Acta Scientiarum

http://www.periodicos.uem.br/ojs/
ISSN on-line: 1807-8621
Doi: 10.4025/actasciagron.v46i1.66986

Genetic characterization of *Zeyheria tuberculosa* progenies and evaluation for formation of a seed orchard

Raul Reis Assunção[1], Adelson Lemes da Silva Júnior[1*], Rodolfo Soares de Almeida[2], Dulcinéia de Carvalho[1] and Lucas Amaral de Melo[1]

[1]Departamento de Ciências Florestais, Escola de Ciências Agrárias de Lavras, Universidade Federal de Lavras, Av. Doutor Sylvio Menicucci, 1001, Aquenta Sol, 37200-000, Lavras, Minas Gerais, Brazil. [2]Departamento de Engenharia Florestal, Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil. *Author for correspondence. E-mail: adelsonlemes@yahoo.com.br

ABSTRACT. *Zeyheria tuberculosa*, a native species of Brazil known for its significant potential in silviculture and genetic improvement, holds prominence among various species. In this study, our objective was to assess the diversity, genetic structure, and feasibility of establishing a seedling seed orchard (SSO) for this species. A total of 71 progenies were collected from different locations and were used in our experiment in Ijaci - MG. We genotyped 92 individuals (nine families with eight individuals, two families with seven individuals, and one family with six individuals), specifically selecting those with the highest predicted genetic values, using ten ISSR primers. The molecular markers employed effectively detected polymorphism (PIC = 0.44). The population exhibited moderate to high genetic diversity, as evidenced by observed (A_O = 2.00) and effective alleles (A_E = 1.61), Nei's diversity index (H^* = 0.35), and Shannon's diversity index (I^* = 0.52). Molecular variance analysis indicated significant genetic differentiation between the progenies (Φst = 0.19), yet the majority of the variation was observed within them (80.1%). Employing a Bayesian approach, we identified the formation of two distinct genetic groups, further confirming the non-genetic structure of the population. These findings affirm the potential of the *Z. tuberculosa* progenies to contribute to the establishment of a seedling seed orchard, supporting genetic improvement strategies and the conservation of the species' genetic diversity.

Keywords: ipê-felpudo; genetic improvement; molecular marker; ISSR; seedlings seed orchard.

Received on February 3, 2023.
Accepted on June 9, 2023.

Introduction

Brazil, renowned for its remarkable tree species diversity, boasts 8,715 registered forest species (Beech, Rivers, Oldfield, & Smith, 2017). However, the country's forestry sector predominantly focuses on two exotic genera, namely *Pinus* spp. and *Eucalyptus* spp., which account for 9.17 million hectares (96%) of commercial forests (Indústria Brasileira de Árvores [IBÁ], 2021). Consequently, there exists untapped potential for native forest species to assume a more prominent role in Brazilian forestry.

Among the diverse native species with considerable potential for silvicultural and genetic improvement, *Zeyheria tuberculosa* (Vell.) Bureau ex Verl., belonging to the Bignoniaceae family, holds significant commercial value. Regrettably, due to historical unsustainable exploitation in native forests, it is classified as vulnerable on the list of endangered plants (Brasil, 2021). Unregulated exploitation of *Z. tuberculosa* has the potential to disrupt gene flow patterns and diminish its effective population size (Ne), ultimately jeopardizing its survival (El-Kassaby et al., 2019).

Z. tuberculosa primarily occurs in the Southeast region of Brazil, with additional records in the Northeast and Midwest (Lohmann, 2015). This semi-deciduous tree reaches heights of 6 to 35 meters and diameters at breast height (DBH) ranging from 30 to 90 centimeters in adulthood (Carvalho, 2005). Its flowering period spans from October to January, while fruit ripening takes place between May and June in Minas Gerais (Lopes, Ferreira, & Brandão, 1996). Being a hermaphroditic plant, it relies on bee pollination (Souza, Nepi, Machado, & Guimarães, 2017) and enters the reproductive stage at three years of age (Carvalho, 2005).

Z. tuberculosa possesses several favorable characteristics for domestication, including rapid growth (1 to 2 meters per year), substantial size, pronounced monopodial growth, natural thinning ability, competitive edge against other pioneer species, high plasticity, easy seed propagation, impressive regrowth rates, and

recognized timber value (Mendonça et al., 2018; Magistrali, Cacau, Nascimento, & Magistrali, 2022). Its wood finds utility in internal works, civil construction, sleepers, tool handles, and agricultural instruments (Lohman, 2015).

Despite the species' significance, scientific studies on the genetic improvement of *Z. tuberculosa* remain limited, and no genetic characterization has been conducted thus far. The closest related work pertains to the genetic diversity of a species from the same genus, namely *Zeyheria montana*, using Random Amplified Polymorphic DNA (RAPD) markers (Bertoni, 2007), which underscores the relevance of our present study.

Inter Simple Sequence Repeats (ISSR) markers are commonly employed in genetic studies of forest populations to assess genetic diversity (Brandão, Vieira, Nazareno, & Carvalho, 2015; Silva Júnior, Souza, Pereira, Caldeira, & Miranda, 2017; Teixeira et al., 2020). They serve as a foundation for studies and programs aimed at producing seeds and clones of superior genetic quality. Despite being dominant markers, ISSRs are considered universal, allowing the genetic characterization of species lacking specific molecular markers, as is the case for *Z. tuberculosa*. Furthermore, ISSRs exhibit high informativeness and reproducibility (Idrees & Irshad, 2014), enabling advancements in genetic studies of native forest species. Against this backdrop, our objective was to quantify the diversity and genetic structure of *Z. tuberculosa* in a progeny test during the juvenile phase and assess its potential for future conversion into seedlings seed orchard (SSO).

Material and methods

Experimental area

The experimental area encompassed approximately 1 hectare and was situated in the Municipality of Ijaci, Minas Gerais State, Brazil, at the Palmital farm (Figure 1) of the *Universidade Federal de Lavras* (UFLA), with coordinates 21º09'3" S and 44º55'55" W. The experiment consisted of 71 progenies arranged in a randomized block design with 15 replications, resulting in a total of 1,065 individual plots. The spacing between progenies was 3 x 3 meters. To enhance height growth, minimize branching, and stem tortuosity, the plants were intercropped with three other fast-growing native forest species, namely *Peltophorum dubium*, *Inga edulis*, and *Guazuma ulmifolia*.

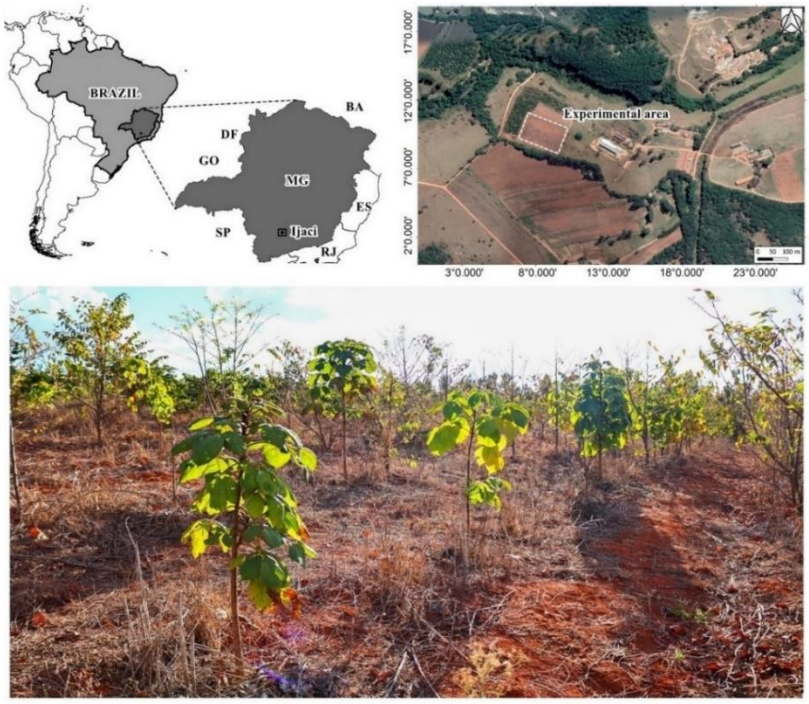

Figure 1. Location map of the experimental area, where the *Zeyheria tuberculosa* progeny test was conducted, situated at the Palmital farm in Ijaci, Minas Gerais State, Brazil.

The experimental area falls within an ecotone zone between the Atlantic Forest and Cerrado biomes, characterized as Semideciduous Seasonal Forest. The soil type is Dark Red Latosol, and the average altitude is 858 meters. As per the Köppen-Geiger classification, the climate is classified as Cwa, representing a

subtropical climate with rainy summers and dry winters. The average annual temperature is 19.9ºC, with variations ranging from 16.9ºC in June and July to 22.8ºC in February. The average annual precipitation is 1486 mm (Instituto Nacional de Meteorologia [INMET], 2021).

Genetic material

The *Zeyheria tuberculosa* seeds were collected from selected trees exhibiting desirable phenotypic traits such as straightness, stem shape, canopy symmetry, and overall plant health. The seeds were sourced from various locations in Minas Gerais State, including Nepomuceno, Itumirim, Carmo da Mata, Mariana, Barra Longa, and Ponte Nova. To ensure genetic diversity, a minimum distance of 50 meters was maintained between the mother trees, and the coordinates of each tree were recorded using a global positioning system (GPS).

Seedlings were produced at the Forestry Nursery of the *Universidade Federal de Lavras* (UFLA). The seeds were extracted from the fruits and air-dried. The substrate consisted of 30% coconut fiber, 30% manure, 20% carbonized rice husks, 20% subsoil earth, and 3 kg of osmocote (15:9:12) per cubic meter. Sowing was carried out in 180 cm^3 tubes placed in trays on suspended beds, receiving irrigation treatments every three hours for five minutes per period. After 60 days, the tubes were alternated to optimize sunlight exposure.

Silvicultural treatments

Prior to planting, the entire area underwent heavy harrowing as a site preparation measure. Planting took place in February 2020. Fertilization was conducted based on eucalyptus recommendations since specific nutritional recommendations for *Z. tuberculosa* are yet to be established. The fertilizers used included 50 g of simple superphosphate at the bottom of each planting hole, 100 g of NPK (6-30-6) in a lateral pit 15 days after planting, and two topdressing applications of 50 g of NPK (20-5-20) during the rainy season in the first year, with additional fertilization in the second year. Ant control measures were implemented before, during, and after planting, employing ant bait near the holes. Ant monitoring rounds were conducted every two months. Weed control was conducted using pre/post-emergent herbicides and mechanical mowing.

Forest measurement and selection of individuals for molecular analysis

At 15 months of age, silvicultural traits related to vegetative growth were measured, including plant height (Ht; in meters) using a ruler, circumference at ground height (CGH; in centimeters) using a tape measure, survival (S; in percentage), and number of branches.

Following data collection, genetic parameters were estimated using the mixed linear model methodology through REML/BLUP (Restricted Maximum Likelihood/Best Linear Unbiased Prediction) with the SELEGEN software (Resende, 2016). Based on the predicted genetic values, 92 individuals were selected for molecular analysis, representing nine families with eight individuals, two families with seven individuals, and one family with six individuals.

DNA extraction and genotyping

Genomic DNA was extracted from leaf tissue of the selected individuals using the CTAB (cationic hexadecyl bromide) method described by Doyle and Doyle (1987), with adjustments to include 1% polyvinylpyrrolidone (PVP) and 2% cetyltrimethylammonium bromide. The samples were precipitated in isopropanol for 12 hours, followed by centrifugation and resuspension of DNA in TE buffer (1 M Tris-HCL pH 8.0 and 0.5 M EDTA pH 8.0). DNA quantification was performed using a 1% agarose gel with known concentration DNA standards.

For genotyping, ten ISSR primers (JOHN, UBC's 808, 810, 822, 835, 840, 841, 842, 880, and 886) were employed. These primers were selected based on their ability to produce distinct and well-defined band patterns, as well as their capacity to detect polymorphisms. Polymerase chain reactions (PCRs) were performed in a thermocycler (GeneAmp PCR System 9700) with a total reaction volume of 20 µL, containing 1X buffer, $MgCl_2$ (2.5 mM), dNTP (1 mM), primer (0.2 µM), 1 unit of Taq DNA polymerase, and 50 ng of genomic DNA. Amplification comprised an initial denaturation step at 94°C for 5 minutes, followed by 37 cycles of denaturation at 94°C for 15 seconds, annealing at 47°C for 30 seconds, and extension at 72°C for 1 minute.

The amplification products were separated by electrophoresis in a 1.5% agarose gel supplemented with Gel Red. Electrophoresis was performed in a horizontal vat (Bio-Rad Sub-Cell®) with 1X TBE buffer (Tris-Borate EDTA) for 2 hours and 30 minutes at 90 volts. A 1 kb ladder marker was used to determine the molecular weight of the loci.

The gels were documented under ultraviolet (UV) light to visually analyze the band patterns, which were translated into a binary matrix where "1" indicated the presence of amplified DNA fragments and "0" indicated their absence.

Statistical analyses

The number of polymorphic bands required to obtain stable associations between individuals was estimated using bootstrap analysis. Simulations were performed for each pair of individuals and progenies, calculating correlation estimates (*r*) between the values of the original similarity matrix and the simulated matrix with different numbers of fragments. The stress value (E) was estimated to evaluate the fit between the matrices. Marker informativeness was assessed using the polymorphic information content (PIC) calculated for each primer. These analyses were conducted using the Genes software (Cruz, 2016).

Descriptive analysis involved determining the total number of bands (TNB), number of polymorphic bands (NPB), and the percentage of polymorphic bands (PPB) per primer. The binary matrix was subjected to statistical analysis, enabling comparisons within and between progenies. Genetic diversity parameters, including the number of observed alleles (A_O), number of effective alleles (A_E), Nei's genetic diversity index (*H**), and Shannon index (*I**), were calculated using the Popgene program (Yeh & Boyle, 1997).

Similarity coefficients were determined between genotypes using the Jaccard coefficient, which provided dissimilarity values. Cluster analysis was performed using the unweighted arithmetic mean grouping method (UPGMA), with the estimated cutoff point determined according to Mojena (1977). The cophenetic correlation coefficient (CCC) was calculated to assess the consistency of the groupings. The Genes program (Cruz, 2016) was used for these analyses. The circular dendrogram based on the dissimilarity matrix was exported to the R software (R Core Team, 2020) using packages such as vegan (Oksanen et al., 2018), cluster (Maechler et al., 2019), dendextend (Galili et al., 2020), factorextra (Kassambara & Mundt, 2020), ggpubr (Kassambara, 2020), cowplot (Wilke, 2019), and gridExtra (Auguie & Antonov, 2017).

To evaluate the genetic structure, molecular analysis of variance (AMOVA) was performed with two hierarchical levels, between and within progenies, using the Genes program (Cruz, 2016). The number of genetic groups (K) present in the population was determined using the Bayesian approach implemented in the Structure software (Falush, Stephens, & Pritchard, 2007). The analysis involved 20 runs for each K value, ranging from one to nine (6 provenances + 3 groups), with 7,500 interactions and a burn-in of 2,500 interactions, resulting in 10,000 Monte Carlo via Markov Chains (MCMC) interactions. The Structure Harvester software (Earl & Vonholdt, 2012) was used to determine the number of genetic groups based on the ad hoc ΔK method (Evanno, Regnaut, & Goudet, 2005).

Results and discussion

Descriptive analysis

A total of 91 bands were observed, of which 82 were polymorphic, accounting for 90.1% polymorphism (Table 1). The number of bands per primer ranged from 5 (UBC 886) to 14 (UBC 810 and UBC 835), with an average of 9 bands per primer. Three primers, UBC 822, UBC 841, and UBC 880, showed 100% polymorphic bands, while the John primer had only 66.7% polymorphic bands.

These results are consistent with previous findings by Lorenzoni, Soares, Santiago, Silva, and Coelho (2014) who reported 81.3% polymorphism and an average of 9 bands per primer using ISSR markers in *Annona mucosa* accessions. Similarly, Ramalho et al. (2016) observed 88.6% polymorphism and an average of 6 bands per primer in a diversity study with *Bertholletia excelsa*. Rajasekharan, Abdul Kareem, Ravish, and and Mini (2017) obtained 88% polymorphism in their work with *Oroxylum indicum*, another species belonging to the Bignoniaceae family.

The percentage of polymorphic bands indicates that the ISSR primers used were effective in detecting polymorphisms among individuals and evaluated progenies. Polymorphisms represent the genetic variability within the species, which is essential for its survival and adaptation to environmental pressures.

The polymorphic information content (PIC), which measures marker efficiency, was found to be 0.44 for the combined data, ranging from 0.28 to 0.41 among the primers. According to Tatikonda et al. (2009) classification, PIC values can be categorized as uninformative (0 to 0.25), moderately informative (> 0.25 to 0.45), and highly informative (> 0.45 to 0.5). Therefore, all ISSR markers used in this study were moderately informative, indicating their suitability for measuring the genetic diversity in the *Z. tuberculosa* population (Table 1).

Bootstrap analysis revealed that 75 ISSR bands were optimal for accurately estimating genetic diversity, as indicated by the high correlation (r = 0.9916) between the original similarity matrix and the simulated matrix, and a low stress value (E = 0.0283, $p < 0.05$). Thus, the 82 polymorphic bands analyzed were more than sufficient for precise estimation of genetic diversity, demonstrating the reliability of the chosen markers.

Table 1. Descriptive analysis of the ISSR primers used in the genetic characterization of *Zeyheria tuberculosa*. TNB: Total number of bands; NPB: Number of polymorphic bands; PPB: Percentage of polymorphic bands; PIC: polymorphism information content; H = (A, T or C); R = (A or G); V = (A, C or G); and Y = (C or T).

Primers	Sequence (5'-3')	TNB	NPB	PPB (%)	PIC
JOHN	(AG)7-YC	9	6	66.66	0.34
UBC 808	(AG)8-C	9	8	88.88	0.40
UBC 810	(GA)8-T	14	13	92.85	0.33
UBC 822	(TC)8-A	9	9	100.00	0.28
UBC 835	(AG)8-YC	14	13	92.85	0.34
UBC 840	(GA)8-YT	7	6	85.71	0.30
UBC 841	(GA)8-YC	10	10	100.00	0.38
UBC 842	(GA)8-YG	7	6	85.71	0.41
UBC 880	(GGAGA)3	7	7	100.00	0.35
UBC 886	VDV-(CT)7	5	4	80.00	0.35
Total	-	91	82	90.10	0.44

Genetic diversity

Progeny 7 exhibited the highest number of observed alleles (A_O) and effective alleles (A_E), with values of 1.70 and 1.49, respectively (Table 2). Higher A_O and A_E values indicate a greater presence and better distribution of alleles within the progeny. Since *Z. tuberculosa* is an allogamous species pollinated by bees and self-incompatible (Souza et al., 2017), these values suggest that the mother trees that produced progeny 7 had better cross-fertilization, possibly due to a higher occurrence of pollinators in the area or the presence of genetically distinct individuals nearby.

Progenies 3 and 73 had the lowest A_O values (1.51), while progeny 6 had the lowest A_E value (1.33). However, even with these lower values, the distribution of alleles within these progenies was still considered fair. Overall, the total number of A_O was 2.00 alleles per locus, while A_E was 1.61 (Table 2). These results suggest that the alleles detected in the population are well distributed among individuals, indicating proper seed collection for the test and the presence of different alleles at each locus.

Table 2. Number of alleles and genetic diversity estimated for the progenies of the *Zeyheria tuberculosa* using ISSR markers. A_O: Number of observed alleles; A_E: Number of effective alleles; H^*: Nei's genetic diversity index; I^*: Shannon index.

Progenies	A_O	A_E	H^*	I^*
3	1.51	1.36	0.20	0.29
6	1.55	1.33	0.19	0.29
73	1.51	1.35	0.19	0.28
74	1.57	1.36	0.21	0.31
22	1.56	1.41	0.22	0.32
7	1.70	1.49	0.27	0.40
12	1.63	1.38	0.22	0.33
59	1.68	1.44	0.25	0.37
41	1.64	1.44	0.25	0.36
10	1.69	1.46	0.26	0.38
70	1.56	1.37	0.21	0.31
77	1.62	1.37	0.21	0.32
Average	1.60	1.39	0.22	0.33
Total	2.00	1.61	0.35	0.52

The Nei (1978) index (H^*) value for genetic diversity was 0.35, and the Shannon index (I^*) was 0.52 (Table 2). The Shannon index ranges from 0 to 1, with higher values indicating greater genetic diversity within the population (Shannon & Weaver, 1949). Although there are no specific studies on the genetic diversity of *Z. tuberculosa* using ISSR markers, comparing with similar methodologies applied to other native forest species, the observed genetic diversity in this population can be considered moderate to high (Borges, Santos,

Maia, Lima, & Valente, 2016; Lopes, Costa, & Arriel, 2020). Similar values were found in *Handroanthus impetiginosus*, another species from the Bignoniaceae family, where ISSR markers yielded the same genetic diversity parameters (H^* = 0.35 and I^* = 0.52) (Pimenta, Felix, Araújo, Fajardo, & Pacheco, 2022).

The H^* and I^* values estimated for individual progenies allow for the identification of their contribution to the overall genetic diversity of the population. Progenies 41, 59, 10, and 7 exhibited the highest H^* and I^* values, while progenies 73, 3, and 6 had the lowest values (Table 2). The significant genetic diversity found within these selected families demonstrates the ability of the species' populations to maintain their genetic base despite historical predatory exploitation and habitat fragmentation. However, further studies on the genetic diversity of *Z. tuberculosa* populations are necessary to support conservation efforts and domestication strategies, enabling the exploration of the species' potential in the forestry sector and reducing its vulnerability to extinction.

Genetic dissimilarity

The UPGMA cluster analysis of individuals resulted in the formation of 10 groups, separated by the cutoff point obtained from the Mojena (1977) method, which considers the mean and standard deviation of dissimilarity values (Figure 2). The largest group, G1, consisted of 54 individuals, followed by G6, with 16 individuals. G1 comprised all progenies, with the most representative being families 3 (8 individuals) and 74 (7 individuals). G6 included six out of the twelve selected families, with families 10 and 59 being the most representative, with four individuals each. Progeny 77 exclusively formed group G5, comprising six individuals. There were also several smaller groups with three individuals (G4 and G10) and individual groups (G3, G7, G8, and G9).

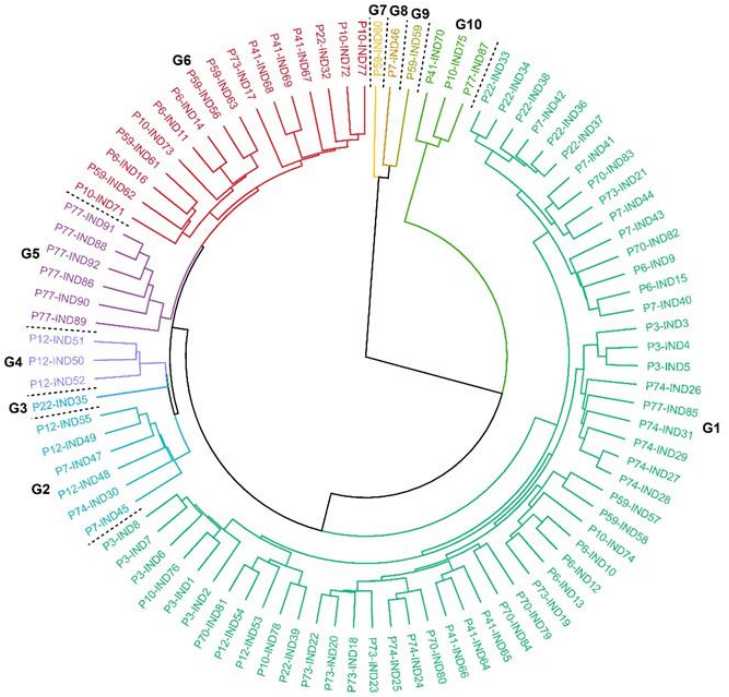

Figure 2. Cluster analysis of individuals of *Zeyheria tuberculosa* evaluated by ISSR markers, obtained using the UPGMA method with a cutoff point of 0.429. Cophenetic correlation coefficient (CCC): 84%. P: progeny; G: Group.

Heterogeneity was observed within some groups (G1, G2, G6, and G10), indicating genetic similarity between individuals from different families. Two possible hypotheses can explain this genetic composition: first, the geographic distance between the mother trees (minimum of 50 m) may not have been sufficient to interrupt gene flow, resulting in genetic similarity between individuals in these groups. The second hypothesis is the presence of historical gene flow when populations were geographically connected or when vegetation cover was extensive. In this case, the genetic similarity has been maintained over time despite forest fragmentation and evolutionary factors (Zambrano et al., 2020).

The cophenetic correlation coefficient (CCC) was 84%, indicating a good correlation between the dissimilarity measures and the dendrogram formed. CCC values equal to or greater than 60% validate the grouping of individuals based on genetic dissimilarity measures (Araújo, Barroso Neto, Pires, & Bertini, 2019).

The UPGMA analysis of progenies identified two distinct groups (Figure 3). The first group consisted of progeny 59, while the second group included all other progenies. The cophenetic correlation coefficient for this grouping was 91%, indicating the consistency of the groups formed based on the dissimilarity matrix.

For future selection and recombination efforts in the same area, aiming to achieve genetic divergence, it is recommended to include individuals from progeny 59. Additionally, priority should be given to including individuals representing the ten groups formed in the individual dendrogram (Figure 2). By considering these factors during individual selection, genetic drift can be prevented, as well as the subsequent increase in inbreeding within future breeding populations (Lefèvre et al., 2014). Moreover, maintaining individuals with greater genetic distances within the same population can enhance the genetic quality of seeds and increase the genetic gains of the selected population (Mijnsbrugge, Bischoff, & Smith, 2010).

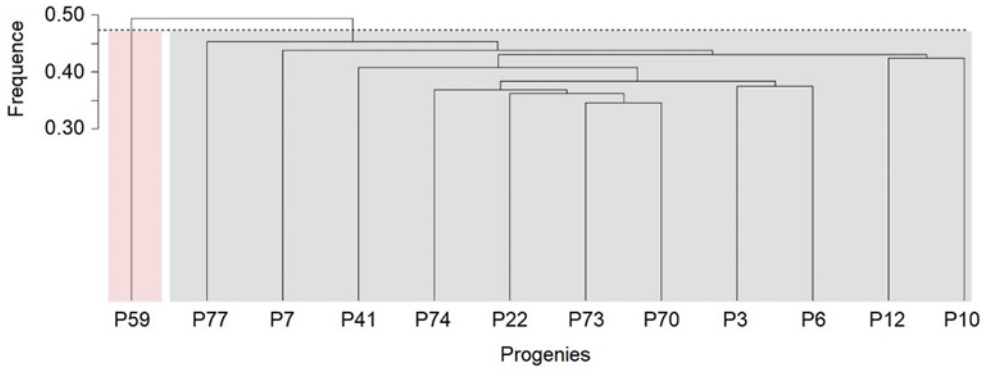

Figure 3. Cluster analysis of progenies of *Zeyheria tuberculosa* evaluated by ISSR markers, obtained using the UPGMA method with a cutoff point of 0.463. Cophenetic correlation coefficient (CCC): 91%.

Genetic structure

The *Φst* value represents the genetic differentiation between progenies, with low (0 to 0.05), moderate (> 0.05 to 0.15), high (> 0.15 to 0.25), and very high differentiation (> 0.25) classifications (Wright, 1978). In this study, a high genetic differentiation (*Φst* = 0.19) was observed between progenies (Table 3). However, when comparing the genetic variation between and within progenies, a larger percentage of variation was found within progenies (80.1%). This result aligns with the formation of a greater number of groups in the individual dendrogram (Figure 2) compared to the progeny dendrogram (Figure 3).

The higher variation within progenies may be associated with allogamy in *Z. tuberculosa*, where individuals prefer cross-fertilization for seed production, resulting in increased genetic variation within each progeny (Souza et al., 2017). The lower variation between progenies indicates that the selected progenies, despite being distributed across different municipalities in the state of Minas Gerais, still exhibit genetic similarity. This similarity may be attributed to past gene flow before fragmentation of the natural landscape when populations were more connected or when extensive vegetation (Slatkin, 2017).

Table 3. Analysis of molecular variance between and within progenies of *Zeyheria tuberculosa* species.

Source of variation	Degrees of freedom	Sum of squares	Variance components	Variation (%)
Between progenies	11	324.60	2.52	19.91
Within progenies	80	812.71	10.15	80.09
Total	91	1137.31	12.67	-
Φst = 0.19				

Our findings have important implications for the selection of genetic material for the establishment of a genetically improved seed orchard. It is recommended to prioritize the selection of the best individuals within each progeny rather than focusing solely on the best progenies. This approach ensures that the maximum post-selection genetic variation is retained, which can enhance the hybrid vigor of the seeds and increase the overall genetic gains in the selected population.

The Bayesian approach revealed the formation of two distinct genetic groups within the *Z. tuberculosa* population (Figure 4a). These groups corresponded to the clusters observed in the UPGMA cluster analysis between the progenies (Figure 3), indicating consistency between the two analyses. Progeny 59 exhibited a

higher degree of genetic divergence from the other progenies, as indicated by its prevalence in the green color in the cluster analysis (Figure 4b).

It is important to note that this study was conducted with a sample of the population, aiming to capture the genetic diversity of the species and simulate an early selection scenario. Despite the selection process, the results suggest that the population of *Z. tuberculosa* will still maintain sufficient levels of genetic diversity to support ongoing genetic improvement programs. Additionally, the use of mother trees can provide a quick way to collect seeds while ensuring the presence of confirmed genetic variability.

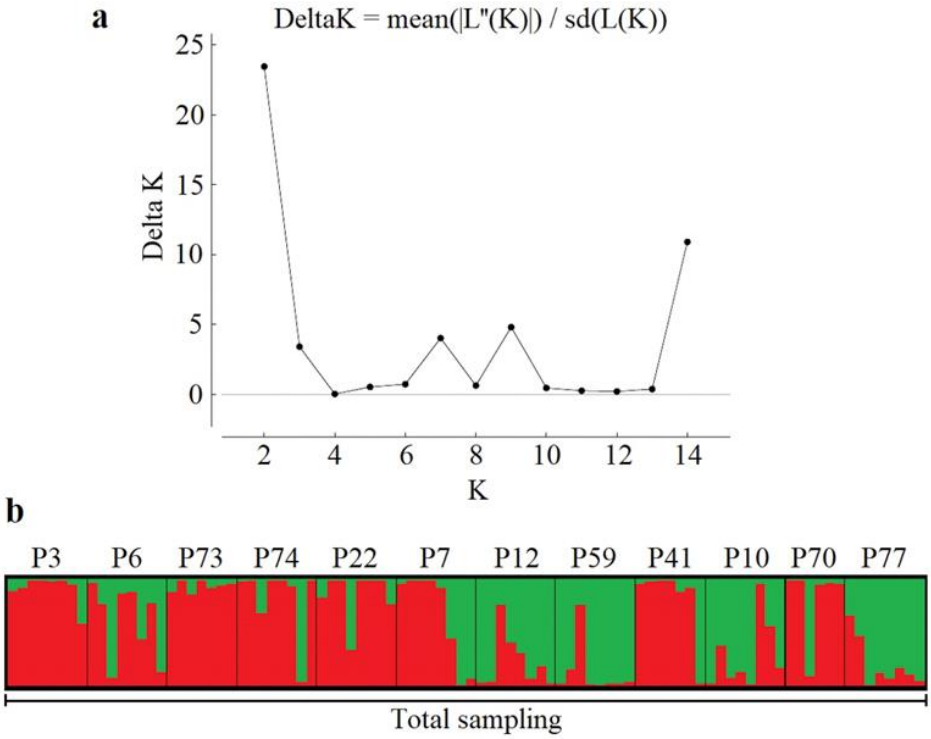

Figure 4. Bayesian analysis of genetic clustering in 12 progenies and 92 individuals of *Zeyheria tuberculosa*: (a) Graph of ΔK values for each K value, determining the optimal number of genetic clusters. Best K (2) based on Evanno et al. (2005) criteria; (b) Bar chart illustrating the distribution of clusters among progenies and the total sampling of individuals.

Conclusion

Zeyheria tuberculosa progeny testing confirmed a moderate to elevated level of genetic diversity. Additionally, the presence of genetically distinct groups and absence of genetic structure indicate the feasibility of maintaining satisfactory genetic variation levels in the resulting seeds and seedlings, even after selective breeding for genetic gain. Thereby, the findings support the viability of establishing an orchard of seedlings through seeds, which represents a promising strategy for both species conservation and genetic enhancement.

Acknowledgements

To the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) and *Fundação de Amparo à Pesquisa do Estado de Minas Gerais* (Fapemig) for financial support. To agreement 213/2019 between the *Universidade Federal de Lavras* (UFLA) and the *Fundação de Desenvolvimento Científico e Cultural* (Fundecc). To the Decentralized Execution Term 01/2018 between the *Universidade Federal de Lavras* and *Serviço Florestal Brasileiro* (SFB).

References

Araújo, L. B. R., Barroso Neto, A. M., Pires, K. R. A., & Bertini, C. H. C. M. (2019). Diversidade genética em pinhão manso com base em marcadores ISSR. *Nativa*, *7*(4), 363-370. DOI: http://dx.doi.org/10.31413/nativa.v7i4.6571

Auguie, B., & Antonov, A. (2017). *Miscellaneous functions for "Grid" graphics*. R package version 2.3. Retrieved on Nov. 1, 2022 from https://cran.r-project.org/web/packages/gridExtra/gridExtra

Beech, E., Rivers, M., Oldfield, S., & Smith, P. P. (2017). GlobalTreeSearch: The first complete global database of tree species and country distributions. *Journal of Sustainable Forestry*, *36*(5), 454-489. DOI: https://doi.org/10.1080/10549811.2017.1310049

Bertoni, B. W., Filho, S. A., Martins, E. R., Filho, C. F. D., França, S. C., Pereira, A. M. S., ... Filho, J. A. F. D. (2007). Genetic variability in natural populations of *Zeyheria montana* mart. from the Brazilian Cerrado. *Scientia Agricola*, *64*(4), 409-415. DOI: https://doi.org/10.1590/S0103-90162007000400012

Borges, R. C., Santos, F. M. G., Maia, M. C. C., Lima, P. S. C., & Valente, S. E. S. (2016). Investigating genetic diversity in sapucaia using inter simple sequence repeat markers. *Genetics and Molecular Research*, *15*(3), 1-13. DOI: http://dx.doi.org/10.4238/gmr.15038565

Brandão, M. M., Vieira, F. A., Nazareno, A. G., & Carvalho, D. (2015). Genetic diversity of neotropical tree *Myrcia splendens* (Myrtaceae) in a fragment-corridor system in the Atlantic rainforest. *Flora - Morphology, Distribution, Functional Ecology of Plants*, *206*, 35-41. DOI: https://doi.org/10.1016/j.flora.2015.07.006

Brasil. (2021). *Portaria MMA Nº 561, de 15 de dezembro de 2021*. Brasília, DF: Ministério do Meio Ambiente.

Carvalho, P. E. R. (2005). *Ipê-felpudo*. Colombo, PR: Embrapa Florestas.

Cruz, C. D. (2016). Genes software-extended and integrated with the R, Matlab and Selegen. *Acta Scientiarum. Agronomy*, *38*(4), 547-552. DOI: https://doi.org/10.4025/actasciagron.v38i4.32629

Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, *19*(1), 11-15.

Earl, D. A., & Vonholdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*, 359-361. DOI: https://doi.org/10.1007/s12686-011-9548-7

El-Kassaby, Y. A., Wang, Q., Wang, T., Ratcliffe, B., Bi, Q. X., Wang, Z., ... Guan, W. (2019). Concept for gene conservation strategy for the endangered Chinese yellowhorn, *Xanthoceras sorbifolium*, based on simulation of pairwise kinship coefficients. *Forest Ecology and Management*, *432*, 976–982. DOI: https://doi.org/10.1016/j.foreco.2018.10.045

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software Structure: a simulation study. *Molecular Ecology, 14*(8), 2611-2620. DOI: https://doi.org/10.1111/j.1365-294X.2005.02553.x

Falush, D., Stephens, M., & Pritchard, J. K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, *7*(4), 574-578. DOI: https://doi.org/10.1111/j.1471-8286.2007.01758.x

Galili, T., Benjamini, Y., Simpson, G., Jefferis, G., Gallota, M., Renaudie, J., ... Bradley, D. (2020). *Extending 'dendrogram' functionality in* R. R package version 1.13.4. Retrieved on Nov. 1, 2022 from https://cran.r-project.org/web/packages/dendextend/dendextend

Idrees, M., & Irshad, M. (2014). Molecular markers in plants for analysis of genetic diversity: A review. *European Academic Research*, *2*(1), 1513-1540.

Indústria Brasileira de Árvores [IBÁ]. (2021). *Relatório anual IBÁ 2021*. Brasilia, DF: Indústria Brasileira de Árvores.

Instituto Nacional de Meteorologia [INMET]. (2021). *Normais climatológicas do Brasil*. Retrieved on Mar. 27, 2022 from https://portal.inmet.gov.br/normais

Kassambara, A. (2020) *'ggplot2' Based Publication Ready Plots*. R package version 0.4.0. Retrieved on Nov. 1, 2022 from https://cran.r-project.org/web/packages/ggpubr/ggpubr

Kassambara, A., & Mundt, F. (2020). *Factoextra: Extract and visualize the results of multivariate data analyses*. R package version 1.0.5. Retrieved on Nov. 1, 2022 from https://CRAN.R-project.org/package=factoextra

Lefèvre, F., Boivin, T., Bontemps, A., Courbet, F., Davi, H., Gillmann, M. D., ... Pichot, C. (2014). Considering evolutionary processes in adaptive forestry. *Annals of Forest Science*, *71*, 723-739. DOI: https://doi.org/10.1007/s13595-013-0272-1

Lohmann, L. G. (2015). *Bignoniaceae in lista de espécies da flora do Brasil*. Rio de Janeiro, RJ: Jardim Botânico do Rio de Janeiro. Retrieved on Mar. 27, 2022 from http://floradobrasil2015.jbrj.gov.br/FB112305

Lopes, B. C., Ferreira, M. B. D., & Brandão, M. (1996). Sombreamento em pastagens: espécies recomendadas para as diversas regiões do Estado de Minas Gerais. *Daphne*, *6*(4), 7-15.

Lopes, J. D. S., Costa, M. R. J., & Arriel, D. A. A. (2020). Genetic diversity of potential mother trees of *Myracrodruon urundeuva* Allemão in a remnant population from Brazilian Cerrado using ISSR. *Advances in Forestry Science*, *7*(2), 1017-1024. DOI: http://dx.doi.org/10.34062/afs.v6i2.5621

Lorenzoni, R. M., Soares, T. C. B., Santiago, V. F., Silva, J. A., & Coelho, R. I. (2014). Utilização de marcadores ISSR na avaliação da diversidade genética entre acessos de biribazeiro. *Revista Brasileira de Fruticultura*, *36*(1), 251-257. DOI: https://doi.org/10.1590/S0100-29452014000500029

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., Studer, M., ... Murphy, K. (2019). *Finding groups in data": Cluster analysis extended Rousseeuw et al*. R package version 2.1.0. Retrieved on Nov. 1, 2022 from https://cran.r-project.org/web/packages/cluster/cluster

Magistrali, P. R., Cacau, F. V. P., Nascimento, J. F., & Magistrali, I. C. (2022). Tamanho de recipientes e luminosidade na produção de mudas *Zeyheria tuberculosa* (Vell.) Bur. (Bignoniaceae). *Revista Brasileira de Ciências da Amazônia*, *11*(1), 1-8. DOI: https://doi.org/10.47209/2317-5729.v.11.n.1.p.1-8

Mendonça, A. R., Silva, J. C., Aozani, T. S., Silva, E. R., Santos, J. S., Binoti, D. H. B., & Silva, G. F. (2018). Estimação da altura total de árvores de ipê felpudo utilizando modelos de regressão e redes neurais artificiais. *Revista Brasileira de Biometria*, *36*(1), 128-139. DOI: https://doi.org/10.28951/rbb.v36i1.154

Mijnsbrugge, K. V., Bischoff, A., & Smith, B. (2010). A question of origin: Where and how to collect seed for ecological restoration. *Basic and Applied Ecology*, *11*(4), 300-311. DOI: https://doi.org/10.1016/j.baae.2009.09.002

Mojena, R. (1977). Hierarchical grouping methods and stopping rules: an evaluation. *The Computer Journal*, *20*(4), 359-363. DOI: https://doi.org/10.1093/comjnl/20.4.359

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, *89*(3), 583-590. DOI: https://doi.org/10.1093/genetics/89.3.583

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., ... Wagner, E. (2018). *Vegan: Community ecology package*. R package version 2.4-5. Retrieved on Nov. 1, 2022 from https://CRAN.R-project.org/package=vegan

Pimenta, J. M. A., Felix, F. C., Araújo, J. S. O., Fajardo, C. G., & Pacheco, M. V. (2022). Selection of ISSR molecular primers for studies of genetic diversity in *Handroanthus impetiginosus* (Mart. Ex DC.) Mattos. *Revista Caatinga*, *35*(1), 231-238. DOI: https://doi.org/10.1590/1983-21252022v35n124rc

R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, AT: R Foundation for Statistical Computing. Retrieved on Nov. 1, 2022 from URL https://.R-project.org/

Rajasekharan, P. E., Abdul Kareem, V. K., Ravish, B. S., & Mini, S. (2017). Genetic diversity in *Oroxylum indicum* (L.) Vent., a threatened medicinal plants from India by ISSR analysis. *Indian Journal of Biotechnology*, *16*(3), 360-368.

Ramalho, A. B., Rossi, A. A. B., Dardengo, J. F. E., Zortéa, K. É. M., Tiago, A. V., & Martins, K. C. (2016). Diversidade genética entre genótipos de *Bertholletia excelsa* por meio de marcadores moleculares ISSR. *Floresta*, *46*(2), 207-214. DOI: http://dx.doi.org/10.5380/rf.v46i2.41970

Resende, M. D. V. (2016). Software Selegen-REML/BLUP: a useful tool for plant breeding. *Crop Breeding and Applied Biotechnology*, *16*(4), 330-339. DOI: https://doi.org/10.1590/1984-70332016v16n4a49

Slatkin, M. (2017). *Gene flow and population structure*. Real, LA: Princeton University Press.

Shannon, C. E., & Weaver, W. A. (1949). Mathematical model of communication. Urbana, IL: University of Illinois Press.

Silva Júnior, A. L., Souza, L. C., Pereira, A. G., Caldeira, M. V. W., & Miranda, F. D. (2017). Genetic diversity of *Schizolobium parahyba* var. *amazonicum* (Huber ex. Ducke) Barneby, in a forest area in Brazil. *Genetics and Molecular Research*, *16*(3), 1-11. DOI: http://dx.doi.org/10.4238/gmr16039774

Souza, C. V., Nepi, M., Machado, S. R., & Guimarães, E. (2017). Floral biology, nectar secretion pattern and fruit set of a threatened Bignoniaceae tree from Brazilian tropical forest. *Flora*, *227*, 46-55. DOI: https://doi.org/10.1016/j.flora.2016.12.007

Tatikonda, L., Wani, S. P., Kannan, S., Beerelli, N., Sreedevi, T. K., Hoisington, ... Varshney, R. K. (2009). AFLP-based molecular characterization of an elite germplasm collection of *Jatropha curcas* L. biofuel plant. *Plant Science*, *176*(4), 505-513. DOI: https://doi.org/10.1016/j.plantsci.2009.01.006

Teixeira, G. C., Konzen, E. R., Faria, J. C. T., Gonçalves, D. S., Carvalho, D., & Brondani, G. E. (2020). Genetic diversity analysis of two *Eucalyptus* species using ISSR markers. *Ciência Florestal*, *30*(1), 270-278. DOI: https://doi.org/10.5902/1980509832804

Wilke, C. O. (2019). *Streamlined plot theme and plot annotations for 'ggplot2'*. R package version 1.0.0. 2019. Retrieved on Nov. 1, 2022 from https://cran.r-project.org/web/packages/cowplot/cowplot

Wright, S. (1978). *Evolution and the genetics of populations: variability within and among natural populations*. Chicago, US: University of Chicago Press.

Yeh, F. C., & Boyle, T. J. B. (1997). Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany*, *129*, 156-157.

Zambrano, J., Cordeiro, N. J., Garzon-Lopez, C., Yeager, L., Fortunel, C., Ndangalasi, H. J., & Beckman, N. G. (2020). Investigating the direct and indirect effects of forest fragmentation on plant functional diversity. *PLoS ONE*, *15*(7), 1-16. DOI: https://doi.org/10.1371/journal.pone.0235210