

# Clonal selection of *Eucalyptus grandis* x *Eucalyptus globulus* for productivity, adaptability, and stability, using SNP markers

D. Torres-Dini<sup>1,2</sup>, A. C. P. Nunes<sup>3</sup>, A. Aguiar<sup>2,4</sup>, N. Nikichuk<sup>1</sup>, C. Centurión<sup>4,5</sup>, M. Cabrera<sup>5</sup>, M. L. T. Moraes<sup>1,2</sup>, M. D. V. Resende<sup>4</sup>; and A. M. Sebbenn<sup>2,6\*</sup>

<sup>1</sup> Instituto Nacional de Investigación Agropecuaria, Ruta 5 Km 386, CEP 45000, Tacuarembó, TB, Uruguay.

<sup>2</sup> Faculdade de Engenharia de Ilha Solteira/UNESP, CP 31, CEP 15385-000, Ilha Solteira, SP, Brazil.

<sup>3</sup> Departamento de Engenharia Florestal, Universidade Federal de Viçosa, Avenida Peter Henry Rolfs, s/n – Campus Universitário, Centro, 36570-000, Viçosa, MG, Brazil.

<sup>4</sup> EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Florestas, Estrada da Ribeira, Km 111 - Bairro Guaraituba, 83411-000, Colombo, PR, Brazil.

<sup>5</sup> UPM, Forestal Oriental, Paysandú, Uruguay.

<sup>6</sup> Instituto Florestal de São Paulo, CP 1322, São Paulo, SP, 01059-970, Brazil.

(\*) Corresponding author: A.M. Sebbenn, Phone: +0055 19 34351681, E-mail: alexandresebbenn@yahoo.com.br

## Abstract

In Uruguay, reforestations with *Eucalyptus* sp. is of fundamental importance to supply paper, pulp, and wood production. This study investigates genetic, productivity, stability, and adaptability parameters in the selection of *Eucalyptus grandis* x *Eucalyptus globulus* full-sib hybrid clones. The study was conducted in a clonal test, repeated in two different soils types, in Rio Negro State, Uruguay. The population was characterized phenotypically for diameter at breast height (dbh) and genotyped for SNP markers (EuCHIP60K chip). Mean dbh was similar between sites and the genotype–environment interaction was simple. We found high genotype correlation in clone performance between environments (0.708), indicating the possibility of selecting the same clones for both study locations. Mean heritability between clones (0.724), coefficient of individual genetic variation (10.9 %), and relative variation (0.916), suggest the possibility of gains (estimated at 3.1 % for both sites together) by selecting clones with higher growth rates. A total of 15,196 SNPs were used to confirm parentage and test a genomic selection model for dbh. The predictive capacity was negative (-0.15) given the small population size (78 individuals). The most adaptable material among the tested study sites presented higher values for SNP heterozygosity. Thus, using molecular markers to identify clones responsive to environmental changes can act as a powerful tool in *Eucalyptus* breeding programs. The hybrid population showed greater adaptability than *E. globulus* for this region.

**Keywords:** *Eucalyptus* hybrid, forest breeding, gene markers; population genetics, quantitative genetics

## Introduction

Worldwide, reforestations with *Eucalyptus* sp. exceed 20 million hectares, distributed across more than 90 countries with a vast range of climates. South America produces 55 % of global production, with Brazil, Uruguay, Chile, Argentina, and Peru as the main players. Together, Asia, China, and India account for 23 % of global production, while Europe produces 7 % of the total, including Spain, Portugal, and to a lesser extent Italy. On the African continent, South Africa is the main producer with 3 % of total global production, and the remaining production is shared among countries around world with levels of less than 3 % (Resquin and Balmelli, 1999; Griffin et al., 2000; Balmelli and Resquin, 2008; Booth, 2013; Paseyro, 2015).

In general, *Eucalyptus* commercial plantations have used hybrid clones due to their productive potential. The creation of new genetic combinations to exploit heterosis effects in interspecific crosses is a common objective in *Eucalyptus* breeding programs (Griffin et al., 2000; Grattapaglia and Kirst, 2008). *Eucalyptus grandis* is known for its adaptability and stability in different environments, while *E. globulus* is renowned for its high-quality wood pulp. However, in Uruguay, the cultivation of *E. globulus* is restricted to the South-eastern region as it exhibits low adaptability, susceptibility to disease, and poor growth in other areas of the country (Balmelli and Resquin, 2005).

To select the best clones, it is necessary to evaluate several promising clones in different environments before making a final recommendation and subsequent multiplication (Rosado et al., 2012). In most cases, interactions between genotype and environment (GxE) affect genetic gains, thus requiring the estimation of the magnitude and nature of this interaction. Although there are many methods to estimate this effect, simple GxE analysis does not provide complete and accurate

information (Resende et al., 2012). Selections performed using the mixed model of the harmonic mean of relative performance of genotypic value (HMRPGV) is advantageous because it classifies genotypes simultaneously by productivity, stability, and adaptability in different locations (Maia et al., 2009; Resende et al., 2012). HMRPGV has been employed to evaluate a range of different crops, including *E. grandis* (Pinto Junior et al., 2006), cashew (Maia et al., 2009), *Eucalyptus* sp. (Rosado et al., 2012), rice (Colombari Filho et al., 2013), *Eucalyptus urophylla* (Pupin et al., 2015), and cotton (Farias et al., 2016).

Another tool that predicts genetic parameters with increased accuracy is genomic selection or Genomic Wide Prediction (GWP, Desta and Ortiz, 2014), a methodology that integrates quantitative genetics with new genotyping technologies, such as DArT (Diversity Arrays Technology) and SNP (Single Nucleotide Polymorphisms) (Aguar et al., 2015; Silva-Junior et al., 2015). These high-density gene markers provide the theoretical basis for the GWP approach (Meuwissen et al., 2001) and can be applied to accelerate breeding cycles, increase the rate of genetic gain per unit of time, and decrease the costs associated with selecting superior genotypes (Heffner et al., 2010). Selection based on GWP focuses on the simultaneous prediction of the genetic effects of thousands of dispersed markers in the genome, making it possible to infer large to small effects on traits for most loci, thus explaining almost all quantitative traits (Meuwissen et al., 2001; Isik, 2014). In addition, these molecular markers support a wide range of applications, including the study of high resolution genetic maps (Bartholomé et al., 2015), molecular diversity among species (Hudson et al., 2015), QTL detection (Resende et al., 2017), and inbreeding depression (Hedrick et al., 2016).

The aims of this study were to select the best *E. grandis* x *E. globulus* clones for increased wood productivity for commercial reforestation at two sites in Uruguay. Specifically, we aim to estimate genetic parameters for diameter at breast height (dbh) to: i) select hybrid clones based on productivity, stability, and adaptability; ii) transfer the adaptability and stability capabilities of *E. grandis* to a population of *E. grandis* x *E. globulus* hybrids; iii) evaluate the predictive response of the GWP model and compare the results with the REML/BLUP method; and iv) use SNP marker information to confirm relatedness among clones.

## Material and Methods

### Clone production and testing

The crossbreeding strategy began with controlled pollination between an *E. grandis* mother, denominated as G50 (CSIR, South Africa), and mixed pollen collected from ten *E. globulus* trees (Figure 1). This crossing produced a F1 single-progeny compound with mixtures of half- and full-sibs. The best tree of this F1 progeny for dbh growth was cloned and identified as BBT01058 (Forestal Oriental SA, Uruguay). This clone was used as the father (pollen donor) in a pseudo-backcrossing with *E. grandis* clone 678.2.1 (Forestal Oriental SA, Uruguay), producing an F2 full-sib single progeny, denominated 2162 (Myburg et al., 2003). Approximately 6,637 seeds were obtained in the

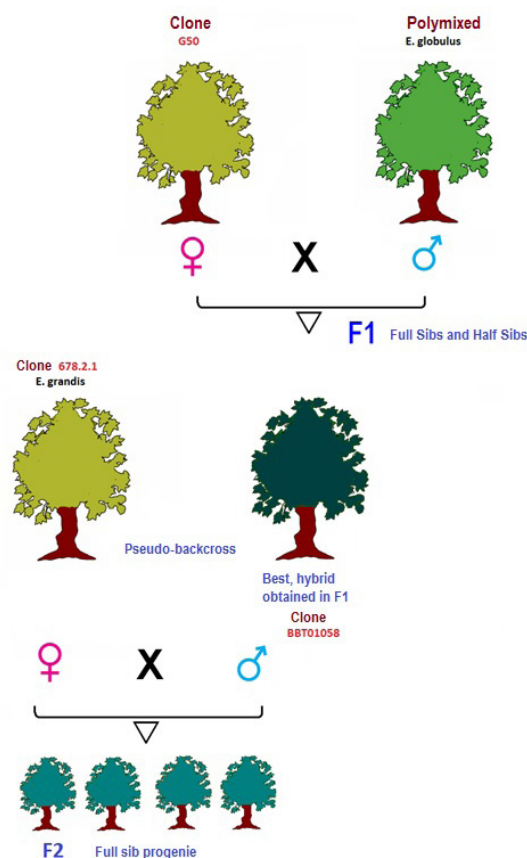


Figure 1  
Representation of the crossing scheme used to produce inter-specific clones

F2 progeny, but only 292 seeds (4.4 %) germinated. The seedlings were selected at the nursery stage, with weak and abnormal phenotypes discarded, resulting in only 99 seedlings that were cloned by micro-cuttings (Griffin et al., 2000). The 99 selected clones were used to establish a clonal test in two sites with contrasting soils of the company Forestal Oriental Uruguay, in Rio Negro State, Uruguay: i) Tres Bocas (-32.7881891 S, -57.8865168 W); and ii) Algorta (-32.423335 S, -57.4606389 W). While the soils of Algorta have better drainage than Tres Bocas, neither of the two sites are suitable for growing *E. globulus*. The clonal tests were established in January 2000, using an alpha lattice experimental design, with 99 treatments (clones), five replications (blocks), and one plant per plot, at a spacing of 3 x 2 m. The *E. grandis* clone 3523 (Forestal Oriental SA, Uruguay) was used as the control genotype. Clones were measured for dbh at 48 months after planting.

### Determination of genetic parameters

Deviance analysis to detect differences between the treatments, estimate of variance components (REML), and predicted clone genetic values (BLUP) were carried out for both sites using mixed mating model 52 in the SELEGEM-REML/BLUP software (Resende, 2016), for clonal tests with single-plant plots:  $y = Xr + Zg + Hb + Wge + e$ , where  $y$  is the vector of

dbh trait,  $r$  is the vector of fixed effects (general mean and experimental effects),  $g$  is the vector of genotypic effects (assumed to be random),  $b$  is the vector of block effects (random),  $ge$  is the vector of GxE interaction effects (random),  $e$  is the vector of errors (random), and  $X$ ,  $Z$ ,  $H$ , and  $W$  are the incidence matrices for the above mentioned effects. From this analysis, the following variance components were estimated: genotypic variance among clones ( $\sigma_g^2$ ), variance among blocks ( $\sigma_b^2$ ), variance of GxE interaction ( $\sigma_{ge}^2$ ), and environmental variance among plots ( $\sigma_e^2$ ). The individual phenotypic variance was estimated as  $\sigma_f^2 = \sigma_g^2 + \sigma_b^2 + \sigma_{ge}^2 + \sigma_e^2$ . The estimated parameters from the variance components were:

i) correlation due to common environment of the block:

$$C_b^2 = \frac{\sigma_b^2}{\sigma_f^2};$$

ii) coefficient of the genotype interaction effects:

$$C_{ge}^2 = \frac{\sigma_{ge}^2}{\sigma_f^2}$$

iii) genotypic correlation of clones within site:

$$r_{gloc} = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ge}^2}$$

iv) broad-sense individual heritability:

$$H_g^2 = \frac{\sigma_g^2}{\sigma_f^2}$$

v) mean site heritability between clones, assuming complete survival:

$$H_c^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_e^2}{r}}$$

vi) accuracy of clone selection, assuming complete survival:

$$r_c = \sqrt{H_c^2}$$

vii) coefficient of individual additive genetic variation corrected to block effects:

$$CV_{gi} = 100\sqrt{\frac{\sigma_g^2}{m}}$$

viii) coefficient of experimental variation:

$$CV_e = 100\sqrt{\frac{\sigma_e^2}{m}}$$

ix) coefficient of relative variation:

$$CV_r = \frac{CV_{gi}}{CV_e}$$

where  $r$  is the number of repetitions and  $m$  is the average dbh.

Predicted genetic gains through selection were estimated based on high selection intensity (10%), considering only ten of the 99 clones tested in both sites that presented the largest dbh, using SELEGEN-REML/BLUP software (Resende, 2016). The selection was based on the predicted HMRPGV method for three strategies: i) selection considering mean dbh performance in the two sites, without interaction effect; ii) selection considering dbh performance of the genotypes in each site, with average interaction effect (HMRPGV); iii) simultaneous selection for production, considering stability and adaptability (HMRPGVxGM: HMRPGV multiplied by the general mean in all environments).

### SNP genotyping

Leaf tissue samples were collected and 500 ng of DNA from each sample was purified using a standard CTAB DNA extraction method. The DNA samples were genotyped using Illumina Infinium technology with the Eucalyptus EuCHIP60K chip, including a total of 64,639 markers (Silva-Junior et al., 2015). To select SNPs, a 95% call rate was used, with a 5% Minor Allele Frequency (MAF). In addition, the Hardy-Weinberg equilibrium test was performed and loci with deviations were excluded. SNP markers were used to plot allele effects, producing a dendrogram based on both heterogeneity and genetic similarity among genotypes, which was associated with the top ten clones. In addition, a HeatMap was constructed to schematically represent the parentage coefficient (Lima, 2014).

### Predictive model for GWP

The predictive model used for the GWP analysis was the random regression-best linear unbiased prediction (rrBLUP, Meuwissen et al., 2001), which considers marker genotypes as random effect covariates. The following linear mixed model was used to estimate the marker effects:  $y = Xb + Zm + e$ , where  $y$  is the vector of phenotypic observations;  $b$  is the vector of fixed effects;  $m$  is the vector of random effects of SNP markers;  $e$  refers to the random vector of residuals; and  $X$  and  $Z$  are the incidence matrices for  $b$  and  $m$ , respectively (Resende et al., 2012). This model was cross validated with the leave-one-out technique, in which one individual is taken in turn for prediction and validation. In each cycle, the remaining N-1 individuals are used to estimate the parameters of the model including the SNP effects. This process was repeated N times, using a different set of individuals for estimation and a different individual for validation each time until all individuals had their phenotypes predicted. Predicted phenotypes were given as  $Zm$ . The

correlation between predicted and observed phenotypes provided the predictive ability of GWP.

## Results

### Clone growth, environment control, and GxE

At 48 months after establishing the trials, 78 of the 99 hybrid clones survived (78.8 %). No significant differences were observed between sites for mean dbh through deviance analysis (chi-square = 2.71,  $P > 0.05$ ). The coefficient of determination of the block effects ( $C_b^2$ ) was very low (0.043), indicating environmental homogeneity within blocks and that the alpha lattice experimental design was efficiently controlled (Table 1). The GxE was also low ( $C_{ge}^2 = 0.151$ ) and the correlation of clone growth between the two sites was high ( $r_{gloc} = 0.708$ ), showing a simple GxE interaction.

Table 1

Estimates of genetic parameters for diameter at breast height (dbh) for both environments

Parameter	Estimate
Mean dbh at Tres bocas (cm)	16.91
Mean dbh at Algorta (cm)	16.64
General mean dbh (cm)	16.76
Coefficient of determination of block effects: $C_b^2$	0.043
Determination coefficient of the genotype interaction effects: $C_{ge}^2$	0.151
Correlation between clone ranking and environments: $r_{gloc}$	0.708
Broad-sense individual heritability: $H_g^2$	$0.367 \pm 0.083$
Adjusted heritability for mean clones: $H_c^2$	0.724
Accuracy of genotype selection: $r_c$	0.851
Coefficient of genotypic variation: $CV_{gt}$ (%)	10.9
Residual variation coefficient: $CV_e$ (%)	11.9
Relative variation coefficient: $CV_r$	0.916

### Heritability and coefficients of variation

The mean site broad-sense individual heritability ( $H_g^2 = 0.367$ ) and clone heritability ( $H_c^2 = 0.724$ ) indicate an extensive degree of genetic control over dbh and the possibility to obtain genetic gains from clone selection (Table 1). The selective accuracy for mean clones ( $r_c$ ) was equally high (0.851), indicating a strong association between true genotypic value and measured phenotype, further supporting a highly favourable scenario for selection. The individual coefficients of genetic variation ( $CV_{gt} = 10.9\%$ ) and experimental variation ( $CV_e = 11.9\%$ ) were similar, which resulted in a high coefficient of relative variation ( $CV_r = 0.916$ ). This also indicates the possibility for genetic gains by selecting clones with higher dbh growth.

### Clone selection and clustering analysis

The mean clone performance ranked higher than control clone 3523 in Tres Bocas, Algorta, and both sites together for three, nine, and six clones, respectively (Table 2). Clone 1744 was

ranked first for each site and both sites together. In the simultaneous selection for productivity, stability, and adaptability using HMRPGV, only clones 1744 and 1756 ranked higher than the control clone. Clustering based on heterozygosity was associated with those clones that were ranked in the top ten for each site, both sites together, and HMRPGV (Figure 3, Table 2). According to these results, the clones that show greater productivity and general adaptation can be selected, as well as some clones that present improved performance in specific environments.

### SNPs

The 78 surviving clones were used for DNA extraction and subsequent analysis of Genomic Wide Prediction (GWP). After quality control with MAF, the number of SNPs was reduced from 64,639 to 15,196 (23.5 %), which were then used in the GWP analysis for the dbh trait. Clones 1778, 1770, 1744, and 1756 were grouped together as part of the same node, supporting the fact that the observed phenotypes with the best traits share similar genotyping (Figure 3). On the other hand, clone 1796 is completely outside the main group. The predictive capacity was expected to be low ( $-0.15$ ) with no precision to predict the dbh trait (Table 3). Relatedness was confirmed by HeatMap analysis (Figure 2). In this analysis we observed that clones 1751 and 1758 are from the same genotype.

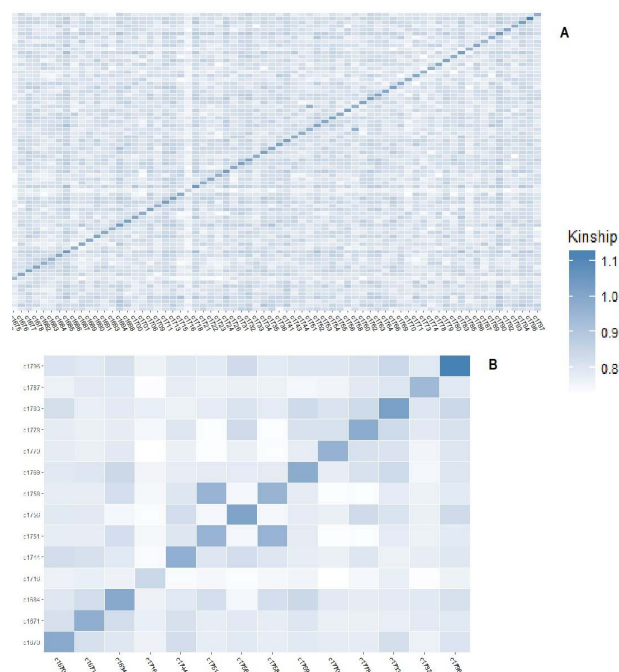


Figure 2

Estimate of pairwise kinship coefficient (HeatMap) represented by variation in color. Strong blue indicates a higher degree of relatedness. As these individuals are full siblings, the degree of kinship is expected to be close to 0.5. A: represents the entire population; B: the ten selected individuals

Table 2

Estimates of genetic values and gain (G) predicted for dbh for the ten best clones in Tres Bocas, Algorta, and both sites together (combined); genetic value for productivity, stability, and adaptability (HMPRGV and HMPRGVxGM) and genomic wide prediction (GWP) value for both environments. The control genotype is *Eucalyptus grandis* clone number 3523.

Ranking	Tres Bocas			Algorta			Combined							
	Clone	$u + g$	G (%)	Clone	$u + g$	G (%)	Clone	$u + g + ge$	G (%)	Clone	MHPRGV	MHPRGVxGM	Clone	GWP
1	1744	20.29	3.4	1744	20.75	4.1	1744	20.51	3.1	1744	1.22	20.46	1701	15.25
2	1670	20.18	3.3	1796	20.06	3.8	1769	19.81	2.8	1756	1.18	19.80	1794	14.57
3	1778	20.00	3.2	1769	19.96	3.6	1670	19.70	2.7	3523	1.17	19.69	1689	14.53
4	3523	19.93	3.2	1756	19.34	3.4	1796	19.64	2.6	1769	1.17	19.63	1676	14.48
5	1769	19.69	3.1	1670	19.25	3.3	1778	19.49	2.5	1670	1.16	19.48	1793	14.45
6	1756	19.65	3.0	1787	19.14	3.1	1756	19.48	2.5	1722	1.16	19.47	1752	14.45
7	1796	19.26	2.9	1716	19.05	3.0	3523	19.43	2.5	1684	1.16	19.42	1780	14.43
8	1684	19.19	2.9	1778	19.00	2.9	1783	19.02	2.4	1770	1.13	19.01	1778	14.42
9	1783	19.18	2.8	1770	18.99	2.9	1716	19.00	2.3	1778	1.13	18.89	1783	14.42
10	1671	19.12	2.7	3523	18.96	2.8	1770	18.88	3.7	1716	1.13	18.88	1755	14.40

$u + g$  is the predicted genotypic value;  $u + g + ge$  is the mean genotypic value in the environments; HMPRGV is the harmonic mean relative performance of genotypic value; HMPRGVxGM is the HMPRGV multiplied by the general mean in all environments

## Discussion

### Clone growth and GxE

The hybrid clones showed greater rates of survival (78.8 %) than was observed for the original *E. globulus* population (20.4–60.6 %) in Rio Negro State (Balmelli and Resquin, 2005). These results confirm that the hybrids show improved development and survival in that region of the country in comparison to the parental *E. globulus* population. In Brazil, the main component of *E. urograndis* hybrid clones is *E. urophylla*, which have enabled the expansion of plantations into previously unsuitable areas (Pupin et al., 2015).

Mean dbh was similar between Tres Bocas (16.91 cm) and Algorta (16.64 cm), but lower than the mean *E. grandis* control clone (3523) both in Tres Bocas (18.40 cm) and Algorta (18.64 cm). The coefficient of genotype effects and environment interaction ( $C_{ge}^2 = 0.151$ ) was low ( $< 0.4$ , Resende et al., 1995). The correlation between clones and the two sites was high ( $r_{gloc} = 0.708$ ), indicating a simple interaction effect (Vencovsky and Barriga, 1992), or that the mean dbh growth of the clones have similar ranking classification in both sites. This suggests that many clones can be selected for productivity in both environments. In addition, we also considered clone classification for productivity, stability, and adaptability (HMPRGV) in comparison to the control. This allows us to evaluate hybrid improvement related to one of the parental species. In this case, the interaction was simple.

### Heritability

The heritability coefficient plays a key role in the selection process because it indicates how much of the phenotypic

variation is based on genetic control (Vencovsky and Barriga, 1992). The mean site heritability between clones ( $H_c^2$ ) was high (0.724), indicating a high degree of genetic control for dbh and a high probability of obtaining genetic gains from clone selection. This result also demonstrates the precision of the experiment, and that the number of repetitions was enough to control for environmental effects (Maia et al., 2009; Rosado et al., 2012). Individual heritability ( $H_g^2$ ) was median (0.367) and the selective accuracy for mean clones ( $r_c$ ) was high (0.851). With greater accuracy, the selection precision is also greater, as are the genetic gains. In the present study, high accuracy indicates a strong association between true genetic value and phenotype, which is highly favourable for selection (Table 1).

### Coefficients of variation

The coefficient of individual genetic variation ( $CV_g = 10.9$  %) was high ( $> 10$  %), based on Resende (2002). This parameter represents the fraction of total phenotypic variation that is genetic in nature (Resende, 2002); high values of  $CV_{gi}$  indicate a greater possibility of obtaining genetic gains through selection. Thus, the studied population has potential to contribute to tree improvement through the selection of clones with larger dbh. Furthermore, the magnitude of accuracy for genotype selection ( $r_c = 0.851$ ) and coefficient of relative variation ( $CV_r = 0.916$ ) were high. The high  $r_c$  indicates a strong correlation between the predicted values and real genotypic values, and a  $CV_r$  similar to unity (1.0) indicates a highly favourable situation for selection and improvement of the dbh trait (Vencovsky and Barriga, 1992).

### SNP markers

The genotyping results were compared to those obtained by Silva-Junior et al. (2015), who evaluated the EUChip60K chip for 12 *Eucalyptus* sp., including *E. grandis* and *E. globulus*. They obtained 30,040 SNPs for *E. grandis* and 19,299 for *E. globulus* and in both cases SNPs were from a “Species-specific cluster”. While the results of the present study resemble those obtained by Silva-Junior et al. (2015) for *E. globulus*, it is reasonable to have a reduction in the number of markers when crossing *E. globulus* x *E. grandis*. The differences in number of loci selected here compared to the study by Silva-Junior et al. (2015) is related to the fact that the authors employed a call rate greater than 98 % and an MAF > 0.01. Furthermore, in the present study only loci in Hardy-Weinberg equilibrium were selected.

In the relatedness analysis, we can see a traceability error in that clones 1751 and 1758 are identified as the same genotype (Figure 2). The importance of this analysis should be highlighted as its use can avoid the mislabelling of parentage, and which consequently compromises the certification of breeding programmes that guarantee the identity of a germplasm via genotyping (Torres-Dini et al., 2011; De Vieira et al., 2013; Shanmugapriya and Yasodha, 2014). Incorporating traditional breeding strategies with genotyping increases the level of knowledge about improvement populations, enabling a better selection of clones and preventing common errors, thus improving the effectiveness of decision-making for breeders.

### Genomic wide prediction model (GWP)

Our model showed no precision to predict the dbh trait and no correlation between the rank classification of GWP and results for HMRPGV (Table 2). These results were consistent with theoretical expectations. According to simulations carried out by Grattapaglia and Resende (2011), for tree species it is necessary to have an improvement population of at least 1,000 individuals characterized for phenotype and genotyped; with 1,000 individuals, an accuracy of more than 0.80 was observed with high marker density. However, a population of over 2,000 individuals has little additional impact on accuracy. The results confirm the underlying assumptions of the simulations carried out by Grattapaglia and Resende (2011). Nevertheless, these are the first results obtained by applying the GWP method in a tree genetic improvement program in Uruguay. With advances in technology and reductions in genotyping costs, new populations may be genotyped and included in this database, and subsequently more accurate results can be obtained with GWP. With wide genomic selection, breeding time can be reduced by half and the efficiency of gains can exceed 100% (Grattapaglia and Resende, 2011). According to Grattapaglia and Resende (2011), by reducing flowering time from 12 to three years, the gain may be greater than 300 %.

The heritability estimated from phenotypes is considered the upper limit that could be explained by the GWP model (Resende et al., 2012). In the present study, the model captured only 37 % of genotypic heritability (Table 3). This result is not consistent when using all markers to estimate the model for complex traits (Meuwissen et al., 2001). With 200 large effect markers, one can capture 80 % of the phenotypic heritability

Table 3

Estimate of parameters from genomic wide prediction (GWP) selection for *Eucalyptus grandis* x *Eucalyptus globulus*

Parameter	Estimate
Genotypic heritability	0.37
Accuracy	0.81
Number of markers	15.196
Genotyped and phenotypically characterized individuals	78
Molecular heritability	0.17
Predictive capacity	-0.15
GS accuracy	-0.25

(Resende et al., 2012). Higher percentages (around 97 %) can be achieved with only 300 to 500 markers (Resende et al., 2012). Despite the hybrid origin of the population, which could be contributing to the high linkage disequilibrium, the effect required to capture the maximum genetic variation was not sufficient due to the number of individuals in the population. Furthermore, the number of markers and the number of individuals in genomic selection is crucial for accurate estimates of genetic parameters. Another effect that has a negative impact on transferability of genomic selection is the GxE interaction, as the GWP is specific for each population (Resende et al., 2012). This information should be considered in future work.

### Clone selection

We focused our analysis on the top ten clones with the best performance for both sites. The selection strategy was based on the following criteria: i) confirmation of parentage and identity (Figure 2); ii) selection of the best clones for productivity (Table 2); and iii) selection of the best clones based simultaneously on productivity, stability, and adaptability (HMRPGV, Table 2). The hybrid population confirmed a level of relatedness close to 0.5 (Figure 2), coinciding with the theoretical expectation for full-sibs.

A total of six clones were ranked higher than the control clone 3523 for the mean genotypic value ( $\mu + g + ge$ ), but only four (1744, 1769, 1670, 1796) obtained genetic gains greater than the control clone (Table 2). Clone 1744 ranked in the first position and the genetic gain was estimated at 3.4 % in Tres Bocas, 4.1 % in Algorta, and 3.1 % for the two sites combined. The value of genetic gain for the control clone (3523) was 3.2, 2.8, and 2.5 %, respectively. For the top ten clones, we observed that the predicted gain is lower for the two sites together than for each site separately (Table 2). This suggests a better use of the effects of GxE in site selection (Rosado et al., 2012). Clones 1744, 1769, and 1670 ranked in the top five most productive in all environments and did not suffer significant environmental effects, suggesting that these genotypes present a limited variation for the GxE interaction. Thus, these three clones are the most highly recommended for commercial reforestation in both sites. However, the simultaneous selection for productivity, stability, and adaptability using HMRPGVxGM (Resende et al., 2012) has several advantages as

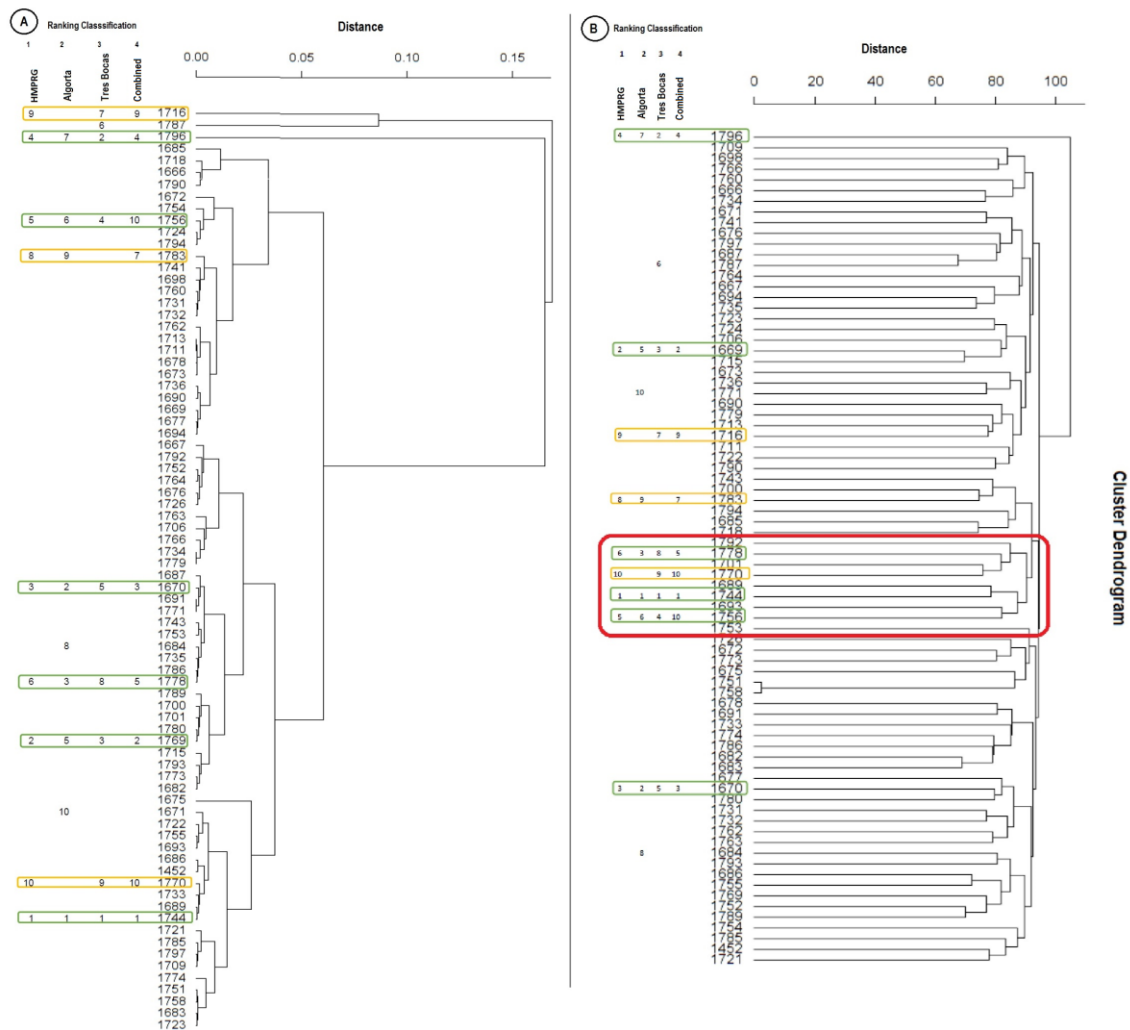


Figure 3

A: Dendrogram based on heterozygosity, showing the top ten ranked positions obtained for Algorta, Tres Bocas, and both sites together; 1: Classification ranking for stability and adaptability of genetic values (HMPRGV); 2, 3, and 4 represent the predicted dbh genetic gain positions. B: Similarity clustering dendrogram, showing the top ten rankings obtained for Tres Bocas, Algorta, and both sites together; 1: Stability and adaptability of genetic values (HMPRGV); 2, 3, and 4 represent the ranking for the predicted genetic gain for dbh

it considers the genotypic effects as random and provides genotypic values of stability and adaptability (not phenotypic values). This model allows us to calculate the combined genetic gains for the three attributes simultaneously. The value of HMPRGVxGM is the result of HMPRGV multiplied by the general means from both study locations and provides a genotypic mean value for the traits, which is reduced due to instability and increased due to adaptability (Rosado et al., 2012). In the HMPRGVxGM classification, clones 1744, 1769, and 1670 appear in the top five, while clone 1756 reached second place (Table 2). The *E. grandis* control clone (3523) was ranked third in the classification. It is important to note that clones 1744 and 1756 are both ranked higher than the *E. grandis* control clone for HMPRGV and HMPRGVxGM. This confirms that these clones are more productive, adaptable, and stable than the *E.*

*grandis* control in a region of Uruguay where the stability and adaptability of *E. globulus* is problematic (Balmelli and Resquin, 2005).

Our hybrid clones have 25 % of the *E. globulus* genome. It would be interesting in future studies to evaluate the wood quality of these clones to determine the *E. globulus* traits inherited in these hybrids, such as the proportion of cellulose pulp capacities. As such, the SNP database reported in this article could be associated with new phenotypic data. Furthermore, the quantitative traits exhibited by the population may be associated with their corresponding genes through Genomic Wide Association (GWAS) and Regional Heritability Mapping (RHM) studies (Resende et al., 2017)

The clustering analysis based on heterozygosity and similarity were associated with the top ten ranked clones for Tres

Bocas, Algorta, the two environments together and HMPRGV (Figure 3). Clones that classified in the top ten for the four rankings are shown in green and clones that were in the top ten in at least in three of the four classifications are shown in yellow. An association is observed among the four rankings with some variation in order. The GxE interaction analysis also indicates that it would be possible to select some of the same clones for the two environments to increase productivity. The top ten hybrids demonstrate a productivity and adaptability similar to the *E. grandis* control. When the hybrid population is compared with *E. globulus* previously studied in Rio Negro State, Uruguay (Balmelli and Resquin, 2005), the results for survival and adaptability exceed those of the *E. globulus* parents.

Finally, it is important to emphasize that the selected clones originated from a full-sib family; as such, they should be preferentially used in reforestation based on cloning, since the crossing of these clones is expected to generate inbreeding due to mating among related individuals. As the endogamy originating from mating between related individuals is equal to the coancestry coefficient among parents (Moraes et al., 2012), which in the present case of full-sibs is 0.25, we would expect at least 25 % inbreeding in descendants of crossings between the clones suggested for selection. Inbreeding in *Eucalyptus* species produces inbreeding depression, resulting in mortality, infertility, among other effects, as has been well documented in the literature (Eldridge and Griffin, 1983; Costa and Silva et al., 2011; Wu et al., 2015; Hedrick et al., 2016). However, the genotype information could be used in controlled crosses with other elite clones from unrelated populations. In such crosses, the comparison of genotypes based on SNPs will allow monitoring of the levels of homozygotes, while maintaining high levels of heterozygosity and avoiding inbreeding in the developed populations.

## Conclusions

We found a genetic variation correlated with dbh that can be exploited through the highly accurate selection of clones. Genotypes 1744, 1796, 1756, 1670, 1778, and 1769 present greater potential for commercial plantations in both studied sites due to higher productivity, stability, and adaptability. Clone 1744 is the most suitable for reforestation at both sites due to its greater genetic gain. The use of the SNP markers helped to confirm the degree of parentage between the clones, as well as provide control over clonal identity. Genotyping data based on the rrBLUP genomic selection model showed limited accuracy due to the small size of the studied population.

## Acknowledgement

This study was part of the Doctoral Thesis of Diego Torres-Dini in the Postgraduate program in Agronomy, UNESP, Ilha Solteira Campus. Alexandre M. Sebbenn, Marcos D.V. Resende, and Mario L.T. Moraes acknowledge CNPq for fellowships in research productivity. The authors thank INIA, Uruguay, for financial support necessary for the realization of this project. We also thank UPM Forestal Oriental for providing genetic

material and technical information. The authors are also grateful to Dr. Daniela Canuto and Dr. Flavia C.R. Lisoni for a previous review of the manuscript and Dr. Hector H. Dini for early suggestions in the development of this manuscript.

## References

- Aguiar AV de, Teixeira-Freitas DMA, de Almeida Filho JE, Sousa VA, Resende MDV, Silva-Junior OB, Grattapaglia D (2015) Genomic prediction of growth traits in *Pinus taeda* using genome-wide sequence-based DArT-seq markers. In: IUFRO TREE Biotechnol Conference 2015, Florence, Italy: Forests: the importance to the planet and society.
- Balmelli G, Resquin F (2005) Evaluación productiva de orígenes de *Eucalyptus globulus* en zonas litoral y norte. *Serie Técnica INIA* 149:1–19.
- Balmelli G, Resquin F (2008) Evaluación productiva de especies de *Eucalyptus* en Zona 9, a los 13 años. *INIA Serie Actividades de Difusión* 535:1–4.
- Bartholomé J, Mandrou E, Mabiala A, Jenkins J, Nabihoudine I, Klopp C, Schutz J, Plomion C, Gion JM (2015) High-resolution genetic maps of *Eucalyptus* improve *Eucalyptus grandis* genome assembly. *New Phytologist* 206(4):1283–1296. Available at <https://doi.org/10.1111/nph.13150>
- Booth TH (2013) Eucalypt plantations and climate change. *Forest Ecology and Management* 301:28–34. Available at <https://doi.org/10.1016/j.foreco.2012.04.004>
- Colombari Filho JM, Resende MDV, Moraes OP, Castro AP, Guimarães EP, Pereira JA, Utumi MM, Bresseghele F (2013) Upland rice breeding in Brazil: a simultaneous genotypic evaluation of stability, adaptability and grain yield. *Euphytica* 192(1):117–129. Available at <https://doi.org/10.1007/s10681-013-0922-2>
- Costa e Silva J, Hardner C, Tilyard P, Potts BM (2011) The effects of age and environment on the expression of inbreeding depression in *Eucalyptus globulus*. *Heredity* 107(1):50–60. Available at <http://doi.org/10.1038/hdy.2010.154>
- De Viera JL, Tavares Filho LFQ, Souza FVD, Alves ACC (2013) Development of interspecific hybrids of cassava and paternity analysis with molecular markers. *The Journal of Agricultural Science* 151(6):849–861. Available at <http://doi.org/10.1017/S0021859612000871>
- Desta ZA, Ortiz R (2014) Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science* 19(9):592–601. Available at <http://doi.org/10.1016/j.tplants.2014.05.006>
- Eldridge K, Griffin AR (1983) Selfing effects in *Eucalyptus regnans*. *Silvae Genetica* 32:216–221.
- Farias FJ, Carvalho LP, Silva Filho JL, Teodoro PE (2016) Usefulness of the HMRP-GV method for simultaneous selection of upland cotton genotypes with greater fiber length and high yield stability. *Genetics and Molecular Research*, 15(3): gmr.15038439. Available at <https://doi.org/10.4238/gmr.15038439>
- Grattapaglia D, Kirst M (2008) *Eucalyptus* applied genomics: from gene sequences to breeding tools. *New Phytologist* 179(4):911–929. Available at <http://doi.org/10.1111/j.1469-8137.2008.02503.x>
- Grattapaglia D, Resende MDV (2011) Genomic selection in forest tree breeding. *Tree Genetics and Genomes* 7:241–255. Available at <http://doi.org/10.1007/s11295-010-0328-4>
- Griffin A, Harbard J, Centurion C, Santini P (2000) Breeding *Eucalyptus grandis* x *globulus* and other inter-specific hybrids with high inviability - problem analysis and experience with Shell Forestry projects in Uruguay and Chile. Dungey HS, Dieters MJ Nikles DG (eds), *Hybrid Breed Genet For Trees Proc QFRI/CRC-SPF Symp 9–14th April 2000 Noosa, Queensland, Aust Dep Prim Ind Brisbane* 1–13.
- Hedrick PW, Hellsten U, Grattapaglia D (2016) Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. *New Phytologist* 209(2):600–611. Available at <https://doi.org/10.1111/nph.13639>
- Heffner EL, Lorenz AJ, Jannink JL, Sorrells ME (2010) Plant breeding with genomic selection: gain per unit time and cost. *Crop Science* 50(5):1681–1690 Available at <http://doi.org/10.2135/cropsci2009.11.0662>

- Hudson CJ, Freeman JS, Myburg AA, Potts BM, Vaillancourt RE (2015) Genomic patterns of species diversity and divergence in *Eucalyptus*. *New Phytologist* 206(4):1378–1390. Available at <https://doi.org/10.1111/nph.13316>
- Isik F (2014) Genomic selection in forest tree breeding: The concept and an outlook to the future. *New Forests* 45(3):379–401. Available at <https://doi.org/10.1007/s11056-014-9422-z>
- Lima BM (2014) Bridging genomics and quantitative genetics of *Eucalyptus*: genome-wide prediction and genetic parameter estimation for growth and wood properties using high-density SNP data. PhD (Dissertation) – Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba. Available at <http://10.11606/T.11.2014.tde-25062014-085814>
- Maia M, Resende MDV, Paiva JR, Cavalcanti JJV, Barros LM (2009) Seleção simultânea para produção, adaptabilidade e estabilidade genotípicas em clones de cajueiro, via modelos misto. *Pesquisa Agropecuária Tropical* 39(1):43–50. Available at <https://www.revistas.ufg.br/pat/article/view/5704/4512>
- Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157(4):1819–1829.
- Moraes MA, Gaino A, Moraes MLT, Freitas MLM, Sebbenn (2012) Estimating coancestry within open-pollinated progenies of a dioecious species: the case study of *Myracrodruon urundeuva*. *Silvae Genetica* 61(6):256–264.
- Myburg AA, Griffen AR, Sederoff RR, Whetten RW (2003) Comparative genetic linkage maps of *Eucalyptus grandis*, *Eucalyptus globulus* and their F1 hybrid based on a double pseudo-backcross mapping approach. *Theoretical and Applied Genetics* 107:1028–1042. Available at <http://doi:10.1007/s00122-003-1347-4>
- Paseyro J (2015) Más vale especie conocida que otra por conocer. *Revista Forestal* [online]. Uruguay, to be found at <<http://www.revistaforestal.uy/silvicultura/mas-vale-especie-conocida-que-otra-por-conocer.html>> [quoted, 24.03.2017]
- Pinto Junior JE, Sturion JA, Resende MDV, Rozelli Junior P (2006) Avaliação simultânea de produtividade, adaptabilidade e estabilidade genotípica de *Eucalyptus grandis* em distintos ambientes do Estado de São Paulo. *Boletim de Pesquisa Florestal* 53:79–108.
- Pupin S, dos Santos VA, Gonzalez-Zaruma UD, Miranda AC, Muller da Silva PH, Marino CL, Sebbenn AM, Moraes MLT (2015) Produtividade, estabilidade e adaptabilidade em progênies de polinização aberta de *Eucalyptus urophylla* S.T. Blake. *Scientia Forestalis* 43(105):127–134.
- Resende MDV (2002) *Genética biométrica e estatística no melhoramento de plantas perenes*. Brasília: Embrapa Informação Tecnológica, 975 p, ISBN 85-7383161-8
- Resende MDV (2016) Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology* 16(4):330–339. Available at <http://doi:10.1590/1984-70332016v16n4a49>
- Resende MDV, Araujo AJ, Sampaio PTB, Wiecheteck MSS (1995) Acurácia seletiva, intervalos de confiança e variâncias de ganhos genéticos associados a 22 métodos de seleção em *Pinus caribaea* var. *hondurensis*. *Floresta* 25(1/2):3–16.
- Resende MDV, Resende MF Jr, Sansaloni CP, Petrolí CD, Missiaggia AA, Aguiar AM, Abad JM, Takahashi EK, Rosado AM, Faria DA, Pappas GJ Jr, Kilian A, Grattapaglia D (2012) Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* 194(1):116–128. Available at <https://doi:10.1111/j.1469-8137.2011.04038.x>
- Resende RT, Resende MDV, Silva FF, Azevedo CF, Takahashi EK, Silva-Junior OB, Grattapaglia D (2017) Regional heritability mapping and genome-wide association identify loci for complex growth, wood and disease resistance traits in *Eucalyptus*. *New Phytologist* 213(3):1287–1300. Available at <https://doi.org/10.1111/nph.14266>
- Resquin F, Balmelli G (1999) Evaluación de especies y orígenes de *Eucalyptus* al quinto año. *INIA Ser Técnica* 106:20.
- Rosado AM, Rosado TB, Alves AA, Laviola BG, Bhering LL (2012) Seleção simultânea de clones de eucalipto de acordo com produtividade, estabilidade e adaptabilidade. *Pesquisa Agropecuária Brasileira* 47:964–971. Available at <http://doi:10.1590/S0100-204X2012000700013>
- Shanmugapriya A, Yasodha R (2014) Unique pattern information combinations for clonal discrimination in *Eucalyptus camaldulensis* Dehnh. using micro-satellite markers. *Indian Journal of Biotechnology* 13:75–80.
- Silva-Junior OB, Faria DA, Grattapaglia D (2015) A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytologist* 206(4):1527–1540. Available at <http://doi:10.1111/nph.13322>
- Torres-Dini D, Bennadji Z, Cabrera M, et al (2011) Use of SSR-Tools for clone certification in Uruguayan *Eucalyptus grandis* and *Eucalyptus dunii* breeding programs. *BMC Proceedings* 5(7):P58. Available at <http://doi:10.1186/1753-6561-5-S7-P58>
- Vencovsky R, Barriga P (1992) *Genética Biométrica no Fitomelhoramento*. Ribeirão Preto: Sociedade Brasileira de Genética.
- Wu SJ, Xu JM, Lu ZH, Li GY, Pan LQ, Han C (2015) Effects of inbreeding on growth and wood properties of selfed *Eucalyptus urophylla* progenies. *Journal of Tropical Forest Science* 279(3):369–375. Available at <http://www.jstor.org/stable/43490295>