

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

¹, Juliette Archambeau¹, Marta Benito Garzón¹, Frédéric Barraquand², Marina de Miguel Vega^{1,3}, Christophe Plomion¹ and Santiago C. González-Martínez¹

¹INRAE, Univ. Bordeaux, BIOGECO, F-33610 Cestas, France

²CNRS, Institute of Mathematics of Bordeaux, F-33400 Talence, France

³EGFV, Univ. Bordeaux, Bordeaux Sciences Agro, INRAE, ISVV, F-33882, Villenave d'Ornon, France

Abstract

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

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

21 1 Introduction

22 Global change is expected to have a profound impact on forests (Franklin et al. 2016, Seidl et al.
23 2017), and whether tree populations will be able to migrate or persist across their current range
24 is uncertain (Aitken et al. 2008). Assessing the potential of populations to accommodate future
25 environmental conditions requires a thorough understanding of the origin of variation in quanti-
26 tative traits subject to natural selection (Shaw and Etterson 2012, Alberto et al. 2013). To this aim,
27 a necessary first step is to quantify the plastic and genetic components of adaptive traits and their
28 interaction in multiple environments (Des Marais et al. 2013, Merilä and Hendry 2014), which has
29 been done extensively in forest trees (Franks et al. 2014). A second step consists in identifying
30 the underlying drivers of these components (Merilä and Hendry 2014). The plastic component
31 corresponds to the ability of one genotype to produce varying phenotypes depending on the envi-
32 ronment (Bradshaw 1965). Phenotypic plasticity can help individuals to overcome new conditions
33 up to a certain threshold (Nicotra et al. 2010), and can be to some extent genetically assimilated
34 and therefore involved in the evolutionary process of adaptation (Pigliucci et al. 2006). The ge-
35 netic component can stem from both neutral (e.g. population demographic history and genetic
36 drift) and adaptive processes (e.g. adaptation to local biotic and abiotic environments), both pro-
37 cesses implying changes in allele frequencies. Populations are locally adapted when they have
38 higher fitness in their own environment than populations from other environments (Kawecki
39 and Ebert 2004). In forest trees, a large amount of work highlighted the importance of climate in
40 driving the plastic and genetic responses of quantitative traits to new environmental conditions
41 (Savolainen et al. 2007, Valladares et al. 2014b). However, it is still unclear how multiple and
42 interacting drivers underlying quantitative trait variation could be combined to improve predic-
43 tions of population responses to global change. The increasing availability of genomic data opens
44 new opportunities to boost prediction accuracy, which is critical for breeding (i.e. genomic se-
45 lection; Grattapaglia and Resende 2011), to anticipate future distribution of natural populations
46 (e.g. Razgour et al. 2019), or to support the ongoing development of assisted gene flow strategies
47 aiming to help populations adapt to future environments (Browne et al. 2019, Mahony et al. 2020,
48 MacLachlan et al. 2021).

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

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

69 The advent and generalization of genomic tools have enhanced our understanding of adaptive
70 and neutral genetic processes resulting in trait variation, and their relationship with climatic
71 gradients (Savolainen et al. 2013, Sork 2018, Leroy et al. 2020). Integrating genomic information
72 into quantitative trait prediction would be highly valuable to consider intraspecific variability at
73 a finer scale than in current models (Mahony et al. 2020), thereby probably improving model ac-
74 curacy, especially for populations not previously planted in commons gardens. More specifically,
75 rapidly growing knowledge on trait-associated alleles identified by Genome-Wide Association
76 Studies (GWAS) is promising for anticipating the genetic response of populations to new envi-
77 ronments (Exposito-Alonso et al. 2018, Browne et al. 2019). For example, Mahony et al. (2020)
78 used counts of alleles positively associated with the traits of interest (PEAs) to describe patterns
79 and identify drivers of local adaptation in lodgepole pine. Recent studies have shown that most
80 quantitative traits are highly polygenic (see reviews in Pritchard et al. 2010, Barghi et al. 2020;
81 and de Miguel et al. 2020 for maritime pine) and that the effect of trait-associated alleles may
82 vary across environments (Anderson et al. 2013, Tiffin and Ross-Ibarra 2014), which complicates
83 the use of genomic information in trait prediction. In addition, patterns in allele frequencies
84 induced by population demographic history are often correlated with environmental gradients
85 (Latta 2009, Alberto et al. 2013, Nadeau et al. 2016), which makes difficult to separate the signature
86 of population structure from that of adaptive processes (Sella and N. H. Barton 2019, Sohail et al.
87 2019). At the species range scale, population structure hinders the use of genomic relationship
88 matrices, which provide more accurate estimates of genetic parameters (e.g. breeding values,
89 additive and non-additive variance) within breeding populations than previously used pedigree-
90 based approaches (Bouvet et al. 2016, El-Dien et al. 2018). Indeed, admixed populations or distinct
91 genetic groups may present different means and variances of their genetic values, which requires
92 new statistical methods to estimate them (e.g. Muff et al. 2019). Thus, integrating genomic infor-
93 mation into quantitative trait prediction in natural populations, while highly valuable, remains
94 challenging.

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

98 to study current and future responses to climate (Savolainen et al. 2007, Alberto et al. 2013).
99 Moreover, forest trees remain largely undomesticated (including those species with breeding
100 programs) and, therefore, genetic variation in natural populations has been little influenced by
101 human-induced selection (Neale and Savolainen 2004). However, forest trees have also large and
102 complex genomes (especially conifers; Mackay et al. 2012), that show a rapid decay of linkage dis-
103 equilibrium (Olson et al. 2010), and extensive genotyping would be needed to identify all (most)
104 relevant polymorphisms underlying (highly polygenic) quantitative traits (Neale and Savolainen
105 2004, Jaramillo-Correa et al. 2015). In addition, although early results have been convincing in
106 predicting trait variation within tree breeding populations (i.e. using populations with relatively
107 low effective population size; Resende Jr et al. 2012, Resende et al. 2012, Jarquín et al. 2014), pre-
108 dicting the genetic component of trait variation across populations or geographical regions of
109 forest trees remains poorly explored.

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

2 Materials & Methods

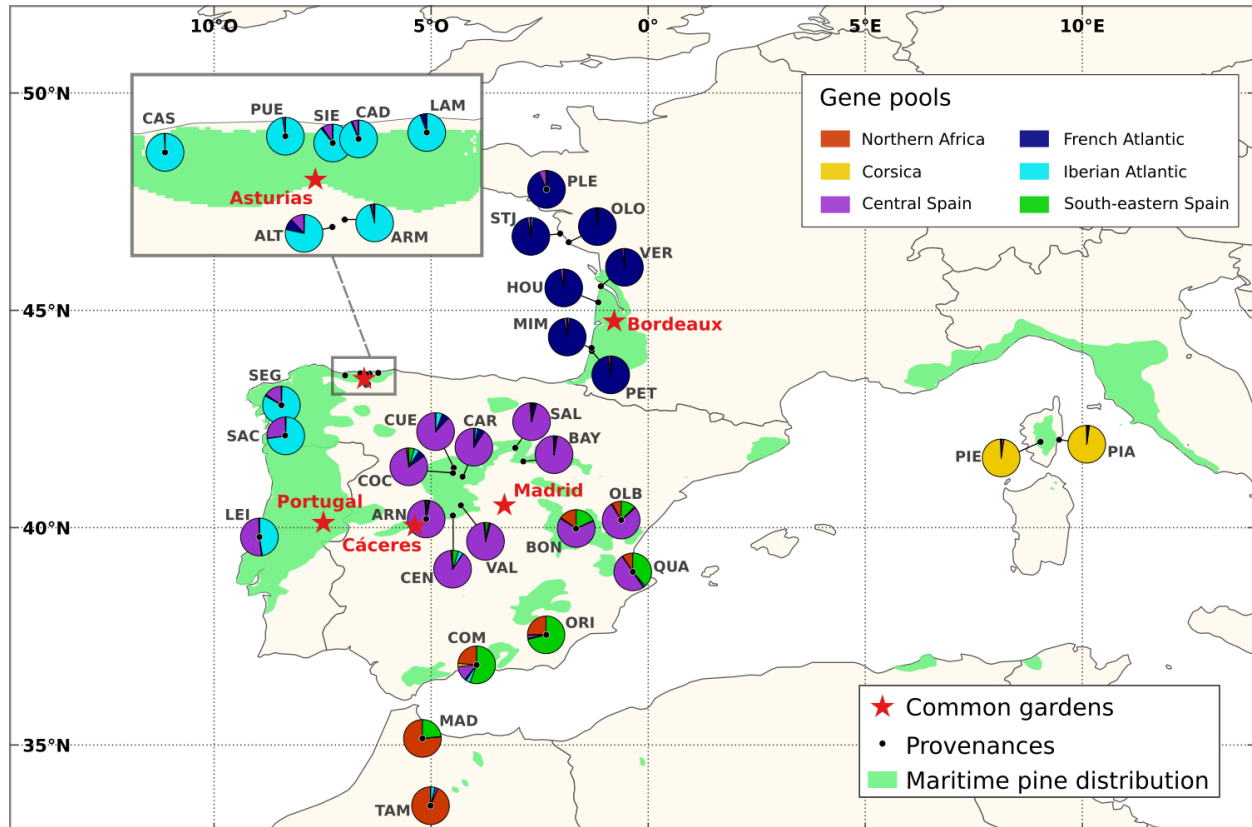
2.1 Plant material and phenotypic measurements

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

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

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

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



173 2.2 Gene pool assignment and positive-effect alleles (PEAs)

174 DNA was extracted from leaves collected in the Asturias common garden and genotyped with a
175 9k Illumina Infinium SNP assay (described in Plomion et al. (2016)), resulting in 5,165 high-quality
176 polymorphic SNPs scored on 523 genotypes. There were on average only 3.3 missing values per
177 genotype (ranging between 0 and 142). For each genotype, the proportion belonging to each gene
178 pool was estimated in Jaramillo-Correa et al. (2015), using nine nuSSRs as well as a subset of the
179 same SNPs as in our study (1,745 SNPs) and the Bayesian approach available in Structure v. 2.3.3
180 (Pritchard et al. 2000; Table S3). This gene pool assignment aimed at reflecting the neutral genetic
181 structure in maritime pine, which results from population demographic history and genetic drift,
182 but may also arise from different selective histories across gene pools.

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

199 2.3 Climatic data

200 In forest trees, large-scale patterns of allele frequencies or quantitative trait variation are known
201 to be associated with climatic variables related to mean temperature and precipitation (e.g. Eckert
202 et al. 2010, McLane et al. 2011, Leites et al. 2019, Fréjaville et al. 2020, Mahony et al. 2020), or
203 episodic climatic conditions, such as summer aridity or maximum temperatures (Rehfeldt et al.
204 2003, Grivet et al. 2011, McLane et al. 2011, Jaramillo-Correa et al. 2015, Fréjaville et al. 2020).
205 As climate change will cause major changes in temperature and precipitation in the near future,
206 particularly in the Mediterranean basin, there is a need to understand the complex influence of
207 climatic variables on quantitative trait variation. We extracted monthly and yearly climatic data
208 from the EuMedClim database with 1 km resolution (Fréjaville and Benito Garzón 2018). The
209 climatic similarity among test sites was described by a covariance matrix Ω including six variables
210 related to both extreme and average temperature and precipitation in the test sites during the year
211 preceding the measurements, and with at most a correlation coefficient of 0.85 among each other
212 (see section 3.1 in the Supplementary Information for more details). The climatic similarity among
213 provenances was described by a covariance matrix Φ including four variables related to the mean
214 temperature and precipitation in the provenance locations over the period from 1901 to 2009 (i.e.
215 representing the climate under which provenances have evolved), and with at most a correlation
216 coefficient of 0.77 among each other (see section 3.2 in the Supplementary Information for more
217 details).

218 2.4 Hierarchical height-growth models

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

240 2.4.1 Baseline *models M1* and *M2*: separating the genetic and plastic components of 241 height-growth variation

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

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

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

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

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

248 between provenance and site ($S_s P_p$). We also performed a model without the genetic component
 249 (called $M0$) whose outputs are reported in the Supplementary Information.

250 2.4.2 Explanatory *models M3* to *M6*: potential drivers underlying height-growth vari- 251 ation

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

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

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

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

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

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

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

276 **2.4.3 Predictive *models M7* to *M12*: combining climatic and genomic information to** 277 **improve predictions**

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

287 were expressed as follows (here for $M7$):

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

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

295 All models were inferred in a Bayesian framework as this approach better handles unbalanced
296 and multilevel designs (Clark 2005) and also to better propagate sources of uncertainty from
297 data and parameter values into the estimates (de Villemereuil 2019). Priors used in the models
298 were weakly informative and are provided in section 4.2 of the Supplementary Information. To
299 build the models, we used the *brms* package (Bürkner 2017), based on the no-U-turn sampler
300 algorithm. Models were run with four chains and between 2,000 and 3,000 iterations per chain
301 depending on the models (including 1,000 warm-up samples not used for the inference). All
302 analyses were undertaken in R version 3.6.3 (R Core Team 2020) and scripts are available at <https://github.com/JulietteArchambeau/HeightPinpinClonapin>.
303

304 2.5 Comparing model goodness-of-fit and predictive ability

305 Three partitions of the data (P1, P2 and P3) were used to evaluate model goodness-of-fit (i.e.
306 in-sample explanatory power, using training datasets) and predictive ability (out-of-sample pre-
307 dictive power, using test datasets). In P1, we aimed to predict new observations, an observation
308 being a height-growth measurement in a given year on one individual. P1 corresponds to a ran-
309 dom split of the data between 75% of observations used to fit the models (the training dataset of
310 24,840 observations) and 25% of observations used to evaluate model predictions (the test dataset
311 of 8,281 observations). Notice that the test dataset of the P1 partition was not totally independent
312 from the training dataset as it belongs to the same genotypes/provenances and blocks/sites. In
313 P2 and P3, we aimed to predict new provenances. P2 corresponds to a random split between a
314 training dataset of 28 provenances and a test dataset containing the remaining 6 provenances. P3
315 corresponds to a non-random split between a training dataset of 28 provenances and a test dataset
316 containing 6 provenances with at least one provenance from each under-represented gene pool

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

320 To evaluate the model goodness-of-fit, we calculated the in-sample (in the training dataset) pro-
321 portion of the variance explained by each model m in each common garden s , conditional on the
322 age effect, such as: $\mathcal{R}_{m,s}^2|age = (V_{\text{pred}_{m,s}} - V_{\text{age}_{2s}})/(V_y - V_{\text{age}_{2s}})$, where $V_{\text{pred}_{m,s}}$ is the variance of the
323 modeled predictive means from model m in site s , V_y the phenotypic variance in the site s and
324 $V_{\text{age}_{2s}}$ the variance explained by the age effect in the model $M2$ in site s . We used V_{age_2} of model
325 $M2$ and not of model m because the variance predicted by the different fixed effects of some of
326 the models ($M7$ to $M12$) could not be properly separated. Moreover, as $M2$ is the model with the
327 highest predictive ability among the models relying only on the common garden design (Table
328 S4), it constitutes an adequate baseline for model comparison. In addition, for baseline models $M1$
329 and $M2$, we also calculated the in-sample proportion of the variance explained by the different
330 model components (i.e. genetic, environment and genetic \times environment) conditional on the age
331 effect, e.g. for the genetic component in $M1$: $\mathcal{R}_{1,gen}^2|age = (V_{\text{pred}_{1,gen}} - V_{\text{age}_1})/(V_y - V_{\text{age}_1})$ where
332 $V_{\text{pred}_{1,gen}}$ is the variance explained by the genetic component (including the provenance and clone
333 effects) in $M1$, V_y the phenotypic variance and V_{age_1} the variance explained by the age effect in
334 $M1$.

335 Finally, to evaluate the model predictive ability, we calculated the out-of-sample (in the test
336 dataset) proportion of the variance predicted by each model m in each common garden s condi-
337 tional on the age effect, that we called *prediction* $\mathcal{R}_{m,s}^2|age$. Details about calculating *prediction*
338 $\mathcal{R}_{m,s}^2|age$ and some supplementary indexes used for model comparison are presented in section
339 5 of the Supplementary Information.

340 **3 Results**

341 **3.1 Underlying drivers of height-growth variation**

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

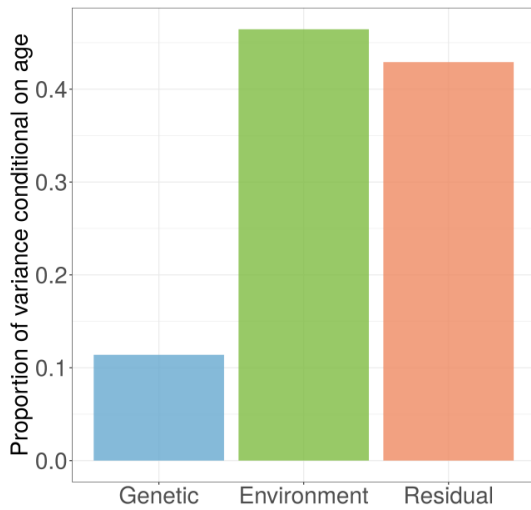
351 credible intervals (Tables S15 & S18). The plastic component was largely driven by the variance
352 among sites (σ_S^2), with very little contribution of the variance among blocks (σ_B^2 ; S15). Trees grew
353 the least in Madrid and the most in Asturias (fig. 3 & Table S16). The genetic component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

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

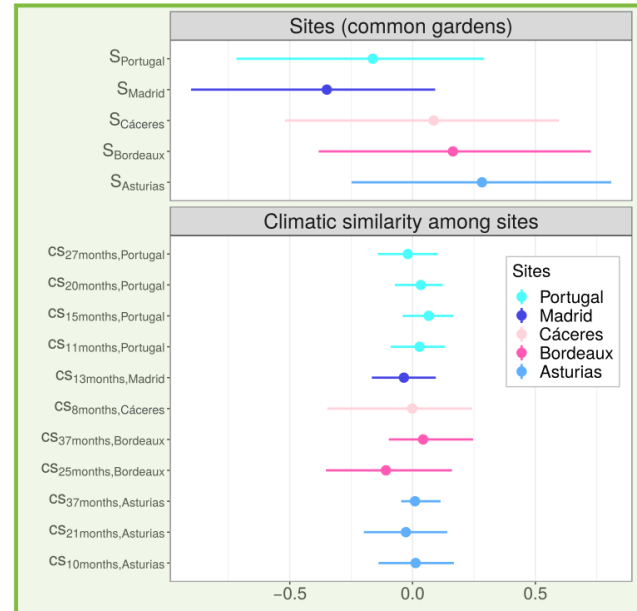
365 Based on *M6*, the genetic component of height growth was mostly determined by the climatic
366 similarity among provenances and to a lesser extent by the gene pool assignment (fig. 2C &
367 Table S29). However, the effects of the gene pools and climatic similarity among provenances
368 were partially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

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

A) Variance partitioning (*model M1*)



B) Plastic (environment) component (*model M3*)



C) Genetic component (*model M6*)

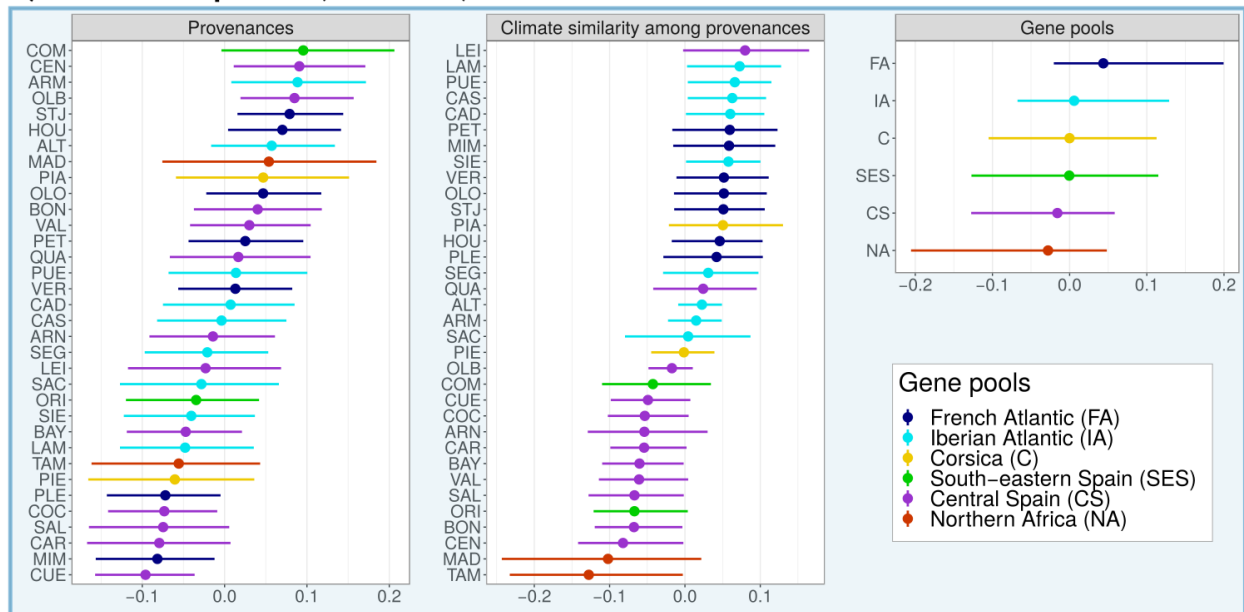
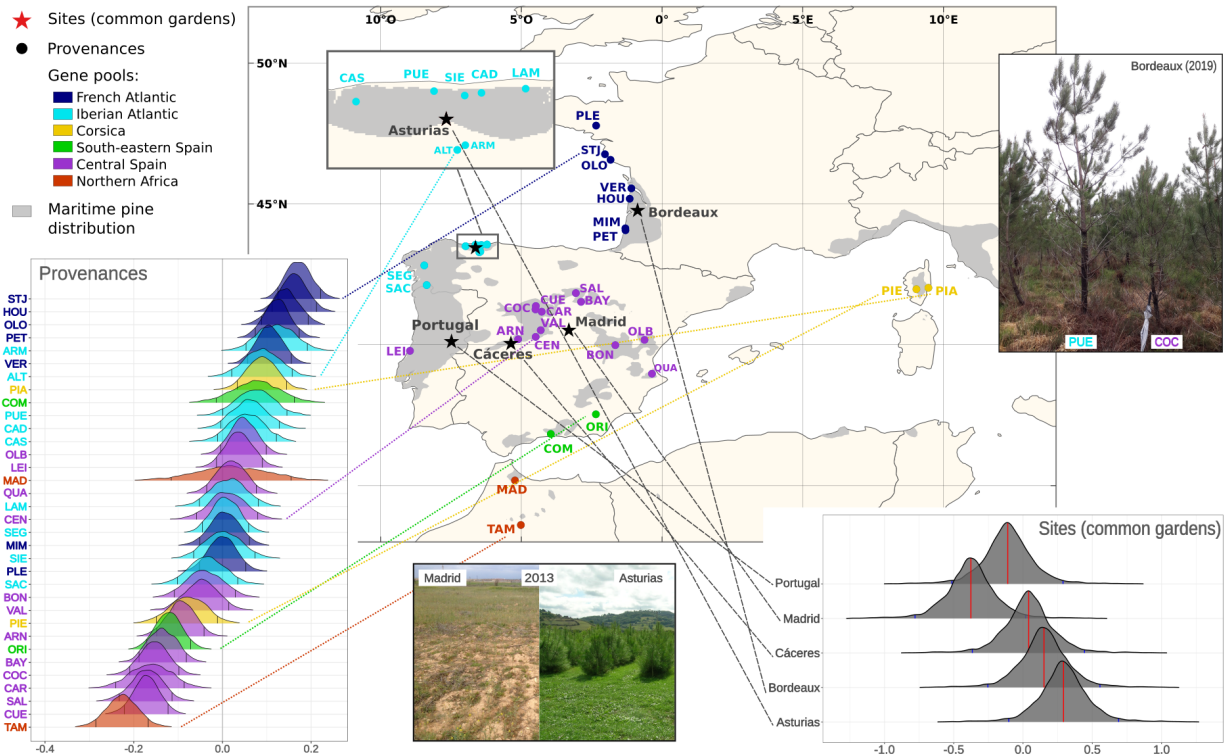


Figure 3. Posterior distributions of the site and provenance intercepts (S_s and P_p) in *model M1* on a map representation. Provenances are colored according to the main gene pool they belong to. The exact values of the median, standard deviation and 0.95 credible interval of the posterior distributions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



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

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

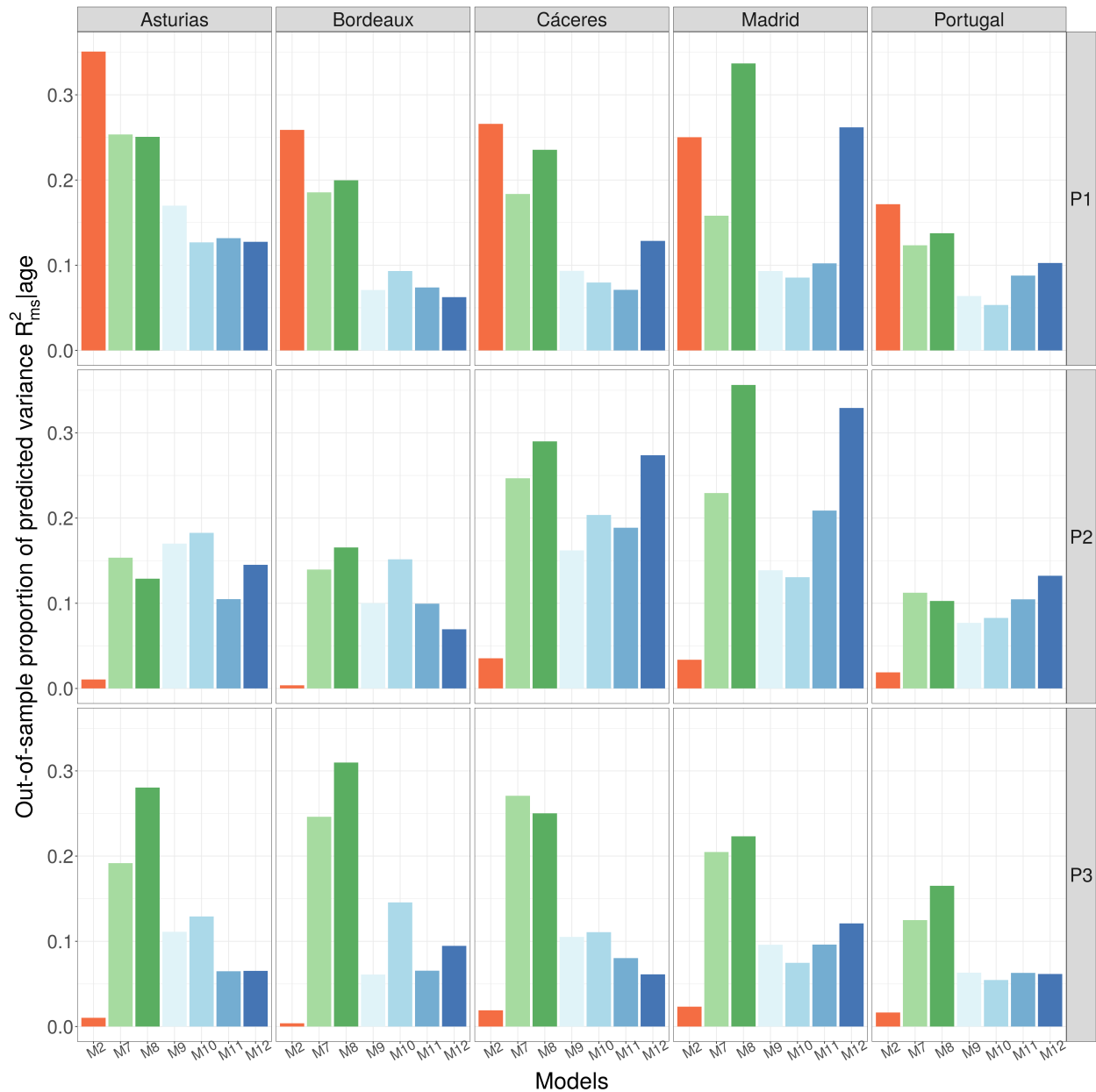
396 in Madrid (fig. S10).

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

411 4 Discussion

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

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



430 **4.1 Predominant role of height-growth plasticity**

431 Plants are known for their remarkable phenotypic plasticity to changing environments (Bradshaw
432 1965). In long-lived forest trees, the plastic component of quantitative trait variation estimated
433 based on the common garden design is generally higher than the genetic component (Franks et
434 al. 2014, Benito Garzón et al. 2019), e.g. in maritime pine (Chambel et al. 2007, Corcuera et al.
435 2010, de la Mata et al. 2012, Vizcaíno-Palomar et al. 2020). This plastic component is also generally
436 associated with the climatic conditions experienced by the trees (Franks et al. 2014, Benito Garzón
437 et al. 2019), allowing them to overcome changing climate up to a certain threshold (Matesanz et
438 al. 2010, Nicotra et al. 2010, Valladares et al. 2014a). In our study, the plastic component of height
439 growth was largely higher than the genetic component (fig. 2) and, although climate plays a role,
440 was likely to be driven by multiple and interacting drivers including the biotic environment, soil
441 quality, and other factors not considered in our study.

442 Plants also present an important genetic variation in plasticity (i.e. the genotype-by-environment
443 interaction, $G \times E$; Des Marais et al. 2013, Sork 2018), often approximated by the family or provenance-
444 by-site interaction in forest tree common gardens, as is the case in our study. $G \times E$ is particularly
445 prevalent for growth traits in trees (Li et al. 2017), as already shown in maritime pine (Alía et al.
446 1997, Corcuera et al. 2010, Correia et al. 2010, de la Mata et al. 2012; but see Chambel et al. (2007)
447 where no provenance-specific responses were observed under two different watering regimes).
448 In our study, provenance-by-site interaction was only weakly associated with height growth and
449 the proportion of variance it explained was not different from zero (*model M2*; Table S5). Pre-
450 vious work in the context of tree breeding argued that $G \times E$ may hinder model transferability
451 across sites and populations (Resende Jr et al. 2012, Resende et al. 2012). In maritime pine, our
452 results suggest that large-scale predictions of height-growth variation will be only marginally
453 impacted by not accounting for provenance-by-environment interaction. However, further work
454 is necessary to assess the importance of the genetic variation of plasticity at the genotype level.

455 **4.2 Potential drivers underlying height-growth genetic component**

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

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

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

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

504 **4.3 Towards integrating genomics into population response functions**

505 Anticipating how provenances will grow in new environments is key to guide forest conserva-
506 tion strategies and population translocations to compensate for rapid climate change (Aitken and
507 Whitlock 2013). To date, population response functions based on the climate in the provenance
508 location have been the most widely used method for anticipating trait values when transplanting
509 provenances in new environments (Rehfeldt et al. 1999, 2003, O’Neill et al. 2008, Wang et al. 2010,
510 Pedlar and McKenney 2017, Rehfeldt et al. 2018, Fréjaville et al. 2020). Genome-informed pre-
511 dictive modeling of key adaptive traits is highly promising as it may provide a mean to further
512 integrate adaptive or neutral genetic variation in the predictions, and to consider intraspecific
513 variability at a finer scale than current models, thus gaining in prediction accuracy (Holliday et
514 al. 2017). In valley oak, Browne et al. (2019) used genomic estimated breeding values (GEBVs; sum
515 of the marker predicted effects, also known as polygenic scores) to identify genotypes that will
516 grow faster under future climates. In lodgepole pine, Mahony et al. (2020) showed that phenotype-
517 associated positive-effect alleles (PEAs, as used in our study) can predict phenotypic traits (e.g.
518 cold injury) as well as climatic or geographical variables. In our study, we investigated whether
519 including genomic information related to past demographic and selective processes resulting in
520 distinct gene pools and counts of trait-associated alleles could improve range-wide height-growth
521 predictions in maritime pine. Models combining climatic conditions in the provenance location,
522 gene pool assignment, and PEAs captured most of the genetic component of height-growth vari-
523 ation (see fig. S10) and better predicted height growth of new provenances, compared to models
524 relying exclusively on the common garden design or models including separately climatic or ge-
525 nomic information (see fig. 4). This suggests that range-wide trait prediction would benefit from
526 jointly considering different sources of information (i.e. climatic and genomic), even though
527 they may have overlapping effects (e.g. confounded effects of provenance climate-of-origin and
528 gene pool assignment), as it may help to embrace the complexity and multidimensionality of the
529 genetic component underlying quantitative traits. Noticeably, regional PEAs were generally bet-
530 ter predictors of height growth in new provenances than gene pool assignment or provenance
531 climate-of-origin as, when they were included alone in the models, they made better predictions
532 in the driest common gardens (Madrid, Cáceres and Portugal) and similar ones to models com-
533 bining multiple drivers of height growth variation in all common gardens except Bordeaux (P2
534 partition in fig. 4). Although this highlights the major role that trait-associated alleles identified
535 using GWAS may play in predictive modeling, predicting traits of new provenances depends also
536 on the number of provenances used to fit the models and the strength of the genetic relation-
537 ship among them (Resende et al. 2012, Jarquín et al. 2014, Moghaddar et al. 2014, Hidalgo et al.
538 2016). This was reflected in our study by better predictive ability on new provenances in the P2
539 partition (random) compared to the P3 partition (containing provenances from underrepresented
540 gene pools) for models including climatic and genomic information separately but not for mod-
541 els considering both jointly (fig. 4). Thus combining multiple sources of information may also be
542 particularly relevant for predicting traits in marginal or difficult-to-access populations, as they

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

545 The high predictive ability of PEAs, both alone and combined with climatic and gene pool infor-
546 mation, was somehow unexpected given the sparse genomic sampling in our study: 5,165 SNPs
547 to cover the 28 Gbp maritime pine genome (Zonneveld 2012). Indeed, conifers have particularly
548 huge genomes, generally ranging from 18 to 35 Gbp (Mackay et al. 2012) and thus rendering the
549 current cost of whole-genome resequencing prohibitive (Holliday et al. 2017). Targeted geno-
550 typing approaches, such as the one used in the present study, select candidate genes based on
551 previous population and functional studies, thus allowing to include potential targets of selec-
552 tion and climate adaptation, but probably inducing an ascertainment bias (Jaramillo-Correa et al.
553 2015). However, as height is a particularly polygenic trait (degree of polygenicity estimated at
554 $\sim 7\%$ in de Miguel et al. 2020), we were able to identify a considerable number of PEAs despite
555 the weak genome coverage of our study. Further genomic sampling would be highly valuable
556 to capture the polygenic architecture of height more broadly, turning PEAs into much better
557 predictors than the provenance climate-of-origin or the gene pool assignment, and ultimately
558 making climatic data redundant, at least for main range populations (see above for marginal pop-
559 ulations). This would also allow to characterize the genetic variation within provenances more
560 precisely, thereby increasing the estimation accuracy and reducing the residual variance. Similar
561 to Mahony et al. (2020) and MacLachlan et al. (2021) who selected the positive-effect alleles as
562 the 1% of SNPs that showed the strongest association with phenotypes (estimated via a GWAS
563 performed on 18,525 SNPs), we used PEA counts instead of the more commonly used polygenic
564 scores (Pritchard et al. 2010, Browne et al. 2019, Fuller et al. 2020). Unlike polygenic scores, PEAs
565 do not account for allele effect sizes, thus minimizing the circularity of the analysis (i.e. effect
566 sizes that are estimated based on the same dataset as the one used for the models, only serve
567 for PEAs identification) and potentially enhancing the prediction accuracy across genetic groups
568 compared to polygenic scores. Indeed, low observed transferability of polygenic scores across
569 genetic groups (Martin et al. 2017, N. Barton et al. 2019, Martin et al. 2019) may stem from vary-
570 ing effect sizes of "peripheral" alleles (i.e. alleles indirectly affecting the phenotype), as suggested
571 in Mathieson 2021).

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

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

584 A last noticeable results was that rPEAs (positive-effect alleles identified in specific geograph-
585 ical regions, i.e. particular environments) had generally a higher predictive ability than gPEAs
586 (positive-effect alleles identified range-wide) (Figure 4). Interestingly, only a small proportion of
587 rPEAs were shared among geographical regions in our study (20% shared between the Iberian and
588 French Atlantic regions, 12% between the French Atlantic and Mediterranean regions, and 24%
589 between the Iberian Atlantic and Mediterranean regions; Figure S2), although we cannot exclude
590 that the proportion of shared rPEAs among regions is a function of the sample size (see details
591 in the section 2.2 of the Supplementary Information). Moreover, those that were shared among
592 different regions showed consistently similar effects across regions (e.g. positive effects in two
593 or more regions rather than antagonist effects). This supports the predominance of conditional
594 neutrality, i.e. alleles that are advantageous in some environments and neutral in others, over
595 antagonistic pleiotropy, i.e. alleles that are advantageous in some environments and disadvanta-
596 geous in others (Tiffin and Ross-Ibarra 2014). Such pattern has already been reported in plants
597 (Prunier et al. 2012, Anderson et al. 2013). Our results show that, despite a high stability in the
598 level of polygenicity for height between the Atlantic and Mediterranean regions (de Miguel et al.
599 2020), height-growth variation in Mediterranean sites is unlikely to be affected by the same loci
600 as in the other regions, probably as a result of genetic divergence in separated southern refugia
601 during the last glaciation. Overall, identifying positive-effect alleles for different geographical
602 regions separately has the potential to greatly improve the predictive ability of the models, but at
603 the cost of reducing GWAS power (due to lower sample size than in global, wide-range analyses).

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

620 5 Conclusion

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

641 6 Acknowledgements

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

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

658 **7 Author contributions**

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

665 **8 Data and script availability**

666 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>.

670 **References**

- 671 Aitken, S. N. and Bemmels, J. B. (2015). Time to get moving: assisted gene flow of forest trees.
672 *Evolutionary Applications*, 9, 271– 290. <https://doi.org/10.1111/eva.12293>.
- 673 Aitken, S. N. and Whitlock, M. C. (2013). Assisted Gene Flow to Facilitate Local Adaptation to
674 Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, 44, 367– 388.
675 <https://doi.org/10.1146/annurev-ecolsys-110512-135747>.
- 676 Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., and Curtis-McLane, S. (2008). Adaptation,
677 migration or extirpation: climate change outcomes for tree populations. *Evolutionary
678 Applications*, 1, 95– 111. <https://doi.org/10.1111/j.1752-4571.2007.00013.x>.
- 679 Alberto, F., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A.,
680 Lefèvre, F., Lenormand, T., Yeaman, S., Whetten, R., and Savolainen, O. (2013). Potential for
681 evolutionary responses to climate change – evidence from tree populations. *Global Change
682 Biology*, 19, 1645– 1661. <https://doi.org/10.1111/gcb.12181>.

- 683 Alía, R., Moro, J., and Denis, J. B. (1997). Performance of *Pinus pinaster* provenances in Spain:
684 interpretation of the genotype by environment interaction. *Canadian Journal of Forest*
685 *Research*, 27, 1548– 1559.
- 686 Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., and Mitchell-Olds, T. (2013). Genetic
687 trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, 22,
688 699– 708. <https://doi.org/10.1111/j.1365-294X.2012.05522.x>.
- 689 Ballouz, S., Pena, M. T., Knight, F. M., Adams, L. B., and Gillis, J. A. (2019). The transcriptional
690 legacy of developmental stochasticity. *bioRxiv*. <https://doi.org/10.1101/2019.12.11.873265>.
- 691 Barghi, N., Hermisson, J., and Schlötterer, C. (2020). Polygenic adaptation: a unifying framework
692 to understand positive selection. *Nature Reviews Genetics*, 21, 1– 13.
693 <https://doi.org/10.1038/s41576-020-0250-z>.
- 694 Barton, N., Hermisson, J., and Nordborg, M. (2019). Why structure matters. *eLife*, 8, e45380.
695 <https://doi.org/10.7554/eLife.45380>.
- 696 Benito Garzón, M., Robson, T. M., and Hampe, A. (2019). Δ TraitSDMs: species distribution
697 models that account for local adaptation and phenotypic plasticity. *New Phytologist*, 222,
698 1757– 1765. <https://doi.org/10.1111/nph.15716>.
- 699 Berg, J. J., Harpak, A., Sinnott-Armstrong, N., Joergensen, A. M., Mostafavi, H., Field, Y.,
700 Boyle, E. A., Zhang, X., Racimo, F., Pritchard, J. K., and Coop, G. (2019). Reduced signal for
701 polygenic adaptation of height in UK Biobank. *eLife*, 8, e39725.
702 <https://doi.org/10.7554/eLife.39725>.
- 703 Bouvet, J.-M., Makouanzi, G., Cros, D., and Vigneron, P. (2016). Modeling additive and
704 non-additive effects in a hybrid population using genome-wide genotyping: prediction
705 accuracy implications. *Heredity*, 116, 146– 157. <https://doi.org/10.1038/hdy.2015.78>.
- 706 Bradshaw, A. D. (1965). Evolutionary Significance of Phenotypic Plasticity in Plants, Page 115–
707 155. in Caspari, E. W. and Thoday, J. M., editors. *Advances in Genetics*. Academic Press.
- 708 Browne, L., Wright, J. W., Fitz-Gibbon, S., Gugger, P. F., and Sork, V. L. (2019). Adaptational lag
709 to temperature in valley oak (*Quercus lobata*) can be mitigated by genome-informed assisted
710 gene flow. *Proceedings of the National Academy of Sciences*, 116, 25179– 25185.
711 <https://doi.org/10.1073/pnas.1908771116>.
- 712 Bucci, G., González-Martínez, S. C., Provost, G. L., Plomion, C., Ribeiro, M. M., Sebastiani, F.,
713 Alía, R., and Vendramin, G. G. (2007). Range-wide phylogeography and gene zones in *Pinus*
714 *pinaster* Ait. revealed by chloroplast microsatellite markers. *Molecular Ecology*, 16, 2137–
715 2153. <https://doi.org/10.1111/j.1365-294X.2007.03275.x>.
- 716 Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of*
717 *Statistical Software*, 80, 1– 28. <https://doi.org/10.18637/jss.v080.i01>.
- 718 Chambel, M. R., Climent, J., and Alía, R. (2007). Divergence among species and populations of
719 Mediterranean pines in biomass allocation of seedlings grown under two watering regimes.
720 *Annals of Forest Science*, 64, 87– 97. <https://doi.org/10.1051/forest:2006092>.
- 721 Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8, 2–
722 14. <https://doi.org/10.1111/j.1461-0248.2004.00702.x>.
- 723 Corcuera, L., Gil-Pelegrin, E., and Notivol, E. (2010). Phenotypic plasticity in *Pinus pinaster*
724 $\delta^{13}\text{C}$: environment modulates genetic variation. *Annals of Forest Science*, 67, 812– 812.
725 <https://doi.org/10.1051/forest/2010048>.
- 726 Correia, I., Alía, R., Yan, W., David, T., Aguiar, A., and Almeida, M. H. (2010). Genotype \times
727 Environment interactions in *Pinus pinaster* at age 10 in a multienvironment trial in

- 728 Portugal: a maximum likelihood approach. *Annals of Forest Science*, 67, 612– 612.
729 <https://doi.org/10.1051/forest/2010025>.
- 730 Csilléry, K., Ovaskainen, O., Sperisen, C., Buchmann, N., Widmer, A., and Gugerli, F. (2020).
731 Adaptation to local climate in multi-trait space: evidence from silver fir (*Abies alba* Mill.)
732 populations across a heterogeneous environment. *Heredity*, 124, 77– 92.
733 <https://doi.org/10.1038/s41437-019-0240-0>.
- 734 De la Mata, R., Voltas, J., and Zas, R. (2012). Phenotypic plasticity and climatic adaptation in an
735 Atlantic maritime pine breeding population. *Annals of Forest Science*, 69, 477– 487.
736 <https://doi.org/10.1007/s13595-011-0173-0>.
- 737 De Miguel, M., Rodríguez-Quilón, I., Heuertz, M., Hurel, A., Grivet, D., Jaramillo-Correa, J.-P.,
738 Vendramin, G. G., Plomion, C., Majada, J., Alía, R., Eckert, A. J., and González-Martínez, S. C.
739 (2020). Polygenic adaptation and negative selection across traits, years and environments in
740 a long-lived plant species (*Pinus pinaster* Ait.) *bioRxiv*.
741 <https://doi.org/10.1101/2020.03.02.974113>.
- 742 De Villemereuil, P. (2019). On the relevance of Bayesian statistics and MCMC for animal models.
743 *Journal of Animal Breeding and Genetics*, 136, 339– 340. <https://doi.org/10.1111/jbg.12426>.
- 744 Des Marais, D. L., Hernandez, K. M., and Juenger, T. E. (2013). Genotype-by-Environment
745 Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic
746 Environment. *Annual Review of Ecology, Evolution, and Systematics*, 44, 5– 29.
747 <https://doi.org/10.1146/annurev-ecolsys-110512-135806>.
- 748 El-Dien, O. G., Ratcliffe, B., Klápště, J., Porth, I., Chen, C., and El-Kassaby, Y. A. (2018).
749 Multienvironment genomic variance decomposition analysis of open-pollinated Interior
750 spruce (*Picea glauca* x *engelmannii*). *Molecular Breeding*, 38, 26.
751 <https://doi.org/10.1007/s11032-018-0784-3>.
- 752 Eckert, A. J., Heerwaarden, J. van, Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J.,
753 González-Martínez, S. C., and Neale, D. B. (2010). Patterns of Population Structure and
754 Environmental Associations to Aridity Across the Range of Loblolly Pine (*Pinus taeda* L.,
755 Pinaceae). *Genetics*, 185, 969– 982. <https://doi.org/10.1534/genetics.110.115543>.
- 756 Exposito-Alonso, M., Burbano, H. A., Bossdorf, O., Nielsen, R., and Weigel, D. (2018). A map of
757 climate change-driven natural selection in *Arabidopsis thaliana*. *bioRxiv*.
758 <https://doi.org/10.1101/321133>.
- 759 Franklin, J., Serra-Diaz, J. M., Syphard, A. D., and Regan, H. M. (2016). Global change and
760 terrestrial plant community dynamics. *Proceedings of the National Academy of Sciences*, 113,
761 3725– 3734. <https://doi.org/10.1073/pnas.1519911113>.
- 762 Franks, S. J., Weber, J. J., and Aitken, S. N. (2014). Evolutionary and plastic responses to climate
763 change in terrestrial plant populations. *Evolutionary Applications*, 7, 123– 139.
764 <https://doi.org/10.1111/eva.12112>.
- 765 Fréjaville, T. and Benito Garzón, M. (2018). The EuMedClim Database: Yearly Climate Data
766 (1901–2014) of 1 km Resolution Grids for Europe and the Mediterranean Basin. *Frontiers in*
767 *Ecology and Evolution*, 6, 31. <https://doi.org/10.3389/fevo.2018.00031>.
- 768 Fréjaville, T., Vizcaíno-Palomar, N., Fady, B., Kremer, A., and Garzón, M. B. (2020). Range margin
769 populations show high climate adaptation lags in European trees. *Global Change Biology*, 26,
770 484– 495. <https://doi.org/10.1111/gcb.14881>.
- 771 Fuller, Z. L., Mocellin, V. J. L., Morris, L. A., Cantin, N., Shepherd, J., Sarre, L., Peng, J., Liao, Y.,
772 Pickrell, J., Andolfatto, P., Matz, M., Bay, L. K., and Przeworski, M. (2020). Population

- 773 genetics of the coral *Acropora millepora*: Toward genomic prediction of bleaching. *Science*,
774 369. <https://doi.org/10.1126/science.aba4674>.
- 775 Grattapaglia, D., Plomion, C., Kirst, M., and Sederoff, R. R. (2009). Genomics of growth traits in
776 forest trees. *Current Opinion in Plant Biology*, 12, 148– 156.
777 <https://doi.org/10.1016/j.pbi.2008.12.008>.
- 778 Grattapaglia, D. and Resende, M. D. V. (2011). Genomic selection in forest tree breeding. *Tree*
779 *Genetics & Genomes*, 7, 241– 255. <https://doi.org/10.1007/s11295-010-0328-4>.
- 780 Grivet, D., Sebastiani, F., Alía, R., Bataillon, T., Torre, S., Zabal-Aguirre, M., Vendramin, G. G.,
781 and González-Martínez, S. C. (2011). Molecular Footprints of Local Adaptation in Two
782 Mediterranean Conifers. *Molecular Biology and Evolution*, 28, 101– 116.
783 <https://doi.org/10.1093/molbev/msq190>.
- 784 Guan, Y. and Stephens, M. (2011). Bayesian variable selection regression for genome-wide
785 association studies and other large-scale problems. *The Annals of Applied Statistics*, 5, 1780–
786 1815. <https://doi.org/10.1214/11-AOAS455>.
- 787 Gwaze, D. P., Wolliams, J. A., Kanowski, P. J., and Bridgwater, F. E. (2001). Interactions of
788 genotype with site for height and stem straightness in *Pinus taeda* in Zimbabwe. *Silvae*
789 *Genetica*, 50, 135– 139.
- 790 Hereford, J. (2009). A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The*
791 *American Naturalist*, 173, 579– 588. <https://doi.org/10.1086/597611>.
- 792 Hidalgo, A. M., Bastiaansen, J. W. M., Lopes, M. S., Calus, M. P. L., and Koning, D. J. de (2016).
793 Accuracy of genomic prediction of purebreds for cross bred performance in pigs. *Journal of*
794 *Animal Breeding and Genetics*, 133, 443– 451. <https://doi.org/10.1111/jbg.12214>.
- 795 Holliday, J. A., Aitken, S. N., Cooke, J. E. K., Fady, B., González-Martínez, S. C., Heuertz, M.,
796 Jaramillo-Correa, J.-P., Lexer, C., Staton, M., Whetten, R. W., and Plomion, C. (2017).
797 Advances in ecological genomics in forest trees and applications to genetic resources
798 conservation and breeding. *Molecular Ecology*, 26, 706– 717.
799 <https://doi.org/10.1111/mec.13963>.
- 800 Jaramillo-Correa, J.-P., Rodríguez-Quilón, I., Grivet, D., Lepoittevin, C., Sebastiani, F.,
801 Heuertz, M., Garnier-Géré, P. H., Alía, R., Plomion, C., Vendramin, G. G., and
802 González-Martínez, S. C. (2015). Molecular Proxies for Climate Maladaptation in a
803 Long-Lived Tree (*Pinus pinaster* Aiton, Pinaceae). *Genetics*, 199, 793– 807.
804 <https://doi.org/10.1534/genetics.114.173252>.
- 805 Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., Piraux, F.,
806 Guerreiro, L., Pérez, P., Calus, M., Burgueño, J., and de los Campos, G. (2014). A reaction
807 norm model for genomic selection using high-dimensional genomic and environmental data.
808 *Theoretical and Applied Genetics*, 127, 595– 607. <https://doi.org/10.1007/s00122-013-2243-1>.
- 809 Kardos, M. and Luikart, G. (2021). The Genetic Architecture of Fitness Drives Population
810 Viability during Rapid Environmental Change. *The American Naturalist*, 197, 511– 525.
811 <https://doi.org/10.1086/713469>.
- 812 Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7,
813 1225– 1241. <https://doi.org/10.1111/j.1461-0248.2004.00684.x>.
- 814 Langlet, O. (1971). Two Hundred Years Genecology. *Taxon*, 20, 653– 721.
815 <https://doi.org/10.2307/1218596>.

- 816 Latta, R. G. (2009). Testing for local adaptation in *Avena barbata*: a classic example of ecotypic
817 divergence. *Molecular Ecology*, 18, 3781– 3791.
818 <https://doi.org/10.1111/j.1365-294X.2009.04302.x>.
- 819 Leimu, R. and Fischer, M. (2008). A Meta-Analysis of Local Adaptation in Plants. *PLOS ONE*, 3,
820 e4010. <https://doi.org/10.1371/journal.pone.0004010>.
- 821 Leites, L. P., Rehfeldt, G. E., Robinson, A. P., Crookston, N. L., and Jaquish, B. (2012a).
822 Possibilities and Limitations of Using Historic Provenance Tests to Infer Forest Species
823 Growth Responses to Climate Change. *Natural Resource Modeling*, 25, 409– 433.
824 <https://doi.org/10.1111/j.1939-7445.2012.00129.x>.
- 825 Leites, L. P., Rehfeldt, G. E., and Steiner, K. C. (2019). Adaptation to climate in five eastern North
826 America broadleaf deciduous species: Growth clines and evidence of the growth-cold
827 tolerance trade-off. *Perspectives in Plant Ecology, Evolution and Systematics*, 37, 64– 72.
828 <https://doi.org/10.1016/j.ppees.2019.02.002>.
- 829 Leites, L. P., Robinson, A. P., Rehfeldt, G. E., Marshall, J. D., and Crookston, N. L. (2012b).
830 Height-growth response to climatic changes differs among populations of Douglas-fir: a
831 novel analysis of historic data. *Ecological Applications*, 22, 154– 165.
832 <https://doi.org/10.1890/11-0150.1>.
- 833 Leroy, T., Louvet, J.-M., Lalanne, C., Provost, G. L., Labadie, K., Aury, J.-M., Delzon, S.,
834 Plomion, C., and Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to
835 climate in European white oaks. *New Phytologist*, 226, 1171– 1182.
836 <https://doi.org/10.1111/nph.16095>.
- 837 Li, Y., Suontama, M., Burdon, R. D., and Dungey, H. S. (2017). Genotype by environment
838 interactions in forest tree breeding: review of methodology and perspectives on research
839 and application. *Tree Genetics & Genomes*, 13, 60. <https://doi.org/10.1007/s11295-017-1144-x>.
- 840 Mackay, J., Dean, J. F. D., Plomion, C., Peterson, D. G., Cánovas, F. M., Pavy, N., Ingvarsson, P. K.,
841 Savolainen, O., Guevara, M. Á., Fluch, S., Vinceti, B., Abarca, D., Díaz-Sala, C., and
842 Cervera, M.-T. (2012). Towards decoding the conifer giga-genome. *Plant Molecular Biology*,
843 80, 555– 569. <https://doi.org/10.1007/s11103-012-9961-7>.
- 844 MacLachlan, I. R., McDonald, T. K., Lind, B. M., Rieseberg, L. H., Yeaman, S., and Aitken, S. N.
845 (2021). Genome-wide shifts in climate-related variation underpin responses to selective
846 breeding in a widespread conifer. *Proceedings of the National Academy of Sciences*, 118.
847 <https://doi.org/10.1073/pnas.2016900118>.
- 848 Mahony, C. R., MacLachlan, I. R., Lind, B. M., Yoder, J. B., Wang, T., and Aitken, S. N. (2020).
849 Evaluating genomic data for management of local adaptation in a changing climate: A
850 lodgepole pine case study. *Evolutionary Applications*, 13, 116– 131.
851 <https://doi.org/10.1111/eva.12871>.
- 852 Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J.,
853 Bustamante, C. D., and Kenny, E. E. (2017). Human Demographic History Impacts Genetic
854 Risk Prediction across Diverse Populations. *The American Journal of Human Genetics*, 100,
855 635– 649. <https://doi.org/10.1016/j.ajhg.2017.03.004>.
- 856 Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., and Daly, M. J. (2019). Clinical use
857 of current polygenic risk scores may exacerbate health disparities. *Nature Genetics*, 51, 584–
858 591. <https://doi.org/10.1038/s41588-019-0379-x>.

- 859 Matesanz, S., Gianoli, E., and Valladares, F. (2010). Global change and the evolution of
860 phenotypic plasticity in plants. *Annals of the New York Academy of Sciences*, 1206, 35– 55.
861 <https://doi.org/10.1111/j.1749-6632.2010.05704.x>.
- 862 Mathieson, I. (2021). The omnigenic model and polygenic prediction of complex traits. *The*
863 *American Journal of Human Genetics*. <https://doi.org/10.1016/j.ajhg.2021.07.003>.
- 864 McLane, S. C., Daniels, L. D., and Aitken, S. N. (2011). Climate impacts on lodgepole pine (*Pinus*
865 *contorta*) radial growth in a provenance experiment. *Forest Ecology and Management*, 262,
866 115– 123. <https://doi.org/10.1016/j.foreco.2011.03.007>.
- 867 Merilä, J. and Hendry, A. P. (2014). Climate change, adaptation, and phenotypic plasticity: the
868 problem and the evidence. *Evolutionary Applications*, 7, 1– 14.
869 <https://doi.org/10.1111/eva.12137>.
- 870 Moghaddar, N., Swan, A. A., and van der Werf, J. H. (2014). Comparing genomic prediction
871 accuracy from purebred, crossbred and combined purebred and crossbred reference
872 populations in sheep. *Genetics Selection Evolution*, 46, 58.
873 <https://doi.org/10.1186/s12711-014-0058-4>.
- 874 Muff, S., Niskanen, A. K., Saatoglu, D., Keller, L. F., and Jensen, H. (2019). Animal models with
875 group-specific additive genetic variances: extending genetic group models. *Genetics Selection*
876 *Evolution*, 51, 7. <https://doi.org/10.1186/s12711-019-0449-7>.
- 877 Nadeau, S., Meirmans, P. G., Aitken, S. N., Ritland, K., and Isabel, N. (2016). The challenge of
878 separating signatures of local adaptation from those of isolation by distance and
879 colonization history: The case of two white pines. *Ecology and Evolution*, 6, 8649– 8664.
880 <https://doi.org/10.1002/ece3.2550>.
- 881 Neale, D. B. and Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends*
882 *in Plant Science*, 9, 325– 330. <https://doi.org/10.1016/j.tplants.2004.05.006>.
- 883 Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P.,
884 Purugganan, M. D., Richards, C. L., Valladares, F., and van Kleunen, M. (2010). Plant
885 phenotypic plasticity in a changing climate. *Trends in Plant Science*, 15, 684– 692.
886 <https://doi.org/10.1016/j.tplants.2010.09.008>.
- 887 O’Neill, G. A., Hamann, A., and Wang, T. (2008). Accounting for population variation improves
888 estimates of the impact of climate change on species’ growth and distribution. *Journal of*
889 *Applied Ecology*, 1040– 1049.
890 [https://doi.org/10.1111/j.1365-2664.2008.01472.x@10.1111/\(ISSN\)1365-2664.CLIMATE:JPE](https://doi.org/10.1111/j.1365-2664.2008.01472.x@10.1111/(ISSN)1365-2664.CLIMATE:JPE).
- 891 Olson, M. S., Robertson, A. L., Takebayashi, N., Silim, S., Schroeder, W. R., and Tiffin, P. (2010).
892 Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*). *New*
893 *Phytologist*, 186, 526– 536. <https://doi.org/10.1111/j.1469-8137.2009.03174.x>.
- 894 Pedlar, J. H. and McKenney, D. W. (2017). Assessing the anticipated growth response of northern
895 conifer populations to a warming climate. *Scientific Reports*, 7, 43881.
896 <https://doi.org/10.1038/srep43881>.
- 897 Pigliucci, M., Murren, C. J., and Schlichting, C. D. (2006). Phenotypic plasticity and evolution by
898 genetic assimilation. *Journal of Experimental Biology*, 209, 2362– 2367.
899 <https://doi.org/10.1242/jeb.02070>.
- 900 Plomion, C., Bartholomé, J., Lesur, I., Boury, C., Rodríguez-Quilón, I., Lagrœulet, H.,
901 Ehrenmann, F., Bouffier, L., Gion, J. M., Grivet, D., Miguel, M. de, María, N. de,
902 Cervera, M. T., Bagnoli, F., Isik, F., Vendramin, G. G., and González-Martínez, S. C. (2016).

- 903 High-density SNP assay development for genetic analysis in maritime pine (*Pinus pinaster*).
904 *Molecular Ecology Resources*, 16, 574– 587. <https://doi.org/10.1111/1755-0998.12464>.
- 905 Postma, F. M. and Ågren, J. (2016). Early life stages contribute strongly to local adaptation in
906 *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences*, 113, 7590– 7595.
907 <https://doi.org/10.1073/pnas.1606303113>.
- 908 Pritchard, J. K., Pickrell, J. K., and Coop, G. (2010). The Genetics of Human Adaptation: Hard
909 Sweeps, Soft Sweeps, and Polygenic Adaptation. *Current Biology*, 20, R208– R215.
910 <https://doi.org/10.1016/j.cub.2009.11.055>.
- 911 Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of Population Structure Using
912 Multilocus Genotype Data. *Genetics*, 155, 945– 959.
- 913 Prunier, J., Gérardi, S., Laroche, J., Beaulieu, J., and Bousquet, J. (2012). Parallel and
914 lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology*,
915 21, 4270– 4286. <https://doi.org/10.1111/j.1365-294X.2012.05691.x>.
- 916 R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria:
917 R Foundation for Statistical Computing.
- 918 Razgour, O., Forester, B., Taggart, J. B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille, S. J.,
919 Novella-Fernandez, R., Alberdi, A., and Manel, S. (2019). Considering adaptive genetic
920 variation in climate change vulnerability assessment reduces species range loss projections.
921 *Proceedings of the National Academy of Sciences*, 116, 10418– 10423.
922 <https://doi.org/10.1073/pnas.1820663116>.
- 923 Rehfeldt, G. E., Leites, L. P., Joyce, D. G., and Weiskittel, A. R. (2018). Role of population genetics
924 in guiding ecological responses to climate. *Global Change Biology*, 24, 858– 868.
925 <https://doi.org/10.1111/gcb.13883>.
- 926 Rehfeldt, G. E., Tchebakova, N. M., Milyutin, L. I., Parfenova, E. I., Wykoff, W. R., and
927 Kouzmina, N. A. (2003). Assessing population responses to climate in *Pinus sylvestris* and
928 *Larix* spp. of Eurasia with climate-transfer models. *Eurasian Journal of Forest Research -*
929 *Hokkaido University (Japan)*.
- 930 Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., and Hamilton, D. A. (1999). Genetic Responses to
931 Climate in *Pinus contorta*: Niche Breadth, Climate Change, and Reforestation. *Ecological*
932 *Monographs*, 69, 375– 407.
933 [https://doi.org/10.1890/0012-9615\(1999\)069\[0375:GRTCIP\]2.0.CO;2](https://doi.org/10.1890/0012-9615(1999)069[0375:GRTCIP]2.0.CO;2).
- 934 Resende Jr, M. F. R., Muñoz, P., Acosta, J. J., Peter, G. F., Davis, J. M., Grattapaglia, D.,
935 Resende, M. D. V., and Kirst, M. (2012). Accelerating the domestication of trees using
936 genomic selection: accuracy of prediction models across ages and environments. *New*
937 *Phytologist*, 193, 617– 624. <https://doi.org/10.1111/j.1469-8137.2011.03895.x>.
- 938 Resende, M. D. V., Resende, M. F. R., Sansaloni, C. P., Petrolí, C. D., Missiaggia, A. A.,
939 Aguiar, A. M., Abad, J. M., Takahashi, E. K., Rosado, A. M., Faria, D. A., Pappas, G. J.,
940 Kilian, A., and Grattapaglia, D. (2012). Genomic selection for growth and wood quality in
941 *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits
942 in forest trees. *New Phytologist*, 116– 128. [https://doi.org/10.1111/j.1469-
943 8137.2011.04038.x@10.1002/\(ISSN\)1469-8137\(CAT\)FeatureIssues\(VI\)Bioenergytrees](https://doi.org/10.1111/j.1469-8137.2011.04038.x@10.1002/(ISSN)1469-8137(CAT)FeatureIssues(VI)Bioenergytrees).
- 944 Rodríguez-Quilón, I., Santos-del-Blanco, L., Serra-Varela, M. J., Koskela, J.,
945 González-Martínez, S. C., and Alía, R. (2016). Capturing neutral and adaptive genetic
946 diversity for conservation in a highly structured tree species. *Ecological Applications*, 26,
947 2254– 2266. <https://doi.org/10.1002/eap.1361>.

- 948 Santos-del-Blanco, L., Climent, J., González-Martínez, S. C., and Pannell, J. R. (2012). Genetic
949 differentiation for size at first reproduction through male versus female functions in the
950 widespread Mediterranean tree *Pinus pinaster*. *Annals of Botany*, 110, 1449– 1460.
951 <https://doi.org/10.1093/aob/mcs210>.
- 952 Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation.
953 *Nature Reviews Genetics*, 14, 807– 820. <https://doi.org/10.1038/nrg3522>.
- 954 Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Gene Flow and Local Adaptation in Trees.
955 *Annual Review of Ecology, Evolution, and Systematics*, 38, 595– 619.
956 <https://doi.org/10.1146/annurev.ecolsys.38.091206.095646>.
- 957 Seidl, R., Thom, D., Kautz, M., Martin-Benito, D., Peltoniemi, M., Vacchiano, G., Wild, J.,
958 Ascoli, D., Petr, M., Honkaniemi, J., Lexer, M. J., Trotsiuk, V., Mairota, P., Svoboda, M.,
959 Fabrika, M., Nagel, T. A., and Reyer, C. P. O. (2017). Forest disturbances under climate
960 change. *Nature Climate Change*, 7, 395– 402. <https://doi.org/10.1038/nclimate3303>.
- 961 Sella, G. and Barton, N. H. (2019). Thinking about the evolution of complex traits in the era of
962 genome-wide association studies. *Annual Review of Genomics and Human Genetics*, 20.
- 963 Shaw, R. G. and Etterson, J. R. (2012). Rapid climate change and the rate of adaptation: insight
964 from experimental quantitative genetics. *New Phytologist*, 195, 752– 765.
965 <https://doi.org/10.1111/j.1469-8137.2012.04230.x>.
- 966 Sohail, M., Maier, R. M., Ganna, A., Bloemendal, A., Martin, A. R., Turchin, M. C., Chiang, C. W.,
967 Hirschhorn, J., Daly, M. J., Patterson, N., Neale, B., Mathieson, I., Reich, D., and Sunyaev, S. R.
968 (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in
969 genome-wide association studies. *eLife*, 8, e39702. <https://doi.org/10.7554/eLife.39702>.
- 970 Sork, V. L. (2018). Genomic Studies of Local Adaptation in Natural Plant Populations. *Journal of*
971 *Heredity*, 109, 3– 15. <https://doi.org/10.1093/jhered/esx091>.
- 972 Stern, A. J., Speidel, L., Zaitlen, N. A., and Nielsen, R. (2020). Disentangling selection on
973 genetically correlated polygenic traits using whole-genome genealogies. *bioRxiv*.
974 <https://doi.org/10.1101/2020.05.07.083402>.
- 975 Tiffin, P. and Ross-Ibarra, J. (2014). Advances and limits of using population genetics to
976 understand local adaptation. *Trends in Ecology & Evolution*, 29, 673– 680.
977 <https://doi.org/10.1016/j.tree.2014.10.004>.
- 978 Valladares, F., Gianoli, E., and Gómez, J. M. (2014a). Ecological limits to plant phenotypic
979 plasticity. *New Phytologist*, 176, 749– 763. <https://doi.org/10.1111/j.1469-8137.2007.02275.x>.
- 980 Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M. B., Balaguer, L., Benito-Garzón, M.,
981 Cornwell, W., Gianoli, E., Kleunen, M. van, Naya, D. E., Nicotra, A. B., Poorter, H., and
982 Zavala, M. A. (2014b). The effects of phenotypic plasticity and local adaptation on forecasts
983 of species range shifts under climate change. *Ecology Letters*, 17, 1351– 1364.
984 <https://doi.org/10.1111/ele.12348>.
- 985 Viñas, R. A., Caudullo, G., Oliveira, S., and de Rigo, D. (2016). *Pinus pinaster* in Europe:
986 distribution, habitat, usage and threats.
- 987 Vizcaíno-Palomar, N., Fady, B., Alía, R., Raffin, A., Mutke, S., and Benito Garzón, M. (2020). The
988 legacy of climate variability over the last century on populations' phenotypic variation in
989 tree height. *Science of The Total Environment*, 749, 141454.
990 <https://doi.org/10.1016/j.scitotenv.2020.141454>.

- 991 Vogt, G. (2015). Stochastic developmental variation, an epigenetic source of phenotypic
992 diversity with far-reaching biological consequences. *Journal of Biosciences*, 40, 159– 204.
993 <https://doi.org/10.1007/s12038-015-9506-8>.
- 994 Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R.,
995 Mock, T., Schmid, K., Schmitt, I., Bataillon, T., Savolainen, O., Bergland, A., Flatt, T.,
996 Guillaume, F., and Pfenninger, M. (2020). Evolutionary genomics can improve prediction of
997 species' responses to climate change. *Evolution Letters*, 4, 4– 18.
998 <https://doi.org/10.1002/evl3.154>.
- 999 Wang, T., O'Neill, G. A., and Aitken, S. N. (2010). Integrating environmental and genetic effects
1000 to predict responses of tree populations to climate. *Ecological Applications*, 20, 153– 163.
1001 <https://doi.org/10.1890/08-2257.1>.
- 1002 Wu, H. X. and Ying, C. C. (2004). Geographic pattern of local optimality in natural populations
1003 of lodgepole pine. *Forest Ecology and Management*, 194, 177– 198.
1004 <https://doi.org/10.1016/j.foreco.2004.02.017>.
- 1005 Zas, R., Merlo, E., Díaz, R., and Fernández López, J. (2003). Stability across sites of Douglas-fir
1006 provenances in northern Spain. 10, 71– 82.
- 1007 Zonneveld, B. J. M. (2012). Conifer genome sizes of 172 species, covering 64 of 67 genera, range
1008 from 8 to 72 picogram. *Nordic Journal of Botany*, 30, 490– 502.
1009 <https://doi.org/10.1111/j.1756-1051.2012.01516.x>.