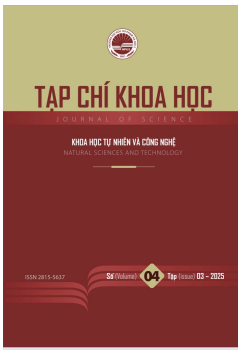
HPU2 Journal of Sciences:
Natural Sciences and Technology

Journal homepage: https://sj.hpu2.edu.vn



Article type: Research article

Genetic diversity of Shan tea (*Camellia sinensis* var. *assamica*) from Cao Son Commune, Lao Cai province using RAPD-PCR

Thuy-Lien Bui[a,b], Thi-Thuong Ngo[a], Xuan-Phong Ong[a], Van-Thiep Nguyen[c], Thi-Xuyen Ngo[d], Viet-Hong La[a*]

[a]*Hanoi Pedagogical University 2, Phu Tho, Vietnam*
[b]*Hoa Lu University, Ninh Binh, Vietnam*
[c]*Northern mountainous Agriculture & Forestry Science Institute, Phu Tho, Vietnam*
[d]*Vinh Phuc College, Phu Tho, Vietnam*

Abstract

Shan tea (*Camellia sinensis* var. *assamica*) represents a valuable indigenous genetic resource predominantly distributed in the mountainous regions of Northwest Vietnam, especially in Lao Cai province. To assess the genetic diversity of these tea populations, this study employed the Random Amplified Polymorphic DNA (RAPD)-PCR technique using six specific primers to analyze 16 leaf samples collected from Muong Khuong district. A total of 38 DNA bands were amplified, of which 32 were polymorphic, demonstrating the high effectiveness of the RAPD primer. The genetic similarity coefficient ranged from 0.62 to 0.92, suggesting a relatively wide genetic variability among the Shan tea individuals. Cluster analysis grouped the samples into four major genotypic clusters. These findings provide important molecular data to support broader conservation planning and guide selective breeding programs aimed at improving tea quality and climate resilience in indigenous Shan tea populations.

Keywords: Shan tea (Shan tuyet tea), *Camellia sinensis* var. assamica, Genetic diversity, RAPD-PCR, UPGMA

1. Introduction

Shan tea (*Camellia sinensis* var. *assamica*) is a highly valuable native plant, as the components in tea exhibit various bioactivities such as antioxidant, anti-tumor, anti-mutagenic, anti-diabetic, lipid-lowering [1], [2], resource found in the mountainous regions of Northwest Vietnam, particularly in Lao

* Corresponding author, E-mail: laviethong@hpu2.edu.vn

https://doi.org/10.56764/hpu2.jos.2025.4.3.72-79

Received date: 25-6-2025 ; Revised date: 30-7-2025 ; Accepted date: 22-11-2025

Cai province. Grown under unique climatic and edaphic conditions, these ancient tea populations exhibit distinct morphological and biochemical characteristics and are considered genetically rich [3]–[5]. Recent studies have highlighted the high genetic diversity of *C. sinensis* var. *assamica*, including Shan-like tea types, which play an essential role in conserving tea genetic resources and enhancing breeding programs [6]. Predictive models indicate that by 2050, the suitable cultivation area for *Camellia sinensis* var. *sinensis* will decrease by 15 – 32%, and for *Camellia sinensis* var. *assamica* by 32–34%. In addition, new agricultural expansion into natural areas creates conflicts with native ecosystems and poses a risk of losing valuable genetic resources [7]. However, increasing anthropogenic pressures-such as the reduction of natural forests, agricultural expansion, and climate change pose serious threats to the genetic integrity of Shan tea populations. This has underscored the urgent need for systematic genetic diversity assessments to inform effective conservation and sustainable utilization strategies [8], [9].

Globally, advances in molecular technologies have facilitated deeper insights into the genetic diversity of tea plants. Liu *et al.* (2019) employed whole-genome resequencing of the "Yunkang 10" cultivar and identified millions of single-nucleotide polymorphisms (SNPs) and insertion-deletion mutations (InDels) [10]. Zhao *et al*. evaluated the genetic and phenotypic diversity of 145 ancient tea plant germplasm samples in Guizhou Province, China, using SSR markers, and found that both genetic and phenotypic diversity were relatively high [11]. An *et al.*, investigated the genetic diversity and differentiation between *Camellia sinensis* var. *sinensis* and *Camellia sinensis* var. *assamica* by performing whole-genome sequencing of 30 cultivated varieties and three closely related wild species [12]. Kong *et al.* (2025) analyzed a global collection of 1,325 tea accessions and confirmed that Southwest China represents a primary center of origin for Shan tea, with selective signatures enriched in flavonoid biosynthetic pathways-traits associated with tea quality [6], [8]. Li *et al,* (2024) identified that *C. sinensis* var. *assamica* tea has a relatively high genetic diversity, but is comparatively less diverse than its wild relative *C. crassicolumna* [8]. The study analyzed elite tea lines using Random Amplified Polymorphic DNA (RAPD) and Inter-Simple Sequence Repeat (ISSR) markers to document genetic diversity and identify genetically distinct superior tea lines [13]. Similarly, Samarina *et al.* (2022) reported high heterozygosity (mean = 0.51) among Russian tea germplasm, reflecting the breeding potential of *Camellia* species. While whole-genome resequencing and SSR markers have revealed high genetic and phenotypic diversity in tea plants, these approaches are technically demanding and resource-intensive [14].

The Random Amplified Polymorphic DNA (RAPD) technique, developed based on polymerase chain reaction (PCR), offers a cost-effective and rapid method for detecting genetic variability without the need for prior sequence information. RAPD has proven effective in distinguishing genotypes and assessing phylogenetic relationships in plant populations [15], [16]. Findings by Naik *et al.* (2017) further indicate that RAPD markers can be used to determine genetic relationships among samples [17]. Despite Vietnam's abundance of ancient Shan tea genetic resources, molecular studies utilizing DNA-based markers remain limited, especially for populations in Lao Cai province. This study was therefore conducted to evaluate the genetic diversity of Shan tea populations in Lao Cai using RAPD markers. The findings aim to provide a scientific basis for germplasm conservation and tea breeding strategies adapted to the ecological conditions of the northern mountainous regions of Vietnam.

2. Materials and Methods

2.1. Plant Materials

A total of 16 leaf specimens from Shan tea trees (*Camellia sinensis* var. *assamica*) were collected from Cao Son commune, Lao Cai province (coded LCTT1 to LCTT16). These individuals were

deliberately selected to capture the ecological heterogeneity and morphological diversity of the Muong Khuong Shan tea population, thereby ensuring representative coverage for the genetic diversity assessment. The sampling strategy encompassed different elevations (900-1400 m), varied microhabitats, and distinct morphological types, which collectively represent the main genetic lineages of the population. Healthy, young leaves without visible signs of pest or disease damage were collected, surface-sterilized using 70% ethanol, and stored at -20°C for subsequent DNA extraction.

2.2. DNA Extraction

Total genomic DNA was extracted from young leaf tissues using a modified CTAB method [18]. The extraction protocol involved liquid nitrogen grinding, cell lysis, removal of proteins and impurities, and DNA precipitation with isopropanol. DNA concentration and purity were assessed using a NanoDrop One spectrophotometer (Thermo Scientific, USA), and the quality was confirmed by 2% agarose gel electrophoresis in 1X TAE buffer.

2.3. RAPD-PCR Amplification

PCR reactions were conducted in a final volume of 10 µL, containing 5 µL of 2X Master Mix (Bio-Helix, Taiwan), 1 µL of RAPD primer (20 µM), 2 µL of template DNA (100 ng), and 2 µL of nuclease-free water. A total of six RAPD primers were used: OPA1, OPA18, OPF, OPM, OPC, and OPO, as listed in Table 1.

Table 1. List of RAPD primers used in this study.

Primer	Sequence (5'-3')	Annealing temperature (°C)
OPA1	CAGGCCCTTC	40
OPA18	GTGCAACGTG	37
OPF	ACGGATCCTG	40
OPM	CAGTGCTGTG	37
OPC	TTCGAGCCAG	35
OPO	GGGACGTTGG	37

The six primers were selected based on previous studies reporting high polymorphism in *Camellia sinensis* and were further confirmed through preliminary screening to ensure clear, reproducible band patterns.

2.4. Agarose Gel Electrophoresis

The amplified products were separated by electrophoresis on a 2% agarose gel at 50 V for 35 minutes. A 100 bp DNA ladder (Bio-Helix) was used as a molecular size marker. Gel images were captured using a BioRad UV 2000 gel documentation system and analyzed with Quantity One 4.6 software. PCR amplification was performed using a Mastercycler X50s (Eppendorf, Germany) under the following thermal profile: initial denaturation at 95°C for 4 minutes; 40 cycles of denaturation at 94°C for 30 seconds, annealing at 36°C for 1 minute, and extension at 72°C for 2 minutes; with a final extension step at 72°C for 10 minutes. PCR products were stored at 4°C prior to electrophoresis.

2.5. Data Scoring and Analysis

The presence or absence of DNA bands was scored as binary data: "1" for presence and "0" for absence. The binary matrix was compiled using Microsoft Excel and analyzed with NTSYSpc version 2.1 [19]. A similarity matrix was generated, and a dendrogram was constructed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm.

3. Results and Discussion

3.1. Distribution and Ecological Characteristics of Shan Tea Population

The Shan tea (*Camellia sinensis* var. assamica) is distributed across several mountainous provinces of northern Vietnam, including Tuyen Quang, Son La, and Lao Cai. Among these regions, the Shan tea population found in Cao Son commune, Lao Cai province, is particularly remarkable. These tea trees are estimated to be several hundred years old and are scattered across high mountain areas at elevations ranging from 900 to 1,400 meters above sea level (Figure 1).

Figure 1. Representative individuals of the Shan tea population (Cao Son commune, Lao Cai province): *a-d: LCTT1, LCTT4, LCTT12, and LCTT16, respectively.*

The region is characterized by a cool climate year-round, high humidity, and frequent cloud cover-conditions that significantly influence the growth and phytochemical composition of the tea plants. These unique ecological conditions contribute to the distinctive quality of Shan tea, which is renowned for its rich aroma, smooth taste, and high content of bioactive compounds. The combination of high altitude, undisturbed soil, and natural biodiversity makes the Shan tea of Ta Thang a valuable genetic and cultural heritage requiring urgent conservation and sustainable utilization.

3.2. DNA Profile

The success of RAPD-PCR largely depends on the quality and integrity of the extracted genomic DNA. As a PCR-based technique, RAPD is highly sensitive to impurities and DNA degradation, which can significantly affect primer binding and amplification efficiency. Therefore, obtaining high-purity DNA is a crucial first step in ensuring reliable and reproducible results. In this study, genomic DNA was successfully extracted from all 16 samples of Shan tea using the modified CTAB protocol. The quality of the extracted DNA was assessed by 2% agarose gel electrophoresis (Figure 2).

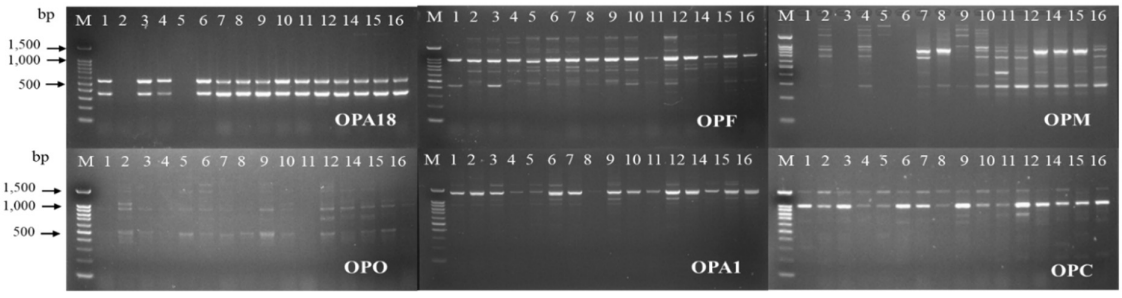


Figure 2. Electrophoresis of RAPD-PCR products in agarose gel
Lane M: 100 bp DNA ladder (Bio-Helix); Lanes 1–16: DNA from Shan tea samples LCTT1 to LCTT16.

The results showed that all DNA bands were sharp, distinct, and free from smearing, indicating high purity and integrity of the extracted DNA, suitable for subsequent RAPD-PCR amplification.

3.3. Analysis of Genetic Polymorphism

Genetic polymorphism among the 16 Shan tea samples was evaluated using six RAPD primers, and the results are summarized in Table 2.

Table 2. Summary of amplified polymorphic DNA fragments by RAPD primers.

No.	Primer	Band size (bp)	Total bands	Polymorphic bands	% Polymorphism	PIC
1	OPA18	350-550	2	0	0	0,5
2	OPF	300-3000	9	8	88,9	0,87
3	OPM	300-3000	11	11	100	0,89
4	OPO	400-1500	6	5	83,3	0,77
5	OPA1	900-2000	5	4	80	0,78
6	OPC	200-1500	5	4	80	0,69

A total of 38 DNA bands ranging in size from 200 to 3000 bp were amplified, of which 32 were polymorphic, corresponding to a polymorphism rate of 84.2%. The average polymorphism information content (PIC) was 0.75, which is classified as high, indicating the effectiveness of the RAPD markers in detecting genetic variability within the Lao Cai Shan tea population. Among the six primers, OPF and OPM produced the highest levels of polymorphism, with PIC values of approximately 0.87 and 0.89, respectively, and should be prioritized for further genetic studies. In contrast, primer OPA18 generated only two monomorphic bands and exhibited no polymorphism (PIC = 0.5), suggesting it may be unsuitable for diversity analysis in this population and could be replaced with a GC-rich primer to enhance resolution. Furthermore, the broad range of amplified fragment sizes (200-3000 bp) indicates that the selected primers target diverse genomic regions, minimizing locus-specific bias and enhancing genome-wide polymorphism coverage. The 84.2% polymorphism observed in this study is comparable to the 83.9% reported in 15 Azerbaijani tea accessions using 10 RAPD primers [20], though slightly lower than the 94.2% recorded in 15 Chinese tea genotypes with 20 primers [21]. These findings confirm the rich genetic variability within the Lao Cai Shan tea population while also suggesting the potential to uncover additional variation through the use of more primers or alternative markers such as SSRs or SNPs. The use of highly informative primers (PIC > 0.70), such as OPF, OPM, OPO, and OPA1, and strategic hybridization among genetically distant genotypes identified in the outer clusters could facilitate the conservation of rare alleles and support breeding efforts aimed at developing elite cultivars adapted to the mountainous regions of Northern Vietnam.

3.4. Genetic Relationships

To elucidate the genetic relationships among the 16 Shan tea accessions from Lao Cai, RAPD data were analyzed and a dendrogram was constructed using the Unweighted Pair Group Method with Arithmetic Mean (UPGMA). This clustering method was widely used in tea genetic studies and provides robust hierarchical clustering based on similarity coefficients. While alternative methods such as neighbor-joining can produce comparable results, UPGMA remains the preferred approach for RAPD-based analyses. The resulting dendrogram is shown in Figure 3.

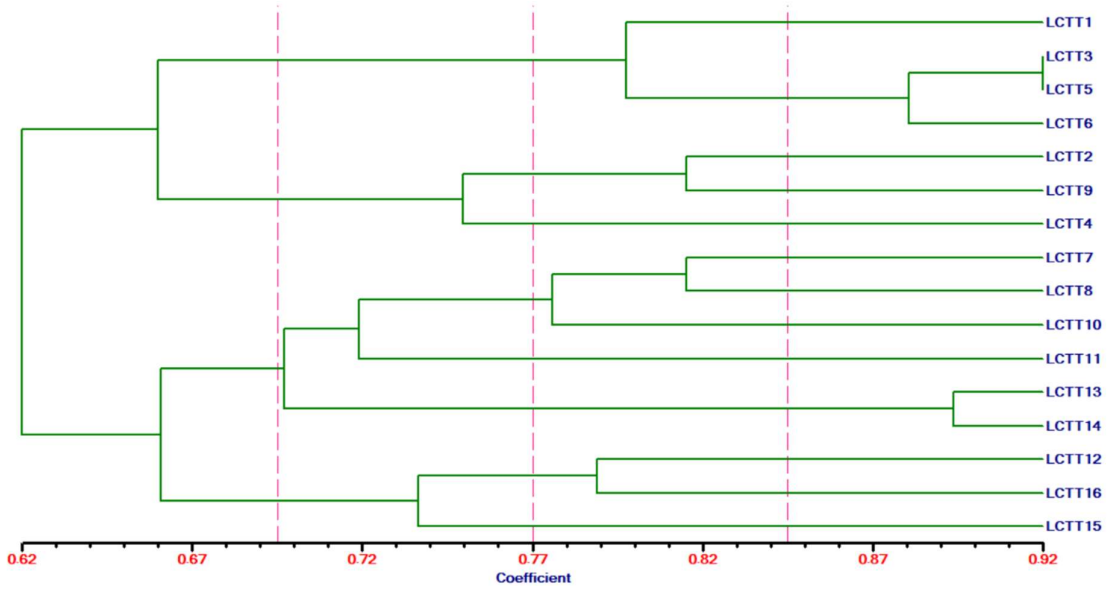


Figure 3. Dendrogram of genetic relationships among Shan tea genotypes based on RAPD markers. *The coefficient indicates the genetic similarity index.*

The genetic similarity coefficients ranged from 0.62 to 0.92, indicating a moderate-to-wide range of genetic variation (~0.30) within the studied population. No pair of individuals exhibited complete genetic identity (similarity = 1.0), suggesting that none of the samples were genetically identical under the resolution of the six RAPD primers [22], [23]. Conversely, the lowest similarity value (0.62) implies the presence of genetically divergent individuals, which may represent valuable sources for future breeding or conservation efforts. At a similarity threshold of 0.695, the 16 samples clustered into four distinct groups: Cluster I included four closely related individuals: LCTT1, LCTT3, LCTT5, and LCTT6; Cluster II comprised three intermediate genotypes: LCTT2, LCTT9, and LCTT4; Cluster III contained six individuals with moderate relatedness: LCTT7, LCTT8, LCTT10, LCTT11, LCTT13, and LCTT14 and Cluster IV consisted of three genetically distinct or peripheral individuals: LCTT12, LCTT15, and LCTT16. These results confirm the existence of significant genetic diversity within the Lao Cai Shan tea population. The presence of highly divergent genotypes in Cluster IV suggests unique alleles or adaptations that should be prioritized in conservation programs and could be strategically incorporated into breeding schemes to enhance genetic resilience.

4. Conclusion

This study employed the RAPD-PCR technique to evaluate the genetic diversity of 16 Shan tea (*Camellia sinensis* var. *assamica*) samples collected from Cao Son Commune, Lao Cai province. The results demonstrated that the selected RAPD primers exhibited strong discriminatory capacity, with a polymorphism rate of 84.2% and an average PIC value of 0.75, indicating high genetic diversity within the population. Cluster analysis using the UPGMA method revealed four major genotype groups, confirming the presence of genetically distinct lineages with potential value for conservation and selective breeding.

Acknowledgments:

This research is funded by Hanoi Pedagogical University 2 Foundation for Sciences and Technology Development under Grant Number: HPU2.2023-UT-16.

References

[1] X.-H. Meng, N. Li, H.-T. Zhu, D. Wang, C.-R. Yang, and Y.-J. Zhang, "Plant resources, chemical constituents, and bioactivities of tea plants from the genus *Camellia* section *Thea*," *J. Agric. Food Chem.*, vol. 67, no. 19, pp. 5318–5349, Nov. 2018, doi: 10.1021/acs.jafc.8b05037.

[2] S. Samanta, "Potential bioactive components and health promotional benefits of tea (*Camellia sinensis*)," *J. Am. Nutr. Assoc.*, vol. 41, no. 1, pp. 1–29, Nov. 2020, doi: 10.1080/07315724.2020.1827082.

[3] N. T. Do, *Flora of Vietnam*, vol. 8. Hanoi, Vietnam: Science and Technology Publishing House, 2007.

[4] P. H. Ho, *Vietnamese Plants and Trees*, vol. 3, p. 495. Ho Chi Minh, Vietnam: Young Publishing House, 2000.

[5] N. T. Ban, *Checklist of Plant Species of Vietnam*, vol. 3. Hanoi, Vietnam: Agriculture Publishing House (in Vietnamese), 2005.

[6] W. Kong *et al.*, "Genomic analysis of 1,325 *Camellia* accessions sheds light on agronomic and metabolic traits for tea plant improvement," *Nat. Genet.*, vol. 57, no. 4, pp. 997–1007, 2025, doi: 10.1038/s41588-025-02135-z.

[7] S. Tibpromma *et al.*, "Climate-fungal pathogen modeling predicts loss of up to one-third of tea growing areas," *Front. Cell. Infect. Microbiol.*, vol. 11, Apr. 2021, doi: 10.3389/fcimb.2021.610567.

[8] M.-M. Li *et al.*, "Genetic analyses of ancient tea trees provide insights into the breeding history and dissemination of Chinese Assam tea (*Camellia sinensis var. assamica*)," *Plant Divers.*, vol. 46, no. 2, pp. 229–237, 2024, doi: 10.1016/j.pld.2023.06.002.

[9] D. Li *et al.*, "Genetic diversity and population structure of wild ancient *Camellia tetracocca* in Pu'an, Guizhou, China," *Plants*, vol. 14, no. 11, p. 1709, Jun. 2025, doi: 10.3390/plants14111709.

[10] S. Liu *et al.*, "Characterization of genome-wide genetic variations between two varieties of tea plant (*Camellia sinensis*) and development of InDel markers for genetic research," *BMC Genomics*, vol. 20, no. 1, p. 935, 2019, doi: 10.1186/s12864-019-6347-0.

[11] Y. Zhao, R. Wang, Q. Liu, X. Dong, and D.-G. Zhao, "Genetic diversity of ancient *Camellia sinensis* (L.) O. Kuntze in Sandu County of Guizhou Province in China," *Diversity*, vol. 13, no. 6, p. 276, 2021, doi: 10.3390/d13060276.

[12] Y. An *et al.*, "Revealing distinctions in genetic diversity and adaptive evolution between two varieties of *Camellia sinensis* by whole-genome resequencing," *Front. Plant Sci.*, vol. 11, Nov. 2020, doi: 10.3389/fpls.2020.603819.

[13] S. Sharma *et al.*, "Evaluation of genetic diversity and population structure in elite south Indian tea [*Camellia sinensis* (L.) Kuntze] using RAPD and ISSR markers," *Genet. Resour. Crop Evol.*, vol. 70, pp. 381–398, 2023, doi: 10.1007/s10722-022-01433-3.

[14] L. S. Samarina *et al.*, "Genetic diversity and genome size variability in the Russian genebank collection of tea plant [*Camellia sinensis* (L.) O. Kuntze]," *Front. Plant Sci.*, vol. 12, Feb. 2022, doi: 10.3389/fpls.2021.800141.

[15] K. N. Babu *et al.*, "Random amplified polymorphic DNA (RAPD) and derived techniques," *Methods Mol. Biol.*, vol. 2222, pp. 219–247, 2021, doi: 10.1007/978-1-0716-0997-2_13.

[16] T. K. Mondal, "Molecular markers," in *Tea: Genome and Genetics*. Singapore: Springer, 2020, doi: 10.1007/978-981-15-8868-6_6.

[17] A. Naik *et al.*, "Assessment of genetic diversity in *Costus pictus* accessions based on RAPD and ISSR markers," *3 Biotech*, vol. 7, p. 70, 2017, doi: : 10.1007/s13205-017-0667-z.

[18] S. O. Rogers and A. J. Bendich, "Extraction of DNA from plant tissues," in *Plant Molecular Biology Manual*, S. B. Gelvin, R. A. Schilperoort, and D. P. S. Verma, Eds. Dordrecht: Springer, 1989, pp. 73–83, doi: 10.1007/978-94-009-0951-9_6.

[19] F. J. Rohlf, *NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System*. New York, NY, USA: Exeter Publishing, 1988, doi: 10.1093/syszoo/37.1.91.

[20] M. Huseynov, Z. Suleymanova, J. Ojaghi, and A. Mammadov, "Characterization and phylogeny analysis of Azerbaijan tea (*Camellia sinensis* L.) genotypes by molecular markers," *Cytol. Genet.*, vol. 56, no. 3, pp. 285–291, 2022, doi: 10.3103/s0095452722030057.

[21] L. Chen, Q.-K. Gao, D.-M. Chen, and C.-J. Xu, "The use of RAPD markers for detecting genetic diversity, relationship and molecular identification of Chinese elite tea genetic resources [*Camellia sinensis* (L.) O. Kuntze] preserved in a tea germplasm repository," *Biodivers. Conserv.*, vol. 14, no. 6, pp. 1433–1444, 2005, doi: 10.1007/s10531-004-9787-y.

[22] B. Martono and S. Syafaruddin, "Genetic variability of 21 tea genotypes [*Camellia sinensis* (L.) O. Kuntze] based on RAPD markers," *J. Teknol. Ind. Pertan.*, vol. 5, no. 2, pp. 77–86, 2018, doi: 10.21082/JTIDP.V5N2.2018.P77-86.

[23] C. Xi, Q. Tang, Y. Wu, J. Xu, H. Chen, and Q. Wu, "Genetic diversity and relationship of 30 tea plant germplasms in Sichuan revealed by SRAP marker," *J. Tea Sci.*, no. 2, 2013, doi: 10.3969/j.issn.1001-3601.2013.02.002.