



RESEARCH PAPER

Open Access



Unlocking genome-based prediction and selection in conifers: the key role of within-family prediction accuracy illustrated in maritime pine (*Pinus pinaster* Ait.)

Victor Papin¹ , Gregor Gorjanc² , Ivan Pocrnic² , Laurent Bouffier^{1*} and Leopoldo Sanchez^{3†}

Abstract

Key message Based on experimental and simulated data for maritime pine (*Pinus pinaster* Ait.) in a genomic selection context, our study reveals that the often-highlighted equivalence between genome-based and pedigree-based prediction accuracies of breeding values is associated with a zero accuracy of genome-based prediction within families, which can be attributed to the still insufficient size of the genomic training sets for conifers.

Context Genomic selection is a promising approach for forest tree breeding. However, its advantage in terms of prediction accuracy over conventional pedigree-based methods is unclear and within-family accuracy is rarely assessed.

Aims We used a pedigree-based model (ABLUP) with corrected pedigree data as a baseline reference for assessing the prediction accuracy of genome-based model (GBLUP) at the global and within-family levels in maritime pine (*Pinus pinaster* Ait.).

Methods We considered 39 full-sib families, each comprising 10 to 40 individuals, to constitute an experimental population of 833 individuals. A stochastic simulation model was also developed to explore other scenarios of heritability, training set size, and marker density.

Results Prediction accuracies with GBLUP and ABLUP were similar, and within-family accuracy with GBLUP was on average zero with large variation between families. Simulations revealed that the number of individuals in the training set was the principal factor limiting GBLUP accuracy in our study and likely in many forest tree breeding programmes. Accurate within-family prediction is possible if 40–65 individuals per full-sib family are included in the genomic training set, from a total of 1600–2000 individuals in the training set.

Conclusions The increase in the number of individuals per family in the training set lead to a significant advantage of GBLUP over ABLUP in terms of prediction accuracy and more clearly justify the switch to genome-based prediction and selection in forest trees.

Keywords Breeding programme, Genomic selection, Maritime pine, Progeny validation, Stochastic simulation, Within-family variability

Handling editor: Ricardo Alia.

[†]Laurent Bouffier and Leopoldo Sanchez contributed equally to this work.

*Correspondence:

Laurent Bouffier

laurent.bouffier@inrae.fr

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

1 Introduction

The use of pedigree information in populations with genealogy records has improved breeding programmes for many species. Pedigree information can be used to infer the expected relatedness between each pair of individuals, making it possible to gauge the extent to which phenotypic values of individuals in the studied population have a genetic basis. This pedigree-based model underpinned the development of predictions of the individual additive genetic value, or the breeding value, via BLUP (Best Linear Unbiased Prediction) methodology (Henderson 1975). The breeding value of each individual can be decomposed into parent average and Mendelian sampling terms. The parent average term captures variation between families and it represents the expected breeding value of progeny given its parents. The Mendelian sampling term captures variation within families and it represents deviation of each individual's breeding value from the parent average due to recombination and segregation of parental genomes. With a pedigree-based model, we need phenotypic values on an individual or its progeny to estimate the Mendelian sampling term of the individual's breeding value. Hence, pedigree-based prediction of breeding values for non-phenotyped individuals (forward prediction) captures only the parent average term. By using genomic data, we observe outcome of recombination and segregation of parental genomes as well as recent or past mutations, meaning that we can in principle estimate parent average and Mendelian sampling terms of breeding value even for non-phenotyped individuals (VanRaden 2008; Hill and Weir 2011).

The advent of affordable genome-wide DNA marker genotyping platforms has enabled such genome-based predictions to pave the way to genomic selection (GS). Albeit initially proposed by Bernardo (1994) and Nejati-Javaremi et al. (1997), genome-based prediction took off with the work of Meuwissen et al. (2001), which showed how regressing individuals' phenotypic values onto their genome-wide marker genotypes captured variation between individuals' breeding values by leveraging linkage-disequilibrium between quantitative trait loci (QTL) affecting traits of interest and the genome-wide markers. Using a training set of individuals that have been phenotyped and genotyped, the model estimates associations between variation in genome-wide markers and variation in phenotypic values. This means that the associations can be used to predict breeding values for non-phenotyped individuals which have genomic information. Such genome-based predictions have revolutionized many breeding programmes, enabling an efficient and early selection of candidates and leading to significant genetic and economic gain per unit of time (Crossa et al. 2017; Hayes et al. 2009a, 2009b; Pryce et al. 2011).

Genome-based prediction is of particular interest in forest trees, as it could decrease the length of breeding cycles, which are long for such species, and cut the cost of phenotyping complex traits, such as drought tolerance and disease resistance (Grattapaglia & Resende 2011, Isik 2014). Driven by the promising results obtained from simulations and first empirical approaches (Grattapaglia et al. 2011; Grattapaglia & Resende 2011; Iwata et al. 2011), increasing numbers of experimental GS studies have been performed in recent years on many forest tree species (see Lebedev et al. 2020 and Beaulieu et al. 2024 for recent reviews). Many studies have highlighted the attractiveness of genome-based prediction by reporting moderate to high prediction accuracies (Durán et al. 2017; Isik et al. 2016; Resende et al. 2012a, 2012b) and improved genetic gain per unit of time due to 20–50% shorter generation interval for GS (Chen et al. 2018; Lenz et al. 2017; Ratcliffe et al. 2015; Resende et al. 2012a, 2012b). However, such a reduction in generation interval is also possible with pedigree-based prediction (if considering forward selection). Thus, GS overcomes pedigree-based selection only if genome-based predictions are more accurate than mid-parental breeding values (i.e. pedigree-based prediction estimated at the seedling age). In fact, several studies in forest tree breeding report that pedigree-based predictions have similar accuracy to genome-based predictions (Beaulieu et al. 2014; Lenz et al. 2020a, 2020b; Thistlethwaite et al. 2017, 2019; Zapata-Valenzuela et al. 2012, 2013; Zhou et al. 2020). Furthermore, the use of an incomplete or error-containing pedigree tends to distort the comparison with genomic data (El-Dien et al. 2018; Li et al. 2019). Such errors may be common in forest tree breeding programmes and this penalizes pedigree-based evaluation (Doerksen and Herbringer 2010; Munoz et al. 2014). The advantage of GS may therefore stem at least partly from the errors inherent to pedigree-based selection (Lenz et al. 2020a, 2020b). Clarifying the conditions under which genome-based models can deliver real benefits remains a prerequisite for full exploitation of the advantages of GS in forest trees.

The access to within-family variability provided by molecular markers should increase the benefits of GS relative to pedigree-based selection, making it possible to improve the management of diversity. Indeed, accurate within-family prediction would facilitate the exploitation of within-family genetic variability rather than inter-familial variability, preventing the over-representation of certain lineages during selection and subsequent drift (Allier et al. 2019; Jannink 2010; Rauf et al. 2010). However, little attention has been paid to the accuracy of genome-based prediction within families in forest trees, mostly due to the limited number of individuals

per half- and full-sib family generally used in progeny trials. This issue has been addressed in only a few studies. Fuentes et al. (2017) and Cros et al. (2019) each studied a single large full-sibling family, making it difficult to extrapolate their results to more general cases. In three other studies (Pégard et al. 2020; Resende et al. 2017; Ukrainetz & Mansfield, 2019), the accuracies of within-family predictions were substantial, but variable.

Maritime pine (*Pinus pinaster* Ait.) covers 4.2 million hectares in south-western Europe (Abad Viñas et al., 2016). A breeding programme based on a recurrent selection scheme was initiated for this species in France in the 1960s (Durel 1992, GIS 2002). Starting from a base population (600 G0 individuals) selected for growth, environmental adaptation and stem straightness, two breeding cycles were performed using the estimated breeding values from a pedigree-based model (Bouffier et al. 2016). The potential of genome-based prediction in maritime pine breeding has already been highlighted in two previous studies (Isik et al. 2016; Bartholomé et al. 2016), but, as for most forest tree species, it is essential to investigate in greater depth the conditions in which genome-based prediction is clearly superior to pedigree-based one.

The aim of this study was to evaluate the ability of genome-based prediction to capture the Mendelian sampling term in a maritime pine breeding population with empirical and simulation approaches. In both cases, the accuracy of genome-based prediction was estimated at the population as well as at the within-family levels and was compared with that of pedigree-based prediction. The real data were obtained for a population of 39 full-sib families with family sizes ranging from 10 to 40 individuals per family. The simulation, designed to mimic the conditions of the maritime pine breeding programme, added other scenarios not observed in the real population, including variations of heritability, training set size and marker density.

2 Materials and methods

2.1 Exploring accuracy of predictions with real data

2.1.1 Maritime pine trial

A maritime pine trial was established in 2011 in the Landes de Gascogne forest at Le Barp (Lat 44.62, Long -0.77). A complete block design was used, with 89 full-sib families and 10 checklots, each containing 48 individuals planted in six-tree plots (1250 trees/ha). The full-sib families considered were taken from the third generation of the French maritime pine breeding programme (i.e., the pedigree of the trees was known to grandparent level).

Preliminary simulations were performed to determine the most appropriate proportion of families and offspring

per family from the total trial population to maximize GS accuracy at both global and within-family levels (Appendix 1). Based on the simulation, an optimal sample of 40 families was obtained, 30 with 20 individuals and 10 with 40 individuals. The selected families were representative of the genetic diversity present in the trial and the within-family samples were representative of the phenotypic variability of each family. The larger families corresponded to five of the best-related and five of the worst-related families (referred to hereafter as “well-related” and “poorly related” families). The average relatedness to the rest of the population was 0.03 for the well-related families and 0.01 for the poorly related families (calculated from pedigree data). Considering families with large numbers of offspring is key to investigating within-family predictive ability. After genotyping, our study set, called POP_R, contained 833 individuals with an effective population size of 25 (Lindgren et al. 1996). Thirty-nine families could be used to assess within-family predictive ability, including nine families of more than 30 individuals each.

2.1.2 Genomic and pedigree information for POP_R

Genomic DNA was extracted from young needles collected from each individual of POP_R. DNA quantification and quality control were performed by fluorimetry (Qubit 2.0, Life Technologies, Thermo Fisher Scientific, USA) and spectrophotometry (NanoDrop Technologies, Wilmington, DE, USA), respectively. Genotyping was performed by Thermo Fisher Scientific (Thermo Fisher Scientific, Santa Clara, CA, USA) with the 4TREE Axiom 50 K single-nucleotide polymorphisms (SNP) multi-species array (Guilbaud et al. 2020). Samples with a call rate below 97% were excluded from further analysis. In addition to the quality control filters at the SNP level (CallRate ≥ 85%, f_{ld}-cutoff ≥ 3.2, het-so-cutoff ≥ -0.3) suggested by Thermo Fisher Scientific, we also excluded SNPs with more than 5% Mendelian segregation errors, SNPs with a repeatability below 98% (estimated with 42 duplicated samples), and SNPs with a minor allele frequency (MAF) below 1%. Finally, 833 individuals characterized for 8235 SNPs were available for this study. Missing genotypes were imputed by assigning the average genotype within the full-sib family. It was not possible to apply more sophisticated imputation methods due to the lack of a genetic map. We computed a realized genomic relationship matrix (**G**) following VanRaden (2008), with the R package AGHmatrix (Amadeu et al. 2016):

$$G = \frac{(M-P)(M-P)'}{2\Sigma p_i(1-p_i)} \quad (1)$$

where **M** and **P** are matrices of dimensions *n* (number of individuals) × *p* (number of markers). **M** gives the genotype at each locus, coded -1 for one of the homozygotes,

0 for heterozygotes and 1 for the other homozygote, and the $\mathbf{P}$ is a matrix of allele frequencies expressed as $2(p_i - 0.5)$, where p_i is the observed allele frequency at marker i for all genotyped individuals.

Pedigree information was also available for POP_R at the parental (41 seed parents and 40 pollen parents) and grandparental (103 initial progenitors from the base population of the breeding programme) levels. Pedigree errors were detected with the R package *pedtools* (Vigeland, 2021), by comparing the genotyping data of POP_R with the genotyping data available for 78 of 81 parents. Pedigree errors were detected for 35 individuals (4% of individuals), i.e. when more than 1% of mismatches were detected, for all SNPs, between these individuals and their theoretical parents. Non-genotyped parents were assumed to be correct. Of these 35 individuals, 4 individuals were reassigned to existing families, 10 individuals initially assigned to one family were identified as new full-sib family, as well as 6 individuals initially assigned to another family were identified as a second new full-sib family, and 15 individuals were considered of unknown parentage. The total number of families has therefore increased to 42, but within-family predictive ability was only assessed for 39 full-sib families since 3 families did not contain enough individuals (<10 individuals) for a reliable assessment. More precisely, within-family predictive ability was assessed for 9 large families with a mean of 34 individuals per family (30 to 40) and 30 families with a mean of 17 individuals per family (10 to 20). A complete corrected version of the pedigree was used to calculate an additive relationship matrix $\mathbf{A}$, for a total of 1014 individuals.

2.1.3 Phenotypic data

All individuals in the original trial, including those of POP_R, were phenotyped at the age of 8 years for height (HT) and deviation of the stem from verticality (DEV). Direct phenotypic values were then used in the subsequent genomic selection models after a step of pre-adjustment (7.1.3.). Narrow-sense heritability for HT and DEV within POP_R were 0.13 and 0.21, respectively, for estimates based on genomic data, and 0.17 and 0.25, respectively, for estimates based on pedigree information.

2.1.4 Genome-based and pedigree-based models

Breeding values were estimated for each trait and for the n individuals using the model:

$$y = 1\mu + Zu + e \quad (2)$$

where y is the vector of adjusted phenotypes (dimension $nx1$), $\mathbf{1}$ is the vector of 1 s, μ is the population mean associated with a vector $\mathbf{1}$ of dimension ($nx1$), $\mathbf{Z}$ is the

incidence matrix (dimension nxn) connecting the phenotypes to the vector of breeding values u (dimension $nx1$) and e is the vector of residuals (dimension $nx1$). u and e are assumed to be independent from each other and to follow a normal distribution of the form $u \sim N(0, X\sigma_u^2)$ and $e \sim N(0, I_n\sigma_e^2)$, where X is either the genome-based (realized) additive relationship matrix $\mathbf{G}$ or pedigree-based (expected) additive relationship matrix $\mathbf{A}$, σ_u^2 is the associated variance of breeding values, I_n is the n -dimensional identity matrix and σ_e^2 is the variance of residual effects. Mixed-model equations were solved to predict the random genetic effects u :

$$\begin{bmatrix} 1'1 & 1'Z \\ Z'1 & Z'Z + X^{-1}\alpha \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} 1'y \\ Z'y \end{bmatrix} \quad (3)$$

where X^{-1} is the inverse of X and $\alpha = \sigma_e^2/\sigma_u^2$ (Henderson 1975; Mrode and Pocrnic 2023). These two model versions (GBLUP and ABLUP) respectively produced genome-based (GEBV) and pedigree-based (EBV) estimates / predictions of breeding values. All model fitting was performed with the R.4.2.2 environment (R Core Team 2022) using the R package *breedR* (Muñoz and Sanchez 2020).

2.1.5 Cross-validation scenarios and assessment of prediction accuracy

Two cross-validation scenarios (CV) were used to assess the prediction accuracies of the GBLUP and ABLUP models (Fig. 1): CV1 was used to assess within-family predictive ability for the 39 families of POP_R, whereas CV2 focused on the nine large families (i.e. the families from which 40 individuals were sampled). In the CV1 scenario, the training set (T_{set}) included 40% of each family and the validation set (V_{set}) included the remaining 60% of each family. The CV2 scenario was divided into six subscenarios. For the first subscenario, the T_{set} consisted of POP_R minus the large families and the V_{set} included all the individuals from large families. The number of individuals from large families included in the T_{set} was progressively increased in the other five subscenarios

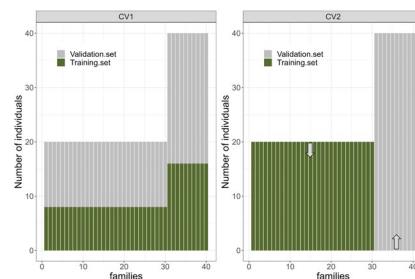


Fig. 1 Cross-validation scenarios CV1 and CV2 performed with ABLUP and GBLUP models

(5, 10, 15, 20 and then 25 individuals per large family), with a corresponding reduction of the contribution of other families to T_{set} so as to maintain a constant T_{set} size (the individuals from non-large families not included in the T_{set} became part of the V_{set} , thereby also keeping the size of this set constant).

Each cross-validation subscenario was replicated 100 times. For each replicate, within-family predictive abilities were defined as Pearson's correlation coefficients between adjusted phenotypes (y) and (G)EBV ($\hat{y}$) for individuals in the V_{set} , considering each family separately. Note that within-family prediction is only meaningful for the GBLUP model. Namely, the ABLUP model will only predict parent average component of the breeding value for non-phenotyped full-sibs, meaning that the resulting EBV will be the same for all full-sibs giving a within-family predictive ability of 0.

At the global level (i.e. considering all individuals in the V_{set} simultaneously), we rather estimated prediction accuracy to facilitate comparison with other studies, which frequently report this metric (Legarra et al. 2008):

$$\text{accuracy} = \frac{r(y, \hat{u})}{\sqrt{\hat{h}_y^2}} \quad (4)$$

where $\hat{h}_y^2$ is the estimated heritability of the trait. We did not use prediction accuracy as the metric for within-family analyses because dividing by the heritability defined at population level gives rise to values much higher than 1 or lower than -1. Instead, we focused on the deviation from 0.

2.2 Identifying key parameters for prediction accuracy with simulations

2.2.1 Description of the simulation model

Stochastic simulations based on an allelic model were performed with the R package AlphaSimR (Gaynor et al. 2021). Briefly, from a base population representative of the maritime pine genome (Chagné et al. 2002; Chancerel et al. 2013; Jaramillo-Correa et al., 2020; Milesi et al. 2023), we simulated successive breeding populations based on a single trait (equivalent to HT), taking into account the characteristics of the real French maritime pine breeding programme. Details of the simulations are provided in 8.1., and the code is available at https://github.com/HighlanderLab/vpapin_pine_gs. Simulated phenotypes and genotypes for the final population, POP_S (the simulated version of POP_R), were used to fit GBLUP and ABLUP models as described for the real data. The analyses were performed for the CV1 scenario (40% of individuals in each family included in the T_{set}). Heritability of simulated phenotypes was set to 0.13, and the number of markers used was 8235 to mimic the real-life data as closely as possible. Ten independent replications of the

entire process described above were performed to ensure that the results were robust.

2.2.2 ABLUP and GBLUP prediction accuracy under different scenarios

Stochastic simulations were used to extend the comparison of prediction accuracy between GBLUP and ABLUP to different scenarios in which trait heritability (h^2), training set size (nT_{set}) and the number of markers ($n\text{SNP}$) were varied. For this purpose, POP_S was extended by generating 100 individuals for each of the 42 initially sampled families. Accuracy was assessed with a unique cross-validation scenario similar to CV1 (all families contributing equally to T_{set}) but with a fixed-size V_{set} of 1200 individuals (evenly distributed between families). The values taken by the three parameters mimic real possibilities for maritime pine breeding:

- The size of T_{set} was set to $nT_{\text{set}} = 400, 600, 1,600$ or $2,600$ individuals, corresponding to 10, 15, 40 and 65 individuals per family, respectively. Such numbers are usually available in most forest tree breeding programmes. These numbers are also economically viable, because they require only 10, 15, 40 and 65% of the population to be phenotyped.
- The number of markers was set to $n\text{SNP} = 8235$ or $17,220$ or $35,000$ SNPs, corresponding to marker densities of 5.7, 12 and 24 markers/cM, respectively. Data for 8235 SNPs are already available in our real maritime pine dataset, but this number could be increased by developing chips with a higher density of markers.
- The heritability of the trait was set to $h^2 = 0.13, 0.33$ or 0.50 . Mean heritability for HT in the maritime pine breeding programme is 0.33, but HT ranges between 0.13 and 0.50, depending on the trial considered.

3 Results

3.1 Global and within-family predictive ability for maritime pine data

3.1.1 Global prediction accuracies of GBLUP and ABLUP

Global prediction accuracies were estimated with the CV1 scenario (Fig. 2). Mean prediction accuracy was similar for the two traits at $0.52 (\pm 0.08)$ and $0.55 (\pm 0.07)$ for HT and DEV, respectively. Mean prediction accuracy was slightly higher for the GBLUP model than for the ABLUP model, at $+0.05 (\pm 0.11)$ for HT and $+0.02 (\pm 0.10)$ for DEV, but the differences were significant only for HT. For this CV1 scenario, we also varied the percentage of individuals from each family included in the T_{set} . The proportion included ranged from 20 to 80%, and accuracy increased with this percentage, from $0.45 (\pm 0.09)$ to 0.62

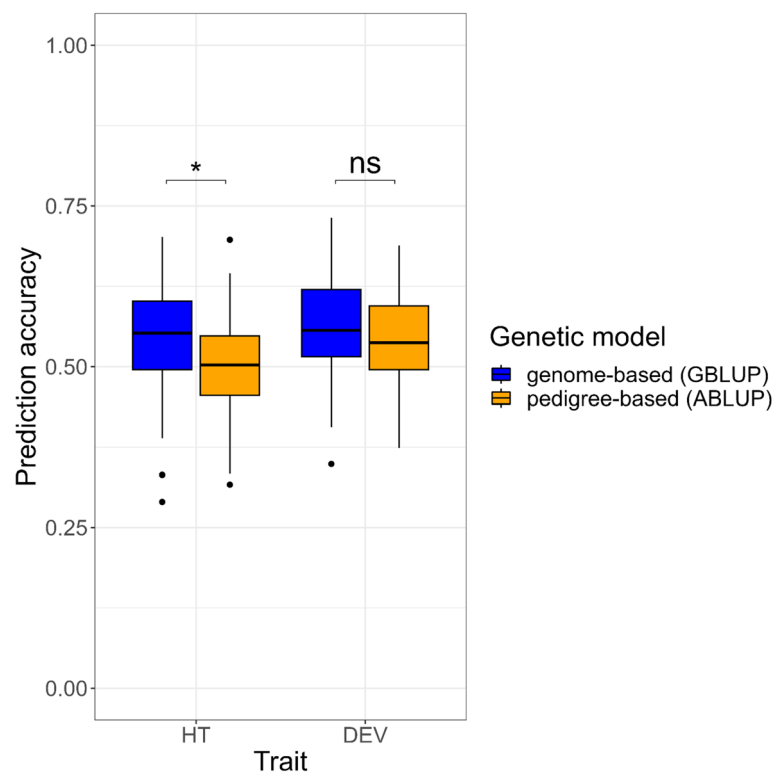


Fig. 2 Global prediction accuracies obtained in the scenario CV1 with ABLUP and GBLUP models, for height (HT) and stem deviation to verticality (DEV)

(± 0.18), but in a similar manner for ABLUP and GBLUP. We present only the data for a proportion of 40% (the mode) in Fig. 2.

In the CV2 scenario (Fig. 3), adding more individuals from the large families and removing individuals from the other families increased mean accuracy, with a maximum value of $0.68 (\pm 0.10)$ obtained for both traits when 10 to 15 individuals from large families were included in the T_{set} . Global accuracy declined if the number of individuals from large families was increased any further. In other words, accuracy was highest when all families were equally represented in the T_{set} . In most subscenarios, the differences between the GBLUP and ABLUP models were not significant.

Repetitions of CV scenarios lead to changes in T_{set} composition, and therefore in the relatedness between T_{set} and V_{set} , which could affect prediction accuracies (Scutari et al. 2016). However, CV scenario structures lead only to very slight changes in this relatedness, which were greater between CV scenarios than between repetitions of the same CV scenario (9.1.).

3.1.2 Within-family genomic prediction accuracy

The large number of individuals per family in our design made it possible to assess within-family predictive ability.

Figure 4 shows the predictive abilities obtained with the CV1 scenario when 40% of the individuals from each family were included in T_{set} , for each of the two traits (Fig. 4A and B). Within-family predictive ability was therefore estimated for the 60% of individuals per family included in the V_{set} (a mean of 20 individuals for each of the 9 large families and 10 individuals for the other 30 families). For each family, the variance of prediction accuracies was very high, indicating that the choice of individuals in the T_{set} (and therefore in the V_{set}) had a major impact on prediction accuracy. Mean prediction accuracy differed between families (Fig. 4), ranging from $-0.43 (\pm 0.20)$ to $+0.45 (\pm 0.18)$ for HT and from $-0.35 (\pm 0.24)$ to $+0.46 (\pm 0.22)$ for DEV. These mean values were centred around 0, with a very low across-family mean ($-0.01 (\pm 0.33)$ for HT and $+0.06 (\pm 0.33)$ for DEV). Despite the similar distribution of within-family predictive ability for the two traits, the ranking of families differed significantly between HT and DEV (Kendall correlation of $+0.01$).

We further investigated within-family predictive ability using the CV2 scenario (Fig. 5). Prediction accuracy was low for both traits, but slightly higher for families well-related to the T_{set} than for poorly related families (mean $+0.12 (\pm 0.08)$ for HT and $+0.16 (\pm 0.09)$ for

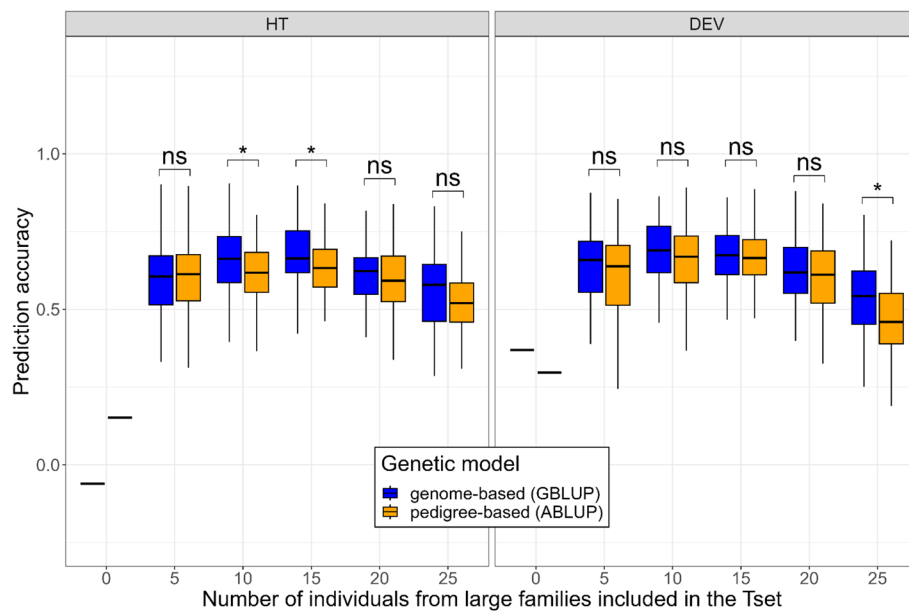


Fig. 3 Global prediction accuracies obtained in the different sub-scenarios CV2 with ABLUP and GBLUP models, for height (HT) and stem deviation to verticality (DEV)

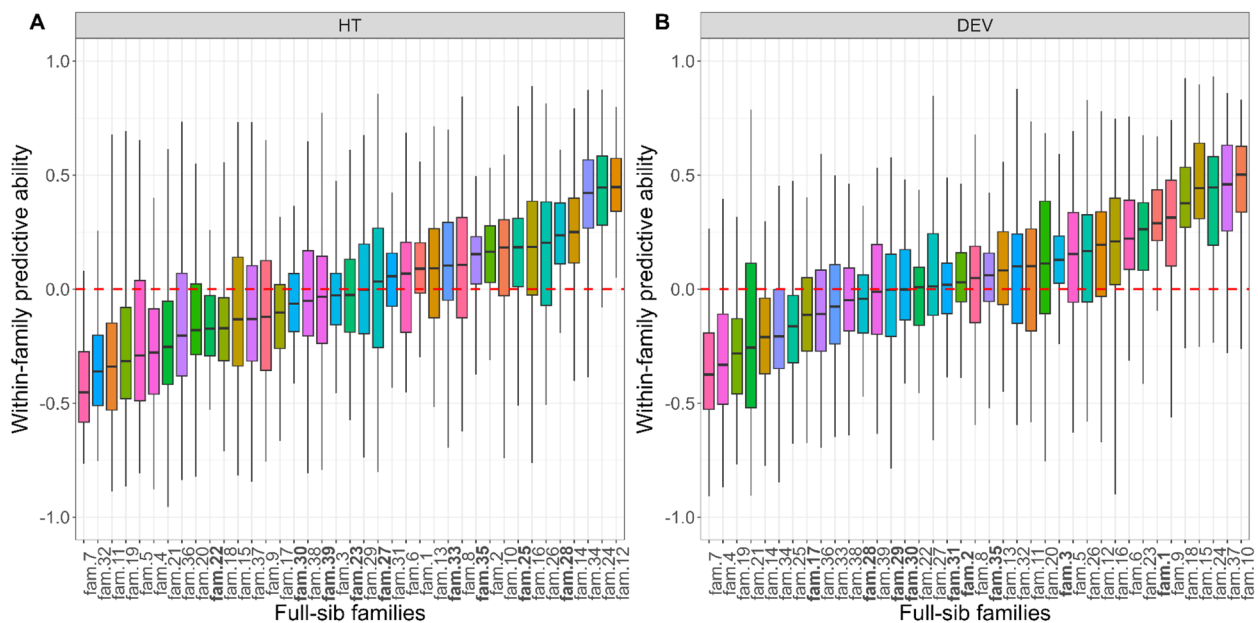


Fig. 4 Genome-based within-family predictive ability obtained in the scenario CV1 for each of the 39 full-sib families, for height (HT) and stem deviation to verticality (DEV)

DEV). Adding more individuals to the T_{set} had no impact, increased or, more surprisingly, decreased within-family predictive ability especially for families poorly related to the T_{set} as observed for both traits in family 22. For both scenarios (CV1 and CV2), within-family predictive abilities remained close to 0 on average, despite strong

variation between families, but could potentially explain the equivalence in terms of global prediction accuracy between the GBLUP and ABLUP models. In this context, simulations can be more informative in investigating the determinants of between- and within-family predictive ability.

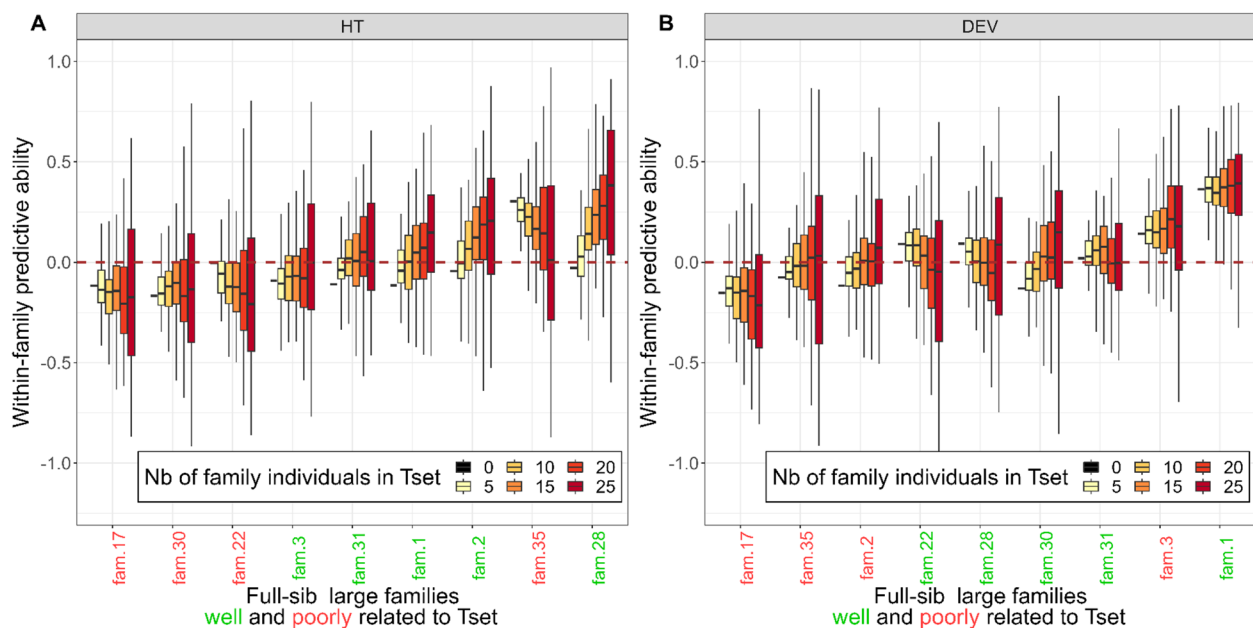


Fig. 5 Genome-based within-family predictive ability obtained in the sub-scenarios CV2 for each of the 9 large full-sib families, for height (HT) and stem deviation to verticality (DEV)

3.2 Use of simulations to explore prediction accuracy

3.2.1 A relevant simulation model

The comparison of simulated and real data is an interesting preliminary step in assessments of the relevance of the simulation model. For both models, GBLUP and ABLUP, the simulated and real data yielded very similar prediction accuracies (10.1.1.) in terms of the mean ($0.54 (\pm 0.07)$ and $0.55 (\pm 0.08)$ for real and simulated data, respectively, with the GBLUP model, $0.49 (\pm 0.07)$ and $0.47 (\pm 0.07)$, respectively with the ABLUP model), and the distribution of values (the coefficient of variation for prediction accuracy being 15% and 14% for real and simulated data, respectively, with the GBLUP model, and 15% and 16%, respectively, with the ABLUP model). With the simulated data, prediction accuracy was slightly better for GBLUP than for ABLUP. The agreement between the simulated and real data was very high when accuracy comparisons were made within families (10.1.1.). The high variance of the within-family predictive abilities in each family reported previously based on real data was also observed with the simulated data, as were the large differences between families in terms of mean within-family predictive ability. Within-family predictive ability ranged from $-0.43 (\pm 0.20)$ to $+0.45 (\pm 0.18)$ with the real data, and from $-0.31 (\pm 0.27)$ to $+0.42 (\pm 0.28)$ with the simulated data. The ranking of families based on within-family predictive ability differed between the real and simulated data, as the genotypes were simulated independently of the real data.

3.2.2 Determining relevant conditions for GS implementation

Starting from the initial conditions defined by the real data ($h^2=0.13$, $n_{\text{SNP}}=8235$ and $n_{T_{set}} \in [167 : 667]$ depending on the CV1 subscenario considered), new scenarios with different conditions were explored through simulations. Figure 6 shows the overall prediction accuracy for different combinations of h^2 , $n_{T_{set}}$ and n_{SNP} . The objective was not only to study the behaviour of accuracy as a function of the variation of these key parameters, but also to compare it with the accuracy of predictions based solely on pedigree, the reference in many forest tree improvement programmes.

Marker density did not appear to be of a critical importance for GS accuracy, as curves associated to genomic models (blue, red and green) overlapped in most situations. The greatest benefits of a higher marker density (12 and 24 markers/cM) in the tested scenarios were observed in situations in which a large T_{set} was combined with high heritability. A similar combination of parameters also resulted in a lower level of variation in accuracy. Similar results were obtained for the highest two densities, suggesting that the response is saturated when there are more than 17,000 markers.

Regardless of heritability, prediction accuracy appeared to be highly dependent on T_{set} size. GS prediction accuracy increased steadily for the first few increases in T_{set} size ($n_{T_{set}}=400$, 600 , 1600) and then tended to stabilize at about $n_{T_{set}}=2600$, reaching a plateau with inflexion

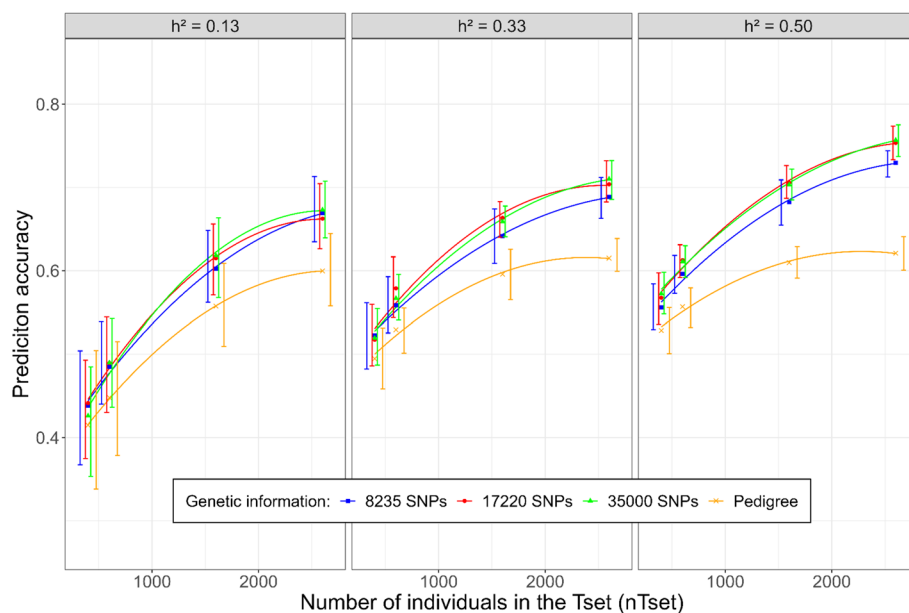


Fig. 6 Global prediction accuracies for height (HT) with simulated data for different combinations of heritabilities (h^2), training set sizes (nT_{set}) and marker densities ($nSNP$)

points between $nT_{set} = 1600$ and $nT_{set} = 2000$. The prediction accuracy of ABLUP models followed a similar pattern, but the advantage of GBLUP models over ABLUP models was greatest for larger T_{set} sizes and higher heritability. For a trait of average heritability ($h^2 = 0.33$), the advantage of GBLUP models in terms of mean prediction accuracy was only $+0.02 (\pm 0.08)$ when $nT_{set} = 400$, whereas it reached $+0.08 (\pm 0.05)$ when $nT_{set} = 1600$ ($+0.13 (\pm 0.03)$ with 2600). The simulation results clearly indicate that the conditions under which the real data were analysed were not optimal for revealing any advantage of the GBLUP model. According to the simulation data, the GBLUP model would be much more advantageous at higher nT_{set} ($nT_{set} \geq 1600$) values, for traits of intermediate or higher heritability and with a larger number of markers.

For two of the situations in which GBLUP performed significantly better than ABLUP — with a minimal difference in one case ($+0.06 (\pm 0.05)$ in mean accuracy) and for the maximum observed difference in the other ($+0.13 (\pm 0.03)$), we present within-family predictive abilities in Fig. 7. As for all the within-family predictive abilities presented above, the variance was high and mean values differed considerably between families. However, in this case, mean prediction ability values were positive for most families, ranging from $-0.03 (\pm 0.18)$ to $0.34 (\pm 0.13)$ for the first situation ($h^2 = 0.33$, $nT_{set} = 1600$, $nSNP = 17,220$) and from $0.07 (\pm 0.14)$ to $0.50 (\pm 0.12)$ for the second situation ($h^2 = 0.5$, $nT_{set} = 2600$, $nSNP = 35,000$). Taking all the families into account, mean within-family predictive

ability was $0.18 (\pm 0.14)$ and $0.29 (\pm 0.17)$, respectively, for the two situations described, indicating a clear advantage of genome-based approach over the reference value of 0 associated with the pedigree-based model. The advantage of the GBLUP model over the ABLUP model in terms of overall prediction accuracy coincides with non-zero within-family predictive ability. This advantage of the GBLUP models was increasingly evident for higher values of within-family predictive ability.

4 Discussion

Genomic selection is particularly interesting for forest trees in order to reduce the length of selection cycles by predicting the breeding values of the next generation (Bartholomé et al. 2016; Thistlethwaite et al. 2019; Simiqueli et al. 2023; Haristoy et al. 2023). However, the maintenance of predictive ability over generations can be hampered by the processes of genetic recombination, selection and drift, which alter the linkage disequilibrium and relatedness between the training population and the population under selection (Grattapaglia 2022). A strategy of “updating” the genomic selection model is recommended by integrating individuals from each generation into the training set (Iwata et al. 2011; Grattapaglia 2022), which implies implementing progeny tests that slow down the selection process. In addition, new crosses are dependent on the time required for trees to reach sexual maturity—between 3 and 10 years depending on the species.

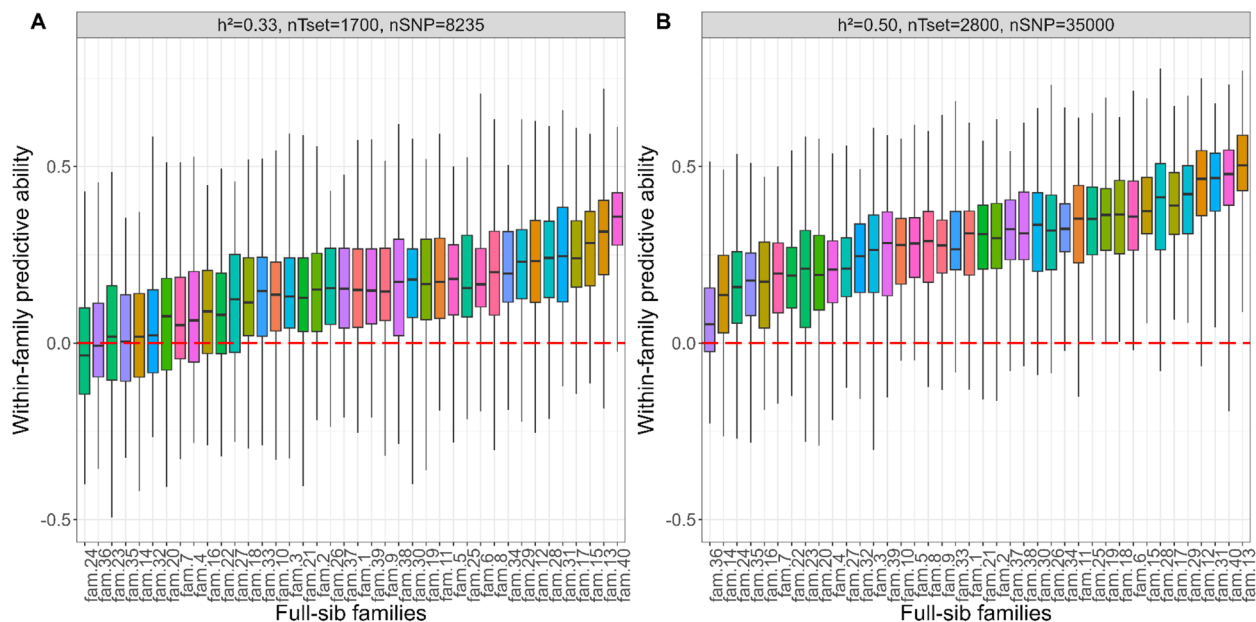


Fig. 7 Genome-based within-family predictive ability for each full-sib family for height (HT), with simulated data and for two different combinations of heritability (h^2), training set size (nT_{set}) and marker density (nSNP)

By concentrating its phenotyping effort on a training population of limited size, genomic selection also makes it possible to increase the selection intensity and to integrate new selection criteria. These criteria can be more complex and better adapted to the challenging environmental conditions (Rivers et al. 2023), targeting for example drought tolerance (Bouvet et al. 2020; Papin et al. 2024) or pathogen resistance (Stocks et al. 2019; Beaulieu et al. 2020; Lenz et al. 2020b; Westbrook et al. 2020). Our study therefore focuses on within-generation and within-family prediction that can be particularly valuable in this context.

Before implementing genome-based prediction and selection in maritime pine and forest trees more generally, we need to clearly define the conditions in which the use of this approach has real advantages over conventional methods, particularly in terms of prediction accuracy (Bartholomé et al. 2016; Beaulieu et al., 2014). We evaluated genome-based prediction accuracies by clear comparison with the corresponding pedigree-based prediction accuracies, and we investigated within-family predictive ability. To justify the genotyping of individuals, genome-based models must provide higher prediction accuracies than pedigree-based models. We found that even with a test design that a priori favoured within-family variability, the advantage of GBLUP over ABLUP models was not clearly significant, and that within-family predictive ability was zero, on average, across families. Stochastic simulations mimicking our selection

scheme showed that our real case was close to the tipping point for recommended training set size, beyond which genome-based predictions would start to show its full potential relative to conventional pedigree-based predictions.

4.1 Equivalence of ABLUP and GBLUP prediction accuracies with a maritime pine dataset

4.1.1 Baseline for comparison of the ABLUP and GBLUP models

Bartholomé et al. (2016) reported a higher prediction accuracy for ABLUP than for GBLUP models in maritime pine, with similar conclusions drawn in several other studies on forest trees (Zapata-Valenzuela et al. 2012; Thistlethwaite et al. 2017; Zhou et al. 2020). An insufficiently high marker density has frequently been identified as the cause of a lack of advantage of genome-based predictions. Marker density was higher in this study than in previous studies in maritime pine (5.7 SNPs/cM versus 2.4 and 1.7 SNPs/cM in the articles by Bartholomé et al. 2016 and Isik et al. 2016, respectively). Furthermore, the original design of POP_R, characterized by a limited number of full-sib families but relatively large numbers of trees per family, was expected to favour the genome-based approach, better capturing Mendelian sampling terms that drive the within-family variation.

However, GBLUP prediction accuracies obtained in this study were moderate (Fig. 2). Prediction accuracies

were equivalent to or slightly lower than the accuracy obtained in previous studies on maritime pine (Isik et al. 2016; Bartholomé et al. 2016) which proposed GS scenarios that were quite different compared to our strategies. Above all, GBLUP models had only a slight advantage over ABLUP models in terms of prediction accuracy for some of the cross-validation subscenarios evaluated (Figs. 2 and 3). The reasons for this slight advantage are difficult to identify and could even be partially due to the way in which prediction accuracy is calculated. Prediction accuracy, defined as the ratio of predictive ability to the square root of heritability of the predicted trait, requires the estimation of variance components, which are subject to additional errors (Legarra et al. 2008). As in other studies, GBLUP estimates of heritability were lower than the estimates obtained with the ABLUP model (Resende et al. 2017; El-Dien et al. 2018; Lenz et al. 2020b). This resulted in higher prediction accuracies for the GBLUP model despite a predictive ability similar to that for the ABLUP model.

We also decided to base our comparison solely on the accuracy, without converting it into a response to selection. Pedigree- and genome-based predictions for non-phenotyped individuals are both usable for early selection without phenotype observations and must be compared in terms of accuracy. We believe that the assumption of a shorter selection cycle for genome-based predictions compared to pedigree-based prediction does not hold and that this assumption introduces a bias, increasing the perceived advantage of GS over traditional approaches (Lenz et al. 2017; Ratcliffe et al. 2015; Resende et al. 2012a, 2012b). Genome-based prediction is clearly advantageous when it enables pedigree recovery and correction (Zapata-Valenzuela et al. 2012), but similar results can routinely be obtained with a very small number of carefully chosen markers across the genome (Vidal et al. 2017). For the use of genome-based prediction to be considered truly advantageous, it must, therefore, provide a higher accuracy than pedigree-based prediction associated with a complete corrected pedigree.

4.1.2 Capture of linkage disequilibrium and relatedness in GBLUP models

Theoretically, the genome-wide markers used for genome-based predictions can provide additional information over and above that obtained from a simple pedigree. They can, for example, capture the effects of nearby QTL in linkage disequilibrium (Habier et al. 2007), and provide more accurate information about the relationship between any two given individuals (Nejati-Javaremi et al. 1997; VanRaden 2008). However, these two advantages of marker use did not emerge clearly in this study. In our case, there are two principal explanations for

the equivalence between ABLUP and GBLUP (Figs. 2 and 3), as outlined by other authors faced with similar results (Beaulieu et al. 2014). First, the estimation of genomic effects may not be very accurate, potentially due to the T_{set} being too small to guarantee a good estimate (Habier et al. 2013). For example, with simulation, Hayes et al. (2009b) showed that the advantage of GBLUP over ABLUP models was apparent from 100 individuals per full-sib family. Second, it is possible that genome-wide markers capture only existing genetic relationships (Legarra et al. 2008), being much less effective at capturing actual linkage disequilibrium (LD) (Habier et al. 2007).

A lack of long-range LD has already been reported for conifers (Eckert et al. 2010; Kujala & Savolainen, 2012), including maritime pine (Plomion et al. 2014; Isik et al. 2016), suggesting that a large number of markers may be required for genome-based predictions (Resende et al. 2017; Thistlethwaite et al. 2019). Grattapaglia et al. (2011) performed simulations for forest trees that suggested that a density of about 10 markers/cM was required to increase the prediction accuracy of GBLUP to such an extent that it surpassed ABLUP when $N_e=100$. Most conifer genomes have a much smaller genetic map than physical size [1435 cM (Chancerel et al. 2013) and 24 Gb (Chagné et al. 2002), respectively, for maritime pine], suggesting that there is very little recombination in large parts of the genome. Thus, rather than the number of markers, it may be their distribution across the genome that counts. However, it can be difficult to improve the distribution of markers in the absence of a physical map for the species, as in maritime pine. The densities used in our real case would, therefore, be sufficiently high to capture pedigree-like relatedness, but they are probably still too low and the markers are probably not distributed appropriately to capture QTL information across LD. One alternative not considered here would be the use of alternative statistical approaches based on Bayesian variable selection, such as Bayes-B. Such methods have been shown to capture population LD more effectively and to give a greater weight to causal SNPs (Habier et al. 2007; Thistlethwaite et al. 2017), but their benefits are often case- and trait-specific, and may even disappear, particularly if the training population exceeds a certain size (Karaman et al. 2016).

4.1.3 Genome-based within-family predictive ability

Unlike ABLUP model, GBLUP model capture realized relatedness between and within families through the **G** matrix. Nevertheless, regardless of the scenario tested here, within-family predictive ability for the real data was, on average, zero when all full-sib families were considered (Fig. 4), indicating that underlying genome-wide

marker associations in GBLUP (Strandén and Garrick 2009) were not estimated accurately. For both traits, within-family predictive abilities were not significantly different from zero in most families, but there was considerable variation across the families, including some (15%) for which accuracy values were, surprisingly, significantly negative. For some relatively small families, the number of individuals included in the V_{set} was probably too small for a robust estimation of correlations, resulting in a very large standard deviation when all CV iterations were taken into account. By contrast, within-family predictive ability for large families was calculated with a mean of 20 individuals (in the V_{set}), resulting in zero or slightly positive values, but with a lower standard deviation. Genome-based prediction accuracy was, therefore, mostly driven by capturing parent average term rather than capturing the Mendelian sampling term of breeding values. This partly explains the observed equivalence with pedigree-based prediction models, as suggested in other studies with similar results (Thistlethwaite et al., 2019).

Some specific scenarios of cross-validation imposing restrictions on relatedness between V_{set} and T_{set} , particularly if V_{set} did not contain relatives of individuals present in T_{set} , resulted in particularly low prediction accuracies at the global (Fig. 3) and within-family levels (Fig. 5). This suggests that relatedness was the main source of information in both GBLUP and ABLUP, with little or no additional information captured from LD to maintain prediction quality in the absence of relatedness. Globally, the relatedness between families in our population was low, with parents and grandparents at the founding level of the population producing a mean of 1.1 and 2.4 full-sib families, respectively. This structure may pose challenges that prevent GBLUP from outperforming ABLUP. In general, the use of a diversified training population is always desirable, to produce robust predictions and a validation set well related to the training set. The first condition could have been met by ensuring the equal representation of different grandparents and parents. The second condition is less easy to satisfy, as the siblings in V_{set} probably had few collaterals other than the remaining siblings already present in T_{set} .

4.2 Use of simulations to identify conditions in which GBLUP has a better prediction accuracy than ABLUP

The stochastic simulation model yielded results similar to those obtained empirically, suggesting that we can have some confidence in the relevance of the results, in terms of both the trends observed and the absolute values obtained. The effect of increasing T_{set} size and the number of genome-wide markers on GS prediction accuracy was tested under three contrasting heritability scenarios, each with an explicit comparison with the

prediction accuracy of ABLUP. These simulations showed that the size of T_{set} was the most important determinant of prediction accuracy in our study. The inflexion point of the curves occurred somewhere between 1500 and 2000 individuals, consistent with the findings of deterministic approaches in other forest tree contexts (Grattapaglia et al. 2011; Grattapaglia & Resende 2011). From this T_{set} size upwards, prediction accuracy begins to be significantly higher for GBLUP than for ABLUP. Increasing the number of individuals per family, as in the simulation scenarios in which T_{set} increased, made it possible to estimate breeding values with greater accuracy (Habier et al. 2013). Thus, the often-highlighted equivalence between GBLUP and ABLUP in terms of prediction accuracy for forest trees, regardless of the species considered, can be explained by the training sets, which rarely exceed 1000 individuals, being too small (Lebedev et al., 2020). By contrast, increasing the number of markers did not increase significantly the genome-based predictive accuracy, at least for traits with moderate heritability, suggesting that our initial marker density would have been sufficient if it had been coupled with a larger training population. It should be noted that prediction accuracy is also impacted by the effective size of the breeding population, although we did not vary this parameter in our study (Beaulieu et al. 2024). Relatedness between T_{set} and V_{set} is one of the factors explaining the success of predictions (Grattapaglia and Resende 2011). Thus, our conclusions could be readily extrapolated to other conifers, which have similar genome sizes and effective breeding population sizes, but it would be more difficult to generalize it to deciduous species, which can have very different genome structures and LD profiles.

Complementary analysis of genome-based prediction accuracy based on deterministic approaches (Daetwyler et al. 2008) with empirical parameters for our population of maritime pine (see 11.) were consistent with our stochastic simulation results.

Within-family prediction accuracy is rarely considered in genomic studies but appears to be key for the superiority of genome-based predictions over pedigree-based predictions to be expressed. Simulations showed that more accurate within-family prediction was associated with a greater accuracy advantage of GBLUP over ABLUP. For our study design, the suggested T_{set} size for efficient within-family prediction corresponds to between 40 and 65 individuals per full-sib family. This is a very important requirement for the implementation of genome-based prediction and selection, as the use of GBLUP models with zero within-family predictive abilities can have several negative consequences.

In the short term, the effectiveness of selection and the response to selection would be reduced if one of the

sources of genetic variation in the population, the within-family variation due to Mendelian sampling, was not exploited. While this source of variation can be measured with genome-wide markers at the genotype level, attaching quantitative value to these genotypes requires sufficiently powered training set of phenotyped and genotyped individuals. Ensuring such sufficiently large training set is important to avoid longer-term consequence related to the fact that selection on underpowered genome-based predictions would be only leveraging variation between families. Such an approach would increase the risk of losing diversity due to the elimination of certain lineages and the co-selection of candidates from the same families. This loss of diversity due to a shift in the weighting between within-family and between-family selection would lead to long-term losses of genetic gain (Jannink 2010) and an accumulation of inbreeding.

This tendency can be counteracted by selection methods based on the optimization of genetic contributions (Meuwissen 1997; Woolliams et al. 2015) — so-called “optimal contribution selection” (OCS) — which allows a trade-off between short-term and longer-term gains through the application of constraints to the balance between parental genetic contributions (Gorjanc et al. 2018). Future studies should assess these optimal strategies which would, presumably, work better if genome-based predictions could discriminate between candidates within families more accurately (Hallander & Waldmann, 2009), with sufficiently large families in the training populations, thereby increasing the efficiency of selection and of the constraints imposed by the OCS.

5 Conclusion

Despite the undeniable potential benefits for forest trees, examples in which genome-based approaches have clearly demonstrated superiority over pedigree-based approaches are lacking. Using an ABLUP model with full and corrected pedigree information as a reference, we evaluated the accuracy of GS in a maritime pine trial with the largest number of individuals per full-sib family to date. Prediction accuracy was found to be similar for the pedigree-based and genome-based models, and within-family genomic accuracy for forward predictions was close to zero. By constructing a relevant simulation model, we were able to demonstrate that the number of individuals per family, and thus the overall size of the training set, is a key parameter for accurately estimating marker associations and for detecting a clear advantage of genome-based approach. This conclusion can be extended to many forestry contexts in which the equivalence between ABLUP and GBLUP prediction accuracies

can be explained by suboptimal training set sizes and structures. Increases in training-set size may be readily achievable in forestry due to the large numbers of individuals commonly used in breeding programmes and decreasing genotyping costs. Effective within-family prediction, based on well-scaled genome-based approaches, will be key to maintaining diversity in the long term and ensuring genetic gain in the challenging years ahead.

Appendix 1

Definition of sampling within the trial for genome-based analysis

The initial maximum capacity for genotyping in this study was 800 individuals. Thus, preliminary simulations were performed to determine a relevant choice of these 800 individuals among the total population in the trial.

Choice of sampling scenario and families

We compared 3 sampling scenarios to select 800 individuals among the 89 families present in the trial: 40 families with 20 individuals each (A), 32 families with 25 individuals each (B) and 20 families with 40 individuals each (C). The best scenario was selected on the basis of the prediction accuracy, following a three-step procedure that was replicated 10 times independently:

- Genotypes and phenotypes were simulated for the families present in the trial using the MoBPS software (Pook et al. 2020).
- For each scenario, the subset of 20, 32 or 40 families was selected with the goal to minimize the average genomic relationships among the subset (with genomic relationships calculated using the simulated genotypes) and with a simple trail-and-error optimization process. The aim was to select the subset that maximizes the genetic diversity.
- Genome-based models were run using the R-package *breedR* for each scenario using the simulated data and the accuracy was assessed through a cross-validation routine (the 800 individuals were selected as the T_{set} and the remaining $\sim 2,000$ individuals from the trial were in the V_{set}). For each sampling scenario we assessed both global and within-family predictive ability (see figures below).

Finally, the scenario using 40 families with 20 individuals each, was chosen since it achieved significantly higher genome-based prediction accuracy at the global level (Fig. 8), and similar within-family predictive ability compared to the other 2 scenarios (Fig. 9).

Choice of individuals within each family

Within each of the 40 selected families, we applied the Kennard-Stone algorithm (Kennard and Stone 1969) implemented in the R-package *prospectr* (Stevens and Ramirez-Lopez 2022) to get the subset of 20 individuals that maximize phenotypic diversity (now based on the real phenotypic values) for the three traits of interest: circumference, height, and deviation to verticality at 8 years.

Addition of 200 individuals

Subsequently, additional 200 individuals could be genotyped and added to the study. As the sampling of the 800 individuals was already performed, we chose to add 20 additional individuals in the 10 out of the 40 selected families. These families (in the main manuscript labeled as the “large families”) were finally made up of 40 individuals each.

These 10 families were chosen to provide a contrast in connectivity within the population sampled. They represent the top 5 and bottom 5 families in terms of connectivity, with their average relatedness to the rest of the families sampled calculated as 0.03 and 0.01, respectively, based on pedigree data. Note that this choice of large families was made a posteriori so that the gap in relatedness with the rest of the population between the two groups of large families may be not very pronounced.

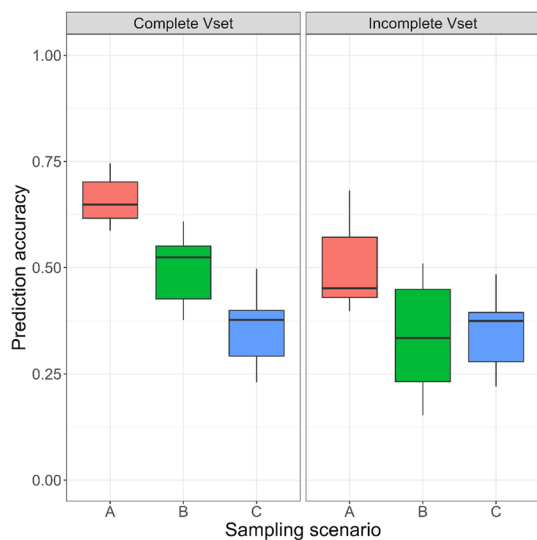


Fig. 8 Genome-based prediction accuracy at the global level for the different sampling scenarios. A: 40 families × 20 individuals, B: 32 families × 25 individuals, C: 20 families × 40 individuals. “Complete Vset” indicates that all individuals in the Vset were considered to assess prediction accuracy, while “incomplete Vset” indicates that only individuals from families not included in the Tset were considered to assess prediction accuracy

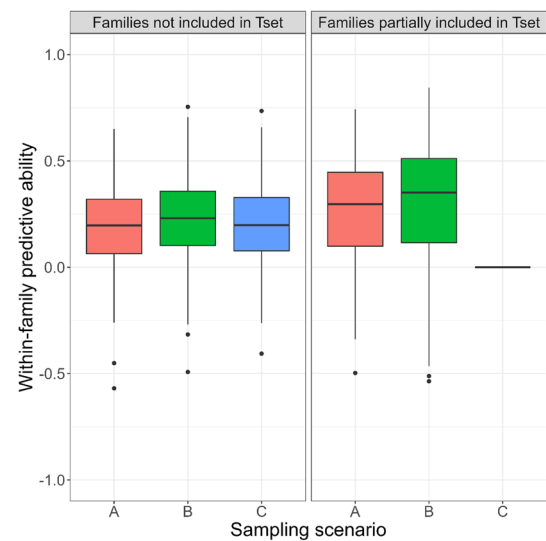


Fig. 9 Genome-based within-family predictive ability for the different sampling scenarios. A: 40 families × 20 individuals, B: 32 families × 25 individuals, C: 20 families × 40 individuals. On the left-hand side, assessing within-family predictive ability for all the families not included in the Tset, and on the right-hand side only for families with some individuals in the Tset. In the latter modality, prediction accuracy is not available considering sampling scenario C since all the individuals in these families are included in the Tset

Appendix 2

Analysis of direct phenotypic values

Direct phenotypic values for HT and DEV are available for all the trees within sites. First, phenotypes were analysed with the following simple model using R package *breedR* (Muñoz & Sanchez 2020):

$$y = 1\mu + Zu + \varepsilon \quad (5)$$

with y the vector of phenotypic observations, μ the population mean, u the vector of regression coefficients for additive genetic random effects and Z the corresponding incidence matrix, ε the vector of residuals. We assumed $u \sim N(0, A\sigma_u^2)$, with A the additive relationship matrix (calculated from pedigree) and σ_u^2 the additive genetic variance. This analysis of phenotypes revealed important spatial autocorrelation within the site, as shown for HT (Figs. 10, 11):

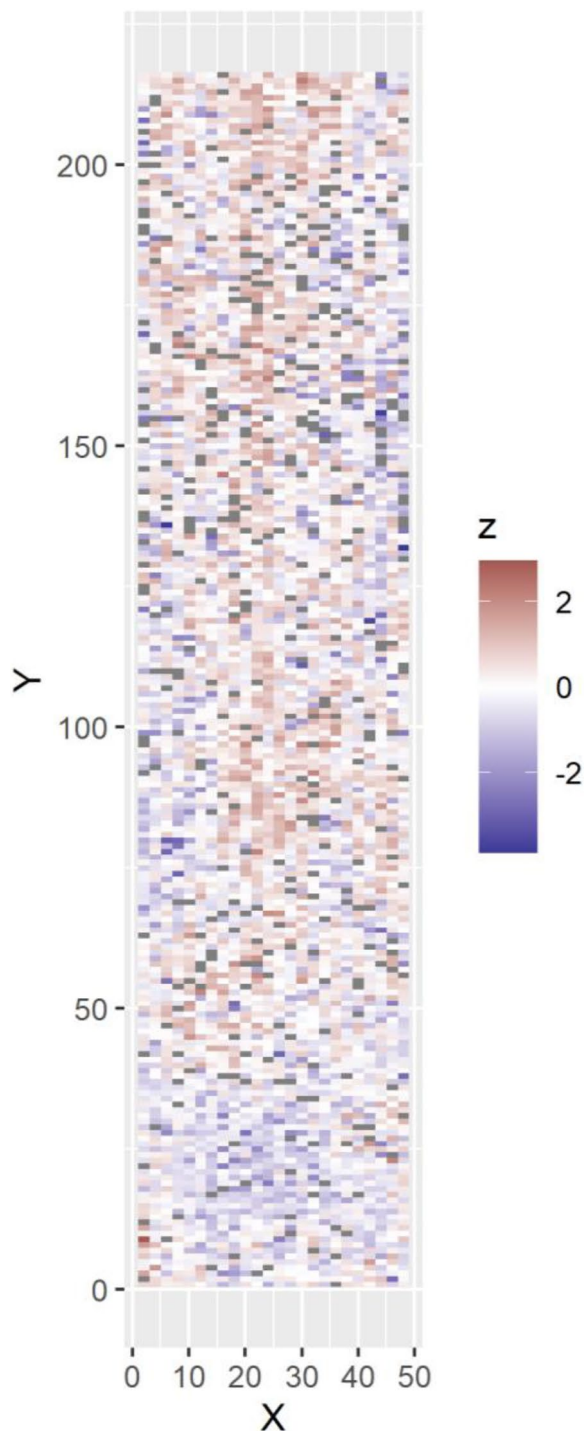


Fig. 10 Spatial representation of residuals from model 1

Each individual is defined by x and y coordinates within the site. There is clearly a spatial autocorrelation, as the residuals are not randomly distributed.

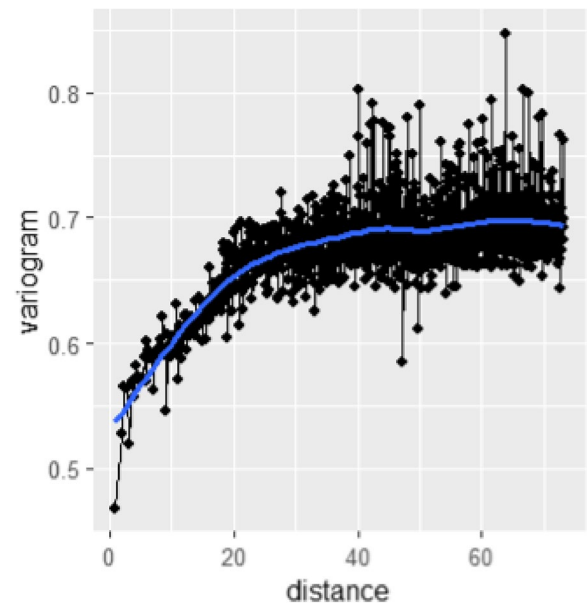


Fig. 11 Isotropic variogram from model 1

The empirical isotropic variogram is built by aggregating all the pairs of points separated by h , no matter the direction (see *breedR* manual). Residuals from nearby individuals are more similar than residuals from distant individuals, pointing to the existence of spatial autocorrelation.

We compared three individual-tree models with *breedR* to estimate the spatial effects affecting phenotypes within the site, as proposed by (Cappa et al. 2019). The models are of the general form:

$$y = 1\mu + Bb + Zu + e \tag{6}$$

with alternative formulations of the spatial random effect **Bb**: block design (block), bidimensional separable first-order autoregressive (AR), and bidimensional spline regression (splines). The goodness-of-fit of the models was assessed by comparing Aikake information criterion (AIC) and the splines approach was found to be the most appropriate.

Model	Model 1	Model 2 block	Model 2 AR	Model 2 splines
AIC	14308	14111	14105	13668

The form of the model chosen therefore corresponds to Eq. 6, with **B** the matrix containing the two-dimensional B-splines basis evaluated in the corresponding row and column for each tree (Fig. 12), and assuming $b \sim N(0, U\sigma_s^2)$ with **U** a fixed spatial structure and σ_s^2 the spatial variance parameter. The residuals of this model are now more randomly distributed (Fig. 13).

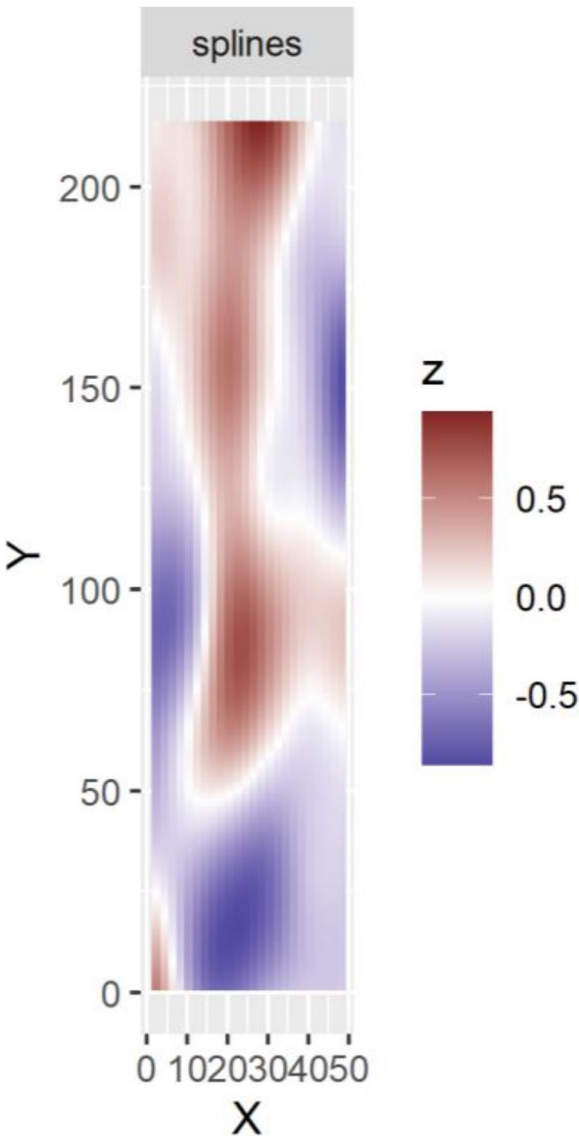


Fig. 12 Spatial component estimated with the model 2 splines

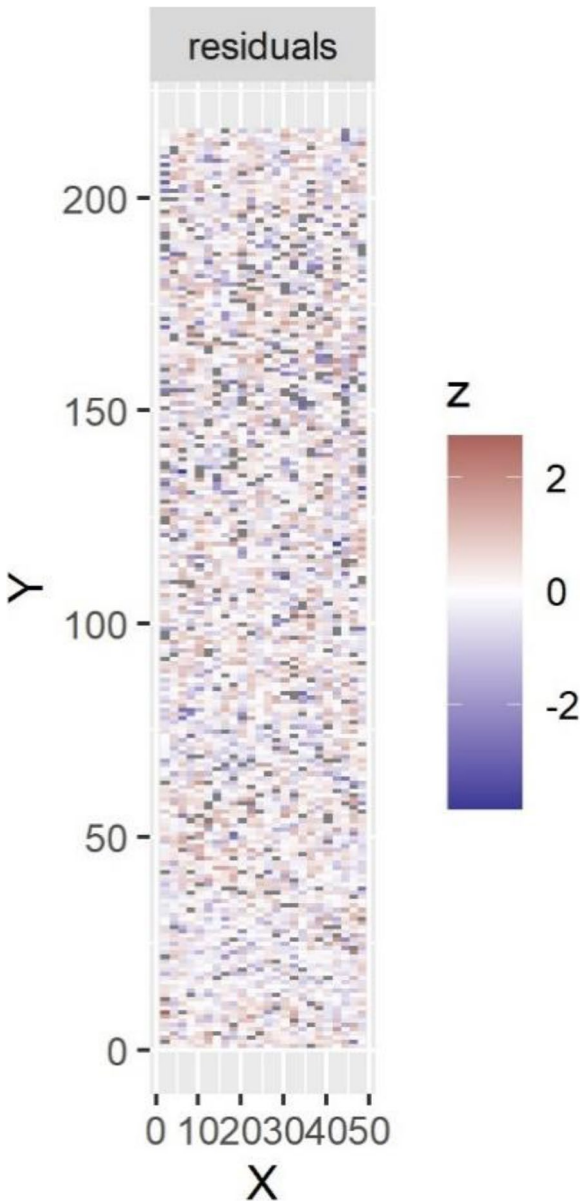


Fig. 13 Spatial representation of residuals from model 2 splines

Finally, we obtained adjusted phenotypic values y' as:
 $y' = y - b$

These phenotypic values y' were subsequently used in the models of the manuscript.

Appendix 3

Simulation of the French maritime pine breeding programme

We simulated a diploid maritime pine genome composed of 12 chromosomes each with a physical length of 2.14×10^9 bp (Chagné et al. 2002), a genetic length of 1.2 M (Chancerel et al. 2013) and a total number of segregating sites of 6910 (10 times the average number of SNP per chromosome available in the real dataset plus 50 theoretical QTL). A simple species-specific demography was used to mimic the selection in the natural environment of the breeding programme base population (popG0): from a wild ancestral population in the Landes forest, we simulated one significant bottleneck reducing effective population size from $\sim 50,000$ (Milesi et al. 2023) to 100 (actual effective size for popG0) and with a mutation rate of 4×10^{-18} per generation (Jaramillo-Correa et al. 2020). We investigated the coherence in LD and allele frequency distributions for the 600 individuals

of popG0, by comparing real genotyping data and simulated SNP array data (see Fig. 14). One phenotypic trait was simulated with reference values equal to those of the “height” trait targeted in the breeding programme (phenotypic average = 6.81 m and genetic variance = 0.12 m^2).

The 833 individuals of POP_R considered in this study were simulated after the two breeding cycles, as closely as possible to the real-life conditions (see Fig. 15). Each individual in popG0 took the identity of a real individual in the pedigree, the one with which it shared the same rank in terms of EBV. The EBV were generated with a correlation of 0.97 with true breeding values (BV) since the accuracy of EBV in real programme for these individuals is very high due to progeny testing. Individuals of popG0 were crossed according to the actual crossing plan to generate popP1 FS families of 130 individuals. Individuals for popG1 were selected based on EBV in each family with an intensity of 1.07. The use of this selection intensity, estimated with real data, made it possible to mimic the multi-character aspect of the actual selection carried out, as well as the diversity constraint actually used. Finally, the families of POP_S (simulated version of POP_R) were obtained by crossing the individuals of popG1 according to the actual pedigree.

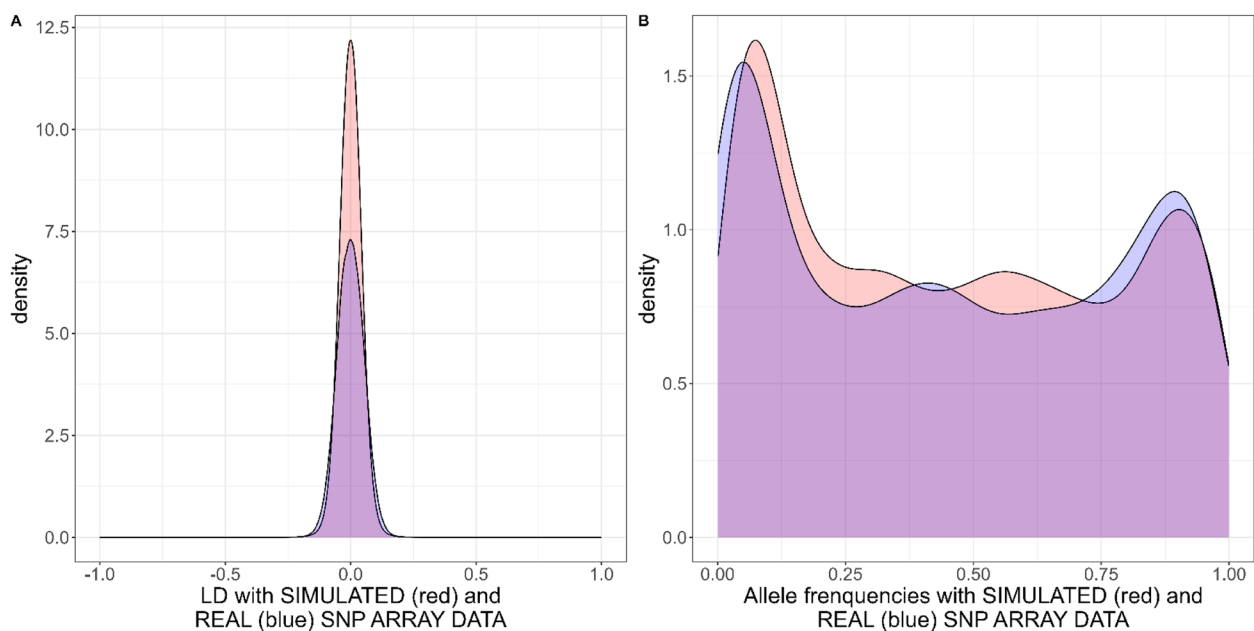


Fig. 14 Comparison of LD (A) and allele frequency (B) distributions between real and simulated SNP array data

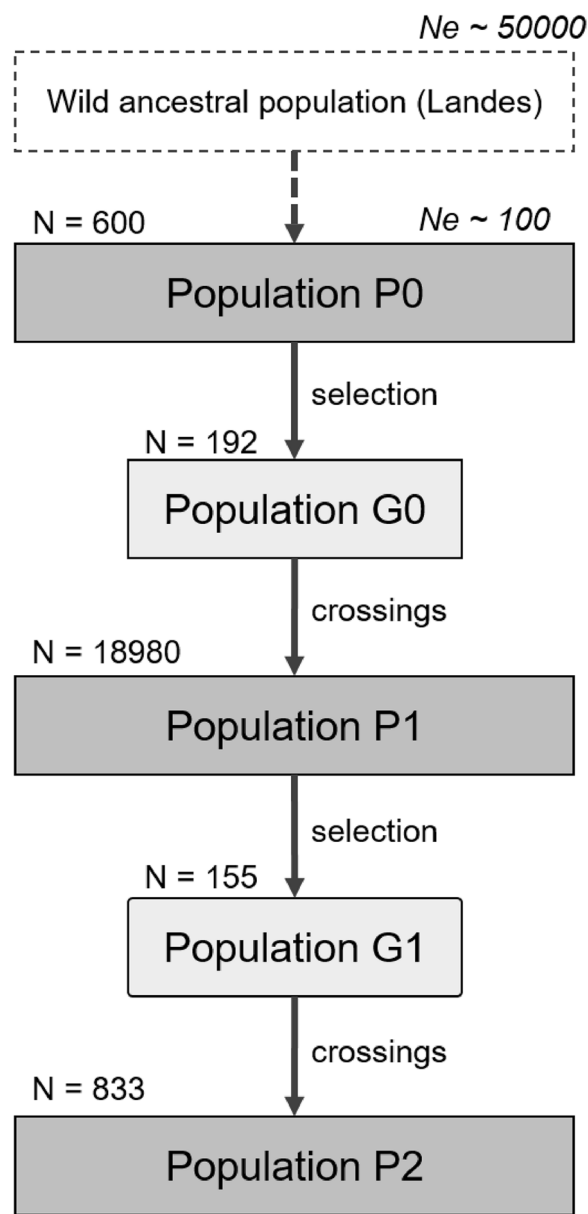


Fig. 15 Breeding cycles simulated to mimic the French maritime pine breeding programme

Appendix 4
Relatedness between training and validation sets
in the different cross-validation scenarios
Mean relatedness between training (T_{set}) and validation (V_{set}) sets in the different cross-validation scenarios

	CV1	CV2					
		Sub-scenario with n individuals from large families included in the T_{set}					
		n=0	n=5	n=10	n=15	n=20	n=25
Mean relatedness between T_{set} and V_{set}	6.2e-04	1.2e-02	1.4e-02	7.4e-03	2.9e-03	9.1e-04	1.2e-03
Standard deviation	2.7e-05	NA	1.0e-04	1.1e-04	5.9e-05	3.7e-05	3.6e-05

Relatedness was calculated for each repetition of the cross-validation scenario as the average of the genomic relationship coefficients in the G matrix between training and validation individuals. Standard deviation is not available for CV2 ($n=0$) as this sub-scenario is repeated only once.

Appendix 5
Verification of the consistency of results obtained with real and simulated data

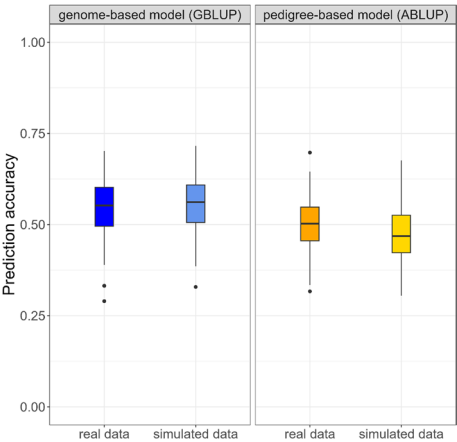


Fig. 16 Global prediction accuracies obtained in the scenario CV1 for height with ABLUP and GBLUP models based on real or simulated data

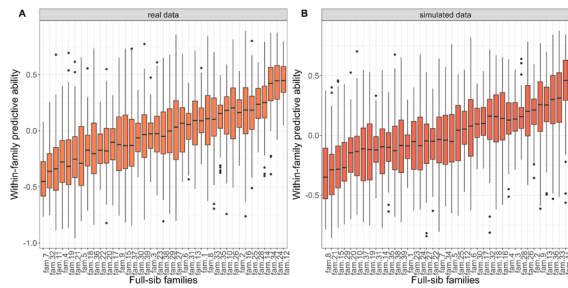


Fig. 17 Genome-based within-family predictive ability obtained in the scenario CV1 for height (HT) for each of the 39 full-sib families with real or simulated data

Appendix 6

Complementary analysis of genome-based prediction accuracy using deterministic approaches

Mathematical modeling approaches are useful tools to predict the accuracy of genome-based prediction as a function of various population parameters (Daetwyler et al. 2008; Hayes et al. 2009b; Goddard 2009). Deterministic formulas have already been applied to forest trees (Grattapaglia & Resende, 2011) and are here adapted with input parameters from our maritime pine case study, following Gorjanc et al. (2015). The accuracy (square root of reliability) of GEBV was obtained by: $r_{EBV} = \sqrt{\frac{\theta}{1+\theta}} b$

Where $\theta = nT_{set}bh^2/M_e$ with nT_{set} the size of the training set and h^2 the heritability of the trait. M_e and b ,

respectively the effective number of chromosome segments and the proportion of genetic variance captured by markers, being defined by $M_e = 2N_eLC/\log(N_eL)$ and $b = nSNP/(nSNP + M_e)$, with $N_e = 100$ the effective size of our breeding population, $L = 1.2$ the average size of maritime pine chromosomes (in Morgans) and $C = 12$ the number of chromosomes. Based on progeny records, the prediction accuracy of non-phenotypes progeny with ABLUP depends solely on the accuracy of EBV in its parents. In case of no selection among parents, we have $r_{EBV}^2_{pedigree} = \frac{1}{4}(r_{EBV}^2_{mother} + r_{EBV}^2_{father})$ with $r_{EBV}^2_{mother} = r_{EBV}^2_{father} = n/(n + \frac{4-h^2}{h^2})$, n being the number of progenies with phenotypic values per parent.

The trends observed (see Fig. 18) are fully consistent with those obtained with stochastic simulations presented in the previous section (Fig. 6). Overall prediction accuracy is mainly impacted by T_{set} size and to a much lesser extent by trait heritability. The advantage of GBLUP over ABLUP models in terms of forward prediction accuracy only becomes apparent above 2,000 individuals in the T_{set} . Although deterministic approaches are very interesting for quickly revealing key parameters for genome-based prediction accuracy, they have been called into question and refined several times (Brard and Ricard 2015; Elsen 2016, 2017). Especially when employing genome-based prediction within full-sibling families, as in this study, accurately forecasting the performance of corresponding accuracy for a specific trait within a particular family remains a challenging endeavour (Schopp et al. 2017).

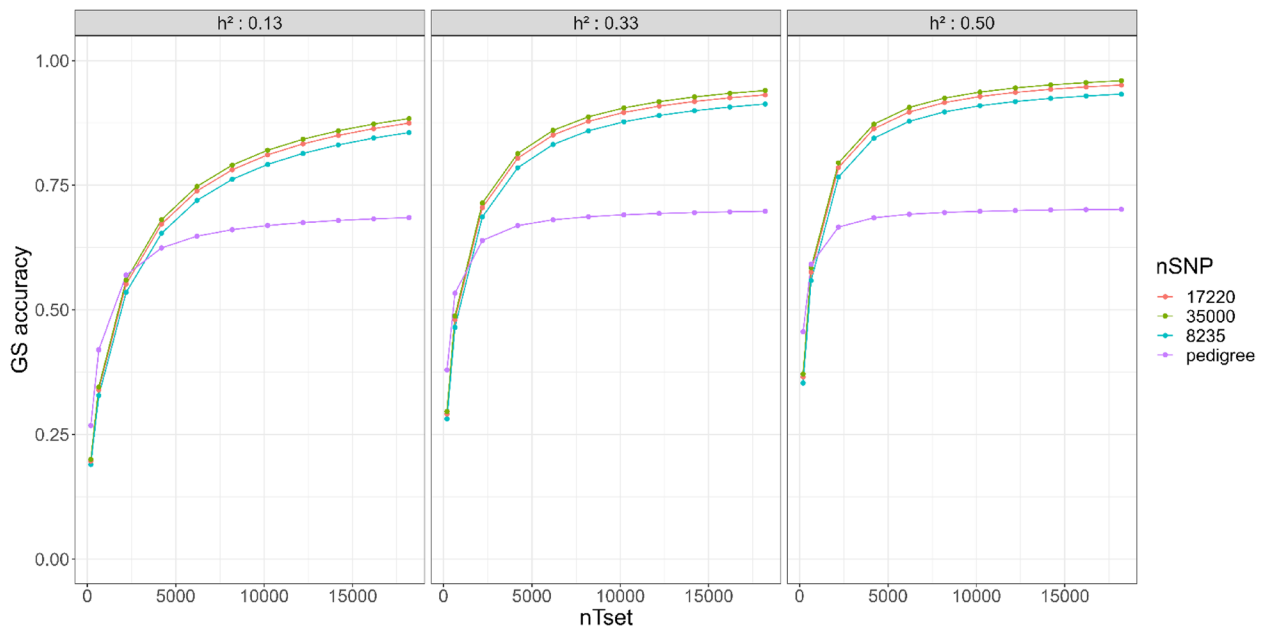


Fig. 18 Genome-based accuracy determined with deterministic formulas

Abbreviations

ABLUP	Pedigree-based best linear unbiased prediction
BLUP	Best linear unbiased prediction
CV	Cross-validation
DEV	Stem deviation to verticality
DNA	Deoxyribonucleic acid
EBV	Pedigree-based estimated breeding value
GBLUP	Genome-based best linear unbiased prediction
GBEV	Genome-based estimated breeding values
GS	Genomic selection
HT	Height
LD	Linkage disequilibrium
nT_{set}	Training set size
$nSNP$	Number of SNP
OCS	Optimum contribution selection
POP_r	Trees sampled for this study
POP_s	Simulated version of POP_r
QTL	Quantitative trait loci
SNP	Single-nucleotide polymorphism
T_{set}	Training set
V_{set}	Validation set

Acknowledgements

The authors would like to thank GIS “Groupe Pin Maritime du Futur” and INRAE—UEFP (<https://doi.org/10.15454/1.5483264699193726E12>) for the installation of the studied sites, the management of the sites, the help to collect data (HT and DEV measurements) and biological material (needles). Part of the experiments (DNA extraction, quantification and manipulation) were also performed at the PGTB (<https://doi.org/10.15454/1.5572396583599417E12>), with the help of Mathilde Flores, Christophe Boury and Céline Lalanne. Authors are also thankful to Christophe Plomion for his help during the conceptualization of this study.

Code availability

The custom code and/or software application generated during the current study are available in the Github.com repository, https://github.com/HighlanderLab/vpapun_pine_gs

Authors' contributions

Conceptualization: Laurent Bouffier, Leopoldo Sanchez; Methodology: Victor Papin, Gregor Gorjanc, Ivan Pocnic, Laurent Bouffier, Leopoldo Sanchez; Formal analysis and investigation: Victor Papin; Writing—original draft preparation: Victor Papin; Writing—review and editing: Gregor Gorjanc, Ivan Pocnic, Laurent Bouffier, Leopoldo Sanchez; Funding acquisition: Laurent Bouffier, Leopoldo Sanchez; Ressources: Laurent Bouffier, Leopoldo Sanchez; Supervision: Laurent Bouffier, Leopoldo Sanchez. The authors read and approved the final manuscript.

Funding

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme Project under grant agreement n°773383 (B4EST). VP was awarded a doctoral fellowship (N°2020-CK-126) from Ecole Nationale Supérieure Des Sciences Agronomiques de Bordeaux-Aquitaine, 1 cours du Général de Gaulle, CS 40201 33175 Gradignan Cedex.

Data availability

The datasets generated during and/or analysed during the current study are available in the DATA INARE repository: <https://doi.org/10.57745/QYBKDE>

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors gave their informed consent to this publication and its content.

Competing interests

The authors declare that they have no competing interests.

Author details

¹INRAE, University of Bordeaux, BIOGECO, UMR 1202, 69 Route d'Arcachon, Cestas 33610, France. ²The Roslin Institute and Royal (Dick) School of Veterinary Medicine, The University of Edinburgh, Edinburgh, UK. ³INRAE-ONF, BioForA, UMR 0588, 2163 Avenue de la Pomme de Pin, CS 40001 Ardon, 45075, Cedex 2 Orléans, France.

Received: 31 January 2024 Accepted: 28 November 2024

Published online: 30 December 2024

References

- Abad Viñas R, Caudullo G, Oliveira S, de Rigo D (2016) Pinus pinaster in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayán, J., de Rigo, D., Caudullo, G., Houston Durrant, T., Mauri, A. (Eds.), European Atlas of Forest Tree Species. Publ. Off. EU, Luxembourg, pp. e012d59+
- Allier A, Teyssède S, Lehermeier C et al (2019) Assessment of breeding programs sustainability: application of phenotypic and genomic indicators to a North European grain maize program. *Theor Appl Genet* 132:1321–1334. <https://doi.org/10.1007/s00122-019-03280-w>
- Amadeu RR, Cellon C, Olmstead JW et al (2016) AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: a blueberry example. *Plant Genome* 9:1–10. <https://doi.org/10.3835/plantgenome2016.01.0009>
- Bartholomé J, Van Heerwaarden J, Isik F et al (2016) Performance of genomic prediction within and across generations in maritime pine. *BMC Genomics* 17:604. <https://doi.org/10.1186/s12864-016-2879-8>
- Beaulieu J, Doerksen TK, MacKay J et al (2014) Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics* 15:1048. <https://doi.org/10.1186/1471-2164-15-1048>
- Beaulieu J, Nadeau S, Ding C et al (2020) Genomic selection for resistance to spruce budworm in white spruce and relationships with growth and wood quality traits. *Evol Appl* 13:2704–2722. <https://doi.org/10.1111/eva.13076>
- Beaulieu J, Lenz PRN, Laverdière J-P et al (2024) A meta-analysis on the effects of marker coverage, status number, and size of training set on predictive accuracy and heritability estimates from genomic selection in tree breeding. *Tree Genet Genomes* 20:1–14. <https://doi.org/10.1007/s11295-024-01653-x>
- Bernardo R (1994) Prediction of Maize Single-Cross Performance Using RFLPs and Information from Related Hybrids. *Crop Sci* 34:1. <https://doi.org/10.2135/cropsci1994.0011183X003400010003x>
- Bouvet J-M, Makouanzi Ekomo CG, Brendel O et al (2020) Selecting for water use efficiency, wood chemical traits and biomass with genomic selection in a *Eucalyptus* breeding program. *For Ecol Manage* 465:118092. <https://doi.org/10.1016/j.foreco.2020.118092>
- Bouffier L, Raffin AA, Dutkowski G (2016) Using pedigree and trait relationships to increase gain in the French maritime pine breeding program. In: IUFRO Conference “Forest Genetics for Productivity”, 14–18 March 2016, Rotorua, New Zealand, 2016. <https://hal.inrae.fr/hal-02801580v1>
- Brard S, Ricard A (2015) Is the use of formulae a reliable way to predict the accuracy of genomic selection? *J Anim Breed Genet* 132:207–217. <https://doi.org/10.1111/jbg.12123>
- Cappa EP, Muñoz F, Sanchez L (2019) Performance of alternative spatial models in empirical Douglas-fir and simulated datasets. *Ann for Sci* 76:1–16. <https://doi.org/10.1007/s13595-019-0836-9>
- Chagné D, Lalanne C, Madur D et al (2002) A high density genetic map of maritime pine based on AFLPs. *Ann for Sci* 59:627–636. <https://doi.org/10.1051/forest:2002048>
- Chancerel E, Lamy J-B, Lesur I et al (2013) High-density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. *BMC Biol* 11:50. <https://doi.org/10.1186/1741-7007-11-50>
- Chen Z-Q, Baisan J, Pan J et al (2018) Accuracy of genomic selection for growth and wood quality traits in two control-pollinated progeny trials using exome capture as the genotyping platform in Norway spruce. *BMC Genomics* 19:946. <https://doi.org/10.1186/s12864-018-5256-y>
- Cros D, Mbo-Nkoulou L, Bell JM et al (2019) Within-family genomic selection in rubber tree (*Hevea brasiliensis*) increases genetic gain for rubber production. *Ind Crops Prod* 138:111464. <https://doi.org/10.1016/j.indcrop.2019.111464>

- Crossa J, Pérez-Rodríguez P, Cuevas J et al (2017) Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends Plant Sci* 22:961–975. <https://doi.org/10.1016/j.tplants.2017.08.011>
- Daetwyler HD, Villanueva B, Woolliams JA (2008) Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. *PLoS ONE* 3:e3395. <https://doi.org/10.1371/journal.pone.0003395>
- Doerksen TK, Herlinger CM (2010) Impact of reconstructed pedigrees on progeny-test breeding values in red spruce. *Tree Genet Genomes* 6:591–600. <https://doi.org/10.1007/s11295-010-0274-1>
- Durán R, Isik F, Zapata-Valenzuela J et al (2017) Genomic predictions of breeding values in a cloned *Eucalyptus globulus* population in Chile. *Tree Genet Genomes* 13:74. <https://doi.org/10.1007/s11295-017-1158-4>
- Durel C-E (1992) Gains génétiques attendus après sélection sur index en seconde génération d'amélioration du Pin maritime. *Revue Forestière Française* 44:341–355. <https://doi.org/10.4267/2042/26331>
- Eckert AJ, van Heerwaarden J, Wegryzn JL et al (2010) Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (*Pinus taeda* L., Pinaceae). *Genetics* 185:969–982. <https://doi.org/10.1534/genetics.110.115543>
- El-Dien OG, Ratcliffe B, Klápště J et al (2018) Multi-environment genomic variance decomposition analysis of open-pollinated Interior spruce (*Picea glauca* × *engelmannii*). *Mol Breeding* 38:26. <https://doi.org/10.1007/s11032-018-0784-3>
- Elsen J-M (2016) Approximated prediction of genomic selection accuracy when reference and candidate populations are related. *Genet Sel Evol* 48:18. <https://doi.org/10.1186/s12711-016-0183-3>
- Elsen J-M (2017) An analytical framework to derive the expected precision of genomic selection. *Genet Sel Evol* 49:95. <https://doi.org/10.1186/s12711-017-0366-6>
- Fuentes-Utrilla P, Goswami C, Cottrell JE et al (2017) QTL analysis and genomic selection using RADseq derived markers in Sitka spruce: the potential utility of within family data. *Tree Genet Genomes* 13:33. <https://doi.org/10.1007/s11295-017-1118-z>
- Gaynor RC, Gorjanc G, Hickey JM (2021) AlphaSimR: an R package for breeding program simulations. *G3: Genes, Genomes, Genet* 11:jkaa017. <https://doi.org/10.1093/g3journal/jkaa017>
- Goddard M (2009) Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* 136:245–257. <https://doi.org/10.1007/s10709-008-9308-0>
- Gorjanc G, Bijma P, Hickey JM (2015) Reliability of pedigree-based and genomic evaluations in selected populations. *Genet Sel Evol* 47:65. <https://doi.org/10.1186/s12711-015-0145-1>
- Gorjanc G, Gaynor RC, Hickey JM (2018) Optimal cross selection for long-term genetic gain in two-part programs with rapid recurrent genomic selection. *Theor Appl Genet* 131:1953–1966. <https://doi.org/10.1007/s00122-018-3125-3>
- Grattapaglia D (2022) Twelve Years into Genomic Selection in Forest Trees: Climbing the Slope of Enlightenment of Marker Assisted Tree Breeding. *Forests* 13:1554. <https://doi.org/10.3390/f13101554>
- Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. *Tree Genet Genomes* 7:241–255. <https://doi.org/10.1007/s11295-010-0328-4>
- Grattapaglia D, Vilela Resende MD, Resende MR et al (2011) Genomic Selection for growth traits in *Eucalyptus*: accuracy within and across breeding populations. *BMC Proc* 5:O16. <https://doi.org/10.1186/1753-6561-5-S7-O16>
- Guilbaud R, Biselli C, Buiteveld J, et al (2020) Development of a new tool (4TREE) for adapted genome selection in European tree species. In: *Proceedings of the Gentree Symposium*. Avignon, France
- Habier D, Fernando RL, Dekkers JCM (2007) The Impact of Genetic Relationship Information on Genome-Assisted Breeding Values. *Genetics* 177:2389–2397. <https://doi.org/10.1534/genetics.107.081190>
- Habier D, Fernando RL, Garrick DJ (2013) Genomic BLUP Decoded: A Look into the Black Box of Genomic Prediction. *Genetics* 194:597–607. <https://doi.org/10.1534/genetics.113.152207>
- Hallander J, Waldmann P (2009) Optimum contribution selection in large general tree breeding populations with an application to Scots pine. *Theor Appl Genet* 118:1133–1142. <https://doi.org/10.1007/s00122-009-0968-7>
- Haristoy G, Bouffier L, Fontes L et al (2023) Genomic prediction in a multi-generation *Eucalyptus globulus* breeding population. *Tree Genet Genomes* 19:8. <https://doi.org/10.1007/s11295-022-01579-2>
- Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009a) Invited review: Genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci* 92:433–443. <https://doi.org/10.3168/jds.2008-1646>
- Hayes BJ, Visscher PM, Goddard ME (2009b) Increased accuracy of artificial selection by using the realized relationship matrix. *Genet Res* 91:47–60. <https://doi.org/10.1017/S0016672308009981>
- Henderson CR (1975) Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics* 31:423–447. <https://doi.org/10.2307/2529430>
- Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genet Res* 93:47–64. <https://doi.org/10.1017/S0016672310000480>
- Isik F (2014) Genomic selection in forest tree breeding: the concept and an outlook to the future. *New Forests* 45:379–401. <https://doi.org/10.1007/s11056-014-9422-z>
- Isik F, Bartholomé J, Farjat A et al (2016) Genomic selection in maritime pine. *Plant Sci* 242:108–119. <https://doi.org/10.1016/j.plantsci.2015.08.006>
- Iwata H, Hayashi T, Tsumura Y (2011) Prospects for genomic selection in conifer breeding: a simulation study of *Cryptomeria japonica*. *Tree Genet Genomes* 7:747–758. <https://doi.org/10.1007/s11295-011-0371-9>
- Jannink J-L (2010) Dynamics of long-term genomic selection. *Genet Sel Evol* 42:35. <https://doi.org/10.1186/1297-9686-42-35>
- Jaramillo-Correa JP, Bagnoli F, Grivet D et al (2020) Evolutionary rate and genetic load in an emblematic Mediterranean tree following an ancient and prolonged population collapse. *Mol Ecol* 29:4797–4811. <https://doi.org/10.1111/mec.15684>
- Karaman E, Cheng H, Firat MZ et al (2016) An Upper Bound for Accuracy of Prediction Using GBLUP. *PLoS ONE* 11:e0161054. <https://doi.org/10.1371/journal.pone.0161054>
- Kennard RW, Stone LA (1969) Computer aided design of experiments. *Technometrics* 11:137–148. <https://doi.org/10.1080/00401706.1969.10490666>
- Kujala ST, Savolainen O (2012) Sequence variation patterns along a latitudinal cline in Scots pine (*Pinus sylvestris*): signs of clinal adaptation? *Tree Genet Genomes* 8:1451–1467. <https://doi.org/10.1007/s11295-012-0532-5>
- Lebedev VG, Lebedeva TN, Chernodubov AI, Shestibratov KA (2020) Genomic selection for forest tree improvement: methods, achievements and perspectives. *Forests* 11:1190. <https://doi.org/10.3390/f11111190>
- Legarra A, Robert-Granié C, Manfredi E, Elsen J-M (2008) Performance of Genomic Selection in Mice. *Genetics* 180:611–618. <https://doi.org/10.1534/genetics.108.088575>
- Lenz PRN, Beaulieu J, Mansfield SD et al (2017) Factors affecting the accuracy of genomic selection for growth and wood quality traits in an advanced-breeding population of black spruce (*Picea mariana*). *BMC Genomics* 18:335. <https://doi.org/10.1186/s12864-017-3715-5>
- Lenz PRN, Nadeau S, Azaiez A et al (2020a) Genomic prediction for hastening and improving efficiency of forward selection in conifer polycross mating designs: an example from white spruce. *Hered* 124:562–578. <https://doi.org/10.1038/s41437-019-0290-3>
- Lenz PRN, Nadeau S, Mottet M-J et al (2020b) Multi-trait genomic selection for weevil resistance, growth, and wood quality in Norway spruce. *Evol Appl* 13:76–94. <https://doi.org/10.1111/eva.12823>
- Li Y, Klápště J, Telfer E et al (2019) Genomic selection for non-key traits in radiata pine when the documented pedigree is corrected using DNA marker information. *BMC Genomics* 20:1026. <https://doi.org/10.1186/s12864-019-6420-8>
- Lindgren D, Gea L, Jefferson PA (1996) Loss of genetic diversity monitored by status number. *Silvae Genet* 45:52–59
- Meuwissen THE (1997) Maximizing the response of selection with a predefined rate of inbreeding. *J Anim Sci* 75:934–940. <https://doi.org/10.2527/1997.754934x>
- Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157:1819–1829. <https://doi.org/10.1093/genetics/157.4.1819>
- Milesi P, Kastally C, Dauphin B, et al (2023) Synchronous effective population size changes and genetic stability of forest trees through glacial cycles. *bioRxiv* 2023–01
- Mrode R, Pocrnic I (2023) *Linear Models for the Prediction of the Genetic Merit of Animals*, 4th Edition. CABI
- Munoz PR, Resende MFR Jr, Huber DA et al (2014) Genomic Relationship Matrix for Correcting Pedigree Errors in Breeding Populations: Impact on Genetic Parameters and Genomic Selection Accuracy. *Crop Sci* 54:1115–1123. <https://doi.org/10.2135/cropsci2012.12.0673>

- Muñoz F, Sanchez L (2020) breedR: statistical methods for forest genetic resources analysts. R package version 0.12–5. <https://github.com/famuv/ie/breedR>
- Papin V, Bosc A, Sanchez L, Bouffier L (2024) Integrating environmental gradients into breeding: application of genomic reactions norms in a perennial species. *Hered* 1–13. <https://doi.org/10.1038/s41437-024-00702-4>
- Nejati-Javaremi A, Smith C, Gibson JP (1997) Effect of total allelic relationship on accuracy of evaluation and response to selection. *J Anim Sci* 75:1738–1745. <https://doi.org/10.2527/1997.7571738x>
- Pégard M, Segura V, Muñoz F et al (2020) Favorable Conditions for Genomic Evaluation to Outperform Classical Pedigree Evaluation Highlighted by a Proof-of-Concept Study in Poplar. *Front Plant Sci* 11:581954. <https://doi.org/10.3389/fpls.2020.581954>
- Plomion C, Chancerel E, Endelman J et al (2014) Genome-wide distribution of genetic diversity and linkage disequilibrium in a mass-selected population of maritime pine. *BMC Genomics* 15:171. <https://doi.org/10.1186/1471-2164-15-171>
- Pook T, Schlather M, Simianer H (2020) MoBPS - Modular Breeding Program Simulator. G3: Genes, Genomes, Genet 10:1915–1918. <https://doi.org/10.1534/g3.120.401193>
- Pryce JE, Daetwyler HD, Pryce JE, Daetwyler HD (2011) Designing dairy cattle breeding schemes under genomic selection: a review of international research. *Anim Prod Sci* 52:107–114. <https://doi.org/10.1071/AN11098>
- Ratcliffe B, El-Dien OG, Klápště J et al (2015) A comparison of genomic selection models across time in interior spruce (*Picea engelmannii* × *glauca*) using unordered SNP imputation methods. *Hered* 115:547–555. <https://doi.org/10.1038/hdy.2015.57>
- R Core Team (2022) R: A Language and Environment for Statistical Computing
- Rauf S, da Silva JT, Khan AA, Naveed A (2010) Consequences of plant breeding on genetic diversity. *Int J Plant Breed* 4:1–21
- Resende JMFR, Muñoz P, Resende MDV et al (2012a) Accuracy of Genomic Selection Methods in a Standard Data Set of Loblolly Pine (*Pinus taeda* L.). *Genetics* 190:1503–1510. <https://doi.org/10.1534/genetics.111.137026>
- Resende MFR Jr, Muñoz P, Acosta JJ et al (2012b) Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments. *New Phytol* 193:617–624. <https://doi.org/10.1111/j.1469-8137.2011.03895.x>
- Resende RT, Resende MDV, Silva FF et al (2017) Assessing the expected response to genomic selection of individuals and families in *Eucalyptus* breeding with an additive-dominant model. *Hered* 119:245–255. <https://doi.org/10.1038/hdy.2017.37>
- Rivers M, Newton AC, Oldfield S, Contributors GTA (2023) Scientists' warning to humanity on tree extinctions. *Plants People Planet* 5:466–482. <https://doi.org/10.1002/ppp3.10314>
- Simiqueli GF, Resende RT, Takahashi EK, et al (2023) Realized genomic selection across generations in a reciprocal recurrent selection breeding program of *Eucalyptus* hybrids. *Front Plant Sci* 14: <https://doi.org/10.3389/fpls.2023.1252504>
- Schopp P, Müller D, Wientjes YCJ, Melchinger AE (2017) Genomic Prediction Within and Across Biparental Families: Means and Variances of Prediction Accuracy and Usefulness of Deterministic Equations. G3: Genes, Genomes, Genet 7:3571–3586. <https://doi.org/10.1534/g3.117.300076>
- Scutari M, Mackay I, Balding D (2016) Using Genetic Distance to Infer the Accuracy of Genomic Prediction. *PLOS Genet* 12:e1006288. <https://doi.org/10.1371/journal.pgen.1006288>
- Stevens A, Ramirez-Lopez L (2022) An introduction to the prospectr package. R package Vignette R package version 0.2.6
- Stocks JJ, Metheringham CL, Plumb WJ et al (2019) Genomic basis of European ash tree resistance to ash dieback fungus. *Nat Ecol Evol* 3:1686–1696. <https://doi.org/10.1038/s41559-019-1036-6>
- Strandén I, Garrick DJ (2009) Technical note: Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. *J Dairy Sci* 92:2971–2975. <https://doi.org/10.3168/jds.2008-1929>
- Thistlethwaite FR, Ratcliffe B, Klápště J et al (2017) Genomic prediction accuracies in space and time for height and wood density of Douglas-fir using exome capture as the genotyping platform. *BMC Genomics* 18:930. <https://doi.org/10.1186/s12864-017-4258-5>
- Thistlethwaite FR, Ratcliffe B, Klápště J et al (2019) Genomic selection of juvenile height across a single-generational gap in Douglas-fir. *Hered* 122:848–863. <https://doi.org/10.1038/s41437-018-0172-0>
- Ukrainetz NK, Mansfield SD (2019) Assessing the sensitivities of genomic selection for growth and wood quality traits in lodgepole pine using Bayesian models. *Tree Genet Genomes* 16:14. <https://doi.org/10.1007/s11295-019-1404-z>
- VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91:4414–4423. <https://doi.org/10.3168/jds.2007-0980>
- Vidal M, Plomion C, Raffin A et al (2017) Forward selection in a maritime pine polycross progeny trial using pedigree reconstruction. *AnnFor Sci* 74:21. <https://doi.org/10.1007/s13595-016-0596-8>
- Vigeland MD (2021) Pedigree analysis in R. Academic Press
- Westbrook JW, Zhang Q, Mandal MK et al (2020) Optimizing genomic selection for blight resistance in American chestnut backcross populations: A trade-off with American chestnut ancestry implies resistance is polygenic. *Evol Appl* 13:31–47. <https://doi.org/10.1111/eva.12886>
- Woolliams J, a., Berg P, Dagnachew B s., Meuwissen T h. e. (2015) Genetic contributions and their optimization. *J Anim Breed Genet* 132:89–99. <https://doi.org/10.1111/jbg.12148>
- Zapata-Valenzuela J, Isik F, Maltecca C et al (2012) SNP markers trace familial linkages in a cloned population of *Pinus taeda*—prospects for genomic selection. *Tree Genet Genomes* 8:1307–1318. <https://doi.org/10.1007/s11295-012-0516-5>
- Zapata-Valenzuela J, Whetten RW, Neale D et al (2013) Genomic Estimated Breeding Values Using Genomic Relationship Matrices in a Cloned Population of Loblolly Pine. G3: Genes, Genomes, Genet 3:909–916. <https://doi.org/10.1534/g3.113.005975>
- Zhou L, Chen Z, Olsson L et al (2020) Effect of number of annual rings and tree ages on genomic predictive ability for solid wood properties of Norway spruce. *BMC Genomics* 21:323. <https://doi.org/10.1186/s12864-020-6737-3>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.