Sains Malaysiana 54(3)(2025): 797-814
http://doi.org/10.17576/jsm-2025-5403-15

Exploring the Diversity of Shan Tea in Vietnam through SSR Markers, Morphological Traits, and Biochemical Content

(Meneroka Kepelbagaian Teh Shan di Vietnam melalui Penanda SSR, Ciri Morfologi dan Kandungan Biokimia)

KHUYNH THE BUI[1], THAI HOANG DINH[1], NGOC-THANG VU[1], VAN THU DANG[2], THI VIET HA DO[3], XINGHUI LI[4] & XUAN HOANG TRAN[3,*]

[1]*Faculty of Agronomy, Vietnam National University of Agriculture, Hanoi 131000, Vietnam*
[2]*The Vietnamese Tea Science – Technology Association, Phu Tho 293823, Vietnam*
[3]*Tea Research and Development Center, Northern Mountainous Agriculture and Forestry Science Institute, Phu Tho 293823, Vietnam*
[4]*Tea Research Institute, Nanjing Agricultural University, Nanjing 210095, China*

Received: 7 June 2024/Accepted: 14 December 2024

ABSTRACT

Shan tea (*Camellia sinensis* cv. Shan), a variety native to the mountainous regions, used to prepare 'Che Shan Tuyet'- a high quality standard tea product with cultural significance, is considered the most precious tea in Vietnam. However, little is known about its genetic diversity, composition, and variation in biochemical content across the central growing regions until now. Here, the use of 30 pairs of SSR primers selected based on their proven high utility in previous studies in tea with high polymorphisms showed that the Shan tea population exhibit rich genetic diversity, with gene diversity (H) varying from 0.47 to 0.82 and the polymorphic information content (PIC) ranging from 0.47 to 0.84. The cluster (UPGMA-based) analysis showed that 60 Shan tea accessions can be categorized into three groups with different origins. Biochemical profiles including tannin and catechins were observed to have high variation by harvest season of which the highest content was recorded during summer. Though the variation in biochemical profiles was not considerably significant among the three groups of origin, accessions from Suoi Giang (Yen Bai) significantly had lower content of tannin, EC, ECG, and EGC compared to Shan tea in Cao Bo (Ha Giang). In addition, morphology-based PCA also showed that it is practical to discriminate three groups of different origins, with the essential traits being leaf blade width, pericarp thickness, leaf area (PC1), fruit length, and fine pluck weight (PC2). The clustering of 60 Shan accessions based on morphological traits also showed consistent results with the genetic diversity analysis conducted using SSR, where accessions from Suoi Giang and Cao Bo had higher similarity levels than accessions from Tua Chua.

Keywords: Biochemical profile; genetic diversity; morphology; Shan Tea germplasm; SSR marker

ABSTRAK

Teh Shan (*Camellia sinensis* kv. Shan), varieti asal di kawasan pergunungan, digunakan untuk menyediakan 'Che Shan Tuyet' - produk teh berkualiti tinggi dengan kepentingan budaya, dianggap sebagai teh paling berharga di Vietnam. Walau bagaimanapun, sedikit yang diketahui tentang kepelbagaian genetiknya, komposisi dan variasi kandungan biokimia di seluruh rantau kawasan tengah berkembang sehingga kini. Di sini, penggunaan 30 pasang primer SSR yang dipilih berdasarkan utiliti tinggi yang terbukti dalam kajian teh terdahulu dengan polimorfisme tinggi menunjukkan bahawa populasi teh Shan menunjukkan kepelbagaian genetik yang kaya dengan kepelbagaian gen (H) berbeza dari 0.47 hingga 0.82 dan kandungan maklumat polimorfik (PIC) antara 0.47 hingga 0.84. Analisis kelompok (berasaskan UPGMA) menunjukkan bahawa 60 aksesi teh Shan boleh dikategorikan kepada tiga kumpulan dengan asal usul yang berbeza. Profil biokimia termasuk tanin dan katekin diperhatikan mempunyai variasi yang tinggi mengikut musim menuai yang mana kandungan tertinggi direkodkan semasa musim panas. Walaupun variasi dalam profil biokimia tidak begitu ketara dalam kalangan tiga kumpulan asal, aksesi daripada Suoi Giang (Yen Bai) secara signifikan mempunyai kandungan tanin, EC, ECG dan EGC yang lebih rendah berbanding teh Shan di Cao Bo (Ha Giang). Di samping itu, PCA berasaskan morfologi juga menunjukkan bahawa adalah praktikal untuk mendiskriminasi tiga kumpulan asal yang berbeza, dengan ciri penting ialah lebar helai daun, ketebalan perikarpa, luas daun (PC1), panjang buah dan petik halus berat (PC2). Pengelompokan 60 aksesi Shan berdasarkan ciri morfologi juga menunjukkan hasil yang tekal dengan analisis kepelbagaian genetik yang dijalankan menggunakan SSR, dengan aksesi daripada Suoi Giang dan Cao Bo mempunyai tahap persamaan yang lebih tinggi daripada aksesi daripada Tua Chua.

Kata kunci: Germplasma Teh Shan; kepelbagaian genetik; morfologi; penanda SSR; profil biokimia

INTRODUCTION

Tea, a widely consumed beverage known for its cultural importance and health advantages (Wambulwa et al. 2021) brings significant revenue to tea-producing countries including Vietnam. Tea genetic resources served as the background for tea breeding and quality improvement. Tea germplasm has been collected and conserved globally, including in major tea-growing countries like China, Japan, India, and Kenya (Chen & Yamaguchi 2002; Takeda 2000). As a self-incompatible and cross-fertilized plant, tea displays significant heterogeneity, resulting in extensive genetic variation and diverse phenotypic traits. Additionally, eco-geographical variances may contribute to the variation in morphological and agronomic characteristics of the tea plant. Hence, it becomes crucial and imperative to evaluate genetic diversities and establish unique fingerprints for tea plants while collecting gerplasm resources and selecting parental types for cross-breeding (Ma et al. 2010).

To date, the diversity of tea germplasm has been assessed using morphological markers (Clarke, Richter & Rathinasabapathi 2023; Pandolfi et al. 2009; Phong et al. 2016; Vo 2006). Also, thanks to the fast development of molecular tools, various markers have been used in plants including SSR, ISSR, RFLD, and AFLP (Nadeem et al. 2018). Among these, SSR markers are highly suitable for many applications in plant research as they provide high reproducibility with comprehensive genome coverage (Le et al. 2023; Powell, Machray & Provan 1996). In tea, SSR markers have been applied for the assessment of genetic relationships and population structure (Guo et al. 2021; Zhao et al. 2022), variety identification (Wang et al. 2016), genome-wide association and functional mapping (Parmar, Seth & Sharma 2022). Interestingly, besides morphology and molecular-based approaches, assessment of tea germplasm diversity was also achieved using biochemical profiles such as catechins and polyphenols (Deka et al. 2021; Kottawa-Arachchi, Gunasekare & Ranatunga 2019; Saravanan et al. 2005).

Apart from being the leading tea producer in the world, Vietnam is well known for its diverse tea culture, and tea is considered an integral part of daily life for many Vietnamese people (Vo 2007). The country has various tea-growing regions, each characterized by unique climate, soil, and altitude characteristics, contributing to a wide variety of tea germplasm. Shan tea (*Camellia sinensis* var. Shan), used to prepare a particular tea product – 'Che Shan Tuyet', known for its high-quality standard, and cultural significance (Nguyen et al. 2022), is considered the most precious tea in Vietnam. Shan tea is mainly cultivated and distributed in mountainous regions, particularly in Ha Giang, Yen Bai, and Dien Bien, which have high elevation (1200-2000 m above sea level). Currently, there are hundreds of ancient Shan tea trees with the age of centuries recognized as national heritage trees, being conserved and managed by local ethnic minority groups. Some were documented to be up to 300 years old with a trunk diameter of 0.8 to 1.2m; plant height ranges from 4 to 15 m. The Shan tea germplasms have become a primary source of income and significantly improved the livelihoods of the local people.

Previous studies have reported the genetic diversity of tea resources in Vietnam. Vo (2007) used 65 ISSR markers to assess 69 accessions, including local, wild, selected, and imported teas and noted the large genetic variance with a similarity coefficient range from 0.13 to 0.95. Phong et al. (2016) accessed a 15-tea accession population, including imported resources from China and India, hybrids, and improved cultivars, using 6 iPBS markers. They found a moderate range from 0.52 to 0.87 in the similarity coefficient. Despite its well-recognized importance and widespread cultivation, research on Shan tea in Vietnam is limited. Until now, there has been no assessment of the genetic diversity of Shan tea either by using morphology or molecular markers. Also, little is known about the composition and variation in the bio-chemical content of Shan tea across the central growing regions in Vietnam, except in the study of Nguyen et al. (2022) where the authors only confirmed the unique biochemical profile of Shan tea compared with that in Chinese tea (*Camellia sinensis* var. Sinensis). Here, we aim to explore the genetic diversity of Shan tea accessions grown in the mountainous areas of Vietnam using SSR markers and morphological characteristics. Also, the bio-chemical content in the leaf including tannin and catechin content, was evaluated to understand the distinctive characteristics of Shan tea across central growing regions.

MATERIALS AND METHODS

PLANT MATERIALS

We collected 60 accessions of Shan tea in 3 mountainous regions in the north of Vietnam, namely (i) Suoi Giang – Yen Bai province, (ii) Tua Chua – Dien Bien province, (iii) Cao Bo– Ha Giang province (S. Figure 1). The description of the Shan tea germplasm used in this study is given in Table 1. For each tea sample, 10 g of fresh leaf tissue (the youngest fully leaf was chosen) was sampled, frozen with liquid nitrogen, and stored at a low temperature (-70 °C) for future use. For total DNA extraction, the CTAB method was employed following instructions given by Doyle and Doyle (1987).

SSR AMPLIFICATION AND DATA ANALYSIS

Thirty pairs of SSR primers (listed in Table 2, and Supplementary Table S2) were selected to evaluate the genetic relationship between 60 Shan tea accessions, based on their proven high utility in previous studies in tea with high polymorphisms (Bali et al. 2013; Freeman et al. 2004; Sharma et al. 2009; Taniguchi et al. 2012; Tan et al. 2013). PCR was used to amplify total DNA using 10 μL reaction mixtures containing 50 ng of template DNA and 10X PCR

TABLE 1. Name, type, and origin of Shan tea germplasm used in this study

Accession code	Name	Type	Origin
T1	TC-ĐB1	Landrace	Tua Chua, Dien Bien
T2	TC-ĐB 2	Landrace	Tua Chua, Dien Bien
T3	TC-ĐB 3	Landrace	Tua Chua, Dien Bien
T4	TC-ĐB 4	Landrace	Tua Chua, Dien Bien
T5	TC-ĐB 5	Landrace	Tua Chua, Dien Bien
T6	TC-ĐB 6	Landrace	Tua Chua, Dien Bien
T7	TC- ĐB 7	Landrace	Tua Chua, Dien Bien
T8	TC-ĐB 8	Landrace	Tua Chua, Dien Bien
T9	TC-ĐB 9	Landrace	Tua Chua, Dien Bien
T10	TC-ĐB 10	Landrace	Tua Chua, Dien Bien
T11	TC–ĐB 11	Landrace	Tua Chua, Dien Bien
T12	TC-ĐB 12	Landrace	Tua Chua, Dien Bien
T13	TC-ĐB 13	Landrace	Tua Chua, Dien Bien
T14	TC- ĐB 14	Landrace	Tua Chua, Dien Bien
T15	TC-ĐB 15	Landrace	Tua Chua, Dien Bien
T16	TC-ĐB 16	Landrace	Tua Chua, Dien Bien
T17	TC-ĐB 17	Landrace	Tua Chua, Dien Bien
T18	TC-ĐB 18	Landrace	Tua Chua, Dien Bien
T19	TC-ĐB 19	Landrace	Tua Chua, Dien Bien
T20	TC-ĐB 20	Landrace	Tua Chua, Dien Bien
T21	SG-YB 1	Landrace	Suoi Giang, Yen Bai
T22	SG-YB 2	Landrace	Suoi Giang, Yen Bai
T23	SG-YB 3	Landrace	Suoi Giang, Yen Bai
T24	SG-YB 4	Landrace	Suoi Giang, Yen Bai
T25	SG-YB 5	Landrace	Suoi Giang, Yen Bai
T26	SG-YB 6	Landrace	Suoi Giang, Yen Bai
T27	SG-YB 7	Landrace	Suoi Giang, Yen Bai
T28	SG-YB 8	Landrace	Suoi Giang, Yen Bai
T29	SG-YB 9	Landrace	Suoi Giang, Yen Bai
T30	SG-YB 10	Landrace	Suoi Giang, Yen Bai
T31	SG-YB 11	Landrace	Suoi Giang, Yen Bai
T32	SG-YB 12	Landrace	Suoi Giang, Yen Bai
T33	SG-YB 13	Landrace	Suoi Giang, Yen Bai
T34	SG-YB 14	Landrace	Suoi Giang, Yen Bai
T35	SG-YB 15	Landrace	Suoi Giang, Yen Bai
T36	SG-YB 16	Landrace	Suoi Giang, Yen Bai
T37	SG-YB 17	Landrace	Suoi Giang, Yen Bai
T38	SG-YB 18	Landrace	Suoi Giang, Yen Bai
T39	SG-YB 19	Landrace	Suoi Giang, Yen Bai
T40	SG-YB 20	Landrace	Suoi Giang, Yen Bai
T41	CB-HG 1	Landrace	Cao Bo, Ha Giang

continue to next page

continue from previous page

T42	CB-HG 2	Landrace	Cao Bo, Ha Giang
T43	CB-HG 3	Landrace	Cao Bo, Ha Giang
T44	CB-HG 4	Landrace	Cao Bo, Ha Giang
T45	CB-HG 5	Landrace	Cao Bo, Ha Giang
T46	CB-HG 6	Landrace	Cao Bo, Ha Giang
T47	CB-HG 7	Landrace	Cao Bo, Ha Giang
T48	CB-HG 8	Landrace	Cao Bo, Ha Giang
T49	CB-HG 9	Landrace	Cao Bo, Ha Giang
T50	CB-HG 10	Landrace	Cao Bo, Ha Giang
T51	CB-HG 11	Landrace	Cao Bo, Ha Giang
T52	CB-HG 12	Landrace	Cao Bo, Ha Giang
T53	CB-HG 13	Landrace	Cao Bo, Ha Giang
T54	CB-HG 14	Landrace	Cao Bo, Ha Giang
T55	CB-HG 15	Landrace	Cao Bo, Ha Giang
T56	CB-HG 16	Landrace	Cao Bo, Ha Giang
T57	CB-HG 17	Landrace	Cao Bo, Ha Giang
T58	CB-HG 18	Landrace	Cao Bo, Ha Giang
T59	CB-HG 19	Landrace	Cao Bo, Ha Giang
T60	CB-HG 20	Landrace	Cao Bo, Ha Giang

buffer (purchased from Thermo Fisher Scientific). The PCR procedures started with an initial denaturation step for 5 min at 95 °C, followed by 35 cycles at 94 °C for 45 s, 1 min at 52 °C and 10 min at 72 °C for the final extension. PCR products were then separated on an 8% polyacrylamide gel, visualized under ultraviolet light, and manaully scored for '1' with the presence of the polymorphic band or '0' for its absence. Power Marker V3.25 software (Liu & Muse 2005) was employed for the estimation of the following parameters: (i) a total number of alleles (N_A), (ii) genotype (N_o), (iii) observed heterozygosity (Ho), (iv) gene diversity (H), (v) polymorphism information content (PIC). Jaccard similarity coefficients were calculated and the clustering analysis of 60 Shan tea accessions was done following the unweighted pair group method of arithmetic mean (UPGMA) via the NTSYS-PC ver.2.10 software (Rohlf 2000).

QUANTIFICATION OF TANNIN AND CATECHIN CONTENT IN LEAF

For each Shan tea accession, the youngest fully developed leaves from 5 plants, counted as one replicate were carefully selected, dried in 72 h until unchanged mass, and used for biochemical profile analysis. The total tannin content in tea leaves was quantified using the Folin-Ciocalteu reagent by the GB/T8313-2008 standard (Chinese National Standard 2008). Briefly, the total tannin was extracted by 70% methanol aqueous solution in a 70 °C water bath, and then the –OH group in tannin was oxidized by Folin–Ciocalteu reagent, the maximum absorption wavelength was 765 nm, and the total tannin content was quantified by gallic acid (GA) as calibration standard as described by Huang et al. (2022).

The contents of catechins in the leaf were measured by high-performance liquid chromatography (HPLC), following the same standard as described by Zhang et al. (2022). Briefly, the catechins, including EGC, EC, GC, EGCG, and ECG in tea leaves samples were ground and extracted in a 70% methanol aqueous solution at 70 °C (in a water bath) and then detected with a wavelength of 278 nm. Two mobile phases were employed: A (containing 9% acetonitrile, 2% acetic acid with 124 10 mg/mL EDTA) and B (with 80% acetonitrile, 2% acetic acid with 10 mg/mL EDTA). The gradient elution program was as follows: A: B (100% in A:0% in B) for 10 min, followed by A: B = (68% in A:32% in B) for 15 min, and A: B (100% in A:0% in B) for another 10 min. Pure chemicals, including EGCG, ECG, EGC, GC, and GCG, were purchased from Sigma (USA) and used as standards. All measurements were made with three replications.

PHENOTYPIC DIVERSITY EVALUATION AND ANALYSIS

The 4th leaf counted from the bud of each Shan tea accession was selected to measure the leaf blade with, leaf blade length, and leaf blade area (single leaf area). In addition, fine pluck (1 bud and 2 true leaves) was

collected to measure the fine pluck length (measured from the 2nd leaf's petiole position on the branch to the highest point of bud), and fine pluck weight (g). Bud pubescence density was also graded on a scale from 1-3 (1: low, 2: moderate, 3: high). For the characteristics of the flower, the following parameters were measured: flower diameter, stamen length, stamen number, and pistil length. Ovary pubescence density level was evaluated on a scale from 1-3 (1: low, 2: moderate, 3: high). Uniform and mature fruits were then collected to measure fruit length, width, and pericap thickness. Seed characteristics were assessed via seed length and seed width. All the measurements were done with 10 replicates and the average values were recorded. All of the measured traits' values were then applied to the principal component analysis (PCA) using the prcomp function, and the clustering analysis based on Euclidean distance in R (version 4.0.5).

RESULTS AND DISCUSSION

GENETIC DIVERSITY SHOWED BY SSR MARKERS

Collection and evaluation of diverse germplasms are imperative for variety selection, crop improvement and breeding (Tao et al. 2023). The genetic diversity assessment of tea resources in Vietnam was previously reported by Vo (2007) using both 150 SSR and ISSR markers and Phong et al. (2016) using iPBS markers. However, to the best of our knowledge, there is no report on the genetic diversity of Shan tea in Vietnam. Here, we first investigated DNA polymorphisms among Shan tea germplasms using SSR markers and aimed to evaluate the genetic relationship among 60 accessions in the Shan tea population. Employing 30 SSR primer pairs, 88 alleles and 122 genotypes were identified (Figure 1, Table 2). In addition, the average number of alleles per locus varied from 2 to 6 (average of 3.90). This was in a considerable range with previous polymorphisms reported in several crops such as tea (Zhou et al. 2019), maize (Labate et al. 2003), and cucumber (Mu et al. 2003). The maximum number of alleles per locus is 6 (recorded in Csin07, TUGMS27), while the minimum is 2 (observed in CsFM1384, MSE0313, Csin46). The maximum number of genotypes was 9 (CsFM1051, TUGMS27), the minimum was 3 (found in CsFM1384, CsFM1509, MSE0313, MSG0423, Csin46).

PIC is essential for evaluating plant population's genetic diversity (Le et al. 2023; Ni, Colowit & Mackill 2002). Here, the PIC values varied from 0.47 (in Csin46) to

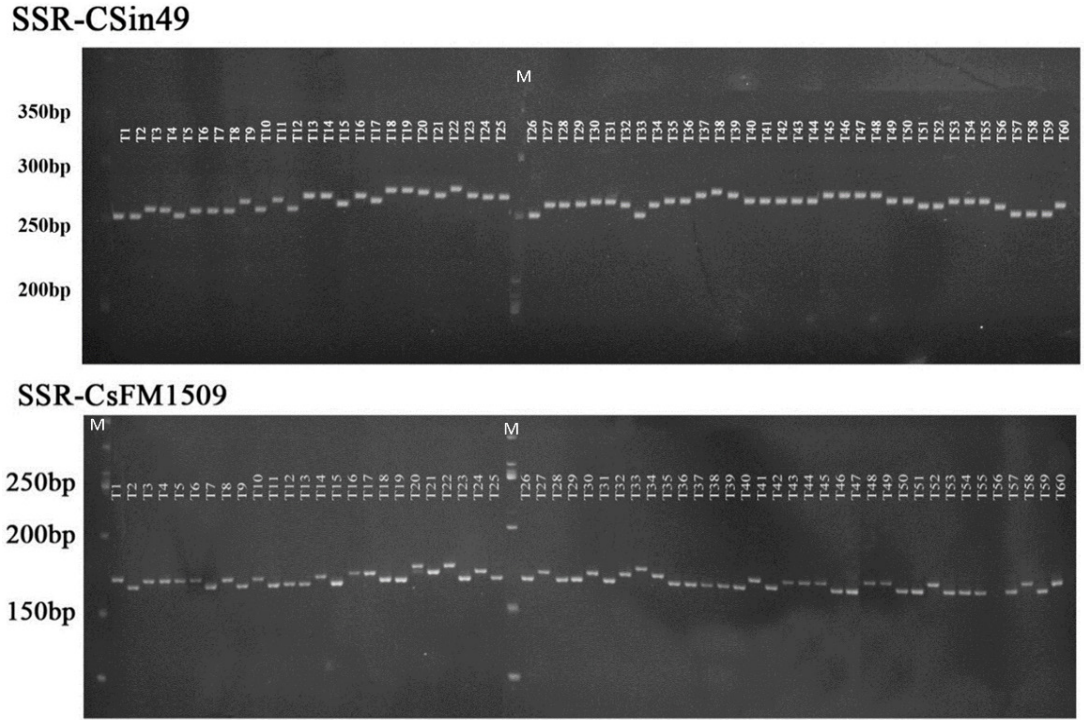

FIGURE 1. The amplification result of primer pair Csin49 and CsFM1509 for 60 Shan tea germplasm accessions. The Lines (T1-T60) are marked as abbreviated in Table 2. The lanes where DNA molecular size marker (Low Range) was loaded are marked as marker (M)

TABLE 2. Amplification information of 30 SSR primer pairs in 60 Shan tea accessions

Primer code	Repeat motif	Expected product size (bp)	T_a (^{0}C)	Allele No(a)	Genotype (No)	Gene diversity (H)	Obs. Heterozygosity (Ho)	*PIC*
CsFM1051	(TTG)8	110-125	46	5.00	9.00	0.80	0.133	0.80
CsFM1550	(TAG)7	108-116	48	4.00	8.00	0.73	0.167	0.74
CsFM1599	(TCC)7	167-178	48	4.00	7.00	0.69	0.117	0.71
CsFM1384	(ACC)8	236-240	46	2.00	3.00	0.49	0.167	0.49
CsFM1509	(GATGAA)6	167-175	47	3.00	3.00	0.57	0.000	0.65
A17	TC)4/T8	183-200	49	3.00	6.00	0.66	0.133	0.66
A28	(TC)18	133-149	44	4.00	7.00	0.75	0.100	0.75
A38	(TC)6	246-254	47	4.00	5.00	0.74	0.033	0.73
A55	(AATCC)3/(CGC)7	135-144	49	5.00	8.00	0.79	0.117	0.80
A166	(TTA)7	267-280	46	3.00	6.00	0.66	0.117	0.66
MSG0533	(AG)18, (GA)3	225-230	50	4.00	5.00	0.73	0.050	0.73
MSG0380	(AG)21, (TA)4	275-281	50	3.00	4.00	0.67	0.033	0.67
MSE0313	(AG)3, (AG)12	225-230	49	2.00	3.00	0.50	0.133	0.50
MSG0423	(TC)4, (TC)13, (AC)7	153-165	51	3.00	3.00	0.65	0.000	0.65
MSE0291	(CA)3, (TC)9	215-230	48	3.00	4.00	0.66	0.033	0.66
MSG0681	(AG)18	212-235	49	5.00	6.00	0.78	0.050	0.79
Csin04	(AAG)20(GAA)10	110-120	46	4.00	7.00	0.74	0.167	0.84
Csin06	(CCA)4(CAG)9	496-550	51	5.00	7.00	0.79	0.133	0.80
Csin07	(GCT)9(GnT)5	320-400	49	6.00	8.00	0.78	0.267	0.80
Csin24	(GCT)7	145-160	45	4.00	5.00	0.73	0.083	0.71
Csin41	(A)17(GT)9	293-300	45	3.00	4.00	0.66	0.117	0.66
Csin46	(GCT)7	244-250	48	2.00	3.00	0.47	0.133	0.47
Csin49	(CAG)4	250-270	47	5.00	5.00	0.77	0.000	0.77
Csin68	(CAG)8	225-250	45	3.00	5.00	0.65	0.267	0.66
Csin71	(GTn)6	400-450	48	4.00	6.00	0.75	0.100	0.75
TUGMS102A	(GGAAA)12	216-237	47	4.00	5.00	0.74	0.133	0.74
TUGMS27	(GA)20	235-258	46	6.00	9.00	0.82	0.167	0.77
TUGMS82	(CAT)8	259-304	46	5.00	7.00	0.77	0.200	0.78
TUGMS73	(TAA)12	256-295	46	5.00	7.00	0.72	0.167	0.74
CamsinM14	(GA)16	156-190	47	4.00	5.00	0.72	0.100	0.71
Mean	-	-	-	3.90	5.67	0.70	0.114	0.71

**PIC*: Polymorphism information content; Ta: Annealing temperature

0.84 (in Csin04). Based on these values, 28 SSR markers are categorized as highly informative (with PIC values >0.5) and 2 SSR markers are considered reasonably informative (with PIC values in the range of 0.25 to 0.5). The average PIC value was 0.71, suggesting that the Shan population in this study had a high genetic diversity. Notably, this value was higher than Fang et al. (2011) and Zhou et al. (2019) for 185 tea accessions with an average PIC of 0.495 and 68 accessions of Dongting Biluochun tea with an average PIC of 0.51, respectively. However, the average PIC recorded in this study was lower an average of 0.862 for the tea plant (Liu et al. 2017). Differences in PIC values among these studies can be explained by the differences in the tea materials selected and the SSR markers employed.

Sixty samples of Shan tea in the study had a genetic similarity ranging from 0.61 to 0.97. This result is different from previous works on the genetic diversity of tea in Vietnam using SSR markers (Vo 2007) and iPBS markers (Phong et al. 2016), which demonstrated a more significant variation, with a range from 0.09 to 1, and 0.39 to 0.86,

respectively. The considerable variation in the similarity index values reported in other studies could be attributed to the various tea germplasm tested compared with focusing only on one type of tea as in our study. For instance, Vo (2007) assessed the genetic variation among tea types, including local types, imported large-leaved and small-leaved China types, imported Assam, cross varieties, and imported lines from Sri Lanka and Japan. However, compared to previous studies focusing on a single type of tea, Vietnam's Shan tea displayed a greater genetic diversity compared with other tea germplasms previously reported, including the 28 Huangjincha cultivar germplasm (Yang et al. 2009), and the 36 clonal tea germplasm in China (Yao et al. 2007). Thus, this result further supports the fact that Vietnam is the origin center of Shan tea.

Based on the similarity index, the cluster analysis successfully categorized 60 Shan accessions into three groups based on their origin. Accessions collected at Tua Chua-Dien Bien (group I) were separated from samples collected at Suoi Giang-Yen Bai (group II) and Cao Bo - Ha Giang (group III) at a cutoff value of 0.61. Within group I, the genetic similarity ranged from 0.75 to 0.94 (Figure 2). With the other two groups, group II had a similarity level of 64% compared to group III. Similar to the first group, accessions in groups II and III had relatively high genetic similarities within groups, ranging from 0.77 to 0.97 and from 0.74 to 0.93, respectively. The Shan tea accessions from Suoi Giang (Yen Bai province) and Tua Chua (Ha Giang province) displayed a higher genetic similarity compared to those from Tua Chua (Dien Bien province). This difference in genetic similarity might be due to the geographical distance, as Yen Bai and Ha Giang are neighboring provinces, while Dien Bien province is isolated from them.

VARIATION IN TOTAL TANNIN AND CATECHIN CONTENT AMONG 60 ACCESSIONS OF SHAN TEA

The quality of tea is influenced mainly by biochemical components such as tannin and catechin present in tea leaves. Here, we examined the tannin and catechin levels in 60 varieties of Shan tea at different harvest times. Our findings showed significant variation in the levels of these components depending on the harvest season. The same result was also reported by Zheng et al. (2008) and Sharma, Joshi and Gulati (2011). The seasonal variation in biochemical profiles in tea could be explained by the seasonal fluctuation of environmental factors including temperature, sunlight intensity, water availability, and rainfall patterns (Ahmed et al. 2019; Paiva et al. 2021), and the interactions among these factors, and thus, it's challenging to isolate these environmental effects from each other (Tounekti et al. 2013). The highest concentrations of either tannin or different types of catechin were observed during the summer harvest, followed by the spring harvest. At the same time, the lowest was recorded during the fall harvest (Figures 3 & 4). Of note, EGCG was the most abundant among different types of catechin, and tea leaves also exhibited higher levels of EGCG in summer than in spring and fall seasons. This aligns with previous studies which suggested that tea leaves accumulated high levels of catechin to prevent themselves from damage caused by ultraviolet rays during summer (Ye et al. 2022; Zagoskina et al. 2005; Zheng et al. 2008). In addition, longer daylight hours in summer, and intense sunlight during summer months were believed to enrich more catechins than in spring and fall (Ahmed et al. 2019; Jhou et al. 2025; Yao et al. 2005).

Apart from seasonal variation, tannin and catechin content did not show significant variations among three groups of Shan tea with different origins. One of the reasons for this is that tea accessions all belong to the same type of tea-Shan tea. However, significant differences among the three groups were noted in the EC, ECG, and EGC content. In contrast, EGCG and GC content did not significantly vary across groups (Figure 4). Overall, germplasm from CB tended to have the highest content of EC, ECG, and EGC, whereas accessions from SG were observed with the lowest values. For the total catechin content, germplasm in CB exhibited a higher value than in SG (at $p=0.05$). At the same time, no significant variation was recorded between CB and TC or TC and SG, respectively (Figure 5). The difference in the level of catechin in Shan tea from Cao Bo and Suoi Giang may be due to the difference in their cultivation altitude. Cao Bo's Shan tea was planted at a lower altitude (around 1200 m above sea level) than Suoi Giang's (over 1300 m above sea level). According to Gong et al. (2020), the catechin level in tea leaves was negatively correlated with the altitude of the cultivation area. Also, Tian et al. (2024), when investigating the effects of 3 elevations (86, 256, and 880 m) on two tea cultivars named Mingke 1' (MK) and 'Fuyun 6' (FY) concluded that the content of catechins decreased with increasing altitude, explaining the expected lower bitterness of high-mountain tea.

PHENOTYPIC VARIATION AMONG 60 SHAN TEA ACCESSIONS

Results from morphology-based PCA showed that the first two dimensions (or PCs) captured 44.5% of the total variation in all studied morphological traits of 60 Shan tea accessions (Figure 6(a)). While leaf blade width, pericarp thickness, and leaf area are key traits contributing to the first dimension, the second dimension was characterized by critical traits, including fruit length, fine pluck weight (weight of shoot with bud and two leaves), seed length and seed width (Figure 6(b) & 6(c)). Notably, the PCA successfully distinguished three groups of Shan tea based on their geographical distribution, as shown in Figure 6(A). Accessions in SG were recorded with negative scores of PC1, contrasting with predominantly positive scores observed in accessions from CB and TC. This was further supported as accessions from SG had significantly higher

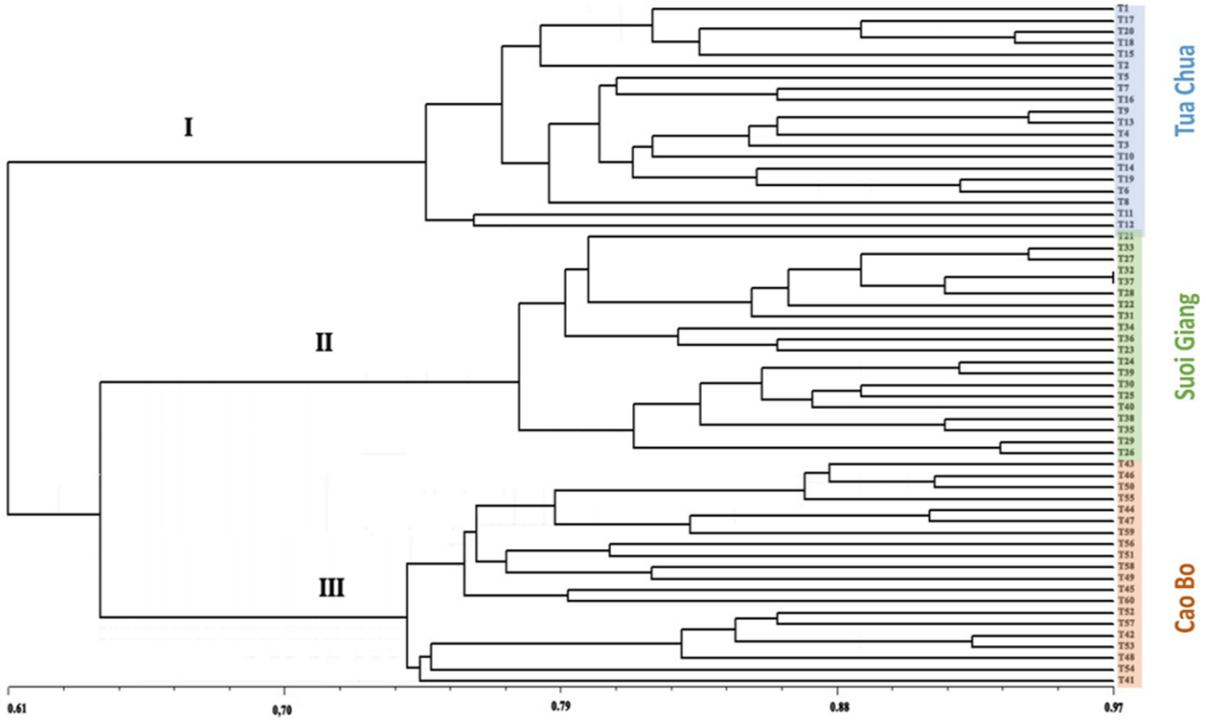

FIGURE 2. Dendrogram of 60 Shan tea germplasm accessions based on SSR data as clustered using the UPGMA method after a 1000 replicate bootstrap analysis.

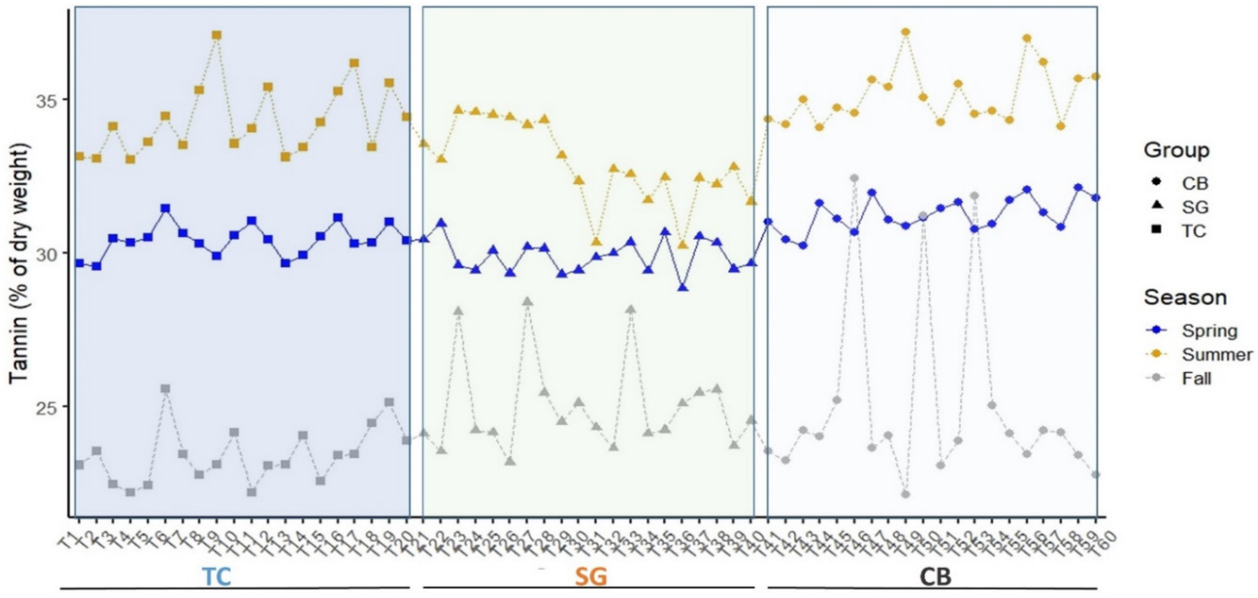

FIGURE 3. Seasonal variation in tannin content (total tannin) among 60 accessions of tea belonging to 3 groups of geographical distribution. CB: Cao Bo, Ha Giang; SG: Suoi Giang, Yen Bai; TC: Tua Chua, Dien Bien

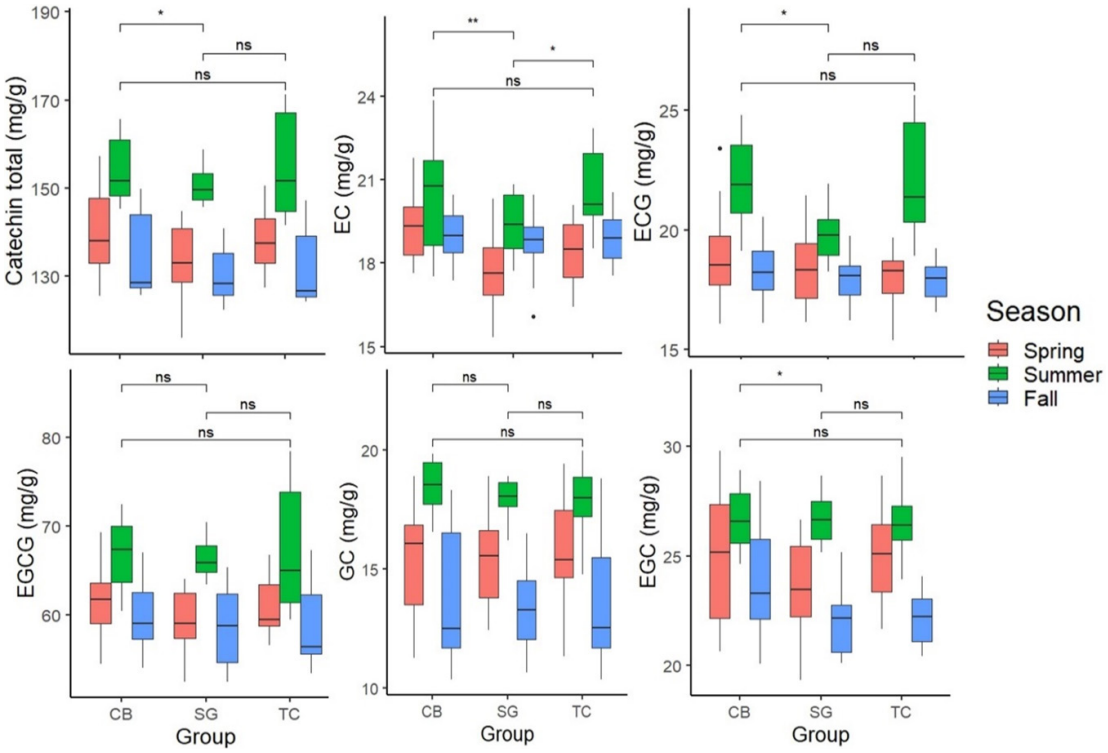

FIGURE 4. Variation in leaf catechin content by groups of origin and seasons of harvest. Significance was applied to compare the mean values among three groups using the combined data in 3 seasons. ***p< 0.001; **p< 0.01; *p< 0.05; ns, not significant

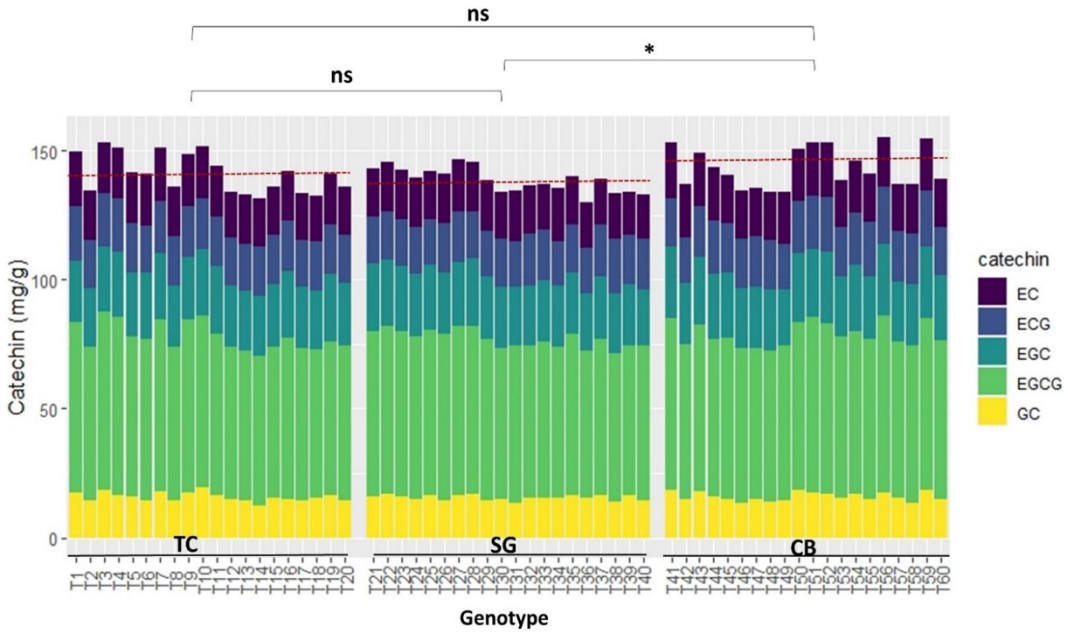

FIGURE 5. Variation in total catechin content among 60 accessions of tea belonging to 3 groups of geographical distribution. The average value of catechin in 3 harvest seasons was used. Significance was applied to compare the mean of values among three groups of geographical distribution. The red dot line indicates the average value of total catechin in each group. Significance levels: ***p< 0.001; **p< 0.01; *p< 0.05; ns, not significant

leaf blade width than germplasm from TC and CB (Figure 7(a) & 7(b)). In addition, the separation between CB and TC was evident in PC2, where accessions in CB exhibited negative scores while accessions in TC showed positive scores. This was confirmed in boxplot analysis as Shan tea in CB exhibited significantly higher fruit length than Shan tea from TC (Figure 7(c)). Thus, these parameters could be considered critical indicators for classifying Shan tea germplasm in Vietnam.

In an attempt to group the 60 accessions in the Shan tea germplasm based on their morphological characteristics (data given in Supplementary Table 1), the Euclidean distance was computed and employed for clustering (Figure 8). The results showed that 60 Shan tea accessions were divided into three groups. Notably, all accessions in group I originated from TC, while group II exclusively comprised six accessions from CB. Group III, consisting of the most significant accessions (44), further branched into III-A and III-B sub-groups. Within III-A, most of the 24 accessions were from CB, with 14 accessions and only two from SG, namely T27 and T35. Noteworthy, the composition of sub-group IIIB was predominantly from SG with 18 out of 20 accessions. Overall, Suoi Giang (Yen Bai) accessions showed high similarity levels with those from Cao Bo (Ha Giang). On the other hand, most accessions from Tua Chua (Dien Bien) were found to be different from those from Suoi Giang and Cao Bo. This observation was consistent with the genetic diversity analysis conducted using SSR markers.

The first tea classification was introduced by Sealy (1958) using basic leaf characteristics and later revised by Wight (1962) incorporating morphological traits such as leaf size, leaf shape pistil length, and flower size. Here, by investigating 13 morphological traits, we showed the critical traits including (leaf blade width) and fruit size (fruit length) which can help to discriminate 3 groups of Shan tea based on origin. This was consistent with the findings of Rajkumar et al. (2010) who reported that leaf and fruit characteristics showed a great degree of variability and were important in differentiating tea accessions. Taken with the genetic diversity showed by SSR markers, the study suggests that using both phenotype and genotype-based clustering can provide a unique and more in-depth understanding of Shan tea, which can help facilitate the breeding and selection of Shan tea in Vietnam.

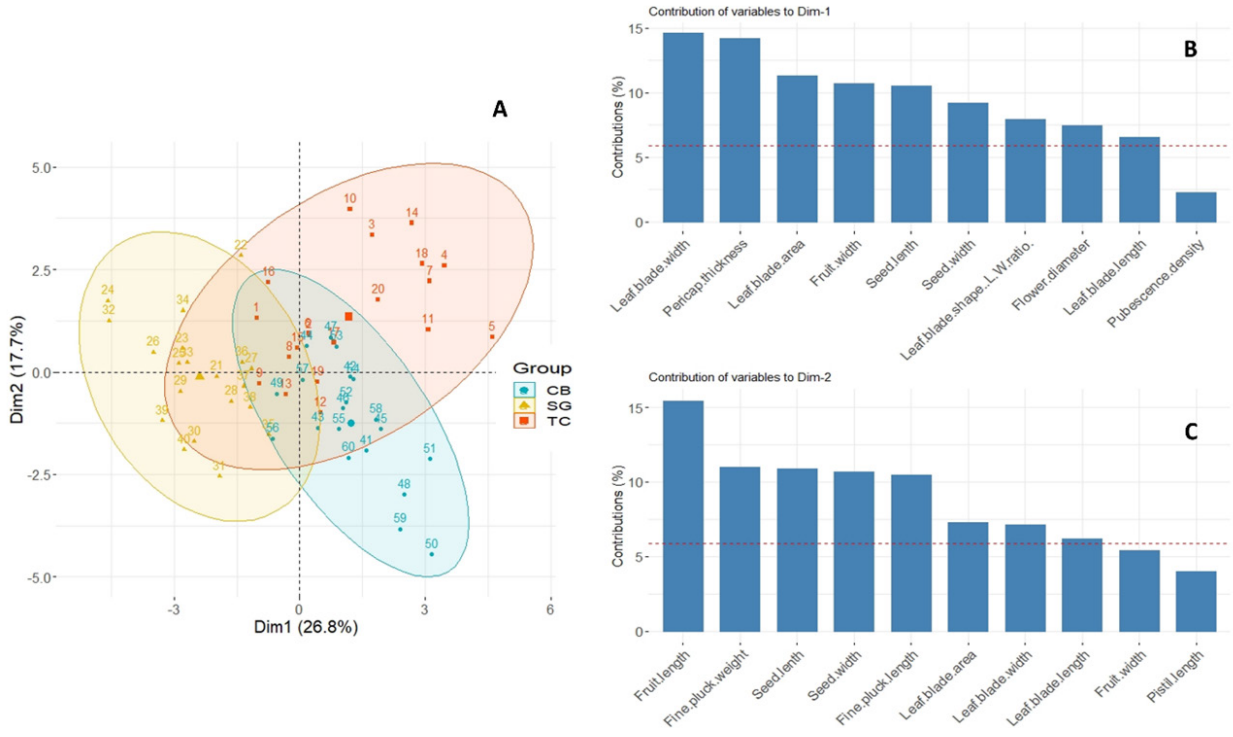

FIGURE 6. Principal component analysis (PCA) using phenotypic traits among 60 Shan tea accessions. Biplot showing phenotypic variation among three groups of Shan tea based on origin (A), and phenotypic traits with highest contribution to the first dimension (B) and the second dimension (C) in PCA

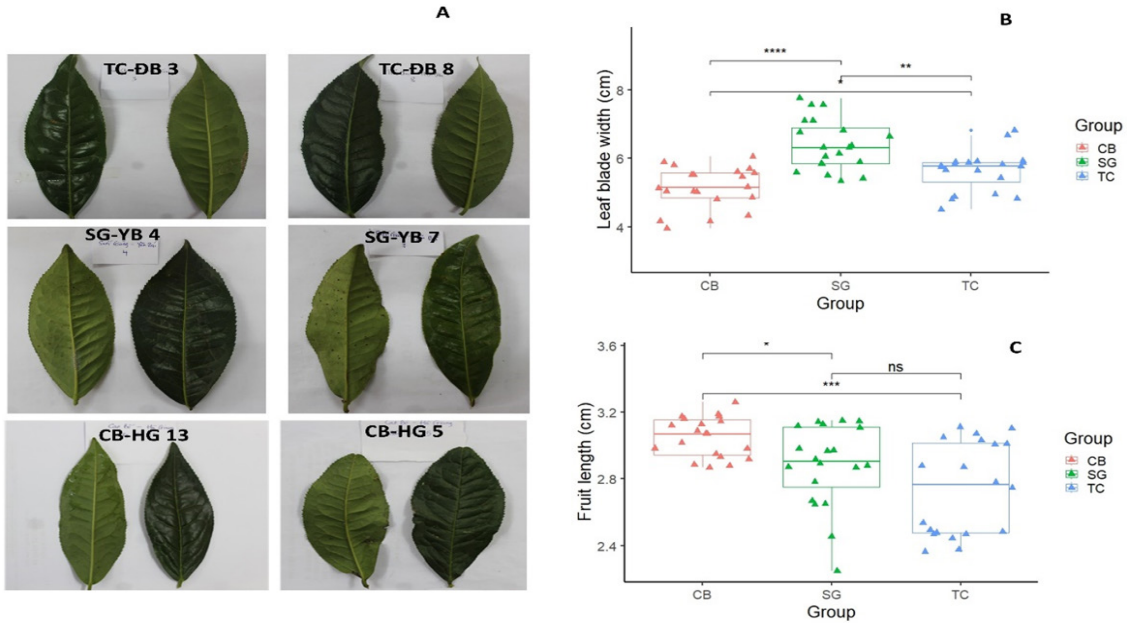

FIGURE 7. Leaf morphology of several Shan accessions (A), box plots showing the comparison of leaf blade width (B) and fruit length (C) among three groups of geographical distribution. Significance levels: ***p< 0.001; **p< 0.01; *p< 0.05; ns, not significant

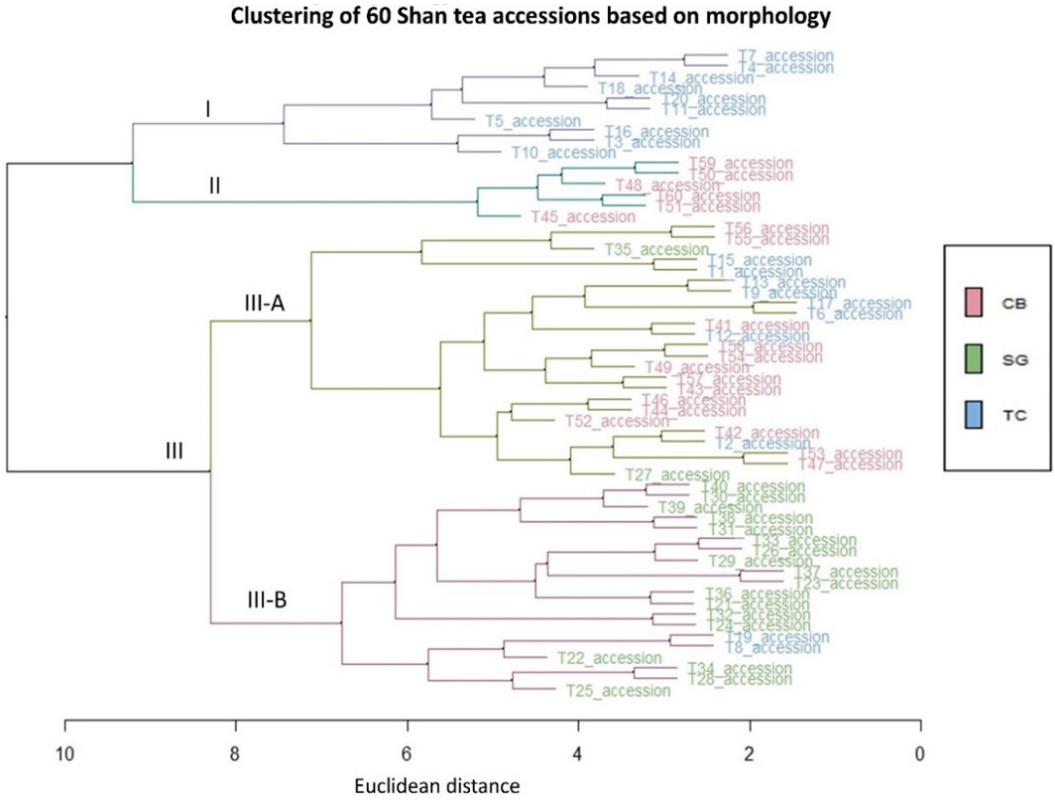

FIGURE 8. Clustering of 60 Shan tea accessions based on morphological traits

CONCLUSIONS

Our first comprehensive study of Shan tea germplasms in Vietnam showed significant genetic and phenotypic diversity among the 60 accessions. Using SSR markers, we observed a wide range of DNA polymorphisms and high polymorphic information content (PIC), suggesting a high genetic variation within the Shan tea population. The cluster analysis based on similarity coefficients identified three distinct groups, reflecting their geographical origins, with accessions from Suoi Giang and Cao Bo (Ha Giang) showing more remarkable genetic similarity than those from Tua Chua (Dien Bien). Biochemical profiles, including tannin and catechin levels, showed seasonal variations, with higher concentrations observed during the summer harvest. Variation in biochemical profiles among the three groups of Shan tea was not considerable. However, significant differences were still noted, particularly between Suoi Giang and Cao Bo tea accessions. These variations may be influenced by cultivation altitude and local environment. Morphology-based principal component analysis (PCA) effectively discriminated three groups based on traits like leaf blade width, pericarp thickness, leaf area, fruit length, and fine pluck weight. These morphological traits provided additional insights into the diversity of Shan tea and supported the genetic clustering results obtained through SSR markers. Overall, integrating genetic, biochemical, and morphological analyses offers an in-depth view of the diversity and relationships among Shan tea germplasms in Vietnam. This study contributes to our understanding of the genetic diversity within Shan tea and provides valuable information for future breeding and selection programs such as developing improved Shan varieties with desired biochemical content and greater adaptability to regional conditions. Besides, the biochemical profile of Shan tea provides chances for optimizing harvest strategies, while the unique genetic characteristics of each Shan tea region could support local tea branding and tea resource conservation purposes. Future research should be expanded with larger genetic resources of Shan tea including ancient Shan tea clones and improved Shan tea cultivars in other tea-growing regions in the country.

ACKNOWLEDGMENTS

The authors thank all research staffs at Vietnam National University and Tea Research and Development Center, Northern Mountainous Agriculture and Forestry, for their continuous support. Khuynh The Bui performed the statistical analysis, prepared the manuscript, and reviewed it. Thai Hoang Dinh, Ngoc-Thang Vu, Van Thu Dang, Thi Viet Ha Do contributed to data collection and curation. Xinghui Li and Xuan Hoang Tran conceptualized the study, reviewed it and edited the final manuscript. All authors read and approved the final manuscript.

REFERENCES

Ahmed, S., Griffin, T.S., Kraner, D., Schaffner, M.K., Sharma, D., Hazel, M., Leitch, A.R., Orians, C.M., Han, W., Stepp, J.R., Robbat, A., Matyas, C., Long, C., Xue, D., Houser, R.F. & Cash, S.B. 2019. Environmental factors variably impact tea secondary metabolites in the context of climate change. *Front in Plant Sci.* 10: 939.

Bali, S., Raina, S.N., Bhat, V., Aggarwal, R.K. & Goel, S. 2013. Development of a set of genomic microsatellite markers in tea (*Camellia* L.) (Camelliaceae). *Molecular Breeding* 32: 735-741.

Chen, L. & Yamaguchi, S. 2002. Genetic diversity and phylogeny of the tea plant (*Camellia sinensis*) and its related species and varieties in the section *Thea* genus *Camellia* determined by randomly amplified polymorphic DNA analysis. *Journal of Horticultural Science and Biotechnology* 77(6): 729-732. https://doi.org/10.1080/14620316.2002.11511564

Chinese National Standard. 2008. GB/T 8313-2008. Determination of total polyphenols and catechins content in tea.

Clarke, C., Richter, B.S. & Rathinasabapathi, B. 2023. Genetic and morphological characterization of United States tea (*Camellia sinensis*): Insights into crop history, breeding strategies, and regional adaptability. *Front. in Plant Sci.* 14: 1149682. https://doi.org/10.3389/fpls.2023.1149682

Deka, H., Barman, T., Dutta, J., Devi, A., Tamuly, P., Paul, R.K. & Karak, T. 2021. Catechin and caffeine content of tea (*Camellia sinensis* L.) leaf significantly differ with seasonal variation: A study on popular cultivars in North East India. *Journal of Food Composition and Analysis* 96: 103684. https://doi.org/10.1016/j.jfca.2020.103684

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

Fang, W., Cheng, H., Duan, Y., Jiang, X. & Li, X. 2011. Genetic diversity and relationship of clonal tea (*Camellia sinensis*) cultivars in China as revealed by SSR markers. *Plant Syst. Evol.* 298: 469-483. https://doi.org/10.1007/s00606-011-0559-3

Freeman, S., West, J., James, C., Lea, V. & Mayes, S. 2004. Isolation and characterization of highly polymorphic microsatellites in tea (*Camellia sinensis*). *Molecular Ecology Notes* 4: 324-326. https://doi.org/10.1111/j.1471-8286.2004.00682.x

Gong, A.D., Lian, S.B., Wu, N.N., Zhou, Y.J., Zhao, S.Q., Zhang, L.M., Cheng, L. & Yuan, H.Y. 2020. Integrated transcriptomics and metabolomics analysis of catechins, caffeine and theanine biosynthesis in tea plant (*Camellia sinensis*) over the course of seasons. *BMC Plant Biology* 20: 294. https://doi.org/10.1186/s12870-020-02443-y

Guo, R., Xia, X., Chen, J., An, Y., Mi, X., Li, R., Zhang, C., Chen, M., Wei, C. & Liu, S. 2021. Genetic relationship analysis and molecular fingerprint identification of the tea germplasms from Guangxi Province, China. *Breeding Science* 71(5): 584-593.

Huang, D.J., Wang, Y.P., Chen, X., Wu, J., Wang, H.J., Tan, R.R., Jiao, L. & Mao, Y.X. 2022. Application of tea-specific fertilizer combined with organic fertilizer improves aroma of green tea. *Horticulturae* 8(10): 950.

Jhou, H.C., Chen, C.I., Shen, J.Y., Wang, C.T. & Chang, F.C. 2025. Influence of natural light intensities on growth and content of catechins and caffeine in *Camellia formosensis* leaves and tea infusion. *BioResources* 20(1): 877-887.

Kottawa-Arachchi, J.D., Gunasekare, M.K. & Ranatunga, M.A. 2019. Biochemical diversity of global tea [*Camellia sinensis* (L.) O. Kuntze] germplasm and its exploitation: A review. *Genetic Resources and Crop Evolution* 66: 259-273. https://doi.org/10.1007/s10722-018-0698-2

Labate, J.A., Lamkey, K.R., Mitchell, S.E., Kresovich, S., Sullivan, H. & Smith, J.S.C. 2023. Molecular and historical aspects of Corn Belt dent diversity. *Crop Science* 43(1): 80-91. https://doi.org/10.2135/cropsci2003.8000

Le, T.L., Jang, D.C., Pham, L.A.M., Le, H.D., Nguyen, T.H., Nguyen, V.K., Nguyen, T-H., Hoang, T-N., Vu, D-T., Bui, T-K., Vu, T-T-H., Dinh, T-H., Tran, A-T., Phan, T-T-H., Tong, V-G., Le, T-T-C. & Vu, N-T. 2023. Genetic diversity and population structure of Canna edulis accessions in Vietnam revealed by ISSR markers. *Vegetos* 37: 1009-1018. https://doi.org/10.1007/s42535-023-00633-1

Liu, K.J. & Muse, S.V. 2005. PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21: 2121-2129. https://doi.org/10.1093/bioinformatics/bti282

Liu, S., Liu, H., Wu, A., Hou, Y., An, Y. & Wei, C. 2017. Construction of fingerprinting for tea plant (*Camellia sinensis*) accessions using new genomic SSR markers. *Mol. Breeding* 37(8): 93. https://doi.org/10.1007/s11032-017-0692-y

Ma, J.Q., Zhou, Y.H., Ma, C.L., Yao, M.Z., Jin, J.Q., Wang, X.C. & Chen, L. 2010. Identification and characterization of 74 novel polymorphic EST-SSR markers in the tea plant, *Camellia sinensis* (Theaceae). *American Journal of Botany* 97(12): e153-e156. https://doi.org/10.3732/ajb.1000376

Mu, S.Q., Gu, X.F., Zhang, S.P., Wang, X.W. & Wang, Y. 2003. Genetic diversity of Cucumber (*Cucumis sativus* L.) germplasm by SSR. *Acta Horticulture Sinica* 35(9): 1323-1330.

Nadeem, M.A., Nawaz, M.A., Shahid, M.Q., Doğan, Y., Comertpay, G., Yildiz, M., Hatipoğlu, R., Ahmad, F., Alsaleh, A., Labhane, N., Özkan, H., Chung, G. & Baloch, F.S. 2018. DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnology & Biotechnological Equipment 32*(2): 261-285.

Ni, J., Colowit, PM. & Mackill, D.J. 2002. Evaluation of genetic diversity in rice subspecies using microsatellite markers. *Crop Science* 42(2): 601-607. https://doi.org/10.2135/cropsci2002.6010

Nguyen, K.O.T., Nguyen, P.L., Le, H.L. & Le, H.T. 2022. Discriminative chemical profiles of Shan Tuyet Tea (*Camellia sinensis* var. Shan) and Sinensis Tea (*Camellia sinensis* var. sinensis) collected in Ta Xua, Son La, Vietnam and their correlation with antioxidant activity. *Natural Product Communications* 17: 9. https://doi.org/10.1177/1934578X221128410

Paiva, L., Lima, E., Motta, M., Marcone, M. & Baptista, J. 2021. Influence of seasonal and yearly variation on phenolic profiles, caffeine, and antioxidant activities of green tea (*Camellia sinensis* (L.) Kuntze) from Azores. *Applied Sciences* 11(16): 7439.

Pandolfi, C., Mugnai, S., Azzarello, E., Bergamasco, S., Masi, E. & Mancuso, S. 2009. Artificial neural networks as a tool for plant identification: A case study on Vietnamese tea accessions. *Euphytica* 166: 411-421. https://doi.org/10.1007/s10681-008-9828-9

Parmar, R., Seth, R. & Sharma, R.K. 2022. Genome-wide identification and characterization of functionally relevant microsatellite markers from transcription factor genes of Tea (*Camellia sinensis* (L.) O. Kuntze). *Scientific Reports* 12(1): 201.

Phong, N.H., Pongnak, W., Soytong, K., Poeaim, S. & Poeaim, A. 2016. Diversity of Tea (*Camellia sinensis*) grown in Vietnam based on morphological characteristics and Inter-primer Binding Sites (iPBS) Marker. *International Journal of Agriculture & Biology* 18(2): 385-392.

Powell, W., Machray, G.C. & Provan, J. 1996. Polymorphism revealed by simple sequence repeats. *Trends in Plant Science* 1(7): 215-222. https://doi.org/10.1016/1360-1385(96)86898-1

Rajkumar, S., Karthigeyan, S., Sud, R.K., Rajkumar, R., Muraleedaran, N., Das, S.C., Hazarika, M. & Ahuja, P.S. 2010. Genetic diversity of Indian tea (*Camellia sinensis* (L.) Kuntze) germplasm detected using morphological characteristics. *Journal Cell and Plant Science* 1(1): 13-22.

Rohlf, F.J. 2000. Statistical power comparisons among alternative morphometric methods. *Am. J. Phys. Anthropol.* 111: 463-478.

Saravanan, M., Maria, J.K.M., Raj Kumar, R., Pius, P.K. & Sasikumar, R. 2005. Genetic diversity of UPASI tea cones (*Camellia sinensis* (L.) O. Kuntze) on the basis of total catechins and their fractions. *Phytochemistry* 66: 561-565. https://doi.org/10.1016/j.phytochem.2004.06.024

Sealy, J.B. 1958. *A Revision of the Genus Camellia*. London: Royal Horticultural Society. p. 239.

Sharma, R.K., Bhardwaj, P., Negi, R., Mohapatra, T. & Ahuja, P.S. 2009. Identification, characterization and utilization of unigene derived microsatellite markers in tea (*Camellia sinensis* L.). *BMC Plant Biology* 9: 1-24.

Sharma, V., Joshi, R. & Gulati, A. 2011. Seasonal clonal variations and effects of stresses on quality chemicals and prephenate dehydratase enzyme activity in tea (*Camellia sinensis*). *European Food Research and Technology* 232: 307-317.

Takeda, Y. 2000. History and development in Japanese tea breeding. In: Tea Culture, Tea Food Industry and Tea Breeding in Korea, China and Japan. *The Korea Tea Society*, Korea Park Y.G. and Shin D.I. (eds). pp. 139-158.

Tan, L.Q., Wang, L.Y., Wei, K., Zhang, C.C., Wu, L.Y., Qi, G.N. & Liang, J.B. 2013. Floral transcriptome sequencing for SSR marker development and linkage map construction in the tea plant (*Camellia sinensis*). *PLoS ONE* 8(11): e81611.

Taniguchi, F., Furukawa, K., Ota-Metoku, S., Yamaguchi, N., Ujihara, T., Kono, I. & Tanaka, J. 2012. Construction of a high-density reference linkage map of tea (*Camellia sinensis*). *Breeding Science* 62(3): 263-273.

Tao, L.L., Ting, Y.J., Chen, H.R., Wen, H.L., Hui, X., Luo, L.Y., Huang, K.L., Zhu, J.Y., Liu, S.R. & Wei, C.L. 2023. Core collection construction of tea plant germplasm in Anhui Province based on genetic diversity analysis using simple sequence repeat markers. *Journal of Integrative Agriculture* 22(9): 2719-2728. https://doi.org/10.1016/j.jia.2023.07.020

Tian, X., Chen, S., Zhong, Q., Wang, J., Chen, J., Chen, L., Moon, D. & Ma, J. 2024. Widely targeted metabolomics analysis reveals the effect of cultivation altitude on tea metabolites. *Agronomy* 14(4): 812. https://doi.org/10.3390/agronomy14040812

Tounekti, T., Joubert, E., Hernández, I. & Munné-Bosch, S. 2013. Improving the polyphenol content of tea. *Critical Reviews in Plant Sciences* 32(3): 192-215.

Vo, T.D. 2007. Assessing genetic diversity in Vietnam tea [*Camellia sinensis* (L.) O. Kuntze] using morphology, inter-simple sequence repeat (ISSR) and microsatellite (SSR) markers. PhD Dissertation. Georg-August University Goettingen (Unpublished).

Wambulwa, M.C., Meegahakumbura, M.K., Kamunya, S. & Wachira, F.N. 2021. From the wild to the cup: Tracking footprints of the tea species in time and space. *Frontiers in Nutrition* 8: 706770. https://doi.org/10.3389/fnut.2021.706770

Wang, R.J., Gao, X.F., Kong, X.R. & Yang, J. 2016. An efficient identification strategy of clonal tea cultivars using long-core motif SSR markers. *SpringerPlus* 5(1): 1152.

Wight, W. 1962. Tea classification revised. *Curr. Sci.* 31: 298-299.

Yang, Y., Liu, Z., Zhao, Y., Liang, G.Q. & Zhao, X. 2009. Genetic diversity and relationship of Huangjincha cultivar based on EST-SSR markers. *J. Tea Sci.* 29: 236-242 (in Chinese with English abstract).

Yao, L., Caffin, N., D'Arcy, B., Jiang, Y., Shi, J., Singanusong, R., Liu, X., Datta, N., Kakuda, Y. & Xu, Y. 2005. Seasonal variations of phenolic compounds in Australia-grown tea (*Camellia sinensis*). *J. Agric. Food Chem*. 53(16): 6477-6483.

Yao, M.Z., Chen, L., Wang, X.C., Zhao, L.P. & Yang, Y.J. 2007. Genetic diversity and relationship of clonal tea cultivars in China revealed by ISSR markers. *Acta Agron. Sin.* 33: 598-604 (in Chinese with English abstract).

Ye, F., Guo, X., Li, B., Chen, H. & Qiao, X. 2022. Characterization of effects of different tea harvesting seasons on quality components, color and sensory quality of "Yinghong 9" and "Huangyu" large-leaf-variety black tea. *Molecules* 27(24): 8720. https://doi.org/10.3390/molecules27248720

Zagoskina, N.V., Alyavina, A.K., Gladyshko, T.O., Lapshin, P.V., Egorova, E.A. & Bukhov, N.G. 2005. Ultraviolet rays promote development of photosystem II photochemical activity and accumulation of phenolic compounds in the tea callus culture (*Camellia sinensis*). *Russian Journal of Plant Physiology* 52: 731-739.

Zhang, Q., Zhang, Y., Xie, J., Ye, J., Pang, X. & Jia, X. 2022. Differences in the analysis of the quality indexes and characteristic amino acids of the different grades of Wuyi Shuixian (*Camellia sinensis*) tea. *Food Science and Technology* 42: e66122.

Zhao, Z., Song, Q., Bai, D., Niu, S., He, Y., Qiao, D., Chen, Z., Li, C., Luo, J. & Li, F. 2022. Population structure analysis to explore genetic diversity and geographical distribution characteristics of cultivated-type tea plant in Guizhou Plateau. *BMC Plant Biology* 22(1): 55.

Zheng, X.Q., Jin, J., Chen, H., Du, Y.Y., Ye, J.H., Lu, J.L., Lin, C., Dong, J.J., Sun, Q.L., Wu, L.Y. & Liang, Y.R. 2008. Effect of ultraviolet B irradiation on accumulation of catechins in tea (*Camellia sinensis* (L) O. Kuntze. *African Journal of Biotechnology* 7(18): 3283-3287.

Zhou, Q., Li, H., Hoang, T.X., Ruan, X., Zhang, Y., Arkorful, E., Chen, X., Sun, K. & Li, X. 2019. Genetic diversity and relationship of Dongting Biluochun Tea gemplasm in Suzhou revealed by SSR markers. *Pak. J. Bot.* 51(3): 895-902. DOI: 10.30848/PJB2019-3(31)

*Corresponding author; email: hoangxuantrannomafsi@gmail.com

SUPPLEMENTARY TABLE S1

Genotype	Group	Leaf blade length	Leaf blade width	Leaf blade area	Leaf blade shape (L:W ratio)	Fine pluck length	Fine pluck weight	Pubescence density	Fruit length	Fruit width	Seed length	Seed width	Pericap thickness	Flower diameter	Stamen length	Stamen number	Pistil length	Ovary pubescence density	EGC	GC	EC	EGCG	ECG	Cattechin total	Tannin
T1	TC	15.72	5.92	64.23	2.655405	10.78	2.38	3	2.88	2.53	1.67	1.5	1.61	3.12	0.92	216	1.3	3	24.14	17.34	20.98	65.98	20.93	149.37	28.63
T2	TC	15.25	5.66	57.9	2.694346	8.76	2.42	2	3.05	2.32	1.53	1.37	1.78	3.41	1.1	188.7	1.29	1	22.64	14.62	19.24	59.3	18.61	134.41	28.73
T3	TC	18.67	5.89	75.97	3.169779	8.4	2.2	3	2.47	1.67	1.16	1.09	1.6	3.28	0.95	231.4	1.27	2	25.29	18.29	19.43	69.11	20.8	152.92	29.03
T4	TC	12.97	4.88	42.87	2.657787	9.53	2.34	3	2.44	1.45	1.24	1.11	1.67	3.39	0.99	196.3	1.33	1	25.31	16.39	19.7	68.9	20.93	151.23	28.53
T5	TC	12.73	4.52	37.86	2.816372	6.76	2.15	2	2.36	1.23	1.17	1.08	1.45	3.16	1.14	312.4	1.36	2	24.93	15.81	19.53	62.01	19.02	141.3	28.85
T6	TC	14.09	5.88	56.36	2.396259	8.62	2.3	2	3.01	2.19	1.42	1.32	1.71	3.19	0.93	195.9	1.3	1	26.01	14.19	20.03	62.52	18.27	141.02	30.49
T7	TC	14.72	5.43	52.88	2.710866	9.38	2.09	3	2.54	1.47	1.22	1.22	1.17	3.51	0.95	212.4	1.32	1	25.64	17.9	20.83	66.83	20.02	151.22	29.2
T8	TC	15.63	5.82	61.4	2.685567	7.7	2.49	2	3.11	2.35	1.56	1.35	1.6	3.18	1.1	210.5	1.35	3	23.84	14.5	19.54	59.28	18.97	136.13	29.47
T9	TC	15.18	5.87	59.6	2.586031	8.66	2.17	3	2.87	2.53	1.68	1.54	1.62	3.16	0.86	187.7	1.3	3	24.16	17.63	19.91	66.98	19.77	148.45	30.03
T10	TC	18.16	6.82	84.19	2.662757	9.92	2.12	3	2.48	1.67	1.19	1.12	1.54	3.17	1.12	280.9	1.52	1	25.74	19.41	20.08	66.83	19.34	151.4	29.43
T11	TC	13.14	4.81	65.46	2.731809	7.13	2.14	3	2.47	1.42	1.34	1.14	1.6	3.29	1.12	207.8	1.45	2	26.22	16.26	19.52	62.78	19.32	144.1	29.1
T12	TC	13.07	4.82	70.69	2.711618	7.48	2.11	2	3.03	2.33	1.51	1.34	1.73	3.2	0.98	206.2	1.32	3	23.58	14.81	17.65	59.18	18.55	133.77	29.65
T13	TC	13.5	5.77	62.62	2.339688	8.12	2.19	3	3.1	2.37	1.62	1.47	1.72	3.49	0.94	215.8	1.3	2	23.15	14.65	19.32	57.86	18.12	133.1	28.64
T14	TC	15.35	5.79	58.88	2.651123	8.36	2.35	2	2.38	1.25	1.18	1.08	1.52	3.38	1.12	218	1.27	1	23.35	12.23	18.34	57.96	19.33	131.21	29.16
T15	TC	15.81	5.91	57.69	2.675127	10.09	2.23	3	2.78	2.5	1.61	1.45	1.56	3.48	0.93	322.5	1.3	3	23.96	15.64	18.71	58.46	19.04	135.81	29.13
T16	TC	17.88	6.67	78.74	2.68066	8.48	2.36	2	2.75	2.45	1.46	1.25	1.46	3.23	0.96	294.5	1.3	2	25.45	14.9	19.05	62.71	19.88	141.99	29.95
T17	TC	13.76	5.65	56.46	2.435398	9.13	2.17	3	3.01	2.19	1.42	1.32	1.72	3.22	0.97	216.2	1.35	1	23.66	14.44	18.3	58.83	18.24	133.47	29.98
T18	TC	13.12	4.95	42.89	2.650505	9.47	2.48	3	2.48	1.68	1.2	1.09	1.63	3.19	0.99	218.5	1.33	3	23	15.57	17.84	57.17	19.01	132.59	29.42
T19	TC	15.55	5.74	55.79	2.709059	7.93	2.19	3	3.07	2.36	1.62	1.32	1.7	3.51	1.13	243.3	1.28	3	26.61	16.31	19.34	59.5	19.16	140.92	30.57
T20	TC	15.57	5.82	57.65	2.675258	7.59	2.13	2	2.49	1.62	1.27	1.09	1.66	3	0.96	176.2	1.48	1	24.13	14.6	18.69	59.82	18.72	135.96	29.57
T21	SG	15.43	5.84	60.13	2.642123	7.61	2.38	1	2.65	2.43	1.68	1.52	2.56	2.82	0.9	215.4	1.33	2	26.18	15.85	18.69	64.01	18.5	143.23	29.38
T22	SG	16.17	6.63	73.14	2.438914	8.25	2.36	3	2.25	2.23	1.48	1.33	2.61	3.01	1.12	187.5	1.28	3	25.79	16.68	19.17	65.48	18.46	145.58	29.19
T23	SG	14.51	6.82	65.93	2.127566	7.95	2.23	1	2.67	2.56	1.66	1.5	2.6	2.98	0.94	201.7	1.26	1	25.1	15.95	18.93	64.18	18.3	142.46	30.78
T24	SG	17.17	7.57	89.31	2.268164	9.16	2.37	1	3.12	2.67	1.69	1.54	2.63	2.79	0.99	195.3	1.23	1	24.23	15.05	18.92	62.83	18.46	139.49	29.42
T25	SG	15.79	7.57	81.4	2.085865	8.16	2.11	3	3.15	2.35	1.65	1.48	2.58	3.11	1.13	310.4	1.4	3	25.32	16.32	19.06	63.98	17.53	142.21	29.58
T26	SG	15.97	7.1	76.08	2.249296	7.28	2.17	2	2.78	2.56	1.58	1.52	2.67	2.69	0.92	192.8	1.28	3	23.76	14.52	19.12	64.38	19	140.78	28.98
T27	SG	16.61	5.41	59.64	3.07024	8.52	2.37	2	2.98	2.76	1.56	1.43	2.56	2.9	0.91	210.6	1.3	1	25.12	16.27	20.15	65.52	19.34	146.4	30.93
T28	SG	16.15	6.14	67.35	2.630293	7.15	2.05	3	3.11	2.35	1.69	1.53	2.45	2.88	1.13	201.2	1.3	2	26.3	16.91	19.31	65.02	18.26	145.8	29.98
T29	SG	13.79	6.31	59.08	2.18542	8.18	2.35	2	2.97	2.56	1.73	1.55	2.57	2.76	0.86	185.5	1.34	3	24.3	14.57	19.34	62.4	17.63	138.24	29
T30	SG	16.37	6.37	71.89	2.569859	7.33	2.18	2	3.13	2.68	1.8	1.57	2.55	3.12	1.02	279.9	1.5	2	23.69	15.05	18.41	58.27	18.56	133.98	28.97
T31	SG	13.83	5.58	50.76	2.478495	7.43	2.13	2	3.15	2.77	1.87	1.58	2.61	2.91	1.02	197.8	1.42	2	22.91	13.21	19.33	61.23	17.57	134.25	28.18

continue to next page

continue from previous page

T32	SG	18.77	7.75	97.97	2.421935	7.63	2.32	1	2.89	2.58	1.69	1.53	2.6	3.2	0.95	205.9	1.3	2	23.45	15.48	18.68	58.93	19.71	136.25	28.8
T33	SG	16.58	6.31	72.95	2.627575	8.25	2.27	2	2.97	2.46	1.67	1.47	2.54	2.69	0.91	212.8	1.32	3	23.63	15.65	17.83	60.4	19.57	137.08	30.36
T34	SG	18.24	7.1	88.34	2.569014	7.94	2.18	3	2.87	2.56	1.64	1.45	2.66	3.18	1.11	216	1.26	1	23.49	15.53	20.23	58.59	17.41	135.25	28.42
T35	SG	12.94	5.34	49.43	2.423221	8.08	2.03	3	2.92	2.67	1.57	1.42	2.65	2.78	0.9	302.2	1.35	3	23.66	16.46	18.63	62.33	19.08	140.16	29.13
T36	SG	15.14	5.84	61.45	2.592466	8.56	2.2	1	2.45	2.38	1.56	1.41	2.67	2.83	0.93	290.4	1.38	3	22.27	15.37	17.73	57.04	17.47	129.88	28.06
T37	SG	13.67	5.89	54.74	2.320883	8.24	2.16	1	2.65	2.28	1.65	1.56	2.6	3.22	0.97	206	1.3	1	24.63	16.57	17.57	60.15	20.09	139.01	29.48
T38	SG	15.5	5.51	54.15	2.813067	8.13	2.06	2	2.87	2.43	1.69	1.48	2.65	2.79	0.99	203.2	1.3	2	23.32	14	17.82	57.44	20.78	133.36	29.38
T39	SG	17.77	6.77	83	2.624815	6.99	1.95	1	2.88	2.75	1.7	1.52	2.59	2.91	1.02	240.6	1.44	3	23.69	16.48	16.89	58.17	18.96	134.19	28.67
T40	SG	16.71	6.06	66.73	2.757426	7.79	2.17	1	3.14	2.57	1.84	1.61	2.65	3	0.95	175.4	1.5	2	21.78	14.35	17.17	60.2	19.57	133.07	28.63
T41	CB	13.39	4.81	43.48	2.783784	6.96	1.95	3	2.98	2.66	1.51	1.28	1.74	3.42	0.96	216.3	1.32	2	27.84	18.65	21.63	66.34	18.78	153.24	29.64
T42	CB	14.82	5.03	51.3	2.946322	8.15	2.25	3	2.89	2.75	1.38	1.38	1.78	3.51	1.07	189.7	1.3	2	23.87	14.87	20.71	59.92	17.68	137.05	29.29
T43	CB	15.61	5.53	58.73	2.822785	7.92	1.81	2	2.92	2.56	1.66	1.34	1.89	3.98	0.95	201.8	1.25	1	26.11	17.96	20.6	64.74	19.63	149.04	29.83
T44	CB	16.24	5.61	61.72	2.894831	9.1	2.28	1	3.13	2.67	1.59	1.14	1.77	3.79	0.99	192.3	1.2	1	25.33	15.89	20.79	61.06	20.45	143.52	29.92
T45	CB	13.89	5.17	49.26	2.686654	8.16	2.1	1	3.02	1.78	1.45	1.38	1.75	4.01	1.12	312.4	1.42	3	25.49	14.98	18.85	62.34	19.02	140.68	30.35
T46	CB	15.85	5.03	54.48	3.151093	7.29	2.18	2	3.16	2.56	1.58	1.12	1.71	3.69	0.96	194.6	1.2	3	23.22	13.59	18.94	59.92	19	134.67	32.57
T47	CB	16.21	5.52	61.04	2.936594	8.56	2.39	3	3.07	2.61	1.46	1.19	1.87	3.9	0.91	211.4	1.25	1	23.73	14.85	18.86	58.76	19.38	135.58	30.42
T48	CB	13.29	4.33	37.84	3.069284	7.15	2.05	3	3.12	2.11	1.79	1.43	1.72	3.88	1.03	301.2	1.34	1	23.9	14.13	18.79	58.28	18.79	133.89	30.19
T49	CB	17.24	5.89	69.49	2.926995	8.14	2.14	2	2.88	2.53	1.48	1.55	1.72	3.76	0.91	198.5	1.37	3	21.98	14.49	20.12	59.84	17.48	133.91	30.07
T50	CB	12.23	3.96	32.28	3.088384	7.36	1.89	2	3.19	2.58	1.48	1.45	1.76	4.12	1.02	269.9	1.55	2	26.71	18.69	20.18	64.69	20.46	150.73	32.48
T51	CB	12.77	4.18	35.47	3.055024	7.45	2.13	3	2.95	1.77	1.47	1.47	1.86	3.91	1	179.8	1.43	2	26.12	17.65	20.65	68.1	20.56	153.08	29.6
T52	CB	13.57	5.13	47.24	2.645224	7.66	2.42	1	3.09	2.53	1.39	1.37	1.73	4.2	0.91	207.9	1.26	1	27.5	16.92	20.85	66.12	21.56	152.95	30.35
T53	CB	15.48	5.7	59.63	2.715789	8.24	2.24	3	2.98	2.44	1.37	1.27	1.72	3.69	0.91	200.8	1.32	1	23.22	15.38	17.86	62.47	19.46	138.39	32.39
T54	CB	16.09	5.58	61.69	2.883513	7.95	2.16	3	3.17	2.15	1.34	1.35	1.82	4.18	1	211.6	1.3	2	25.7	16.81	20.16	63.16	20.42	146.25	30.2
T55	CB	14.59	5.47	54.36	2.667276	8.07	2.08	3	3.07	2.13	1.55	1.39	1.76	3.78	0.89	304.2	1.4	3	24.24	15.03	18.35	62.06	21.14	140.82	30.05
T56	CB	15.14	5.81	60.02	2.605852	8.57	2.16	3	3.26	2.75	1.86	1.37	1.86	3.83	0.97	291.4	1.35	3	27.95	17.25	19.61	68.81	21.76	155.38	30.85
T57	CB	15.95	6.05	66.69	2.636364	8.26	2.11	1	2.93	1.92	1.65	1.38	1.72	4.22	0.98	207.1	1.28	2	23.14	15.46	18.23	60.71	19.4	136.94	30.58
T58	CB	14.41	5.03	50.01	2.864811	8.13	2.06	2	3.15	1.75	1.28	1.58	1.63	3.79	0.98	200.3	1.34	2	23.89	13.59	18.88	60.9	19.47	136.73	29.71
T59	CB	13.89	4.17	39.56	3.330935	7.03	1.95	1	3.17	2.01	1.67	1.47	1.76	3.91	1	243.5	1.41	2	27.84	18.4	20.23	66.58	21.39	154.44	30.4
T60	CB	13.35	4.86	44.38	2.746914	7.96	2.18	1	2.87	2.14	1.64	1.56	1.94	4	0.98	212.3	1.45	1	25.26	14.66	18.83	61.74	18.48	138.97	30.11

SUPPLEMENTARY TABLE S2. The informa tion of SSR markers used in the study

No.	Marker	Forward primer (5'-3')	Reverse primer (3'-5')	Reference
1	CsFM1051	AACCCATTTCGTCTTTGTGC	AGAATCAACAACACCCTGGC	Tan et al. (2013)
2	CsFM1550	CGAGACATCGAACACCACAG	CGTATCGTAGCGGTGAAGGT	Tan et al. (2013)
3	CsFM1599	GGCCCTGTTTTTACACTCCA	GATTGGTTTCTGGTTCGCAT	Tan et al. (2013)
4	CsFM1384	CGAATCATGATGACCCACTG	CAGCGAGGGAGAAGAATGAG	Tan et al. (2013)
5	CsFM1509	GACGATGGACCCTTCTTTGA	CATCATCATCATCCTCACCG	Tan et al. (2013)
6	A17	ACTAAGGCGGTCACGAAGTT	AAGGGATACAGCAGATCCAAAT	Tan et al. (2013)
7	A28	AATAAGAATCGGTGACCTCTG	CTTCATTAACCCCTAAACTAAAAC	Tan et al. (2013)
8	A38	CCAAAACCCTAGTTTCACTCCA	ATCAAACGCTCTGTATCGGTG	Tan et al. (2013)
9	A55	GCTTCCTCTTCTCCTTCCCC	CCCCTCCTCCTCTGTTTGAT	Tan et al. (2013)
10	A166	TTGGCAGATTACCTTGGAGA	GACCAACAACGGATCACATA	Tan et al. (2013)
11	MSG0533	AGACCTAGCCAAGACAACCACACC	GTTTCCCCTATTTTCCCGACTGTCT	Taniguchi et al. (2012)
12	MSG0380	ACAGACCTTCACCCTCTCCATTTC	GTTTACCTCTGCCTTCGTTCTTCAGC	Taniguchi et al. (2012)
13	MSE0313	TGCTATGCCGCCTAACAAAAACTT	ACCACCAACAACAATTCCCACTCT	Taniguchi et al. (2012)
14	MSG0423	ACTCCATGTGCTGCTCTGTAGTTC	GTTTGCAGGAAGTTGAGCCAGAC	Taniguchi et al. (2012)
15	MSE0291	AATCAAATAACACTTGCACCCGC	AAAAAGAGAAAGTCACGTCCACGG	Taniguchi et al. (2012)
16	MSG0681	AGGGTTTGCGTCTTCAAAGAGAGA	GTTTGTAACACTTGCCACGTTTCG	Taniguchi et al. (2012)
17	Csin04	ATTTTGAAGTCCTCTCAGAACCAT	CATCGTGAACCGCATCTGTAG	Bali et al. (2013)
18	Csin06	CGGGCACTCAATGGAAAGCAC	TGGCATCTGTTGGCGTGGTG	Bali et al. (2013)
19	Csin07	CCAACCCAACTCAGGCAGAT	GCTACAACCACCTTCAACACCT	Bali et al. (2013)
20	Csin24	CCAAGTAGAAGGACGCACTC	GGAGCATAGCATAGCATAGC	Bali et al. (2013)
21	Csin41	CCCTCAACTCCATCAGCAAT	CCCAAAACGAAAACCGACTA	Bali et al. (2013)
22	Csin46	CAGGGAGAGGACGGTGATTA	GCACGAAAAGTCAGGCTACA	Bali et al. (2013)
23	Csin49	CTCCAGCAGCAACATTATTACG	GACCTCAGAAAACTCCCCTTG	Bali et al. (2013)
24	Csin68	GTGGCATGGAAATGGGATAC	AGATGCTATCATAACAAAGAAACAAT	Bali et al. (2013)
25	Csin71	GTTGCTGCTGTTGTCAGTTGC	CCAACCACAATCAGCCACTAC	Bali et al. (2013)
26	TUGMS102A	CGTAGCTCGCACACAACAC	CGTCCCCTCCGAAATGA	Sharma et al. (2009)
27	TUGMS27	GGGGATAGTACAAACACACAAC	GCTCCTCTTTCTTCACCACTT	Sharma et al. (2009)
28	TUGMS82	AAGTTAGAGAGAGAGAAGTGGC	AATGCCACACCAGTCCTAG	Sharma et al. (2009)
29	TUGMS73	GTCAAGACGCCCACTACAGT	GACTGTGTAACCTGCCAAGAC	Sharma et al. (2009)
30	CamsinM14	TGGACTTGGAAGGACTGAGG	ACAAAGCTCAACCTGCCATT	Freeman et al. (2004)