, and $(popBLUP + popEnvBLUP)_{ij}$ are the sum of the population and population × environment BLUPs from model (3) as the response for models (13-15), respectively, μ is the overall mean for models (8-15), $Prov_i$ is the random effect of population provenance i in models (8, 10, 12, 13, and 15), Env_j is the random effect of trial environment j

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

in models (9, 11, 12, 14, and 15), and ε_i is the random error for models (8-15). The random effects were distributed as $Prov_i \sim N(0, \boldsymbol{S}_{Prov}\sigma_p^2)$, $Env_j \sim N(0, \boldsymbol{S}_{Env}\sigma_e^2)$, and $\varepsilon_i \sim N(0, \boldsymbol{I}\sigma_\varepsilon^2)$. Each model was five-fold cross-validated with 20 replications. Predictive ability was estimated within each fold of each replication as the Pearson's correlation between the observed and predicted values of the test set.

Results

Climatic modeling increases heritable signal for juvenile tree growth

In the "naïve" across-year phenotypic model (3), we found that population and population × environment explained a smaller proportion of the total variance in tree height (9% each) than environment, which accounted for 75% of the total variance (Table 1). Although population variance was low (9%), the highly replicated experimental design allowed for a high estimation of broad-sense heritability for tree height ($H^2 = 0.89$) (Table 1). Year had an effect estimate of 13.46 ± 0.73 cm.

We also fit a "climatic" across-year phenotypic model (4) in which the variance-covariance of the population provenance and environment random effects were modeled as the climatic similarity matrices between population provenances and trial environments, respectively. The climatic across-year model increased the proportion of the total variance in tree height explained by environment (92%) compared to the naïve across-year model (75%) (Table 1). Although population and population × environment variance were smaller (population = 5%, population × environment = 2%) than in the naïve across-year model (population = 9%; population × environment = 9%), broad-sense heritability increased ($H^2_{naive} = 0.89$; $H^2_{climatic} = 0.96$) (Table 1). Year had no effect on tree height in the across-year climatic model.

In the naïve within-year analysis (model 5), environmental variance increased while broad-sense heritability and population variance tended to decrease over time from 1980 (seven-year-old trees, $H^2 = 0.81$, genetic variance = 33%, environmental variance = 46%) to 1983 (10-year-old trees, $H^2 = 0.49$, genetic variance = 4%, environmental variance = 83%) (Table 1). In contrast, specifying the variance-covariance of the population and environment terms with climatic similarity matrices ("climatic" model 6) led to more stable estimates of broad-sense heritability ($H^2_{1980\rightarrow1983} = 0.94 \rightarrow 0.77$) and population × environment variance (population × environment variance$_{1980\rightarrow1983}$ = 4 → 2%) over time (Table 1). Within-year climatic modeling (6) demonstrated an increase in environmental variance and a decrease in population variance over time (genetic variance$_{1980\rightarrow1983}$ = 36 → 4%, environmental variance$_{1980\rightarrow1983}$ = 58 → 92%) similar to that of the naïve models (Table 1).

Climate explains environmental variance in tree height

Because the common garden experiment consisted of even-aged stands planted in the same year across trial locations, tree age (physiological growth stage) and trial environment (location-year) may have been confounded. As such, we modeled the environment BLUPs for tree height with climatic similarity between trial environments (model 9). The environmental similarity

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

matrix, estimated from climatic and geospatial data across trial location-year environments, explained 71% of the variance in environmental BLUP estimates for tree height and had a predictive ability of r = 0.48 (Table 2). Complementarily, climatic phenotypic modeling including climatic variance-covariance matrices within (model 6) and across (model 4) years increased the variance explained by the environment term (Table 2) and removed the effect of year on tree height when compared the naïve phenotypic modeling (models 3 and 5). Combined, these results suggest that the variation in tree height estimates among trial location-year environments was largely driven by climatic variation rather than by the confounding effects of tree growth stage.

Provenance climate is highly predictive of inter-population and population × environment variation in tree height

The population provenance similarity matrix, estimated from historical climatic and geospatial data across population provenances, demonstrated strong population structure (Figure 1) and explained 45% of the variance in populations BLUPs for tree height (model 8), with a predictive ability of r = 0.44 (Table 2). Although population × environment BLUPs were poorly modeled (r < 0; explained variance ≤ 1%) by provenance and/or environment climatic similarity (models 10-12), provenance climatic similarity explained nearly all of the variance (>91%) in and was highly predictive (r = 0.68) of the sum of population BLUPs plus population × environment BLUPs (models 13 and 15) (Table 2). In addition, climatic phenotypic modeling (models 4 and 6) demonstrated more stable estimation of population × environment effects when compared to the naïve phenotypic models (3 and 5). These results indicate a strong signal of local adaption, where populations from provenances with similar geography/climate have similar tree growth, regardless of the trial environment in which they are observed.

Climatic and geospatial variation among trial environments and population provenances

In a PCA using geospatial and weather data from trial environments (location-years), the first and second principal components (PCs) explained 28% and 21% of the variance, respectively (Figure 2). The first PC demonstrated a separation between environments based on altitude and summer/winter temperature (Figure 2). On the second PC, environments were separated by year, spring/winter temperature, and summer precipitation (Figure 2).

We also conducted PCA for population provenances using geospatial and weather data. The first and second PCs explained 44% and 23% of the variance, respectively (Figure 1). The eigenvectors of the PCA predictor variables demonstrated that variation in PC 1 was driven by altitude, temperature, and snowfall (solid precipitation), while variation in PC 2 was driven by rainfall (liquid precipitation). Hierarchical clustering of the PCA revealed six clusters, which were geographically distributed as Carinthia, Austria and Slovenia; Styria, Austria; Lower Austria, Germany, Czech Republic, Slovakia, and Poland; and three alpine regions (Figure 1).

Conclusions

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We used agricultural statistical methods to dissect the heritable and environmental components of juvenile tree growth in Central European Norway Spruce. The highly replicated design of the common garden experiment allowed for precise estimation of genetic (population and population × environment) and environmental (trial location-year) effects. Population explained a relatively small proportion of the total variance in tree height compared to environment, similar to previous findings in this dataset (Kapeller et al. 2017; Messina et al. 2018; Washburn et al. 2020) and consistent with low estimates of population differentiation between Norway spruce populations (Androsiuk et al. 2013). However, broad-sense heritability was moderate to high and previous studies on tree height in clonal and progeny trials of Norway spruce (Hannrup et al. 2004; Chen et al. 2017) have reported similar heritability estimates. Modeling the variance-covariance of the population provenance and environment terms using climate data led to higher estimates of heritability and environmental variance and more stable estimation of heritability and population × environment for tree height across developmental time. The higher heritability estimates with climatic modeling were likely the result of increased environmental signal and subsequently decreased error variance. Tree height becomes less predictive of total biomass with age and growth habit (branching pattern etc.) may be a better indicator of tree growth after the transition between the juvenile and adult stages (Zianis et al. 2005), which may partially explain the reduction in heritability and genetic variance for tree height over time in this study. Drought conditions in Central Europe increased in severity from the beginning to the end of the trials (1983-1988), with a large historical drought event beginning in 1987 (Spinoni et al. 2015), supporting the PCA separation between the 1988 and 1981-1983 trial years reported here. In addition, the increasing environmental variance for tree height with developmental age may be partially explained by increasing drought severity across the trial period.

Although genetic marker data was not available in this dataset, we modeled inter-population relationships using historical climatic data from the population provenances. Plant and animal breeders frequently use climate and/or genetic marker data to model relationships among individuals and/or environments in order to improve the prediction accuracy of phenotypes in new (genetically related) material and even in new environments (Heslot et al. 2014; Jarquín et al. 2014; Crossa et al. 2016; Chen et al. 2017; Messina et al. 2018; Bustos-Korts et al. 2019; Gillberg et al. 2019; Washburn et al. 2020; Costa-Neto et al. 2021). Previous studies have demonstrated that local adaptation is prevalent in tree species, including Norway spruce, and that ancestral environment is a strong predictor of growth phenotypes in Norway spruce (Alberto et al. 2013; Berg and Coop 2014; Aitken and Bemmels 2016; Milesi et al. 2019). We found that climatic relationships among population provenances (a) were highly predictive of population + population × environment tree height variation and (b) demonstrated population structure nested within geography, indicating that these populations have strong local adaptation (Aitken and Bemmels 2016). Our findings complement previous results from ecological modeling of the dataset, where within-population variation was driven by population provenance temperature (Kapeller et al. 2017). The modeling used in this study results in refined tree growth estimates for population and population × environment, which could be used to identify potential sources of material adapted to new environments.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figures and tables

Figure 1. Plots of the first and second principal components (PCs) from principal component analysis (PCA) of trial environments based on geospatial and climatic data. Trial environments were labeled based on (A) a combination of altitude, longitude, and latitude and (B) year. Eigenvectors for the geospatial and weather predictors of the PCA are shown in (B).

Figure 2. (A) Plot of the fist and second principal components (PCs) from principal component analysis (PCA) of population provenances based on geospatial and climatic data. (B) Map of population provenances. Points in (A) and (B) were labeled based on the five clusters identified by hierarchical clustering. (C) Population provenance climatic distance matrix, organized by hierarchical clustering.

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 1. Broad-sense heritability (H^2) and number of levels and estimated variances for population/provenance, environment, and residual random effects from tree height models across (1980-1988) and within (1980-1983) years.

Year	Model	Term	Estimated variance ± standard error	Proportion of total variance	Levels	H^2
Across (1980-1988)	Naive	Population	35.67 ± 3.67	0.094	278	0.89
		Environment	284.38 ± 39.62	0.747	107	
		Pop × Env	32.55 ± 1.42	0.086	2702	
		Error	27.91 ± 0.63	0.073		
	Climatic	Provenance	289.49 ± 44.30	0.050	278	0.96
		Environment	5325.40 ± 717.78	0.920	107	
		Prov × Env	115.60 ± 4.15	0.020	2702	
		Error	59.77 ± 1.29	0.010		
1980	Naive	Population	14.06 ± 1.63	0.333	278	0.81
		Environment	19.33 ± 5.85	0.457	24	
		Pop × Env	5.78 ± 0.59	0.137	612	
		Error	3.09 ± 0.14	0.073		
	Climatic	Provenance	60.57 ± 9.81	0.362	278	0.94
		Environment	96.66 ± 29.47	0.578	24	
		Prov × Env	6.84 ± 0.60	0.041	612	
		Error	3.08 ± 0.14	0.018		
1981	Naive	Population	14.91 ± 1.99	0.169	278	0.72
		Environment	56.16 ± 16.80	0.638	24	
		Pop × Env	8.50 ± 1.01	0.097	612	
		Error	8.46 ± 0.39	0.096		
	Climatic	Provenance	60.17 ± 11.20	0.160	278	0.91
		Environment	297.05 ± 89.45	0.792	24	
		Prov × Env	9.57 ± 0.96	0.026	612	
		Error	8.43 ± 0.39	0.022		
1982	Naive	Population	18.42 ± 3.23	0.073	278	0.57
		Environment	191.41 ± 56.97	0.762	24	
		Pop × Env	20.31 ± 2.42	0.081	612	
		Error	20.97 ± 0.96	0.083		
	Climatic	Provenance	73.21 ± 16.98	0.069	278	0.84
		Environment	938.58 ± 281.11	0.890	24	
		Prov × Env	21.55 ± 2.21	0.020	612	
		Error	20.93 ± 0.97	0.020		
1983	Naive	Population	29.73 ± 6.24	0.042	278	0.49
		Environment	582.74 ± 173.03	0.825	24	
		Pop × Env	47.06 ± 5.49	0.067	612	
		Error	46.68 ± 2.17	0.066		
	Climatic	Provenance	109.13 ± 30.34	0.041	278	0.77
		Environment	2427.60 ± 724.66	0.922	24	
		Prov × Env	50.90 ± 5.07	0.019	612	
		Error	46.56 ± 2.16	0.018		

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Table 2. Predictive ability and proportion of total variance in population, environment, and/or population × environment best linear unbiased predictors (BLUPs) for tree height explained by population provenance and/or trial environment climatic similarity matrices.

Response (BLUPs)	**Model**	**Provenance (prop. var.)**	**Environment (prop. var.)**	**Predictive ability**
Population	8	0.450 ± 0.068		0.441 ± 0.083
Environment	9		0.706 ± 0.115	0.476 ± 0.157
Population × Environment	10	$7x10^{-7} \pm 5x10^{-7}$		-0.068 ± 0.061 b
	11	$1x10^{-6} \pm 6x10^{-7}$		-0.169 ± 0.037 a
	12	0.014 ± 0.049	$5x10^{-17} \pm 3x10^{-17}$	-0.090 ± 0.073 b
Population + Population × Environment	13	0.911 ± 0.091		0.678 ± 0.064 a
	14		0.024 ± 0.090	0.063 ± 0.084 b
	15	0.921 ± 0.009	$6x10^{-18} \pm 6x10^{-17}$	0.676 ± 0.028 a

All models were five-fold cross-validated with 20 replications. Predictive ability was defined as the Pearson's correlation between the observed and predicted values of the test set.

Funding information

This work was partially supported by funding from the Austrian Academy of Sciences.

Author contributions

Conceptualization, LM; Data curation, LM; Formal analysis, LM; Funding acquisition, KS; Methodology, LM; Project administration, LM; Software, LM; Validation, LM; Visualization, LM; Writing-original draft, LM; Writing-review and editing, LM, KS

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Aitken SN, Bemmels JB (2016) Time to get moving: Assisted gene flow of forest trees. Evol Appl 9:271–290. https://doi.org/10.1111/eva.12293

Alberto FJ, Aitken SN, Alía R, et al (2013) Potential for evolutionary responses to climate change - evidence from tree populations. Glob Chang Biol 19:1645–1661. https://doi.org/10.1111/gcb.12181

Androsiuk P, Shimono A, Westin J, et al (2013) Genetic status of Norway spruce (*Picea abies*) breeding populations for Northern Sweden. Silvae Genet 62:127–136. https://doi.org/10.1515/sg-2013-0017

Avanzi C, Piermattei A, Piotti A, et al (2019) Disentangling the effects of spatial proximity and genetic similarity on individual growth performances in Norway spruce natural populations. Sci Total Environ 650:493–504. https://doi.org/10.1016/j.scitotenv.2018.08.348

Berg JJ, Coop G (2014) A Population Genetic Signal of Polygenic Adaptation. PLoS Genet 10: e1004412. https://doi.org/10.1371/journal.pgen.1004412

Bustos-Korts D, Boer MP, Malosetti M, et al (2019) Combining Crop Growth Modeling and Statistical Genetic Modeling to Evaluate Phenotyping Strategies. Front Plant Sci 10:1–21. https://doi.org/10.3389/fpls.2019.01491

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Chen ZQ, Karlsson B, Wu HX (2017) Patterns of additive genotype-by-environment interaction in tree height of Norway spruce in southern and central Sweden. Tree Genet Genomes 13:25. https://doi.org/10.1007/s11295-017-1103-6

Chimani B, Böhm R, Matulla C, Ganekind M (2011) Development of a longterm dataset of solid/liquid precipitation. Adv Sci Res 6:39–43. https://doi.org/10.5194/asr-6-39-2011

Chimani B, Matulla C, Böhm R, Hofstätter M (2013) A new high resolution absolute temperature grid for the Greater Alpine Region back to 1780. Int J Climatol 33:2129–2141. https://doi.org/10.1002/joc.3574

Costa-Neto G, Fritsche-Neto R, Crossa J (2021) Nonlinear kernels, dominance, and envirotyping data increase the accuracy of genome-based prediction in multi-environment trials. Heredity 126:92–106. https://doi.org/10.1038/s41437-020-00353-1

Crossa J, De Los Campos G, Maccaferri M, et al (2016) Extending the marker × Environment interaction model for genomic-enabled prediction and genome-wide association analysis in durum wheat. Crop Sci 56:2193–2209. https://doi.org/10.2135/cropsci2015.04.0260

Crossa J, Pérez-Rodríguez P, Cuevas J, et al (2017) Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends Plant Sci 22:961–975. https://doi.org/10.1016/j.tplants.2017.08.011

Efthymiadis D, Jones PD, Briffa KR, et al (2006) Construction of a 10-min-gridded precipitation data set for the Greater Alpine Region for 1800-2003. J Geophys Res Atmos 111:6120. https://doi.org/10.1029/2005JD006120

Gillberg J, Marttinen P, Mamitsuka H, Kaski S (2019) Modelling G3E with historical weather information improves genomic prediction in new environments. Bioinformatics 35:4045–4052. https://doi.org/10.1093/bioinformatics/btz197

Hannrup B, Cahalan C, Chantre G, et al (2004) Genetic Parameters of Growth and Wood Quality Traits in *Picea abies*. Scand J For Res 19:14–29. https://doi.org/10.1080/02827580310019536

Heslot N, Akdemir D, Sorrells ME, Jannink JL (2014) Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. Theor Appl Genet 127:463–480. https://doi.org/10.1007/s00122-013-2231-5

Heslot N, Jannink J-L, Sorrells ME (2015) Perspectives for genomic selection applications and research in plants. Crop Sci 55:1–12

Hickey JM, Chiurugwi T, Mackay I, Powell W (2017) Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. Nat Genet 49:1297–1303. https://doi.org/10.1038/ng.3920

Hiebl J, Auer I, Böhm R, et al (2009) A high-resolution 1961-1990 monthly temperature climatology for the greater Alpine region. Meteorol Zeitschrift 18:507–530. https://doi.org/10.1127/0941-2948/2009/0403

Holland JB, Nyquist WE, Cervantes-Martinez CT (2003) Estimating and interpreting heritability for plant breeding. Plant Breed Rev 22:9–112

Isidro y Sánchez J, Akdemir D (2021) Training Set Optimization for Sparse Phenotyping in Genomic Selection: A Conceptual Overview. Front Plant Sci 12:1–14. https://doi.org/10.3389/fpls.2021.715910

Jarquín D, Crossa J, Lacaze X, et al (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor Appl Genet 127:595–607. https://doi.org/10.1007/s00122-013-2243-1

Kapeller S, Dieckmann U, Schueler S (2017) Varying selection differential throughout the climatic range of Norway spruce in Central Europe. Evol Appl 10:25–38. https://doi.org/10.1111/eva.12413

Kapeller S, Dieckmann U, Schueler S (2016) Data from: Varying selection differential throughout the climatic range of Norway spruce in Central Europe. Dryad https://doi.org/10.5061/dryad.877ts

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Kapeller S, Lexer MJ, Geburek T, et al (2012) Intraspecific variation in climate response of Norway spruce in the eastern Alpine range: Selecting appropriate provenances for future climate. For Ecol Manage 271:46–57. https://doi.org/10.1016/j.foreco.2012.01.039

Lê S, Josse J, Husson F (2008) FactoMineR: an R package for multivariate analysis. J Stat Softw 25:1–18. https://doi.org/10.18637/jss.v025.i01

Leslie AB, Beaulieu JM, Rai HS, et al (2012) Hemisphere-scale differences in conifer evolutionary dynamics. Proc Natl Acad Sci USA 109:16217–16221. https://doi.org/10.1073/pnas.1213621109

Lindner M, Maroschek M, Netherer S, et al (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. For Ecol Manage 259:698–709. https://doi.org/10.1016/j.foreco.2009.09.023

Messina CD, Technow F, Tang T, et al (2018) Leveraging biological insight and environmental variation to improve phenotypic prediction: Integrating crop growth models (CGM) with whole genome prediction (WGP). Eur J Agron 100:151–162. https://doi.org/https://doi.org/10.1016/j.eja.2018.01.007

Milesi P, Berlin M, Chen J, et al (2019) Assessing the potential for assisted gene flow using past introduction of Norway spruce in southern Sweden: Local adaptation and genetic basis of quantitative traits in trees. Evol Appl 12:1946–1959. https://doi.org/10.1111/eva.12855

Muñoz F, Sanchez L (2020) breedR: Statistical Methods for Forest Genetic Resources Analysts. http://famuvie.github.io/breedR/

Nather J, Holzer K (1979) Über die Bedeutung und die Anlage von Kontrollflächen zur Prüfung von anerkanntem Fichtenpflanzgut. In: Informationsdienst Forstliche Bundesversuchsanstalt Wien 181

Oleksyn J, Modrzyński J, Tjoelker MG, et al (1998) Growth and physiology of *Picea abies* populations from elevational transects: Common garden evidence for altitudinal ecotypes and cold adaptation. Funct Ecol 12:573–590. https://doi.org/10.1046/j.1365-2435.1998.00236.x

Piepho HP, Moehring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. https://doi.org/10.1534/genetics.107.074229

Piepho HP, Möhring J, Schulz-Streeck T, Ogutu JO (2012) A stage-wise approach for the analysis of multi-environment trials. Biometrical J 54:844–860. https://doi.org/10.1002/bimj.201100219

Piepho HP, Tempelman RJ, Williams ER (2020) Guest Editors' Introduction to the Special Issue on "Recent Advances in Design and Analysis of Experiments and Observational Studies in Agriculture." J Agric Biol Environ Stat 25:453–456. https://doi.org/10.1007/s13253-020-00417-z

Pierce D (2021) ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files. http://cirrus.ucsd.edu/~pierce/ncdf/

R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rodríguez-Álvarez MX, Boer MP, van Eeuwijk FA, Eilers PHC (2018) Correcting for spatial heterogeneity in plant breeding experiments with P-splines. Spat Stat 23:52–71. https://doi.org/https://doi.org/10.1016/j.spasta.2017.10.003

Schulze U (1985) Fichtenherkunftsprüfung 1978. In: Internationaler Verband Forstlicher Versuchsanstalten IUFRO, Arbeitsgruppe S 2.02-11 Fichtenherkünfte. pp 1–12

Seidl R, Thom D, Kautz M, et al (2017) Forest disturbances under climate change Europe PMC Funders Group. Nat Clim Chang 7:395–402. https://doi.org/10.1038/nclimate3303.Forest

Spinoni J, Naumann G, Vogt J V., Barbosa P (2015) The biggest drought events in Europe from 1950 to 2012. J Hydrol Reg Stud 3:509–524. https://doi.org/10.1016/j.ejrh.2015.01.001

bioRxiv preprint doi: https://doi.org/10.1101/2022.01.10.475611; this version posted May 4, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

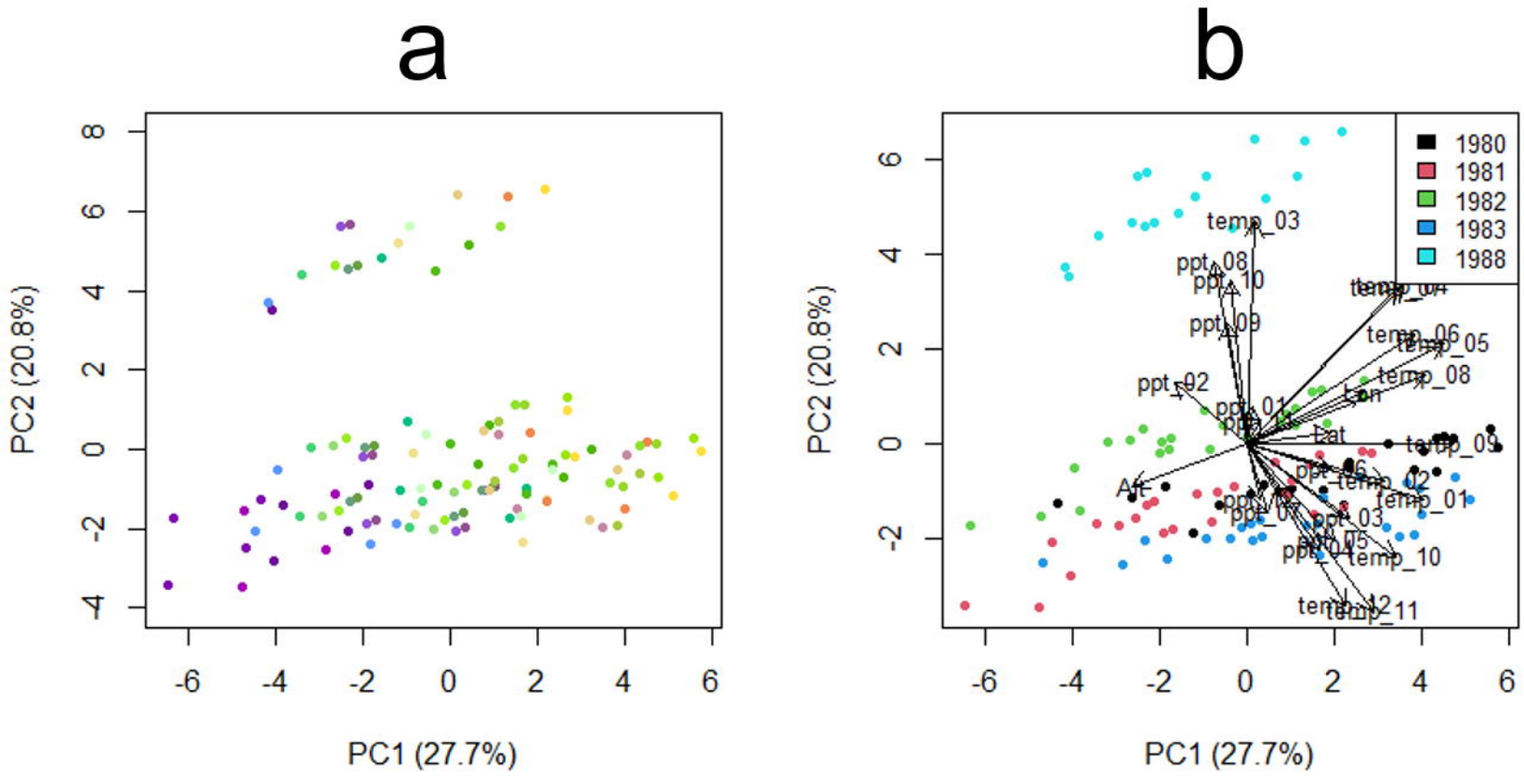
Wang X, Bernhardsson C, Ingvarsson PK (2020) Demography and Natural Selection Have Shaped Genetic Variation in the Widely Distributed Conifer Norway Spruce (*Picea abies*). Genome Biol Evol 12:3803–3817. https://doi.org/10.1093/gbe/evaa005

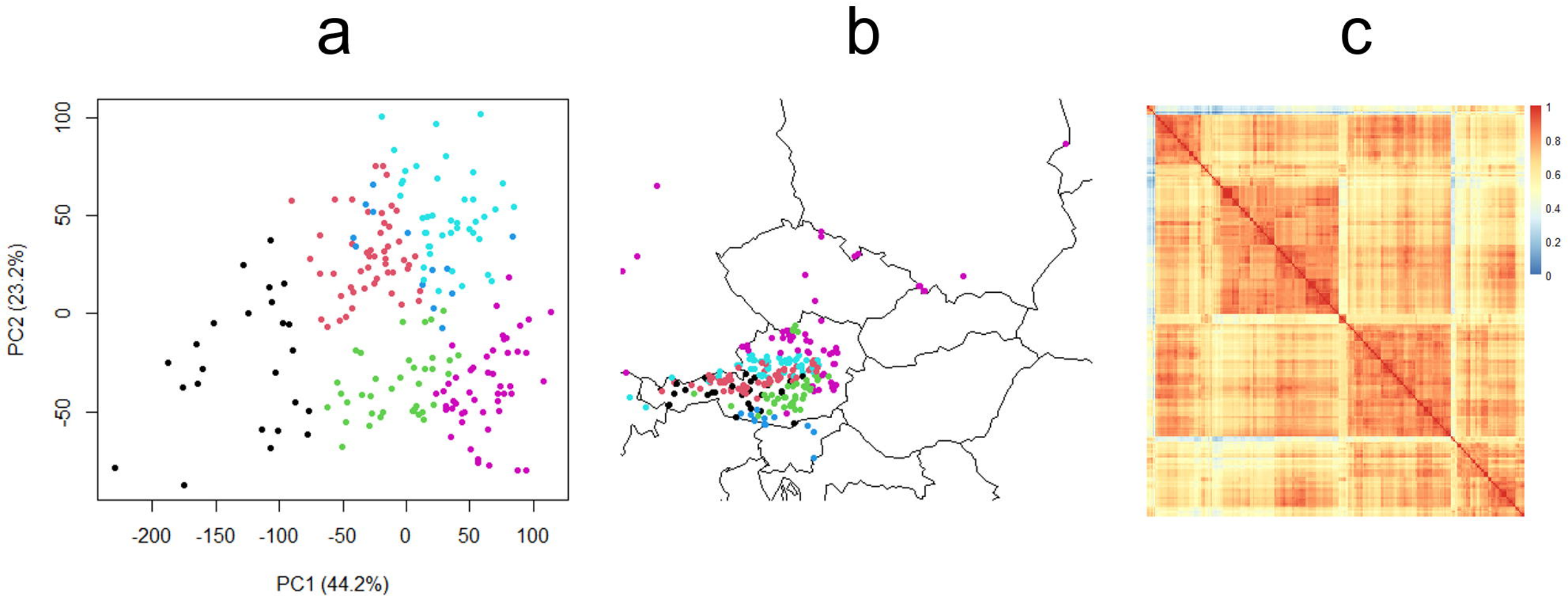
Washburn JD, Burch MB, Franco JAV (2020) Predictive breeding for maize: Making use of molecular phenotypes, machine learning, and physiological crop models. Crop Sci 60:622–638. https://doi.org/10.1002/csc2.20052

Williams MI, Dumroese RK (2013) Preparing for Climate Change: Forestry and Assisted Migration. J For 111:287–297. https://doi.org/http://dx.doi.org/10.5849/jof.13-016

Yeaman S, Hodgins KA, Lotterhos KE, et al (2016) Convergent local adaptation to climate in distantly related conifers. Science 353:1431–1433. https://doi.org/10.1126/science.aaf7812

Zianis D, Muukkonen P, Mäkipää R, et al (2005) Biomass and stem volume equations for tree species in Europe. Silva Fenn Monogr 4:1–63. https://doi.org/10.14214/sf.sfm4


a
b
PC1 (27.7%)
PC2 (20.8%)
1980
1981
1982
1983
1988
temp_03
ppt_08
ppt_10
ppt_09
ppt_02
temp_06
temp_05
temp_08
Lon
Lat
temp_09
temp_02
temp_01
Alt
ppt_03
temp_10
temp_11


a
b
c
PC2 (23.2%)
PC1 (44.2%)
100
50
0
-50
-200
-150
-100
-50
0
50
100
1
0.8
0.6
0.4
0.2
0