CROP BREEDING AND APPLIED BIOTECHNOLOGY

ARTICLE

Validated multiplex PCR of microsatellites for breeding purposes, clonal fidelity and controlled crosses in teak

Dawa Méndez-Álvarez[1], Emanuel Araya-Valverde[1,2*], Yorleny Badilla-Valverde[1] and Olman Murillo-Gamboa[1]

Crop Breeding and Applied Biotechnology
25(2): e51202523, 2025
Brazilian Society of Plant Breeding.
Printed in Brazil
http://dx.doi.org/10.1590/1984-70332025v25n2a03

Abstract: *Teak (Tectona grandis L.f), a valuable hardwood species, is widely cultivated for its superior wood properties. Effective breeding and clonal propagation programs require robust genetic verification methods such as clonal fidelity and parentage verification when these programs include controlled cross-pollination and propagation of elite clones. This study validates the use of ten optimized microsatellite markers in a multiplex PCR system for clonal identification and parentage testing. The system showed high discrimination power, with probabilities of identity (PI) and exclusion (PE), confirming its efficacy. The ten markers accurately identified clonal fidelity and ensured genetic fidelity of progenies. Overall, the study confirms effectiveness of microsatellites for genetic analysis and highlights a cost effective and precise approach for breeding and clonal propagation of teak.*

Keywords: *Capillary electrophoresis, parentage analysis,* Tectona grandis*, probability of exclusion*

INTRODUCTION

Teak occurs naturally in India, Laos, Burma and Thailand and it has been introduced to countries in Asia, Africa, and the Americas (Thakor et al. 2019). Small-scale forestry represents a large portion of teak planters (Kollert et al. 2024). Its hardwood is highly valuable because of wood density, mechanical properties, high resistance to physical and biological agents, and wood color (Ramasamy et al. 2021). Traits related to wood properties (Naranjo et al. 2012, Moya et al. 2013) and stem growth (Mandal and Rambabu 2001, Callister and Collins 2008, Murillo et al. 2019, Badilla and Murillo 2022, Callister et al. 2024) have received significant attention. Breeding programs in teak have involved provenance and family trials, as well as selection of plus trees for further establishment of different types of seed sources (Palanisamy et al. 2009). Advanced tree breeding programs involve controlled cross-pollination through selected elite individuals, where contributions of parents must be assured. Quality control in tree breeding such as clonal fidelity, correct parentage of full-sib crosses and lineage of breeding lines across generations are important (Bell et al. 2004).

Molecular markers have a pivotal role in tree species to estimate genetic relatedness and mating systems, germplasm identification, controlled crosses

***Corresponding author:**
E-mail: earaya@cenat.ac.cr
ORCID: 0000-0002-0025-5420

Received: 20 November 2024
Accepted: 2 February 2025
Published: 18 March 2025

[1] Instituto Tecnológico de Costa Rica, 1, km Sur de la Basílica de los Ángeles, C. 15, Cartago, Dulce Nombre, 30109, Costa Rica
[2] Centro Nacional de Alta Tecnología, 1174-1200 Calle, San Jose, 10109, Costa Rica

(Mahajan and Gupta 2012, Jhariya et al. 2014) and genomic selection (Grattapaglia 2022). Molecular markers in teak have been used to study its genetic diversity and structure (Vaishnaw et al. 2015, Maisuria et al. 2022, Anjos et al. 2023) and genetic relatedness in locations where teak naturally occurs or has been introduced (Verhaegen et al. 2010, Minn et al. 2014, Hansen et al. 2017, Balakrishnan et al. 2021). Tambarussi et al. (2017) used microsatellites to evaluate the genetic variability and mating system in seeds employed to induce *in vitro* organogenesis. Recently, Callister et al. (2024) reported the first genomic selection of teak clones using single nucleotide polymorphism (SNP) markers of five economically important traits. These results are promising but still require a broader genetic sample to increase prediction certainty and reliability. SNP markers in teak will adapt longer at an operational scale and meanwhile other molecular marker systems should be adapted.

Despite next-generation sequencing (NGS) technologies lead to an increasing use of SNP markers for applications mentioned above, microsatellites are still a viable option (Flanagan and Jones 2019). These molecular markers display high polymorphism, are reproducible and their scoring is relatively easy. Besides, true multiplexing is possible because they are amenable to automation using genetic analyzers with capillary electrophoresis (CE) coupled with laser-induced DNA fluorescence, which allows a high precision in DNA genotyping (Butler et al. 2004). It reduces both labor time and cost of microsatellite genotyping (Guichoux et al. 2011), which is especially useful in breeding programs. CE has been used in multiple plant species for purposes of conservation of genetic diversity and breeding (Guimarães et al. 2019, Sanchez et al. 2020, Maia et al. 2022, Koltun et al. 2024).

For clonal fidelity and parentage analyses, parameters such as probability of identity (PI) and probability of exclusion (PE) of microsatellite markers are relevant for both identifying clonally propagated individuals and individual identification in parentage testing (Bell et al. 2004, Queiroz et al. 2023). Studies describing molecular markers for clonal identification in teak are scarce (Huang et al. 2016, Mahesh et al. 2016, Queiroz et al. 2023), and therefore, the present study contributes with validated results in a teak breeding program. The most used microsatellites in teak for reporting genetic diversity, PI and PE parameters, are those developed by Verhaegen et al. (2005). Hence, the objectives of our study were to validate the use of ten microsatellites optimized for their amplification in multiplex PCR reactions. We also demonstrate that this set of ten microsatellites are reliable for routine applications in clonal fidelity and parentage testing in teak, which has not been previously described for this important tree species.

MATERIAL AND METHODS

Plant material

An international teak breeding program is being carried out since early 2000 by GENFORES, a tree improvement and gene conservation cooperative created in the School of Forest Engineering at Instituto Tecnológico de Costa Rica with more than 400 genotypes (Murillo et al. 2019). A set of 25 elite genotypes are being used for control pollination, monoclonal field-testing, as well as for optimizing gene markers applications. Validation of multiplex PCR was performed through clonal purity testing in large monoclonal production beds, as well as by analyzing consistency in control pollination full-sib progenies.

Through a systematic sampling procedure with a random initiation, twenty-eight mother stock ramets were selected in monoclonal production beds under greenhouse conditions, for two different clones, hereinafter named as clone 1M and clone 35E located at Instituto Tecnológico campus (lat 10° 22’ 3.20’’ N, long 84° 30’ 46.45’’ W) in San Carlos, Costa Rica. The second set of samples consisted of three groups of putative full-sib progenies (seedlings) from teak controlled crosses obtained at Panamerican Woods S.A. breeding program, located at Nandayure (lat 9° 50’ 45.96’’ N, long 85° 13’ 33.02’’ W), Guanacaste, Costa Rica. The numbers of individuals were 20, 12 and 16, respectively, for putative progenies named Cross C, Cross E and Cross O. A leaf portion of about 20 cm in length per seedling was collected and placed in a paper bag with desiccant silica (Merck, Darmstadt, Germany). Samples were transported to the Forest Genetics Laboratory, located at the Forest Research and Innovation Center in the Forest Engineering School, Instituto Tecnológico de Costa Rica in Cartago city.

DNA isolation

For DNA isolation, 50 mg of dried leaf were ground in a homogenizer Fastprep-24 (MP Biomedicals, Irvine, USA) using lysing

matrix A (MP Biomedicals, Irvine, USA). A CTAB-based method was employed as described by Quirós-Guerrero et al. (2019). DNA concentration was measured with a spectrophotometer NanoDrop 2000 (Thermo Scientific, Delaware, USA) and diluted (~40 ng μL^{-1}) for PCR amplification.

Microsatellite genotyping using multiplex PCR

Ten simple sequence repeats (SSR) regions previously reported for *T. grandis* were amplified by PCR (Verhaegen et al. 2005). Forward and reverse primers for loci CIRAD3TeakE06, CIRAD3TeakB02, CIRAD2TeakC03 were redesigned to allocate them into PCR multiplex reactions to avoid overlapping of alleles between the different microsatellites (Table 1).

The primer pairs of the ten microsatellites were accommodated in two multiplex PCR reactions. Several primer concentrations were tested and, after each multiplex PCR assay, capillary electrophoresis in ABI SeqStudio Genetic Analyzer (Applied Biosystems, California, USA) was performed and electropherograms were manually inspected in the GeneMapper® V4.0 software. Finally, two multiplex PCR reactions (final volume 25 µL) were optimized, each one containing 1X DreamTaq PCR buffer (ThermoFisher Scientific, Delaware, USA), 2.5 mM $MgCl_2$ (Thermo Scientific, Delaware, USA), 0.2 mM dNTPs (Thermo Scientific, Delaware, USA), 2 Units of DreamTaq DNA Polymerase (Thermo Scientific, Delaware, USA), 0.4 mg bovine serum albumin (Thermo Scientific, Delaware, USA), 2% DMSO (Thermo Scientific, Delaware, USA) and 2 uL DNA (~80 ng). Primer concentrations were optimized as described in Table 1.

The amplification was performed in a thermal cycler (Veriti™, Applied Biosystems, USA) with the following profile: one cycle at 95 ºC for 5 min followed by 35 cycles at 95 ºC for 30 s, 55 ºC for 30 s, 72 ºC for 30 s, with a final extension at 72 °C for 8 min. Multiplex PCR products were diluted 1/10 and visualized by capillary electrophoresis in an ABI SeqStudio Genetic Analyzer (Applied Biosystems, California, USA) with the combination of 1.5 µL of the diluted multiplex PCR product, 0.4 µL of GeneScan 600LIZ™ (Applied Biosystems, California, USA), and 8.5 µL of Hi-Di Formamide (Applied Biosystems, California, USA). After capillary electrophoresis, allele binning and genotyping were manually inspected with GeneMapperTM v4.0 software (Applied Biosystems, USA) for further data analysis.

Table 1. Primer sequences with their respective forward fluorescent dye label of 10 microsatellites and the final concentration used for the multiplex PCR amplification in DNA of *T. grandis*

Locus name	Accession nr.	Fluorescent dye-Sequence of forward primer (5` - 3`)	Sequence of reverse primer (5` - 3`)	Primer concentration (nM)	Multiplex PCR
CIRAD3TeakE06*	AJ968939	VIC-AGGTAGATAGAGAAGCTACGGTTGA	CCTATTTTCTTCCCCTCCCTT	150	1
CIRAD1TeakA06	AJ968929	NED-CAAAACAAAACCAATAGCCAGAC	TTTCATCATCATCATCAACATCC	100	1
CIRAD4TeakF02	AJ968942	PET-CCGGTAAAAAGGTGTGTCA	GAGTGGAAGTGCTAATGGA	100	1
CIRAD3TeakB02*	AJ968937	6FAM-CTGGGACTACAATCTATCTTCACA	GCCTAGGAAGACTGGGGAATAA	100	1
CIRAD3TeakA11	AJ968936	VIC-AAACCATGACAGAAACGAATC	TTGGGAATGGGAGGAGAAGT	100	1
CIRAD2TeakB07	AJ968934	VIC-GGGTGCTGATGATTTTGAGTT	CTAAGGAGTGAGTGGAGTTTT	100	2
CIRAD3TeakF01	AJ968940	NED-GCTCTCCACCAACCTAAACAA	AAAACGTCTCACCTTCTCACT	100	2
CIRAD1TeakF05	AJ968931	FAM-CTTCTGCAACCCTTTTTCAC	AGCCATATCTTCCTTTCTCT	100	2
CIRAD1TeakH10	AJ968933	PET-CGATACCTGCGATGCGAAGC	CGTTGAATACCCGATGGAGA	160	2
CIRAD2TeakC03*	AJ968935	6FAM-AGGTGGGATGTGGTTAGAAGC	AAATGGTCATCAGTGTCAGAA	100	2

* Primers redesigned for multiplex PCR

Data analysis and validation

The ability of the 10 microsatellites for individual identification and parentage testing was determined with the genotyping data of both clonally propagated plants from clone 1M and clone 35E and the three groups of putative full-sib progenies (Cross C, Cross E and Cross O). The following parameters were estimated using GENALEX 6.5 (Peakall and Smouse 2012): 1. Number of alleles per locus (Na); 2. Observed (Ho) and expected (He) heterozygosity; 3. Probability of Identity (PI), which is a crucial parameter for validation in plant breeding and clonal propagation, to test the probability that the same genotype is found in two unrelated individuals; and 4. Probability of Exclusion (PE), which estimates the probability that an individual tree will be incorrectly identified as the parent of an offspring. Three PE estimations were obtained Polymorphic information content (PIC) was estimated in the software CERVUS (Kalinowski et al. 2007).

DNA fingerprinting of clonally propagated ramets of both clones (1M and 35E) were inspected to verify its purity. On the other hand, segregation patterns of the three progenies were verified against the fingerprinting profile of their female and male parents. Finally, genotype frequencies obtained from one controlled cross pollination was analyzed by chi-square procedures, to test deviations from expected segregation contrasted with their parental genotypes. Hypotheses on Mendelian segregation of observed phenotype frequencies were tested for each of the 10 SSR investigated. Goodness-of-fit tests were utilized to verify the fulfillment of the quantitative tests (the χ^2 test yielded non-significant deviation between observed and expected segregation proportions, at 5% level of significance). Expected offspring frequencies were determined based on both parental genotypes as follows:

	M_i	M_j
F_i	¼ (F_i M_i) N	¼ (F_i M_j) N
F_j	¼ (F_j M_i) N	¼ (F_j M_j) N

N = total number of offspring observed

F_i = observed female genotype of the i^{th} allele

F_j = observed female genotype of the j^{th} allele

M_i = observed male genotype of the i^{th} allele

M_j = observed female genotype of the j^{th} allele

Significance of tests was expressed in terms of the p-value, as the likelihood of obtaining different results in the observed data under the null hypothesis. Its calculation was determined as regularly done, based on the deviation between the observed value and the expected alternative reference value. A greater difference between these two values corresponded to a lower p-value. A p-value = 0.05 or lower was considered statistically significant (1-0.05 = 95% of confidence).

Sampling procedures in finite clonal populations

To develop a sampling procedures for testing clonal authenticity or purity in mother stock populations, we took a systematic sample with random initiation from two putative monoclonal populations of teak in greenhouse (clones 1M and 35E). Both clonal populations had an amount of 350 mother stock plants (N), in a single bed for clonal production under greenhouse conditions. A systematic sample with random initiation was taken, where 28 plants were collected in each monoclonal batch. All of them were genotyped as mentioned before, using our multiplex PCR.

Based on conventional sampling analysis for binomial populations (Steel and Torrie 1980) minimum sampling size was estimated for different accuracy scenarios, population size and probability of purity as follows:

$$n = \left(\frac{t\sqrt{p(1-p)}}{Error} \right)^2 \quad (1)$$

where:

n = sampling size

t = Student's t-distribution for populations > 120

p = proportion of true genotypes

error = expected error estimation

"n_c" corrected for population size

$$nc = \frac{n}{1+\frac{n}{N}} \quad (2)$$

N = population size

Sampling scenarios were based on off-type allowed proportions of 10 and 20%, sampling error estimation of 5% and 10%, and monoclonal population sizes of *N* = 500, *N* = 1000 and *N* = 2000 mother stock plants in clonal gardens.

RESULTS AND DISCUSSION

Characterization microsatellites´ capacity for breeding purposes

The total number of alleles amplified with ten microsatellites was 50 (average of 5 alleles per locus) ranging from 3 to 7 alleles. The observed heterozygosity (Ho) ranged from 0.127 to 0.873 (mean of 0.509) and the expected heterozygosity (He) per locus ranged from 0.121 to 0.772 (mean of 0.580). PIC values ranged from 0.117 to 0.737 with similarity to He values. The probability of identity of both scenarios, when individuals are unrelated (PI) or when individuals are full-sibs (PIsibs), were 9.616 x 10^{-8} and 0.00109, respectively (Table 2). The combined expected number of individuals with the same genotype across the ten SSRs is 0.0000106 and 0.1196400, respectively. We also found that clonal identification among unrelated individuals can be reached with a power of exclusion of > 0.999 with the most informative loci (CIRAD1TeakH10, CIRAD1TeakA06, CIRAD2TeakC03, CIRAD3TeakA11, and CIRAD3TeakB02) as shown in Table 2. On the other hand, if the clonal identification in teak is pursued in full-sib progenies (PIsibs), the ten microsatellites are needed to identify clonal purity with a power of exclusion > 0.999. The highest discrimination values of PI of the five most informative SSR of our study have also been reported previously (Huang et al. 2016, Queiroz et al. 2023). Twenty-six clones of teak were unequivocally identified when a set of 15 SSRs were employed with a PI value of 3.248 x 10^{-16} (Huang et al. 2016). Queiroz et al. (2023) described a PI value of 1.843 x 10^{-23} and a PIsibs value of 4.823 10^{-9} with 21 teak microsatellites in a sample of 50 clones planted in Brazil. These authors also found that, with a subset of the five microsatellites, combined PI and PIsibs are for identity analysis and parentage testing. However, a minimum set of 12 highly informative markers is recommended, consistent with our findings.

The combined probability of paternity exclusion (PE) was > 0.999 when the genotypes of the two parents are known (PE1) and when the genotype of only one parent is known (PE2). The probability of exclusion when two putative parents are excluded (PE3) was 0.995. The five most informative SSRs for PI also displayed the highest PE1 values (Table 2), which also aligns with high exclusion power described before (Huang et al. 2016, Queiroz et al. 2023). PE1 value is the most relevant for our study since the exclusion of parents was tested in progenies of controlled crosses. Nonetheless, our set of 10 SSRs is also suitable to test paternity in open pollinated progenies when the female parent is known since PE2 is > 0.999. Future applications must include the 10 loci to determine parent exclusion unequivocally.

Teak breeding programs based on clonal forestry can be assisted with reliable and low-cost gene markers. Several constraints can happen in any organization that promotes large-scale clonal forestry. From the initial propagation of

Table 2. Genetic diversity parameters of ten microsatellites tested in clonally propagated plants and progenies from teak controlled crosses. Na: number of alleles, He: expected heterozygosity, Ho: observed heterozygosity, PIC: Polymorphic Information Content. PI: probability of identity; PIsibs: probability of identity of sibs; PE1: probability of exclusion when the genotypes of the two parents are known: PE2: probability of exclusion when the genotype of only one parent is known; PE3: probability of exclusion when two putative parents are excluded

Locus name	**Allele size range** (bp)	**Na**	**Ho**	**He**	**PIC**	**PI**	**PIsibs**	**PE1**	**PE2**	**PE3**
CIRAD3TeakE06	126-138	4	0.873	0.647	0.592	0.180	0.471	0.391	0.228	0.564
CIRAD1TeakA06	171-189	5	0.473	0.733	0.687	0.117	0.413	0.493	0.319	0.673
CIRAD4TeakF02	205-211	4	0.327	0.301	0.272	0.518	0.729	0.146	0.046	0.248
CIRAD3TeakB02	234-243	7	0.545	0.627	0.591	0.176	0.480	0.405	0.228	0.599
CIRAD3TeakA11	252-260	6	0.682	0.694	0.662	0.125	0.434	0.483	0.298	0.684
CIRAD2TeakB07	123.-127	3	0.127	0.121	0.117	0.777	0.884	0.060	0.007	0.113
CIRAD3TeakF01	188-198	6	0.518	0.528	0.499	0.252	0.549	0.326	0.156	0.513
CIRAD1TeakF05	232-247	3	0.236	0.651	0.576	0.197	0.474	0.357	0.212	0.506
CIRAD1TeakH10	219-233	6	0.564	0.772	0.737	0.087	0.386	0.560	0.381	0.744
CIRAD2TeakC03	296-308	6	0.745	0.730	0.684	0.119	0.415	0.489	0.313	0.668
	Mean	5	0.509	0.580	0.542	0.255	0.523	0.371	0.219	0.531
	Combined probability					9.616 x 10^{-8}	0.00109	>0.999	>0.999	0.995

elite materials to be deployed, human errors commonly occur in labeling and mishandling inside greenhouse facilities. Testing and certifying true clonal purity are essential procedures nowadays. Identity reliability is mandatory in any clonal program, which is especially true at larger scale. The other main concern involves genetic testing, which is becoming a highly professional and delicate issue in any breeding program. The whole breeding strategy is dependent on good genetic values obtained from well designed and conducted genetic tests. Thus, certainty on the identity of those top elite materials outperforming in field tests is simply crucial. The possibility of using reliable and low-cost gene markers will speed up breeding progress. Our multiplex PCR system reduces the cost of PCR reactions by 80% in both laboratory consumables and human labor because only two multiplex PCR reactions are performed instead of 10 individual reactions. The mixing of separate PCR reactions for capillary electrophoresis can be cumbersome as the amount of each individual PCR reaction must be adjusted to obtain signals that are easy to analyze. In the multiplex PCR reaction this adjustment is provided by the optimized concentration of each primer.

Validation of microsatellites (clonal fidelity)

Three out of the 28 individuals clonally propagated for both clone 1M and clone 35E displayed a mismatched DNA fingerprinting with the remaining 25 individuals, which corresponds to 89.3% of clonal purity. The individual 1M-15 was identified as a clone of 35E, which may suggest a sampling error, and the other two individuals (1M-3 and 1M-6) have the same DNA profile. For clone 35M, the three disparate individuals (35E-1, 35E-7 and 35E-15) share the same DNA profile. Based on the power of discrimination of 10 SSRs (high PI), the mismatched individuals showed different DNA profiles in 8 and 7 loci, respectively, for 1M and 35E.

Validation of microsatellites (progeny testing)

The ten SSRs were able to detect full sibs and clones inside the progenies. The expected segregation pattern of all 20 individuals in Cross C at all 10 loci consistently matched the genotype of their female and male parents and hence 100% full sibs were detected. None of the 10 microsatellites were inconsistent with expected offspring segregation in

Table 3. X^2 analysis based on eight teak microsatellites, for offspring deviation from its parental genotypes in controlled crosses

SSR*	Genotype Parent 1	Genotype Parent 2	Genotype of progenies		Observed	Expected	x^2	p value
CIRAD1TeakH10	219	219	219	219	7	10	1.8 ns	0.180
	219	229	219	229	13	10		
CIRAD3TeakE06	128	138	128	130	10	10	0	
	130	130	130	138	10	10		
CIRAD4TeakF02	207	207	205	207	8	10	0.8	0.371
	205	207	207	207	12	10		
CIRAD3TeakA11	258	260	256	258	5	5	1.2	0.753
			256	260	3	5		
	256	258	258	258	6	5		
			258	260	6	5		
CIRAD2TeakC03	300	308	296	300	7	5	4.8	0.187
			296	308	1	5		
	296	298	298	300	5	5		
			298	308	7	5		
CIRAD1TeakA06	171	189	171	185	10	5	6.8	0.079
			171	187	4	5		
	185	187	185	189	3	5		
			187	189	3	5		
CIRAD2TeakB07	127	127	123	127	10	10	0	NA
	123	127	127	127	10	10		
CIRAD3TeakF01	188	188	188	188	13	10	1.8	0.180
	188	194	188	194	7	10		

* Microsatellites CIRAD3TeakB02 and CIRAD1TeakF05 were not analyzed because both parents are homozygous for the same allele.

Table 4. Sampling size scenarios for clonal purity estimation in teak, based on different expected errors, proportions of off-types, and clonal population size

Population size	**500**	**1000**	**2000**	**500**	**1000**	**2000**	**500**	**1000**	**2000**
t value for n >120	1.658	1.658	1.658	1.658	1.658	1.658	1.658	1.658	1.658
p	0.1	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2
error	0.1	0.1	0.1	0.05	0.05	0.05	0.1	0.1	0.1
n	25	25	25	99	99	99	44	44	44
n corrected	24	24	24	83	90	94	40	42	43

Cross C (Table 3). Conversely, in Cross E only four individuals (full sibs) matched their SSR with their parents and two of them are off-types. In Cross E, we also detected clonality of individuals derived from two full sibs (named FSA and FSB). Individuals obtained from Cross E were previously cloned and included as blind samples for conducting the SSR analyses. Nine clones were detected as FSA and three clones as FSB. Finally, in Cross O 3 out of the 12 individuals departed from their parental expected segregation. At four loci, at least one allele was not carried by any of the parents. Although the number of samples in each cross seems low, the multiplex PCR system allowed the detection of full siblings and off-types. Controlled pollination in teak has reported low success and pollen viability, which are major constraints for fruit and seed production in teak (Palupi et al. 2010, Badilla-Valverde et al. 2023).

Sampling size scenarios

Based on different scenarios, Table 4 shows the results for sampling size scenarios. With an accepted estimation error of 10% and an expected presence of 10% off-type plants (p) in population, it was determined that a random sample of 24 plants is accurate enough for detecting them, in populations size of N = 500, 1000 and 2000 mother plants, respectively. However, if off-types are present at a 20% rate in a monoclonal population (p), sample size needs to be increased up to 40 to 43 plants. Conversely, if desired sampling error is set to around 5%, a substantial increment in sampling size is required and must consider the inclusion of at least 83 to 94 plants from the population. In this study a sample of 28 real data (putative clonal plants) was gathered in each of two monoclonal batches in the greenhouse, which revealed that 3 out of 28 (10.7%) were off-types. These results suggest that the sampling procedures followed allow for accurate and reliable information.

To the best of our knowledge, our results represent the first report detailing multiplex PCR of microsatellites for applications in a teak breeding program. This includes assessing both the purity of clonal gardens and conducting paternity testing. The two multiplex PCR reactions of all 10 SSRs reduce genotyping costs, while achieving a higher probability of identity. Paternity testing is recommended to be performed with all 10 SSRs to achieve the highest probability of exclusion, particularly for progeny testing of both full siblings and half-siblings. The number of microsatellites of our study are similar to those of other approaches followed in *Pinus radiata* (Bell et al. 2004) and *Eucalyptus grandis* (Kirst et al. 2005) for parentage analysis. Since the microsatellites used in our study have proven to be highly informative for clonal fingerprinting, protection of clonally propagated trees (Huang et al. 2016, Queiroz et al. 2023), and assessing genetic diversity and structure of teak germplasm (Verhaegen et al. 2010, Minn et al. 2014), this set of 10 SSRs can be adopted for further cross-validation by other researchers working on clonal teak breeding programs.

ACKNOWLEDGEMENTS

The authors wish to thank the Research and Extension Division of the Instituto Tecnológico de Costa Rica and GENFORES the Tree Improvement Cooperative for their financial support of this investigation. We also wish to thank Panamerican Woods S.A. for providing plant material from their second-generation controlled crosses, and Novelteak company for their support on SSR initial improvements. CENIBiot-CONARE is also acknowledged for supporting

DATA AVAILABILITY

The datasets of the microsatellite DNA fingerprinting profiles generated for all the samples during the current research are available from the corresponding author upon request. Supplementary information supporting both segregation patterns and clonal fidelity of crosses is also available.

REFERENCES

Anjos IV, Gilio TAS, Amorim AFS, Jesus JG, Chimello AM, Takizawa FH, Araujo KL and Neves LG (2023) Reassessing the genetic variability of *Tectona grandis* through high-throughput genotyping: Insights on its narrow genetic base. **PLoS ONE 18**: e0285518.

Badilla Y and Murillo O (2022) Selección clonal de *Tectona grandis* L. f. para el Pacífico seco de Costa Rica. **Uniciencia 36**: 1e14910.

Badilla-Valverde Y, Murillo-Gamboa O and Espinoza-Pizarro M (2023) Optimization of controlled pollination techniques in *Tectona grandis*, Linn f. **Revista Forestal Mesoamericana Kurú 20:** 1-23.

Balakrishnan S, Dev SA, Sakthi AR, Vikashini B, Bhasker RT, Magesh NS and Ramasamy Y (2021) Gene-ecological zonation and population genetic structure of *Tectona grandis* L.f. in India revealed by genome-wide SSR markers. **Tree Genetics and Genomes 17**: 33

Bell JC, Powell M, Devey ME and Moran GF (2004) DNA profiling, pedigree lineage analysis and monitoring in the australian breeding program of radiata pine. **Silvae Genetica 53**: 130-134.

Butler JM, Buel E, Crivellente F and McCord BR (2004) Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. **Electrophoresis 25**: 1397-1412.

Callister AN and Collins SL (2008) Genetic parameter estimates in a clonally replicated progeny test of teak (*Tectona grandis* Linn. f.). **Tree Genetics & Genomes 4**: 237-245.

Callister AN, Jiménez-Madrigal JP, Whetten RW and Murillo O (2024) Genomic predictions and candidate single nucleotide polymorphisms for growth, form, and wood properties of teak clones. **Silvae Genetica 73**: 13-23.

Flanagan SP and Jones AG (2019) The future of parentage analysis: From microsatellites to SNPs and beyond. **Molecular Ecology 28**: 544-567.

Grattapaglia D (2022) Twelve years into genomic selection in forest trees: Climbing the slope of enlightenment of marker assisted tree breeding. **Forests 13**: 1554.

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F and Petit RJ (2011) Current trends in microsatellite genotyping. **Molecular Ecology Resources 11**: 591-611.

Guimarães RA, Miranda KMS, Mota EES, Chaves LJ, Telles MPC and Soares TC (2019) Assessing genetic diversity and population structure in a *Dipteryx alata* germplasm collection utilizing microsatellite markers. **Crop Breeding and Applied Biotechnology 19**: 329-336.

Hansen OK, Changtragoon S, Ponoy B, Lopez J, Richard J and Kjaer ED (2017) Worldwide translocation of teak - origin of landraces and present genetic base. **Tree Genetics and Genomes 13**: 87.

Huang GH, Liang KN, Zhou ZZ and Ma HM (2016) SSR genotyping - Genetic diversity and fingerprinting of teak (*Tectona grandis*) clones. **Journal of Tropical Forest Science 28**: 48-58

Jhariya MK, Raj A, Sahu P, Parmeshwar S, Singh NR and Sahu KP (2014) Molecular marker - a new approach for forest tree improvement. **Ecology Environment and Conservation 20**: 1101-1107.

Kalinowski ST, Taper ML and Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. **Molecular Ecology 16**: 1099-1106.

Kirst M, Cordeiro CM, Rezende GD and Grattapaglia D (2005) Power of microsatellite markers for fingerprinting and parentage analysis in Eucalyptus grandis breeding populations. **Journal of Heredity 96**: 161-166.

Kollert W, Sandeep S and Sreelakshmy MP (2024) **Global teak resources and market assessment 2022**. IUFRO World Series, Vienna, 92p.

Koltun A, Silva PA, Torres IY, Bonifácio-Anacleto F and Yassitepe JECT (2024) Microsatellite markers in maize: challenges and guidelines for implementing multiplex SSR analyses. **Crop Breeding and Applied Biotechnology 24**: e46392411.

Mahajan R and Gupta P (2012) Molecular markers: Their use in tree improvement. **Journal of Forest Science 58**: 137-144.

Mahesh S, Vaishnav V and Kumar P (2016) Characterization and validation of teak plus trees ramets of national teak germplasm bank through microsatellites. **Tropical Plant Research 3**: 213-220.

Maia AKS, Terto J, Oliveira IF, Nascimento WF, Almeida C and Silva EF (2022) Genetic diversity and structure of *Hancornia speciosa* Gomes populations characterized by microsatellites markers. **Crop Breeding and Applied Biotechnology 22**: e38012228.

Maisuria HJ, Dhaduk HL, Kumar S, Sakure A and Thounaojam (2022) Teak population structure and genetic diversity in Gujarat, India. **Current Plant Biology 32**: 100267.

Mandal AK and Rambabu N (2001) Quantitative genetic aspects of teak improvement. In Mandal AK and Ansari SA (eds) **Genetics and silviculture of teak**. International Book Distributors, Dehra Dun, p. 134-145.

Minn Y, Prinz K and Finkeldey R (2014) Genetic variation of teak (*Tectona grandis* linn. F.) in Myanmar revealed by microsatellites. **Tree Genetics and Genomes 10**: 1435-1449.

Moya R, Marin JD, Murillo O and Leandro L (2013) Wood physical properties, color, decay resistance and stiffness in Tectona grandis clones with evidence of genetic control**. Silvae Genetica 62**: 142-152.

Murillo O, Resende MDV, Badilla Y and Gamboa JP (2019) Genotype by environment interaction and teak (*Tectona grandis* L.) selection in Costa Rica. **Silvae Genetica 68**: 116-121.

Naranjo SS, Moya R and Chauhan S (2012) Early selection of morphology and some wood properties of *Tectona grandis* L. clones. **Silvae Genetica 61**: 58-65.

Palanisamy K, Hegde M and Yi JS (2009) Teak (*Tectona grandis* Linn. f.): A renowned commercial timber species. **Journal of Forest and Environmental Science 25**: 1-24.

Palupi ER, Owens JN, Sadjab S, Sudarsono S and Solihin DD (2010) The importance of fruit set, fruit abortion, and pollination success in fruit

production of teak (*Tectona grandis*). **Canadian Journal of Forest Research 40:** 2204-2214.

Peakall R and Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. **Bioinformatics 28**: 2537-2539.

Queiroz MDM, Arriel DAA, Caldeira SF, Higa AR, Araújo SP, Sebbenn AM and Grattapaglia D (2023) Microssatellite markers for plant variety protection of clonally propagated forest trees: a case study in teak (*Tectona grandis* L.f.). **Silvae Genetica 72**: 189-199.

Quirós-Guerrero L, Albertazzi F, Araya-Valverde E, Romero RM, Villalobos H, Poveda L, Chavarría M and Tamayo-Castro G (2019) Phenolic variation among Chamaecrista nictitans subspecies and varieties revealed through UPLC-ESI(-)-MS/MS chemical fingerprinting. **Metabolomics** 15:14.

Ramasamy Y, Galeano E and Win TT (2021) **The teak genome**. Springer International Publishing, Cham, 265p.

Sanchez E, Solano W, Gatica-Arias A, Chavarría M and Araya-Valverde E (2020) Microsatellite DNA fingerprinting of *Coffea* sp. germplasm conserved in Costa Rica through singleplex and multiplex PCR. **Crop Breeding and Applied Biotechnology 20**: e27812013.

Steel R and Torrie JH (1980) **Principles and procedures of Statistics. A biometrical approach**. McGraw-Hill Book Company, New York, 633p.

Tambarussi1 EV, Rogalski M, Galeano E, Brondani GE, de Martin VF, Silva LAC and Carrer H (2017) Efficient and new method for Tectona grandis in vitro regeneration. **Crop Breeding and Applied Biotechnology 17**: 124-132

Thakor MC, Fougat RS, Kumar S and Sakure AA (2019) Sequence-related amplified polymorphism (SRAP) analysis of teak (*Tectona grandis* L.) germplasm. **Ecological Genetics and Genomics 12**: 100041.

Vaishnaw V, Mohammad N, Wali SA, Kumar R, Tripathi SB, Negi MS and Ansari SA (2015) AFLP markers for analysis of genetic diversity and structure of teak (*Tectona grandis*) in India. **Canadian Journal of Forest Research 45**: 297-306.

Verhaegen D, Fofana IJ, Logossa ZA and Ofori D (2010) What is the genetic origin of teak (*Tectona grandis* L.) introduced in Africa and in Indonesia? **Tree Genetics and Genomes 6**: 717-733.

Verhaegen D, Ofori D, Fofana I, Poitel M and Vaillant A (2005) Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f). **Molecular Ecology Notes 5**: 945-947.