Bayesian analysis of growth, stem straightness and branching quality in full-sib families of *Eucalyptus globulus*

Freddy Mora¹, Paulina Ballesta¹*, Nicolle Serra²

¹Universidad de Talca - Instituto de Ciências Biolôgicas - Talca, Chile.
²Semillas Imperial SpA - Los Ángeles, Chile.

**ABSTRACT:** *Eucalyptus globulus* is one of the most commonly planted hardwood species for industrial use in various temperate regions around the world. The present study aimed to evaluate 62 full-sib families of *E. globulus* in one of the southernmost progeny trials of the species in the south of Chile. Estimates of genetic parameters for stem straightness, branching quality and growth traits were based on a Bayesian modelling approach using Gibbs sampling. A Bayes Factor (BF) analysis supported the hypothesis of significant additive genetic variation for all traits under study. Conversely, the BF supported a model with significant dominance effects for the diameter at breast height and stem volume, which explained up to 25% of the phenotypic variation. The greatest narrow-sense heritability estimates were found for the tree height and stem straightness, which were 0.15 (0.08 to 0.26) and 0.18 (0.10 to 0.28), respectively (mean of posterior distributions and 90% credible sets). In turn, the branching quality had a low heritability (narrow-sense) that varied from 0.05 to 0.10 (90% Bayesian credible region). The mean posterior estimate of genetic correlation between both quality traits was 0.22 (0.01 to 0.63, 90% Bayesian credible set from a bi-trait threshold model), which indicates that stem straightness is positively related to branching quality. Our findings reveal that the study population responds to common patterns of breeding populations of *E. globulus*. This information is valuable for the development of improved seeds in the southern zone of Chile.

**Key words:** bayes factor, credible region, gibbs sampling, quality traits, threshold models.

https://doi.org/10.1590/1678-4499.20180317
INTRODUCTION

Trees of the genus *Eucalyptus* L’Hér. are recognized for their high biomass production, rapid growth rate, good adaptation to diverse environmental conditions and excellent wood quality for the production of paper and products derived from solid wood (Schmit et al. 2015; Mora and Arriagada 2016). In particular, *E. grandis*, *E. urophylla* and *E. globulus* are the most commonly planted hardwood species for industrial uses in tropical (or subtropical) and temperate zones, respectively (Carocha et al. 2015; Quang et al. 2010). Therefore, they stand out as the targets of multiple breeding programs and silvicultural management (Rosado et al. 2010; Carocha et al. 2015; Carbonari et al. 2016).

*Eucalyptus globulus* Labill. is a genetically diverse species with different geographic races (Foster et al. 2007) that allow it to tolerate environments with a certain degree of drought and/or low temperatures. Naturally, *E. globulus* is dominant in the coastal forests of southeastern Australia (Foster et al. 2007), where the climate is oceanic. Nevertheless, the species has been successfully and extensively planted in several temperate regions worldwide (e.g., Chile, Portugal, Australia, Spain) (Thavamanikumar et al. 2011; Águas et al. 2014; Mora and Serra 2014; Larcombe et al. 2013). Notably, *E. globulus* has been positively grown in a widespread range of environmental conditions that are adverse for plant establishment (Dutkowski and Potts 1999; Tibbits et al. 2006). Particularly in Chile, several studies have been focused on the mechanisms through which *E. globulus* trees respond to abiotic stresses (e.g., Navarrete-Campos et al. 2013; 2017; Aguayo et al. 2016).

*E. globulus* is the second most important woody species in the Chilean forest industry (Ballesta et al. 2018), where it comprises approximately 21% of national plantations (Morales et al. 2015). Traditionally, the species has experienced a rotation period that varies between eight and 12 years. However, breeding and intensive forest management have allowed production optimization, reducing the rotation time to five years (Morales et al. 2015). Due to the economic importance of *Eucalyptus*, several tree breeding programs have achieved substantial gains in various traits, such as growth (Blackburn et al. 2013; Mora and Serra 2014; Hamilton et al. 2015), wood properties (Stackpole et al. 2010; Hamilton et al. 2010), flowering-related traits (Ballesta et al. 2015; Contreras-Soto et al. 2016) pulp yield (Stackpole et al. 2010; Hamilton et al. 2010) and stem quality traits (Mora and Serra 2014; Arriagada et al. 2018). Estimation of the individual breeding value of a trait is closely related to the degree of precision with which the (co)variance components are estimated. In fact, lower precision is directly proportional to a greater deviation of the real genetic value of a genotype (Faria et al. 2007). In this context, the Bayesian method is a useful alternative for scientific inference of the genetic merit of trees because it considers levels of uncertainty in the estimated parameters and, generally, the credibility regions are more accurate than the confidence intervals obtained with frequentist inference (Gazola et al. 2016). Bayesian inference has been increasingly used in plant breeding and genetic studies in general (Fresnedo-Ramírez et al. 2017; Torres et al. 2018) by Markov Chain Monte Carlo (MCMC) methods, which use Markov sequences to effectively simulate complex (or not mathematically addressable) distributions. For instance, the joint density distribution that considers known environmental effects (e.g., experimental design), (co)variance components and additive and non-additive genetic effects, among others. An advantage of Bayesian procedures is the use of prior information in the analysis, which is particularly important when data are scarce (Cappa et al. 2012; Amaral Junior et al. 2016; Mora et al. 2016). Therefore, this method of statistical inference has been used in the analysis and estimation of genetic parameters in different tree species, including *Eucalyptus* (Vargas-Reeve et al. 2013; Mora and Serra 2014), *Pinus* (Cappa and Cantet 2006; 2008) and *Populus* (Wang et al. 2014).

The present study aimed to genetically evaluate full-sib families in one of the southernmost progeny trials of the species in the south of Chile. The estimation of genetic parameters (heritability, additive genetic correlations and variance components) of growth, stem straightness and branching quality (or branch-related defects) was based on Bayesian principles, including the selection of genetic models based on the Bayes Factor as a method of statistical inference.

MATERIALS AND METHODS

Field trial and phenotypic measurements

The study included 62 full-sib families resulting from an incomplete factorial mating design that involved 15 and 21
male and female parental trees, respectively. The trial was conducted in 2012 in the Chilean administrative region of Los Lagos, district of Purranque (40°58' S, 73°30' W, 326 m above sea level), under a randomized complete block design with 30 blocks and single-tree plots. The site features an oceanic climate with an average annual temperature of 13 °C; the average temperatures in the coldest and warmest months are 6 °C and 16 °C, respectively. The area was plowed to a depth of 0.60 m, and was fertilized with Di-Ammonium Phosphate (15% N, 15% P, and 15% K; Basacote® Plus).

The following traits of interest were evaluated in 4-year-old trees in 2016: tree height (H), diameter at breast height (DBH), wood volume (VOL), stem straightness (STR) and branching quality (BQ, or branch-related defects). STR was measured initially on a categorical 6-level scale according to Cameron et al. (2012) (1 = very crooked to 6 = completely straight without loss of productivity). BQ was also measured on a categorical 6-level scale (1 = tree with serious limitations to 6 = tree with all branching variables in good condition without loss of productivity). In addition, an analysis of tree survival was included, which was recorded as a binary response at 4 years old.

Bayesian genetic analysis

For estimation of the variance components and genetic parameters, a Bayesian analysis was carried out using the following base model (complete model) (Eq. 1):

\[ y = X\beta + Z_1a + Z_2s + \varepsilon \]  

where: \( y \) corresponds to the observed values for a specific trait (phenotype) and \( X, Z_1 \) and \( Z_2 \) are the known incidence matrices that relate the observation vector (\( y \)) to vectors \( \beta, a \) and \( s \), respectively; \( B \) is the vector of the block effect; \( a \) is the additive effect vector of the individual trees; \( s \) is the effect vector of the full-sib families (accounts for dominance); and \( \varepsilon \) is the residual effect vector.

The Bayes Factor (BF) was used to test the significance of the effects of vectors \( a \) and \( s \) by model comparison, i.e., models M1 versus M2 in each of the following cases:

To evaluate the significance of the additive effects (Eqs. 2 and 3):

M1: \[ y = X\beta + \varepsilon \]  

M2: \[ y = X\beta + Z_1a + Z_2s + \varepsilon \]  

To evaluate the significance of the effects of the full-sib families (dominance) (Eqs. 4 and 5):

M1: \[ y = X\beta + Z_1a + \varepsilon \]  

M2: \[ y = X\beta + Z_1a + Z_2s + \varepsilon \]  

The BF based on the null hypothesis \( H_0: a = 0 \) or \( H_0: s = 0 \) comparing models M1 or M2, respectively, corresponds to Eq. 6:

\[ BF = \frac{f(M_1 | y)}{f(M_2 | y)} \left( \frac{f(M_1)}{f(M_2)} \right)^{-1} \]  

where: \( f(M_1 | y) \) and \( f(M_2 | y) \) correspond to the relative posterior probabilities of the competing models M1 and M2, respectively; and \( f(M_1) \) and \( f(M_2) \) correspond to the prior probabilities for the competing models M1 and M2, respectively.

Subsequently, the posterior distributions of the variance components of the selected models were determined using the Gibbs sampling algorithm implemented in the MTGSAM program (Van Tassell and Van Vleck 1996). For the traits measured on the categorical scale (STR and BD), the likelihood-based threshold version of MTGSAM was used (Van Tassell et al. 1998). The following model (Eq. 7) was considered for the analysis of genetic associations between pairs of traits:

\[
\begin{bmatrix}
  y_1 \\
  y_2
\end{bmatrix} =
\begin{bmatrix}
  X_1 & \theta & Z_{11} & \theta \\
  \theta & X_2 & \theta & Z_{12}
\end{bmatrix}
\begin{bmatrix}
  \beta_1 \\
  \beta_2 \\
  a_1 \\
  a_2
\end{bmatrix}
+ \begin{bmatrix}
  Z_{21} & \theta & s_1 \\
  \theta & Z_{22} & s_2
\end{bmatrix}
\begin{bmatrix}
  \varepsilon_1 \\
  \varepsilon_2
\end{bmatrix}
\]  

whose (co)variance components assumed an Inverted Wishart (IW) prior distribution. A uniform prior distribution (Flat) was considered for known environmental effects (blocks), whereas a normal distribution was used for the vector of observed responses, additive genetic effects and full-sib family effects. The convergence and auto-correlation of the Gibbs chains were assessed using the tests available in the CODA library of the R program, version 2.6.2 (R Development Core Team 2011).
Parameter estimates

Posterior estimates for individual-tree heritability \((h^2)\), coefficient of additive genetic variation \((CVa)\), dominance ratio \((d^2)\) and coefficient of genetic variation of dominance \((CVd)\) were estimated from the uni-trait analysis using the following expressions (Eqs. 8 to 11):

\[
\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_s^2 + \hat{\sigma}_e^2} \quad (8)
\]

\[
CVa = \frac{\sqrt{\hat{\sigma}_a^2}}{\hat{x}} \cdot 100 \quad (9)
\]

\[
\hat{d}^2 = \frac{4 \cdot \hat{\sigma}_s^2}{\hat{\sigma}_a^2 + \hat{\sigma}_s^2 + \hat{\sigma}_e^2} \quad (10)
\]

\[
CVd = \frac{4 \cdot \hat{\sigma}_s^2}{\hat{x}} \cdot 100 \quad (11)
\]

where \(\hat{\sigma}_a^2\), \(\hat{\sigma}_s^2\) and \(\hat{\sigma}_e^2\) correspond to the additive genetic, full-sib family and residual variances, respectively. If the effect of full-sib families was not significant according to the BF, only the heritability and the corresponding coefficient of additive genetic variation were considered in the analysis. The additive genetic association between each pair of traits measured in the same tree was calculated from a bi-trait model using Eq. 12:

\[
\hat{r} = \frac{\hat{\sigma}_{xy}}{\sqrt{\hat{\sigma}_{ax}^2 + \hat{\sigma}_{ay}^2}} \quad (12)
\]

in which \(\hat{\sigma}_{xy}\) corresponds to the estimation of the additive covariance component between two pairs of traits \(X\) and \(Y\) (posterior mode values), \(\hat{\sigma}_{ax}^2\) and \(\hat{\sigma}_{ay}^2\) are the posterior variances for each pair of traits in the analysis.

RESULTS AND DISCUSSION

Significant additive genetic variation was observed for all traits under study (Table 1). Specifically, each trait had a BF < 0.01, or Log BF < 0, which provided decisive evidence against the model that considered an additive effect equal to zero (i.e., \(\hat{\sigma}_a^2 = 0\)). Conversely, the BF indicated a significant family effect (i.e., dominance variance) for the DBH and the VOL, whose dominance ratios were moderate and were superior to that of heritability (Table 2) for both traits; that is, VOL \((d^2 = 0.23, CVd = 23.7\%)\) and DBH \((d^2 = 0.25, CVd = 9.7\%)\). The coefficients of experimental variation were 14.8\% (H), 18.5\% (DBH), 47.4\% (VOL), 35.5\% (STR) and 20.4\% (BQ), with H and DBH showing moderate values according to the proposed classification by Mora and Arriagada (2016). Additionally, no significant difference for tree survival was detected among families (BF > 1).

The Gibbs sampling chains converged for all posterior distributions of the parametric estimates using 10,000 burn-in iterations and a total of 50,000 Gibbs sampling rounds, in which 4,000 samples were withdrawn to estimate the marginal posterior distributions. Based on the models selected according to the BF (Table 1), the traits under study ranged from low to moderate heritability (estimates from the posterior distributions) (Casell 2009; Bush et al. 2011; Poke et al. 2006; Costa e Silva et al. 2009) with variation between 0.03 (VOL) and 0.15 (STR), and the coefficients of additive variation ranged between CVa = 4.4\% and CVa = 14.2\% for DBH and STR, respectively (Table 2). In a genetic evaluation of twelve-month-old \(E.\) globulus clones also in southern Chile, Mora et al. (2013) found moderate to high values of heritability of tree height with a range from 0.12 to 0.41 (mode value of the posterior distribution of heritability). In a frequentist approach, Costa e Silva et al. (2005) reported that the heritability of height growth for \(E.\) globulus in early stages (at 3 years old) ranged from 0.15 to 0.4 (evaluated under six different environmental conditions), which is also consistent with our findings. Previous studies and our results revealed that the height of plants of \(E.\) globulus is controlled by an additive genetic control and, in a minor proportion, by dominance effects. Contrarily, our results suggested that VOL and DBH could be moderately controlled by dominance effects, a result also reported by Hamilton et al. (2015) in \(E.\) globulus at age 10 (DBH: \(d^2 = 0.13;\) VOL: \(d^2 = 0.24\)).

The stem straightness and branch quality (or the proportion of defective branches) are important traits

---

Bragantia, Campinas, v. 78, n.3, p.328-336, 2019
for the production of timber of *E. globulus*, whose quality is sensitive to the market (Blackburn et al. 2013; Mora and Serra 2014). In the present study, STR showed a moderate heritability ranging from 0.10 to 0.28 (90% Bayesian credible set). These results matched the values obtained by Mora and Serra (2014) in a trail of half-sib families of *E. globulus* of 15 years old, in which they found low to moderate heritability varying from 0.03 to 0.21 (90% credible set). Conversely, Callister et al. (2011) found relatively higher estimates of heritability in full-sib families of *E. globulus* in Australia at an age similar to that of the present study (3.5 years) with a range from 0.10 to 0.46 and a mean of 0.28 (the upper limit of the 90% credible set of the posterior distribution found in the present study). In turn, the branch-related defects had a low heritability that varied from 0.05 to 0.1 (CVa = 4.7%). Although this finding indicated that the genetic progress for this trait would be modest in the present breeding population, we could emphasize that only 2.2% of the trees in the test were in the lower BQ category, with the great majority showing an intermediate level of quality (74%). Callister et al. (2011) determined that the branch thickness (a trait related to BQ) featured low genetic control ($h^2 = 0.05$), which was consistent with our results.

The estimates of the additive genetic correlations between pairs of traits (i.e., the point estimate of the posterior distribution, mode and 90% credible regions) are shown in Table 3. According to the 90% credible set, 6 of the 10 estimates were significantly different from zero. As expected, the growth traits showed a positive association and were significantly different from zero, which agreed with previous studies of *E. globulus* (Hamilton et al. 2010; Mora and Serra 2014). Genetic correlation values between growth traits suggest that the genetic selection based on one growth trait would have a positive impact on other trait in early stages of *E. globulus*. Several studies have reported quantitative trait loci (QTLs) associated with DBH and H (Arriagada et al. 2018; Thumma et al. 2010; Ballesta et al. 2015), supporting to the hypothesis of genetic relation between these growth traits.

### Table 1. Model assessment based on Bayes factor for the total height (H), diameter at breast height (DBH), volume (VOL), stem straightness (STR) and branching quality (BQ) in full-sib families of *Eucalyptus globulus*.

<table>
<thead>
<tr>
<th>Competitive models</th>
<th>H</th>
<th>DBH</th>
<th>VOL</th>
<th>STR</th>
<th>BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y = \mathbf{X}\beta + \mathbf{Z}a + \mathbf{e}$</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>$y = \mathbf{X}\beta + \mathbf{Z}a + \mathbf{Z}s + \mathbf{e}$</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

According to Jeffreys (1961), if BF > 1 or LogBF > 0, model M1 is supported. If BF < 0.01 or LogBF < 0, decisive evidence against model M1.

### Table 2. Bayesian estimates of the individual-tree heritability ($h^2$), coefficient of additive genetic variation (CVa), dominance ratio ($d^2$) and coefficient of genetic variation of dominance (CVd) with lower and upper cutoffs for 90% credible sets the total height (H), diameter at breast height (DBH), volume (VOL), stem straightness (STR) and branching quality (BQ) in full-sib families of *Eucalyptus globulus*.

<table>
<thead>
<tr>
<th>Estimates</th>
<th>H</th>
<th>DBH</th>
<th>VOL</th>
<th>STR</th>
<th>BQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.15</td>
<td>0.07</td>
<td>0.28</td>
<td>0.06</td>
<td>0.18</td>
</tr>
<tr>
<td>Median</td>
<td>0.14</td>
<td>0.06</td>
<td>0.27</td>
<td>0.04</td>
<td>0.17</td>
</tr>
<tr>
<td>Mode</td>
<td>0.12</td>
<td>0.05</td>
<td>0.25</td>
<td>0.03</td>
<td>0.15</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.06</td>
<td>0.05</td>
<td>0.08</td>
<td>0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Lower cutoff</td>
<td>0.08</td>
<td>0.01</td>
<td>0.16</td>
<td>0.01</td>
<td>0.10</td>
</tr>
<tr>
<td>Upper cutoff</td>
<td>0.26</td>
<td>0.16</td>
<td>0.41</td>
<td>0.14</td>
<td>0.28</td>
</tr>
<tr>
<td>CVa (%)</td>
<td>5.2</td>
<td>4.4</td>
<td>8.6</td>
<td>14.2</td>
<td>4.7</td>
</tr>
<tr>
<td>CVd (%)</td>
<td>0.0</td>
<td>9.7</td>
<td>23.7</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Bayesian analysis of economic traits in bluegum

Cameron et al. (2012) found a similar genetic correlation between stem straightness and branching habit (scored for quality) in *Picea sitchensis* ($r = 0.28$). Our findings could be promising in the context of an early selection program of *E. globulus*. According to the results, stem straightness was more heritable than branching quality at age 4, but both traits were genetically related. Therefore, the selection based on stem straightness could promote the selection of other important economic traits with low heritability, such as branching quality.

The stem straightness presented a negative association with the total height ($-0.51$), whose value was significantly different from zero ($-0.67$ to $-0.14$), whereas the associations between straightness with VOL and DBH were not significantly different from zero even though the point estimates were negative. These findings were in agreement with those of Callister et al. (2011) in *E. globulus*, who found an additive genetic correlation between DBH and STR with an average value of $-0.18$ and variation from $-0.71$ to $0.33$. Mora and Serra (2014) also found negative additive genetic correlations between growth (DBH, H and VOL) and STR in a 15-year-old *E. globulus* trial, which were not significantly different from zero. In another *Eucalyptus* species, Vargas-Reeve et al. (2013) found that the additive genetic correlations between stem straightness with height and diameter were not significantly different from zero (95% Bayesian credible set: $-0.04$ to $0.37$ and $-0.16$ to $0.35$, respectively) when measured on 9-year-old trees. Coincidentally, Callister et al. (2008) also found non-significant genetic correlations between growth and STR in *E. cladocalyx* trees at 3.5 and 5.5 years of age.

**CONCLUSION**

Our findings revealed that the study population responds to common patterns of breeding populations of *E. globulus*. The total height of the trees and the stem straightness were the traits showing the greatest narrow-sense heritability, with moderate values and a negligible dominance effect. On the other hand, the dominance effect should be considered for future research and breeding prospects based on the wood diameter and/or volume. In addition, the dominance effect was negligible for branch-related defects, but is controlled by minor genetic additive factors at early stages.

**ACKNOWLEDGEMENTS**

The study was supported by FONDECYT (Grant no. 1170695) and Semillas Imperial SpA. Paulina Ballesta thanks CONICYT for a doctoral fellowship (CONICYT-PCHA/Doctorado Nacional/año 2016-folio 21160624).

**AUTHOR’S CONTRIBUTION**

Conceptualization, Mora F.; Methodology, Mora F.; Investigation, Mora F.; Writing – Original Draft, Mora F.; Writing – Review and Editing, Ballesta P.; Funding Acquisition, Mora F.; Resources, Serra N.; Supervision, Mora F. and Ballesta P.

**ORCID IDs**

F. Mora  
https://orcid.org/0000-0002-4845-3549  
P. Ballesta  
https://orcid.org/0000-0003-1428-0456  
N. Serra  
https://orcid.org/0000-0002-9889-2355

---

Table 3. Posterior means of genetic correlations (above the diagonal) and 90% credible sets from marginal posterior distributions (below the diagonal), calculated between pairs of traits for the total height (H), diameter at breast height (DBH), volume (V), stem straightness (STR) and branching quality (BQ).

<table>
<thead>
<tr>
<th>Traits</th>
<th>H</th>
<th>DBH</th>
<th>VOL</th>
<th>STR</th>
<th>BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td>0.45</td>
<td>0.61</td>
<td>$-0.43$</td>
<td>0.30</td>
</tr>
<tr>
<td>DBH</td>
<td>[0.14, 0.71]</td>
<td>1</td>
<td>0.94</td>
<td>$-0.20$</td>
<td>$-0.34$</td>
</tr>
<tr>
<td>VOL</td>
<td>[0.36, 0.82]</td>
<td>[0.89, 0.98]</td>
<td>1</td>
<td>$-0.24$</td>
<td>$-0.22$</td>
</tr>
<tr>
<td>STR</td>
<td>$[-0.67, -0.14]$</td>
<td>$[-0.50, 0.11]$</td>
<td>$[-0.55, 0.10]$</td>
<td>1</td>
<td>0.22</td>
</tr>
<tr>
<td>BD</td>
<td>[0, 0.71]</td>
<td>$[-0.70, -0.01]$</td>
<td>$[-0.65, 0.24]$</td>
<td>[0.01, 0.63]</td>
<td>1</td>
</tr>
</tbody>
</table>
REFERENCES


and pulp yield in *Eucalyptus globulus*. Tree Genetics & Genomes, 5, 291-305. https://doi.org/10.1007/s11295-008-0174-9


Hamilton, M. G., Potts, B. M., Geaves, B. L. and Dutkowski, G. W. (2010). Genetic correlations between pulpwod and solidwood selection and objective traits in *Eucalyptus globulus*. Annals Forest Science, 67, 511-511. https://doi.org/10.1051/forest/2010013


