Plant Genetic Resources: Characterization and Utilization

cambridge.org/pgr

Genetic resources diversity of tea (*Camellia sinensis* (L.) Kuntze) in the southern region of the Caspian Sea

Samira Montahae Dargah[1], Mohammad Bagher Rezaei[2], Marzieh Ghanbari Jahromi[1], Sepideh Kalateh Jari [1] and Shahin Jahangirzadeh Khiavi[3]

[1]Department of Horticultural Science and Agronomy, Science and Research Branch, Islamic Azad University, Tehran, Iran; [2]Department of Medicinal Plants, Research Institute of Forests and Rangelands, Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran and [3]Tea Research Center, Horticultural Sciences Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Lahijan, Iran

Research Article

Cite this article: Montahae Dargah S, Rezaei MB, Ghanbari Jahromi M, Kalateh Jari S, Jahangirzadeh Khiavi S (2023). Genetic resources diversity of tea (*Camellia sinensis* (L.) Kuntze) in the southern region of the Caspian Sea. *Plant Genetic Resources: Characterization and Utilization* **21**, 97–106. https://doi.org/10.1017/S147926212300045X

Received: 15 November 2022
Revised: 21 June 2023
Accepted: 22 June 2023
First published online: 28 July 2023

Keywords:
Camelia sinensis; cluster analysis; heatmap; molecular diversity; PCA; PIC; primer

Corresponding author:
Marzieh Ghanbari Jahromi;
Email: ghanbari@srbiau.ac.ir

Abstract

In the fields of agricultural, food and medical research, the potential impacts of tea on human health are of special interest because tea (*Camellia sinensis* (L.) Kuntze) is continuously consumed by many people in the world. The evaluation of the diversity of genotypes found in tea germplasm can aid in the improvement of the breeding programme. In this study, the genetic diversity of 30 tea genotypes from two commercial sites of tea production in Iran was investigated by using morphological and 12 ISSR markers. Morphological analysis showed that the diversity between samples of tea was limited, and the narrow matching range was calculated. In cluster analysis at level 0.63, samples were divided into four groups. The application of 12 ISSR primers produced 91 polymorphic bands. PIC test showed a range of 0.41–0.48. Based on the ISSR data, the matching range was obtained in the range of 0.24–0.93. In cluster analysis, samples at level 0.58 were divided into five groups. According to the results, it can be understood that these series of traits and primers can very well recognize genetic differences. Using these markers, genetic diversity was observed among tea genotypes, but this diversity was not such as to be able to separate genotypes of various regions from each other. The results showed that the tea genotypes in Iran had high genetic diversity. As a consequence, the findings of our study will help the development of tea germplasm conservation strategies and their sustainable use in breeding programmes.

Introduction

Tea (*Camellia sinensis* (L.) Kuntze) from the family Theaceae is an evergreen, perennial and cross-pollinate plant (Mondal *et al.*, 2003; Shen *et al.*, 2019), and its young leaves are becoming one of the oldest caffeinated soft drinks in the world. This plant originates from the Yunnan area of China (Chen *et al.*, 2005) and it has been widely cultivated in different parts of the world like Iran. Most tea plants are diploid (2n = 2x = 30). The genetic basis of tea under cultivation in Iran is including three seed varieties with the names of Betjan, Dhonjan and Rajghur (Ahmadishad *et al.*, 2009). Classification and hierarchy of tea plant phytology in scientific societies are debatable due to the self-incompatible, prolonged natural and artificial hybridization of the plant that leads to the creation of a large number of hybrids with various morphological states (Sealy, 1958; Li *et al.*, 2023; Sharma *et al.*, 2023). Based on leaf and growth characteristics, the tea plant is divided into three groups: Chinese tea (small leaf) with the scientific name *C. sinensis* (L.) O. Kuntze, Assam tea (big leaf) with the scientific name of *C. assamica* (Masters) and Wight and Cambodia tea (broad leaf) with the scientific name of *C. assamica* ssp. *lasiocalyx* (Planchon ex Watt). The Chinese variety has the highest self-incompatible amount and the Cambodia variety has the lowest amount (Wight, 1956, 1959; Sealy, 1958).

Tea leaves contain several chemical components that have major physiological benefits, including flavones, flavonols, phenolic acids, caffeine, amino acids, organic acids, monopolysaccharides, lignin, photosynthetic pigments, ash and aromatic chemicals (Li *et al.*, 2020; Samadi and Fard, 2020; Vastrad *et al.*, 2022). Polyphenols are significant aromatic components in tea that have a lot of antioxidant activity and are good for our health (Kottawa-Arachchi *et al.*, 2019). Flavonoids, notably flavanols and organic acids, were the main polyphenols found in tea leaves. Catechins, which make about 60–80% of tea's polyphenols, are one of the most significant of them. The number of polyphenols in tea fluctuates based on temperature and soil changes. Tea has several medical qualities, including cardiovascular protection, cancer prevention, diabetes prevention and obesity prevention (Pan *et al.*, 2017; Tang *et al.*, 2019). Tea leaves'

photosynthetic pigments are also intimately linked to their medicinal effects and color (Wang *et al.*, 2019). Tea leaves from different cultivars and geographical locations may have varied compositions (Wen *et al.*, 2020). Despite the occurrences described, only less research on the genetic arrangement has been conducted. Crop improvement in tea has mostly relied on selecting elite genotypes through controlled cross-pollination, vegetative propagation in the vault and distribution as a commercial clone to date (Raina *et al.*, 2012; Azka *et al.*, 2019; Lee, 2019; Yadav *et al.*, 2020). Recent research studies have focused on DNA markers such as restriction fragment length polymorphism (RFLP), randomly amplified polymorphic DNA (RAPD), inter simple sequence repeats (ISSR), cleaved and amplified polymorphic sequences (CAPS) and amplified fragment length polymorphism (AFLP) for determining the characterization and identification of germplasm, and cases related to genetics in the tea (Raina *et al.*, 2012; Azka *et al.*, 2019; Lee, 2019; Yadav *et al.*, 2020). The identification of different groups of tea by morphological and biochemical descriptors is highly developed, but there are limited levels for identifying differences due to environmental factors. Hence, molecular DNA markers are a more accurate tool for assessing the genetic diversity and genetic association of the tea plant (Lee, 2019). In addition, morphological markers have been widely used to study tea plant diversity, and these studies have been able to identify the available diversity (Azka *et al.*, 2019). To date, various molecular markers such as RFLP (Matsumoto *et al.*, 2004), RAPD (Chen *et al.*, 2005; Ahmadishad *et al.*, 2009; Ramakrishnan *et al.*, 2009; Roy and Chakraborty, 2009; Huseynov *et al.*, 2022), AFLP (Balasaravanan *et al.*, 2003; Kafkas *et al.*, 2009; Raina *et al.*, 2012), SSR (Azka *et al.*, 2019; Lee *et al.*, 2019) and ISSR (Yao *et al.*, 2008; Liu *et al.*, 2009; Roy and Chakraborty, 2009; Ben-Ying, *et al.*, 2010), SRAP (Zhang *et al.*, 2018; Khiavi *et al.*, 2020a) and SCoT (Zhang *et al.*, 2018; Chaeikar *et al.*, 2020; Samarina *et al.*, 2022) have been used to investigate the genetic diversity and genetic relations of tea plant for different goals.

In Iran, studies on the genetic diversity of tea have been performed on a limited number of clones and genotypes in the north of the nation. Identifying and investigating the diversity of genotypes in main areas under tea cultivation is one of the important management challenges of the tea genetic breeding programmes in the country, hence the present study focused on extensively the genetic diversity of genotypes from two main commercial sites of tea production that had not been investigated before by morphological and molecular markers.

Materials and methods

Plant material

In this study, 30 genotypes of tea plant from two main commercial sites (Kohbijar and Bazkiagorab areas) of tea production in the southern region of the Caspian Sea were tested that had not been investigated before (Table 1). Two-year-old tea plants were propagated from single-leaf nodal cuttings from the desired areas and cultivated in a completely randomized block design (CRBD) with five replicates in a greenhouse with a photoperiod of 16/8 (lightness/darkness) and relative humidity of 70–80% in the Tea Research Institute in Lahijan city. The plants were grown in three-litre pots and filled with cocopeat and perlite (at a ratio of 2:1). The morphological and molecular diversity of plants was studied 1 year after their establishment.

Morphology analysis

The average of five propagated plants from each genotype were selected and statistical analysis was done to assess the morphological characteristics. To investigate morphological diversity, a total of 30 quantitative and qualitative traits were studied based on morphological data (Table S1). The studied traits were selected based on the description letter introduced by the International Plant Genetic Resources Institute (IPGRI) (IPGRI. 2000). The data were first standardized by the Ntsys software and the YBAR coefficient, and then, the matching matrix was calculated by the same software based on the Euclidean distance coefficient. The

Table 1. Tea samples used in morphological and ISSR analysis

No.	Scientific name	Location	Plant code	No.	Scientific name	Location	Plant code
1	*C. sinensis*	Kohbijar	K1	16	*C. sinensis*	Bazkiagorab	B3
2	*C. sinensis*	Kohbijar	K2	17	*C. sinensis*	Bazkiagorab	B4
3	*C. sinensis*	Kohbijar	K3	18	*C. sinensis*	Bazkiagorab	B5
4	*C. sinensis*	Kohbijar	K4	19	*C. sinensis*	Bazkiagorab	B6
5	*C. sinensis*	Kohbijar	K5	20	*C. sinensis*	Bazkiagorab	B7
6	*C. sinensis*	Kohbijar	K6	21	*C. sinensis*	Bazkiagorab	B8
7	*C. sinensis*	Kohbijar	K7	22	*C. sinensis*	Bazkiagorab	B9
8	*C. sinensis*	Kohbijar	K8	23	*C. sinensis*	Bazkiagorab	B10
9	*C. sinensis*	Kohbijar	K9	24	*C. sinensis*	Bazkiagorab	B11
10	*C. sinensis*	Kohbijar	K10	25	*C. sinensis*	Bazkiagorab	B12
11	*C. sinensis*	Kohbijar	K11	26	*C. sinensis*	Bazkiagorab	B13
12	*C. sinensis*	Kohbijar	K12	27	*C. sinensis*	Bazkiagorab	B14
13	*C. sinensis*	Kohbijar	K13	28	*C. sinensis*	Bazkiagorab	B15
14	*C. sinensis*	Bazkiagorab	B1	29	*C. sinensis*	Bazkiagorab	B16
15	*C. sinensis*	Bazkiagorab	B2	30	*C. sinensis*	Bazkiagorab	B17

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

UPGMA algorithm was used in the Ntsys software to design clusters based on the matching matrix. The studied traits are presented in Table 2. Data analysis was done using SAS (Ver. 9.0) statistical software and mean comparison was done by Duncan's test.

DNA extraction and molecular analysis

The method used in the study by Dellaporta *et al.* (1983) was utilized with a little change to extract DNA. The spectrophotometer method and Nano Drop device were used to be determined the quantity of the DNA extracted and to be determined its quality, DNA was electrophoresed in a 0.8% agarose gel and a voltage of 100 V.

The number of 21 ISSR primers for this study was investigated on the samples to be optimized and selected the best primers, ultimately, 12 numbers of them that had the best, strongest and most repeatable propagating were selected to continue the investigation. The sequence of the used ISSR primers is presented in Table S1. To run the polymerase chain reaction (PCR), a mix of the material within each reaction contained the DNA template (50 ng), PCR buffer (10×) (25.1 µl), magnesium chloride (2 mM), each dNTP at 0.2 mM, Taq polymerase enzyme (1 U) and primer (0.75 µM), which finally, these were added to distilled water until it reached the volume of 5.12 µl. The conditions of the polymerase chain reaction for the primers were also set as follows: Firstly, the reaction solution was placed for 4 min at 94°C. Then, this solution entered the second step, which the step was 35 cycles, and the conditions were set as follows: 94°C for 1 min, 52°C for 1 min, and 72°C for 1 min and 30 s, and the reaction solution was kept at 72°C for 7 min to finally propagate. Then, the samples were kept at 4°C until the electrophoresis time. A Bio-Rad thermal cycler system (i-Cycler) was used.

Table 2. Investigated morphological characteristics

No.	Characters	No.	Characters
1	Plant height	16	Leaf shape
2	Internode length	17	Leaf upper surface
3	Length of mature leaf	18	Leaf apex shape
4	Width of mature leaf	19	Leaf apex habit
5	Length/ Width of mature leaf	20	Leaf base shape
6	Length of mature leaf petiole	21	Leaf margin
7	Tree habit	22	Leaf size
8	Growth habit	23	Leaf angle
9	Stem type	24	Leaf venation
10	Stem color	25	Leaf pose (angle)
11	Branch angle	26	Leaf waxiness
12	Pigmentation in young leaves (in growth season)	27	Petiole color
13	Pigmentation in young leaves (in off season)	28	Shoot density
14	Immature leaf color	29	Young shoot color
15	Mature leaf color	30	Mature shoot color

Statistical analysis

Products propagated were separated into the ratio of 5 µl of PCR products, 3 µl of loading buffer and 2 µl of gel red (Biotium, USA) in a 1.5% agarose gel under a constant voltage of 90 V for 120 min and under UV light, the gel was photographed by the gel documentation system (Biometra). The high-resolution and unambiguous polymorphic bands in the agarose gel of primers according to the existence and absence were scored as 0 and 1, respectively. Uncertain bands with a low resolution, which are a sign of lack of desirable propagating for different reasons (such as lack of correct primer connection or competition between the positions of primer connection, etc.) and also the monomorphic bands were absent in the scoring. Polymorphic information content (PIC) was calculated by the Excel software based on the equation of $PICi = 2fi\ (1-fi)$ that has been recommended for dominant markers. In the equation, fi is the frequency of ith fragment marker in a state of existence and (1−fi) is the frequency of ith fragment marker in a state of absence (Roldan-Ruiz *et al.*, 2000). After forming the matrix of zero and one and transferring them to the Ntsys software, the SM matching coefficient matrix was formed and cluster analysis of samples was done by the UPGMA algorithm. In addition, separating into original components was done by the SPSS software. The data were used to calculate allele frequencies (na), the number of effective alleles (ne), Shannon index (I), diversity within population (Hs), total diversity (Ht) and the ratio of diversity between population to the total diversity (Gst) by the Popgene software.

The dataset analyzed included morphology and ISSR (binary) data. Principal variance analysis (PVA) was utilized to identify the most influential variables contributing to the overall variance, thereby revealing the key factors influencing the dataset. Principal coordinate analysis (PCoA) was used to visualize the relationships between samples in a reduced-dimensional space, thereby facilitating the identification of clusters and patterns. For the analysis, the R statistical software with dedicated utilities for PVA and PCoA was utilized, allowing for accurate data processing and visualization.

Results

Morphology analysis

Identifying the morphology was done based on the description letter recommended by the IPGRI Institute with the application of 30 quantitative and qualitative traits in some tea genotypes collected from two main tea-growing areas of Kohbijar and Bazkiagorab, Iran. Comparison analysis in 30 tea plant samples showed little variety. In this study, traits related to leaf and stem were used (Table S1).

Based on the morphology data and the Euclidean distance coefficient, and the UPGMA algorithm, the samples were divided into four groups at the different level of 0.63, as observed in Fig. 1. The first and third groups (A and C), each had three members. Group two (B) was the largest, so that included about 53% of the total samples. This group was divided into four subgroups at the level of 0.50, the first (B-1), second (B-2), and third (B-3) subgroups, have two, three, and four members, respectively, but the fourth subgroup had seven members was the largest

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

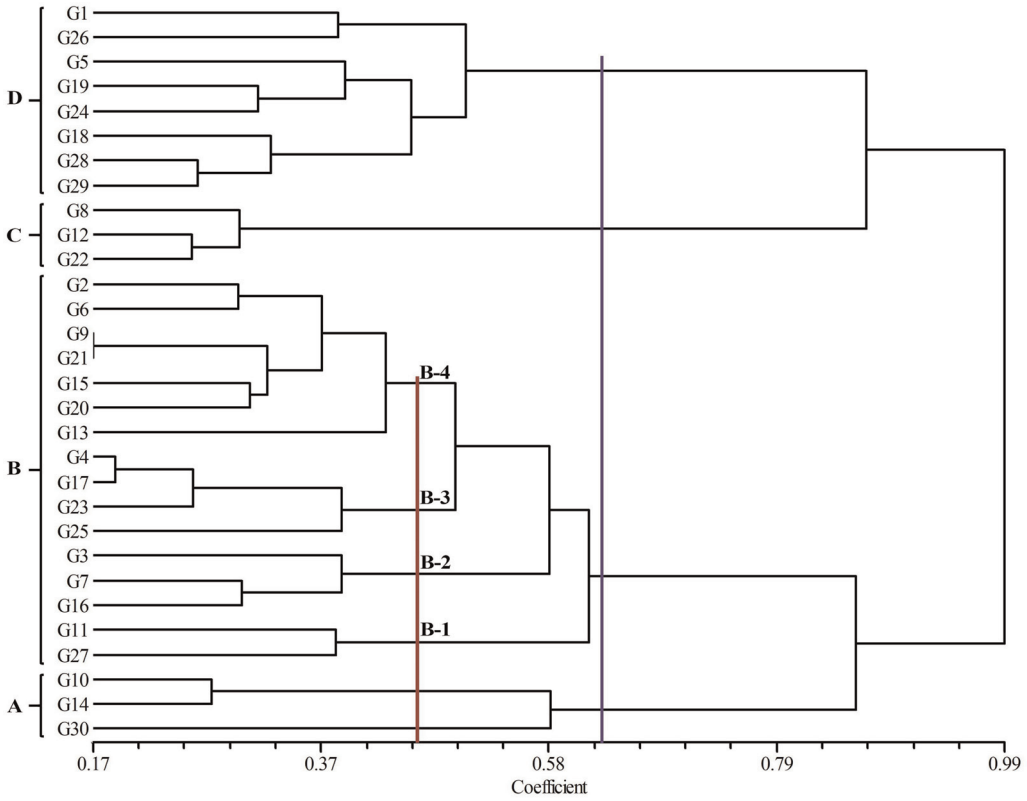

Figure 1. Dendrogram of 30 tea samples using morphological markers based on UPGMA.

subgroup of this main group. The fourth group with eight members included about 27% of the study samples. Based on the amount of matching calculated between the samples observed that the difference range between the samples was very limited and the samples showed little variation based on the morphological data.

Molecular analysis

12 ISSR primers were used in four tea samples based on the reproductive power and repeatability to be investigated the genome diversity of 30 tea extensions collected from two commercial tea sites in Iran, after studying 21 primers that were absent in this study. The used primer information is presented in Table S1. Among the 12 primers used, 111 bands that had desirable characteristics, were identified for study, and 91 bands showed polymorphic mode (81.98% of bands were polymorphic).

The average number of bands for each primer was 9.25, and the average number of polymorphic bands for each primer was 7.58. Primer P9 with the production of 16 bands and primers P2 and 911 with the production of 6 bands had the highest and lowest number of total produced bands. Primers P9 and P5, with the number of 11 and 5 bands (respectively), had the highest number of polymorphic bands. One hundred percent of bands produced in four primers of P9, P7, P10 and P11 also had a polymorphic mode. Changes in polymorphism percentage ranged from 68.75 to 100%. The reproduced fragments were in size from 100 to 2000 bp, but the range of fragments to be examined was 200–1500 bp.

To identify the potential of primers used in the present study, the amount of the polymorphic information content (PIC) was calculated for the primers, the maximum amount of PIC was for primer P2 at the amount of 0.48 and the minimum amount for primers P1 and P10 (0.41). The total polymorphic information content was calculated at 0.45. Table S1 clearly shows that all primers have high PIC, demonstrating their strong ability to identify and distinguish differences. Some samples that have been spread using the ISSR4 marker are shown in Figure S1.

The cophenetic coefficient was calculated for three coefficients of DICE, Jaccard and simple matching to determine the best coefficient for calculating the 2-by-2 matching matrix and cluster analysis. Furthermore, the simple matching coefficient had the lowest rate of cophenetic coefficients (0.81). This coefficient indicates how much the data of the calculated matching matrix has been transferred to the designed cluster. On this basis, the simple matching coefficient to calculate the matching matrix and the UPGMA algorithm to design the cluster was used. The amount of matching of the samples was done by using the simple matching coefficient, and the maximum and minimum matching obtained was 0.93 and 0.24, respectively, and the average amount of matching was calculated as 0.62. The maximum amount of matching was between the two samples of G12 and G13, and the minimum amount was between the two samples of G9 and G24.

The samples were analyzed under the cluster based on the data obtained from the ISSR markers by the simple matching coefficient (SM) and the UPGMA algorithm. In the cluster obtained at the matching level of about 58%, the samples were divided into five groups. Figure 2 shows graphs resulting from the ISSR data and how to place the samples.

According to the molecular data in the cluster analysis, samples were divided into five groups at the matching level of 0.58, in which the first group (A) had a member and was isolated at

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

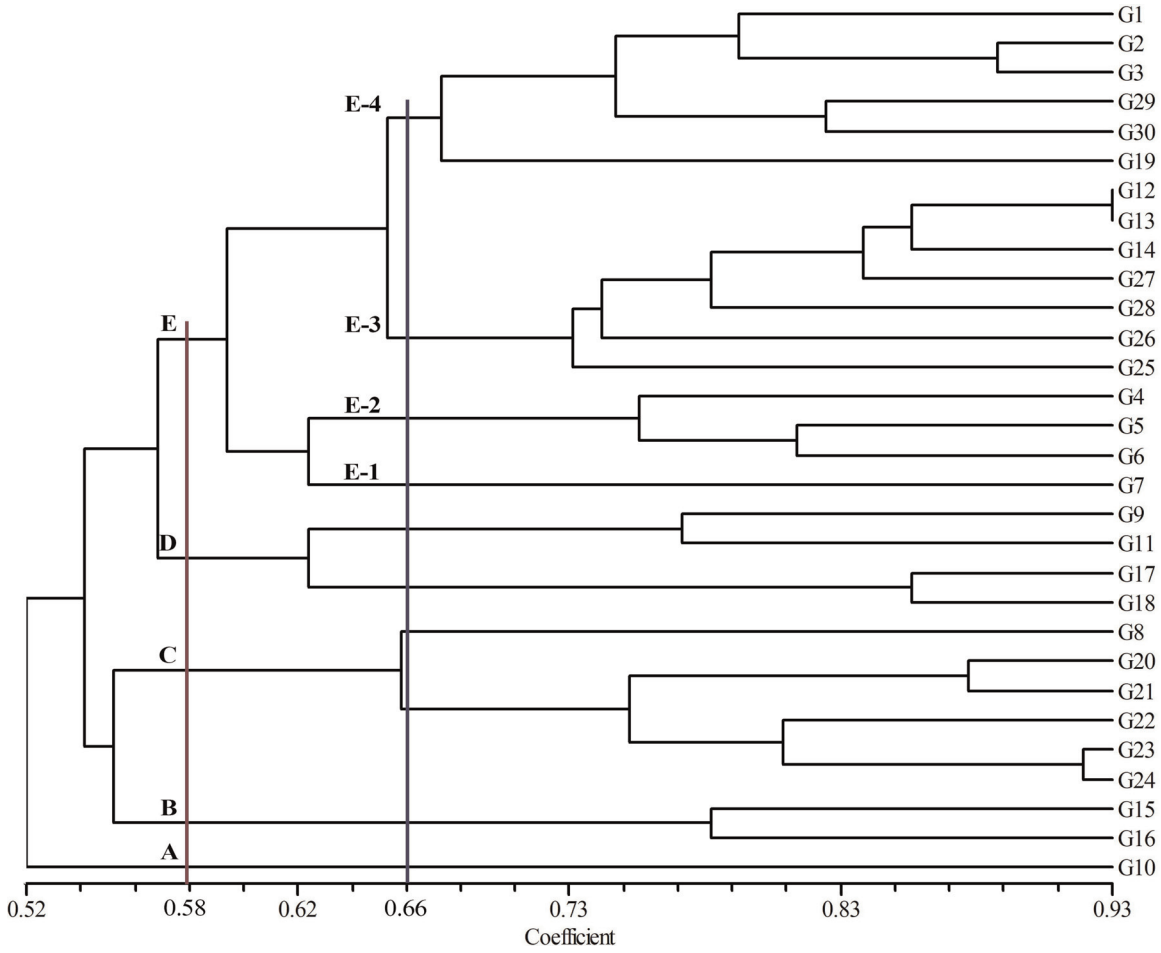

Figure 2. Dendrogram of 30 tea samples using the ISSR markers based on UPGMA.

the greatest distance from the other samples at the level of 54%. The second group (B) which had two members was isolated at the level of approximately 0.55 from other samples. The third group (C) which included the six samples was isolated at the level of 0.55 from other samples. The fourth group (D) also had four samples, which were isolated at the level of 57% from the others. The fifth group (E) was the main group, which was identifiable in the cluster analysis with the number of 17 members (56.66% of the total examined samples). This group was also divided into four subgroups at the matching level of 0.66. The first two subgroups (E-1 and E-2) were limited in terms of the number of members (E-1 = 1 member, E-2 = 3 members), but the third and fourth two subgroups (E-3 and E-4), each having seven and six members (respectively), totally included about 45% of the total examined samples.

PCA analysis

The PCA analysis was performed to evaluate the genetic diversity of tea genotypes based on morphological traits. Six quantitative and 24 qualitative characteristics, for a total of 30, were evaluated. The analysis revealed a small quantity of variation in the morphological characteristics, indicating a low degree of diversity among the tea samples. Although there were minor differences in characteristics such as branch angle, leaf tip and blade margin shape, internode length, and blade size (length and breadth), the overall morphology was similar. The samples were separated into four distinct categories using cluster analysis with a similarity threshold of 0.50 (Fig. 3).

PCoA analysis

A PCoA analysis was conducted using ISSR markers to investigate the genetic diversity of tea genotypes. Utilizing a total of 12 ISSR primers resulted in the formation of 91 polymorphic bands. With a range of 0.41–0.48, the Polymorphic Information Content (PIC) test identified informative markers for assessing genetic diversity. The PCoA analysis based on the ISSR data revealed a higher level of genetic diversity, with the tea genotypes exhibiting distinct clustering patterns. At a similarity level of 0.58, the samples were organized into five separate clusters (Fig. 4).

Heatmap

The heatmap data provide a visual representation of the coordinates of tea genotypes in a multi-dimensional space based on five dimensions (Dim.1 to Dim.5) obtained from the analysis (Figure S2). Each dimension corresponds to a combination of morphological or ISSR parameters. The values in the heatmap indicate the relative positions of the tea genotypes along each dimension, allowing for the observation of patterns and trends in their distribution (Figure S2).

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

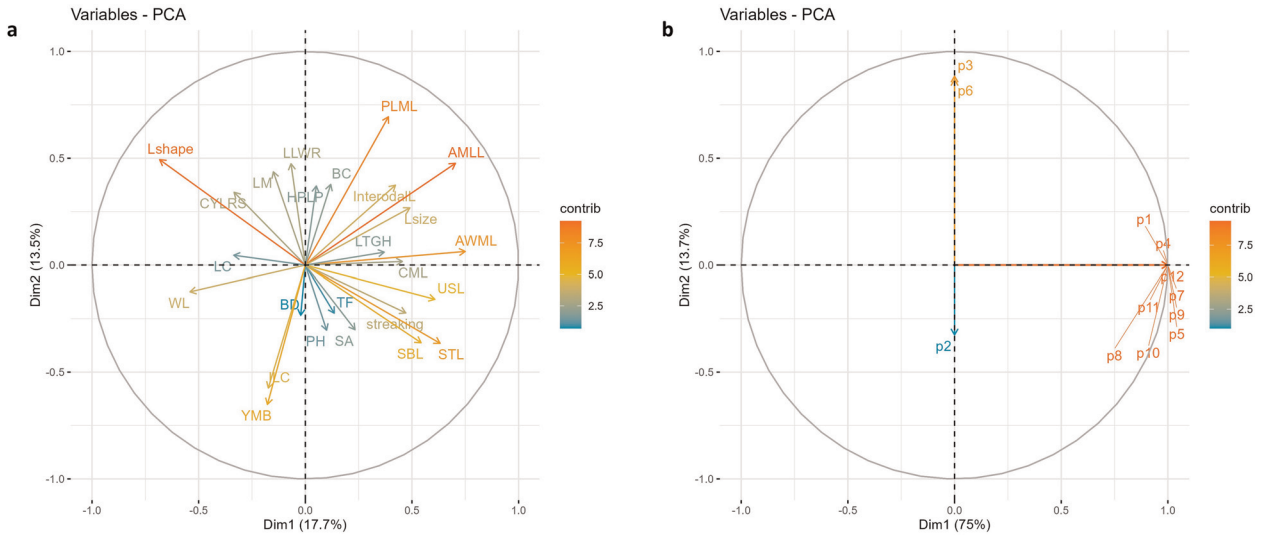

Figure 3. Combined PCA plots showcasing the genetic diversity of tea genotypes. (a) PCA plot based on morphological parameters highlights the limited variance among the tea samples, indicating a morphologically homogeneous population. PH, Plant height; Interodal L, Internodal length; AMLL, Average mature leaf length; AWML, Average width of mature leaf; LLWR, Leaf length to width ratio; PLML, Petiole length of mature leaf; TF, tree form; GH, Growth habit; BF, branch form; BC, branch color; SA, stem angle; CYLDGS, Coloring in young leaves during the growing season; CYLRS, Coloration in young leaves in the recession season; ILC, Immature leaf color; CML, Color of mature leaves; Lshape, leaf shape; lea, The upper surface of the leaf; STL, The shape of the tip of the leaf; LTGH, Leaf tip growth habit; SBL, The shape of the base of the leaf; LM, leaf margin; Lsize, leaf size; LA, leaf angle, streaking; HPLP, How to place leaves on the plant; WL, Waxiness of the leaf; LC, Leaf color; BD, Branch density; CYB, The color of the young branch; YMB, Young mature branch. (b) PCA plot using ISSR primers demonstrates the effectiveness of molecular markers in identifying genetic differences among the tea genotypes.

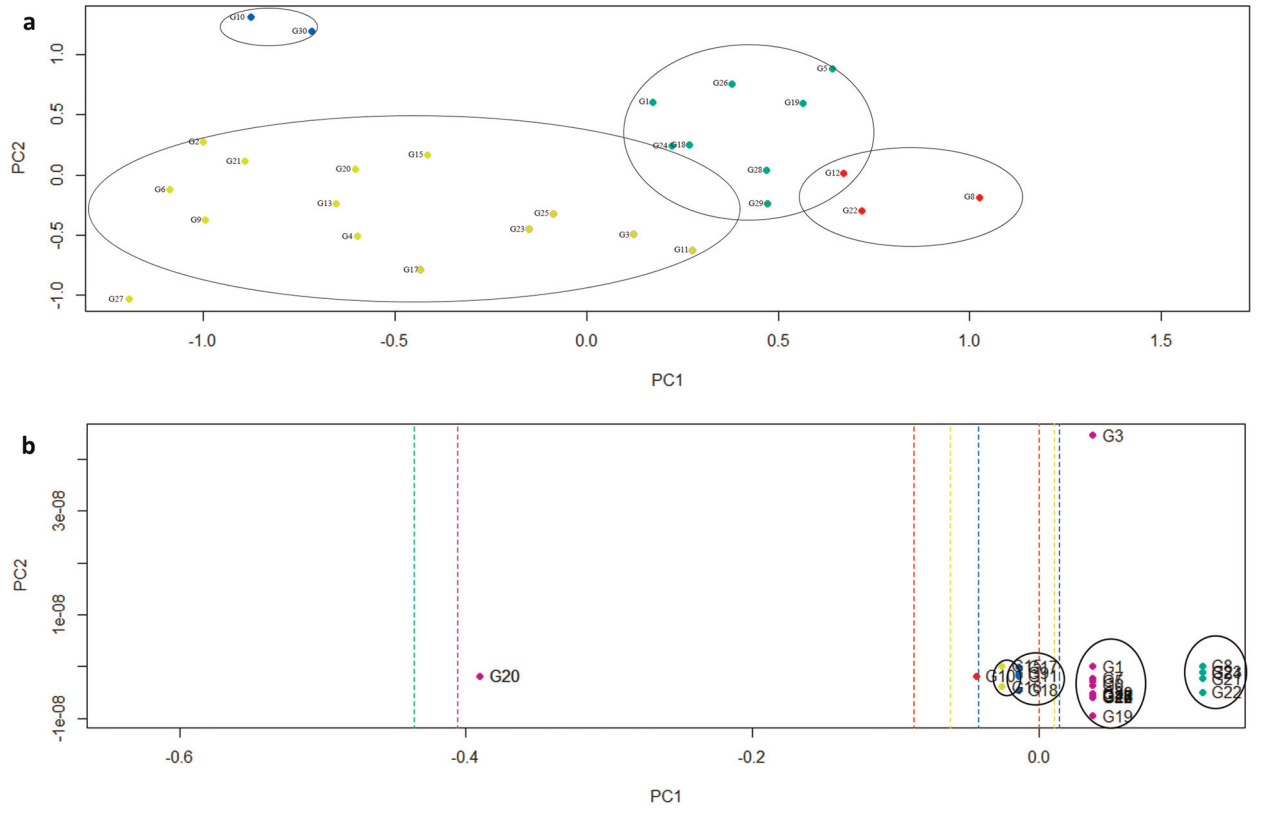

Figure 4. Combined PCoA plots revealing the clustering patterns of tea genotypes. (a) PCoA plot based on morphological parameters displays distinct clusters, providing insights into the genetic diversity of the tea genotypes. (b) PCoA plot using ISSR primers highlights discrete clusters, indicating significant genetic variation among the tea genotypes.

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

Discussion

Morphological analysis

Comparative analysis of 30 morphological traits in tea genotypes showed that there is only a small amount of variance. Although morphological and phytochemical characteristics can vary depending on climate and genetic changes (Ghanbari *et al*., 2022), morphological features continue an important and fundamental step in the accurate classification of plants (Kaouther *et al*., 2017; Ghanbari *et al*., 2023). However, in some circumstances, morphological data are insufficient to generate trustworthy conclusions that are consistent with molecular results (Zare Hoseini *et al*., 2022). Based on the Euclidean distance coefficient, the dual matching matrix of the samples was calculated, and this revealed a little difference. The narrow difference range for morphology was found in earlier investigations on the tea germplasm (Beris *et al*., 2016; Khiavi *et al*., 2020b) Conversely, Chen *et al*. (2005) showed a significant level of variation within several *Camellia* species. According to the results of the research of Khiavi *et al*. (2020b) on the diversity of tea plants in Lahijan, the range of differences between samples based on the Euclidean distance coefficient is much more limited than the range calculated in the present study. They expressed the maximum and minimum difference of 1.74 and 0.19, respectively, the main reason for which can be due to the limited sampling area and also the smaller number of studied traits in their study. In research looking at the morphological diversity of Mexican limes lime (*Citrus aurantifolia* (Christm) Swingle) in the southern region of Iran, whose propagation is mostly by seeds, Khiavi *et al*. (2016) observed considerable variations. These findings support the findings of the current study because the grown tea in Iran is reproduced using seeds from a single species (*C. sinensis* (L.) O. Kuntze).

According to the results of the present study, it can be expressed that the amount of difference range obtained in the studied collection is accepted. Various reasons for limiting the observed differences are noteworthy, which is the most important reason back to enter the plant tea in the past that the origin of all tea plants available in the germplasm of tea plant in Iran is resulting from seed propagation obtained from open pollination of three initial masses. A high matching range has also been reported with the use of AFLP markers (Kafkas *et al*., 2009) and ISSR markers (Beris *et al*., 2016) in studying the genetic diversity of tea plants in Turkey, where tea plants are imported similar to Iran. On the other hand, since the propagation tea plant was through seed in Iran in the past, seedlings were selected that had suitable characteristics for cultivation in the original place, such as starting to grow early in the early season or getting into the late dormancy at the end of the growing season, longer internode length and larger leaf size. Furthermore, due to the limited genetic diversity of the imported plants, reducing the genetic diversity is generally observed, which in these cases is known as a heterozygosity reduction (H) (Doebley, 1989).

The samples were sorted into four groups at the different level of 0.63 based on the morphological data, the Euclidean distance coefficient, and the UPGMA algorithm, as shown in Fig. 1. The estimated degree of matching between the samples revealed that the range of difference between them was extremely limited and that the samples showed just a little amount of variance based on the morphological data. These four groups' formation and separation may be explained by the fact that some morphological traits might change as a result of free confluence, as well as geographic and environmental factors. Despite the use of different accessions originating from Iran's tea-growing commercial areas, which are geographically wide, the lack of distribution of samples based on geographical areas was observed in cluster analysis. This result is similar to previous findings and the reason is the same primary origin of tea plants (*C. sinensis*) in the past (Chen *et al*., 2005; Ben-Ying *et al*., 2010; Chaeikar *et al*., 2020).

This type of clustering algorithm has received approval in earlier research (Rajanna *et al*., 2011; Khiavi *et al*., 2020b), and this grouping is generally consistent with the morphology and systematics of the species of *Camellia* that are currently known (Wright, 1962; Balasaravanan *et al*., 2003; Ramakrishnan *et al*., 2009; Rajanna, *et al*., 2011).

In general, it is possible to identify the samples using these markers based on the morphological information in the Iranian germplasm, according to earlier publications and the findings of the current study. However, more exact methods, such as molecular markers, are required to detect with more accuracy.

Molecular analysis

Out of the 12 primers utilized, 111 bands with desired properties were found for further analysis, and 91 bands displayed polymorphic mode (81.98% of bands were polymorphic). According to the results of other similar studies on tea plants (Thomas *et al*., 2006; Roy and Chakraborty, 2009; Khiavi, 2020b), the rate calculated percentage is acceptable. In a study of the diversity of tea plants obtained from the somaclonal variation, Thomas *et al*. (2006) reported a percentage of polymorphism of around 53%, while in research by Roy and Chakraborty (2009) to analyze the genetic diversity, this amount has been reported to be 88.54%. Yao *et al*. (2008) reported a very high percentage of polymorphism (99.7%) in their study of the genetic diversity of tea plants in China, Japan and Korea, which is due to the size of the sampling areas. As a result, their findings cannot be used to reject the polymorphism percentage found in the current study, because when we refer to the results of each country alone, polymorphism percentages are reduced and approached the range of the present study, so that, for China, where four tea growing areas have been sampled, the range of polymorphism percentage ranges from 54.2 to 88.2, and due to that Iranian tea plants originate from China, it was observed that the polymorphism percentage obtained in the study was strongly close to the maximum obtained in China.

The average number of bands and polymorphic bands for each primer was 9.25 and 7.58, respectively. Changes in polymorphism percentage varied from 68.75 to 100%; a similar range was reported in the research by Thomas *et al*. (2006). The size range of the replicated fragments was 100–2000 bp, whereas the range of the fragments that were to be inspected was 200–1500 bp. The range of the reproduced fragments by the ISSR markers was likewise the same in the research by Ben-Ying *et al*. (2010).

The total polymorphic information content was calculated at 0.45. It should be noted that about the amount of polymorphic information content for co-dominant markers such as ISSR, the maximum amount of polymorphic information content is 0.50 and the higher the calculated PIC value is closer to this, the greater the marker power used to distinguish and identify sample differences (Roldan-Ruiz *et al*., 2000).

Given that polymorphism information content must be between 0 and 0.5 and that the amount of polymorphic information content calculated is closer to 0.5, which indicates a better

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

sample resolution, it can be mentioned that the markers used have a very good ability to separate the tea samples. Khiavi *et al.* (2016) reported the range of changes of PIC from 0.42 to 0.49 in evaluating the genetic diversity of limes of Iran in the southern region that is the same genus and species and their propagation generally is through seed. They stated that the range of the amount of PIC indicated the high ability of this marker to examine the genetic diversity in the plant under study. As is clear in Table S1, all the primers have a high PIC, indicating their high power in separating and identifying the differences. Previous studies also have confirmed that this group of markers is suitable for evaluating the genetic diversity of the tea plant (Liu *et al.*, 2009; Roy and Chakraborty, 2009; Ben-Ying, *et al.*, 2010).

Using the basic matching coefficient, the amount of matching between the samples was determined. The greatest and minimum matching values were 0.93 and 0.24, respectively, and the average amount of matching was found to be 0.62. This level of matching is acceptable according to the results of the morphological section and these two indicators confirm each other. Therefore, it can be concluded that the two methods of morphology and ISSR markers are suitable for identifying differentiation within the genus *Camellia*. Liu *et al.* (2009) reported an average amount of matching of 50% to determine the genetic relationships related to the germplasm of tea by using the ISSR markers. Yao *et al.* (2008) calculated the matching range around 0.54– 0.16 in studying the genetic diversity of tea plants in China, Japan and Korea, this amount is less than that of the present research, due to extending the range investigated in their study. The high rate of matching obtained (0.93) is consistent with the studies by Kafkas *et al.* (2009) in the investigation of the genetic diversity of tea plants in Turkey using the AFLP marker. They calculated the maximum amount of matching 92%, which is accepted due to importing tea plants to these two countries (Iran and Turkey), and the high amount of matching between the samples obtains concerning seed propagation of the plant in the past. In addition, Ben-Ying *et al.* (2010) obtained the same amount of matching (the average matching of 0.51) with the application of ISSR markers in studying tea plants in Yunnan area, China. Given that the tea plant cultivated in Iran origins from China, this shows the close relationship of samples cultivated in Iran.

Samples were sorted into five groups at the matching level of 0.58 based on the molecular information in the cluster analysis. An extremely significant finding from the cluster analysis of the samples using ISSR markers was the lack of a clear geographic trend in the groups, which was also seen in the morphological analysis. In general, ISSR markers are highly useful for analyzing the genetic diversity of tea germplasm sources, and their application in this context is extremely logical (Liu *et al.*, 2009; Roy and Chakraborty, 2009; Ben-Ying *et al.*, 2010). This molecular marker is sensitive to detect the genetic relationship between people with close relationships (Wolfe and Liston, 1998). The results of using the ISSR marker to examine the degree of genetic diversity among the genotypes of the teas under investigation in this study show that there are not many differences between single tea samples from tea growing regions, and if there are any differences at all, they are not related to geographic distribution. Of course, to fully describe the topic, it is necessary to look at regional demographics, and ever-more-accurate markers like SSR and AFLP. The findings of this study are anticipated to assist scientists in designing breeding programs for certain goals, such as the introduction of novel, high-performing cultivars, managing biotic and abiotic challenges, and calculating and defining more appropriate crossings.

PCA and PCoA

The results of the PCA analysis revealed that the morphological traits of the tea genotypes exhibited a restricted range of variation. This discovery suggests a population with relatively uniform morphological characteristics. In contrast, the PCoA analysis based on ISSR markers revealed the genetic diversity of the tea genotypes, indicating a greater level of variation. The PCoA plot's clustering patterns provided additional evidence for the existence of discrete genetic groups within the tea population.

The divergent outcomes of the PCA and PCoA analyses emphasize the necessity of using molecular markers, such as ISSRs, for a comprehensive evaluation of genetic diversity. The ISSR markers revealed a larger genetic diversity than may be captured by morphological characteristics alone. This indicates that molecular markers can provide a more precise and exhaustive comprehension of the genetic structure and diversity of tea genotypes.

The findings indicate that tea genotypes from the southern region of the Caspian Sea contain valuable genetic resources. ISSR analysis revealed a high level of genetic diversity, which enables the use of these genetic resources in breeding programs aimed at developing tea cultivars with desirable traits.

Overall, the results of the PCA and PCoA analyses highlight the need to integrate morphological and molecular marker-based approaches in order to acquire a comprehensive understanding of the genetic diversity of tea genotypes. These findings have implications for the conservation of germplasm, breeding programs and the sustainable use of tea genetic resources in the southern Caspian Sea region.

Morphology heatmap

The heatmap based on morphology (Figure S2) demonstrates the effect of specific morphological traits on the positioning of tea genotypes. Internode length, leaf size, leaf shape, and leaf color all substantially contribute to the segregation of genotypes along distinct dimensions. This indicates that these traits are essential for distinguishing tea genotypes on the basis of their morphological characteristics.

ISSR-based heatmap

The ISSR-based heatmap (Figure S2) illustrates the effect of ISSR markers on tea genotype clustering. The clustering of genotypes with comparable ISSR banding patterns suggests their genetic relatedness. Genotypes are positioned along various dimensions based on the presence or absence of specific ISSR markers. This demonstrates the significance of ISSR markers for capturing genetic variation and distinguishing tea genotypes according to their genetic profiles.

The combined heatmap analysis combines morphological and genetic data to provide a comprehensive view of the relationships between tea genotypes. It demonstrates that the combination of morphological traits and ISSR markers improves our knowledge of the genetic diversity and relatedness of tea genotypes. The integration of these datasets contributes to the characterization and differentiation of tea genotypes by enabling the identification of distinct groups and patterns. These data are useful for breeding programs, germplasm preservation, and the selection of elite tea individuals. Additional research incorporating additional molecular markers and phenotypic traits would enhance our

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press

comprehension of tea genetics and facilitate the creation of improved tea varieties.

Based on the findings of the analysis of 30 morphological traits in tea genotypes, including six quantitative traits and 24 qualitative traits, it has been shown that there is only a small amount of variance. Although some of the studied tea samples displayed modest and subtle changes in traits like branch angle, the shape of the leaf tip and blade margin, internode length, blade size (length and width), etc., they were all morphologically comparable. The samples were split into four groups according to the cluster analysis at the 0.5 level. By using 12 ISSR primers, 91 polymorphic bands were produced. The range of the PIC test was between 0.41 and 0.48. The matching range was determined in the range of 0.24–0.93 using the ISSR data. Samples at level 0.58 were divided into five groups using cluster analysis.

The findings suggest that these traits and primers are particularly effective in identifying genetic differences. These markers allowed for the observation of genetic variation among tea genotypes, but this diversity was insufficient to distinguish between accessions from different geographical locations. The findings demonstrated the considerable genetic diversity of Iranian tea genotypes. As a result, more focus should be placed on identifying and preserving new elite tea individuals in gene banks to access appropriate selection strategies, breeding programmes and germplasm.

Conclusion

Based on the findings of the analysis of 30 morphological traits in tea genotypes, including 6 quantitative traits and 24 qualitative traits, it has been shown that there is only a small amount of variance. Although some of the studied tea samples displayed modest and subtle changes in traits like branch angle, the shape of the leaf tip and blade margin, internode length, blade size (length and width), etc., they were all morphologically comparable.

The samples were split into four groups according to the cluster analysis at the 0.5 level. By using 12 ISSR primers, 91 polymorphic bands were produced. The range of the PIC test was between 0.41 and 0.48. The matching range was determined in the range of 0.24–0.93 using the ISSR data. Samples at level 0.58 were divided into 5 groups using cluster analysis. The findings suggest that these traits and primers are particularly effective in identifying genetic differences. These markers allowed for the observation of genetic variation among tea genotypes, but this diversity was insufficient to distinguish between accession from different geographical locations. The findings demonstrated the considerable genetic diversity of Iranian tea genotypes.

In conclusion, the PCA and PCoA analyses of tea genotypes from the southern region of the Caspian Sea revealed key insights into their genetic diversity. The PCA analysis of morphological characteristics revealed restricted variation among the tea samples, indicating a morphologically homogeneous population. The PCoA analysis based on ISSR markers, on the other hand, revealed a greater level of genetic diversity and discrete clustering patterns among the tea genotypes. These findings highlight the significance of molecular markers in obtaining a more complete picture of genetic diversity than morphological traits alone. The observed genetic diversity among tea genotypes is indicative of the region's rich genetic resources. These results have implications for the development of tea breeding programs as well as the conservation and sustainable use of tea germplasm. For a more comprehensive understanding and application of tea's genetic resources, it would be beneficial to conduct additional research that integrates multiple methods. By leveraging the genetic diversity present in Iranian tea genotypes, appropriate selection strategies, breeding programmes and germplasm conservation efforts can be implemented to enhance the future of tea cultivation and utilization.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S147926212300045X.

Competing interest. The authors declare no competing interests.

References

Ahmadishad MA, Kazemitabar SK, BabaeianJelodar NA, Gholami M and KazemiPoshtmasari H (2009) An Assessment of genetic diversity in cultivated Tea (*Camellia sinensis* L.) clones in Iran using RAPD markers. *Journal of Crop Breeding* **1**, 65–76, (In Persian).

Azka NA, Widhianata H and Taryono T (2019) Morphological and molecular characterization of 5 accessions of tea (*Camellia sinensis* (L.) O. Kuntze) exploited to develop high quality and quantity yield. Paper presented at the AIP Conference Proceedings.

Balasaravanan T, Pius P, Kumar RR, Muraleedharan N and Shasany A (2003) Genetic diversity among south Indian tea germplasm (*Camellia sinensis, C. Assamica and C. Assamica spp. Lasiocalyx*) using AFLP markers. *Plant Science* **165**, 365–372.

Ben-Ying L, You-Yong L, Yi-Chun T, Li-Yuan W, Cheng H and Ping-Sheng W (2010) Assessment of genetic diversity and relationship of tea germplasm in Yunnan as revealed by ISSR markers. *Acta Agronomica Sinica* **36**, 391–400.

Beris FS, Pehlivan N, Kac M, Haznedar A, Coşkun F and Sandallı C (2016) Evaluation of genetic diversity of cultivated tea clones (*Camellia sinensis* (L.) kuntze) in the eastern Black Sea coast by inter-simple sequence repeats (ISSR). *Genetika-Belgrade* **48**, 87–96.

Chaeikar SS, Falakro K, Rahimi M, Jahangirzadeh Khiavi S and Ashourpour M (2020) The investigation of genetic diversity based on SCoT markers, morphological, and chemical characters in tea (*Camellia sinensis* L.) clones. *Journal of Horticulture and Postharvest Research* **3**, 269–284.

Chen L, Wang PS, Xia YM, Xu M and Pei SJ (2005) Genetic diversity and differentiation of *Camellia sinensis* L. (cultivated tea) and its wild relatives in Yunnan province of China, revealed by morphology, biochemistry and allozyme studies. *Genetic Resources and Crop Evolution* **52**, 41–52.

Dellaporta SL, Wood J and Hicks JB (1983) A plant DNA minipreparation: version II. *Plant Molecular Biology Reporter* **1**, 19–21.

Doebley J (1989) Isozymic evidence and the evolution of crop plants. In Soltis DE and Soltis PS (eds), *Isozymes in Plant Biology*. Portland, OR: Dioscorides, pp. 165–191.

Ghanbari MA, Salehi H and Moghadam A (2022) Genetic diversity assessment of Iranian Kentucky bluegrass accessions: I. ISSR markers and their association with habitat suitability within and between different ecoregions. *Molecular Biotechnology* **64**, 1244–1258.

Ghanbari MA, Salehi H and Jowkar A (2023) Genetic diversity assessment of Iranian Kentucky bluegrass accessions: II. Nuclear DNA content and its association with morphological and geographical features. *Molecular Biotechnology* **65**, 84–96.

Huseynov M, Suleymanova Z, Ojaghi J and Mammadov A (2022) Characterization and phylogeny analysis of Azerbaijan Tea (*Camellia sinensis* L.) genotypes by molecular markers. *Cytology and Genetics* **56**, 285–291.

IPGRI (2000) Descriptors for tea (*Camellia Sinensis*). Rome, Italy: International Plant Genetic Resources Institute, p. 50.

Kafkas S, Ercisxli S, Dogan Y, Erturk Y, Haznedar A and Sekban R (2009) Polymorphism and genetic relationships among tea genotypes from Turkey revealed by amplified fragment length polymorphism markers. *Journal of the American Society for Horticultural Science* **134**, 428–434.

Kaouther M, Soumaya K, Wahbi J, Wolfgang F and Abdellah C (2017) Morphological and molecular evaluation of genetic diversity of wild

tunisian oregano, *Origanum vulgare* L. subsp. *glandulosum* Desf. Ietswaart. *Asian Journal of Biology* **3**, 1–5.

Khiavi SJ, Hamidoghli Y, Golein B and Sabouri A (2016) Assessment of lime genetic diversity in three regions of Iran using morphological and ISSR markers. *Agricultural Communications* **4**, 18–29.

Khiavi SJ, Azadi Gonbad R and Falakro K (2020a) Identification of genetic diversity and relationships of some Iranian tea genotypes using SRAP markers. *Journal of Horticulture and Postharvest Research* **3**, 25–34.

Khiavi SJ, Falakro K, Safaei Chaeikar S, Ramzi S and Kahneh E (2020b) Usage of morphological and ISSR markers for investigation of tea genotypes. *Plant Production Research Journal* **26**, 131–147, (In Persian).

Kottawa-Arachchi JD, Gunasekare MK and Ranatunga MA (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.

Lee KJ, Lee J-R, Sebastin R, Shin M-J, Kim S-H, Cho G-T and Hyun DY (2019) Assessment of genetic diversity of tea germplasm for its management and sustainable use in Korea genebank. *Forests* **10**, 780.

Li J, Wang J, Yao Y, Hua J, Zhou Q, Jiang Y and Dong C (2020) Phytochemical comparison of different tea (*Camellia sinensis*) cultivars and its association with sensory quality of finished tea. *LWT* **117**, 108595.

Li JW, Li H, Liu ZW, Wang YX, Chen Y, Yang N, Hu ZH, Li T and Zhuang J (2023) Molecular markers in tea plant (*Camellia sinensis*): applications to evolution, genetic identification, and molecular breeding. *Plant Physiology and Biochemistry* **198**, 107704.

Liu B, Wang L, Li Y, He W, Zhou J, Wang P and Cheng H (2009) Genetic diversity in tea (*Camellia sinensis*) germplasms as revealed by ISSR markers. *Indian Journal of Agricultural Sciences* **79**, 715–721.

Matsumoto S, Kiriiwa Y and Yamaguchi S (2004) The Korean Tea plant (*Camellia sinensis*): RFLP analysis of genetic diversity and relationship to Japanese tea. *Breeding Science* **54**, 231–237.

Mondal T, Satya P and Medda P (2003) India needs national tea germplasm repository. International Conference on Global Advances in Tea Science. Calcutta, India, pp. 58–59.

Pan H, Wang F, Rankin GO, Rojanasakul Y, Tu Y and Chen YC (2017) Inhibitory effect of black tea pigments, theaflavin 3/3'-gallate against cisplatin-resistant ovarian cancer cells by inducing apoptosis and G1 cell cycle arrest. *International Journal of Oncology* **51**, 1508–1520.

Raina SN, Ahuja P, Sharma R, Das S, Bhardwaj P and Negi R (2012) Genetic structure and diversity of India hybrid tea. *Genetic Resources and Crop Evolution* **59**, 1527–1541.

Rajanna L, Ramakrishnan M and Simon L (2011) Evaluation of morphological diversity in south Indian tea clones using statistical methods. *Maejo International Journal of Science and Technology* **5**, 1–12.

Ramakrishnan M, Rajanna L, Narayanaswamy P and Simon L (2009) Assessment of genetic relationship and hybrid evaluation studies in tea (*Camellia sp.*) by RAPD. *International Journal of Plant Breeding* **3**, 144–148.

Roldan-Ruiz I, Calsyn E, Gilliland T, Coll R, Van Eijk M and De Loose M (2000) Estimating genetic conformity between related ryegrass (*Lolium*) varieties. 2. Aflp characterization. *Molecular Breeding* **6**, 593–602.

Roy S and Chakraborty B (2009) Genetic diversity and relationships among tea (*Camellia sinensis*) cultivars as revealed by RAPD and ISSR based fingerprinting. *Indian Journal of Biotechnology* **8**, 370–376.

Samadi S and Fard FR (2020) Phytochemical properties, antioxidant activity and mineral content (Fe, Zn and Cu) in Iranian produced black tea, green tea and roselle calyces. *Biocatalysis and Agricultural Biotechnology* **23**, 101472.

Samarina LS, Matskiv AO, Shkhalakhova RM, Koninskaya NG, Hanke M-V, Flachowsky H and Liu S (2022) Genetic diversity and genome size variability in the Russian genebank collection of tea plant [*Camellia sinensis* (L.) O. Kuntze]. *Frontiers in Plant Science* **12**, 800141.

Sealy JR (1958) *A Revision of the Genus Camellia*. London: Royal Hort. Soc.

Sharma S, Kumar A, Rajpal VR, Singh A, Babbar S and Raina SN (2023) Evaluation of genetic diversity and population structure in elite south Indian tea [*Camellia sinensis* (L.) Kuntze] using RAPD and ISSR markers. *Genetic Resources and Crop Evolution* **70**, 381–398.

Shen W, Li H, Teng R, Wang Y, Wang W and Zhuang J (2019) Genomic and transcriptomic analyses of HD-Zip family transcription factors and their responses to abiotic stress in tea plant (*Camellia sinensis*). *Genomics* **111**, 1142–1151.

Tang GY, Zhao CN, Xu XY, Gan RY, Cao SY, Liu, Q and Li HB (2019) Phytochemical composition and antioxidant capacity of 30 Chinese teas. *Antioxidants* **8**, 180.

Thomas J, Vijayan D, Joshi SD, Lopez SJ and Kumar RR (2006) Genetic integrity of somaclonal variants in tea (*Camellia sinensis* (L.) O. Kuntze) as revealed by inter simple sequence repeats. *Journal of Biotechnology* **123**, 149–154.

Vastrad JV, Badanayak P and Goudar G (2022) Phenolic compounds in Tea: phytochemical, biological, and therapeutic applications. In *Phenolic Compounds-Chemistry, Synthesis, Diversity, Non-Conventional Industrial, Pharmaceutical and Therapeutic Applications*. IntechOpen, p. 452.

Wang X, Liu BY, Zhao Q, Sun X, Li Y, Duan Z, Miao X, Luo S and Li J (2019) Genomic variance and transcriptional comparisons reveal the mechanisms of leaf color affecting palatability and stressed defense in tea plant. *Genes* **10**, 929.

Wen B, Ren S, Zhang Y, Duan Y, Shen J, Zhu X and Fang W (2020) Effects of geographic locations and topographical factors on secondary metabolites distribution in green tea at a regional scale. *Food Control* **110**, 106979.

Wight W (1956) Commercial selection and breeding of tea in India. *World Crops* **8**, 263–268.

Wight W (1959) Nomenclature and classification of the tea plant. *Nature* **183**, 1726–1728.

Wolfe AD and Liston A (1998) Contributions of the polymerase chain reaction to plant systematics. In *Molecular systematics of plants II: DNA sequencing* , Boston, MA: Springer US. pp. 43–86.

Wright W (1962) Tea classification revised. *Current Science Bangalore* **31**, 298–299.

Yadav KC, Parajuli A, Khatri BB and Shiwakoti LD (2020) Phytochemicals and quality of green and black teas from different clones of tea plant. *Journal of Food Quality* **6**, 1–13.

Yao M, Chen L and Liang YJ (2008) Genetic diversity among tea cultivars from China, Japan and Kenya revealed by ISSR markers and its implication for parental selection in tea breeding programmes. *Plant Breeding* **127**, 166–172.

Zare Hoseini R, Mehregan I, Ghanbari Jahromi M, Mousavi A and Salami SA (2022) Evaluating molecular and morphological diversity of *Phlomis olivieri* Benth (Lamiaceae) populations in Iran. *Biodiversity* **23**, 81–95.

Zhang Y, Zhang X, Chen X, Sun W and Li J (2018) Genetic diversity and structure of tea plant in Qinba area in China by three types of molecular markers. *Hereditas* **155**, 1–12.

https://doi.org/10.1017/S147926212300045X Published online by Cambridge University Press