Accelerating the domestication of trees using genomic selection: accuracy of prediction models across ages and environments

M. F. R. Resende Jr[1,3], P. Muñoz[2,3], J. J. Acosta[3], G. F. Peter[3,4], J. M. Davis[3,4], D. Grattapaglia[5,6], M. D. V. Resende[7,8] and M. Kirst[3,4]

[1]Genetics and Genomics Graduate Program, University of Florida, PO Box 103610, Gainesville, FL 32611, USA; [2]Plant Molecular and Cellular Biology Graduate Program, University of Florida, PO Box 110690, Gainesville, FL 32611, USA; [3]School of Forest Resources and Conservation, University of Florida, PO Box 110410, Gainesville, FL 32611, USA; [4]University of Florida Genetics Institute, University of Florida, PO Box 103610, Gainesville, FL 32611, USA; [5]Plant Genetics Laboratory, Embrapa – Recursos Genéticos e Biotecnologia, Parque Estação Biológica, Brasília, DF 70770-970, Brazil; [6]Graduate Program in Genomic Sciences and Biotechnology, Universidade Católica de Brasília–SGAN 916 modulo B, Brasília, DF 70790-160, Brazil; [7]EMBRAPA Forestry, Estrada da Ribeira, km 111 Caixa Postal 319, Colombo, PR 83411-000 Brazil; [8]Department of Forest Engineering, Universidade Federal de Viçosa, Viçosa, MG 36571-000 Brazil

Author for correspondence:
Matias Kirst
Tel: +1 352 846 0900
Email: mkirst@ufl.edu

Received: *11 August 2011*
Accepted: *24 August 2011*

New Phytologist (2012) **193**: 617–624
doi: 10.1111/j.1469-8137.2011.03895.x

Key words: genomic selection, genotype by environment interaction, model estimation, *Pinus taeda*, plant breeding.

Summary

- Genomic selection is increasingly considered vital to accelerate genetic improvement. However, it is unknown how accurate genomic selection prediction models remain when used across environments and ages. This knowledge is critical for breeders to apply this strategy in genetic improvement.
- Here, we evaluated the utility of genomic selection in a *Pinus taeda* population of *c.* 800 individuals clonally replicated and grown on four sites, and genotyped for 4825 single-nucleotide polymorphism (SNP) markers. Prediction models were estimated for diameter and height at multiple ages using genomic random regression best linear unbiased predictor (BLUP).
- Accuracies of prediction models ranged from 0.65 to 0.75 for diameter, and 0.63 to 0.74 for height. The selection efficiency per unit time was estimated as 53–112% higher using genomic selection compared with phenotypic selection, assuming a reduction of 50% in the breeding cycle. Accuracies remained high across environments as long as they were used within the same breeding zone. However, models generated at early ages did not perform well to predict phenotypes at age 6 yr.
- These results demonstrate the feasibility and remarkable gain that can be achieved by incorporating genomic selection in breeding programs, as long as models are used at the relevant selection age and within the breeding zone in which they were estimated.

Introduction

The growing worldwide demand for food and fiber (FAO, 2002) and increasing evidence of climate change (IPCC, 2007) creates a pressing need for the development of more productive germplasm that is adapted to existing and novel sources of biotic and abiotic stress. Marker-assisted selection (MAS) has been proposed as an approach to accelerate plant genetic improvement (Lande & Thompson, 1990; Paterson *et al.*, 1991), and is expected to be particularly valuable for species with long generation times, for characteristics that display low heritability and for selection of traits expressed late in the life-cycle. Early identification of markers for MAS relied largely on quantitative trait loci (QTL) studies derived from the analysis of a few segregating populations. However, weak linkage disequilibrium between markers and traits across different genetic backgrounds and the narrow allelic range captured in QTL studies limited their use to within-family selection strategies (Strauss *et al.*, 1992; Dekkers, 2004; Grattapaglia & Kirst, 2008). Many of the limitations of the QTL approach are being overcome by association genetics. However, because quantitative traits are often controlled by many loci of small effect (Visscher, 2008; Buckler *et al.*, 2009) large populations are required to achieve sufficient statistical power to detect all relevant marker-trait associations (Long & Langley, 1999) – a limitation of many current association studies in non-model crops. Populations designed to achieve high mapping resolution and statistical power, while capturing broad allelic variation, have been established in maize and *Arabidopsis* (Kover *et al.*, 2009; Tian *et al.*, 2011). However, their development is impractical for many crops, particularly those that require several years to reach reproductive age.

Genomic selection (GS) has recently been proposed as an alternative to MAS in crop improvement (Bernardo & Yu, 2007; Heffner *et al.*, 2009), after becoming widely adopted in animal breeding (Hayes *et al.*, 2009; Daetwyler *et al.*, 2010; Hayes & Goddard, 2010). Genomic selection relies on phenotyping and high-density genotyping of a sufficiently large and representative sample of the target breeding population, so that the majority of loci that regulate a quantitative trait are in linkage disequilibrium with one or more molecular markers. Contrary to MAS, in GS the effects of all available genetic markers are estimated simultaneously in a training population, and models are developed to predict the genomic breeding value of progeny in future generations (Meuwissen *et al.*, 2001). In plant species, simulations suggest that GS provides superior efficiency relative to traditional breeding and marker-assisted recurrent selection (Bernardo & Yu, 2007; Heffner *et al.*, 2009; Grattapaglia & Resende, 2010; Iwata *et al.*, 2011). Simulations also indicate that the accuracy of prediction models depends primarily on the level of linkage disequilibrium (LD) in the training population – higher LD improves the likelihood of linkage between markers and a quantitative trait locus. Larger training populations also provide more accurate estimates of marker effects on phenotypes. Finally, the heritability of the trait and the number of QTL regulating its variation also affect model accuracies, because a simpler genetic architecture (i.e. fewer loci that regulate larger fractions of the phenotypic variance) is more easily captured relative to more complex traits. While the importance of marker density, training population size and trait genetic architecture is well established, it is still largely unknown how accurate prediction models remain across different environments. Also, for long-lived perennial species, the suitability of applying one prediction model across multiple ages is unclear. Previous QTL studies in agricultural and forest tree species have suggested that age and genotype × environment interaction effects largely affect the genetic control of complex traits (Paterson *et al.*, 1991, 2003; Kaya *et al.*, 1999).

Here we report the development of prediction models for genomic selection in a breeding population of loblolly pine (*Pinus taeda*), one of the most economically and ecologically important tree species in North America. There are over 30 million acres of pine stands in the south-eastern USA (Wear *et al.*, 2007) – in these stands, loblolly pine is the dominant species, providing over 60% of the timber produced in the USA. However, breeding of pines is costly and time-consuming, and a single breeding cycle can extend for over two decades (White & Carson, 2004). Even genetic improvement programmes that use advanced methods of breeding and propagation require over a decade for completion of a breeding cycle, largely because of the extended period necessary for reliable phenotyping and selection to occur. Incorporation of early-selection based on GS prediction models can virtually eliminate the long period of field-testing. Genomic selection is also expected to be useful for genetic improvement of loblolly pine and other conifers because dramatic changes in wood chemical and physical properties occur from the juvenile to mature phase, delaying selection of superior genotypes. In this study we developed genomic prediction models for growth traits measured at multiple sites, to evaluate the impact of genotype by environment interactions on their accuracy. Training populations were also measured over multiple ages, and models were developed to assess their value in predicting breeding values later in the lifecycle. We show that while a high proportion of the heritable variation in diameter at breast height (DBH) and total height (TH) can be predicted with moderate to high accuracy, the use of models across different climate zones and ages is limited.

Materials and Methods

Population, phenotypes and genotypes

This study was carried out in a structured population of loblolly pine (*P. taeda* L.) derived by crossing 32 parents in a circular mating design, resulting in 61 full-sib families with an average of 15 individuals per family (Baltunis *et al.*, 2005, 2007). This population of 926 individuals is referred hereafter as CCLONES (Comparing Clonal Lines ON Experimental Sites). Parents of CCLONES were sampled to represent Atlantic Coastal Plain, Florida and Lower Gulf provenances of loblolly pine. The 926 individuals in CCLONES were clonally propagated, and eight ramets (i.e. clones) of each individual were planted on each of four sites in the southeastern USA: Palatka and Nassau (Florida, USA), Cuthbert and the B.F. Grant Forest (Georgia, USA). Owing to tree mortality, 790–840 individuals (out of the 926) are represented in each site, and 711 are consistently represented in all four sites. Tests were established using single-tree plots in eight replicates (one ramet of each individual is represented in each replicate) using a resolvable α incomplete block design (Williams *et al.*, 2002). In each test, four replicates were grown under high-intensity culture and four were grown under standard-intensity culture. The traits analysed in this study were DBH and HT. Total height was measured when trees were 1, 2, 3, 4 and 6 yr old, and DBH was measured when trees were 3, 4 and 6 yr old.

The CCLONES population was genotyped using the Illumina Infinium assay (Illumina, San Diego, CA, USA) with 7216 single-nucleotide polymorphisms (SNPs), each representing a unique pine expressed sequences tag (EST) contig (Eckert *et al.*, 2010). A total of 4825 SNPs selected based on quality and reliability of the genotyping calls, according to the BEADSTUDIO ver. 3.1.3.0 software (Illumina), were used in this study. Allele frequency was not considered a criterion to discard SNPs. Genotypic data is publicly available at http://loblolly.ucdavis.edu/bipod/ftp/Genotype_Population_CCLONES.txt.

Variance component estimation and individual breeding values prediction

Analyses were carried out using ASREML v.2 (Gilmour *et al.*, 2006) under the following mixed linear model:

$$y = Xb + Z_1 i + Z_2 a + Z_3 n + Z_4 f + Z_5 d_1 + Z_6 d_2 + e$$

y, measure of the trait being analysed; b, a vector of fixed effects (i.e. culture type and replication within culture type); i, a vector of random incomplete block effect within replication $\sim N(0, I\sigma^2_{\text{iblk}})$, which captures the common environmental block effect. The estimated breeding values (EBVs) are obtained from a, which is a vector of random additive effects of clones $\sim N(0, A\sigma^2_{a})$. The vector n of random nonadditive effects of clones

$\sim N(0,\ I\sigma^2{}_n)$ includes epistasis and within-family dominance effects. The specific combining ability, or among-families dominance effects were captured in f, which is a vector of random family effect $\sim N(0,\ I\sigma^2{}_f)$. The vector d_1 is a vector of random additive by culture type interaction $\sim N(0,\ \mathrm{DIAG}\sigma^2{}_{d1})$, d_2 is a vector of random family by culture type interaction $\sim N(0,\ \mathrm{DIAG}\sigma^2{}_{d2})$, and e is the random residual effect $\sim \mathrm{IDD}(0,\ \mathrm{DIAG}\sigma^2{}_e)$. X and Z_1–Z_6 are incidence matrices and I, A and DIAG are the identity, numerator relationship and block diagonal matrices, respectively. Predicted additive genetic values were deregressed as previously described (Garrick *et al.*, 2009) to obtain the adjusted phenotypic values used in the genomic prediction. Narrow-sense heritability was calculated by $h^2 = \sigma^2_a/(\sigma^2_a + \sigma^2_n + \sigma^2_f + \sigma^2_{d1} + \sigma^2_{d2} + \sigma^2_e)$.

Estimation of the effects

Individual markers had their effects estimated adjusting all the allelic effects simultaneously using the random regression best linear unbiased predictor (RR-BLUP) (Meuwissen *et al.*, 2001). The linear mixed model used was:

$$y = X\beta + Zm + \varepsilon$$

y, vector of adjusted phenotypic values (deregressed additive genetic values); β, vector of fixed effects; m, vector of random marker effects; ε, vector of random error effects. X and Z are the incidence matrices for β and m, respectively. The structure of means and variances of this model are described next, as previously defined (Resende *et al.*, 2008):

$$\begin{array}{ll} m \sim N(0,\ \ G) & E(y) = X\beta \\ \varepsilon \sim N(0,\ \ R = I\sigma^2_\varepsilon) & \mathrm{Var}(y) = V = \mathrm{ZGZ}' + R \\ G = I\sigma^2_m & \end{array}$$

n, the number of marker loci.

Under these settings the genomic mixed models equation for the prediction of m through the genomic BLUP (GBLUP) is equivalent to:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + I\frac{\sigma^2_\varepsilon}{(\sigma^2_a/\eta)} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{m} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

where σ^2_a refers to the total genetic variance of the trait and σ^2_ε is the residual variance.

By RR-BLUP, an estimate of the effect of each marker is obtained by the regression of the adjusted phenotypic records on marker genotypes. The predicted genomic breeding value (GBV) of the individual j is given by $\hat{g}_j = \sum_i Z_{\mathrm{ij}}\hat{m}_i$. The matrix Z was built from the number of alleles observed in each SNP marker (0, 1 or 2) and was standardized to have mean of zero and variance of 1, as previously described by Resende *et al.* (2010). Therefore, the value Z_{ij} for the ith marker in the jth individual can have the values:

$Z_{\mathrm{ij}} = \frac{(0-2p_i)}{\sqrt{\mathrm{Var}(Z_i)}}$ if the individual is homozygous for the first allele (mm);

$Z_{\mathrm{ij}} = \frac{(1-2p_i)}{\sqrt{\mathrm{Var}(Z_i)}}$ if the individual is heterozygous (Mm);

$Z_{\mathrm{ij}} = \frac{(2-2p_i)}{\sqrt{\mathrm{Var}(Z_i)}}$ if the individual is homozygous for the second allele (MM) or;

$Z_{\mathrm{ij}} = 0$ if the individual is a missing data.

This prediction equation assumes a priori that all loci explain an equal amount of the genetic variation, as previously described by Meuwissen *et al.* (2001) and applied by others (Bernardo & Yu, 2007; Muir, 2007). Without standardization the genetic variation of each locus is given by σ^2_a/η where η is related to the number of markers used and is given by $\eta = 2\sum p_i(1-p_i)$ (Resende *et al.*, 2008; Gianola *et al.*, 2009), where p_i is the frequency of one of the alleles of loci i. The total additive genetic variance σ^2_a was estimated by restricted maximum likelihood (REML).

Partitioning the data in training (estimation) and test (validation) data sets

The estimated effects of the markers were validated using a 10-fold cross validation approach using random subsampling replication. Briefly, each population was divided in two subsets. The first population was composed by the majority of the individuals and was used to estimate the marker effects. The second population (the validation population) had their phenotypes predicted based on the marker effects estimated in the training set. Random samples of $N = (9/10) \times N_{\mathrm{T}}$ individuals, were used as training sets while the remaining random set of $N = (1/10) \times N_{\mathrm{T}}$ individuals were used as validation sets, where N_{T} is the total number of individuals in each population. This process was repeated 10 times using a different set of individuals as the validation population each time. Therefore, each fold did not overlap with the others, and at the end of the process all individuals had their phenotypes predicted by genomic selection, as previously described (Legarra *et al.*, 2008; Usai *et al.*, 2009; Verbyla *et al.*, 2010).

Predicted phenotype and accuracy of GS

During the validation step, each individual had its genomic estimated breeding values (GEBV) predicted through GS by multiplying the incidence matrix Z for the marker by the vector of estimated marker effects and summing the estimated general mean according to the expression:

$$\hat{y}_j = \hat{u} + \sum_i Z_{\mathrm{ij}}\hat{m}_i$$

The accuracy of GS (r_{GS}) to predict breeding values was calculated by the correlation of the vector with the GEBV of all the individuals in the population with their EBV. The accuracy at each validation group of the cross-validation was also calculated and the standard error of the accuracy was computed. These analyses were performed across all sites, traits and ages. The selection gain of genomic selection was compared with classical phenotypic selection considering a reduced breeding cycle as a result of early selection, as previously described by Grattapaglia & Resende (2010). Therefore, the selection response was obtained as the ratio between the selection accuracy and the time in years. For GS it is $\mathrm{SR}_{\mathrm{GS}} = \frac{r_{\mathrm{GS}}}{C_{\mathrm{GS}}}$ and for traditional BLUP based selection (TS) it is $\mathrm{SR}_{\mathrm{TS}} = \frac{r_{\mathrm{TS}}}{C_{\mathrm{TS}}}$ where (r_{GS}) and (r_{TS}) are the selection accuracies of GS and TS respectively, and (C_{TS}) and (C_{GS}) are the breeding cycle length for TS and GS, respectively. The ratio of these two selection responses gave the selection efficiency of GS over TS computed using the expression $R_{\mathrm{GS:TS}} = \frac{r_{\mathrm{GS}}C_{\mathrm{TS}}}{r_{\mathrm{TS}}C_{\mathrm{GS}}}$.

Validation of the models across ages and sites and estimation of type B genetic correlation

The GS models developed at each age were evaluated for their accuracy in predicting breeding values across ages. For this analysis, accuracy was calculated by the correlation of the GEBV derived from data collected at early ages, with the EBV at age 6 yr. As the same plant is compared across ages, there is a permanent dependency between a plant in two different ages. Therefore, a 10-fold cross validation was performed the same way as described previously. Next, GS models developed based on measurements made at age 6 yr in each site were validated across sites. For this validation, only 711 individuals that are consistently represented in all four sites were used. The model was estimated using the entire population of one site without jackknife cross-validation. The type B genetic correlation across sites and across ages was calculated as described previously (Yamada, 1962). For that analysis, the initial linear mixed model used to estimate the variance components and predict the individual breeding values was fitted by adding a fixed site/age effect and random interaction effects.

Results

Trait heritability

To assess the extent to which the phenotypic variation is genetically controlled and amenable to GS, we initially estimated the narrow-sense heritability (h^2) for HT and DBH. Heritabilities ranged from 0.09 to 0.32 for both traits (Table 1; see the Supporting Information, Table S1), consistent with previous reports for this population (Baltunis *et al.*, 2005, 2007).

High accuracy of genomic selection prediction models

The CCLONES population is clonally replicated on four sites in the Southeastern USA, for which there is phenotypic data from consecutive years (Palatka and Nassau in Florida, Cuthbert and B.F. Grant in Georgia). Initially, prediction models for GS were developed using phenotypic data measured in each site at year 6 – an age typically used to make early selections in pine breeding – to assess if there was significant variation in the estimated accuracies across sites. Accuracies at all four sites ranged from 0.65 to 0.75 for DBH and 0.63–0.74 for HT (Tables 1, S1).

Table 1 *Pinus taeda* diameter at breast height (DBH) and total height (HT) heritability, accuracy of genomic selection (GS) and standard error, based on data measured at age 6 yr in four sites

Trait	Site	Heritability	Accuracy of GS	Standard Error
DBH	B.F. Grant	0.23	0.73	0.04
	Cuthbert	0.22	0.72	0.05
	Nassau	0.32	0.65	0.07
	Palatka	0.21	0.68	0.04
HT	B.F. Grant	0.17	0.74	0.03
	Cuthbert	0.13	0.72	0.05
	Nassau	0.26	0.64	0.06
	Palatka	0.26	0.67	0.04

Data from all ages are described in the Supporting Information, Table S1.

Improved efficiency of genomic selection compared with BLUP-based phenotypic selection

To evaluate the performance of GS relative to traditional breeding methods, we estimated the accuracies of BLUP-based selection (Resende *et al.*, 2008) and used it as a benchmark for comparing accuracies obtained by GS (Grattapaglia & Resende, 2010). The increase in efficiency per unit of time in the selection response of GS was 53–92% higher for DBH, and 58–112% higher for HT, assuming a conservative reduction of 50% in the length of the breeding cycle (Table 2). The accuracy of GS was comparable to the accuracy of clone selection based on phenotypic BLUP predicted from clone trials at two sites (B.F. Grant and Cuthbert). At the other two sites (Nassau and Palatka) the accuracies of genomic selection were slightly lower. But the genetic gains per unit of time were higher for genomic selection at all sites.

Validation of GS prediction models across ages

In loblolly pine, phenotypic selection occurs at age 5–7 yr, when traits are considered sufficiently predictive of performance at

Table 2 Efficiency of genomic selection when compared with traditional phenotypic selection in *Pinus taeda*

Trait	Site	*h* (BLUP)	*h* (GS)	Efficiency	Increase relative to phenotypic selection (%)
DBH	B.F. Grant	0.79	0.73	1.85	85
	Cuthbert	0.75	0.72	1.92	92
	Nassau	0.85	0.65	1.53	53
	Palatka	0.81	0.68	1.68	68
HT	B.F. Grant	0.74	0.74	2.00	100
	Cuthbert	0.68	0.72	2.12	112
	Nassau	0.80	0.64	1.60	60
	Palatka	0.85	0.67	1.58	58

Accuracies (*h*) of traditional best linear unbiased predictor (BLUP) and genomic selection (GS) were estimated (Grattapaglia & Resende, 2010), and the efficiency calculated assuming a reduction in the length of the breeding cycle by 50%. DBH, diameter at breast height; HT, total height.

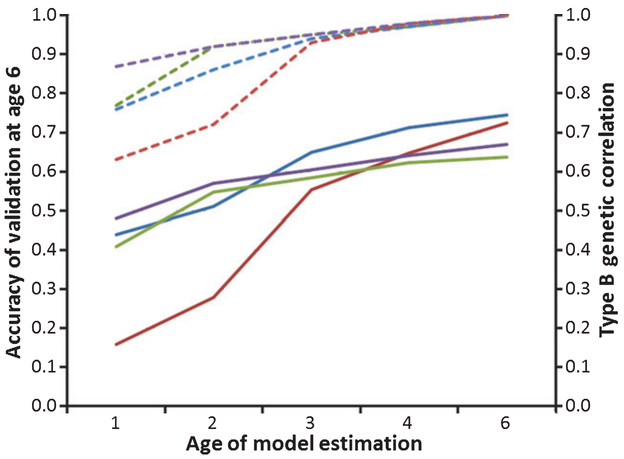

Fig. 1 Accuracy of total height (HT) prediction models estimated at ages 1–4 yr, and validated at age 6 yr for *Pinus taeda*, in B.F. Grant (blue line), Cuthbert (red line), Nassau (green line) and Palatka (purple line). Dotted lines represent the genetic correlation between HT measurements at ages 1–4 yr, and age 6 yr, in each site.

harvest. To evaluate if this delay is also necessary for development of prediction models, we assessed the accuracy of models developed for HT based on data collected at ages 1–4 yr, but validated with measurements from the same populations at age 6 yr. Genetic correlations between measurements made at these ages were also estimated. As expected, using the accuracy of models developed based on data collected at age 6 yr as the benchmark, we observed decreasing accuracies as measurements at younger ages were used (Fig. 1). Models developed based on HT data from year 1 were approximately only half as accurate in predicting HT at age 6 yr. These accuracies were particularly low in Cuthbert (0.16), reflecting the lowest genetic correlation between HT measurements made in year 1 and 6 (0.63) in this site (Fig. 1). By contrast, in Palatka the accuracy of models developed based on HT data from year 1 were far more predictive of year 6 measurements, where the highest genetic correlation (0.87) between the traits was observed (Fig. 1). Similar trends were observed when prediction models developed based on DBH measurements at ages 3 yr and 4 yr were validated at age 6 yr (Fig. S1).

GS prediction models have limited accuracy across sites

Next, we tested the suitability of using models estimated based on data from each individual site, in predicting phenotypes across different sites. Because CCLONES is clonally replicated along a North–South transect, the extent by which prediction models were accurate could be tested across highly distinct environmental conditions. As expected, the accuracies of models predicting GEBV were higher for the same site and declined for different sites (Table 3). The decrease in accuracy parallels the increase in geographic distance between the site for which models were estimated, and the site where they were validated. For example, for models estimated in B.F. Grant (Northern Georgia), and validated in Palatka (Central Florida), the accuracies were only 0.22 and 0.33 for DBH and HT, respectively. Conversely, models estimated from DBH and HT data collected in the most southern site, Palatka, were relatively accurate in Nassau (0.60 and 0.66, respectively). As observed in the analysis of stability of prediction models across ages, these results paralleled the genetic correlation estimated between sites (Figs 2, S2). Palatka and Nassau are located in regions with a similar climate, which likely

Table 3 Accuracy of genomic selection prediction models estimated and validated across sites for *Pinus taeda* (a, diameter at breast height (DBH); b, total height (HT))

	Validation			
Estimation	B.F. Grant	Cuthbert	Nassau	Palatka
(a)				
B.F. Grant	0.73	0.50	0.52	0.22
Cuthbert	0.49	0.72	0.48	0.34
Nassau	0.51	0.50	0.65	0.60
Palatka	0.18	0.32	0.60	0.68
(b)				
B.F. Grant	0.74	0.43	0.55	0.33
Cuthbert	0.41	0.72	0.38	0.24
Nassau	0.50	0.37	0.64	0.64
Palatka	0.29	0.23	0.66	0.67

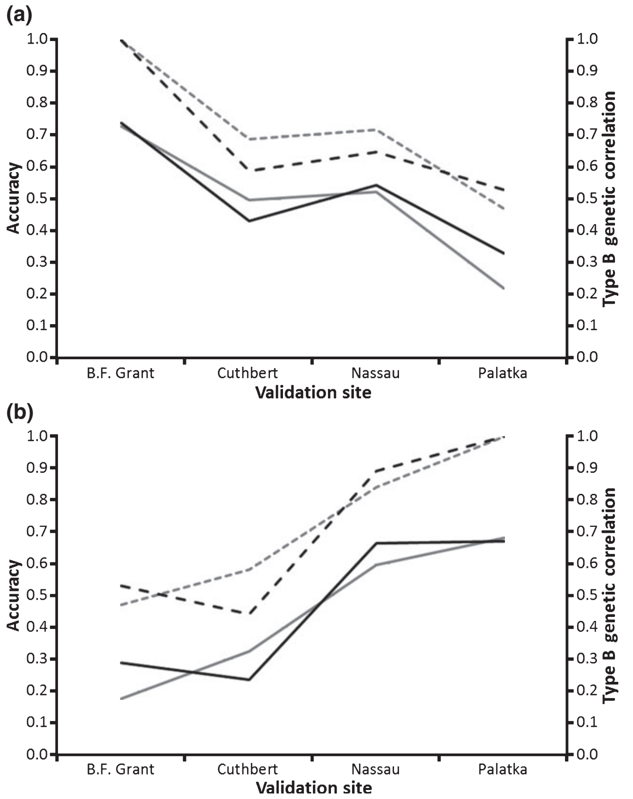

Fig. 2 Accuracy of prediction models estimated in B.F. Grant (a) and Palatka (b) and validated in the other sites, for *Pinus taeda* diameter at breast height (DBH, grey line) and total height (HT, black line). Dotted lines represent the genetic correlation between DBH and HT measurements made in B.F. Grant (a) and Palatka (b), relative to the other sites.

contributes to lower genotype × environment interactions, and higher accuracy of models in predicting GEBV across sites.

Discussion

We developed prediction models for genomic selection of growth properties in the conifer loblolly pine, and assessed their accuracy over distinct environments and ages. Prediction models were estimated based on a training population of *c.* 800 individuals derived from 32 parents, crossed in a circular mating design. Accuracies ranged from 0.63 (HT) to 0.75 (DBH), with a mean of 0.68, largely in agreement with the expectation for a training population with the genetic properties of CCLONES – that is, effective population size (N_e) ≈ 40, genotyped at a marker density of 3 SNP per cM (Grattapaglia & Resende, 2010). Under these conditions, and for traits with low to moderate heritabilities that are regulated by a large number (50+) of loci of small effect, accuracies in the range of 0.55 and 0.80 would be expected.

Accuracy of genomic selection prediction models over time

For DBH and HT, measurements were obtained over multiple years, allowing for the development of separate prediction models for each year. For both traits, estimated accuracies were similar across the years of measurement (Table S1). We also tested how models developed at an early age perform in predicting phenotypes later in the life-cycle – accelerating model estimation is beneficial because the sooner models that accurately predict phenotypes at rotation age can be developed, the faster genomic selection can be adopted. However, models developed for DBH and HT early in the rotation (age 1 yr and 2 yr) had limited accuracy in predicting phenotypes at age 6 yr (Fig. 1). Thus, phenotyping of the training population at mid-rotation or later, as done in traditional tree breeding (White *et al.*, 2007), seems necessary for accurate prediction of growth performance. These results likely reflect the significant physiological changes that occur as conifers transition from the juvenile to mature stage, which translate into noticeable differences in growth rates and other properties. Maker-trait associations detected in QTL studies have been previously shown to be largely unstable over time (Emebiri *et al.*, 1998; Kaya *et al.*, 1999; Lerceteau *et al.*, 2001) in tree species, supporting the conclusion that changes in the genetic control of growth and development negatively affect early model development.

Accuracy of prediction models within and between breeding zones

Loblolly pine breeding populations are typically established to support genetic improvement over broad geographic regions, or 'breeding zones'. Therefore, from a tree improvement standpoint, it is critical that prediction models be accurate across sites, at least within a breeding zone. CCLONES is biologically replicated on four sites spanning multiple breeding zones in a north-south transect, from central Florida (Palatka site, latitude 29.65° N) to northwest Georgia (B.F. Grant site, latitude 33.74° N). Palatka (FL) and Nassau (FL) are in the same breeding zone (Florida), while Cuthbert and B.F. Grant are in the Upper Coastal Plain and Piedmont zones, respectively. As expected from the estimated genetic correlation, the comparison of accuracies derived by estimating models in one Florida site, and validated in the other, indicates a small loss in predictability. The accuracy of prediction models for HT and DBH (age 6 yr) were 0.64–0.68 in Palatka and Nassau, decreasing marginally (≤ 0.08) when validated in the reciprocal site (Table 3). However, much lower accuracies (0.18–0.32) were observed when models estimated in the most southern Florida population (Palatka) were validated based on data from sites in the Upper Coastal Plain (Cuthbert) and Piedmont (B.F. Grant), suggesting that environment × genotype interactions severely affect the transferability of models across breeding zones. Therefore, we anticipate that prediction models will be primarily applicable within breeding zones, and new training population will have to be established if genomic selection is to be applied to different zones. Nonetheless, the robust relationship between genetic correlations and the stability of prediction models across sites provides a valuable indication to breeders as to when these models may be widely applicable.

Implementing genomic selection in conifer breeding programs

In conifers, the time-frame between the beginning of a breeding cycle and the production of improved seeds for commercial plantations can span multiple decades, divided essentially into three phases: (1) breeding, (2) testing and (3) propagation. For loblolly pine, the long period (8+ yr) required for plants to become sexually mature can be drastically reduced by top-grafting, a strategy used to stimulate male and female strobili production from seedlings as young as 1 yr. As a consequence, the breeding phase may be completed in < 4 yr, between top grafting, pollination and seed production for field testing in the second phase. Similarly, the last phase of propagation can be reduced to one or a few years by implementing advanced methods of clonal propagation, such as somatic embryogenesis or rooted cuttings. Therefore, testing has remained the most time-consuming phase in genetic improvement of loblolly pine, lasting typically for 6–10 yr in loblolly pine. Incorporating genomic selection could dramatically reduce the time required for completion of a cycle of genetic improvement by eliminating the testing phase, significantly accelerating the genetic gain relative to traditional breeding. One can envision a scenario where every *c.* 4 yr selected individuals are top-grafted, crossed and seeds are produced (Fig. 3). Seedlings are then genotyped and selected for top-grafting, initiating the next breeding cycle. In this scenario, rapid pyramiding of favorable alleles can be pursued by establishing crosses that create the best allelic complement across quantitative trait loci throughout the genome. In parallel, selected seedlings can be clonally replicated and established in clonal trials to verify their performance relative to elite material. Because breakdown of the accuracy of GS is known to occur across generations, the application of this methodology will require monitoring of the accuracy of prediction models and their recalibration, as necessary.

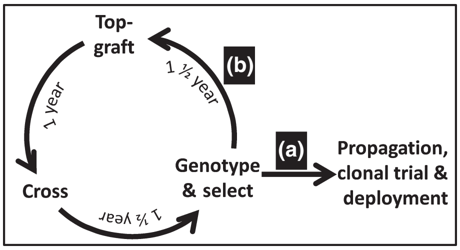

Fig. 3 Pine breeding with genomic selection. Assuming the exclusion of a testing phase aimed at selecting the best individuals (replaced by genomic selection), and the use of clonal propagation to multiply selected genotypes, a single pine breeding cycle could be reduced to *c.* 4 yr. At the end of each cycle, selected individuals can be (a) clonally propagated for clonal trials and deployment and/or (b) used in a new breeding cycle.

An alternative and even more aggressive use of GS in loblolly pine breeding would be to apply it in combination with somatic embryogenesis (SE), to deploy clonally propagated elite genotypes in a similar way as cattle breeders are now advancing the use of GS in combination with reproductive technologies (Humblot *et al.*, 2010). In this scenario, elite parents with high specific combining ability, determined by conventional quantitative genetics methods or by GS-based prediction, can be crossed and large quantities of seeds produced. As somatic embryogenesis is initiated, genotyping can be carried out to eliminate zygotic embryos predicted to perform poorly based on their GBV. Furthermore, as cultures of zygotic embryos with highest GBV are identified, more effort can be dedicated to optimize the methods to culture somatic embryos of the best performers. A set of predicted elite genotypes are immediately propagated and established in clonal trials or even deployed commercially depending on the acceptable level of risk. Combining GS with SE offers several advantages compared with the current implementation of clonal propagation in loblolly pine, notably: a significant reduction in the cost associated with maturing SE cell lines and establishing costly field trials; a remarkable increase in the selection intensity applied to the initial seed population derived from elite crosses, allowing a better capture of additive and nonadditive effects; and the elimination of large-scale, long clonal testing trials, if elite genotypes can be confidently selected at the SE cell line stage, and immediately propagated at a commercial scale.

Integration of genome-wide association data and genomic selection in pine breeding

Identification of quantitative trait nucleotides in genome-wide association studies will be a daunting task in conifers because of technical challenges associated with the need to genotype exceedingly large numbers of polymorphisms, and establishing sufficiently large populations to identify genetic variants with small effects. Despite these difficulties, progress has been made in the last decade, and a few loci significantly associated with traits have been discovered, although invariably controlling small proportions of the total phenotypic variation (Gonzalez-Martinez *et al.*, 2007, 2008; Eckert *et al.*, 2010). It can be anticipated that over the next decade the genetic basis of complex trait variation will be gradually uncovered in loblolly pine and other conifers, explaining an increasingly greater portion of the phenotypic variance. How can the existing and future association genetic studies in loblolly pine and other conifers be integrated with breeding programs relying on genomic selection? We anticipate that, as causative quantitative trait polymorphisms are discovered, and their effects on phenotypes are adequately assessed in the relevant environments and ages, they will be incorporated into GS training models, further improving their accuracies. Ultimately, one can anticipate that models will be solely based on causative polymorphisms. However, until a detailed description of the genetic architecture of the most economically important traits becomes available, strategies such as genomic selection are likely to remain valuable for integrating molecular markers into breeding and selection in traditional conifer genetic improvement programs.

Acknowledgements

The authors thank members of the Forest Biology Research Cooperative (FBRC) for their support in establishing, maintaining and measuring field trials used in this study. The work was supported by the National Science Foundation Plant Genome Research Program (award no. 0501763 to GFP, JMD and MK) and the US Department of Agriculture National Institute of Food and Agriculture Plant Breeding and Education Program (award no. 2010-85117-20569 to MK, GFP and JMD). We also acknowledge the valuable input from three anonymous reviewers.

References

Baltunis BS, Huber DA, White TL, Goldfarb B, Stelzer HE. 2005. Genetic effects of rooting loblolly pine stem cuttings from a partial diallel mating design. *Canadian Journal of Forest Research* **35**: 1098–1108.

Baltunis BS, Huber DA, White TL, Goldfarb B, Stelzer HE. 2007. Genetic analysis of early field growth of loblolly pine clones and seedlings from the same full-sib families. *Canadian Journal of Forest Research* **37**: 195–205.

Bernardo R, Yu J. 2007. Prospects for genomewide selection for quantitative traits in maize. *Crop Science* **47**: 1082–1090.

Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC *et al.* 2009. The genetic architecture of maize flowering time. *Science* **325**: 714–718.

Daetwyler HD, Hickey JM, Henshall JM, Dominik S, Gredler B, van der Werf JHJ, Hayes BJ. 2010. Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. *Animal Production Science* **50**: 1004–1010.

Dekkers JCM. 2004. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. *Journal of Animal Science* **82** (E-Suppl): E313–E328.

Eckert AJ, van Heerwaarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, Gonzalez-Martinez SC, Neale DB. 2010. Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics* **185**: 969–982.

Emebiri LC, Devey ME, Matheson AC, Slee MU. 1998. Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theoretical and Applied Genetics* **97**: 1053–1061.

FAO. 2002. *World agriculture: towards 2015/2030. Summary Report.* Rome, Italy: Food and Agriculture Organization of the United Nations.

Garrick DJ, Taylor JF, Fernando RL. 2009. Deregressing estimated breeding values and weighting information for genomic regression analyses. *Genetics Selection Evolution* **41**: 55.

Gianola D, Campos G, Hill WG, Manfredi E, Fernando R. 2009. Additive genetic variability and the Bayesian alphabet. *Genetics* **183**: 347–363.

Gilmour AR, Cullis BR, Harding SA, Thompson R. 2006. *ASRemlUpdate: what's new in Release 2.00.* Hemel Hempstead, UK: VSN International Ltd.

Gonzalez-Martinez SC, Huber D, Ersoz E, Davis JM, Neale DB. 2008. Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity* **101**: 19–26.

Gonzalez-Martinez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB. 2007. Association genetics in *Pinus taeda* L. I. Wood property traits. *Genetics* **175**: 399–409.

Grattapaglia D, Kirst M. 2008. Eucalyptus applied genomics: from gene sequences to breeding tools. *New Phytologist* **179**: 911–929.

Grattapaglia D, Resende M. 2010. Genomic selection in forest tree breeding. *Tree Genetics and Genomes* **7**: 241–255.

Hayes BJ, Bowman PJ, Chamberlain AC, Verbyla K, Goddard ME. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution* **41**: 51.

Hayes B, Goddard M. 2010. Genome-wide association and genomic selection in animal breeding. *Genome* **53**: 876–883.

Heffner EL, Sorrells ME, Jannink JL. 2009. Genomic selection for crop improvement. *Crop Science* **49**: 1–12.

Humblot P, Le Bourhis D, Fritz S, Colleau JJ, Gonzalez C, Guyader Joly C, Malafosse A, Heyman Y, Amigues Y, Tissier M *et al.* 2010. Reproductive technologies and genomic selection in cattle. *Veterinary Medicine International* **2010**: 192787.

IPCC. 2007. Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, eds. *Climate change 2007: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change.* Cambridge, UK & New York, NY, USA: Cambridge University Press.

Iwata H, Hayashi T, Tsumura Y. 2011. Prospects for genomic selection in conifer breeding: a simulation study of *Cryptomeria japonica*. *Tree Genetics and Genomes* **7**: 1–12.

Kaya Z, Sewell MM, Neale DB. 1999. Identification of quantitative trait loci influencing annual height- and diameter-increment growth in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics* **98**: 586–592.

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R. 2009. A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetics* **5**: e1000551.

Lande R, Thompson R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* **124**: 743–756.

Legarra A, Robert-Granie C, Manfredi E, Elsen JM. 2008. Performance of genomic selection in mice. *Genetics* **180**: 611–618.

Lerceteau E, Szmidt AE, Andersson B. 2001. Detection of quantitative trait loci in *Pinus sylvestris* L. across years. *Euphytica* **121**: 117–122.

Long AD, Langley CH. 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Research* **9**: 720–731.

Meuwissen TH, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**: 1819–1829.

Muir WM. 2007. Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *Journal of Animal Breeding and Genetics* **124**: 342–355.

Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ. 2003. QTL analysis of genotype × environment interactions affecting cotton fiber quality. *Theoretical and Applied Genetics* **106**: 384–396.

Paterson AH, Tanksley SD, Sorrells ME. 1991. DNA markers in plant improvement. *Advances in Agronomy* **46**: 39–90.

Resende MDV, Lopes PS, Silva RL, Pires IL. 2008. Seleção genômica ampla (GWS) e maximização da eficiência do melhoramento genético. *Pesquisa Florestal Brasileira* **56**: 63–77.

Resende MDV, Resende MFRJ, Aguiar AM, Abad JIM, Missiaggia AA, Sansaloni CP, Petroli CD, Grattapaglia D. 2010. Computação da seleção genômica ampla (GWS). *Documentos EMBRAPA 210 – ISSN 1679-2599.* Colombo, Brazil: EMBRAPA Florestas.

Strauss SH, Lande R, Namkoong G. 1992. Limitations of molecular-marker-aided selection in forest tree breeding. *Canadian Journal of Forest Research* **22**: 1050–1061.

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics* **43**: 159–162.

Usai MG, Goddard ME, Hayes BJ. 2009. Lasso with cross-validation for genomic selection. *Genetics Research* **91**: 427–436.

Verbyla KL, Calus MPL, Mulder HA, de Haas Y, Veerkamp RF. 2010. Predicting energy balance for dairy cows using high-density single nucleotide polymorphism information. *Journal of Dairy Science* **93**: 2757–2764.

Visscher PM. 2008. Sizing up human height variation. *Nature Genetics* **40**: 489–490.

Wear DN, Carter DR, Prestemon J. 2007. *The U.S. South's timber sector in 2005: a prospective analysis of recent change.* Asheville, NC, USA: USDA Forest Service, Southern Research Station. General Technical Report No. SRS-99.

White TL, Adams WT, Neale DB. 2007. *Forest genetics.* Wallingford, UK and Cambridge, MA, USA: CABI Publishing.

White TL, Carson M. 2004. Breeding program of conifers. In: Walter C, Carson M, eds. *Plantation forest biotechnology for the 21st century.* Kerala, India: Research Signpost, 61–85.

Williams ER, Matheson AC, Harwood CE. 2002. *Experimental design and analysis for tree improvement, 2nd edn.* Collingwood, Vic., Australia: CSIRO Publishing.

Yamada Y. 1962. Genotype by environment interaction and genetic correlation of the same trait under different environments. *Japan Journal of Genetics* **37**: 498–509.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 Accuracy of diameter at breast height (DBH) prediction models estimated at ages 3 yr and 4 yr, and validated at age 6 yr, in B.F. Grant, Cuthbert, Nassau and Palatka.

Fig. S2 Accuracy of prediction models estimated in Cuthbert and Nassau, and validated in the other sites, for diameter at breast height (DBH) and total height (HT).

Table S1 Diameter at breast height and total height heritability, accuracy of genomic selection and standard error, based on data measured across all sites and ages

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.