Combining climatic and genomic data improves range-wide tree height growth prediction in a forest tree

[1], Juliette Archambeau[1], Marta Benito Garzón[1], Frédéric Barraquand[2], Marina de Miguel Vega[1,3], Christophe Plomion[1] and Santiago C. González-Martínez[1]

[1]INRAE, Univ. Bordeaux, BIOGECO, F-33610 Cestas, France
[2]CNRS, Institute of Mathematics of Bordeaux, F-33400 Talence, France
[3]EGFV, Univ. Bordeaux, Bordeaux Sciences Agro, INRAE, ISVV, F-33882, Villenave d'Ornon, France

Abstract

Population response functions based on climatic and phenotypic data from common gardens have long been the gold standard for predicting quantitative trait variation in new environments. However, prediction accuracy might be enhanced by incorporating genomic information that captures the neutral and adaptive processes behind intra-population genetic variation. We used five clonal common gardens containing 34 provenances (523 genotypes) of maritime pine (*Pinus pinaster* Aiton) to determine whether models combining climatic and genomic data capture the underlying drivers of height-growth variation, and thus improve predictions at large geographical scales. The plastic component explained most of the height-growth variation, probably resulting from population responses to multiple environmental factors. The genetic component stemmed mainly from climate adaptation, and the distinct demographic and selective histories of the different maritime pine gene pools. Models combining climate-of-origin and gene pool of the provenances, and positive-effect height-associated alleles (PEAs) captured most of the genetic component of height-growth and better predicted new provenances compared to the climate-based population response functions. Regionally-selected PEAs were better predictors than globally-selected PEAs, showing high predictive ability in some environments, even when included alone in the models. These results are therefore promising for the future use of genome-based prediction of quantitative traits.

Keywords: Climate change, local adaptation, phenotypic plasticity, population response functions, positive-effect alleles, range-wide predictive models, maritime pine.

1 Introduction

Global change is expected to have a profound impact on forests (Franklin et al. 2016, Seidl et al. 2017), and whether tree populations will be able to migrate or persist across their current range is uncertain (Aitken et al. 2008). Assessing the potential of populations to accommodate future environmental conditions requires a thorough understanding of the origin of variation in quantitative traits subject to natural selection (Shaw and Etterson 2012, Alberto et al. 2013). To this aim, a necessary first step is to quantify the plastic and genetic components of adaptive traits and their interaction in multiple environments (Des Marais et al. 2013, Merilä and Hendry 2014), which has been done extensively in forest trees (Franks et al. 2014). A second step consists in identifying the underlying drivers of these components (Merilä and Hendry 2014). The plastic component corresponds to the ability of one genotype to produce varying phenotypes depending on the environment (Bradshaw 1965). Phenotypic plasticity can help individuals to overcome new conditions up to a certain threshold (Nicotra et al. 2010), and can be to some extent genetically assimilated and therefore involved in the evolutionary process of adaptation (Pigliucci et al. 2006). The genetic component can stem from both neutral (e.g. population demographic history and genetic drift) and adaptive processes (e.g. adaptation to local biotic and abiotic environments), both processes implying changes in allele frequencies. Populations are locally adapted when they have higher fitness in their own environment than populations from other environments (Kawecki and Ebert 2004). In forest trees, a large amount of work highlighted the importance of climate in driving the plastic and genetic responses of quantitative traits to new environmental conditions (Savolainen et al. 2007, Valladares et al. 2014b). However, it is still unclear how multiple and interacting drivers underlying quantitative trait variation could be combined to improve predictions of population responses to global change. The increasing availability of genomic data opens new opportunities to boost prediction accuracy, which is critical for breeding (i.e. genomic selection; Grattapaglia and Resende 2011), to anticipate future distribution of natural populations (e.g. Razgour et al. 2019), or to support the ongoing development of assisted gene flow strategies aiming to help populations adapt to future environments (Browne et al. 2019, Mahony et al. 2020, MacLachlan et al. 2021).

In forest trees, a long history of common gardens (Langlet 1971) has provided a unique framework to associate population-specific quantitative trait variation with large environmental or geographical gradients, and thus identify populations at risk under climate change (Rehfeldt et al. 1999, 2003, Savolainen et al. 2007, Pedlar and McKenney 2017, Rehfeldt et al. 2018, Fréjaville et al. 2020). The development of population response functions was a step forward to evaluate the relative contribution of plasticity -associated to current climatic conditions (i.e. the climate in the common gardens)- and genetic adaptation -associated to the past climatic conditions under which the populations have evolved (i.e. the climate-of-origin of the provenances tested)- in explaining quantitative trait variation (O'Neill et al. 2008, Wang et al. 2010). These models have now been applied to a large variety of traits (Leites et al. 2012a,b, Benito Garzón et al. 2019, Vizcaíno-

Palomar et al. 2020) and one of their main conclusions is that trait variation across species ranges is mostly associated with the climate in the common garden (i.e. related to the plastic component) and, only to a much lesser extent, with the climate-of-origin of the provenances (i.e. related to the genetic component) (Leites et al. 2012b, Benito Garzón et al. 2019). Importantly, these models do not allow to determine to what extent associations between trait variation and provenance climate-of-origin, or the higher trait values of local compared to foreign populations, are caused by adaptive or neutral processes (Leimu and Fischer 2008, Hereford 2009, Franks et al. 2014). This limits our understanding of the genetic processes that led to the current patterns of quantitative trait variation, and therefore our ability to predict trait variation of new (untested in common gardens) populations under new environments.

The advent and generalization of genomic tools have enhanced our understanding of adaptive and neutral genetic processes resulting in trait variation, and their relationship with climatic gradients (Savolainen et al. 2013, Sork 2018, Leroy et al. 2020). Integrating genomic information into quantitative trait prediction would be highly valuable to consider intraspecific variability at a finer scale than in current models (Mahony et al. 2020), thereby probably improving model accuracy, especially for populations not previously planted in commons gardens. More specifically, rapidly growing knowledge on trait-associated alleles identified by Genome-Wide Association Studies (GWAS) is promising for anticipating the genetic response of populations to new environments (Exposito-Alonso et al. 2018, Browne et al. 2019). For example, Mahony et al. (2020) used counts of alleles positively associated with the traits of interest (PEAs) to describe patterns and identify drivers of local adaptation in lodgepole pine. Recent studies have shown that most quantitative traits are highly polygenic (see reviews in Pritchard et al. 2010, Barghi et al. 2020; and de Miguel et al. 2020 for maritime pine) and that the effect of trait-associated alleles may vary across environments (Anderson et al. 2013, Tiffin and Ross-Ibarra 2014), which complicates the use of genomic information in trait prediction. In addition, patterns in allele frequencies induced by population demographic history are often correlated with environmental gradients (Latta 2009, Alberto et al. 2013, Nadeau et al. 2016), which makes difficult to separate the signature of population structure from that of adaptive processes (Sella and N. H. Barton 2019, Sohail et al. 2019). At the species range scale, population structure hinders the use of genomic relationship matrices, which provide more accurate estimates of genetic parameters (e.g. breeding values, additive and non-additive variance) within breeding populations than previously used pedigree-based approaches (Bouvet et al. 2016, El-Dien et al. 2018). Indeed, admixed populations or distinct genetic groups may present different means and variances of their genetic values, which requires new statistical methods to estimate them (e.g. Muff et al. 2019). Thus, integrating genomic information into quantitative trait prediction in natural populations, while highly valuable, remains challenging.

Forest trees are remarkable models to study the genetic and plastic components of quantitative trait variation. Forest tree populations often have large effective population size and are distributed along a large range of environmental conditions, which makes them especially suitable

to study current and future responses to climate (Savolainen et al. 2007, Alberto et al. 2013). Moreover, forest trees remain largely undomesticated (including those species with breeding programs) and, therefore, genetic variation in natural populations has been little influenced by human-induced selection (Neale and Savolainen 2004). However, forest trees have also large and complex genomes (especially conifers; Mackay et al. 2012), that show a rapid decay of linkage disequilibrium (Olson et al. 2010), and extensive genotyping would be needed to identify all (most) relevant polymorphisms underlying (highly polygenic) quantitative traits (Neale and Savolainen 2004, Jaramillo-Correa et al. 2015). In addition, although early results have been convincing in predicting trait variation within tree breeding populations (i.e. using populations with relatively low effective population size; Resende Jr et al. 2012, Resende et al. 2012, Jarquín et al. 2014), predicting the genetic component of trait variation across populations or geographical regions of forest trees remains poorly explored.

In the present study, we aim to identify the potential drivers of the plastic and genetic components of height growth in distinct maritime pine gene pools (i.e. genetic clusters) and investigate how common garden data can be combined with genomics to efficiently predict height-growth variation across the species range. We compared Bayesian hierarchical mixed models that inferred height-growth variation in maritime pine as a function of climatic and genomic-related variables, using a clonal common garden network (CLONAPIN) consisting of five sites and 34 provenances (523 genotypes and 12,841 trees). First, we evaluated the relative importance of potential drivers underlying height-growth variation. We expected that: (i) the plastic component explains most trait variation and is associated with climate in the common gardens, (ii) the genetic component is driven by both adaptive processes, such as adaptation to climate, and neutral processes, such as population demographic history. Second, we compared the out-of-sample predictive ability (on unknown observations or provenances) of models based exclusively on the common garden design and models including (either separately or jointly) potential predictors of the genetic component of trait variation, notably those related to climate and positive-effect height-associated alleles (PEAs). We expected that the distinct demographic history of maritime pine gene pools, the provenance climate-of-origin and the counts of PEAs, either combined or alone, may improve height-growth predictions of unknown provenances. We also expected that height-associated alleles selected regionally, i.e. in particular environments, would have a better predictive ability than globally-selected alleles. Our study is a step towards integrating the recent knowledge brought by large genomic datasets to the modeling of quantitative trait variation in forest trees. Combining common gardens with genomic tools hold great promise for speeding up and improving trait predictions at large scales and for a wide range of species and populations. However, a robust framework is needed to make reliable predictions and to determine when and to what extent genomics can help in making decisions in conservation strategies or in anticipating population responses to climate change.

2 Materials & Methods

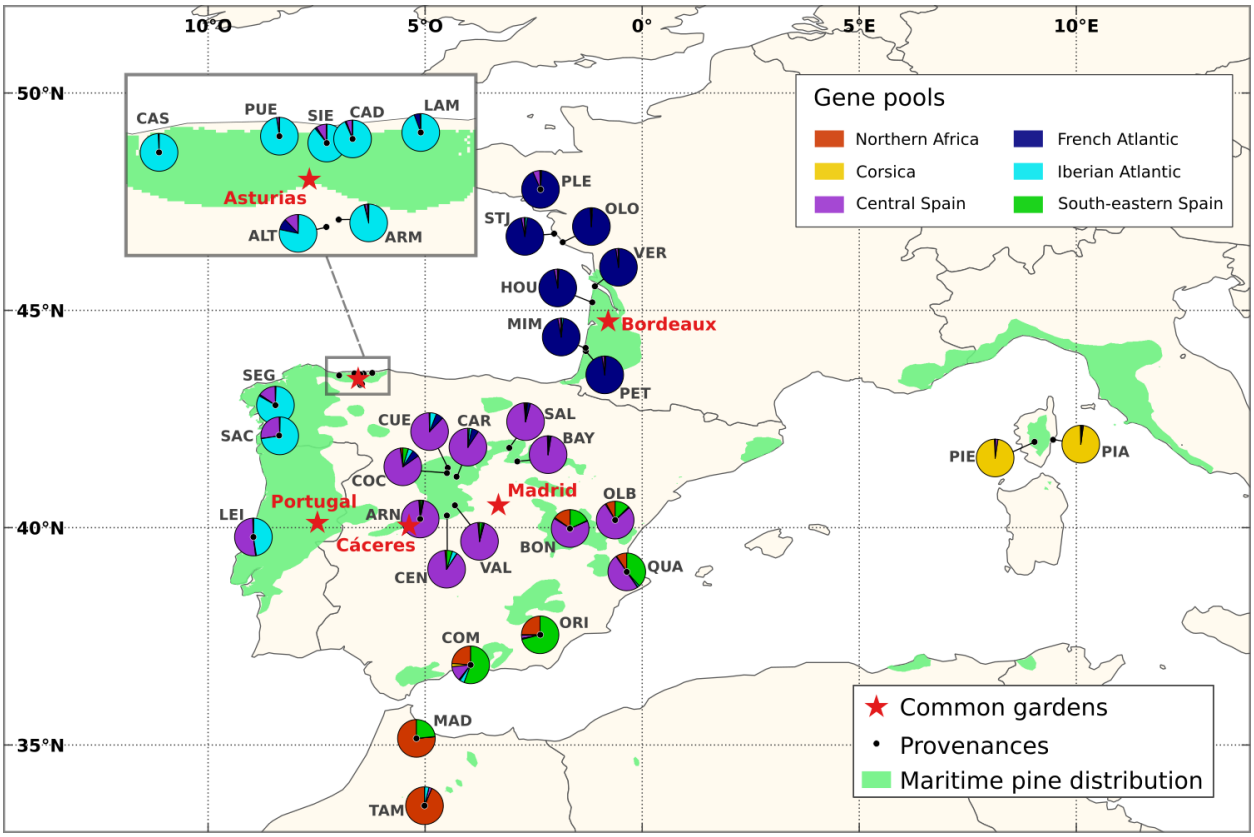
2.1 Plant material and phenotypic measurements

Maritime pine (*Pinus pinaster* Ait., Pinaceae) is an economically important forest tree, largely exploited for its wood (Viñas et al. 2016). It has also an important ecological function stabilizing coastal and fossil dunes and as keystone species supporting forest biodiversity. Native to the western part of the Mediterranean Basin, the Atlas mountains in Morocco, and the south-west Atlantic coast of Europe, its natural distribution spans from the High Atlas mountains in the south (Morocco) to French Brittany in the north, and from the coast of Portugal in the west to western Italy in the east. Maritime pine is a wind-pollinated, outcrossing and long-lived tree species that can grow on a wide range of substrates, from sandy and acidic soils to more calcareous ones. It can also withstand many different climates: from the dry climate of the Mediterranean Basin to the highly humid climate of the Atlantic Europe region, and the continental climate of central Spain. Maritime pine populations are highly fragmented and can be grouped into six gene pools (Alberto et al. 2013, Jaramillo-Correa et al. 2015; see fig. 1), that is genetic clusters that cannot be differentiated on the basis of neutral genetic markers and that probably derive from a common glacial refuge (Bucci et al. 2007, Santos-del-Blanco et al. 2012).

Height growth is a key adaptive trait in forest trees, including maritime pine. Height can be seen as the end-product of multiple ecophysiological processes that are both genetically regulated and affected by multiple environmental effects (Grattapaglia et al. 2009). As such, taller trees compete more efficiently for light, water and nutrients, and are also more likely to have high fecundity (Rehfeldt et al. 1999, Wu and Ying 2004, Aitken and Bemmels 2015). We obtained height data from the clonal common garden network CLONAPIN, consisting of five common gardens located in different environments (also referred as test sites; fig. 1). Three sites are located in the Atlantic Europe region, with mild winters, high annual rainfall and relatively wet summers: Bordeaux in the French part, and Asturias and Portugal in the Iberian part, the Portugal site experiencing slightly colder winters and half the summer precipitation than the site in Asturias. The two other sites, Cáceres and Madrid, are located in the Mediterranean region with high temperatures and intense summer drought, as well as large precipitation differences between summer and winter. In 2010 or 2011 depending on the test site, clonal replicates from 34 provenances were planted in a randomized complete block design with eight blocks. For each provenance, trees represent between 2 and 28 genotypes (clones), on average about 15 (see Rodríguez-Quilón et al. 2016 for details). Genotypes were originally sampled from natural populations, with enough distance among trees (over 50 m) to avoid sampling related individuals. Depending on the site, height was measured from one to four times, when the trees were between 13 and 41 month old (Table S1). Only survivors were measured for height, which resulted in a strongly unbalanced design as 92% and 75% of the trees died in Cáceres and Madrid, respectively (partly due to the clay soils and a strong summer drought). After removing genotypes for which we had no genomic information,

we analyzed 33,121 height observations from 12,841 trees and 523 genotypes (Table S2).

Figure 1. The five common gardens and 34 provenances of maritime pine (CLONAPIN common garden network) used in this study. The distribution of maritime pine is also shown (based on EUFORGEN map, http://www.euforgen.org/). Pie charts represent the proportions belonging to each gene pool for each provenance (see legend) as estimated in Jaramillo-Correa et al. (2015). Provenance names can be found in Table S2.



.

2.2 Gene pool assignment and positive-effect alleles (PEAs)

DNA was extracted from leaves collected in the Asturias common garden and genotyped with a 9k Illumina Infinium SNP assay (described in Plomion et al. (2016)), resulting in 5,165 high-quality polymorphic SNPs scored on 523 genotypes. There were on average only 3.3 missing values per genotype (ranging between 0 and 142). For each genotype, the proportion belonging to each gene pool was estimated in Jaramillo-Correa et al. (2015), using nine nuSSRs as well as a subset of the same SNPs as in our study (1,745 SNPs) and the Bayesian approach available in Structure v. 2.3.3 (Pritchard et al. 2000; Table S3). This gene pool assignment aimed at reflecting the neutral genetic structure in maritime pine, which results from population demographic history and genetic drift, but may also arise from different selective histories across gene pools.

Based on the 523 genotypes for which there were both genotypic and phenotypic data, we performed four GWAS following the Bayesian variable selection regression (BVSR) methodology implemented in the piMASS software (Guan and Stephens 2011), correcting for population structure and using the height BLUPs reported in de Miguel et al. (2020), that accounted for site and block effects. First, a global GWAS was performed to identify SNPs that have an association with height at range-wide geographical scales, thus using the combined phenotypic data from the five common gardens. Second, three regional GWAS were performed to identify SNPs that have a local association with height in a particular geographical region r (i.e. in a particular environment), thus using separately data from the Iberian Atlantic common gardens (Asturias and Portugal), the French Atlantic common garden (Bordeaux) and the Mediterranean common gardens (Madrid and Cáceres). For each of the four GWAS, we selected the 350 SNPs ($\sim$7% top associations) with the highest absolute Rao-Blackwellized estimates of the posterior effect size, corresponding approximately to the estimated number of SNPs with non-zero effects on height in a previous multi-trait study using the same SNP marker set (de Miguel et al. 2020). These SNPs were used to compute the counts of global and regional positive-effect alleles (gPEAs and rPEAs) for each genotype (see section 2.1 of the Supplementary Information for more details).

2.3 Climatic data

In forest trees, large-scale patterns of allele frequencies or quantitative trait variation are known to be associated with climatic variables related to mean temperature and precipitation (e.g. Eckert et al. 2010, McLane et al. 2011, Leites et al. 2019, Fréjaville et al. 2020, Mahony et al. 2020), or episodic climatic conditions, such as summer aridity or maximum temperatures (Rehfeldt et al. 2003, Grivet et al. 2011, McLane et al. 2011, Jaramillo-Correa et al. 2015, Fréjaville et al. 2020). As climate change will cause major changes in temperature and precipitation in the near future, particularly in the Mediterranean basin, there is a need to understand the complex influence of climatic variables on quantitative trait variation. We extracted monthly and yearly climatic data from the EuMedClim database with 1 km resolution (Fréjaville and Benito Garzón 2018). The climatic similarity among test sites was described by a covariance matrix Ω including six variables related to both extreme and average temperature and precipitation in the test sites during the year preceding the measurements, and with at most a correlation coefficient of 0.85 among each other (see section 3.1 in the Supplementary Information for more details). The climatic similarity among provenances was described by a covariance matrix Φ including four variables related to the mean temperature and precipitation in the provenance locations over the period from 1901 to 2009 (i.e. representing the climate under which provenances have evolved), and with at most a correlation coefficient of 0.77 among each other (see section 3.2 in the Supplementary Information for more details).

2.4 Hierarchical height-growth models

Twelve height-growth models were compared. We first built two baseline models relying exclusively on the common garden design and aimed at quantifying the relative contribution of the genetic and plastic components of height-growth variation (*models M1* and *M2*; Table 1). Second, we used climatic and genomic data to detect association of height-growth variation with potential underlying drivers related to plasticity, adaptation to climate or gene pool assignment (i.e. a proxy of the population demographic history and genetic drift experienced by the populations), and estimated gene pool-specific total genetic variances (*models M3* to *M6*; Table 1). Third, we built models either including separately or combining potential drivers of the genetic component of height-growth variation to predict unknown observations and provenances without relying on the common garden design (*models M7* to *M12*; Table 1). In every model, the logarithm of height ($\log(h)$) was used as a response variable to stabilize the variance. Tree age at the time of measurement i was included as a covariate to account for the average height-growth trajectory. This implies that all models shared the form $\log(h_i) = f(\text{age}_i) + m(\text{covariates})$, where $m(\text{covariates})$ is the rest of the model. Therefore, all models can also be written $h_i = \exp(f(\text{age}_i))\ \exp(m(\text{covariates}))$, which explains why covariates in our models affect height growth (i.e. modulate the height-growth trajectory) rather than simply height. We used a second-degree polynomial to account for tree age ($f(\text{age}_i + \text{age}_i^2)$) because the logarithm of height first increases linearly with age and then reaches a threshold (fig. S11). Each tree was measured between one and four times (14% of the trees were measured only once), but we did not include a varying intercept for each tree as it resulted in model miss-specification warnings and strong overfitting. A description of each model specification follows.

2.4.1 Baseline *models M1* and *M2*: separating the genetic and plastic components of height-growth variation

In the baseline *model M1*, height h was modeled as a function of tree age, varying intercepts for the sites S_s and blocks nested within sites $B_{b(s)}$ (i.e. the plastic component), and varying intercepts for the provenances P_p and genotypes within provenances $G_{g(p)}$ (i.e. the genetic component):

$$\begin{aligned} \log(h_{isbpg}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sbpg}\ ,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\text{age}_i + \beta_{age2}\text{age}_i^2 \\ \mu_{sbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} \end{aligned} \tag{1}$$

where $\mathbf{X}$ is the 3-column design matrix and β is a vector including the intercept β_0 and the coefficients β_{age} and β_{age2} of the fixed effect variables (age and age^2, respectively). μ_{sbpg} is the vector of varying intercepts. *Model M2* was based on *model M1* but including an interaction term

Table 1: Variables included in the height-growth models. Baseline *models M1* and *M2* separate the genetic and plastic components of height-growth variation via varying intercepts relying exclusively on the common garden design. Explanatory models (*models M3* to *M6*) test different hypotheses regarding the potential drivers underlying height-growth variation. Predictive models (*models M7* to *M12*) are used to compare the predictions on new observations and provenances when combining or including separately genomic and climatic drivers of height-growth variation. The provenance climate-of-origin is evaluated using the precipitation of the driest month, *min.pre*, and the maximum temperature of the warmest month, *max.temp*. gPEAs and rPEAs correspond to the counts of height-associated positive-effect alleles, selected either globally (across all common gardens) or regionally (in specific common gardens). The provenance climate-of-origin and the PEAs were included in the predictive models with site-specific slopes. All models also contained the age effect, not shown in the table.

Variables	**Baseline**		**Explanatory models**				**Predictive models**					
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12
Site/Block	×	×	×	×	×	×	×	×	×	×	×	×
Provenance	×	×	×	×	×	×						
Genotype	×	×	×	×		×						
Site × Provenance		×										
Climatic similarity among sites			×	×	×	×						
Proportion belonging to each gene pool				×	×	×	×	×	×			
Gene pool-specific genetic variance					×							
Climatic similarity among provenances						×						
Provenance climate-of-origin							×	×		×		
Global PEAs (gPEAs)							×				×	
Regional PEAs (rPEAs)								×				×

between provenance and site (S_sP_p). We also performed a model without the genetic component (called *M0*) whose outputs are reported in the Supplementary Information.

2.4.2 Explanatory *models M3* to *M6*: potential drivers underlying height-growth variation

In *model M3*, we hypothesized that the plastic component of height growth was influenced by the climatic similarity among test sites during the year preceding the measurements. This model can be expressed with the same likelihood as *M1* but with the vector of varying intercepts equal to:

$$\begin{aligned} \mu_{isbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} \\ cs_{is} &\sim \mathcal{N}(0, \Omega\, \sigma^2_{cs_{is}}) \end{aligned} \tag{2}$$

where Ω is the covariance matrix describing the climatic similarity between test sites s during the year i preceding the measurements (fig. S6) and cs_{is} are varying intercepts associated with the climatic conditions in each test site s during the year i. In *M3*, the plastic component was partitioned between the regression on the climatic covariates (cs_{is}) and the deviations related to

block and site effects due to the local environmental conditions that are not accounted for by the selected climatic covariates.

In *models M4*, *M5* and *M6*, we investigated the drivers of the genetic component of height growth. In *M4*, we hypothesized that the genetic component was influenced by the proportion belonging to each gene pool j. *M5* extends *M4* by estimating different total genetic variances in each gene pool while accounting for admixture among gene pools, following Muff et al. (2019). Equations for *M4* and *M5* can be found in section 4 of the Supplementary Information. In *M6*, we hypothesized that populations are genetically adapted to the climatic conditions in which they evolved. Thus, we quantified the association between height growth and the climatic similarity among provenances, while still accounting for the gene pool assignment, such as:

$$\begin{aligned} \mu_{ijsbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + cp_p + \sum_{j=1}^{6} q_{gj} g_j \\ cp_p &\sim \mathcal{N}(0, \Phi\, \sigma^2_{cp_p}) \end{aligned} \tag{3}$$

where q_{gj} corresponds to the proportion belonging of each genotype g to the gene pool j, g_j is the mean relative contribution of gene pool j to height growth, Φ is the covariance matrix describing the climatic similarity between provenances p (fig. S9) and cp_p are varying intercepts associated with the climate in each provenance p. Therefore, in *M6*, the genetic component was partitioned among the regression on the climatic covariates (cp_p), the gene pool covariates (g_j), and the deviations related to the genotype ($G_{g(p)}$) and provenance (P_p) effects (resulting, for example, from adaptation to environmental variables not measured in our study).

2.4.3 Predictive *models M7* to *M12*: combining climatic and genomic information to improve predictions

In this last set of models, we replaced the provenance and genotype intercepts by different potential drivers of height-growth variation that do not rely directly on the common garden design, namely the gene pool assignment (as in *M4*), two variables describing the climate in the provenance locations (*min.pre* the precipitation of the driest month and *max.temp* the maximum temperature of the warmest month) and either global or regional PEAs. This allowed us to determine whether these potential drivers were able to predict the height-growth genetic component as accurately as the provenance and genotype intercepts (i.e. the variables relying directly on the common garden design). In *models M7* and *M8*, the potential predictors were all included together in the models to quantify their predictive performance conditionally to the other predictors, and

were expressed as follows (here for *M7*):

$$\mu_{jsbpg} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj} g_j + \beta_{min.pre,s} min.pre_p + \beta_{max.temp,s} max.temp_p + \beta_{gPEA,s} gPEA_g \tag{4}$$

where $min.pre_p$ and $max.temp_p$ are the climatic variables in the provenance locations, $\beta_{min.pre,s}$ and $\beta_{max.temp,s}$ their site-specific slopes, $gPEA_g$ the counts of global PEAs and $\beta_{gPEA,s}$ its site-specific slopes. *M8* is identical to *M7*, except that the counts of gPEAs were replaced by counts of rPEAs (i.e. regionally-selected alleles, with positive effects in specific geographical regions/environments). We also performed models in which the potential predictors were included individually to determine their specific predictive performance: the gene pool assignment in *M9*, the provenance climate-of-origin in *M10* and the counts of gPEAs and rPEAs, in *M11* and *M12*, respectively.

All models were inferred in a Bayesian framework as this approach better handles unbalanced and multilevel designs (Clark 2005) and also to better propagate sources of uncertainty from data and parameter values into the estimates (de Villemereuil 2019). Priors used in the models were weakly informative and are provided in section 4.2 of the Supplementary Information. To build the models, we used the *brms* package (Bürkner 2017), based on the no-U-turn sampler algorithm. Models were run with four chains and between 2,000 and 3,000 iterations per chain depending on the models (including 1,000 warm-up samples not used for the inference). All analyses were undertaken in R version 3.6.3 (R Core Team 2020) and scripts are available at https://github.com/JulietteArchambeau/HeightPinpinClonapin.

2.5 Comparing model goodness-of-fit and predictive ability

Three partitions of the data (P1, P2 and P3) were used to evaluate model goodness-of-fit (i.e. in-sample explanatory power, using training datasets) and predictive ability (out-of-sample predictive power, using test datasets). In P1, we aimed to predict new observations, an observation being a height-growth measurement in a given year on one individual. P1 corresponds to a random split of the data between 75% of observations used to fit the models (the training dataset of 24,840 observations) and 25% of observations used to evaluate model predictions (the test dataset of 8,281 observations). Notice that the test dataset of the P1 partition was not totally independent from the training dataset as it belongs to the same genotypes/provenances and blocks/sites. In P2 and P3, we aimed to predict new provenances. P2 corresponds to a random split between a training dataset of 28 provenances and a test dataset containing the remaining 6 provenances. P3 corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool

(i.e. northern Africa, south-eastern Spain and Corsican gene pools; see section 6.3 of the Supplementary Information for details). Therefore, the test datasets of the P2 and P3 partitions represent fully independent sets of provenances.

To evaluate the model goodness-of-fit, we calculated the in-sample (in the training dataset) proportion of the variance explained by each model m in each common garden s, conditional on the age effect, such as: $\mathcal{R}^2_{ms}|age = (V_{\text{pred}_{ms}} - V_{\text{age}_{2s}})/(V_{\text{y}_s} - V_{\text{age}_{2s}})$, where $V_{\text{pred}_{ms}}$ is the variance of the modeled predictive means from model m in site s, V_{y_s} the phenotypic variance in the site s and $V_{\text{age}_{2s}}$ the variance explained by the age effect in the model *M2* in site s. We used V_{age_2} of model *M2* and not of model m because the variance predicted by the different fixed effects of some of the models (*M7* to *M12*) could not be properly separated. Moreover, as *M2* is the model with the highest predictive ability among the models relying only on the common garden design (Table S4), it constitutes an adequate baseline for model comparison. In addition, for baseline models *M1* and *M2*, we also calculated the in-sample proportion of the variance explained by the different model components (i.e. genetic, environment and genetic $\times$ environment) conditional on the age effect, e.g. for the genetic component in *M1*: $\mathcal{R}^2_{1,gen}|age = (V_{\text{pred}_{1,gen}} - V_{\text{age}_1})/(V_{\text{y}} - V_{\text{age}_1})$ where $V_{\text{pred}_{1,gen}}$ is the variance explained by the genetic component (including the provenance and clone effects) in *M1*, V_{y} the phenotypic variance and V_{age_1} the variance explained by the age effect in *M1*.

Finally, to evaluate the model predictive ability, we calculated the out-of-sample (in the test dataset) proportion of the variance predicted by each model m in each common garden s conditional on the age effect, that we called *prediction* $\mathcal{R}^2_{ms}|age$. Details about calculating *prediction* $\mathcal{R}^2_{ms}|age$ and some supplementary indexes used for model comparison are presented in section 5 of the Supplementary Information.

3 Results

3.1 Underlying drivers of height-growth variation

In this part, we disentangled the different components of height-growth variation and provided insights on their underlying drivers. Baseline and explanatory models (i.e. *models M1* to *M6*) explained ∼81.5% of height-growth variation, including 57% due to the age effect (Table S4). Based on *M1*, ∼47% (45-48% CIs) of the variation that was not explained by the age effect (i.e. deviating from the growth trajectory) came from the plastic component, ∼11% (11-12% CIs) from the genetic component and ∼43% (42-44% CIs) remained unexplained (fig. 2A & Table S5). In *M2* (same model as *M1* but adding the provenance-by-site interaction), the proportion of variance explained by the provenance-by-site interaction was not different from zero (Table S5). Therefore, we mostly interpret parameter estimates of *M1* (fig. 3), whose results are very similar to *M2*, but with smaller

credible intervals (Tables S15 & S18). The plastic component was largely driven by the variance among sites (σ^2_S), with very little contribution of the variance among blocks (σ^2_B; S15). Trees grew the least in Madrid and the most in Asturias (fig. 3 & Table S16). The genetic component was equally attributed to the variance among provenances (σ^2_P) and genotypes (σ^2_G; Table S15), with the average height of the provenances appearing to be influenced by their belonging to particular gene pools (fig. 3; and more details in section 6.1.1 of the Supplementary Information).

Based on *M3*, the plastic component of height-growth came only marginally from the variance associated with climate similarity among sites, which was more than five times lower than the variance associated with site intercepts (fig. 2B & Table S19). However, *M3* may be unable to separate the effect of these two components (see section 6.1.2 in the Supplementary Information). Indeed, when estimating the effect of the climate similarity among sites in a model that did not include varying intercepts for the sites, we found that height growth was positively associated with the climatic conditions in Bordeaux and Asturias, and negatively with those in Madrid and Cáceres, the two Mediterranean sites, and to a lesser extent also in Portugal (Table S24).

Based on *M6*, the genetic component of height growth was mostly determined by the climatic similarity among provenances and to a lesser extent by the gene pool assignment (fig. 2C & Table S29). However, the effects of the gene pools and climatic similarity among provenances were partially confounded, so that the association between height growth and the gene pools was stronger when the climatic similarity among provenances was not included in the models (i.e. *model M4*; Table S25). Populations from climatic regions neighboring the Atlantic Ocean, and mainly belonging to the French and Iberian Atlantic gene pools, were generally the tallest (e.g. CAD, SIE, PUE, LAM and CAS in northwestern Spain; all provenances along the French Atlantic coast; Figs. 2 & 3). Interestingly, the Leiria (LEI) provenance, which has a strong Iberian Atlantic component (Table S3) and had the highest climate intercept estimate (similar to that of the French Atlantic provenances; fig. 2C), was not among the tallest provenances (fig. 3), probably due to its mixed ancestry with the central Spain gene pool (Table S3). Also, the Corsican provenances showed contrasted climate intercepts (fig. 2), with a positive influence on height growth for Pinia (PIA) but not for Pineta (PIE), located under more Mediterranean conditions, which could explain their large differences in height growth (fig. 3). Finally, the four provenances from south-eastern Spain and northern Africa gene pools, under harsh Mediterranean climates, showed all negative climate intercepts (fig. 2). Noticeably, the total genetic variance of the Iberian and French Atlantic gene pools were likely to be lower than that of the Corsican and south-eastern Spain gene pools, and to a lesser extent the central Spain gene pool, thus resulting in gene pool-specific heritabilities (*model M5*; Table S28 and fig. S13A).

Figure 2. Understanding the genetic and plastic bases of height-growth variation and their potential underlying drivers. A) shows the variance partitioning conditional on age from *model M1* in the P1 partition. B) displays the partitioning of the plastic (i.e. environment) component in *model M3* among the intercepts of the sites (common gardens) (S_s) and the intercepts associated with the climatic similarity among sites during the year preceding the measurements (cs_{is}). C) displays the partitioning of the genetic component in model *M6* among the intercepts of the provenances (P_p), the intercepts associated with the climatic similarity among provenances (cp_p) and the intercepts of the the gene pools (g_j). The median and 0.95 credible intervals shown in B) and C) were obtained by fitting the *models M3* and *M6* on the P1 partition. Provenance names can be found in Table S2.

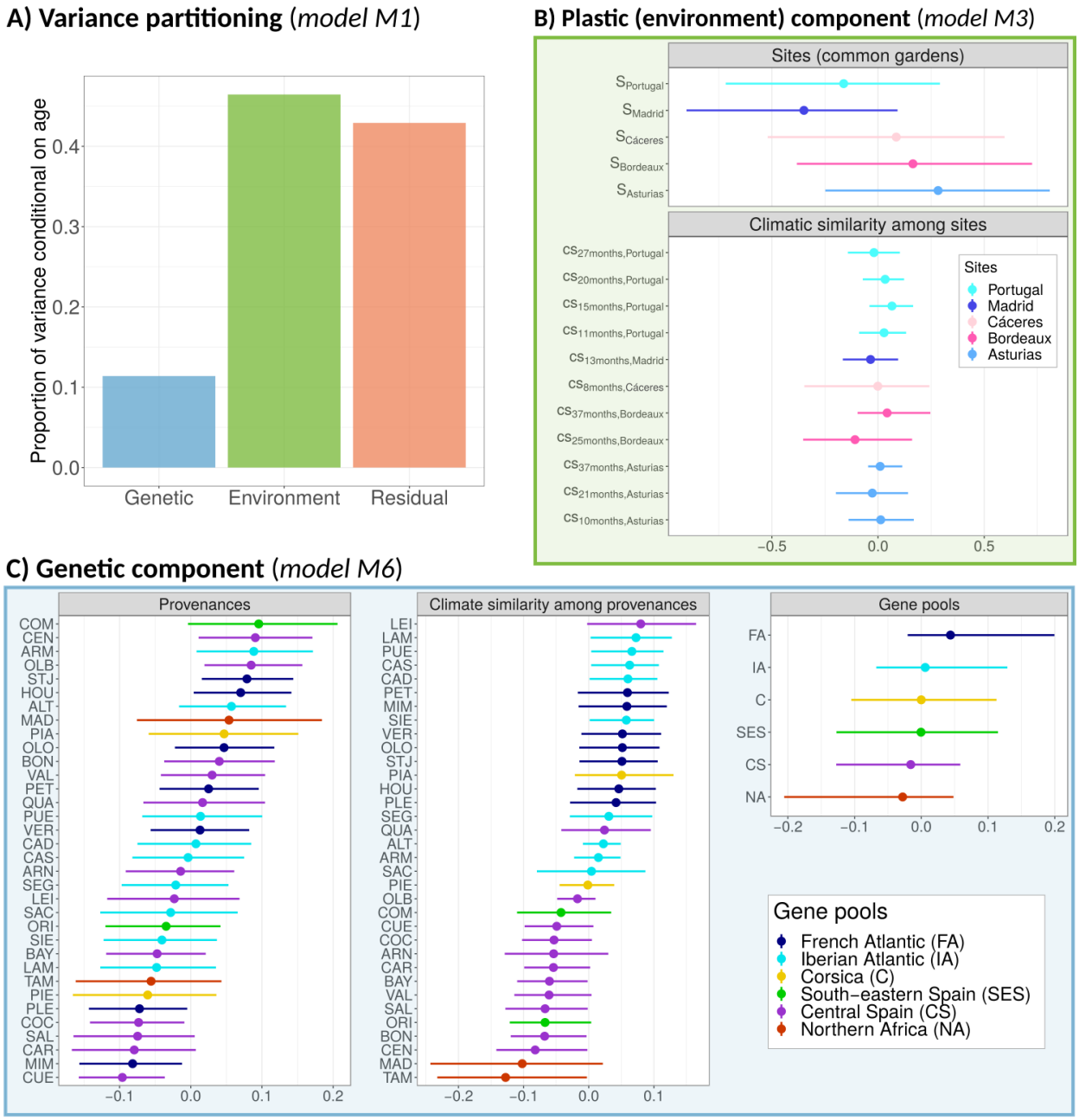
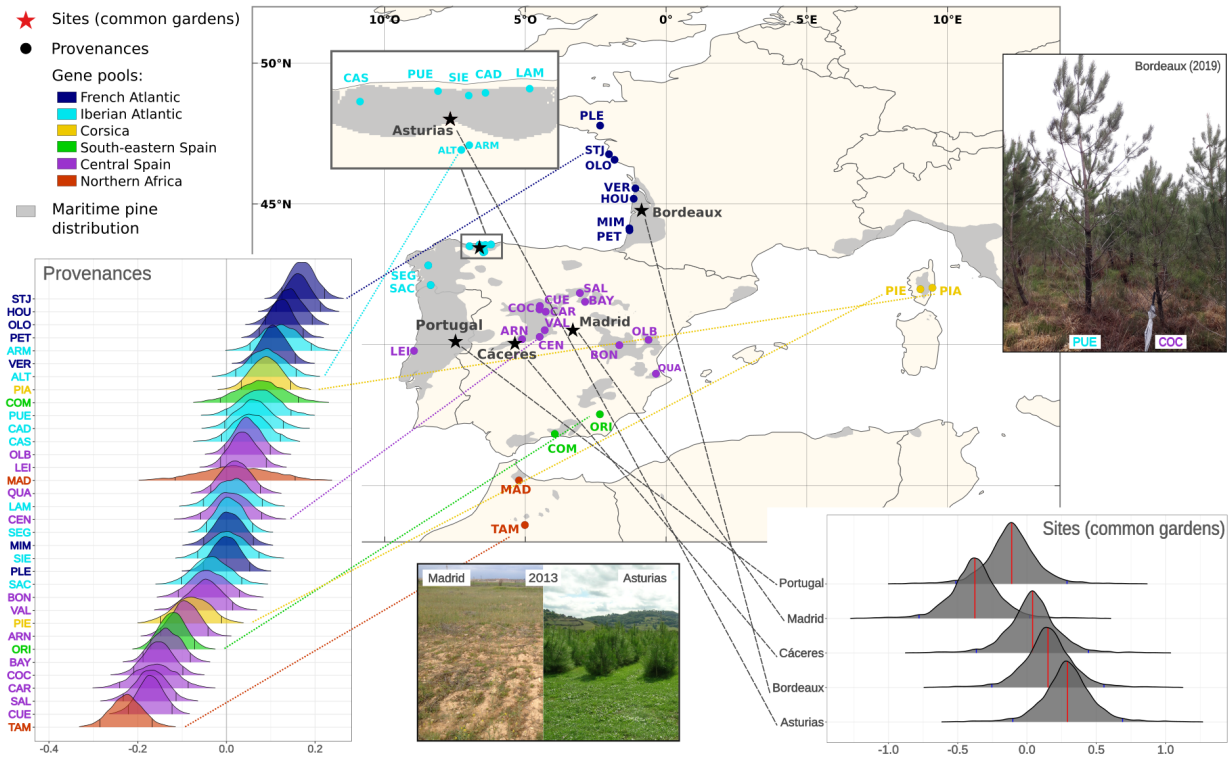


Figure 3. Posterior distributions of the site and provenance intercepts (S_s and P_p) in *model M1* on a map representation. Provenances are colored according to the main gene pool they belong to. The exact values of the median, standard deviation and 0.95 credible interval of the posterior distributions of the site and provenance intercepts are shown in Tables S16 and S17, respectively. The top right picture shows the height difference in 2019 between one tree from Coca in central Spain (COC) and another from Puerto de Vega in the Iberian Atlantic region (PUE) growing next to each other in the Bordeaux common garden. The bottom picture shows the height difference between the trees growing in Madrid and Asturias, under highly contrasted environments, three years after plantation (2013). Provenance names can be found in Table S2.



3.2 Improved prediction of new observations and provenances by combining climatic and genomic data

In this part, we compared the baseline model *M2* (relying exclusively on the common garden design) to the predictive models that either combine genomic and climatic drivers of height-growth variation (i.e. *models M7* and *M8*) or include each driver separately (i.e. *models M9* to *M12*). Models combining genomic and climatic data generally explained in-sample variation almost as well as *M2*, and sometimes even better; e.g. *model M8* (which includes regional PEAs, rPEAs) in the Mediterranean sites (Madrid and Cáceres) (fig. S10). Models including each driver of height-growth variation separately had a lower goodness-of-fit (for all common gardens) than both *M2* and the models combining the genomic and climatic data, except for *M12* (the model including only rPEAs), which explained in-sample variation almost as well as *M2* and even better than *M7*

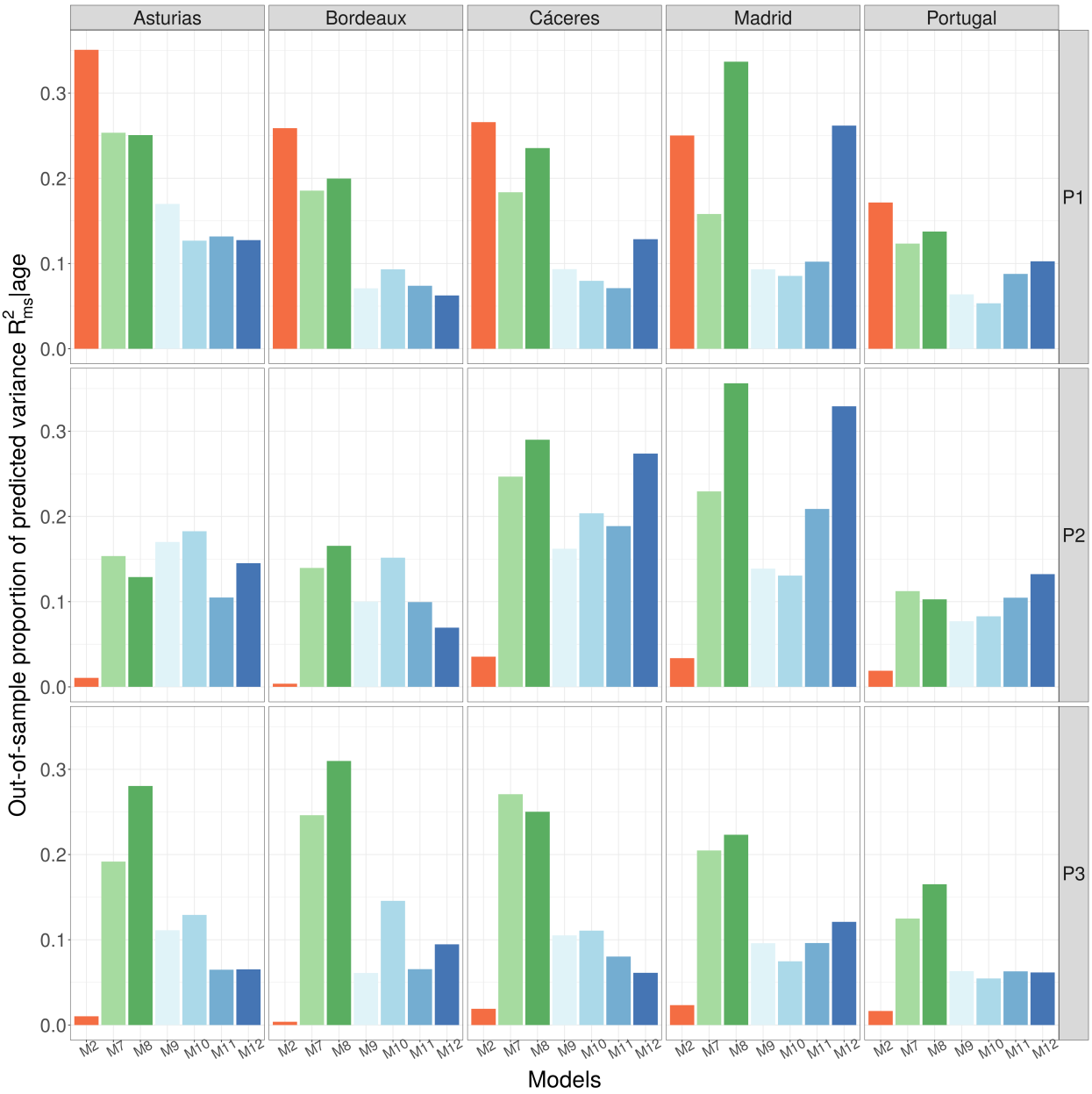
in Madrid (fig. S10).

Model differences in their predictive ability on new observations (observations not used to fit the models; test dataset of the P1 partition) showed similar patterns than for the goodness-of-fit (Table 4), which was expected as the new observations were sampled among the same provenances and genotypes. However, importantly, models combining genomic and climatic data provided much better predictions of height-growth on new provenances (provenances not used to fit the models; test datasets of the P2 and P3 partitions) than did *M2*, with *M8* having a better predictive ability than *M7* in the Mediterranean sites in the P2 partition and in the Atlantic sites in the P3 partition (Table 4). Models including each driver of height-growth variation separately had also a higher predictive ability on new provenances than *M2*, albeit lower than models combining genomic and climatic data, except *model M12* that showed a higher predictive ability than *M7* in the Mediterranean sites in the P2 partition (Table 4). In *model M12*, one standard deviation increase in rPEAs was associated, on average, with 19.0% increase in height in Madrid, 12.7% in Cáceres, 13.0% in Portugal, 10.4% in Asturias and 9.6% in Bordeaux (section 6.4 of the Supplementary Information). More details on model comparisons are given in section 5 of the Supplementary Information.

4 Discussion

We combined genomic, climatic and phenotypic data from five common gardens and 34 provenances of maritime pine (over 30,000 observations) to predict range-wide variation in height growth, a key adaptive trait in forest trees. The plastic component explained the largest part of the deviation from the mean height-growth trajectory ($\sim$47%), probably due to multiple (confounded) environmental factors, including climate. The genetic component explained $\sim$11% of the deviation from the mean height-growth trajectory and was mainly associated with the provenance climate-of-origin (a proxy of adaptation to climate), whose effect was partially confounded with the proportion belonging to distinct gene pools (a proxy for population demographic history and genetic drift, probably reflecting also the different selective histories of the gene pools). Importantly, we showed that models combining climatic drivers of adaptation, gene pool assignment and counts of height-associated positive-effect alleles (PEAs) captured well the genetic component underlying height-growth variation. They also better predicted height growth of new provenances than models relying exclusively on the common garden design or models including separately climatic and genomic information (e.g. the widely used climate-based population response functions). Interestingly, PEAs that show a regional association with height growth (rPEAs) had a higher predictive ability than PEAs identified globally across the species range (gPEAs). These results pave the way towards integrating genomics into large-scale predictive models of quantitative trait variation.

Figure 4. Model predictive ability on new observations (P1 partition) or new provenances (P2 and P3 partitions) based on the out-of-sample proportion of predicted variance conditional on the age effect (*prediction* $\mathcal{R}^2_{ms}|age$) in the test datasets (data not used to fit the models). In the P1 partition, the training dataset was obtained by randomly sampling 75% of the observations and the test dataset contains the remaining 25% observations. In the P2 partition, the training dataset was obtained by randomly sampling 28 provenances and the test dataset contains the remaining 6 provenances. The P3 partition corresponds to a non-random split between a training dataset of 28 provenances and a test dataset containing 6 provenances with at least one provenance from each under-represented gene pool. The exact values of the *prediction* $\mathcal{R}^2_{ms}|age$ estimates and their associated credible intervals can be found in Tables S4 (P1 partition), S9 (P2 partition) and S12 (P3 partition).



4.1 Predominant role of height-growth plasticity

Plants are known for their remarkable phenotypic plasticity to changing environments (Bradshaw 1965). In long-lived forest trees, the plastic component of quantitative trait variation estimated based on the common garden design is generally higher than the genetic component (Franks et al. 2014, Benito Garzón et al. 2019), e.g. in maritime pine (Chambel et al. 2007, Corcuera et al. 2010, de la Mata et al. 2012, Vizcaíno-Palomar et al. 2020). This plastic component is also generally associated with the climatic conditions experienced by the trees (Franks et al. 2014, Benito Garzón et al. 2019), allowing them to overcome changing climate up to a certain threshold (Matesanz et al. 2010, Nicotra et al. 2010, Valladares et al. 2014a). In our study, the plastic component of height growth was largely higher than the genetic component (fig. 2) and, although climate plays a role, was likely to be driven by multiple and interacting drivers including the biotic environment, soil quality, and other factors not considered in our study.

Plants also present an important genetic variation in plasticity (i.e. the genotype-by-environment interaction, G×E; Des Marais et al. 2013, Sork 2018), often approximated by the family or provenance-by-site interaction in forest tree common gardens, as is the case in our study. G×E is particularly prevalent for growth traits in trees (Li et al. 2017), as already shown in maritime pine (Alía et al. 1997, Corcuera et al. 2010, Correia et al. 2010, de la Mata et al. 2012; but see Chambel et al. (2007) where no provenance-specific responses were observed under two different watering regimes). In our study, provenance-by-site interaction was only weakly associated with height growth and the proportion of variance it explained was not different from zero (*model M2*; Table S5). Previous work in the context of tree breeding argued that G×E may hinder model transferability across sites and populations (Resende Jr et al. 2012, Resende et al. 2012). In maritime pine, our results suggest that large-scale predictions of height-growth variation will be only marginally impacted by not accounting for provenance-by-environment interaction. However, further work is necessary to assess the importance of the genetic variation of plasticity at the genotype level.

4.2 Potential drivers underlying height-growth genetic component

Our study shows that the height-growth genetic component in maritime pine is mostly associated with adaptation to climate, whose effect is partially confounded with the effect of gene pool assignment, reflecting both adaptive (different selective histories) and neutral processes (population demographic history and genetic drift) (fig. 2; see also Jaramillo-Correa et al. 2015). For example, the higher growth of most provenances from the French Atlantic gene pool (known for their high growth under a wide range of conditions, including Mediterranean sites in our study; see also Alía et al. 1997, Corcuera et al. 2010, de la Mata et al. 2012) was both associated with the provenance climate-of-origin and the gene pool assignment. As another example, in the northern Africa gene pool, the Madisouka (MAD) provenance was taller than the Tamrabta

(TAM) provenance, which could be both explained by its noticeable ancestry proportion (23.3%) from the south-eastern Spanish gene pool (Jaramillo-Correa et al. 2015) or its adaptation to lower elevation (300 m lower than TAM). As a last example, the Leiria (LEI) provenance grew well in Asturias and Bordeaux as was the case for French Atlantic provenances (that share similar climates) but unlike them, it did not maintain growth in drier and warmer sites, probably due to a different genetic background (this provenance has a strong central Spain gene pool component; Table S3). Nevertheless, in contrast to the three examples above, for some provenances, the effects of the gene pool assignment and adaptation to climate on height growth could be clearly separated. This was the case, for example, for the Corsican provenances: the higher growth of Pinia (PIA) than Pineta (PIE) can only be explained by adaptation to different environmental conditions (and in particular climate), as both belong to the same gene pool. Indeed PIA is at the sea level under a climate similar to that of provenances from Central and south-eastern Spain whereas PIE is located at an altitude of 750 m a.s.l. in the mountains under a climate similar to that of the Atlantic provenances (fig. S9). These different adaptations within a same gene pool calls for a more targeted investigation of the Corsican gene pool. More generally, a $Q_{ST} - F_{ST}$ analysis supported adaptive differentiation of height growth in maritime pine (see details in section 7 of the Supplementary Information).

The entanglement of the effect of climate adaptation and gene pool assignment to explain the genetic component of height-growth variation may partly stem from the distinct selective histories experienced in different parts of maritime pine range, despite gene pools being identified using genetic markers considered neutral (Jaramillo-Correa et al. 2015). This is supported by the estimation of gene pool-specific heritabilities in our study (*model M5*): the Corsican gene pool, and to a lesser extent the south-eastern Spain gene pool, have higher heritabilities than the French and Iberian Atlantic gene pools (Fig. S13; and see section 6.1.3 for a potential explanation of this pattern).

Overall, maritime pine proved to be a particularly suitable model species to study the joint influence of genetic neutral (population demographic history, genetic drift) and adaptive (climate adaptation) processes on quantitative traits. Further work on provenances that have different demographic histories but are exposed to similar climates (e.g. the LEI provenance and provenances from the Atlantic gene pools) would be relevant for understanding how a given genetic background guides population adaptation. Conversely, targeting provenances that have a similar demographic history but are found in highly contrasted environments (e.g. the Corsican provenances) would be valuable to identify signatures of adaptation while avoiding common issues due to confounding population structure (Berg et al. 2019, Sella and N. H. Barton 2019, Sohail et al. 2019). Likewise, investigating trait genetic architecture will also help better understand how adaptive and neutral processes have shaped the genotype-phenotype map and how this will influence future responses to selection (e.g. Kardos and Luikart 2021; see de Miguel et al. 2020 for maritime pine). Finally, it would also be critical to consider drivers of adaptation other than climate, such as resistance to pathogens or other biotic-related traits.

4.3 Towards integrating genomics into population response functions

Anticipating how provenances will grow in new environments is key to guide forest conservation strategies and population translocations to compensate for rapid climate change (Aitken and Whitlock 2013). To date, population response functions based on the climate in the provenance location have been the most widely used method for anticipating trait values when transplanting provenances in new environments (Rehfeldt et al. 1999, 2003, O'Neill et al. 2008, Wang et al. 2010, Pedlar and McKenney 2017, Rehfeldt et al. 2018, Fréjaville et al. 2020). Genome-informed predictive modeling of key adaptive traits is highly promising as it may provide a mean to further integrate adaptive or neutral genetic variation in the predictions, and to consider intraspecific variability at a finer scale than current models, thus gaining in prediction accuracy (Holliday et al. 2017). In valley oak, Browne et al. (2019) used genomic estimated breeding values (GEBVs; sum of the marker predicted effects, also known as polygenic scores) to identify genotypes that will grow faster under future climates. In lodgepole pine, Mahony et al. (2020) showed that phenotype-associated positive-effect alleles (PEAs, as used in our study) can predict phenotypic traits (e.g. cold injury) as well as climatic or geographical variables. In our study, we investigated whether including genomic information related to past demographic and selective processes resulting in distinct gene pools and counts of trait-associated alleles could improve range-wide height-growth predictions in maritime pine. Models combining climatic conditions in the provenance location, gene pool assignment, and PEAs captured most of the genetic component of height-growth variation (see fig. S10) and better predicted height growth of new provenances, compared to models relying exclusively on the common garden design or models including separately climatic or genomic information (see fig. 4). This suggests that range-wide trait prediction would benefit from jointly considering different sources of information (i.e. climatic and genomic), even though they may have overlapping effects (e.g. confounded effects of provenance climate-of-origin and gene pool assignment), as it may help to embrace the complexity and multidimensionality of the genetic component underlying quantitative traits. Noticeably, regional PEAs were generally better predictors of height growth in new provenances than gene pool assignment or provenance climate-of-origin as, when they were included alone in the models, they made better predictions in the driest common gardens (Madrid, Cáceres and Portugal) and similar ones to models combining multiple drivers of height growth variation in all common gardens except Bordeaux (P2 partition in fig. 4). Although this highlights the major role that trait-associated alleles identified using GWAS may play in predictive modeling, predicting traits of new provenances depends also on the number of provenances used to fit the models and the strength of the genetic relationship among them (Resende et al. 2012, Jarquín et al. 2014, Moghaddar et al. 2014, Hidalgo et al. 2016). This was reflected in our study by better predictive ability on new provenances in the P2 partition (random) compared to the P3 partition (containing provenances from underrepresented gene pools) for models including climatic and genomic information separately but not for models considering both jointly (fig. 4). Thus combining multiple sources of information may also be particularly relevant for predicting traits in marginal or difficult-to-access populations, as they

normally belong to underrepresented geographical areas/gene pools in ecological and genetic studies.

The high predictive ability of PEAs, both alone and combined with climatic and gene pool information, was somehow unexpected given the sparse genomic sampling in our study: 5,165 SNPs to cover the 28 Gbp maritime pine genome (Zonneveld 2012). Indeed, conifers have particularly huge genomes, generally ranging from 18 to 35 Gbp (Mackay et al. 2012) and thus rendering the current cost of whole-genome resequencing prohibitive (Holliday et al. 2017). Targeted genotyping approaches, such as the one used in the present study, select candidate genes based on previous population and functional studies, thus allowing to include potential targets of selection and climate adaptation, but probably inducing an ascertainment bias (Jaramillo-Correa et al. 2015). However, as height is a particularly polygenic trait (degree of polygenicity estimated at $\sim$7% in de Miguel et al. 2020), we were able to identify a considerable number of PEAs despite the weak genome coverage of our study. Further genomic sampling would be highly valuable to capture the polygenic architecture of height more broadly, turning PEAs into much better predictors than the provenance climate-of-origin or the gene pool assignment, and ultimately making climatic data redundant, at least for main range populations (see above for marginal populations). This would also allow to characterize the genetic variation within provenances more precisely, thereby increasing the estimation accuracy and reducing the residual variance. Similar to Mahony et al. (2020) and MacLachlan et al. (2021) who selected the positive-effect alleles as the 1% of SNPs that showed the strongest association with phenotypes (estimated via a GWAS performed on 18,525 SNPs), we used PEA counts instead of the more commonly used polygenic scores (Pritchard et al. 2010, Browne et al. 2019, Fuller et al. 2020). Unlike polygenic scores, PEAs do not account for allele effect sizes, thus minimizing the circularity of the analysis (i.e. effect sizes that are estimated based on the same dataset as the one used for the models, only serve for PEAs identification) and potentially enhancing the prediction accuracy across genetic groups compared to polygenic scores. Indeed, low observed transferability of polygenic scores across genetic groups (Martin et al. 2017, N. Barton et al. 2019, Martin et al. 2019) may stem from varying effect sizes of "peripheral" alleles (i.e. alleles indirectly affecting the phenotype), as suggested in Mathieson 2021).

Although combining climatic and genomic information allowed us to capture most of the genetic component of height-growth variation (fig. S10), the residual variance remained high in our study. As already mentioned, this may be partly related to the models' difficulty in accounting for genetic variation within provenances, which might be improved by denser genomic sampling. However, this unexplained variance may also originate from developmental stochasticity, which can play an important role in explaining differences between individuals with the same genotype (Vogt 2015, Ballouz et al. 2019). Height growth may also be influenced by the correlative effects of other traits. For example, Stern et al. (2020) recently showed that variation in some human traits (hair color and educational attainment), previously thought to be under selection, can instead be explained by indirect selection via a correlated response to other traits. Therefore, multi-trait

models may be the next necessary step to improve our understanding and predictive ability of quantitative trait variation at large geographical scales (e.g. Csilléry et al. 2020).

A last noticeable results was that rPEAs (positive-effect alleles identified in specific geographical regions, i.e. particular environments) had generally a higher predictive ability than gPEAs (positive-effect alleles identified range-wide) (Figure 4). Interestingly, only a small proportion of rPEAs were shared among geographical regions in our study (20% shared between the Iberian and French Atlantic regions, 12% between the French Atlantic and Mediterranean regions, and 24% between the Iberian Atlantic and Mediterranean regions; Figure S2), although we cannot exclude that the proportion of shared rPEAs among regions is a function of the sample size (see details in the section 2.2 of the Supplementary Information). Moreover, those that were shared among different regions showed consistently similar effects across regions (e.g. positive effects in two or more regions rather than antagonist effects). This supports the predominance of conditional neutrality, i.e. alleles that are advantageous in some environments and neutral in others, over antagonistic pleiotropy, i.e. alleles that are advantageous in some environments and disadvantageous in others (Tiffin and Ross-Ibarra 2014). Such pattern has already been reported in plants (Prunier et al. 2012, Anderson et al. 2013). Our results show that, despite a high stability in the level of polygenicity for height between the Atlantic and Mediterranean regions (de Miguel et al. 2020), height-growth variation in Mediterranean sites is unlikely to be affected by the same loci as in the other regions, probably as a result of genetic divergence in separated southern refugia during the last glaciation. Overall, identifying positive-effect alleles for different geographical regions separately has the potential to greatly improve the predictive ability of the models, but at the cost of reducing GWAS power (due to lower sample size than in global, wide-range analyses).

Finally, caution has to be taken when generalizing our results to older trees as the drivers of height growth in young trees may differ from that of adult trees. For example, G×E on tree height can be age-dependant (Gwaze et al. 2001, Zas et al. 2003, Rehfeldt et al. 2018) and the plastic component may be higher in younger trees, especially in maritime pine (Vizcaíno-Palomar et al. 2020). Nevertheless, a recent measurement in the Bordeaux common garden (2018) showed a high correlation between young saplings and 10-year old trees for height (Pearson's correlation coefficient of 0.893 based on height BLUPs; see de Miguel et al. 2020 for details on BLUP estimation). Moreover, our study remains indicative of how trees respond to varying environmental conditions during establishment and early-growing stages, a critical phase where most mortality (i.e. selection) is expected to take place (Postma and Ågren 2016). In addition to ontogenic effects, high mortality in the Mediterranean common gardens (Cáceres and Madrid), after a marked summer drought, may have biased estimates of some parameters of interest. Indeed, if this environmental filtering was not independent of tree height, it could have resulted in an underestimation of the genetic variance. Nonetheless, height distributions in Cáceres and Madrid were only slightly right-skewed, suggesting uniform selection across height classes (fig. S21), and thus no bias due to high mortality in these common gardens.

5 Conclusion

The present study connects climate-based population response functions that have been extensively used in predictive models for forest trees (Rehfeldt et al. 1999, 2003, Wang et al. 2010, Leites et al. 2012a) with recent genomic approaches to investigate the potential drivers behind the genetic and plastic components of height-growth variation and predict how provenances will grow when transplanted into new climates. The integration of genomic data into range-wide predictive models is in its infancy and still lacks a well-established framework, especially for non-model species such as forest trees. We showed that combining climatic and genomic information (i.e. provenance climate-of-origin, gene pool assignment and trait-associated positive-effect allele counts) can improve model predictions for a highly polygenic adaptive trait such as height growth, despite sparse genomic sampling. Further genomic sampling may help to improve the accuracy of the estimates, notably through improved characterization of within-provenance genetic variation. Moreover, comparative studies between maritime pine and more continuously distributed species (e.g. Scots pine; Alberto et al. 2013) and/or living under stronger climatic limitations, would be highly valuable to determine whether our findings can be generalized to species with contrasted population demographic and selective history. Finally, our study focuses specifically on the height-growth genetic component of standing populations, but considering evolutionary processes (e.g. genetic drift in small populations, extreme selection events, etc.) into the predictions would be necessary to anticipate the response of future forest tree generations to changing climatic conditions and thus provide a much-needed longer-term vision (Waldvogel et al. 2020)

6 Acknowledgements

We thank A. Saldaña, F. del Caño, E. Ballesteros and D. Barba (INIA) and the 'Unité Expérimentale Forêt Pierroton' (UEFP, INRAE; doi:10.15454/1.5483264699193726E12) for field assistance (plantation and measurements). Data used in this research are part of the Spanish Network of Genetic Trials (GENFORED, http://www.genfored.es). We thank all persons and institutions linked to the establishment and maintenance of field trials used in this study. We are also very grateful to Ricardo Alía who contributed to the design and establishment of the CLONAPIN network and provided comments on the manuscript, and to Benjamin Brachi, Thibault Poiret, Andrew J. Eckert and one anonymous reviewer who provided constructive and valuable comments on a previous version of the manuscript. Thanks are extended to Juan Majada for initiating and supervising the establishment of the CLONAPIN network. JA was funded by the University of Bordeaux (ministerial grant). This study was funded by the Spanish Ministry of Economy and Competitiveness through projects RTA2010-00120-C02-02 (CLONAPIN), CGL2011-30182-C02-01 (AdapCon) and AGL2012-40151-C03-02 (FENOPIN). The study was also supported by the 'Initiative d'Excellence

(IdEx) de l'Université de Bordeaux - Chaires d'installation 2015' (EcoGenPin) and the European Union's Horizon 2020 research and innovation programme under grant agreement No 773383 (B4EST).

7 Author contributions

SCG-M and CP designed the experiment and supervised the curation of field data. MdM cleaned and formatted the phenotypic and genomic data, and produced the BLUPs used in GWAS. SCG-M and MdM ran the GWAS to identify the positive-effect alleles. SCG-M, MBG, JA and FB conceived the paper methodology. JA conducted the data analyses. SCG-M, MBG, JA and FB interpreted the results. JA led the writing of the manuscript. All authors contributed to the manuscript and gave final approval for publication.

8 Data and script availability

Data are publicly available. SNP data were deposited in the Dryad repository at http://dx.doi.org/10.5061/dryad.8d6k1. Height data have been deposited in GENFORED, the Spanish Network of Genetic Trials (http://www.genfored.es). Scripts are available at https://github.com/JulietteArchambeau/HeightPinpinClonapin.

References

Aitken, S. N. and Bemmels, J. B. (2015). Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications*, 9, 271– 290. https://doi.org/10.1111/eva.12293.

Aitken, S. N. and Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, 44, 367– 388. https://doi.org/10.1146/annurev-ecolsys-110512-135747.

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., and Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1, 95– 111. https://doi.org/10.1111/j.1752-4571.2007.00013.x.

Alberto, F., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., and Savolainen, O. (2013). Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology*, 19, 1645– 1661. https://doi.org/10.1111/gcb.12181.

Alía, R., Moro, J., and Denis, J. B. (1997). Performance of Pinus pinaster provenances in Spain: interpretation of the genotype by environment interaction. *Canadian Journal of Forest Research*, 27, 1548– 1559.

Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., and Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, 22, 699– 708. https://doi.org/10.1111/j.1365-294X.2012.05522.x.

Ballouz, S., Pena, M. T., Knight, F. M., Adams, L. B., and Gillis, J. A. (2019). The transcriptional legacy of developmental stochasticity. *bioRxiv*. https://doi.org/10.1101/2019.12.11.873265.

Barghi, N., Hermisson, J., and Schlötterer, C. (2020). Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics*, 21, 1– 13. https://doi.org/10.1038/s41576-020-0250-z.

Barton, N., Hermisson, J., and Nordborg, M. (2019). Why structure matters. *eLife*, 8, e45380. https://doi.org/10.7554/eLife.45380.

Benito Garzón, M., Robson, T. M., and Hampe, A. (2019). ΔTraitSDMs: species distribution models that account for local adaptation and phenotypic plasticity. *New Phytologist*, 222, 1757– 1765. https://doi.org/10.1111/nph.15716.

Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y., Boyle, E. A., Zhang, X., Racimo, F., Pritchard, J. K., and Coop, G. (2019). Reduced signal for polygenic adaptation of height in UK Biobank. *eLife*, 8, e39725. https://doi.org/10.7554/eLife.39725.

Bouvet, J.-M., Makouanzi, G., Cros, D., and Vigneron, P. (2016). Modeling additive and non-additive effects in a hybrid population using genome-wide genotyping: prediction accuracy implications. *Heredity*, 116, 146– 157. https://doi.org/10.1038/hdy.2015.78.

Bradshaw, A. D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants, Page 115– 155. *in* Caspari, E. W. and Thoday, J. M., editors. *Advances in Genetics*. Academic Press.

Browne, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., and Sork, V. L. (2019). Adaptational lag to temperature in valley oak (Quercus lobata) can be mitigated by genome-informed assisted gene flow. *Proceedings of the National Academy of Sciences*, 116, 25179– 25185. https://doi.org/10.1073/pnas.1908771116.

Bucci, G., González-Martínez, S. C., Provost, G. L., Plomion, C., Ribeiro, M. M., Sebastiani, F., Alía, R., and Vendramin, G. G. (2007). Range-wide phylogeography and gene zones in Pinus pinaster Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, 16, 2137– 2153. https://doi.org/10.1111/j.1365-294X.2007.03275.x.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80, 1– 28. https://doi.org/10.18637/jss.v080.i01.

Chambel, M. R., Climent, J., and Alía, R. (2007). Divergence among species and populations of Mediterranean pines in biomass allocation of seedlings grown under two watering regimes. *Annals of Forest Science*, 64, 87– 97. https://doi.org/10.1051/forest:2006092.

Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8, 2– 14. https://doi.org/10.1111/j.1461-0248.2004.00702.x.

Corcuera, L., Gil-Pelegrin, E., and Notivol, E. (2010). Phenotypic plasticity in Pinus pinaster δ13C: environment modulates genetic variation. *Annals of Forest Science*, 67, 812– 812. https://doi.org/10.1051/forest/2010048.

Correia, I., Alía, R., Yan, W., David, T., Aguiar, A., and Almeida, M. H. (2010). Genotype × Environment interactions in Pinus pinaster at age 10 in a multienvironment trial in

Portugal: a maximum likelihood approach. *Annals of Forest Science*, 67, 612– 612. https://doi.org/10.1051/forest/2010025.

Csilléry, K., Ovaskainen, O., Sperisen, C., Buchmann, N., Widmer, A., and Gugerli, F. (2020). Adaptation to local climate in multi-trait space: evidence from silver fir (Abies alba Mill.) populations across a heterogeneous environment. *Heredity*, 124, 77– 92. https://doi.org/10.1038/s41437-019-0240-0.

De la Mata, R., Voltas, J., and Zas, R. (2012). Phenotypic plasticity and climatic adaptation in an Atlantic maritime pine breeding population. *Annals of Forest Science*, 69, 477– 487. https://doi.org/10.1007/s13595-011-0173-0.

De Miguel, M., Rodríguez-Quilón, I., Heuertz, M., Hurel, A., Grivet, D., Jaramillo-Correa, J.-P., Vendramin, G. G., Plomion, C., Majada, J., Alía, R., Eckert, A. J., and González-Martínez, S. C. (2020). Polygenic adaptation and negative selection across traits, years and environments in a long-lived plant species (Pinus pinaster Ait.) *bioRxiv*. https://doi.org/10.1101/2020.03.02.974113.

De Villemereuil, P. (2019). On the relevance of Bayesian statistics and MCMC for animal models. *Journal of Animal Breeding and Genetics*, 136, 339– 340. https://doi.org/10.1111/jbg.12426.

Des Marais, D. L., Hernandez, K. M., and Juenger, T. E. (2013). Genotype-by-Environment Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic Environment. *Annual Review of Ecology, Evolution, and Systematics*, 44, 5– 29. https://doi.org/10.1146/annurev-ecolsys-110512-135806.

El-Dien, O. G., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., and El-Kassaby, Y. A. (2018). Multienvironment genomic variance decomposition analysis of open-pollinated Interior spruce (Picea glauca x engelmannii). *Molecular Breeding*, 38, 26. https://doi.org/10.1007/s11032-018-0784-3.

Eckert, A. J., Heerwaarden, J. van, Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., and Neale, D. B. (2010). Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (Pinus taeda L., Pinaceae). *Genetics*, 185, 969– 982. https://doi.org/10.1534/genetics.110.115543.

Exposito-Alonso, M., Burbano, H. A., Bossdorf, O., Nielsen, R., and Weigel, D. (2018). A map of climate change-driven natural selection in Arabidopsis thaliana. *bioRxiv*. https://doi.org/10.1101/321133.

Franklin, J., Serra-Diaz, J. M., Syphard, A. D., and Regan, H. M. (2016). Global change and terrestrial plant community dynamics. *Proceedings of the National Academy of Sciences*, 113, 3725– 3734. https://doi.org/10.1073/pnas.1519911113.

Franks, S. J., Weber, J. J., and Aitken, S. N. (2014). Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evolutionary Applications*, 7, 123– 139. https://doi.org/10.1111/eva.12112.

Fréjaville, T. and Benito Garzón, M. (2018). The EuMedClim Database: Yearly Climate Data (1901–2014) of 1 km Resolution Grids for Europe and the Mediterranean Basin. *Frontiers in Ecology and Evolution*, 6, 31. https://doi.org/10.3389/fevo.2018.00031.

Fréjaville, T., Vizcaíno-Palomar, N., Fady, B., Kremer, A., and Garzón, M. B. (2020). Range margin populations show high climate adaptation lags in European trees. *Global Change Biology*, 26, 484– 495. https://doi.org/10.1111/gcb.14881.

Fuller, Z. L., Mocellin, V. J. L., Morris, L. A., Cantin, N., Shepherd, J., Sarre, L., Peng, J., Liao, Y., Pickrell, J., Andolfatto, P., Matz, M., Bay, L. K., and Przeworski, M. (2020). Population

genetics of the coral Acropora millepora: Toward genomic prediction of bleaching. *Science*, 369. https://doi.org/10.1126/science.aba4674.

Grattapaglia, D., Plomion, C., Kirst, M., and Sederoff, R. R. (2009). Genomics of growth traits in forest trees. *Current Opinion in Plant Biology*, 12, 148– 156. https://doi.org/10.1016/j.pbi.2008.12.008.

Grattapaglia, D. and Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, 7, 241– 255. https://doi.org/10.1007/s11295-010-0328-4.

Grivet, D., Sebastiani, F., Alía, R., Bataillon, T., Torre, S., Zabal-Aguirre, M., Vendramin, G. G., and González-Martínez, S. C. (2011). Molecular Footprints of Local Adaptation in Two Mediterranean Conifers. *Molecular Biology and Evolution*, 28, 101– 116. https://doi.org/10.1093/molbev/msq190.

Guan, Y. and Stephens, M. (2011). Bayesian variable selection regression for genome-wide association studies and other large-scale problems. *The Annals of Applied Statistics*, 5, 1780– 1815. https://doi.org/10.1214/11-AOAS455.

Gwaze, D. P., Wolliams, J. A., Kanowski, P. J., and Bridgwater, F. E. (2001). Interactions of genotype with site for height and stem straightness in Pinus taeda in Zimbabwe. *Silvae Genetica*, 50, 135– 139.

Hereford, J. (2009). A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist*, 173, 579– 588. https://doi.org/10.1086/597611.

Hidalgo, A. M., Bastiaansen, J. W. M., Lopes, M. S., Calus, M. P. L., and Koning, D. J. de (2016). Accuracy of genomic prediction of purebreds for cross bred performance in pigs. *Journal of Animal Breeding and Genetics*, 133, 443– 451. https://doi.org/10.1111/jbg.12214.

Holliday, J. A., Aitken, S. N., Cooke, J. E. K., Fady, B., González-Martínez, S. C., Heuertz, M., Jaramillo-Correa, J.-P., Lexer, C., Staton, M., Whetten, R. W., and Plomion, C. (2017). Advances in ecological genomics in forest trees and applications to genetic resources conservation and breeding. *Molecular Ecology*, 26, 706– 717. https://doi.org/10.1111/mec.13963.

Jaramillo-Correa, J.-P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F., Heuertz, M., Garnier-Géré, P. H., Alía, R., Plomion, C., Vendramin, G. G., and González-Martínez, S. C. (2015). Molecular Proxies for Climate Maladaptation in a Long-Lived Tree (Pinus pinaster Aiton, Pinaceae). *Genetics*, 199, 793– 807. https://doi.org/10.1534/genetics.114.173252.

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F., Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., and de los Campos, G. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theoretical and Applied Genetics*, 127, 595– 607. https://doi.org/10.1007/s00122-013-2243-1.

Kardos, M. and Luikart, G. (2021). The Genetic Architecture of Fitness Drives Population Viability during Rapid Environmental Change. *The American Naturalist*, 197, 511– 525. https://doi.org/10.1086/713469.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225– 1241. https://doi.org/10.1111/j.1461-0248.2004.00684.x.

Langlet, O. (1971). Two Hundred Years Genecology. *Taxon*, 20, 653– 721. https://doi.org/10.2307/1218596.

Latta, R. G. (2009). Testing for local adaptation in Avena barbata: a classic example of ecotypic divergence. *Molecular Ecology*, 18, 3781– 3791. https://doi.org/10.1111/j.1365-294X.2009.04302.x.

Leimu, R. and Fischer, M. (2008). A Meta-Analysis of Local Adaptation in Plants. *PLOS ONE*, 3, e4010. https://doi.org/10.1371/journal.pone.0004010.

Leites, L. P., Rehfeldt, G. E., Robinson, A. P., Crookston, N. L., and Jaquish, B. (2012a). Possibilities and Limitations of Using Historic Provenance Tests to Infer Forest Species Growth Responses to Climate Change. *Natural Resource Modeling*, 25, 409– 433. https://doi.org/10.1111/j.1939-7445.2012.00129.x.

Leites, L. P., Rehfeldt, G. E., and Steiner, K. C. (2019). Adaptation to climate in five eastern North America broadleaf deciduous species: Growth clines and evidence of the growth-cold tolerance trade-off. *Perspectives in Plant Ecology, Evolution and Systematics*, 37, 64– 72. https://doi.org/10.1016/j.ppees.2019.02.002.

Leites, L. P., Robinson, A. P., Rehfeldt, G. E., Marshall, J. D., and Crookston, N. L. (2012b). Height-growth response to climatic changes differs among populations of Douglas-fir: a novel analysis of historic data. *Ecological Applications*, 22, 154– 165. https://doi.org/10.1890/11-0150.1.

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

Li, Y., Suontama, M., Burdon, R. D., and Dungey, H. S. (2017). Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genetics & Genomes*, 13, 60. https://doi.org/10.1007/s11295-017-1144-x.

Mackay, J., Dean, J. F. D., Plomion, C., Peterson, D. G., Cánovas, F. M., Pavy, N., Ingvarsson, P. K., Savolainen, O., Guevara, M. Á., Fluch, S., Vinceti, B., Abarca, D., Díaz-Sala, C., and Cervera, M.-T. (2012). Towards decoding the conifer giga-genome. *Plant Molecular Biology*, 80, 555– 569. https://doi.org/10.1007/s11103-012-9961-7.

MacLachlan, I. R., McDonald, T. K., Lind, B. M., Rieseberg, L. H., Yeaman, S., and Aitken, S. N. (2021). Genome-wide shifts in climate-related variation underpin responses to selective breeding in a widespread conifer. *Proceedings of the National Academy of Sciences*, 118. https://doi.org/10.1073/pnas.2016900118.

Mahony, C. R., MacLachlan, I. R., Lind, B. M., Yoder, J. B., Wang, T., and Aitken, S. N. (2020). Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. *Evolutionary Applications*, 13, 116– 131. https://doi.org/10.1111/eva.12871.

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J., Bustamante, C. D., and Kenny, E. E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *The American Journal of Human Genetics*, 100, 635– 649. https://doi.org/10.1016/j.ajhg.2017.03.004.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., and Daly, M. J. (2019). Clinical use of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, 51, 584– 591. https://doi.org/10.1038/s41588-019-0379-x.

Matesanz, S., Gianoli, E., and Valladares, F. (2010). Global change and the evolution of phenotypic plasticity in plants. *Annals of the New York Academy of Sciences*, 1206, 35– 55. https://doi.org/10.1111/j.1749-6632.2010.05704.x.

Mathieson, I. (2021). The omnigenic model and polygenic prediction of complex traits. *The American Journal of Human Genetics*. https://doi.org/10.1016/j.ajhg.2021.07.003.

McLane, S. C., Daniels, L. D., and Aitken, S. N. (2011). Climate impacts on lodgepole pine (Pinus contorta) radial growth in a provenance experiment. *Forest Ecology and Management*, 262, 115– 123. https://doi.org/10.1016/j.foreco.2011.03.007.

Merilä, J. and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, 7, 1– 14. https://doi.org/10.1111/eva.12137.

Moghaddar, N., Swan, A. A., and van der Werf, J. H. (2014). Comparing genomic prediction accuracy from purebred, crossbred and combined purebred and crossbred reference populations in sheep. *Genetics Selection Evolution*, 46, 58. https://doi.org/10.1186/s12711-014-0058-4.

Muff, S., Niskanen, A. K., Saatoglu, D., Keller, L. F., and Jensen, H. (2019). Animal models with group-specific additive genetic variances: extending genetic group models. *Genetics Selection Evolution*, 51, 7. https://doi.org/10.1186/s12711-019-0449-7.

Nadeau, S., Meirmans, P. G., Aitken, S. N., Ritland, K., and Isabel, N. (2016). The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution*, 6, 8649– 8664. https://doi.org/10.1002/ece3.2550.

Neale, D. B. and Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends in Plant Science*, 9, 325– 330. https://doi.org/10.1016/j.tplants.2004.05.006.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M. D., Richards, C. L., Valladares, F., and van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, 15, 684– 692. https://doi.org/10.1016/j.tplants.2010.09.008.

O'Neill, G. A., Hamann, A., and Wang, T. (2008). Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology*, 1040– 1049. https://doi.org/10.1111/j.1365-2664.2008.01472.x@10.1111/(ISSN)1365-2664.CLIMATE˙JPE.

Olson, M. S., Robertson, A. L., Takebayashi, N., Silim, S., Schroeder, W. R., and Tiffin, P. (2010). Nucleotide diversity and linkage disequilibrium in balsam poplar (Populus balsamifera). *New Phytologist*, 186, 526– 536. https://doi.org/10.1111/j.1469-8137.2009.03174.x.

Pedlar, J. H. and McKenney, D. W. (2017). Assessing the anticipated growth response of northern conifer populations to a warming climate. *Scientific Reports*, 7, 43881. https://doi.org/10.1038/srep43881.

Pigliucci, M., Murren, C. J., and Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology*, 209, 2362– 2367. https://doi.org/10.1242/jeb.02070.

Plomion, C., Bartholomé, J., Lesur, I., Boury, C., Rodríguez-Quilón, I., Lagraulet, H., Ehrenmann, F., Bouffier, L., Gion, J. M., Grivet, D., Miguel, M. de, María, N. de, Cervera, M. T., Bagnoli, F., Isik, F., Vendramin, G. G., and González-Martínez, S. C. (2016).

High-density SNP assay development for genetic analysis in maritime pine (Pinus pinaster). *Molecular Ecology Resources*, 16, 574– 587. https://doi.org/10.1111/1755-0998.12464.

Postma, F. M. and Ågren, J. (2016). Early life stages contribute strongly to local adaptation in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 113, 7590– 7595. https://doi.org/10.1073/pnas.1606303113.

Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*, 20, R208– R215. https://doi.org/10.1016/j.cub.2009.11.055.

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155, 945– 959.

Prunier, J., Gérardi, S., Laroche, J., Beaulieu, J., and Bousquet, J. (2012). Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology*, 21, 4270– 4286. https://doi.org/10.1111/j.1365-294X.2012.05691.x.

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

Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J., Novella-Fernandez, R., Alberdi, A., and Manel, S. (2019). Considering adaptive genetic variation in climate change vulnerability assessment reduces species range loss projections. *Proceedings of the National Academy of Sciences*, 116, 10418– 10423. https://doi.org/10.1073/pnas.1820663116.

Rehfeldt, G. E., Leites, L. P., Joyce, D. G., and Weiskittel, A. R. (2018). Role of population genetics in guiding ecological responses to climate. *Global Change Biology*, 24, 858– 868. https://doi.org/10.1111/gcb.13883.

Rehfeldt, G. E., Tchebakova, N. M., Milyutin, L. I., Parfenova, E. I., Wykoff, W. R., and Kouzmina, N. A. (2003). Assessing population responses to climate in Pinus sylvestris and Larix spp. of Eurasia with climate-transfer models. *Eurasian Journal of Forest Research - Hokkaido University (Japan)*.

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., and Hamilton, D. A. (1999). Genetic Responses to Climate in Pinus Contorta: Niche Breadth, Climate Change, and Reforestation. *Ecological Monographs*, 69, 375– 407. https://doi.org/10.1890/0012-9615(1999)069[0375:GRTCIP]2.0.CO;2.

Resende Jr, M. F. R., Muñoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D., Resende, M. D. V., and Kirst, M. (2012). Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytologist*, 193, 617– 624. https://doi.org/10.1111/j.1469-8137.2011.03895.x.

Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petroli, C. D., Missiaggia, A. A., Aguiar, A. M., Abad, J. M., Takahashi, E. K., Rosado, A. M., Faria, D. A., Pappas, G. J., Kilian, A., and Grattapaglia, D. (2012). Genomic selection for growth and wood quality in Eucalyptus: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist*, 116– 128. https://doi.org/10.1111/j.1469-8137.2011.04038.x@10.1002/(ISSN)1469-8137(CAT)FeatureIssues(VI)Bioenergytrees.

Rodríguez-Quilón, I., Santos-del-Blanco, L., Serra-Varela, M. J., Koskela, J., González-Martínez, S. C., and Alía, R. (2016). Capturing neutral and adaptive genetic diversity for conservation in a highly structured tree species. *Ecological Applications*, 26, 2254– 2266. https://doi.org/10.1002/eap.1361.

Santos-del-Blanco, L., Climent, J., González-Martínez, S. C., and Pannell, J. R. (2012). Genetic differentiation for size at first reproduction through male versus female functions in the widespread Mediterranean tree Pinus pinaster. *Annals of Botany*, 110, 1449– 1460. https://doi.org/10.1093/aob/mcs210.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14, 807– 820. https://doi.org/10.1038/nrg3522.

Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, 38, 595– 619. https://doi.org/10.1146/annurev.ecolsys.38.091206.095646.

Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J., Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M. J., Trotsiuk, V., Mairota, P., Svoboda, M., Fabrika, M., Nagel, T. A., and Reyer, C. P. O. (2017). Forest disturbances under climate change. *Nature Climate Change*, 7, 395– 402. https://doi.org/10.1038/nclimate3303.

Sella, G. and Barton, N. H. (2019). Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annual Review of Genomics and Human Genetics*, 20.

Shaw, R. G. and Etterson, J. R. (2012). Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist*, 195, 752– 765. https://doi.org/10.1111/j.1469-8137.2012.04230.x.

Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., Chiang, C. W., Hirschhorn, J., Daly, M. J., Patterson, N., Neale, B., Mathieson, I., Reich, D., and Sunyaev, S. R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife*, 8, e39702. https://doi.org/10.7554/eLife.39702.

Sork, V. L. (2018). Genomic Studies of Local Adaptation in Natural Plant Populations. *Journal of Heredity*, 109, 3– 15. https://doi.org/10.1093/jhered/esx091.

Stern, A. J., Speidel, L., Zaitlen, N. A., and Nielsen, R. (2020). Disentangling selection on genetically correlated polygenic traits using whole-genome genealogies. *bioRxiv*. https://doi.org/10.1101/2020.05.07.083402.

Tiffin, P. and Ross-Ibarra, J. (2014). Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, 29, 673– 680. https://doi.org/10.1016/j.tree.2014.10.004.

Valladares, F., Gianoli, E., and Gómez, J. M. (2014a). Ecological limits to plant phenotypic plasticity. *New Phytologist*, 176, 749– 763. https://doi.org/10.1111/j.1469-8137.2007.02275.x.

Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M., Cornwell, W., Gianoli, E., Kleunen, M. van, Naya, D. E., Nicotra, A. B., Poorter, H., and Zavala, M. A. (2014b). The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecology Letters*, 17, 1351– 1364. https://doi.org/10.1111/ele.12348.

Viñas, R. A., Caudullo, G., Oliveira, S., and de Rigo, D. (2016). Pinus pinaster in Europe: distribution, habitat, usage and threats.

Vizcaíno-Palomar, N., Fady, B., Alía, R., Raffin, A., Mutke, S., and Benito Garzón, M. (2020). The legacy of climate variability over the last century on populations' phenotypic variation in tree height. *Science of The Total Environment*, 749, 141454. https://doi.org/10.1016/j.scitotenv.2020.141454.

Vogt, G. (2015). Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *Journal of Biosciences*, 40, 159– 204. https://doi.org/10.1007/s12038-015-9506-8.

Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., Mock, T., Schmid, K., Schmitt, I., Bataillon, T., Savolainen, O., Bergland, A., Flatt, T., Guillaume, F., and Pfenninger, M. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters*, 4, 4– 18. https://doi.org/10.1002/evl3.154.

Wang, T., O'Neill, G. A., and Aitken, S. N. (2010). Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological Applications*, 20, 153– 163. https://doi.org/10.1890/08-2257.1.

Wu, H. X. and Ying, C. C. (2004). Geographic pattern of local optimality in natural populations of lodgepole pine. *Forest Ecology and Management*, 194, 177– 198. https://doi.org/10.1016/j.foreco.2004.02.017.

Zas, R., Merlo, E., Díaz, R., and Fernández López, J. (2003). Stability across sites of Douglas-fir provenances in northern Spain. 10, 71– 82.

Zonneveld, B. J. M. (2012). Conifer genome sizes of 172 species, covering 64 of 67 genera, range from 8 to 72 picogram. *Nordic Journal of Botany*, 30, 490– 502. https://doi.org/10.1111/j.1756-1051.2012.01516.x.