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

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

Running title: Integrating genomics in trait prediction

Keywords: Climate change, local adaptation, range-wide predictive models, positive-effect alleles, maritime pine.

Abstract

Predicting adaptive-trait variation across species ranges is essential to assess the potential of populations to survive under future environmental conditions. In forest trees, multi-site common gardens have long been the gold standard for separating the genetic and plastic components of trait variation and predicting population responses to new environments. However, relying on common gardens alone limits our ability to extrapolate predictions to populations or sites not included in these logistically expensive and time-consuming experiments. In this study, we aimed to determine whether models that integrate large-scale climatic and genomic data could capture the underlying drivers of tree adaptive-trait variation, and thus improve predictions at large geographical scales. Using a clonal common garden consisting of 34 provenances of maritime pine (523 genotypes and 12,841 trees) planted in five sites under contrasted environments, we compared twelve statistical models to: (i) separate the genetic and plastic components of height growth, a key adaptive trait in forest trees, (ii) identify the relative importance of factors underlying height-growth variation across individuals and populations, and (iii) improve height-growth prediction of unknown observations and provenances. We found that the height-growth plastic component exceeded more than twice the genetic component. The plastic component was likely due to multiple environmental factors, including annual climatic variables, while the genetic component was driven by the confounded effects of past demographic history and provenance adaptation to the climate-of-origin. Distinct gene pools were characterized by different total genetic variance, with broad-sense heritability ranging from 0.104 (95% CIs: 0.065-0.146) to 0.223 (95% CIs: 0.093-0.363), suggesting different potential response to selection along the geographical distribution of maritime pine. When predicting height-growth of new observations, models combining population demographic history, provenance climate-of-origin, and positive-effect height-associated alleles (PEAs; previously identified by GWAs) explained as much variance as models relying directly on the common garden design. Noteworthy, these models explained substantially more variance when predicting height-growth in new provenances, particularly in harsh environments. Predicting quantitative traits of ecological and/or economic importance across species ranges would therefore benefit from integrating ecological and genomic information.

1 Introduction

Global change is expected to have a profound impact on forests (Franklin et al. 2016, Seidl et al. 2017). However, it remains uncertain whether tree populations will become extinct, migrate or persist across their current range (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 in 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 neutral population demographic history or adaptation to local biotic and abiotic conditions, or both (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 population responses to global change could be jointly integrated to improve the predictions of current and future adaptive-trait variation across species ranges. The increasing availability of genomic data opens new opportunities to boost adaptive-trait prediction accuracy, which is critical for breeding (i.e. genomic selection; Grattapaglia and Resende 2011), to anticipate future distributions 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).

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. 2019). The development of the *universal response function* was a step forward to jointly 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 climate conditions under which the populations have evolved (i.e. the climate-of-origin of the populations tested)- in explaining quantitative trait variation (O'Neill et al. 2008, Wang et al. 2010). These models, based exclusively on climatic and phenotypic data from common gardens, are now widely applied in the mixed model framework to generate future projections for a large variety of adaptive traits (Leites et al. 2012a,b, Benito Garzón et al. 2019, Vizcaíno-Palomar et al. 2020). One of the main conclusions suggested by these models is that trait variation across species ranges is largely 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 populations (i.e. related to the genetic component) (Leites et al. 2012b, Benito Garzón et al. 2019). Besides, these models do not allow determining to what extent associations between adaptive-trait variation and climate in the provenance location, or the higher trait values of local compared to foreign populations, are caused by adaptive processes or population demographic history (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 adaptive-trait variation, and therefore our ability to predict the genetic component of new (untested in common gardens) populations and forecast future trait variation under new environments at large geographical scales.

The advent and generalization of genomic tools have enhanced our understanding of adaptive and neutral genetic processes resulting in adaptive-trait variation, and their relationship with climatic gradients (Savolainen et al. 2013, Sork 2018, Leroy et al. 2020). Further integration of genomic information in pre-

dictive models for adaptive traits would be highly valuable as it would provide a mean for including intraspecific variability at a finer scale than in current models (Mahony et al. 2020). It would also allow developing more precise predictive models for species and populations for which hard-to-collect phenotypic data from common gardens are not available. 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 Barton 2019, Sohail et al. 2019). Such genetic population structure hinders the use at the species range scale of genomic relationship matrices, which now provide more accurate estimates of genetic parameters (e.g. breeding values, additive and non-additive variance) within breeding populations than previous pedigree-based approaches (Bouvet et al. 2016, El-Dien et al. 2018). Indeed, admixed populations or distinct gene pools 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 to predict current and future adaptive-trait variation in natural populations, while highly valuable, remains challenging.

Forest trees are remarkable models to study the genetic and plastic components of adaptive-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 (de Villemereuil et al. 2014). 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) adaptive traits (Neale and Savolainen 2004, Jaramillo-Correa et al. 2015). For example, de Miguel et al. (2020) found that the average proportion of genome-wide variants with non-zero effects on height growth in maritime pine, i.e. the degree of polygenicity, could be as high as ~7%. 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 in large populations or geographical regions of forest trees remains to be explored.

In the present study, we aim to identify the potential drivers of the plastic and genetic height-growth components in distinct maritime pine gene pools and determine whether genomic information can help improve predictions across the species range. To this aim, 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 tested the commonly accepted hypotheses in forest trees that most adaptive-trait variation is associated with the plastic component, and much less with the genetic component, and that the provenance-by-environment interaction displays footprints of local adaptation. Second, we evaluated the relative importance of potential drivers of adaptive-trait variation. We expected that: (i) the plastic and genetic components are both associated with climate; (ii) population demographic history explains a substantial amount of the genetic component of height growth, as maritime pine is known for its strong population genetic structure (Jaramillo-Correa et al. 2015), (iii) provenance climate-of-origin and height growth are correlated, with correlation strength and direction varying among common gardens, suggesting local adaptation; and (iv) distinct gene pools have specific total genetic variances, suggesting a potential variability in the response to selection in different parts of maritime pine range. Third, we compared the predictive performance -on unknown observations or provenances- of models based exclusively on the common garden design and models including the potential predictors of the genetic component considered above, in particular those relying on genomic information. We expect that demographic history, provenance climate-of-origin and height-associated positive-effect alleles (PEAs), either combined or alone, may improve height-growth predictions of provenances for which phenotypic data from common gardens are not available, at least in some environments. Our study is a step towards integrating the recent knowledge brought by large genomic datasets to the modeling of large-scale adaptive-trait variation in forest trees. Genomic tools hold great promise for predicting future adaptive-trait variation without using phenotypic measurements, however, a robust framework is needed to make reliable predictions across species ranges 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 has also been used in commercial plantations outside its natural distribution (e.g. in Australia and South Africa). 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 region, and the continental climate of central Spain. Maritime pine populations are highly

fragmented. Based on different molecular markers and morphological traits, six maritime pine gene pools have been identified for this species (Alberto et al. 2013, Jaramillo-Correa et al. 2015; see Fig. 1), presumably resulting from the expansion of as many glacial refugia (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 (also referred as test sites; Fig. 1). These five sites are located within the main climates in which maritime pine occurs: the Iberian Atlantic region, with mild winters and high annual rainfall (two sites in Asturias and Portugal), the French Atlantic region with colder winters (one site in Bordeaux), and the Mediterranean region with high temperatures and intense summer drought, as well as large precipitation differences between summer and winter (two sites in Cáceres and Madrid). 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 allowing enough distance between trees (over 50 m) as 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).

2.2 Single-nucleotide polymorphism (SNP) genotyping and positive-effect alleles (PEAs)

Genotypes collected in the Asturias common garden were genotyped with the Illumina Infinium 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). Details about DNA extraction and genotyping can be found in Plomion et al. (2016). The proportion of each genotype g belonging to the gene pool j was based on Jaramillo-Correa et al. (2015), who used nine nuSSRs as well as a subset of the same SNPs as in our study (1,745 SNPs) to infer population genetic structure in maritime pine (referred as the population demographic history) using the Bayesian approach available in Structure v. 2.3.3 (Pritchard et al. 2000; Table S3).

Based on the 523 genotypes for which there were both genotypic and phenotypic data, we performed four GWAS analyses under a polygenic framework to select groups of SNPs with the strongest associations with height. First, to identify SNPs that have an association with height at range-wide geographical scales, a global GWAS was performed using the combined phenotypic data from the five common

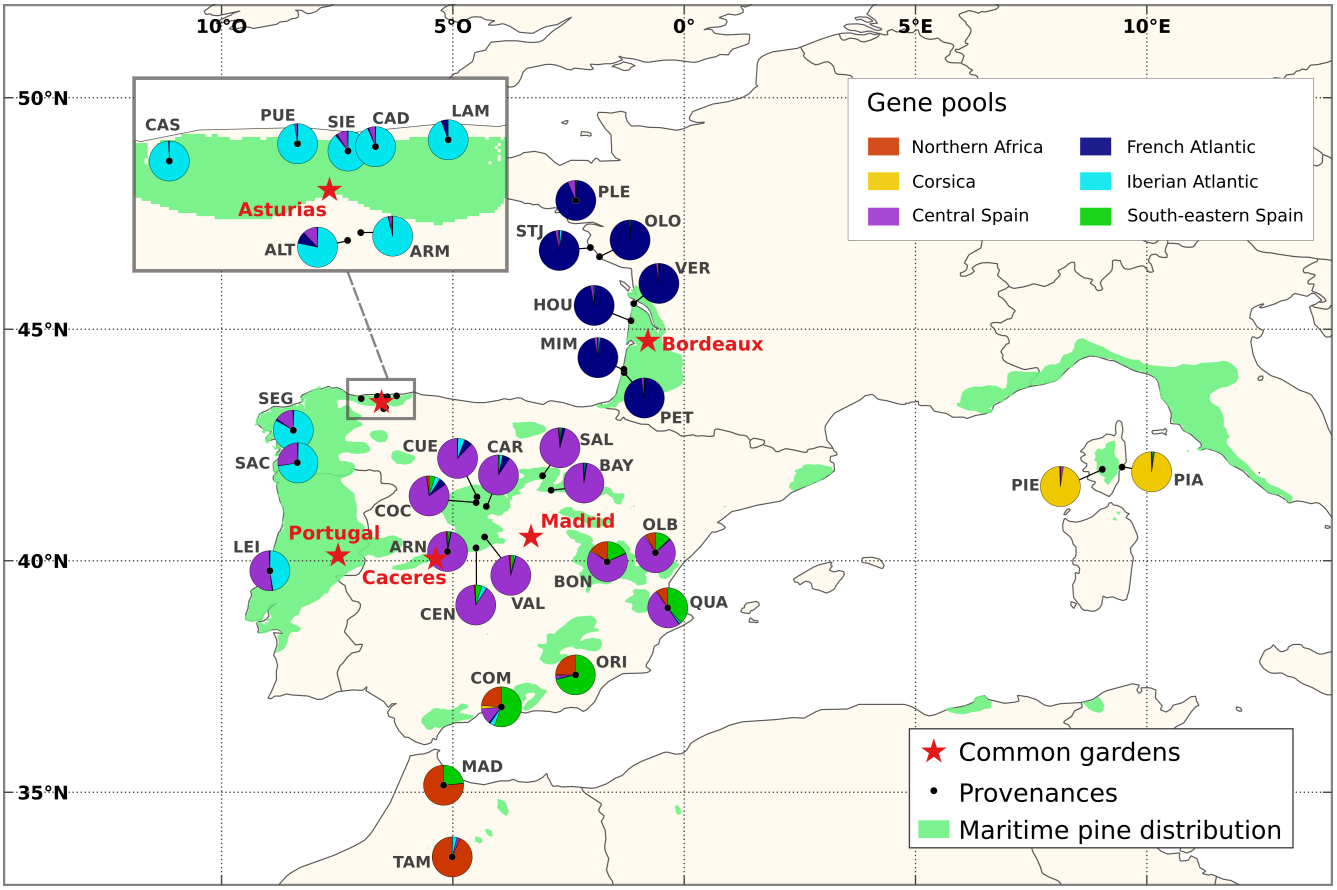

Figure 1. The five common gardens and 34 provenances of maritime pine (CLONAPIN common garden network) used in this study. The distribution of maritime pine is also shown (based on EUFORGEN map, http://www.euforgen.org/). Pie charts represent the proportions of belonging to each gene pool for each provenance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

gardens. Second, to identify SNPs that have a local association with height in a particular geographical region r (i.e. in a particular environment), three regional GWAS were performed 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). All GWAS were performed following the Bayesian variable selection regression (BVSR) methodology implemented in the piMASS software (Guan and Stephens 2011), including corrections for population structure and using the height BLUPs reported in de Miguel et al. (2020). For each of the four GWAS, we selected the 350 SNPs with the highest absolute estimates of the posterior effect size (i.e. Rao-Blackwellized estimates), corresponding approximately to the levels of polygenicity estimated for height in maritime pine (de Miguel et al. 2020). These SNPs were used to compute the number of global and regional positive-effect alleles (PEAs) for each genotype ($gPEA_g$ and $rPEA_{gr}$, respectively; see Fig. S1). Interestingly (but not surprisingly, see Prunier et al. 2012, Tiffin and Ross-Ibarra 2014), regional PEAs were generally not the same across regions (only 69 PEAs shared between Iberian and French Atlantic regions, 41 shared between French Atlantic and Mediterranean regions, and 83 shared between the Iberian Atlantic and Mediterranean regions).

Finally, SNPs that did not show any association with height at range-wide geographical scales were used to calculate genomic relationship matrices $\mathbf{A_j}$, specific to each gene pool j (see Muff et al. 2019 for

details on $\mathbf{A_j}$ calculation)

2.3 Climatic data

Both genomic and modeling studies in forest trees have shown that large-scale patterns of allele frequencies or adaptive-trait variation are associated with climatic variables related to temperature and precipitation, such as their annual means (Eckert et al. 2010, McLane et al. 2011, Fréjaville et al. 2019, Leites et al. 2019, Mahony et al. 2020), and 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. 2019). In addition, as climate change will cause major changes in temperature and precipitation conditions in the near future, particularly in the Mediterranean basin, there is a need to understand the complex influence of temperature and precipitation variables on adaptive-trait variation.

We extracted monthly and yearly climatic data from the EuMedClim database with 1 km resolution (Fréjaville and Benito Garzón 2018). First, we calculated six variables that describe both extreme and average temperature and precipitation conditions in the test sites during the year preceding the measurements: mean of monthly precipitation (mm), minimum of monthly minimum temperatures (°C), minimum of monthly precipitation during summer -June to September- (mm), mean of monthly maximum temperatures (°C), maximum of monthly precipitation (mm), and maximum of monthly maximum temperatures (°C). These variables had at most a correlation coefficient of 0.85 among each other (Fig. S2). Due to the unbalanced number of measurements among test sites (trees were measured only once in the hottest and driest sites, Cáceres and Madrid, as survival was very low), some of these variables were slightly correlated with tree age (at most with a correlation coefficient of 0.56 for the mean of the monthly precipitation; Fig. S2). We decided not to include soil variables in the analyses as they were highly correlated to some of the climatic variables (see section 3.1 of the Supplementary Information for details). Second, we calculated four variables that describe the mean temperature and precipitation in the provenance locations over the period from 1901 to 2009, representing the climate under which provenances have evolved: the average of the annual daily mean temperature (°C), the average of the maximum temperature of the warmest month (°C), the average of the annual precipitation (mm) and the average of the precipitation of the driest month (mm) (see section 3.2 of the Supplementary Information for details). These variables had at most a correlation coefficient of 0.77 among each other (Fig. S5), and three of them were correlated to the population genetic structure ($|\rho| \geq 0.6$). Indeed, the provenance proportion of belonging to the French Atlantic gene pool was positively correlated (ρ=0.83) to the average of the precipitation during the driest month, whereas the proportion of belonging to the Central Spain gene pool was negatively correlated (ρ=-0.68) to the average of the annual precipitation and positively correlated (ρ=0.6) to the average of the maximum temperature of the warmest month (Fig. S5). However, the confounding effect introduced by these correlations was mitigated by some provenances belonging to different gene pools but occurring in similar climates (i.e. French and Iberian Atlantic provenances), and some provenances occurring in different climates but belonging to the same gene pool (i.e. Corsican provenances).

The selected climatic variables were used to build two covariance matrices, Ω and Φ, describing the climatic similarity between test sites and provenances, respectively (Figs. S4 & S7), following Jarquín et al. (2014); see also a similar approach but using Euclidean distance matrices in Thomson et al. (2018).

2.4 Hierarchical height-growth models

Twelve hierarchical models for height growth were compared. First, we built models to quantify the relative contribution of the genetic and plastic components of height-growth variation, and their interaction, based exclusively on the common garden design (*models M0*, *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 and genetic adaptation to climate, as well as population demographic history; in addition, we estimated gene pool-specific total genetic variances (*models M3* to *M6*; Table 1). Third, we used some of the potential drivers of the genetic height-growth component (two temperature and precipitation variables, population demographic history, and global and regional PEAs) to predict unknown observations and provenances without directly relying on the common garden design (*models M7* to *M12*; Table 1).

Variables	**Genetic/plastic bases**			**Potential drivers**				**Predicting new obs. and prov.**					
	M0	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**
Site/Block	×	×	×	×	×	×	×	×	×	×	×	×	×
Prov		×	×	×	×	×	×						
Gen		×	×	×	×		×						
Site:Prov			×										
SiteClimSim				×	×	×	×						
DH					×	×	×	×	×	×			
GR						×							
ProvClimSim							×						
ProvClimVar								×	×		×		
gPEA								×				×	
rPEA									×				×

Table 1. Variables included in the different height-growth models, from *M0* (baseline model) to *M12*. *Models M0*, *M1* and *M2* separate the genetic and plastic components of height-growth variation: Site/Block, sites and blocks nested within sites; Prov, provenances; Gen, genotypes (clones); Site:Prov, interaction between sites and provenances. *Models M0* to *M2* are based exclusively on height data from the common gardens. *Models M3* to *M6* test different hypotheses regarding the potential drivers underlying height-growth variation: SiteClimSim, climatic similarity between test sites during the year preceding the measurement; DH, population demographic history; GR, gene pool-specific total genetic variances; ProvClimSim, climatic similarity between provenances from 1901 to 2009. *Models M7* to *M12* predict new observations and provenances without relying directly on the common garden design: ProvClimVar, two temperature and precipitation-related variables describing the climatic conditions in the provenance locations (*min.pre*, precipitation of the driest month, and *max.temp*, maximum temperature of the warmest month), gPEA and rPEA, global and regional height-associated positive-effect allele counts, respectively. ProvClimVar, gPEA and rPEA were included in the models with site-specific slopes. All models also contained the age effect, not shown in the table.

All models shared the following common features. 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. S10). The plastic component was inferred with varying intercepts for the sites S_s and blocks nested within sites $B_{b(s)}$. 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 *Models M0, M1* and *M2*: genetic and plastic bases of height-growth variation

In our baseline *model M0*, height h was modeled as a function of tree age and the environmental conditions of the five test sites (i.e. the plastic component):

$$\begin{aligned} \log(h_{isb}) &\sim \mathcal{N}(\mathbf{X}\beta + \mu_{sb}\,,\ \sigma^2) \\ \mathbf{X}\beta &= \beta_0 + \beta_{age}\mathrm{age}_i + \beta_{age2}\mathrm{age}_i^2 \\ \mu_{sb} &= S_s + B_{b(s)} \end{aligned} \tag{1}$$

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_{sb} is the vector of varying intercepts.

In *model M1*, we considered both the plastic and genetic components for height growth. The genetic component was inferred with varying intercepts for provenances P_p and for genotypes within provenances $G_{g(p)}$. Height h was then expressed as follows:

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

Model M2 was based on *model M1* but including an interaction term between provenance and site as follows:

$$\mu_{sbpg} = S_s + B_{b(s)} + P_p + G_{g(p)} + S_s P_p \tag{3}$$

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

In *model M3*, we hypothesized that the height-growth plastic component was influenced by the annual climatic conditions in the test sites (modeled as the climatic similarity between test sites during the year preceding the measurements). This model can be expressed with the same likelihood function as *M1*

but with μ_{sbpg} equal to:

$$\begin{aligned} \mu_{isbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} \\ cs_{is} &\sim \mathcal{N}(0, \Omega\sigma^2_{cs_{is}}) \end{aligned} \tag{4}$$

where Ω is the covariance matrix describing the climatic similarity between test sites s during the year i preceding the measurements (Fig. S4) and cs_{is} are varying intercepts associated with the annual climatic conditions each test site s. In *M3*, the plastic component was partitioned between the regression on the annual 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* and *M5*, we hypothesized that the genetic component of height growth was influenced by the population demographic history and thus conditioned by the existence of differentiated gene pools in maritime pine. In *M4*, following Wolak and Reid (2017), gene pools j were allowed to vary in their mean relative contribution g_j on height growth as follows:

$$\begin{aligned} \mu_{ijsbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + \sum_{j=1}^{6} q_{gj} g_j \\ g_j &\sim \mathcal{N}(0, \sigma^2_{g_j}) \end{aligned} \tag{5}$$

where q_{gj} corresponds to the proportion of each genotype g belonging to the gene pool j (as estimated in Jaramillo-Correa et al. 2015). In this model, trees from the same gene pool are considered to be unrelated, which is a reasonable assumption given the sampling scheme (see Materials & Methods).

M5 extends *M4* by allowing gene pools j to vary in their total genetic variance $\sigma^2_{A_j}$, following Muff et al. (2019). This involves replacing the genotype varying intercepts $G_{g(p)}$ in *M4* by partial genetic values a_{gj} standing for the relative contribution of gene pool j to the genetic value a_g of genotype g (approximated in *M4* by the genotype intercepts $G_{g(p)}$). Thus, *M5* can be expressed as *M4* but with μ_{ijsbpg} equal to:

$$\begin{aligned} \mu_{ijsbpg} &= S_s + B_{b(s)} + P_p + cs_{is} + \sum_{j=1}^{6} q_{gj} g_j + \sum_{j=1}^{6} a_{gj} \\ \mathbf{a}_j^{\mathsf{T}} &= (a_{1j}, ..., a_{nj})^{\mathsf{T}} \sim \mathcal{N}(0, \sigma^2_{A_j}\mathbf{A_j}) \end{aligned} \tag{6}$$

with $\mathbf{A_j}$ the genomic relationship matrix specific to the gene pool j and $\sigma^2_{A_j}$, the total genetic variance in gene pool j. Using the modeled residual variance σ^2 and gene-pool specific total genetic variances $\sigma^2_{A_j}$, we calculated the gene-pool specific broad-sense heritability as:

$$H^2_j = \frac{\sigma^2_{A_j}}{\sigma^2_{A_j} + \sigma^2}$$

In *model M6*, we hypothesized that populations were genetically adapted to the climatic conditions in which they evolved. Thus, we aimed to quantify the association between height growth and the climatic conditions in the provenance locations (modeled as the climatic similarity between provenances from 1901 to 2009), while accounting also for the population demographic history. We kept the genotype varying intercepts (like in *M1* to *M4*) but not the gene pool-specific total genetic variances (unlike *M5*). Thus, *M6* extends *M4* as:

$$\begin{aligned} \mu_{ijsbpg} &= S_s + B_{b(s)} + P_p + G_{g(p)} + cs_{is} + cp_p + \sum_{j=1}^{6} q_{gj} g_j \\ cp_p &\sim \mathcal{N}(0, \Phi\sigma^2_{cp_p}) \end{aligned} \tag{7}$$

where Φ is the covariance matrix describing the climatic similarity between provenances p (Fig. S7) and cp_p are varying intercepts associated with each provenance p. In *M6*, the genetic component was partitioned among the regression on the climatic covariates (cp_p), the population structure covariates (g_j), and the deviations related to the genotype ($G_{g(p)}$) and provenance (P_p) effects (resulting, for instance, from adaptation to climatic variables not measured in our study, or from trans-generational epigenetic effects).

2.4.3 *Models M7* to *M12*: combining climatic and genomic data to predict new observations and provenances

In this set of models, we replaced the provenance and genotype (clone) intercepts by potential predictors of trait variation that do not rely exclusively on the common garden design (namely provenance climate-of-origin and genomic information). More specifically, we replaced P_p and $G_{g(p)}$ intercepts by the mean relative contribution g_j to height growth of each gene pool j (as in *M4*, Table 1), two variables related to the provenance climate (*min.pre* the precipitation of the driest month and *max.temp* the maximum temperature of the warmest month) and either global or regional PEA counts ($gPEA_g$ and $rPEA_{gr}$, respectively). This allowed us to determine whether these variables were able to predict the height-growth genetic component as accurately as provenance and genotype intercepts (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. In *models M9* to *M12*, the potential predictors were included individually in the models to determine their specific predictive performance.

M7 can be expressed as follows:

$$\begin{aligned} \mu_{jsbpg} = S_s + B_{b(s)} + \sum_{j=1}^{6} q_{gj} g_j + \beta_{min.pre,s} min.pre_p \\ + \beta_{max.temp,s} max.temp_p + \beta_{gPEA,s} gPEA_g \end{aligned} \tag{8}$$

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 count of global PEAs and $\beta_{gPEA,s}$ its site-specific slopes. *M8* is identical to *M7*, except that $gPEA_g$ was replaced by $rPEA_{gr}$ (the count of regional PEAs, i.e. the alleles with positive effects in specific geographical regions, and thus environments, of maritime pine range) and its corresponding site-specific slopes. In *M9*, only the demographic history was considered through the relative contribution g_j of each gene pool j. In *M10*, only the variables related to the provenance climate-of-origin were included ($min.pre_p$ and $max.temp_p$). And in *M11* and *M12*, only the gPEAs and rPEAs counts were included, 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 5 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 Model evaluation and comparison

Two partitions of the global dataset, P1 and P2, were used to evaluate the model goodness-of-fit on the fitted data (training datasets) and predictive performance on data not used to fit the models (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 global dataset 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, we aimed to predict new provenances. P2 corresponds to a random split between a training dataset of 28 provenances and a test data set containing the remaining 6 provenances. In the P2 partition, however, while observations are also drawn from the same blocks/sites, the partitions represent independent sets of provenances. Besides, we checked that the results obtained on the P2 partition were not due to provenance sampling by replicating the analyses on another partition (the P3 partition, composed of 6 other provenances, see section 9 of the Supplementary Information).

To evaluate the model goodness-of-fit, we calculated the residual based Bayesian $\mathcal{R}^2$ on both P1 and P2 training datasets (Gelman et al. 2019):

$$\mathcal{R}^2 = \frac{Var_\mu}{Var_\mu + Var_{res}}$$

where $Var_{\mu} = V_{n=1}^{N} \hat{y}_n$ is the variance of modeled predictive means and $Var_{res} = V_{n=1}^{N} y_n - \hat{y}_n$ the variance of draws from the residual distribution. For each model, $\mathcal{R}^2$ was calculated using modeled predictive means conditioned either on both varying and fixed effects ($\mathcal{R}^2_{all}$) or only on fixed effects ($\mathcal{R}^2_{fix}$). We also calculated the model mean predictive error (mean of observed minus predicted responses, PE) on the training datasets.

To evaluate the model predictive ability, the expected log predictive density (ELPD) was estimated for each model by Bayesian leave-one-out cross-validation (Vehtari et al. 2017). We also calculated the standard error of each pairwise comparison of the models' ELPD: when the difference between the models' ELPD was more than four times the standard error, the models were considered to be significantly different. Lastly, we calculated PE and $\mathcal{R}^2_{all}$, and site-specific PE_s and $\mathcal{R}^2_s$, on P1 and P2 test datasets (data not used to fit the models). For the P2 partition, we also calculated the provenance-specific PE_p and $\mathcal{R}^2_p$ for the six provenances of the test dataset.

3 Results

3.1 Genetic and plastic bases of height-growth variation

M0 and *M1* explained on average 76.8% [95% CIs: 76.6-77.1] and 81.5% [95% CIs: 81.3-81.7] of the variance in height growth (Table 2). As expected, a large amount of the variance (~65%) was attributed to tree age (see $\mathcal{R}^2_{fix}$ in Table 2). Comparing $\mathcal{R}^2_{fix}$ and $\mathcal{R}^2_{all}$ of *M0*, *M1* and *M2* allows to deduce that ~11% of the variance was attributed to plasticity, ~5% to genetic components and only 0.3% to the provenance-by-site interaction (Table 2). PE and $\mathcal{R}^2_{all}$ calculated on both the training and test datasets of the P1 partition showed that *M0* had a lower goodness-of-fit and predictive ability on new observations than *M1* and *M2*, the latter being almost similar (Table 2). A similar ranking of *M0*, *M1* and *M2* was obtained using the ELPD (see section 6.1.1 of the Supplementary Information).

In *M1*, the plastic component was mainly attributed to the variance σ^2_S between sites (median of 0.108), while the variance σ^2_B between blocks was almost null (median of 0.002) (Table S8 & Fig. S9). Some sites showed heights deviating strongly from the global mean: Madrid, where trees grew the least (median of -0.376), and Asturias where they grew particularly well (median of 0.272) (Fig. 2 & Table S9). The genetic component was equally attributed to the variance between provenances σ^2_P and genotypes σ^2_G, with a median of 0.013 and 0.012, respectively (Table S8 & Fig. S9). In *M2*, parameter estimates were similar to *M1* and are thus not presented herein, and the variance σ^2_{Inter} of the provenance-by-site interaction was much smaller than the variances among provenances or genotypes (median of 0.004 and 95% CIs: 0.003-0.006; Table S11 & Fig. S11).

Provenance-specific intercepts are shown in Fig. 2 (global intercepts across all sites from *M1*) and Fig. S12 (site-specific intercepts from *M2*). Most provenances from the French Atlantic gene pool (e.g. STJ, HOU, PET, OLO and VER) were taller on average than other provenances in all sites. Most provenances from the Iberian Atlantic gene pool (e.g. ARM or PUE) were among the tallest in Asturias and Bordeaux

Models	Training P1			Test P1		Test P2	
	$\mathcal{R}^2_{all}$	$\mathcal{R}^2_{fix}$	PE	$\mathcal{R}^2_{all}$	PE	$\mathcal{R}^2_{all}$	PE
M0	0.768 [0.001]	0.654 [0.001]	0.267 [0.226]	0.767 [0.001]	0.269 [0.229]	0.768 [0.001]	0.264 [0.003]
M1	0.815 [0.001]	0.651 [0.001]	0.231 [0.206]	0.808 [0.001]	0.236 [0.212]	0.769 [0.001]	0.264 [0.003]
M2	0.818 [0.001]	0.651 [0.001]	0.229 [0.205]	0.810 [0.001]	0.235 [0.210]	0.768 [0.003]	0.264 [0.003]
M3	0.816 [0.001]	0.642 [0.001]	0.230 [0.206]	0.810 [0.001]	0.235 [0.211]	-	-
M4	0.816 [0.001]	0.643 [0.001]	0.230 [0.206]	0.810 [0.001]	0.235 [0.211]	-	-
M5	0.816 [0.001]	0.643 [0.001]	0.231 [0.206]	0.809 [0.001]	0.236 [0.211]	-	-
M6	0.816 [0.001]	0.643 [0.001]	0.230 [0.206]	0.810 [0.001]	0.235 [0.211]	-	-
M7	0.804 [0.001]	0.652 [0.001]	0.241 [0.214]	0.801 [0.001]	0.243 [0.216]	0.794 [0.001]	0.247 [0.003]
M8	0.807 [0.001]	0.652 [0.001]	0.238 [0.212]	0.806 [0.001]	0.239 [0.213]	0.800 [0.001]	0.246 [0.003]
M9	0.785 [0.001]	0.654 [0.001]	0.253 [0.222]	0.783 [0.001]	0.255 [0.224]	0.771 [0.001]	0.264 [0.003]
M10	0.783 [0.001]	0.653 [0.001]	0.255 [0.223]	0.781 [0.001]	0.257 [0.225]	0.771 [0.001]	0.265 [0.003]
M11	0.784 [0.001]	0.652 [0.001]	0.257 [0.220]	0.782 [0.001]	0.259 [0.223]	0.782 [0.001]	0.258 [0.003]
M12	0.787 [0.001]	0.652 [0.001]	0.255 [0.218]	0.785 [0.001]	0.257 [0.220]	0.784 [0.001]	0.263 [0.003]

Table 2. Summary of model performance. $\mathcal{R}^2_{all}$ and $\mathcal{R}^2_{fix}$ correspond to the proportion of explained variance by the model and by the fixed variables only, respectively. PE corresponds to the mean predictive error (mean of observed minus predicted responses). For $\mathcal{R}^2$ and PE, the mean (over all iterations) and the standard deviation (in brackets), and the mean (over all observations) and the standard error of the mean (in brackets) are given, respectively. $\mathcal{R}^2_{all}$ and PE are given for both the P1 and P2 partition. 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 data set contains the remaining 6 provenances. See Table 1 and main text for description of model the components.

but the shortest in Madrid and Cáceres. In contrast, most provenances from the Central Spain gene pool were among the shortest, especially in Asturias and Bordeaux. The two provenances from the Corsican gene pool showed highly contrasted intercepts. The provenance Pinia (PIA) was taller than other provenances on average, especially in Asturias and Bordeaux. In contrast, the second Corsican provenance (Pineta, PIE), was shorter than other provenances on average but grew particularly well in Portugal. Provenances belonging to the south-eastern Spain gene pool (i.e. ORI and COM) also showed contrasted intercepts (trees from ORI were shorter than those from COM on average), but did not show strong differences between sites. Lastly, regarding provenances belonging to the northern African gene pool, the TAM provenance had the lowest growth in our dataset and was more likely to be taller in Cáceres and Madrid (the harsh Mediterranean sites) than in the other sites.

3.2 Potential drivers underlying height-growth variation

According to PE and $\mathcal{R}^2_{all}$ from both the training and test datasets of P1, *models M3* to *M6* had similar goodness-of-fit and predictive ability on new observations than *M1* and *M2* (Table 2 and section 6.1.1 of the Supplementary Information).

In *M3* (first model to include the climate similarity between sites) and all subsequent models until *M6*, the plastic component was mostly associated with height-growth variation between sites (σ^2_S with a median of 0.126), and only marginally with variation in site annual climate ($\sigma^2_{cs_{is}}$ with a median of 0.023; Fig. S13 & Table S12). However, the estimation of the intercepts and variances of the sites and site annual climate were uncertain (Figs. S13 to S15 & Tables S12 to S14), and the median and the credible interval of the site variance σ^2_S increased from *M1* to *M3* (where the climatic similarity between sites was included), suggesting that *M3* may hardly separate between these two variances (Tables S8 & S12).

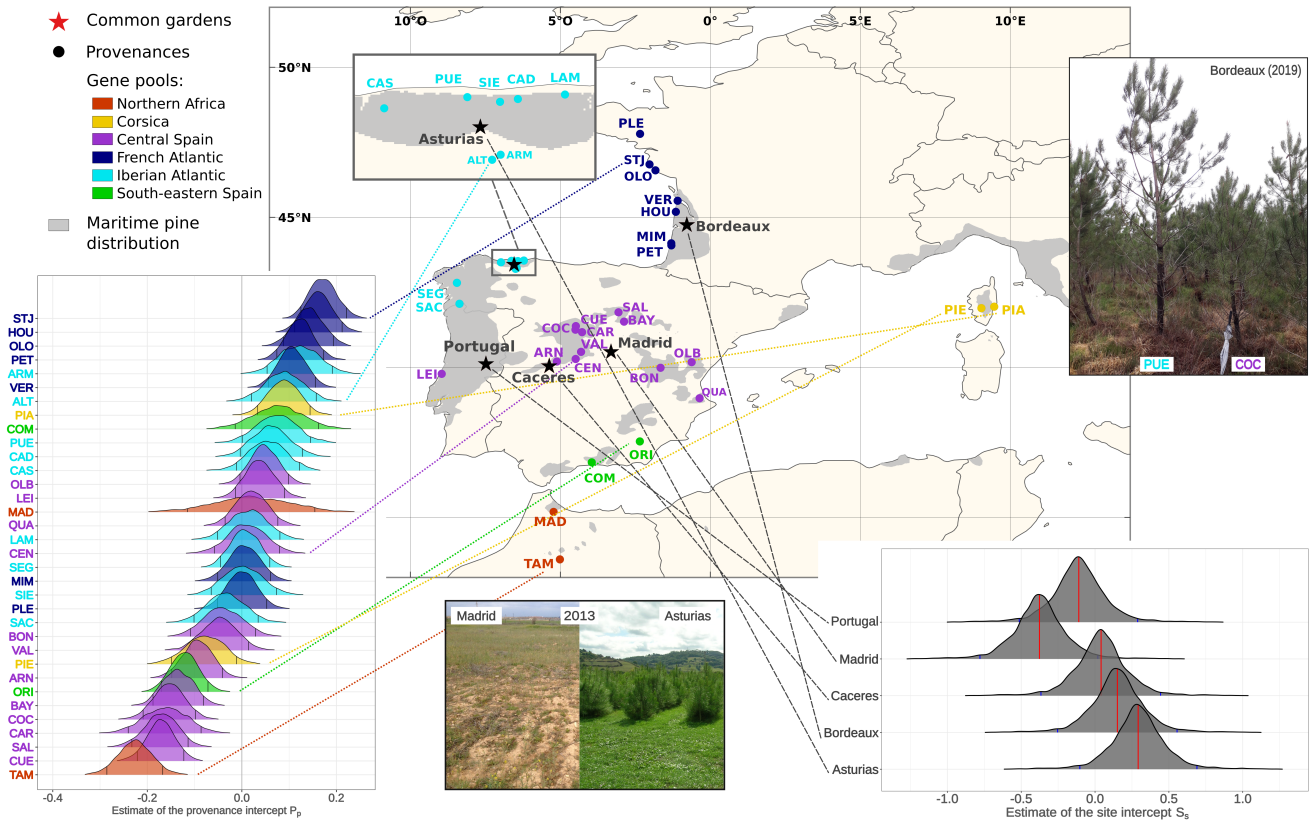

Figure 2. Posterior distributions of the site and provenance intercepts (S_S and P_P) in *model M1* on a map representation. Provenances are colored according to the main gene pool they belong to. The exact values of the median, standard deviation and 0.95 credible interval of the posterior distributions of the site and provenance intercepts are shown in Tables S9 and S10, respectively. The top left 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 range (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 being planted (2013).

To check this, we ran a supplementary model identical to *M3* but without the site intercepts S_s (see *model M3bis* in section 7 of the Supplementary Information). In this model, the variance related to the site annual climate was nearly equal and as uncertain as in *M3*, suggesting that our variance estimation of the plastic component in *M3* was robust (Table S27). However, the posterior distributions of the intercepts in *M3bis* were different from *M3*: height growth was positively associated with the annual climate in Bordeaux and Asturias, and negatively with the annual climate in Madrid and Cáceres, the two Mediterranean sites, and to a lesser extent in Portugal (Fig. S23 & Table S28).

In *M4*, the variance $\sigma^2_{g_j}$ between gene pools was as important as the variance σ^2_G between genotypes, but had a very wide posterior distribution (Table S15 & Fig. S16). From *M3* to *M4*, adding the gene pool intercepts g_j resulted in a decreased variance between provenances (from a median of 0.013 to a median of 0.006, Tables S12 & S15). This indicates redundant information between gene pools and provenances. Genotypes belonging to the French Atlantic gene pool, and to a lesser extent to the Iberian Atlantic gene pool, were on average taller than genotypes belonging to the northern Africa gene pool, and to a lesser extent to the Central Spain gene pool (Fig. S19; see also Fig. 5).

In *M5*, the total genetic variance of the Iberian Atlantic gene pool had a probability higher than 0.95 of being lower than that of the Corsican and south-eastern Spain gene pools, and a probability higher

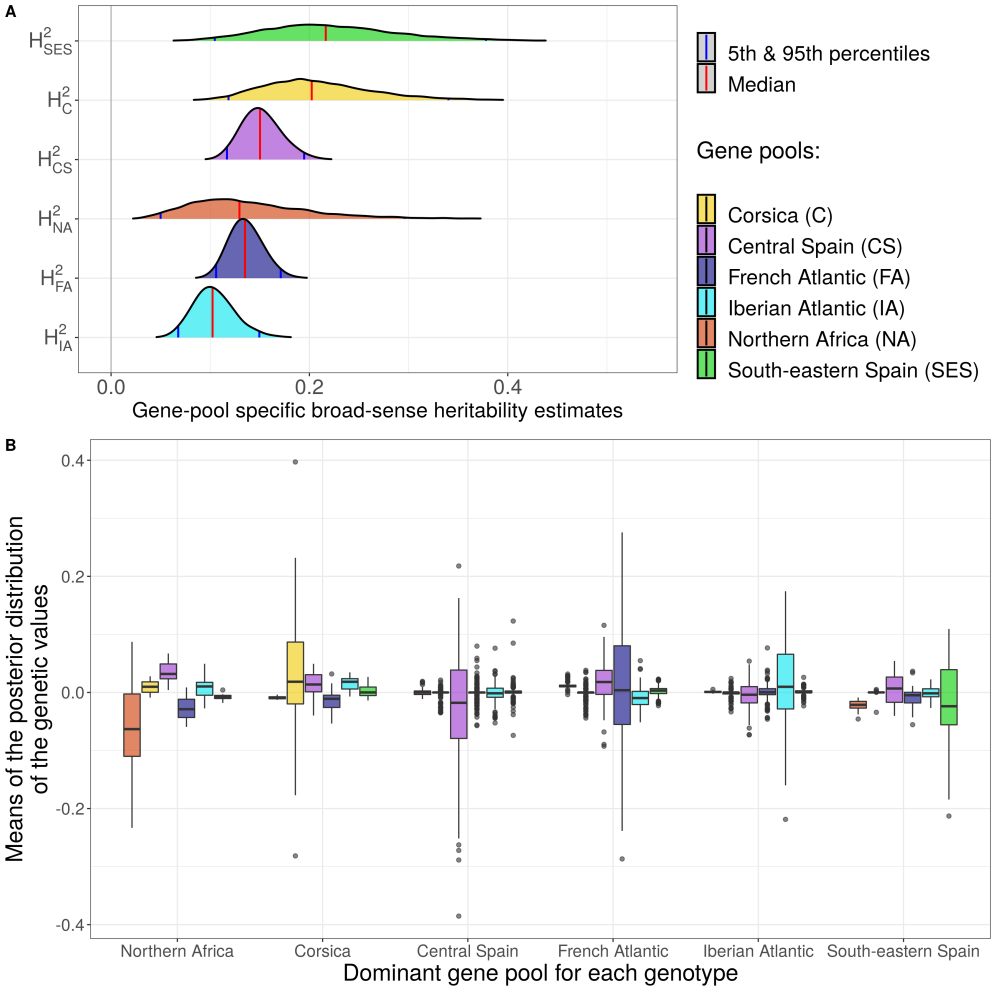

Figure 3. Posterior distribution of H^2_j, the gene pool-specific heritabilities obtained from *model M5* (A). Boxplot of the gene pool partial genetic values for genotypes grouped according to the main gene pool to which they belong (B). The colors represent the gene pool from which the partial genetic values were estimated. The lower, middle and upper parts of the hinges correspond to the 25%, 50% and 75% quantiles, respectively. The lower/upper whiskers extend from the hinge to the smallest/largest value no further than 1.5 time the IQR from the hinge (where IQR is the distance between the first and third quartiles).

than 0.90 of being lower than that of the Central Spain gene pool, and the total genetic variance of the French Atlantic gene pool had a probability higher than 0.90 of being lower than that of the Corsican and south-eastern Spain gene pools (Table S17; see also Fig. 3). However, this should be taken with caution as the total variance of the gene pools from south-eastern Spain, Corsica and northern Africa had wide posterior distributions, which is probably due to the small number of genotypes from these gene pools (see Table S3). In line with the genetic variance estimates, the medians of the gene-pool specific estimates of heritability H^2_j varied between 0.104 in the Iberian Atlantic gene-pool (95% CIs: 0.065-0.146) and 0.223 in the south-eastern Spain gene pool (95% CIs: 0.093-0.363) (Table S18 and Fig. 3A). Genotypes displayed genetic values mostly determined by the dominant gene pool from which they belong to, but also, to a lesser extent, by other gene pools (Fig. 3B). This is particularly relevant for admixed provenances, such as those from south-eastern Spain (see Fig. 1). For instance, for these provenances, the genetic values were partly determined by a negative influence of the partial genetic values associated with the northern Africa gene pool and a positive influence of those from the Central Spain gene pool (Fig. 3B).

In *M6*, the variance $\sigma^2_{cp_p}$ associated with the provenance climate-of-origin had a higher median than the variances attributed to either gene pools or provenances ($\sigma^2_{g_j}$ and σ^2_P; Table S19 & Fig. S18). Including the provenance climate-of-origin in *M6* resulted in a decreased variance of the provenances and gene pools, suggesting confounded effects (Tables S15 & S19). Once these confounded effects are taking into account, trees from climatic regions neighboring the Atlantic Ocean were generally the tallest (e.g. CAD, SIE, PUE, LAM and CAS in northwestern Spain; all provenances along the French Atlantic coast; Fig. S19). 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), was not among the tallest provenances (Fig. S19). Also, the Corsican provenances showed contrasted climate intercepts, with a positive influence on height growth for Pinia (PIA) but not for Pineta (PIE), which could explain their striking differences in height-growth patterns (Fig. 2). Finally, the four provenances from south-eastern Spain and northern Africa gene pools showed all negative climate intercepts (Fig. S19).

3.3 Combining climatic and genomic data to predict new observations and provenances

According to $\mathcal{R}^2_{all}$ and PE calculated on the training dataset (data used to fit the models) from P1, *models M7* and *M8* (combining genomic and climatic data) had a similar goodness-of-fit than *M1* and *M2*, which rely exclusively on the common garden design (Table 2). More precisely, *M7* and *M8* explained 90.7% and 79.7% respectively of the genetic variance explained in *M2*. *Models M9* to *M12* (including only one factor of those associated with the genetic component of height growth) had a lower goodness-of-fit than *M1* and *M2*, but still better than *M0*, which does not include the provenance/genotype intercepts (Table 2). Specifically, *M9* (demographic history), *M10* (provenance climate-of-origin), *M11* (global PEA counts) and *M12* (regional PEA counts) explained 37.2%, 32.6%, 34.9% and 41.9% respectively of the genetic variance explained in *M2*. Same ranking was obtained based on $\mathcal{R}^2_{all}$ and PE calculated on the test dataset from the P1 partition (observations not used to fit the models, but sampled from the same genotypes/provenances), confirming that the predictive ability on new observations was close to the model goodness-of-fit. Similar results were obtained based on the ELPD (see section 6.1.1 of the Supplementary Information).

Interestingly, the models ranked differently when comparing the predictive ability on new provenances, based on $\mathcal{R}^2_{all}$ and PE on the test dataset from the P2 partition (provenances not used to fit the models; Table 2), and which gives insight into the model generalizability across the species range. *Models M1* and *M2* failed to integrate the genetic component into height-growth predictions on new provenances (similar $\mathcal{R}^2_{all}$ and PE to that of *M0*), whereas *M7* and *M8* stand out clearly from the other models by having higher $\mathcal{R}^2_{all}$ (0.794 and 0.800, respectively, compared to 0.768 for *M0*) and lower PE (0.247 and 0.246, respectively, compared to 0.264 for *M0*; Table 2). *M11* and *M12* had a lower predictive ability on new provenances than *M7* and *M8*, but still better than the other models (*M0*, *M1*, *M2*, *M9* and *M10*; see Table 1). Results were similar with the P3 partition, where six other unknown provenances were included in the test dataset (Table S42).

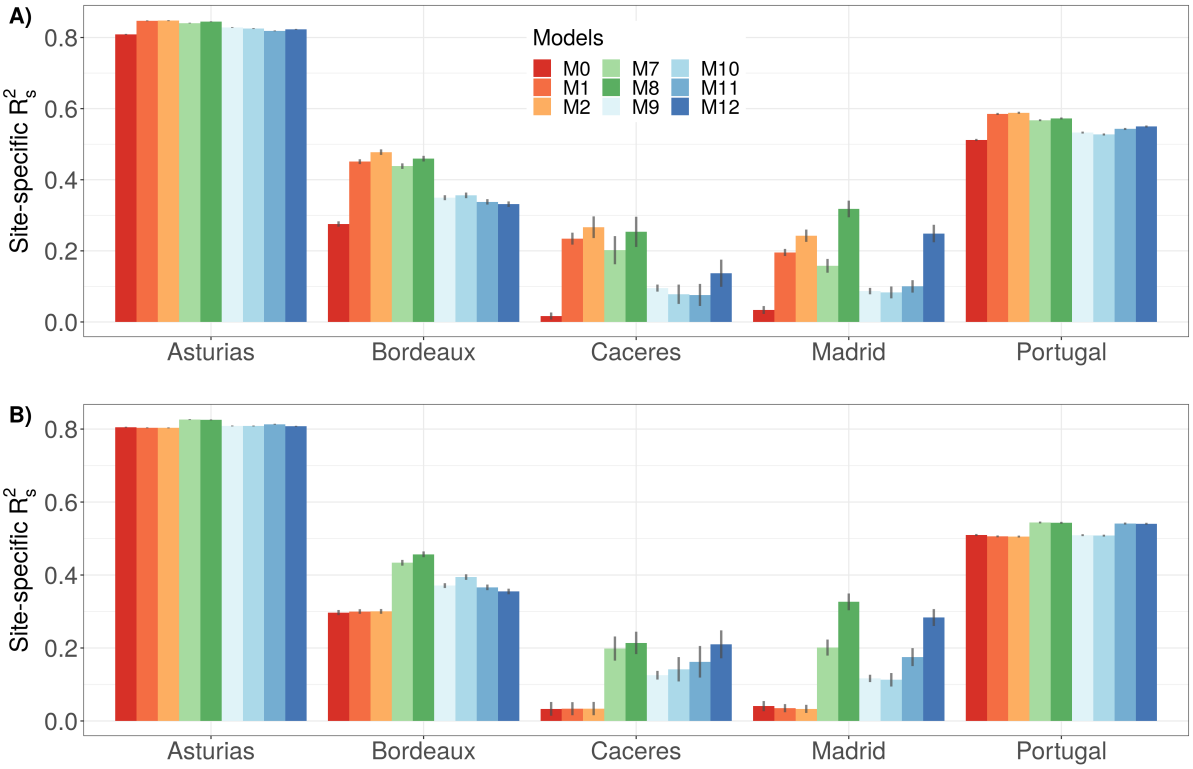

Figure 4. Site-specific explained variance ($\mathcal{R}^2_s$) of the test datasets (data not used to fit the models) in the P1 (A) and P2 (B) partitions, corresponding to new observations and new provenances, respectively. Bar heights correspond to the mean $\mathcal{R}^2_s$ of each model (over all iterations), and the gray lines correspond to the standard deviation intervals.

To assess whether models had the same predictive ability in different environments, we compared site-specific explained variance $\mathcal{R}^2_s$ in the test datasets from P1 and P2 partitions (Fig. 4 and Tables S6 & S31). Site-specific $\mathcal{R}^2_s$ varied greatly across sites, from around 0.8 in Asturias, under mild Atlantic climate, to very low values in Madrid and Cáceres (between 0 and 0.3 depending on the models), under harsher Mediterranean climate. Furthermore, differences in predictive ability across models were higher in Cáceres and Madrid, and to a lesser extent in Bordeaux. *M8* was the best model to explain height growth on new observations in Madrid and on new provenances in all sites. *M12* had also a good predictive ability overall, in particular in Cáceres and Madrid. We also compared the provenance-specific explained variance $\mathcal{R}^2_p$ in the test datasets from the P2 partition (Fig. S25 and Table S32). Provenance-specific $\mathcal{R}^2_p$ varied little across provenances, differences across models being also far less noticeable than for site-specific $\mathcal{R}^2_s$. Site and provenance-specific predictive errors (PE_s & PE_p) ranked the models similarly to $\mathcal{R}^2_s$ and $\mathcal{R}^2_p$ in each site and provenance (Tables S6, S31 & S32 and Figs. S8, S24 & S26).

Finally, in *models M7* and *M8* combining genomic and climatic data, population demographic history and PEA counts better explained the genetic component of height growth than the provenance climate-of-origin (Figs. 5 & S20). The provenance climate showed the largest association with height growth in Madrid, and to a lesser extent in Cáceres, where provenances from locations with warmer temperatures and higher precipitation in summer grew more (Figs. 5 and S20). These relationships did not hold when the provenance climate was included alone in *M10* (Fig. S21). In *M9* (including only population demographic history as predictor), the gene pool intercepts were similarly ranked as in *M4*, but the

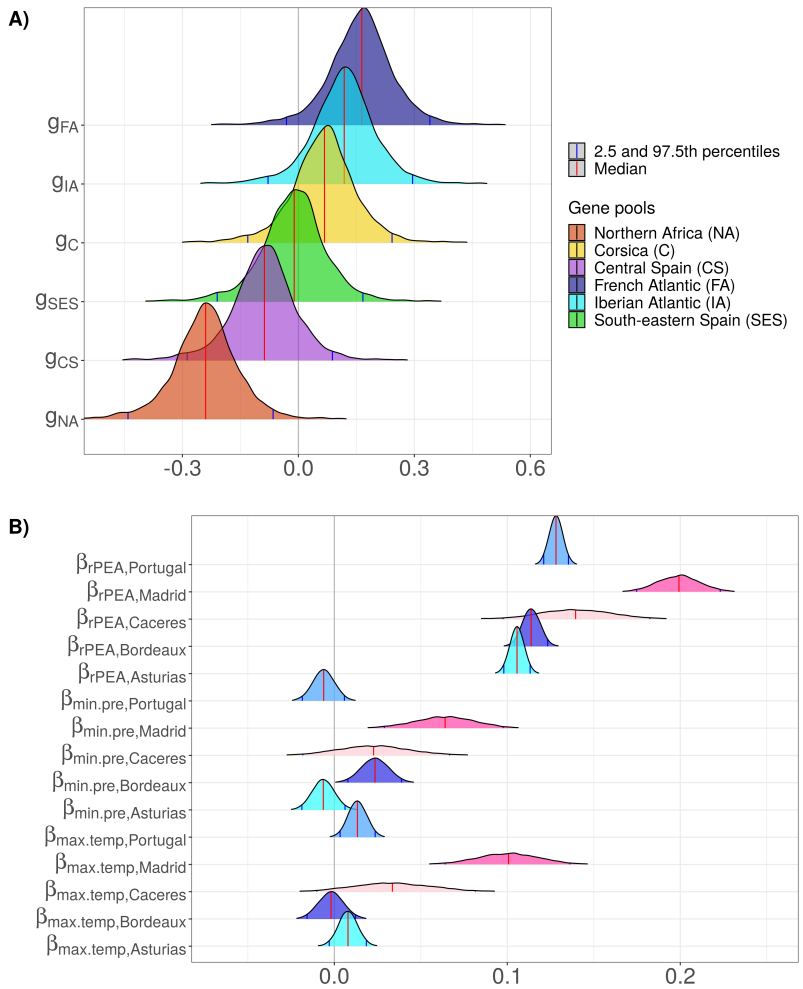

Figure 5. Posterior distributions of parameters from *model M8* fitted on the P1 partition, including the gene-pool intercepts g_j (A), and the site-specific slopes of the regional positive-effect alleles counts $\beta_{rPEA,s}$, the minimum precipitation during the driest month $\beta_{min.pre,s}$ and the maximum temperature of the warmest month $\beta_{max.temp,s}$ (B). Parameters specific to the same site are colored the same to aid visualization.

strength of their association with height was stronger, probably because there was no redundancy with the provenance and genotype intercepts (included in *M4* but not in *M9*; Figs. S19 & S21). Overall, population demographic history and PEA counts showed a robust association with height growth that did not change whether or not they were included in the models alone or with other predictors (Figs. S21, 5 & S20). Global PEA counts (gPEAs) were positively associated with height growth almost uniformly across all sites, while regional ones (rPEAs) had a much stronger positive association in Madrid, while remaining positive for the other sites too (Figs. 5 & S20).

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 the range-wide variation of height growth, a key adaptive trait in forest trees. First, our study adds further evidence that plasticity explains a large

part of height-growth variation (more than twice the genetic component in our study), and may be driven by multiple confounding factors including climate. Second, population demographic history and adaptation to climate, evaluated using provenance climate-of-origin as a proxy, showed confounded associations with height growth for most provenances. We also showed that provenances belonging to the Corsican and south-eastern Spain gene pools had higher heritabilities than provenances belonging to the French and Iberian Atlantic gene pools, which suggests that southern populations may adapt faster to changing selection pressures. Third, our study suggests that models combining population demographic history, provenance climate-of-origin and height-associated positive-effect alleles (PEAs) capture the genetic component underlying height-growth variation almost as well as models relying directly on the common garden design (via varying intercepts), and better predict height-growth variation of new provenances. 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), in particular in harsh environments (i.e. Madrid and Cáceres), suggesting different adaptive histories for different provenances. These results pave the way towards integrating genomics into large-scale predictive models of quantitative trait variation in maritime pine.

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 to 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, phenotypic plasticity explained more than twice height-growth variation than the genetic component (*model M1*). Interestingly, although climate plays a role, the plastic height-growth component of maritime pine is 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, often measured in common gardens by the genotype-by-environment interaction (G×E) (Des Marais et al. 2013, Sork 2018). In trees, G×E is particularly prevalent for growth traits (Li et al. 2017). Likewise, height growth in maritime pine is often associated with provenance or family-by-site interaction (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, we detected provenance-by-site interaction that was associated with height growth, but its magnitude was lower than the provenance, genotype or site main effects (*model M2*). There was no global pattern of local adaptation: some provenances had an optimal height growth close to their origin (e.g. provenances PUE, SIE and ARM close to Asturias), but others did not (e.g. provenances PET, VER and HOU close to Bordeaux; Fig. S12). Interestingly, some provenances from the same gene pools did not have a common plastic response across sites. In forest tree breeding programs, previous works argued that G×E may hinder the model transferability

across sites and populations (Resende Jr et al. 2012, Resende et al. 2012). In maritime pine, although our results suggest that large-scale predictions of height-growth variation will be only slightly impacted by not accounting for provenance-by-environment interaction, future work is necessary to assess the importance of the genetic variation of plasticity at the genotype level.

4.2 Potential drivers underlying the genetic component of height-growth variation

In maritime pine, population demographic history and adaptation to climate are partially confounded (Jaramillo-Correa et al. 2015, Nadeau et al. 2016). For example, in our study, 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 their demographic history and provenance climate-of-origin (Fig. S19). 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). Nevertheless, in contrast to the three examples above, for some provenances, the effect of demographic history and climate adaptation 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 be explained by adaptation to different environmental conditions (and in particular climate), but not by demographic history, 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. S7). These different adaptations within a same gene pool calls for a more targeted investigation of the Corsican gene pool.

Overall, maritime pine proved to be a particularly suitable model species to study the joint influence of population demographic history and climate adaptation 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 Barton 2019, Sohail et al. 2019). Nevertheless, it would be necessary to go beyond correlative approaches to better understand how demographic history and adaptation processes have shaped current patterns of quantitative trait variation at the species range scale (see, for example, Ovaskainen et al. 2011). Likewise, investigating trait ge-

netic 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 (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 and other biotic-related traits.

4.3 Gene pool-specific heritabilities: the telltale of distinct adaptive histories

Heritability variation across populations or gene pools can inform on differences in the drivers underlying their adaptive histories, such as evolutionary mechanisms (e.g. capacity of dispersion, selection strength, migration) and local environmental constraints (e.g. environmental heterogeneity).

Using CLONAPIN data from all sites except Bordeaux, Rodríguez-Quilón et al. (2016) showed that populations from the Mediterranean gene pools had a higher heritability than those from the Atlantic gene pools. Nonetheless, their heritability estimates did not consider population admixture, a notable feature in maritime pine (Fig. 1). In our study, we applied the recent methodology proposed by Muff et al. (2019) to calculate gene pool-specific heritabilities in a single model that accounts for population admixture, and showed that the Corsican gene pool, and to a lesser extent the south-eastern Spain gene pool, have higher total genetic variances (and so higher heritabilities) than the French and Iberian Atlantic gene pools (Fig. 3). Less markedly, the central Spain gene pool had also a higher total genetic variance (and heritability) than the Iberian Atlantic gene pool, despite their geographical proximity. Interestingly, provenances from the Corsican, south-eastern Spain, and central Spain gene pools are also the ones facing more contrasted climates. Indeed, Corsica and south-eastern Spain are mountainous areas, with strong environmental heterogeneity at small spatial scales. Central Spain is less contrasted spatially but experiences a high continentality, and thus strong daily and annually climatic variation. Taken together, these results may suggest that gene pools from regions with high environmental heterogeneity (in space and/or time) have also higher heritability. Indeed, theoretical and empirical works have proposed that high levels of adaptive genetic variance can be maintained in regions of high environmental heterogeneity (mainly spatial, but to a lesser extent also temporal) and some degree of gene flow between populations (Yeaman and Jarvis 2006, Yeaman and Otto 2011, McDonald and Yeaman 2018). Further research involving multiple adaptive traits and more detailed environmental data would be needed to confirm this hypothesis.

4.4 Towards predicting height growth outside common gardens

Anticipating how provenances will respond to new climatic conditions is key to guide forest conservation strategies and population translocations to compensate for climate change velocity (Aitken and Whitlock 2013). Predicting quantitative trait variation of provenances without planting them in costly and time-consuming experiments would increase the chances of fast adaptive decision-making and would allow to broaden the predictions to a larger number of species, outside the few for which common gardens currently exist. In our study, models combining genomic and climate data (*M7* and *M8*)

explained almost as much of the genetic component underlying height variability as models (*M1* and *M2*) based exclusively on the common garden design (similar explained variance when predicting new observations), and were much better predicting height growth of new provenances (see Table 2). This suggests that *M7* and *M8* were able to capture the drivers behind the genetic component of height growth and could thus be used for predictions at range-wide geographical scales. The joint use of phenotypic, genomic and climatic data, facilitated by freely-available large climatic datasets (e.g. Fréjaville and Benito Garzón 2018) and the advances in high-throughput genotyping, is therefore a major opportunity to anticipate how our forests will cope with climate change (Sork et al. 2013, de Villemereuil et al. 2016).

Models including only one genomic or climatic-related driver (*M9* to *M12*) predicted height-growth variation worse (both on new observations and provenances) than those combining them (*M7 & M8*), highlighting the complexity and multidimensionality of the genetic component underlying adaptive traits. In agreement with Mahony et al. (2020), these findings suggest that including genomic information (population demographic history and PEAs) in the *population response functions* commonly used in forest trees may considerably improve their predictive ability range-wide, since they have been based so far almost exclusively on the relationship of phenotypic traits in common gardens with the climate in the provenance location (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. 2019). We showed a high predictive ability of the population demographic history in maritime pine, which was expected as this species is known for its strong population structure and scattered distribution (Alberto et al. 2013, Jaramillo-Correa et al. 2015). Nonetheless, in species with more continuous populations (e.g. Scots pine; Alberto et al. 2013) and/or living under stronger climatic limitations, a higher influence of adaptation to climate would be expected. Further studies in forest tree species with contrasted population demographic history and life-history traits are needed to generalize our findings.

Interestingly, integrating genomic and climatic data in predictive models proved more useful in harsh environments, despite harsh environments having also the lowest proportion of explained variance (both for predictions of new observations and new provenances; Fig. 4). Harsh conditions can cause an increase in both environmental variance (resulting in hardly predictable trait variation; Charmantier and Garant 2005) and genetic variance (allowing to better discriminate among models capturing or not the genetic component), which can be connected to generally higher heritability in maritime pine common gardens or experiments under water limitation (Aranda et al. 2010, Rodríguez-Quilón et al. 2016, de Miguel et al. 2020). However, these studies are based on young plants and genetic variation could be expressed at later ontogenic stages in favorable environments. This is supported by the increasing heritability estimates in the French Atlantic common garden (Bordeaux) across years reported in de Miguel et al. (2020). In this line, a meta-analysis for wild animal populations reported higher heritability for morphometric traits in favorable environments (but not for life-history traits; Charmantier and Garant 2005). In forest trees, studies comparing heritability estimates across different environments are still scarce (Geber and Griffen 2003) and thus, determining whether the increased heritability in harsh environments observed in maritime pine is a global pattern would require further investigation.

Finally, the higher predictive ability of models including rPEAs over gPEAs suggests that most height-

associated alleles have mainly a local effect, i.e. conditionally neutral alleles (Tiffin and Ross-Ibarra 2014). This is especially noticeable in the Mediterranean region, where tree populations are expected to be under stronger selective pressure. Remarkably, only a few rPEAs were shared among geographical regions (i.e. environments) in our study (20% shared between Iberian and French Atlantic regions, 12% between the French Atlantic and Mediterranean regions, and 24% between the Iberian Atlantic and Mediterranean regions). The predominance of conditional neutrality over antagonistic pleiotropy has already been reported in plants (Prunier et al. 2012, Anderson et al. 2013). Interestingly, a previous study on maritime pine showed a high stability in polygenicity of height-associated markers between Atlantic and Mediterranean regions (de Miguel et al. 2020). Our results complement this study by suggesting that, despite having similar polygenicity, height growth variation in Mediterranean sites are unlikely to be affected by the same loci as in the other regions, probably as a result of the different gene pools that diverged in southern refugia during the last glaciation and then recolonized Europe during the Holocene.

4.5 Limitations and challenges

In forest trees most mortality (i.e. selection) takes place in early developmental stages. However, 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 Arregui 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 Bordeaux (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 natural 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. S35).

The integration of genomic data into range-wide adaptive-trait predictive models is in its infancy and still lacks a well-established framework, especially for non-model species such as forest trees. We used relatively few SNPs (5,165 SNPs) to cover the huge genome of maritime pine (28 Gbp; Zonneveld 2012), but given the high polygenicity of this trait (~7%; de Miguel et al. 2020), we still expected to identify a substantial number of height-associated positive-effect alleles (PEAs). We preferred to use PEA counts, as recently done in Mahony et al. 2020, to more commonly used polygenic scores (Pritchard et al. 2010, Browne et al. 2019, Fuller et al. 2020), to minimize the circularity of the analysis (i.e. we did not have to estimate allele-effect weights for each PEA based on the same dataset as the one used for model fitting). Using this approach, we obtained good predictive ability for both new observations and provenances. Resende et al. (2012) pointed out the difficulty of predicting height of provenances not included in the

training models; this prediction ability depends on the number of provenances used to fit the models and the strength of the genetic relationship between provenances (Jarquín et al. 2014). In our study, sampling a relatively large number of provenances representing all known gene pools in the species may partly explain the good predictive ability of the models. In addition, the integration in the models of the provenance climate-of-origin and the PEAs with site-specific slopes probably helped to better capture the G×E, one common difficulty to predict the genetic component of unknown provenances (Jarquín et al. 2014).

Lastly, about 18% of height-growth variation remained unexplained in our study. Recent studies have highlighted the important role that developmental stochasticity can play in 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).

5 Conclusion

In this study, we linked the *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 understand population vulnerability to future climates (Jaramillo-Correa et al. 2015, Rellstab et al. 2016). We showed that combining climatic and genomic information (i.e. provenance climate-of-origin, population demographic history and trait-associated positive-effect allele counts) can improve the predictions on unsampled provenances of a highly polygenic adaptive trait such as height, in particular in harsh environments. These results are promising towards anticipating the forest tree responses when faced with rapidly changing conditions or when transplanted to new environments, without having to evaluate them in common gardens. Our study focuses specifically on the height-growth genetic component of current populations, but the next step could be to integrate evolutionary processes (e.g. genetic drift in small populations, strong selection events, etc.) into the models, which would allow us to also 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 are very grateful to Ricardo Alía who contributed to the

design and establishment of the CLONAPIN network and provided comments on the manuscript. We thank all persons and institutions linked to the establishment and maintenance of field trials used in this study. Thanks are extended to Juan Majada for initiating and supervising the establishment of the CLONAPIN network. We thank Benjamin Brachi and Thibault Poiret for constructive comments on a previous version of the manuscript. 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 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.

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.

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.

Alberto, F. J., 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.

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.

Aranda, I., Alía, R., Ortega, U., Dantas, Â. K., and Majada, J. 2010. Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from four Pinus pinaster populations. *Tree Genetics & Genomes* **6**: 169– 178.

Ballouz, S., Pena, M. T., Knight, F. M., Adams, L. B., and Gillis, J. A. 2019. The transcriptional legacy of developmental stochasticity. *bioRxiv*.

Barghi, N., Hermisson, J., and Schlötterer, C. 2020. Polygenic adaptation: a unifying framework to understand positive selection. *Nature Reviews Genetics* **21**: 1– 13.

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.

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.

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.

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.

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.

Bürkner, P.-C. 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* **80**: 1– 28.

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.

Charmantier, A. and Garant, D. 2005. Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences* **272**: 1415– 1425.

Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecology Letters* **8**: 2– 14.

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.

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.

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.

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.

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*.

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

De Villemereuil, P. 2019. On the relevance of Bayesian statistics and MCMC for animal models. *Journal of Animal Breeding and Genetics* **136**: 339– 340.

De Villemereuil, P., Frichot, É., Bazin, É., François, O., and Gaggiotti, O. E. 2014. Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology* **23**: 2006– 2019.

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.

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.

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.

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*.

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.

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.

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.

Fréjaville, T., Vizcaíno-Palomar, N., Fady, B., Kremer, A., and Garzón, M. B. 2019. Range margin populations show high climate adaptation lags in European trees. *Global Change Biology* **26**: 484– 495.

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**.

Geber, M. A. and Griffen, L. R. 2003. Inheritance and Natural Selection on Functional Traits. *International Journal of Plant Sciences* **164**: S21– S42.

Gelman, A., Goodrich, B., Gabry, J., and Vehtari, A. 2019. R-squared for Bayesian Regression Models. *The American Statistician* **73**: 307– 309.

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.

Grattapaglia, D. and Resende, M. D. V. 2011. Genomic selection in forest tree breeding. *Tree Genetics & Genomes* **7**: 241– 255.

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.

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.

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.

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.

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.

Kawecki, T. J. and Ebert, D. 2004. Conceptual issues in local adaptation. *Ecology Letters* **7**: 1225– 1241.

Langlet, O. 1971. Two Hundred Years Genecology. *Taxon* **20**: 653– 721.

Latta, R. G. 2009. Testing for local adaptation in Avena barbata: a classic example of ecotypic divergence. *Molecular Ecology* **18**: 3781– 3791.

Leimu, R. and Fischer, M. 2008. A Meta-Analysis of Local Adaptation in Plants. *PLOS ONE* **3**: e4010.

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.

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.

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.

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.

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.

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.

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.

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.

McDonald, T. K. and Yeaman, S. 2018. Effect of migration and environmental heterogeneity on the maintenance of quantitative genetic variation: a simulation study. *Journal of Evolutionary Biology* **31**: 1386– 1399.

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.

Merilä, J. and Hendry, A. P. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications* **7**: 1– 14.

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.

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.

Neale, D. B. and Savolainen, O. 2004. Association genetics of complex traits in conifers. *Trends in Plant Science* **9**: 325– 330.

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.

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.

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.

Ovaskainen, O., Karhunen, M., Zheng, C., Arias, J. M. C., and Merilä, J. 2011. A New Method to Uncover Signatures of Divergent and Stabilizing Selection in Quantitative Traits. *Genetics* **189**: 621– 632.

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.

Pigliucci, M., Murren, C. J., and Schlichting, C. D. 2006. Phenotypic plasticity and evolution by genetic assimilation. *Journal of Experimental Biology* **209**: 2362– 2367.

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.

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.

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.

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.

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

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.

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.

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.

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

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.

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.

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.

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.

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

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.

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.

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.

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.

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., and Neale, D. B. 2013. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes* **9**: 901– 911.

Sork, V. L. 2018. Genomic Studies of Local Adaptation in Natural Plant Populations. *Journal of Heredity* **109**: 3– 15.

Stern, A. J., Speidel, L., Zaitlen, N. A., and Nielsen, R. 2020. Disentangling selection on genetically correlated polygenic traits using whole-genome genealogies. *bioRxiv*.

Thomson, C. E., Winney, I. S., Salles, O., and Pujol, B. 2018. A guide to using a multiple-matrix animal model to disentangle genetic and nongenetic causes of phenotypic variance. *bioRxiv*.

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.

Valladares, F., Gianoli, E., and Gómez, J. M. 2014a. Ecological limits to plant phenotypic plasticity. *New Phytologist* **176**: 749– 763.

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.

Vehtari, A., Gelman, A., and Gabry, J. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* **27**: 1413– 1432.

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.

Vogt, G. 2015. Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. *Journal of Biosciences* **40**: 159– 204.

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.

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.

Wolak, M. E. and Reid, J. M. 2017. Accounting for genetic differences among unknown parents in microevolutionary studies: how to include genetic groups in quantitative genetic animal models. *Journal of Animal Ecology* **86**: 7– 20.

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.

Yeaman, S. and Jarvis, A. 2006. Regional heterogeneity and gene flow maintain variance in a quantitative trait within populations of lodgepole pine. *Proceedings of the Royal Society B: Biological Sciences* **273**: 1587– 1593.

Yeaman, S. and Otto, S. P. 2011. Establishment and Maintenance of Adaptive Genetic Divergence Under Migration, Selection, and Drift. *Evolution* **65**: 2123– 2129.

Zas Arregui, 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.