Genetic diversity and SSR markers development of ancient *Camellia sinensis* in Sandu County of Guizhou Province

Yichen Zhao (✉ yczhao@gzu.edu.cn)

Guizhou University <https://orcid.org/0000-0002-4778-2930>

Runying Wang

Guizhou University

Qing Liu

Guizhou University

Xuan Dong

Guizhou University

De-Gang Zhao

Guizhou University

Research Article

Keywords: Ancient *Camellia sinensis*, Genetic diversity, Phenotypic characters, SSR molecular markers

Posted Date: February 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-208220/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The genetic abundance of ancient tea germplasm has been preserved in the long evolution process, which provides valuable resources for the protection and breed selection of ancient tea germplasm. However, the limited studies related to the genetic diversity of ancient tea germplasm restrict protection and breed selection. Therefore, the genetic diversity of ancient tea germplasm in Sandu county of Guizhou Province was explored in this study. **II.Methods and Results:** The genetic diversity was analyzed using phenotypes and SSR markers. The ranges for the variation coefficients of the six quantitative and seven qualitative characters were 17.76%-60.37% and 18.58%-50.64%, respectively. The ranges of diversity indices of the six quantitative and seven qualitative characters were 1.72-2.74 and 0.55-0.84, respectively. Ninety-six bands were amplified using 15 pairs of SSR primers from the 145 samples, and the average polymorphism information index was 0.66. The analysis revealed that the average values of Nei's genetic diversity index (H) and the Shannon information index (I) are 0.26 and 0.41, respectively. Further, a genetic similarity coefficient of 0.734 shown by UPGMA dendrogram classified the 145 samples decreased into four groups. **III.Conclusions:** This study revealed the rich phenotypic variation and high molecular genetic diversity and the genetic diversity of the arbor is higher than that of the shrub of the ancient tea germplasm in Sandu of Guizhou province. Thus, this study not only provides a theoretical basis for the protection and breed selection but also promotes further research of ancient tea germplasm.

Introduction

Camellia sinensis is a perennial cross-pollinating plant having high genetic diversity and good genetic differentiation {Karthigeyan, 2008 #2}{Karthigeyan, 2008 #2}(Karthigeyan et al. 2008). It is a shrub or small tree from the genus *Camellia* in the family *Camellia*. It has high biodiversity and is an abundant resource of germplasm, and an essential strategic resource for the sustainable development of the tea industry(Danhui 2013; Lin et al. 2020). The ancient tea germplasm is a very good treasure of resources (Liu et al. 2010). The exploration of the genetic diversity of this germplasm may facilitate effective protection and utilization of these tea resources (Niu et al. 2019). However, due to man-made and natural causes, these valuable resources are gradually diminishing (Tripathi and Negi 2006).

Guizhou is one of the origin places of tea; Guizhou Sandu Aquarium Autonomous County is located in the Yunnan-Guizhou Plateau, the territory of mountain peaks, fertile soil, mineral elements and organic material rich, year-round cloud and mist, frost-free period long, abundant rainfall, year-round average temperature up to 18 °C, is the natural growth of tea forest (Zheng 2017). It is reported that Sandu Aquarium Autonomous County is rich in ancient tea germplasm group. Therefore, samples of ancient tea germplasm from this area were collected for this study, hoping to have a certain understanding of the biodiversity of ancient tea germplasm, and to screen and get the molecular markers that can distinguish the ancient tea germplasm in this region, and to develop the molecular markers. The most effective detection method for plant variety identification is DNA molecular marker technology (Jiao et al. 2018). Good polymorphism, ease of operation, and standardization, in addition to being unaffected by

environmental factors, crop growth period, and other factors, are the advantages of this technology (Koebner et al. 2001; Lu et al. 2018). Furthermore, advantages, such as codominance, simple operation, good universality, and high reproducibility, have made the SSR marker an essential agricultural industry standard for the identification of wheat (Thungo et al. 2020), watermelon (Mujaju et al. 2011), pepper (Lu et al. 2011), rice (Raza et al. 2020) and other crop varieties in China, among various molecular marker methods (Raza et al. 2020).

In this study, the genetic diversity of ancient tea germplasm in Sandu of Guizhou was analyzed using phenotypic characters and SSR markers combined with capillary electrophoresis technology to enable the exploration of these high-quality resources. Additionally, this study also provides a theoretical basis for the protection and promotes further research of ancient tea germplasm.

Materials And Methods

Materials

The materials collected from four plots of ancient tea germplasm in May 2019 in Sandu aquatic autonomous county of Guizhou Province, for this study, are given in Table S1.

Phenotypic statistical method

Phenotypic characters of the resources of ancient tea germplasm viz., tree shape, leaf color, leaf length, leaf width, leaf tip, leaf tooth depth, leaf tooth density, leaf vein logarithm, leaf area, leaf shape, altitude and height of individual trees were recorded immediately after sample collection (M et al. 2016). It should be noted that the above-mentioned data was surveyed by a single person to minimize the visual errors for certain phenotypic characters.

Standardized assignment of data of phenotypic character

Phenotypic characters used in the study are categorized as quantitative and qualitative characters, based on either direct or descriptive methods of measurement, respectively. Examples of quantitative characters are leaf length and leaf width, whereas those of qualitative characters are leaf color, leaf shape, and leaf tip (Kim et al. 2012). The data values for the quantitative characters were calculated following the direct measurement of collected samples. The 'Description Specification and Data Standard of Tea Germplasm Resources' was used to standardize the qualitative characters of the collected samples in this study (Table S2) (Steel 1992).

Statistical analysis of phenotypic characters

Following data standardization, the SPSS 19.0 software was used to analyze the genetic diversity of phenotypic characters (Inês et al. 2015). The genetic diversity index was calculated using the formula for Shannon-Weaver index (H): **see formula 1 in the supplementary files section**, where P_j is the frequency of the J code of a certain character (Bajracharya et al. 2006).

Cluster analysis of phenotypic characters

Cluster analysis was carried out by using standardized data of phenotypic characters of ancient tea germplasm. Initially, the Euclidean distance between 145 materials was calculated. This was followed by using the intergroup connection method in SPSS software (default) to obtain the cluster diagram (Benedetti et al. 2007).

SSR molecular marker technology

Genomic DNA of ancient tea germplasm was extracted using the CTAB method (Tamari et al. 2013). A total of 15 pairs of primers with good polymorphism, stable amplification, and interspecific differences were used (Table 1). The components of the PCR mixture for amplification are 1.5 μ L DNA (2 ng/ μ L), 10.72 μ L ddH₂O, 1.5 μ L 10*Buffer (Mg²⁺), 0.6 μ L dNTP (2.5 mmol/L), 0.08 μ L Easy Taq (5U/ μ L), 0.3 μ L Primer F (10 μ mol/L), and 0.3 μ L Primer R (10 μ mol/L). The Touchdown PCR program was used. This program is divided into two stages; First stage-4 min pre-denaturation at 94 °C, 40s denaturation at 94 °C, 30s annealing at Tx+8 °C (Tx is the optimal annealing temperature for the primer), 72 °C extension for 90s, 16 cycles (each cycle is reduced by 0.5 °C). The second stage- the annealing temperature is reduced to Tx, 94 °C modified 40s, Tx °C annealing 30s, 72 °C extensions for 90s, 24 cycles, 72 °C extensions for 12 min and 12 °C storage. The SSR-PCR products were detected by capillary electrophoresis, and the original matrix of 0–1 was derived (Gomes et al. 2018). This experiment using SSR markers was done in triplicate. POPGENE 32 was used to compute the observed number of alleles (Na), the effective number of alleles (Ne), Nei gene diversity index (H), Shannon diversity information index (I), genetic distance, and genetic consistency (Yeh et al. 2000). The NTSYS PC-2.1 software was used for cluster analysis following the unweighted average method (UPGMA), which was used to construct the cluster diagram (Rohlf et al. 2000).

Results

Diversity analysis of phenotypic characters

Results of statistical analysis of seven qualitative characters of 145 ancient tea germplasm resources are tabulated in Table S3. It can be observed that leaf tooth depth displayed the highest coefficient of variation (50.64%), whereas leaf quality showed the lowest coefficient of variation. The coefficients of variation for phenotypic characteristics of leaves such as their tooth depth, tip, color, size, shape, quality, as well as for that of tree shape ranged within the minimum and maximum values of coefficients of variation (Table S3). Further, the diversity indices for all phenotypic characters except leaf quality were found to be higher than 0.75. The diversity index for leaf quality was 0.55. Leaf color displayed the highest diversity index of 0.84.

A statistical analysis of six quantitative characters shown in Table S4 revealed that tree height had the highest (60.37%) and leaf vein logarithm had the lowest (17.76%) coefficients of variation, respectively. The coefficients of variation decreased in the following manner of quantitative characters-tree height, leaf

area, the logarithm of leaf teeth, leaf length, leaf width, and logarithm of leaf vein. The diversity indices of all quantitative characters were above 1.5, the lowest Leaf vein logarithm showed the lowest (1.72), whereas the leaf area displayed the highest (2.74) diversity indices, respectively. Resultant values of the statistical analysis indicate the rich genetic diversity in the phenotypic characters of ancient tea germplasm in Sandu County, Guizhou province.

Phenotypic character diversity analysis of five ancient tea germplasm populations

The variation coefficients of ancient tea germplasm were observed to significantly vary among different populations and for phenotypic characteristics (Table S5). The tree type for the ancient tea germplasm collected from Landong and Guqi villages were shrubs, and hence the coefficient of variation was not included in the study. Among the ancient tea germplasm of arbors, the variation coefficient of leaf tip in Yangmeng village was observed to be the highest (41.7%), whereas that in Zenya village, was found to be the lowest (30.84%). There was no obvious difference in the coefficient of variation of leaf shape between 19.57% and 25.66%. The largest variation coefficient of leaf tooth depth of the ancient tea germplasm population in Zenya village was 57.77%, and the smallest was 34.54% of the ancient tea germplasm population in Landong village. The largest variation coefficient of leaf color was 55.8% of the ancient tea germplasm in Yangmeng village, and the smallest was 34.81% of the ancient tea germplasm in Landong village. The observed variation coefficients of leaf vein logarithm, leaf teeth logarithm, leaf length, and leaf width were small. The variation coefficient of tree height was the highest in the ancient tea population in Zenya village. Among the two shrub groups of ancient tea germplasm, the variation coefficient range of Guqi village was the widest and ranged from 13.05% to 55.36%.

It can be observed that the average genetic diversity index of the arboreal ancient tea germplasm populations is relatively high (Table S6). The genetic diversity index of leaf tip was the highest in Yangmeng village (1.0889) and the lowest in Landong village (0.6089). The highest genetic diversity of leaf shape and leaf size was observed in the ancient tea germplasm population of Landong village, and the values were 1.0209 and 0.7653, respectively. The highest diversity index of leaf tooth depth was observed for the ancient tea germplasm group in Yangmeng village. The highest genetic diversity index of leaf color was observed to be 1.0067 in Zenya village. The highest genetic diversity index of leaf veins was 2.6921 in Zenya village, and the lowest value, i.e., 1.8867, was observed for Yangmeng village. The highest genetic diversity index of leaf tooth logarithmic was observed for the ancient tea germplasm in Landong village, and the lowest was noted for Yangmeng village. The highest genetic diversity index of tree height belongs to the ancient tea germplasm in Landong village, which is 3.3431. The genetic diversity index of leaf length ranged from 2.3026 to 3.7226, with the highest diversity. The genetic diversity index of leaf width is the highest (3.3193) in Landong village. Further, the lowest was the ancient tea group in Yangmeng village. The genetic diversity index of leaf area was reported to be the highest in Zenya village, as 3.9890, followed by Landong village (arbor) and Yangmeng village. Between the two shrub ancient tea germplasm groups, the genetic diversity indices of leaf size, leaf color, leaf tooth logarithm, tree height, leaf length, leaf width, and leaf area in Guqi village were higher than those of the

shrub ancient tea germplasm group in Landong village. This indicates that the genetic diversity index of the ancient tea germplasm in Guqi village is high.

Principal component analysis of phenotypic characters

In order to clarify the role of each character in the phenotypic diversity of ancient tea germplasm, 13 phenotypic characters were subjected to principal component analysis (Upadyayula et al. 2006). The contribution rate of the first six components reached 85.48%, indicating that they can reflect the features of 13 phenotypic characters (Table S7) (Tehrim et al. 2012). The correlation analysis between the original characteristic variables and the six principal components (Table 2) found that PC1 mainly represented the leaf size, leaf length, leaf width, and leaf area. Its principal component can be defined as the size factor of the leaf. PC2 mainly represents leaf shape and leaf shape index, while PC4 mainly represents leaf tip and leaf vein logarithm, both of which are related to the leaf shape. These two principal components are defined as the leaf shape factor. PC3 mainly represents the depth and logarithm of leaf teeth, which are defined as the leaf teeth factors. PC5 mainly represented leaf quality and was defined as a leaf quality factor. In a similar way, PC6 was defined as the vein factor on behalf of the leaf vein logarithm.

The principal component analysis of ancient tea germplasm (Table S8) indicated that the high positive value of PC1 reflects that the leaf size of ancient tea germplasm resources was dominated by middle and large leaves. LDT8, ZYT59, ZYT62, ZYT67, ZYT69, ZYT70, ZYT71, ZYT73, ZYT74, ZYT76, ZYT81, ZYT82, ZYT84, ZYT85, ZYT86, ZYT87, ZYT90, ZYT91, ZYT94, ZYT95, ZYT97, ZYT98, YMT99, and YMT101 belong to this category. On the other hand, a lower negative value of PC1 indicates that the leaves of this ancient tea germplasm are smaller. High positive values of PC2 indicated that the leaf shape of these materials tended to belong to elliptic or lanceolate. LDT14, LDT20, LDT33, LDT35, LDT36, LDT54, ZYT55, ZYT56, ZYT57, ZYT58, ZYT61, and ZYT63 can be classified into this category. A low negative value of PC2 indicates that the leaf shape of these materials tends to be round, and LDT21, LDT22, ZYT71, GQB121, GQB123, and GQB124 belong to this category. A higher positive PC3 value indicates that this kind of material has deeper leaf teeth and more logarithm of leaf teeth, ZYT69, YMT100, YMT101, GQB109, GQB110, GQB111, GQB112, GQB113, and GQB115 belong to this kind of material. A lower negative PC3 value indicated that the leaf teeth of this kind of material are shallow, and the logarithm of leaf teeth has a lower value. The high positive value of PC4 indicated that there were many vein logarithms of this material, and the leaf tip is obtuse, and LDT1, LDT2, LDT3, LDT4, LDT14, LDT35, YMT100, YMT101, YMT102, LDB139, and LDB143 can be assigned to this category. A lower PC4 negative value reflects that the leaf tip of these materials is sharp, and the logarithm of the leaf vein is less. The high positive values of PC5 indicated that the leave of such materials are relatively hard, such as LDT10, LDT16, LDT46, LDT47, ZYT58, ZYT64, ZYT66, ZYT68, ZYT82, ZYT83, ZYT96, YMT99, LDB134, LDB136, LDB137, etc. A low negative value of PC5 indicates that the leaves of these materials are soft, LDT5, LDT8, LDT9, LDT11, LDT12, LDT20, LDT29, LDT43, LDT44, LDT45, LDT48 belong to this category. A high positive value of PC6 indicates that this kind of material has more leaf vein logarithms, and LDT2, LDT5, LDT7, LDT9, LDT12, LDT20, LDT33 are in this category. A low negative value of PC6

indicates that the leaf vein logarithms are less, including LDT21, LDT29, LDT40, LDT43, ZYT63, and LDT52.

Cluster analysis of phenotypic characters

The cluster analysis with a genetic distance of 10, divided the 145 experimental samples used in this study into 4 categories. The first class comprised 31 ancient tea germplasm, which included ancient tea germplasm from trees in Landong village and ancient tea germplasm from all shrubs. The second class consisted of a total of 38 ancient tea germplasm, composed of arboreal ancient tea germplasm in Landong village and all ancient tea germplasm in Yangmeng village. Ancient tea germplasm of a total of 22 shrubs from the Guqi village formed the third class, whereas the fourth class comprised of 54 arboreal ancient tea germplasm in Zenya village.

Genetic diversity analysis of germplasm resources of ancient tea germplasm

Three sets of duplicate data were analyzed by POPGENE 32 (Table S9 and Table S10) (Liu et al. 2016). The results showed that the genetic consistency between 145 samples of ancient tea germplasm was 0.5765–0.9529. The genetic consistency of arbor lies between 0.5882–0.9529, and the largest one is for the ancient tea germplasm numbered 25 and 37. It indicates that the genetic background is similar, and the similarity is high. The lowest genetic consistency was numbered 48 and 102, indicating their genetic distance is relatively far. The genetic consistency of ancient tea germplasm of shrubs ranges from 0.6118 to 0.9294. The trees numbered 124 and 126 have the highest genetic consistency, indicating that their genetic backgrounds are quite similar. The two shrubs numbered 113 and 132 have the lowest genetic consistency, indicating that their genetic distance is relatively far. The genetic distance between trees and shrubs consistency is 0.5765–0.9176. The maximum genetic consistency was observed between the arbor numbered 105, and the shrub numbered 108. The minimum genetic consistency was noted between the arbor no. 48 and the shrub no. 127, indicating that the genetic distance between these two ancient tea germplasms was relatively farther. POPGENE 32 software was used to analyze the genetic diversity index (Table 3) of the 145 ancient tea resources. The results showed that the average observed allele number (N_a) was 2.0000, the average effective allele number (N_e) was 1.3984, the average Nei's genetic diversity index (H) was 0.2584, and the average Shannon information diversity index (I) was 0.4119. The results indicated that the genetic diversity among the 145 ancient tea germplasm materials is relatively high.

Genetic diversity parameter analysis of ancient tea germplasm at different sampling points

Comparing the same genetic parameters between the samples collected at different locations (Table S11) enabled the researchers to understand the differences between them. Compared with the arboreal ancient tea germplasm population in Landong village, the allele number (N_a) of the arboreal ancient tea germplasm population in Zenya village increased by 2.21%, the N_a of the shrubby ancient tea germplasm population in Guqi village decreased by 2.65%, and the N_a of the shrubby ancient tea germplasm population in Landong village decreased by 7.52%. In contrast to the Shannon information index (I) of the

ancient arboreal tea group in Landong village, the value of *I* for the ancient arboreal tea group in Zenya village was increased by 4.02%. The largest decrease was noted as 8.46% in the arboreal ancient tea group in Yangmeng village. On the other hand, in contrast to the polymorphism points of the arboreal ancient tea germplasm population in Landong village, in addition to the 3.80% increase in Zenya village, the others were decreased by 27.85%, 6.33%, and 16.46%, respectively. Among them, the largest decrease occurred in the ancient tea germplasm group in Yangmeng village. In contrast to with the percentage of polymorphism sites (PPB) in the arboreal ancient tea population in Landong village, the PPB in the arboreal ancient tea population in Zenya village increased by 4.61%, the PPB in the arboreal ancient tea population in Yangmeng Village decreased by 27.53%, the PPB in the shrubby ancient tea population in Guqi Village decreased by 5.53%, and the PPB in the shrubby ancient tea population in Landong Village decreased by 15.71%. Overall, the genetic diversity of arboreal ancient tea germplasm was richer than that of shrubby ancient tea germplasm.

UPGMA cluster analysis

Cluster analysis of 145 ancient tea germplasm was carried out with NTSYS-PC 2.1 ([Rout et al. 2009](#)), and the corresponding UPGMA tree diagram (Figure 2) was obtained. It was found that a genetic similarity coefficient of 0.734 divided the 145 ancient tea germplasm into four categories. The first category included 44 ancient tea germplasm of Landong village. The second category comprised 54 ancient tea germplasm in Zenya village, and the third category had nine ancient tea germplasm in Yangmeng village. The fourth category was further subdivided into two subclasses. The first subclass included 22 parts of the shrub ancient tea germplasm of Guqi village. The second subclass comprised 15 parts of the shrub ancient tea germplasm of Landong village as well as 1 part of the shrub ancient tea germplasm of Yangmeng village. The results from the cluster analysis suggest that the shrub ancient tea germplasm and the arboreal ancient tea germplasm have a greater genetic distance and far relationship, which conforms to the phenotypic clustering results. Moreover, the clustering results are closely related to the distribution region. Hence all ancient tea germplasm collected from a specific region were grouped under one category. Thus, it can be said that SSR markers could be used to classify and identify the germplasm resources of ancient tea germplasm and study the relationship between them.

Discussion

There is an obvious connection between different phenotypic characters and genetic variation ([Rode and Morrow 2010](#)). Abundant genetic variation has led to the abundant phenotypic characters in ancient tea germplasm, found in Sandu of Guizhou province ([Li-Jie et al. 2019](#)). In a study conducted by Niu Suzhen et al. (2012), which analyzed the diversity in 10 phenotypic characters of 144 ancient tea germplasm resources in Guizhou, it was reported that in addition to tree appearance, the variation coefficients of the other 9 phenotypic characters were all above 35%. Further, the study reported that the diversity indices of all the 10 phenotypic characters were above 0.85. The variation coefficients and diversity indices computed in this study were found to be lower than those reported by Niu Suzhen et al. (2012). The latter's sample collection of ancient tea germplasm from 32 ancient tea germplasm resources distributed

in Guizhou, could account for the difference in the diversity indices and variation coefficients between the two studies.

This study, which also focuses on the phenotypic diversity of ancient tea germplasm reported higher phenotypic diversity in comparison to previous studies conducted by Huang Haitao et al. (2019) and Xie Wengang et al. (2019). Both these studies analyzed 30 phenotypic characters; however, Huang Haitao et al. (2019) sampled 100 Longjing tea plants from 10 natural villages in Xihu District, Hangzhou, while Xie Wengang et al. (2019) sampled 109 middle and small-leaf tea plants in Sichuan. The reasons for reporting a lower phenotypic diversity than this study could be viz., i) The tea germplasm sampled in this study might have experienced a long time of natural selection, which might have led to an accumulation of genetic variation, resulting in richer phenotypic characteristics (Rossetto et al. 1999), and ii) the environment in a particular region might impact the phenotypic characteristics of the tea germplasm (Weifeng et al. 2018).

Previous studies indicate that ancient tea germplasm in 26 county's Guizhou province has rich genetic diversity (Niu et al. 2013). This genetic diversity observed after analysis of 145 ancient tea germplasm in Sandu County was lower than that reported in the literature. This indicates that gene exchange in ancient tea germplasm is lower in comparison to an entire province. Further, the UPGMA clustering tree separates the arbors and shrubs. The clustering results connect the clusters and region of collected samples, which in turn indicates the stability and aggregation of genes of ancient tea germplasm in a specific region (K. et al. 2003). Being surrounded by mountains, geographical isolation, and traffic inconvenience could block gene exchange between resources of ancient tea germplasm located in different regions. This could ultimately lead to the phenomenon of genetic stability and aggregation in a particular region

Conclusion

This study used phenotypes and SSR markers technology to analyze the genetic diversity of ancient tea germplasm in Sandu Aquatic Autonomous County. The results revealed that the genetic diversity of the arbor is higher than that of the shrub. Further, considerable genetic stability and aggregation among the ancient tea groups in the studied area, along with a rich genetic diversity within these groups, have been observed. Thus, this study not only provides a theoretical basis for the protection but also promotes further research of ancient tea germplasm.

Declarations

Acknowledgements: The authors would like to acknowledge Chen Xianxiang and Pan Xuanhua, teachers from Sandu County Tea Office, for leading us to collect samples of ancient tea germplasm.

Declarations

Funding: This study was funded by Key Laboratory of Mountain Plant Resources Protection and Planting Innovation, Ministry of Education (Guizhou University), Open Project funded by independent Project

"Survey, Identification and Utilization of Ancient tea germplasm Planting Resources in Sandu"; Science and Technology Support Program (Agriculture) of Guizhou, China. QianKeHeZhiCheng No. [2020]1Y001 .

Authors' contributions: ZYC, ZDG and DX conceived and planned the experiments. WRY and LQ carried out the experiments. WRY wrote the manuscript, all authors edited and approved the manuscript.

Data availability statements: The data that support the findings of this study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest: No competing or conflict of interests declared by all the authors.

Research involving human and animal rights: No human participants or animals were involved in this research.

References

- Bajracharya J, Steele KA, Jarvis DI, Sthapit BR, Witcombe JR (2006) Rice landrace diversity in Nepal: Variability of agro-phenotypic traits and SSR markers in landraces from a high-altitude site. *Field Crops Research* 95:327-335
- Benedetti S et al. (2007) Phenotypic clustering of lamin A/C mutations in neuromuscular patients. *Neurology* 69:1285-1292
- Danhui Q (2013) Role of ancient tea gardens in biodiversity conservation revealed.
- Gomes S, Breia R, Carvalho T, Carnide V, Martins-Lopes P (2018) Microsatellite High-Resolution Melting (SSR-HRM) to Track Olive Genotypes: From Field to Olive Oil. *Journal of food science* 83:págs. 2415-2423
- Huang H T, Zheng X X, MAO Y X, Li H L, Yu J Z (2019) Analysis of phenotypic traits and diversity of West Lake Longjing Tea plant resources. *Zhejiang Agricultural Sciences* 60(10): 1793-1797 (**In Chinese**)
- Inês L, Teresa R, António R, Adelaide A (2015) Incidence and Diversity of Antimicrobial Multidrug Resistance Profiles of Uropathogenic Bacteria. *Biomed Research International* 2015:354084
- Jiao J et al. (2018) Species identification of polygonati rhizoma in China by both phenotypic and molecular marker methods. *Comptes Rendus Biologies*:S1631069118300027
- K, NkongoloL., Nsapato (2003) Genetic diversity in *Sorghum bicolor*(L.) Moench accessions from different ecogeographical regions in Malawi assessed with RAPDs. *Genetic Resources & Crop Evolution*
- Karthigeyan S, Rajkumar S, Sharma RK, Gulati A, Sud RK, Ahuja PS (2008) High Level of Genetic Diversity Among the Selected Accessions of Tea (*Camellia sinensis*) from Abandoned Tea Gardens in Western

Himalaya. Biochemical Genetics 46:810

Kim YD, Jeong MJ, Song HJ, Kim JC, Choi MS (2012) Phenotypic characters and genetic relationship between catechins-rich and -poor tea tree (*Camellia sinensis* L.) lines. Forest Science and Technology 8

Koebner RMD, Powell W, Donini P (2001) Contributions of DNA Molecular Marker Technologies to the Genetics and Breeding of Wheat and Barley. John Wiley & Sons, Inc.

Li-Jie C et al. (2019) SNP Analysis of the Genetic Evolution of Ancient *Camellia sinensis* Trees from Huaxi, Guiyang. Journal of Southwest University(Natural Science Edition)

Lin Y, Yu W, Zhou L, Fan X, Ye N (2020) Genetic diversity of oolong tea (*Camellia sinensis*) germplasms based on the nanofluidic array of single-nucleotide polymorphism (SNP) markers. Tree Genetics & Genomes 16

Liu BY, You-Yong LI, Tang YC, Wang LY, Wang PS (2010) Assessment of Genetic Diversity and Relationship of Tea Germplasm in Yunnan as Revealed by ISSR Markers. Acta Agronomica Sinica 36:391-400

Liu Y, Zhang P, Zhang R, Song M, Hou J (2016) Analysis on Genetic Diversity of *Radix Astragali* by ISSR Markers. Advances in Bioscience and Biotechnology 07:381-391

Lu FH et al. (2011) Transcriptome analysis and SNP/SSR marker information of red pepper variety YCM334 and Taaan. Scientia Horticulturae 129:38-45

Lu X et al. (2018) Identification of high-efficiency SSR markers for assessing watermelon genetic purity. Journal of Genetics 97:1295-1306

M S, N CK, S A, R RK (2016) Molecular characterization of tea mosquito bug from tea growing regions of India. Mitochondrial DNA Part A, DNA mapping, sequencing, and analysis 27

Mujaju C, Werlemark G, Garkava-Gustavsson L, Nybom H (2011) HIGH LEVELS OF RAPD AND SSR MARKER DIVERSITY IN LANDRACES OF WATERMELON (*CITRULLUS LANATUS*) IN SOUTHERN AFRICA. Acta Horticulturae 918:291-295

Niu Suzhen, LIU Yuqian, Yang Jin-biao, Fan Weiguo(2012) Analysis of EST-SSR genetic diversity of guizhou tea tree species. Acta Agriculturae Zhejiang 24(05): 836-841 **(In Chinese)**

Niu S, Song Q, Fan W (2013) Analysis on site soil nutrients of ancient tea trees in the Southwest State of Guizhou Province. Acta Agriculturae zhejiangensis 25:0-1353

Niu S et al. (2019) Genetic diversity, linkage disequilibrium, and population structure analysis of the tea plant (*Camellia sinensis*) from an origin center, Guizhou plateau, using genome-wide SNPs developed by genotyping-by-sequencing. BMC Plant Biology 19

- Raza Q, Riaz A, Saher H, Bibi A, Sabar M (2020) Grain Fe and Zn contents linked SSR markers based genetic diversity in rice. *PLoS ONE* 15:e0239739
- Rode NO, Morrow EH (2010) An examination of genetic variation and selection on condition in *Drosophila melanogaster* males. *Entomologia Experimentalis Et Applicata* 131:167-177
- Rohlf F, Rohlf F, Rohlf F, Rohlf F, Rohlf FJ (2000) NTSYS-pc, Numerical Taxonomy and Multivariate Analysis System, version 2.1
- Rossetto M, Slade RW, Baverstock PR, Henry RJ, Lee LS (1999) Microsatellite variation and analysis of genetic structure in tea tree (*Melaleuca alternifolia* Myrtaceae). *Molecular Ecology* 8:633-643
- Rout GR, Senapati SK, Aparajita S, Palai SK (2009) Studies on genetic identification and genetic fidelity of cultivated banana using ISSR markers. *Plant Omics* 2:250-258
- Steel M (1992) The complexity of reconstructing trees from qualitative characters and subtrees. *Journal of Classification* 9:91-116
- Tamari F, Hinkley CS, Ramprashad N (2013) A Comparison of DNA Extraction Methods using *Petunia hybrida* Tissues. *Journal of Biomolecular Techniques* Jbt:113-118
- Tehrim S, Pervaiz ZH, Mirza MY, Rabbani MA, Masood, Shahid M (2012) Assessment of phenotypic variability in rice (*Oryza sativa* L.) cultivars using multivariate analysis. *Pakistan Journal of Botany* 44:999-1006
- Thungo Z, Shimelis H, Odindo A, Mashilo J, Shayanowako AIT (2020) Genetic relationship among selected heat and drought tolerant bread wheat genotypes using SSR markers, agronomic traits and grain protein content. *Acta Agriculturae Scandinavica, Section B - Soil & Plant Science* 70:594–604
- Tripathi SB, Negi MS Molecular markers as tools for characterization and improvement of tea germplasm. In: *Oceans*, 2006.
- Upadyayula N, Wassom J, Bohn MO, Rocheford TR (2006) Quantitative trait loci analysis of phenotypic traits and principal components of maize tassel inflorescence architecture. *Theoretical & Applied Genetics* 113:1395-1407
- Weifeng, Faxing, Chen, Kaiwun, Yeh, Jianjun (2018) ISSR Analysis of Genetic Diversity and Structure of Plum Varieties Cultivated in Southern China. *Biology*
- Xie W G, Li X S, Li W, Tang X (2019) Phenotypic genetic diversity of small leaf ancient tea germplasm in Sichuan province. *Acta Agriculturae Zhejiangensis* 31(09): 1405-1415 (**In Chinese**)
- Yeh F, Yang RC, Boyle T (2000) POPGENE - for the analysis of genetic variation among and within populations using co-dominant and dominant markers.

Tables

Table1 Names and sequence information of the 15 SSR markers used for genetic analysis

Number	Primer ID	Forward Primer Sequence(5'-3')	Reverse Primer Sequence(5'-3')	Repeat Moti	Expected Product Size(bp)
1	Cs1	ATGCCCTCTACATGCCTTTG	ATGACGTAGGCGGAAACAAC	(TC)6	179
2	Cs2	GAGCTGAGGCAGTCCATAGG	AAAAGGGAGAAAGACGTGGG	(GT)7	217
3	Cs3	GCCATCATAGACTGCTCGGT	GGTTGGCTTGACAAAAGGA	(TA)10	278
4	Cs4	CTCTTCCTCAGCCACCAAAG	TGAGGAGTTGTGGCAGAATG	(AAG)6	199
5	Cs5	GGGATCAGATATGGAGCCAA	CAGCAAATTCTTGAGGAAAC	(TTC)5	240
6	Cs6	GCTCATTGGCTTTGGCTTT	AATGCATTCCGTAAGCTTGG	(CTA)5	180
7	Cs7	GAACAGGCGAACAAGTAGGG	CGACCTCTGAGGCAATCTTC	(GAA)6	253
8	Cs8	GGCTTGCATGCCAGTTTATT	ACGTGGGGTTGGAAGACATA	(GT)7	195
9	Cs9	CCACGCTTTCTAACACCCAT	AAGGCTCCAAATGCTGAAGA	(TC)6	183
10	Cs10	CTGAGTCGGGACAGTTTGGT	CCCAAGAGGTGGAATAGCA	(CT)6	253
11	Cs11	CGTGCAATTGAGAATGCTGT	GGGTCGCTGTCTCTACTTGG	(TC)7	222
12	Cs12	AACCATGCAGCAAGACACTG	CCCGTAGGAGGTGCATAAGA	(CAA)5	129
13	Cs13	AAAGTGGTCGGTGTCCAAAG	TAACAGGTTTCATCCCTGGC	(GAT)5	235
14	Cs14	CTTTTGGCCATTGTCAAGGT	CAGACCTATCGAAAACCCGA	(CAC)6	242
15	Cs15	CCCACTCCTAAACTCACCCA	AGCCATCACATTGTCCAACA	(CAC)5	254

Table 2 Principal component scores of each phenotypic characters measured of 145 samples of ancient tea germplasm

Variable	Factor loading					
	PC1 (Leaf size)	PC2 (Leaf shape)	PC3 (Leaf teeth)	PC4 (Leaf shape)	PC5 (Leaf texture)	PC6 (Veins logarithmic)
Height of tree[m]	0.555	0.208	-0.495	-0.118	0.222	0.096
Tree shape	0.457	0.126	-0.686	0.357	0.057	0.186
Leaf texture	0.172	-0.042	0.200	0.174	0.890	-0.270
Leaf opex	-0.123	-0.356	0.037	0.591	-0.307	-0.450
Leaf shape	0.339	0.799	0.343	0.125	-0.105	-0.135
Leaf size	0.871	-0.209	0.109	-0.027	-0.093	-0.190
Leaf tooth depth	-0.143	-0.301	0.568	0.355	0.091	0.058
Length of leaf[cm]	0.960	0.055	0.179	-0.065	-0.096	-0.067
Width of leaf[cm]	0.859	-0.467	-0.056	-0.078	-0.040	-0.024
Leaf index	0.291	0.841	0.371	-0.006	-0.064	-0.057
Leaf area	0.946	-0.231	0.081	-0.086	-0.076	-0.064
Veins logarithmic	0.351	0.031	0.197	0.632	0.011	0.580
Leaves teeth logarithmic	0.087	-0.351	0.584	-0.376	0.032	0.345

Table 3 Genetic diversity index estimated for 145samples of ancient tea germplasm using 15 SSR markers

Sample number	Na	Ne	H	I
145	2.0000±0	1.3984±0.0123	0.2584±0.0069	0.4119±0.0106

Na-Number of different alleles; Ne-number of effective alleles = $1/(\sum p_i^2)$; H-Nei's genetic diversity index = $1-(\sum p_i^2)/N$; I-Shannon's Information Index = $-1 \times \sum (p_i \times \ln(p_i))$; Where P_i is the frequency of the i th allele for the population & $\sum p_i^2$ is the sum of the squared population allele frequencies. N-Number of loci in the population.

Figures

Figure 1

Cluster diagram of phenotypic traits of ancient tea germplasm showing clustering relationship of 145 samples of ancient tea germplasm

Figure 2

Classification and identification of 145 samples of ancient tea germplasm showing genetic relationship based on 15 SSR markers

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [formula.docx](#)
- [SupplementaryMaterial.docx](#)