

GGE biplot-based genetic selection to guide interspecific crossing in *Corymbia* spp.

Filipe Manoel Ferreira¹ , João Romero do Amaral Santos de Carvalho Rocha¹ , Rodrigo Silva Alves² , Renan Garcia Malikuski¹ , Marco Antônio Peixoto¹ , Suellen Sales de Oliveira¹ , Aurélio Mendes Aguiar³ , Leonardo Lopes Bhering^{1,*} 

1. Universidade Federal de Viçosa – Departamento de Biologia Geral – Viçosa (MG), Brazil.
2. Universidade Federal de Viçosa – Instituto Nacional de Ciência e Tecnologia do Café – Viçosa (MG), Brazil.
3. Suzano S/A – Centro de Tecnologia – São Paulo (SP), Brazil.

Received: Jun. 30, 2021 | **Accepted:** Jul. 28, 2021

Section Editor: Carlos Alberto Scapim

***Corresponding author:** leonardo.bhering@ufv.br

How to cite: Ferreira, F. M., Rocha, J. R. A. S. C., Alves, R. S., Malikuski, R. G., Peixoto, M. A., Oliveira, S. S., Aguiar, A. M. and Bhering, L. L. (2021) GGE biplot-based genetic selection to guide interspecific crossing in *Corymbia* spp. *Bragantia*, 80, e5221. <https://doi.org/10.1590/1678-4499.20210195>

ABSTRACT: Tree breeders are prioritizing the increase of the interspecific diversity in the breeding programs due to the pronounced environmental changes, and expansion of the forest frontier to less productive areas. In this scenario, there is a growing interest for species of the genus *Corymbia* due to their good growth, resistance to biotic and abiotic factors, and industrial properties. Also, *Corymbia* spp. can cross with each other, generating interspecific hybrids. However, there are great difficulties on defining the best crosses. Thus, the objectives of this study were to evaluate the genetic variability of 64 progenies of *Corymbia maculata* and 64 progenies of *Corymbia torelliana* and to use the GGE biplot method to point the most appropriated crosses to broad and specific environments in order to generate stable and productive interspecific hybrids of *C. maculata* and *C. torelliana*. The GGE biplot was an efficient method for evaluating the genetic variability between progenies, for visualizing the genotype by environment interaction pattern and for identifying the best crosses in order to generate interspecific hybrids of the species. By adopting a selection intensity of 10%, the recommended progenies to generate interspecific hybrids for the Aracruz municipality (ARA) were 10, 40, 49, 64, 20, 9 and 47 of *C. maculata* and 24, 56, 41, 64, 43, 31 and 12 of *C. torelliana*. For Três Lagoas municipality (TLA), the recommended progenies were 11, 17, 1, 48, 7, 63 and 21 of *C. maculata* and 42, 46, 61, 44, 10, 63 and 28 of *C. torelliana*.

Key words: *Corymbia* hybrids, genotypes, environment interaction, heterosis, genetic selection.

INTRODUCTION

Forest plantations are a strong ally to preserve natural resources and provide social and economic benefits (Carle et al. 2020). Also, they are fundamental to meet the growing demand for forest products. By 2050, it is expected that the manufacture of forest products will have to triple (Payn et al. 2015) due to the growing on human population. In this context, there is an increase interest for species more resistant to adverse conditions, such as those belonging to the genus *Corymbia* (Lee 2007; Lee et al. 2009). *Corymbia maculata* and *Corymbia torelliana* present good industrial and silvicultural characteristics (Shepherd et al. 2007). However, there is little information about the selection strategies for these species and its hybrids (Darrow 1985; Dickinson et al. 2013, Lee et al. 2009; Sato et al. 2010).

Corymbia interspecific hybrids have high biomass production and are faster growing than open pollinated seed lots (Loureiro et al. 2019). In addition, they often show heterosis, which is an important phenomenon to improve forest productivity, once it enables the combination of alleles present in different species (Goulet et al. 2017). The manifestation of heterosis on hybrids requires the choice of parents that will compose high-performance crosses (Malosetti et al. 2013). An efficient way to choose

the best parents is via progeny tests, that can be conducted in more than one environment, allowing to estimate the genotype by environment interaction (GEI) effect (Smith et al. 2001). After performing these tests, good analytical methods are fundamental to predict the performance of genotypes as accurately as possible (Araujo et al. 2019). The restricted maximum likelihood/best linear unbiased prediction (REML/BLUP) procedure (Henderson 1975; Patterson and Thompson 1971) is a powerful method for estimating genetic parameters and predicting genetic values in plant breeding (Xiang and Li 2003).

Genotype by environment interaction is the main reason for conduction multiple environment trials (MET), since stable genotypes for many environments are commonly desired. Some methodologies have been applied on plant breeding, such as harmonic mean of the relative performance of the genotypic values (HMRPGV) (Resende 2002), a method for ranking genotypes that considers simultaneously their stability, adaptability and productivity, and additive main effects and multiplicative interaction (AMMI) (Gauch Junior and Zobel 1988) a methodology that uses variance analysis (ANOVA) and principal component analysis (PCA) to decompose the source of variation into additive effects of genotypes and environments and multiplicative effects of GEI (Gauch Junior and Zobel 1988). Likewise, the GGE biplot (Yan et al. 2000) is based on ANOVA and PCA. It uses a site regression model allowing a graphical representation of the genotype and GEI interaction pattern on MET. Thus, the objectives of this study were to evaluate the genetic variability of 64 progenies of *C. maculata* and 64 progenies of *C. torelliana* and build the GGE biplot based on genetic values obtained via REML/BLUP procedure. Additionally, to point the more stable and productive genotypes to clone selection and the most appropriate crosses in order to generate interspecific hybrids of *C. maculata* and *C. torelliana*.

MATERIAL AND METHODS

Experimental design and genetic material

In December 2014, two half-sibling progeny trials were implemented in the Aracruz municipality, Espírito Santo state (19.8188°S, 40.2736°W, hereafter named ARA), the first one of *C. torelliana* and the second of *C. maculata*. Both trials were replicated in the Três Lagoas municipality, Mato Grosso do Sul state (20.7882°S, 51.7030°W, hereafter named TLA), totaling four trials. The ARA altitude is 60 m a.s.l. The soil is classified as medium-textured yellow latosol, flat relief and strongly drained (Embrapa 2000). According to Köppen and Geiger classification, updated by Kottek et al. (2006), the ARA climate is classified as tropical Aw, with an average temperature of 24.4 °C and an average annual rainfall of 1157 mm. The TLA altitude is 319 m a.s.l. The soil is classified as very porous dark-red latosol, strongly drained. According to Köppen and Geiger classification, updated by Kottek et al. (2006), TLA climate is classified as tropical Aw, with an average temperature of 24.7 °C and 1340 mm of average annual rainfall (Climate-data.org 2021).

The experimental design was an augmented block design, with 64 progenies of *Corymbia* and six commercial checks of *Eucalyptus*, totaling 70 treatments. It was used 40 repetitions. Into each repetition, the treatments were randomized into seven blocks, being each block with 10 treatments. In this design, each treatment (tree) represents a plot, being named single tree plot (STP) (Sandon 1958). The progenies for both species were codified from 1 to 70, being the treatments 65, 66, 67, 68, 69 and 70, the checks. At 29 months old, the circumference at breast height (cm), at 1.30 m from ground level, were measured on all trees using a flexible and graduated tape. The diameter at breast height (DBH) was obtained from the π division of the circumference at breast height.

STATISTICAL ANALYSES

Individual analyses

The restricted maximum likelihood (REML) (Henderson 1975) and best linear unbiased prediction (BLUP) (Patterson and Thompson 1971) procedures were used to estimate the components of variance and to predict the genetic values, respectively. The best progenies in each experiment were selected based on their predicted genetic values by the model (Eq. 1):

$$y = Xr + Za + Wb + e \quad (1)$$

where: y is the phenotypic data vector; r is the vector of the repetition effects added to the general mean and the checks effects (fixed); a is the vector of the additive genetic effects (assumed to be random, $a \sim \text{NID}(0, \sigma_a^2)$, where σ_a^2 is the additive genetic variance; b is the vector of the blocks effects (assumed to be random, $b \sim \text{NID}(0, \sigma_b^2)$, where σ_b^2 is the block variance; and e is the vector of residuals effects (random), ($e \sim \text{NID}(0, \sigma_e^2)$, where σ_e^2 is the residual variance). The capital letters (X , Z and W) represent the incidence matrices for r , a , and b effects, respectively.

Joint analysis

The best progenies of each species were selected based on their predicted genetic values by the model (Eq. 2)

$$y = Xr + Za + Wb + Ti + e \quad (2)$$

where: y is the phenotypic data vector; r is the vector of replication-environment combinations (assumed to be fixed), which comprises the effects of environment and replication within environment, added to the overall mean and the checks effects (fixed); a is the vector of the individual additive genetic effects (assumed to be random, $a \sim \text{NID}(0, \sigma_a^2)$; b is the vector of block effects (assumed to be random, $b \sim \text{NID}(0, \sigma_b^2)$; i is the vector of GEI effects (assumed to be random, $i \sim \text{NID}(0, \sigma_i^2)$, where σ_i^2 is the GEI variance; and e is the vector of residuals effects (random) - ($e \sim \text{NID}(0, \sigma_e^2)$). The capital letters (X , Z and W) represent the incidence matrices for r , a , and b effects, respectively.

Likelihood ratio test

The significance of the additive genetic effects and GEI effects were tested using the likelihood ratio test (LRT) (Wilks 1938) (Eq. 3):

$$LRT = -2(\text{Log}L_F - \text{Log}L_R) \quad (3)$$

where: $\text{Log}L_F$ is the logarithm of the restricted likelihood function of the full model, and $\text{Log}L_R$ is the logarithm of the restricted likelihood function of the reduced model, which is the one without the individual additive genetic effects or without the GEI effect. The models were compared at the level of 1% probability and one degree of freedom by the chi-square test.

Genetic parameters

Phenotypic variance (σ_{phen}^2), narrow-sense heritability (h_a^2), coefficient of determination of block effects (c_b^2), coefficient of determination of GEI effects (c_{ge}^2) and the genotypic correlation of genotypes among environments (r_{ge}) were obtained, respectively, by Eqs. 4–8.

$$\sigma_{phen}^2 = \sigma_a^2 + \sigma_b^2 + \sigma_{gm}^2 + \sigma_e^2 \quad (4)$$

$$h_a^2 = \sigma_a^2 / \sigma_{phen}^2 \quad (5)$$

$$c_b^2 = \sigma_b^2 / \sigma_{phen}^2 \quad (6)$$

$$c_{ge}^2 = \sigma_{ge}^2 / \sigma_{phen}^2 \quad (7)$$

$$r_{ge} = h_g^2 / (h_g^2 + c_{ge}^2) \quad (8)$$

The coefficient of genetic variation (CV_g) and the coefficient of experimental variation (CV_e) were obtained, respectively, by Eqs. 9 and 10:

$$CV_g = \frac{\sqrt{\sigma_g^2}}{\mu} * 100 \quad (9)$$

$$CV_e = \frac{\sqrt{\sigma_e^2}}{\mu} * 100 \quad (10)$$

where μ is the overall mean.

The standard errors of the estimates of variance components and of the genetic parameters were obtained through mixed model output and post-processing “pin” function of the ASReml software (Butler et al. 2017).

Accuracy was estimated by Eq. 11 (Resende et al. 2014):

$$r_{\hat{g}g} = \sqrt{1 - \frac{PEV}{\hat{\sigma}_g^2}} \quad (11)$$

where PEV is the prediction error variance, extracted from the diagonal of the generalized inverse of the coefficient matrix of the mixed model equation.

GGE biplot

The biplots were created by plotting the first principal component scores of the genotypes and the environments against their respective scores for the second principal component. These scores were obtained by singular value decomposition (SVD) of environment-centered data using Eq. 12, proposed by Yan et al. (2000).

$$y_{ij} - \bar{y}_j = \lambda_1 \xi_{1i} \eta_{1j} + \lambda_2 \xi_{2i} \eta_{2j} + e_{ij} \quad (12)$$

where: y_{ij} is the additive genetic values for genotype i at environment j ; \bar{y}_j is the overall additive genetic mean of all families in the environment j ; λ_1 and λ_2 are the singular values of the first and second principal components (PC1 and PC2, respectively); ξ_{1i} and ξ_{2i} are the values of eigenvectors for genotype i , for PC1 and PC2, respectively; η_{1j} and η_{2j} are the values of eigenvectors for environment j , for PC1 and PC2, respectively; and e_{ij} is the error associated with the y_{ij} observations. The biplot was based on singular value decomposition of the standardized table. Scaled by standard deviation = 1, centering = 2 and row metric preserving (SVP = 1) was used to generate the biplots (Frutos et al. 2014).

Progenies selection

The joint selection for both species was made based on the performance and stability of the families on both environments, which could be accessed on the “mean versus stability” biplot. Two likely approaches were considered, the first one would be cloning the more stable and high mean genotypes for planting as pure species. The second would be the generation of interspecific hybrids where two strategies could be adopted: choosing both stable and good average genotypes to maximize the expression of favorable genes to both environments, or crossing a stable genotype with high performance with a high mean genotype (without considering stability). Both approaches can favor the hybrid vigor expression due to allelic interaction.

The gain with selection (GS(%)) for the progenies, on the joint analysis were estimated by Eq. 13:

$$GS (\%) = ((X_{si} - X_{oi}) * 100) / X_{oi} \quad (13)$$

where: X_{si} is the mean of additive values for the selected progenies for DBH; X_{oi} is the mean of additive values for the original population for DBH.

Software

The analyses were performed using the Selegen REML/BLUP software (Resende 2016), and the GGEbiplotGUI (Frutos et al. 2014) and ASReML-R package (Butler et al. 2017) in R software (R Core Team 2020).

RESULTS

Likelihood ratio test

The LRT tested the null hypothesis, which assumes that there is no difference between the complete and reduced models by removing the additive genetic effect (individual and joint analyses) and the GEI effect (joint analysis). All effects were significant ($p < 0.01$) on both models and for both species, except the block effects (Table 1).

Table 1. Deviance and likelihood ratio test (LRT) for the progenies of *C. maculata* and *C. torelliana* for diameter at breast height (cm), planted on two environments.

| | Individual analyses | | | | Joint analysis | |
|-------------|---------------------|---------|----------------------|---------|--------------------|----------------------|
| | <i>C. maculata</i> | | <i>C. torelliana</i> | | <i>C. maculata</i> | <i>C. torelliana</i> |
| | ARA | TLA | ARA | TLA | | |
| CM | 6411.36 | 7187.23 | 4166.04 | 6303.27 | 13598.22 | 10760.55 |
| RM_g | 6848.50 | 7374.71 | 4639.25 | 6426.92 | 13610.99 | 10766.99 |
| RM_b | 6411.36 | 7187.35 | 4166.38 | 6303.83 | 13598.39 | 10761.58 |
| RM_{gxe} | | | | | 13744.68 | 10834.19 |
| LRT_g | 437.14* | 187.48* | 473.21* | 123.65* | 12.77* | 6.44* |
| LRT_b | 0 | 0.12 | 0.34 | 0.56 | 0.17 | 1.03 |
| LRT_{gxe} | | | | | 146.46* | 73.64* |

Note. * Significant at 5% probability level according to the Chi-square distribution. ARA is the Aracruz municipality and TLA is the Três Lagoas municipality. CM is the complete model, RM_g is the reduced model without the additive genetic effect, RM_b is the reduced model without the block effect, RM_{gxe} is the reduced model without GEI effect, and LRT_g , LRT_b and LRT_{gxe} are their respective likelihood ratio test.

Genetic parameters

The additive genetic and phenotypic variance were greater for *C. maculata* (Table 2). The narrow-sense heritability between progenies was 0.37 and 0.34 for *C. maculata* and *C. torelliana*, respectively. The coefficient of genetic variation (CV_g) and the coefficient of residual variation (CV_e) and the mean were greater for *C. maculata*. The genotypic correlation among environments was greater for *C. torelliana*. The selective accuracy (r_{gg}) was 0.77 and 0.79 for *C. maculata* and *C. torelliana*, respectively (Table 2).

Table 2. Variance components and genetic and non-genetic parameters for the joint analysis of *C. maculata* and *C. torelliana* progenies for diameter at breast height on two environments.

| Components/Parameters | <i>C. maculata</i> | <i>C. torelliana</i> |
|-----------------------|--------------------|----------------------|
| σ_a^2 | 4.22 | 1.11 |
| σ_b^2 | 0.04 | 0.02 |
| σ_{gxe}^2 | 1.11 | 0.23 |
| σ_e^2 | 6.07 | 1.86 |
| σ_{phen}^2 | 11.44 | 3.21 |
| h_{ab}^2 | 0.37 | 0.34 |
| c_b^2 | 0.00 | 0.01 |
| c_{gxe}^2 | 0.10 | 0.07 |
| r_{ge} | 0.47 | 0.55 |
| CV_g (%) | 15.55 | 9.82 |
| CV_e (%) | 18.67 | 12.73 |
| r_{gg} | 0.76 | 0.79 |
| μ | 13.20 | 10.71 |

Note. σ_a^2 is the additive genetic variance; σ_{gxe}^2 is the genotypes by environments interaction variance; σ_e^2 is the residual variance; σ_{phen}^2 is the phenotypic variance; h_{ab}^2 is the narrow-sense heritability between progenies; c_b^2 is the coefficient of determination of blocks; c_{gxe}^2 is the coefficient of determination of genotypes by environments interactions; r_{ge} is the genotypic correlation between performance in different environments; CV_g (%) is the coefficient of genetic variation in percentage, CV_e (%) is the coefficient of residual variation in percentage, r_{gg} is the selective accuracy, and μ is the overall mean.

GGE biplot and genetic selection

For *C. maculata* and *C. torelliana*, the PC1 and the PC2 explained a total of 100% of genotype and GEI variation (Fig. 1). For both, *C. maculata* and *C. torelliana*, the “which-won-where” biplot divided the graph into ten and six distinct environments, respectively (Fig. 1). For both species, the studied environments (ARA and TLA) were in different sectors, meaning that they cannot be clustered into a single environment. For each polygon, the genotypes in each vertex represents those with best performance in each sector. For *C. maculata*, the progeny 10 was the one that had better performance in ARA and the progeny 42 was the one that had best performance on TLA (Fig. 1). For *C. torelliana*, no families have stood out on ARA; however, family 18 was the one that had best performance on TLA (Fig. 1).

The “mean versus stability” graph allows the visualization of the mean and stability of genotypes in all environments (Fig. 2). An average environment is calculated from the mean of PC1 and PC2 scores for all environments and is represented by a red point. The average environment abscissa (AEA) is represented by a single head arrow that passes through the biplot origin and the average environment and is pointing towards higher yield. The approximated mean of each evaluated genotype can be estimated by projecting the genotypes onto the average environment abscissa (Yan 2002). The narrow is pointed to the higher mean performance; therefore, it is possible to rank the best genotypes. For *C. maculata*, the five higher mean progenies were 42, 14, 38, 16, and 21 (Fig. 2). For *C. torelliana*, the progenies 18, 55, 13, 16, and 2 presented the higher means (Fig. 2). The average environment ordinate (AEO) is represented as a double-headed arrow perpendicular to AEA abscissa. It is pointed towards lower stability in both directions and associates the GEI with each genotype. Thus, greater projection onto the AEA ordinate, means greater instability of the genotypes, regardless of the direction. Therefore, for *C. maculata*, the five more stable genotypes were 38, 2, 61, 34, 54 and 25 (Fig. 2). For *C. torelliana*, the five more stable genotypes were 16, 4, 26, 48 and 14, once they are on the top of the average environment abscissa (Fig. 2).

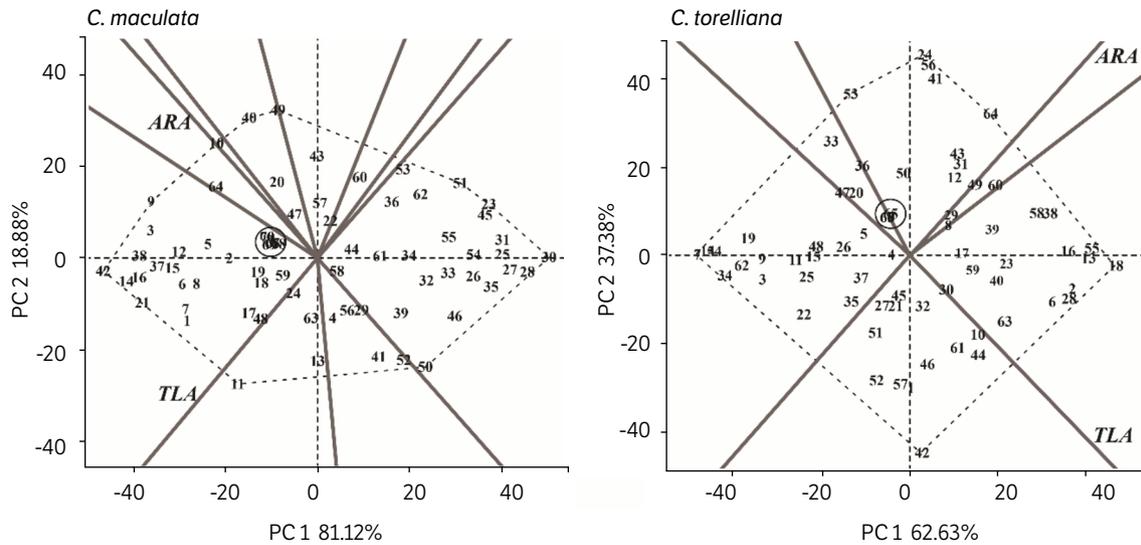


Figure 1. GGE biplot representing the “which-won-where” graph for diameter at breast height (cm) for 64 progenies of *C. torelliana* and 64 progenies of *C. maculata* planted on two environments (ARA, TLA).

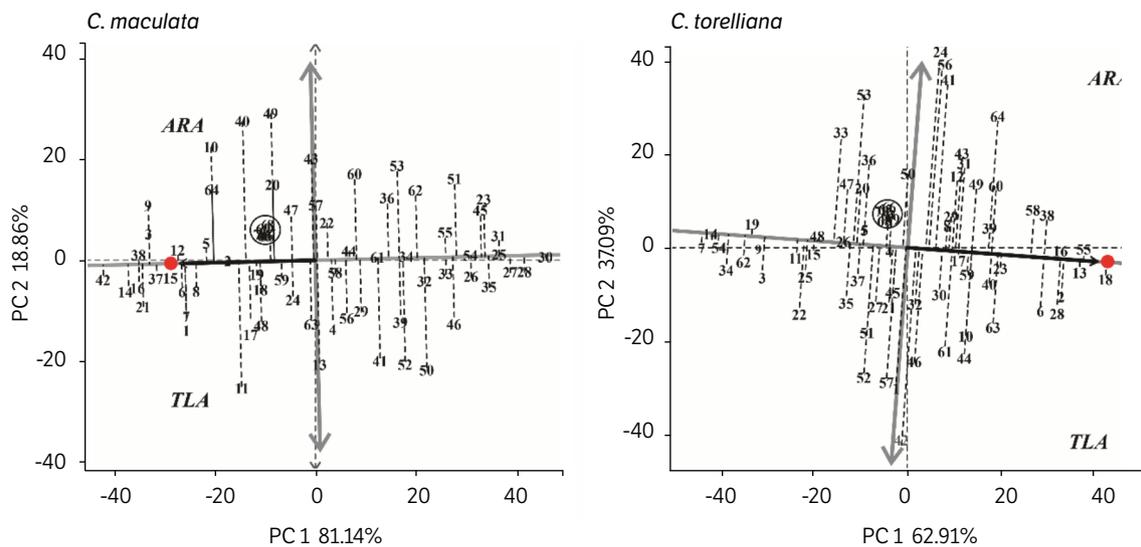


Figure 2. GGE biplot representing “mean versus stability” graph for diameter at breast height (cm) for 64 progenies of *C. torelliana* and 64 progenies of *C. maculata* planted on two environments (ARA, TLA).

Note. The black circles and the red points indicate the group of check and the average environments, respectively.

By adopting a selection intensity of 10%, it is possible to indicate the potential genotypes to be selected. In case the objective is cloning the best individuals for planting as pure species in both environments, it is recommended to select genotypes that simultaneously have high means and good stability, once the desire is to maximize genetic gains (Borém et al. 2013). In this scenario, for *C. maculata*, the progenies 42, 38, 16, 14, 37, 15 and 12 and, for *C. torelliana*, the progenies 18, 55, 13, 16, 23, 17 and 59 could be selected (Fig. 2). However, if the objective is to generate interspecific hybrids it is recommended to cross more genetically distant progenies in order to expand the possibility of high-performance individuals in the population due to heterosis (Borém et al. 2013). Therefore, in the ARA environment, the progenies 10, 40, 49, 64, 20, 9 and 47 of *C. maculata* can be crossed with the progenies 24, 56, 41, 64, 43, 31 and 12 of *C. torelliana* (Fig. 2). On the other hand, in the TLA environment, the progenies 11, 17, 1, 48, 7, 63 and 21 of *C. maculata* can be crossed with the progenies 42, 46, 61, 44, 10, 63 and 28 of *C. torelliana* (Fig. 2).

The biplot indicated 32 and 33 progenies with genetic values above average for *C. maculata* and *C. torelliana*, respectively (Fig. 2). The gain with selection (%) was 14.33% for *C. maculata* and 6.63% for *C. torelliana* when the best 20 families were selected, which is recommended by Resende (2002) to start a breeding program. The representativeness of a given environment was measured by the angle between the vector of a given environment and the AEC axis (the line that passes through the biplot origin and the average environment). The average environment is represented by a star (Fig. 3). Hence, both ARA and TLA were not representative. The “ideal” environment is represented by a point on the AEC axis, with an arrow point to it. In theory, it is the most discriminating and representative environment. It can be seen that ARA and TLA are not far from the “ideal” environment concentric line, meaning that they are discriminative (Fig. 3).

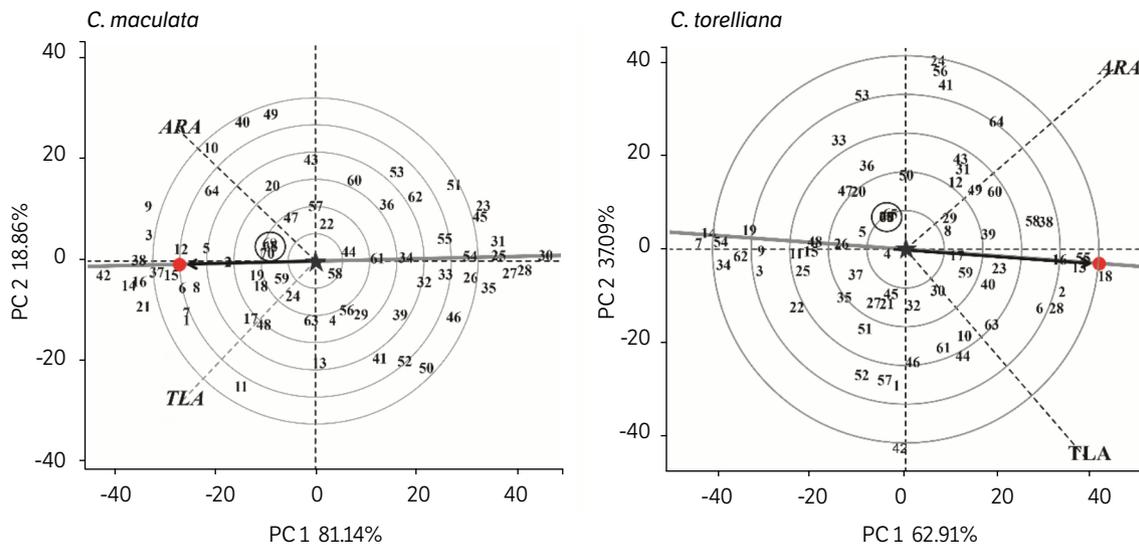


Figure 3. GGE biplot representing discrimination and representativeness of environments for diameter at breast height (cm) for 64 progenies of *C. torelliana* and 64 progenies of *C. maculata* planted on two environments (ARA, TLA).

Note. The star and the red point indicate the average and “ideal” environment, respectively.

DISCUSSION

Likelihood ratio test

The significance of the additive genetic effects ($p < 0.01$) indicates that there is genetic variability between the progenies for both species, allowing the genetic selection (Zimback et al. 2011). Moreover, there is significant GEI, which can difficult the selection of the same genotypes for both environments. One strategy to deal with the GEI effect is to choose genotypes with high adaptability and stability for the environments analyzed.

Genetic parameters

The narrow-sense heritability between progenies (h_{ab}^2) values for DBH in half-sibling progenies of *Corymbia* was lower than the ones reported by Araujo et al. (2020), which varied from 0.41 to 0.69. According to Lima et al. (2019), the traditional mixed model methodology commonly overestimate genetic parameters estimates and consequently narrow-sense heritability for growth traits in comparison to methodologies that consider the genomic relationship matrix. Despite this fact, the heritability values obtained by the present study was similar to those observed by genomic studies for half-sibling families, which varied from 0.29 to 0.48 (Tambarussi et al. 2018).

The coefficient of genetic variation allowed to make inferences about the magnitude of the variability present in the populations, once it is a measure of the fraction of phenotypic variation that is genetic, normalized by the mean (Downie et al., 2020). The lowest coefficient of genetic variation was obtained for the *C. torelliana* (9.82 %). According to Resende (2002) coefficients of genetic variation of around 10% or more are sufficient for practicing effective selection. The CVe (%) was classified as high magnitude for *C. maculata* and very high for *C. torelliana* (Mora et al. 2016), which is common in progeny trials of forest species (Bertii et al. 2011; Costa et al. 2015). The selective accuracy is a measure of the proximity between the predicted and true genetic values. The greater the accuracy, the greater is the reliability on the selection, because the deviations between the parametric and predicted genetic values are smaller. The selective accuracy for DBH is greater than 0.7, being classified as high magnitude (Resende and Alves 2020). Thus, there is a favorable scenario for the genetic selection for both species.

GGE biplot and selection

Commonly, high productive genotypes are preferred by breeders (Bornhofen et al. 2017). In the selection of parents for hybridization it is not different (Assis 2000). The GGE biplot methodology selects based on the progeny pattern that considers mean and stability (Santos et al. 2016). It extracts information about GEI and crossover interaction in multi environment trials and display it on a simple, elegant and visual way. Thus, GGE biplot would be considered a suitable methodology to study GEI.

Understanding GEI is crucial for the progress of any breeding program (van Eeuwijk et al., 2016). The “which-won-where” patterns of an MET (Fig. 1) allow the identification of more responsive families for each environment, which are the ones located farthest from the origin (Yan et al. 2007). The genotypes close to the origin are less responsive in their respective directions (Yan and Kang 2002). The progenies that are not placed into any environment are considered unfavorable to the tested environments (Karimizadeh et al. 2013). The biplot is divided into sectors by perpendicular lines. Each sector has a vertex genotype, which mean that it presents the highest mean for that particular environment (Yan and Kang 2002). Additionally, the large number of mega environments formed indicates the genetic potential of these families to be used in other environments not yet tested.

The first two principal components account for 100% of the total variation for both species, indicating that the GGE biplot methodology was efficient on explaining the variation of genotypes and the GEI effects (Yan and Holland 2010). In addition, the right angle formed between the two environments on the GGE biplot suggests moderate correlation between the environments (Yan and Holland 2010), indicating the impossibility of forming mega-environments containing ARA and TLA. Therefore, they belong to different agronomical areas, which can be explained due to the different geographic coordinates, altitudes and edaphoclimatic conditions among the two environments. Additionally, they have significant GEI, which explain the changes in the ranking of genotypes from one environment to another (Alwala et al. 2010).

Environment ARA and TLA were not the best environments for selecting superior cultivars, once they are not representative, what explains why most of the genotypes were not allocated in these environments (Fig. 1). However, according to Yan and Kang (2002), they are discriminating environments, being useful for selection stable cultivars, which was an important task for this work. For future works, it is recommended to use a larger number of environments to increase the representativeness.

This work aimed to select progenies to be planted on both environments, therefore it seeks to avoid the GEI by selecting based on mean and stability (Yan and Kang 2002). Stability is often overlooked by plant breeders due to the difficulty to weigh between mean performance and stability. However, some investigation has been made considering it on different cultures and using different methods, such as HMRPGV (Ferreira et al. 2021), AMMI (Malosetti et al. 2013) and GGE biplot (Xu et al. 2014). The advantage of the GGE biplot is the graphical representation of the genotype and GEI patterns on multi environmental trials, considering stability. Despite that, no work has been found linking GGE biplot analysis and *Corymbia* progenies in MET.

Therefore, in possession of the selected progenies, an interesting alternative for the future would be starting an interpopulation recurrent selection (Comstock et al. 1949) aiming at the improvement of both species and the increase of heterosis between them. In this way, it will be possible to obtain, assess and recombine the best individuals, improving

crosses between *C. torelliana* and *C. maculata* populations (Chen et al. 2013). In these crosses, *C. torelliana* is a better fit to be the female genitor, since this specie has a 30% higher rooting capacity compared to many other species of the genus (Reis et al. 2014). Therefore, exploiting this advantage can increase the chance of crossbreeding and hybrid propagation.

CONCLUSION

There is genetic variability for diameter at breast height for *C. maculata* and *C. torelliana* suggesting a positive scenario for breeding.

The GGE biplot analysis discriminate the more stable and high productive progenies of each specie in both environments. The patterns provided by the biplot allowed to point the best progenies to clone selection and the best progenies to be crossed in order to generate interspecific hybrids of *C. maculata* and *C. torelliana*.

AUTHORS' CONTRIBUTION

Conceptualization: Ferreira, F. M., Bhering L. L. and Aguiar, A. M.; **Methodology:** Ferreira, F. M., Rocha, J. R. A. S. C., Alves, R. S., Peixoto, M. A. and Malikuski, R. G.; **Investigation,** Ferreira, F. M. and Rocha, J. R. A. S. C.; **Writing – Original Draft:** Ferreira, F. M., Oliveira, S. S. and Malikuski, R. G.; **Writing – Review and Editing:** Ferreira, F. M., Rocha, J. R. A. S. C., Alves, R. S., and Peixoto, M. A.; **Funding Acquisition:** Bhering, L. L.; **Resources:** Aguiar, A. M.; **Supervision:** Bhering, L. L.

DATA AVAILABILITY STATEMENT

The data will be available upon request.

FUNDING

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
[<https://doi.org/10.13039/501100002322>]
Finance Code 001

Suzano Celulose S/A

ACKNOWLEDGMENTS

Not applicable.

REFERENCES

Alwala, S., Kwolek, T., McPherson, M., Pellow, J. and Meyer, D. (2010). A comprehensive comparison between Eberhart and Russell joint regression and GGE biplot analyses to identify stable and high yielding maize hybrids. *Field Crops Research*, 119, 225-230. <https://doi.org/10.1016/j.fcr.2010.07.010>

- Araujo, M. J., Paula, R. C., Campoe, O. C. and Carneiro, R. L. (2019). Adaptability and stability of eucalypt clones at different ages across environmental gradients in Brazil. *Forest Ecology and Management*, 454, 117631. <https://doi.org/10.1016/j.foreco.2019.117631>
- Araujo, M. J., Lee D. J., Tambarussi, E. V., Paula, R. C. and Silva, P. H. M. (2020). Initial productivity and genetic parameters of three *Corymbia* species in Brazil, designing a breeding strategy. *Canadian Journal of Forest Research*, 51, 25-30. <https://doi.org/10.1139/cjfr-2019-0438>
- Assis, T. F. (2000). Production and use of *Eucalyptus* hybrids for industrial purposes [Paper presentation]. In Hybrid breeding and genetics of forest trees. Proceedings of QFRI/CRCOSPF Symposium, 9-14 April. Noosa, Queensland, Australia.
- Bertii, C. L. F., Freitas, M. L. M., Zanatto, A. C. S., Moraes, E., Moraes, M. L. T. and Sebbenn, A. M. (2011). Variação genética, herdabilidades e ganhos na seleção para caracteres de crescimento e forma em teste de progênes de polinização aberta de *Eucalyptus cloeziana*. *Revista do Instituto Florestal*, 23, 13-26.
- Borém, A., Miranda, G. V., Fritsche-Neto, R. (2013). *Melhoramento de Plantas*. Viçosa: UFV.
- Bornhofen, E., Benin, G., Storck, L., Woyann, L. G., Duarte, T., Stoco, M. G. and Marchioro, S. V. (2017). Statistical methods to study adaptability and stability of wheat genotypes. *Bragantia*, 76, 1-10. <https://doi.org/10.1590/1678-4499.557>
- Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J. and Thompson, R. (2017). *ASReml-R Reference Manual Version 4*. Hemel Hempstead: VSN International.
- Carle, J. B., Duval, A. and Ashford, S. (2020). The future of planted forests. *International Forestry Review*, 22, 65-80.
- Climate-data.org. (2021). Climate Três Lagoas (Brazil). Climate-Data.org. [Accessed Jul. 16, 2021]. Available at: <https://en.climate-data.org/south-america/brazil/mato-grosso-do-sul/tres-lagoas-34651/>
- Comstock, R. E., Robinson, H. F. and Harvey, P. H. (1949). A breeding procedure designed to make maximum use of both general and specific combining ability. *Agronomy Journal*, 41, 360-367. <https://doi.org/10.2134/agronj1949.00021962004100080006x>
- Costa, R. B., Martinez, D. T., Silva, J. C. and Almeida, B. C. (2015). Variabilidade e ganhos genéticos com diferentes métodos de seleção em progênes de *Eucalyptus camaldulensis*. *Revista de Ciências Agrárias*, 58, 69-74. <https://doi.org/10.4322/rca.1668>
- Darrow, W. K. (1985). Provenance Trials of *Eucalyptus maculata* In South Africa: Tenth Year Results. *South African Forestry Journal* 133, 12-17. <https://doi.org/10.1080/00382167.1985.9629567>
- Dickinson, G. R., Wallace, H. M. and Lee, D. J. (2013). Reciprocal and advanced generation hybrids between *Corymbia citriodora* and *Corymbia torelliana*: forestry breeding and the risk of gene flow. *Annals of Forest Science*, 70, 1-10. <https://doi.org/10.1007/s13595-012-0231-2>
- Downie, J., Silvertown, J., Cavers, S. and Ennos, R. (2020). Heritable genetic variation but no local adaptation in a pine-ectomycorrhizal interaction. *Mycorrhiza*, 30, 185-195. <https://doi.org/10.1007/s00572-020-00941-3>
- [Embrapa] Empresa Brasileira de Pesquisa Agropecuária. (2000). Levantamento generalizado e semi detalhado de solos da Aracruz Celulose S.A. no estado do Espírito Santo e no extremo sul do Estado da Bahia e sua aplicação aos plantios de eucalipto. Rio de Janeiro: Embrapa Solos.
- Ferreira, F. M., Rocha, J. R. A. S. C., Bhering, L. L., Fernandes, F. D., Lédo, F. J. S., Rangel, J. H. A., Kopp, M., Câmara, T. M. M., Silva, V. Q. R. and Machado, J. C. (2021). Optimal harvest number and genotypic evaluation of total dry biomass, stability and adaptability of elephant grass clones for bioenergy purposes. *Biomass and Bioenergy*, 149, 106104. <https://doi.org/10.1016/j.biombioe.2021.106104>
- Frutos, E., Galindo, M. P. and Leiva, V. (2014). An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stochastic Environmental Research and Risk Assessment*, 28, 1629-1641. <https://doi.org/10.1007/s00477-013-0821-z>
- Gauch Junior, H. G. and Zobel, R. W. (1988). Predictive and postdictive success of statistical analyses of yield trials. *Theoretical and Applied Genetics*, 76, 1-10. <https://doi.org/10.1007/BF00288824>

- Goulet, B. E., Roda, F. and Hopkins, R. (2017). Hybridization in Plants, Old Ideas, New Techniques. *Plant Physiology*, 173, 65-78. <https://doi.org/10.1104/pp.16.01340>
- Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423-447. <https://doi.org/10.2307/2529430>
- Karimizadeh, R., Mohammadi, M., Sabaghni, N., Mahmoodi, A. A., Roustami, B., Sayyedi, F. and Akbari, F. (2013). GGE Biplot Analysis of Yield Stability in Multi-environment Trials of Lentil Genotypes under Rainfed Condition. *Notulae Scientia Biologicae*, 5, 256-262. <https://doi.org/10.15835/nsb529067>
- Kottek, M., Grieser, J., Beck, C., Rudolf, B. and Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15, 259-263. <https://doi.org/10.1127/0941-2948/2006/0130>
- Lee, D. J. (2007). Achievements in forest tree genetic improvement in Australia and New Zealand 2, Development of *Corymbia* species and hybrids for plantations in eastern Australia. *Australian Forestry*, 70, 11-16. <https://doi.org/10.1080/00049158.2007.10676256>
- Lee, D. J., Huth, J. R., Brawner, J. T. and Dickinson, G. R. (2009). Comparative performance of *Corymbia* hybrids and parental species in subtropical Queensland and implications for breeding and deployment. *Silvae Genetica*, 58, 205-212. <https://doi.org/10.1515/sg-2009-0026>
- Lima, B. M., Cappa, E. P., Silva-Junior, O. B., Garcia C., Mansfield S. D. and Grattapaglia D. (2019). Quantitative genetic parameters for growth and wood properties in *Eucalyptus* “urograndis” hybrid using near-infrared phenotyping and genome-wide SNP-based relationships. *PLoS ONE*, 14, e0218747. <https://doi.org/10.1371/journal.pone.0218747>
- Loureiro, B. A., Vieira, T. A. S., Costa, L. J., Silva, A. B., Assis, M. R. and Trugilho, P. F. (2019). Selection of superior clones of *Corymbia* hybrids based on wood and charcoal properties. *Maderas. Ciencia y Tecnología*, 21, 619-630. <https://doi.org/10.4067/S0718-221X2019005000417>
- Malosetti, M., Ribaut, J.-M. and van Eeuwijk, F. A. (2013). The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology*, 4, 44. <https://doi.org/10.3389/fphys.2013.00044>
- Mora, F. and Arriagada, O. (2016). A classification proposal for coefficients of variation in *Eucalyptus* experiments involving survival, growth and wood quality variables. *Bragantia*, 75, 263-267. <https://doi.org/10.1590/1678-4499.458>
- Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58, 545-554. <https://doi.org/10.1093/biomet/58.3.545>
- Payn, T., Carnus, J.-M., Freer-Smith, P., Kimberley, M., Kollert, W., Liu, S., Orazio, C., Rodriguez, L., Silva, L. N. and Wingfield, M. J. (2015). Changes in planted forests and future global implications. *Forest Ecology and Management*, 352, 57-67. <https://doi.org/10.1016/j.foreco.2015.06.021>
- R Development Core Team (2020). R: A language and environment for statistical computing [Computer software]. R Foundation.
- Reis, C. A. F., Assis, T. F., Santos, A. M. and Paludzyszyn Filho, E. (2014). *Corymbia maculata*, estado da arte de pesquisas no Brasil. Colombo: Embrapa Florestas.
- Resende, M. D. V. (2002). Genética biométrica e estatística no melhoramento de plantas perenes. Brasília: Embrapa.
- Resende, M. D. V., Silva, F. F. and Azevedo, C. F. (2014). Estatística matemática, biométrica e computacional: modelos mistos, multivariados, categóricos e generalizados (REML/BLUP), inferência bayesiana, regressão aleatória, seleção genômica, QTI-GWAS, estatística espacial e temporal, competição, sobrevivência. Viçosa: UFV.
- Resende, M. D. V. (2016). Software Selegen-REML/BLUP, a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology* 16, 330-339. <https://doi.org/10.1590/1984-70332016v16n4a49>
- Resende, M. D. V. and Alves, R. S. (2020). Linear, generalized, hierarchical, Bayesian and random regression mixed models in genetics/genomics in plant breeding. *Functional Plant Breeding Journal*, 2, 1. <https://doi.org/10.35418/2526-4117/v2n2a1>

- Sandon, F. (1958). *Experimental Designs*. By W. G. Cochran and G. M. Cox. 2nd ed. Pp. xiv, 611. 82s. 1957. (John Wiley and Sons, New York; Chapman and Hall, London). *The Mathematical Gazette*, 42, 334. <https://doi.org/10.2307/3610494>
- Santos, A., Ceccon, G., Teodoro, P. E., Correa, A. M., Alvarez, R. C. F., Silva, J. F., Alves, V. B. (2016). Adaptability and stability of erect cowpea genotypes via REML/BLUP and GGE Biplot. *Bragantia*, 75, 299-306. <https://doi.org/10.1590/1678-4499.280>
- Sato, A. S., Freitas, M. L. M., Lima, I. L., Zimback, L., Toniato, M. T. Z. and Sebbenn, A. M. (2010). Genetic variation among and within provenances and progenies of *Corymbia maculata* (Hook.) K. D. Hill and L. A. S. Johnson, in Pederneiras, SP. *Cerne*, 16, 60-67. <https://doi.org/10.1590/S0104-77602010000100007>
- Shepherd, M., Pomroy, P., Dieters, M. and Lee, D. (2007). Genetic control of propagation traits in a single *Corymbia torelliana* × *Corymbia variegata* family. *Canadian Journal of Forest Research*, 37, 2563-2574. <https://doi.org/10.1139/X07-111>
- Smith, A., Cullis, B. and Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*, 57, 1138-1147. <https://doi.org/10.1111/j.0006-341X.2001.01138.x>
- Tambarussi, E. V., Pereira, F. B., Silva, P. H. M., Lee, D. and Bush, D. (2018). Are tree breeders properly predicting genetic gain? A case study involving *Corymbia* species. *Euphytica*, 214, 150. <https://doi.org/10.1007/s10681-018-2229-9>
- Van Eeuwijk, F. A., Bustos-Korts, D. V. and Malosetti, M. (2016). What should students in plant breeding know about the statistical aspects of genotype × Environment interactions? *Crop Science*, 56, 2119-2140. <https://doi.org/10.2135/cropsci2015.06.0375>
- Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, 9, 60-62. <https://doi.org/10.1214/aoms/1177732360>
- Xiang, B. and Li, B. (2003). Best linear unbiased prediction of clonal breeding values and genetic values from full-sib mating designs. *Canadian Journal of Forest Research*, 33, 2036-2043. <https://doi.org/10.1139/x03-118>
- Xu, N.-y., Fok, M., Zhang, G.-w., Li, J. And Zhou, Z.-g. (2014). The Application of GGE Biplot Analysis for Evaluating Test Locations and Mega-Environment Investigation of Cotton Regional Trials. *Journal of Integrative Agriculture*, 13, 1921-1933. [https://doi.org/10.1016/S2095-3119\(13\)60656-5](https://doi.org/10.1016/S2095-3119(13)60656-5)
- Yan, W., Hunt, L. A., Sheng Q. and Szlavnicz Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Science*, 40, 597-605. <https://doi.org/10.2135/cropsci2000.403597x>
- Yan, W. and Kang, M. S. (2002). *GGE Biplot Analysis: A graphical tool for breeders, geneticists, and agronomists*. Taylor & Francis, London, 271p
- Yan, W., Kang, M. S., Ma B., Woods S. and Cornelius P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Science*, 47, 643-653. <https://doi.org/10.2135/cropsci2006.06.0374>
- Yan, W. and Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica*, 171, 355-369. <https://doi.org/10.1007/s10681-009-0030-5>
- Chen, Z.-h., Zhu, Y.-f., Wang, A.-g., Guo, X.-y., Zhao, L. and Hu, X. (2013). Two Maize Populations of Tuxpeno-Reid and Suwan-Lancaster by Reciprocal Recurrent Selection and the Heterosis. *Journal of Maize Sciences*, 10, 5-10.
- Zimback, L., Mori, E. S., Brizolla, T. F. and Chaves, R. (2011). Correlações entre caracteres silviculturais durante o crescimento de *Eucalyptus grandis* Hill ex Maiden. *Revista do Instituto Florestal*, 23, 57-67. <https://repositorio.unesp.br/handle/11449/141157>