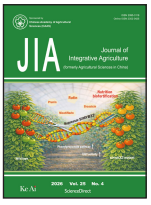
Journal of Integrative Agriculture 2026, 25(4): 1519–1530

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Identification of *CsAK* as a critical caffeine-related upstream gene in tea accessions through genome-wide association study

Kaixin Rao[1], Yuting Ouyang[1], Yanjun Chen[1], Xiaojing Wang[1, 2], Ting Liu[3], Qinfei Song[1], Shaojuan Zhang[1, 2], Biao Xiong[1], Suzhen Niu[1, 2#]

[1] *College of Tea Science/Institute of Agro-Bioengineering, Guizhou University, Guiyang 550025, China*
[2] *The Key Laboratory of Plant Resources Conservation and Germplasm Innovation in the Mountainous Region (Ministry of Education), Guiyang 550025, China*
[3] *Zhengzhou Foreign Language School, Zhengzhou 450006, China*

Highlights

- GWAS identified 19 SNPs and a key gene (*CsAK*) that are significantly associated with caffeine (CAF) content in tea accessions.
- The *CsAK* gene is involved in CAF biosynthesis by catalyzing L-aspartate to generate the key intermediate-L-methionine.

Abstract

Caffeine (CAF), a primary flavor component in tea, is one of the main reasons for the popularity of tea beverages. As an important secondary metabolite in tea plants, the CAF content varied greatly among different tea accessions. However, the genetic mechanisms underlying the CAF biosynthesis were still unclear. In this study, we performed a genome-wide association study (GWAS) on 359 tea accessions in the Guizhou Plateau to identify genetic variation associated with CAF content. A total of 19 significant single nucleotide polymorphisms (SNPs) and key gene (*CsAK*) involved in CAF biosynthesis were identified. Subcellular localization revealed that the CsAK-GFP fusion protein was located on the cell membrane. Antisense oligodeoxynucleotide (AsODN) targeting the *CsAK* gene to the buds and leaves revealed that the expression levels of the *CsAK* gene were significantly reduced, and the corresponding CAF content was also decreased in AsODN-treated tea plants. Overexpression of the *CsAK* gene in eukaryotic cells resulted in the accumulation of key intermediate product (L-methionine) during CAF biosynthesis process. These findings offered a theoretical foundation for future tea breeding programs aimed at cultivating excellent germplasm with high or low levels of CAF.

Keywords: tea accessions, caffeine (CAF), single nucleotide polymorphism (SNP), genome-wide association study (GWAS), Guizhou Plateau

1. Introduction

Tea plant, as one of the important commercial crops, is widely cultivated in China and around the world (Xia *et al.* 2020a). The plentiful flavor and health benefits provided by secondary metabolites in tea are the reasons for the popularity of tea beverages (Jin *et al.* 2018; Niu *et al.* 2020; Fang *et al.* 2021). Caffeine (CAF; 1,3,7-trimethylxanthine), a methylxanthine alkaloid compound accounting for 2–5% of dry matter content, is an important characteristic secondary metabolite in tea plants (Zhang S R *et al.* 2022; Reddy *et al.* 2024). CAF is mainly distributed in young tissues with the highest content in leaf buds, which can generate complexes with theaflavins and thearubigins through hydrogen bonding to decrease the bitterness and increase briskness of tea (Xia *et al.* 2020b; Kim *et al.* 2021; Ahammed and Li 2022). The beneficial effects of CAF intake on health have been reported in many studies, including refreshing, aiding digestion, anti-diabetic properties and the prevention of cardiovascular disorders (Van Dam and Hu 2022; Yao *et al.* 2023).

The CAF biosynthesis is primarily mediated through the adenine nucleotide pathway, contains sequential three-

Received 8 January 2025; Received in revised form 28 April 2025; Accepted 20 May 2025; Available online 30 June 2025
Kaixin Rao, E-mail: 1778424029@qq.com; #Correspondence Suzhen Niu, E-mail: niusuzhen@163.com

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2025.06.026

step methylation of xanthine derivatives at positions 7-N, 3-N, and 1-N (Fig. 1) (Misako and Kouichi 2004; Teng *et al.* 2020; Zhang Y R *et al.* 2022; Jin *et al.* 2023). These steps are catalyzed by N-methyltransferases (NMTs) that utilize S-adenosylmethionine (SAM) as methyl donors (Koshiishi *et al.* 2001; Jin *et al.* 2016; Xia *et al.* 2017; Gong *et al.* 2020; Zhou *et al.* 2020). Methionine is the immediate precursor of SAM, aspartate kinase (AK) is involved in generating the biosynthesis of methionine, threonine, isoleucine, and lysine (Ravanel *et al.* 1998; Jander and Joshi 2010; Wang *et al.* 2018). For instance, it has been discovered that aspartate kinase is the initial enzyme in the pathway leading to the production of methionine, threonine, and isoleucine, and five genes in *Arabidopsis* have been shown to encode aspartate kinase activity (Jander and Joshi 2010). Overexpression of aspartate kinase and cystathionine γ-synthase in tobacco leaves resulted in enhanced accumulation of methionine and threonine (Hacham *et al.* 2008).

Many markers, including classical markers (morphological, cytological, and biochemical traits) and DNA/molecular markers (RFLP, RAPD, SSR, AFLP and SNPs), were applied in tracking the tea population's genetic diversity and identifying the candidate genes' involvement in regulating the trait of interest. SNPs, as third-generation DNA molecular markers, were successfully used to investigate genetic variation among different species (Helyar *et al.* 2011; Grover and Sharma 2016; Tang *et al.* 2016). Genome-wide association study (GWAS) represents an effective approach for mining genetic loci and screening functional genes related to traits of interest *via* linking phenotype and genotype (Uffelmann *et al.* 2021; Wang *et al.* 2023). For example, GWAS performed on 248 maize varieties and 202 rice varieties identified 29 grain quality-related genes and two genes involved in regulating kernel traits, respectively (Zheng *et al.* 2021; Yoshida *et al.* 2023). Shi *et al.* (2024) performed GWAS on 202 *Brassica napus* accessions and totally identified seven candidate genes related to crude fiber traits. GWAS has been widely used for detecting genetic variations among different tea accessions and identifying key genes involved in regulating the biosynthesis of characteristic secondary metabolites. Zhang *et al.* (2020) conducted GWAS on 176 tea accessions and screened out three genes associated with catechin content. Chen Y J *et al.* (2023) used GWAS to obtain six genes associated with mature leaf size, texture, color and shape respectively in 338 tea plants. Deng *et al.* (2025) carried out GWAS in 299 tea accessions, identified a salicylic acid (SA)-related gene.

In this study, we identified and analyzed high-quality SNPs of 359 tea accessions collected from Guizhou Plateau using the genotyping-by-sequencing (GBS) technology. GWAS analysis was used to investigate the relationships between dynamic changes of CAF content and genetic variants. AsODN targeting candidate gene and overexpression of target gene in eukaryotic cell assays were applied to verify the roles of candidate gene in the CAF biosynthesis. The results of this study may provide a reference for breeding and quality improvement of tea cultivars with high or low CAF.

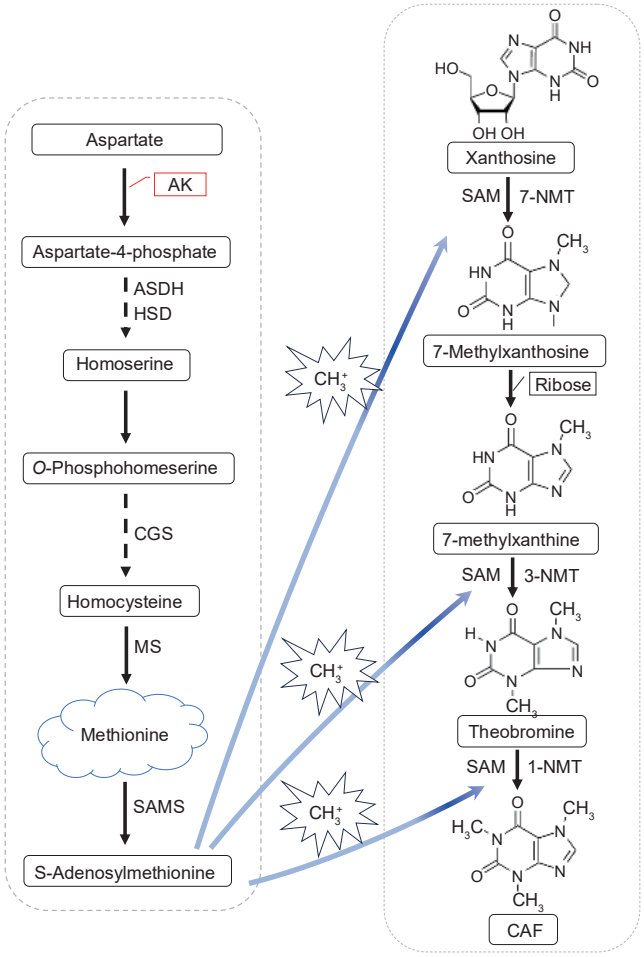


Fig. 1 Biosynthesis of methionine and biosynthetic pathway of caffeine from xanthosine in tea plants. AK, aspartate kinase; ASDH, aspartate semialdehyde dehydrogenase; HSD, homoserine dehydrogenase; CGS, cystathionine gamma-synthase; MS, methionine synthase; SAMS, S-adenosylmethionine synthetase; 7-NMT, xanthosine methyltransferase; 3-NMT, 7-methylxanthine methyltransferase (theobromine synthase, TS); 1-NMT, 3, 7-dimethylxanthine methyltransferase (CAF synthase, CS).

2. Materials and methods

2.1. Plant materials

In this study, 359 tea accessions were collected from 32 counties in Guizhou Plateau (Appendices A and B). All accessions were cultivated in the same tea genetic resource nursery in Guizhou University, China, which was managed uniformly according to the conventional tea plantation water and fertilizer management methods. The samples (one bud and two leaves) were collected from different tea accessions in the spring of 2022 (Environment I) and 2023 (Environment II). The collected samples were frozen in liquid nitrogen. Statistical analysis was performed using SPSS software (v.25) (IBM Corp., Armonk, NY, USA). The generalized heritability (h_B^2) was estimated using the variance component method.

2.2. Identification of SNPs and genetic diversity

Illumina's HiSeq-X platform was applied for sequencing the isolated genomic DNA from 359 tea accessions retrieved from published data (Niu *et al.* 2019) (ID: CRA001438, https://ngdc.cncb.ac.cn/gsa/browse/CRA001438). After removing adapters

and low-quality reads, the clean data were mapped to the reference genome of tea variety 'Shuchazao' using BWA-MEM v. 0.7.10 (http://tpia.teaplants.cn/) (Xia *et al.* 2020a). SNP calling was carried out using the GATK v. 3.7.0 software, and the results were filtered using McKenna's research-proposed methodology (McKenna *et al.* 2010). For each inferred group, the following metrics were computed using Plink v.1.90: minor allele frequency (*MAF*), expected heterozygosity (*He*), inbreeding coefficient (*Fis*), and observed heterozygosity (*Ho*) (Zhao Z F *et al.* 2022). For each inferred group, nucleotide diversity (*Pi*) and Tajima's D were calculated using VCFtools (v.0.1.16) (Danecek *et al.* 2011).

2.3. Analysis of population structure

ADMIXTURE v.1.3.0 was applied to analyze the population structure of the 359 tea accessions using the best-fit cluster numbers (K) by calculating the minimum cross-validation (CV) error (Chen *et al.* 2017; Wang *et al.* 2022). Membership coefficient (0.8) was applied to distinguish the pure and admixed groups (Chen Y J *et al.* 2023). Principal component analysis (PCA) was performed using Tassel (v.5.2.72) software (Bradbury *et al.* 2007). Neighbor-joining method in MEGA (v.10.2.4) software was used to construct an unrooted phylogenetic tree with default parameters (Stecher *et al.* 2020). Linkage disequilibrium (LD) analysis for these SNP markers was performed with PopLDdecay-3.40 based on the squared correlation coefficient (r^2) of alleles (Zhang *et al.* 2019).

2.4. GWAS analysis

The SNPs associated with CAF content were identified by GWAS using TASSEL v.5.2.72. Three general models (GLM, MLM, and CMLM) were applied for GWAS analysis. In the GLMs models, either the population structure (Q) or principal component (P) was employed as fixed effects to mitigate potential false positives resulting from population structure mixing. In MLMs and CMLMs, the kinship (K) was used as a covariate to address unequal correlations between genotypes. The optimal model was determined by analyzing the Q-Q plots. A threshold of $-\log_{10}(P)\geq 4.0$ (Chen Y J *et al.* 2023; Deng *et al.* 2025) was selected to identify SNPs associated with CAF in tea accessions. These SNPs were subsequently mapped to regions within 200 kb of candidate genes. The Tea Plant Information Archive (TPIA, http://tpia.teaplants.cn/) and the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) databases were used to functionally annotate the candidate genes.

2.5. Expression patterns of candidate genes and subcellular localization

Using Trizol DNA-free kits (Sangon Biotech Co. Ltd., Shanghai, China) in accordance with the manufacturer's instructions, samples (one bud and two leaves) from nine tea accessions with varying CAF contents (from high to low) were utilized for RNA extraction. The extracted total RNA was reverse-transcribed into cDNA, and RT-qPCR was subsequently conducted using the ChamQ Universal SYBR qPCR Master Mix Kit (Vazyme Biotech Co. Ltd., Nanjing, China) with cDNA as the template. The expression levels of these genes were assessed using the $2^{-\Delta\Delta CT}$ technique, with the *GAPDH* gene serving as the reference gene (Appendix C) (Li *et al.* 2023). Each experiment was performed in triplicate with three technical replicates.

The coding sequences of *CsAK* gene were inserted into the pCAMBIA2300-GFP vector (Li *et al.* 2021). The *Agrobacterium tumefaciens* strain (GV3101) transformed with the binary vectors was transiently transformed into the leaf epidermis of *Nicotiana benthamiana*. The fluorescence signal in tobacco leaves was detected after three days of incubation using confocal laser scanning microscopy. All primers used and reaction system in this study can be found in Appendix D.

2.6. Gene suppression in tea accessions

Antisense oligodeoxynucleotides (AsODN) experiment: the Soligo software was used to design candidate sense oligonucleotides (sODN) and AsODN using the target gene as input sequence (Yu *et al.* 2021; Zhao M Y *et al.* 2022; Chen Y F *et al.* 2023). The sODNs and AsODNs of target genes were biosynthesized in Sangon Biotech Co. Ltd., Shanghai, China. Buds and two leaves from tea plants were immersed in 1 mL of 20 μmol L^{-1} AsODN solution and incubated for 24 h under 16 h light/8 h dark. The sODN was used as negative control. This experiment was replicated three times (Appendix E).

2.7. Overexpression of *CsAK* gene in eukaryotic cells

The coding sequences of *CsAK* gene were inserted into the pcDNA3.1 vector to construct recombinant plasmids. DNA plasmids were extracted *via* EndoFree Maxi Plasmid Kit (Tiangen Biotechnology Co. Ltd., Beijing, China). DMEM media supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin was used to cultivate 293T cells, which were then kept at 37°C in a humid environment with 5% CO_2. Cells were transfected with plasmids using PEI (Changsheng Biomedical Technology Co. Ltd., Shenzhen, China) following the previously described method. L-aspartate was subsequently added to the culture system, and the transfected cells were harvested at 12, 24, and 36 h after transfection, respectively. Untransfected cells were used as the control (Devitt *et al.* 2007; Jiang *et al.* 2021). Then, L-aspartate and L-methionine of the experimental and control groups were determined by S-433D Amino Acid Analyzer (Sykam, Germany). The inclusion of a minimum of three biological replicates was imperative for obtaining precise results.

2.8. Determination of the contents of CAF and theobromine in the leaf buds

The CAF and theobromine were determined using HPLC (Thermo Fisher Scientific, America) based on the previously described method with some modifications (Srdjenovic *et al.* 2008). About 200 mg dry weight samples were ground into fine powder and added to 5 mL 70% methanol preheated to 70°C for 10 min. After centrifugation for 10 min at 3,500 r min^{-1}, the supernatant from the first round of centrifugation was centrifuged again. For the CAF content, 300 μL supernatant was added to 2,700 μL 70% methanol and

thoroughly mixed, which was filtered through 0.45 μm sterile filters. HPLC-UV system: injection volume of 5 μL, SB-C18 column (250 mm×4.6 mm, 5 μm) and 278 nm of detection wavelength. Mobile phase: A (100% methanol), B (100% acetonitrile) and D (0.5% glacial acetic acid).

For the theobromine, about 100 mg dry weight samples were ground into fine powder and added to 3 mL water solution (0.2 g MgO, 70°C) preheated to 70°C for 10 min. After centrifugation for 10 min at 4,000 r min^{-1}, the supernatant was removed and added to 2 mL of 70°C pure water and centrifuged again. The extraction solutions were combined, diluted with deionized water to 5 mL. The mixture was thoroughly mixed and filtered through a 0.45 μm filter for further use. Selected SB-C18 column (250 mm×4.6 mm, 5 μm) and 271 nm of detection wavelength. Mobile phases: A (100% acetonitrile) and B (0.5% formic acid solution).

2.9. Determination of the contents of metabolites in the leaf buds

Metabolites (L-aspartate, L-homocysteine, L-methionine, S-adenosylmethionine, xanthosine and 7-methylxanthine) were detected using UPLC-MS/MS, referring to previous research (Zou *et al.* 2021). About 400 mg fresh weight samples were ground into fine powder, added to 2 mL of pre-cooled 50% DMSO water solution and shaken for 1 min. Ultrasonic extraction at 4°C for 30 min, and centrifuged at 13,000 r min^{-1} for 10 min. Kept the supernatant and repeated the above process again. Diluted the extraction solutions with deionized water to 4 mL and thoroughly mixed, which was filtered through a 0.22 μm sterile filter. UPLC-MS/MS system: ACQUITY Ultra-Performance-LC (Waters, America) and AB Sciex 4000 QTRAP LC-MS/MS (Kexiao, China). HSS T3 column: 2.1 mm×100 mm, 1.8 μm. Mobile phase: A (0.1% formic acid-water solution) and B (100% acetonitrile). MSMS conditions: ion source gas I, gas II, curtain gas, and collision gas were set at 40, 50, 35, and 7 psi, respectively; ion spray voltage (IS) of ESI positive and negative modes was set at +5,500 and –4,500 V, respectively; the source temperature was set at 450°C.

3. Results

3.1. Sequencing and SNP discovery

In this study, GBS of 359 tea accessions totally yielded 29,393,327 SNPs. After applying missing data thresholds, 105,370 high-quality SNPs with *MAF*>0.05 and max-missing<20% were retained (Appendices F and G), which were unevenly distributed throughout the 15 chromosomes (Appendices H and I). The 15 chromosomes contained the number of SNPs ranging from 4,348 to 9,564, with average number of 6,424 in each chromosome. Chromosomes 1 and 15 contained the largest (9,564 SNPs) and smallest (4,348 SNPs) number of high-quality SNPs, respectively (Appendix J). The highest and lowest SNPs density was detected on chromosome 8 (44 SNPs/Mb) and chromosome 5 (33 SNPs/Mb), respectively (Appendix J).

3.2. Analysis of population structure

To further reveal the relationship among 359 tea accessions, we performed the population structure analysis based on 105,370 high-quality SNPs using the Admixture software (Appendix K). The result indicated that 359 tea accessions were further classified into six inferred groups with a membership coefficient>0.8 when the CV error was smallest at K=5 (Fig. 2-A and B), including five pure subgroups (G-I, G-II, G-III, G-IV, and G-V) and one admixed subgroup (G-VI). G-I contained 91 individuals, of which 81.32% were *Camellia sinensis* accessions. G-II included 36 *C. tachangensis* individuals and one *C. gymnogyna* accession. G-III comprised 12 *C. sinensis*, six *C. tachangensis*, and one *C. gymnogyna*, and two unknown accessions. G-IV contained 25 individuals, of which 84% were *C. sinensis* accessions. G-V comprised four *C. gymnogyna* accessions. G-VI contained 181 tea accessions, including 45 *C. gymnogyna*, 59 *C. tachangensis*, 74 *C. sinensis* and three unknown individuals (Appendix L). PCA further supported this division, revealing some degree of overlap among the groups (Fig. 2-C). An unrooted phylogenetic tree also classified the accessions into six clades, which was aligned with the results of population structure analysis and PCA (Fig. 2-D).

3.3. Determination of CAF contents

The CAF contents of 359 tea accessions were evaluated under two environmental conditions (Appendix M). Our results showed that the range variation in CAF contents was 2.02–4.07% and 1.81–4.29% in environment I (EI) and environment II (EII), respectively (Fig. 3-A; Table 1). The variation coefficient of the CAF content under two environments was 14.59 and 15.48%, respectively. In both environments, frequency distribution of the CAF content in 359 tea accessions exhibits a normal distribution or approximate normal distribution (Fig. 3-B). We further compared and analyzed the distribution of CAF content in different inferred subgroups. Results revealed that the highest CAF content was detected in G-III subgroup, followed by G-VI, G-V, G-I, and G-IV, while the CAF content was the lowest in G-II subgroup (Fig. 2-E). The influence of genotype (G) and environment (E) over CAF was evaluated in 359 accessions during two environments by ANOVA. Results revealed that the h_B^2 of CAF was estimated to be 94.22%, and stable heritability above 90%, exhibiting high broad sense heritability (Wang *et al.* 2019) (Appendix N).

3.4. Genetic diversity of inferred subgroups

Genetic diversity reflects the balance between the emergence and disappearance of genetic variants (alleles). The *Pi*, *MAF*, *Ho*, *He*, and *Fis* of 359 tea populations were 0.2260, 0.1438, 0.0730, 0.2255 and 0.6893, respectively (Appendix O). The values of *Pi*, *MAF*, *Ho*, *He*, and *Fis* in the G-VI were significantly higher than those in other inferred subgroups, and *Pi*, *MAF*, *He*, and *Fis* values in G-IV subgroup were significantly lower than those in other inferred subgroups. The *Ho* value was the lowest in G-II subgroup (Appendix O). Previous studies indicated that when a population has a positive Tajima's D value, this population has undergone a population bottleneck and/or balancing selection (Pandey *et al.* 2021; He *et al.* 2023). The six inferred groups showed

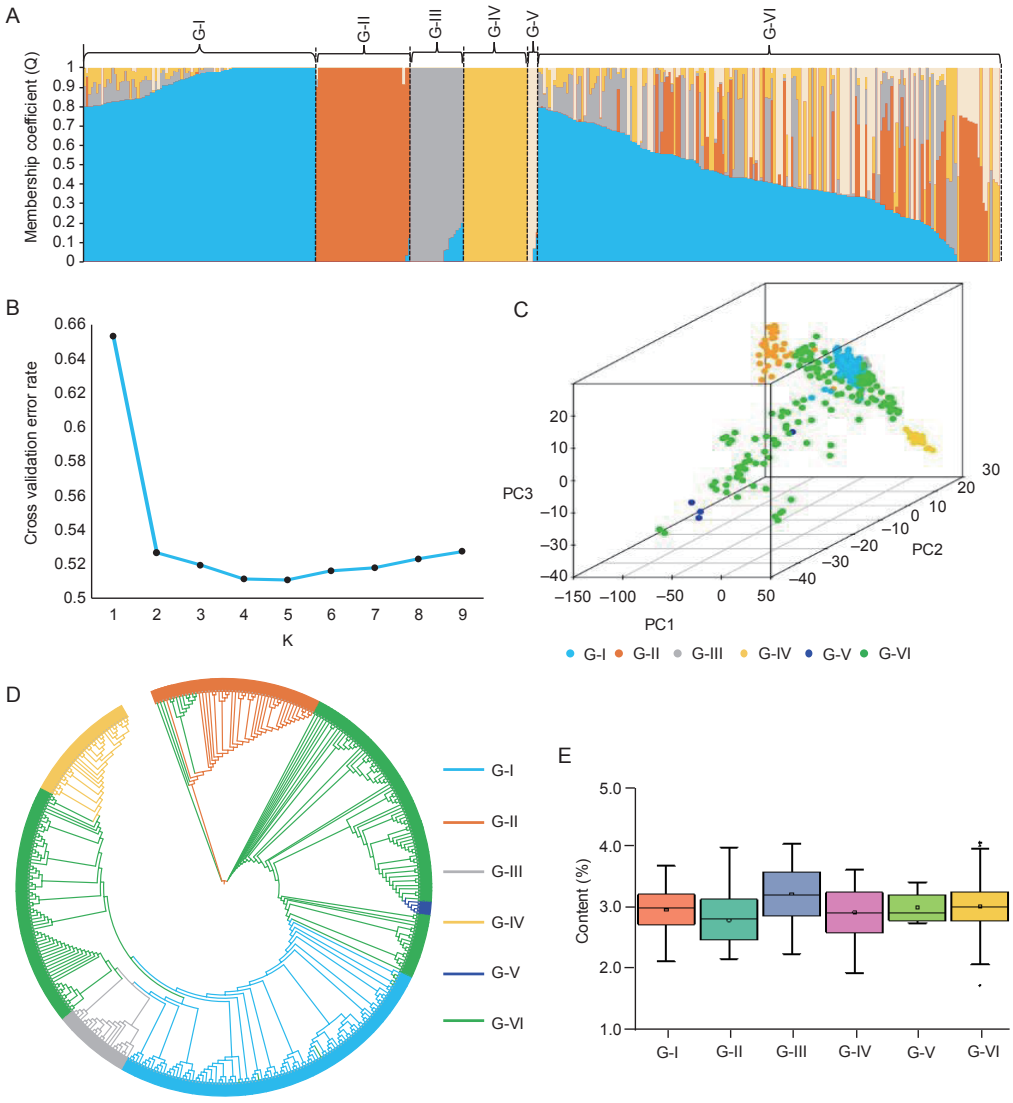


Fig. 2 Analyzing the population structure, PCA, and NJ tree establishment of the 359 tea accessions. A, separation of accessions into six inferred subgroups (G-I, G-II, G-III, G-IV, G-V, and G-VI) by population structure. B, the minimum CV error value of 0.51 at a K-value of five. C, the representation of individuality by each dot on the scatter plot of the first three principal components. D, the establishment of the NJ tree by 105,370 high-quality SNPs. E, box plot of caffeine (CAF) contents in six inferred groups.

positive Tajima's D values, indicating that all of them have experienced population bottlenecks and/or balancing selection (Appendix O).

3.5. GWAS for identifying SNPs, LD and candidate genes associated with CAF content

Based on Q-Q plot comparisons, the GLM-P model was identified as the most suitable for GWAS of CAF (Appendix P). A total of 19 high-quality SNPs significantly associated with CAF were detected, with $-\log_{10}(P) \geq 4.0$. The contribution rate of phenotypic variation (R^2) for the SNPs was from 6.27 to 10.27% (Fig. 4-A and B; Appendix Q). The genome-wide LD decay for the 359 accessions was measured at 3.5 kb, with r^2 declining to half decay distance (Fig. 4-C), indicating strong linkage disequilibrium between SNP markers within the 3.5 kb region, making these SNPs reliable for association analysis. Among these, two high-quality SNPs (P-83385711 and P-157358411) located in chromosomes 4 and 7 were remarkably associated with the CAF content. Two genes (CSS0003534 and CSS0028755) were identified as respectively involved in encoding aspartate kinase and conserved oligomeric Golgi complex subunit 7 by using three databases (NCBI, TPIA, and KEGG) (Appendix R).

Table 1 Statistical analysis of caffeine (CAF) in two environments

Trait	Environment	Min (%)	Max (%)	Mean±SD (%)	CV (%)
CAF	EI	2.02	4.07	2.97±0.43	14.59
CAF	EII	1.81	4.29	2.98±0.46	15.48

To further investigate the expression levels of *CsAK* (CSS0003534) gene and *CsCOG7* (CSS0028755) gene, we performed RT-qPCR in three groups of tea plants with high, medium, and low CAF concentrations (designated as H, M, and L). Results revealed that the relative expression level of the *CsAK* gene positively correlated with the

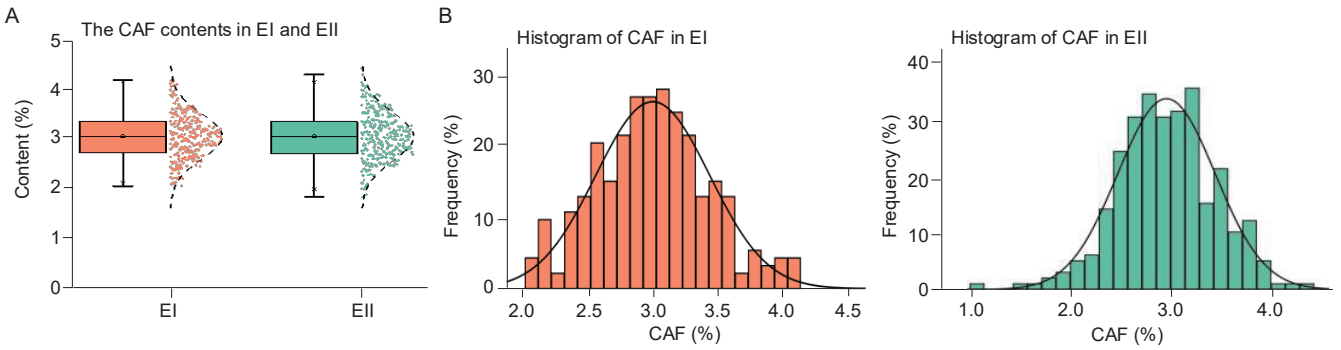


Fig. 3 The contents of caffeine (CAF) (A) and frequency distribution (B) in two environments (EI and EII).

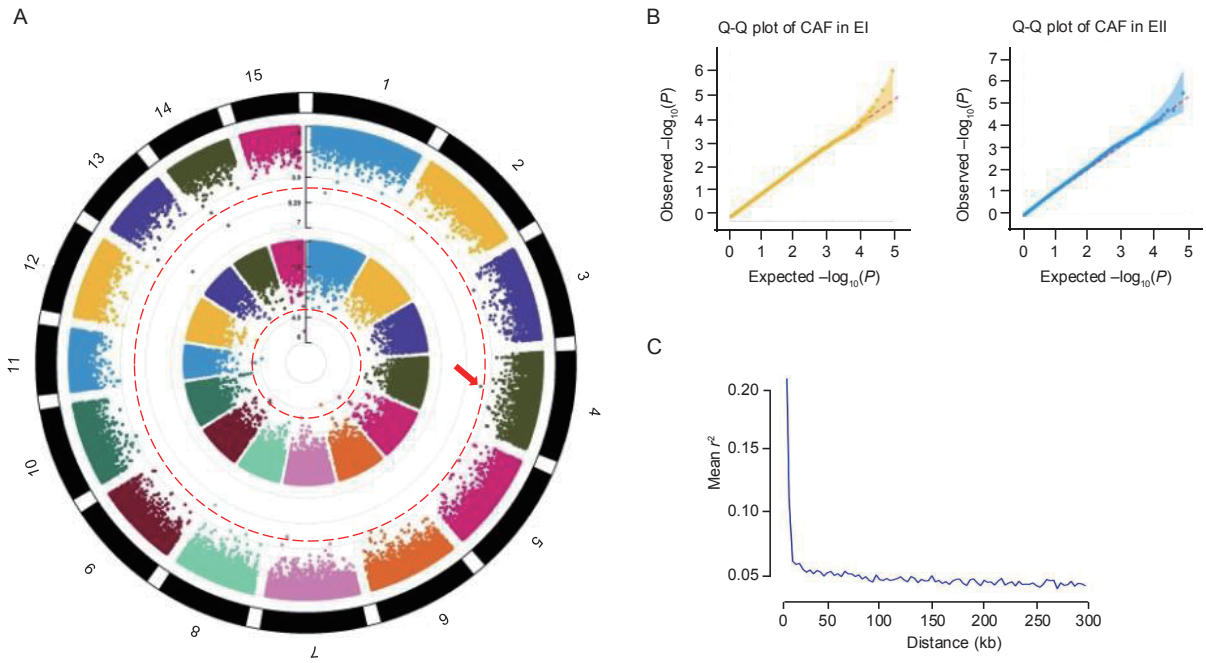


Fig. 4 GWAS for the phenotype of caffeine content and genotype from 359 tea accessions. A, Manhattan plots of CAF *via* GLM-P model in two environments (inner circle: EII; middle circle: EI; outer circle: chromosomes). B, Q-Q plots of two environments. C, LD decay analysis of 359 accessions. Red dashed line indicates the significance threshold.

dynamic changes of CAF in tea plant, demonstrating that the gene contributed to regulating the CAF biosynthesis (Fig. 5-A; Appendices S, T, and U). However, the expression patterns of *CsCOG7* gene were not related to the CAF content (Fig. 5-A). And the expression level of *CsAK* gene in tea plants with high CAF content was consistent with the transcriptome analysis of tea plants with high content of CAF, and *vice versa* (Appendix V).

3.6. Subcellular localization and suppression of *CsAK* gene by AsODN ways in tea accessions

Subcellular localization refers to the identification of where a protein is located within a cell, which provides insights into its potential functions. Transient expression of the CsAK-GFP fusion protein in tobacco leaves was conducted. Result indicated that the *CsAK* protein was mainly localized in the cell membrane (Fig. 6).

To further verify the function of *CsAK* gene in CAF biosynthesis in tea accessions, we applied AsODN targeting the *CsAK* gene to the buds and leaves of tea plants with high CAF content. The results revealed that the expression levels of the *CsAK* gene in the buds and leaves treated with AsODN were significantly reduced compared to the control group, and the corresponding CAF contents were dramatically decreased in AsODN-treated tea plants (Fig. 5-C). This result indicated that *CsAK* gene might be involved in CAF biosynthesis in tea accessions. KEGG analysis and gene function annotation revealed that the *CsAK* gene might be involved in the aspartate metabolic pathway, and aspartate kinase encoded by the *CsAK* gene was involved in generating the L-methionine (L-met) and S-adenosylmethionine (SAM) (Fig. 7-A). Previous studies revealed that SAM could provide the methyl donors for the CAF biosynthesis. Thus, we further investigated the main metabolites of the aspartate

metabolic pathway. Results showed that the content of L-aspartate increased, while the contents of L-homocysteine, L-methionine, and S-adenosylmethionine were significantly decreased in buds and leaves treated with AsODN (Fig. 7-B–E). Further analysis showed that the contents of xanthosine, 7-methylxanthine, theobromine, and CAF were significantly decreased in buds and leaves treated with AsODN (Figs. 5-C and 7-F–H), indicating that *CsAK* gene played positive roles in regulating the CAF biosynthesis in tea accessions.

3.7. Overexpression of *CsAK* in eukaryotic cells

Previous studies revealed that the CAF in tea plants was biosynthesized from xanthosine in a crucial pathway that has three methylation steps catalyzed by SAM-dependent N-methyltransferases (NMTs) with L-methionine as a substrate (Xia *et al*. 2017; Gong *et al*. 2020). To further investigate the roles of the *CsAK* gene in L-methionine accumulation, we performed overexpression of the *CsAK* gene in eukaryotic cell to determine the dynamic changes of L-methionine. The results showed that the L-aspartate content was significantly decreased in comparison with the control group (CK) after 12, 24, and 36 h of cultivation, respectively, while the content of downstream product, L-methionine, was increased (Fig. 8-B–D). This result indicated that the *CsAK* gene was involved in catalyzing L-aspartate to generate the key intermediate-L-methionine.

4. Discussion

Tea plant, originated in Southwest China, is an important economic crop with self-incompatibility. Due to the unique natural environment and long-term cross-pollination, the Guizhou Plateau is rich in tea germplasm accessions with high nutrient contents, including polyphenols, CAF, and theanine. These main secondary metabolites typically determine the component quality and flavor of tea (Qiao *et al.* 2019; Xia *et al.* 2020b; Fang *et al.* 2021). CAF, as the most abundant purine alkaloid in tea plant, is closely related to the bitterness of tea (Mohanpuria *et al.* 2010; Zhang S R

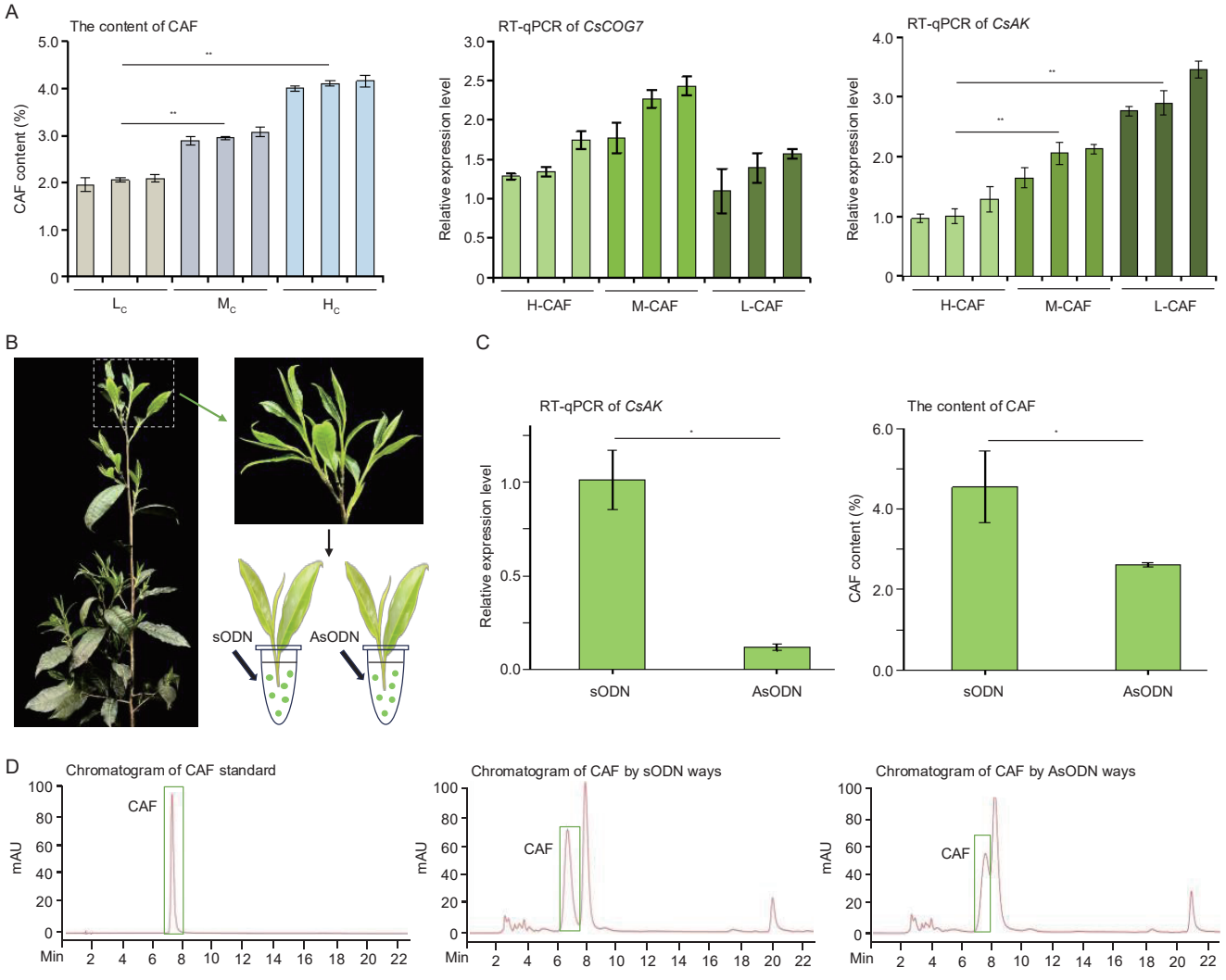


Fig. 5 Candidate genes on expression levels and effects of *CsAK* gene silencing on expression and caffeine accumulation in tea accessions. A, the content of CAF and the relative expression level of *CsCOG7* and *CsAK* genes. B, macro-photograph of the tea plant and morphological features of the bud-leaves. C, the relative expression levels and contents of CAF bud-leaves of treatment with *CsAK*-AsODN in *Camellia sinensis*. D, chromatogram of CAF content. *, $P<0.05$ and **, $P<0.01$.

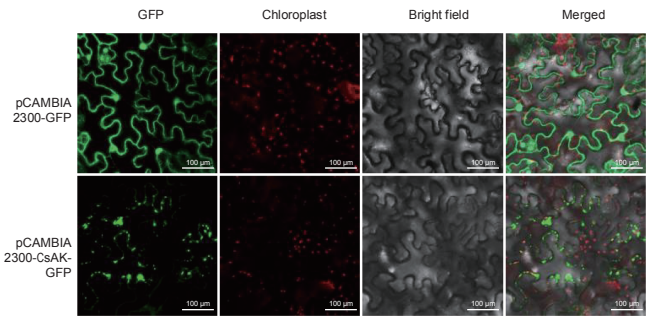


Fig. 6 Subcellular location of *CsAK*.

et al. 2022). However, the dynamic changes in CAF contents in different tea accessions of Guizhou Plateau were still unclear. This study investigated the CAF contents in 359 tea accessions collected from Guizhou Plateau. Results revealed that the distribution of CAF contents in 359 tea accessions followed a normal distribution, suggesting that there is obvious difference in CAF contents among 359 different tea accessions.

SNPs are efficient and powerful molecular markers for genotyping breeding populations, enabling the application of GWAS, studying genetic diversity, analyzing genetic linkage, and mining elite genes. In this study, we obtained 105,370 high-quality SNPs of 359 tea accessions to perform genetic variation analysis. Population structure analysis revealed that 359 tea accessions were clustered into six groups: five inferred pure groups (G-I–G-V) and one-mixed group (G-VI). The population structure was further validated using PCA and NJ tree analyses, which confirmed the results, although

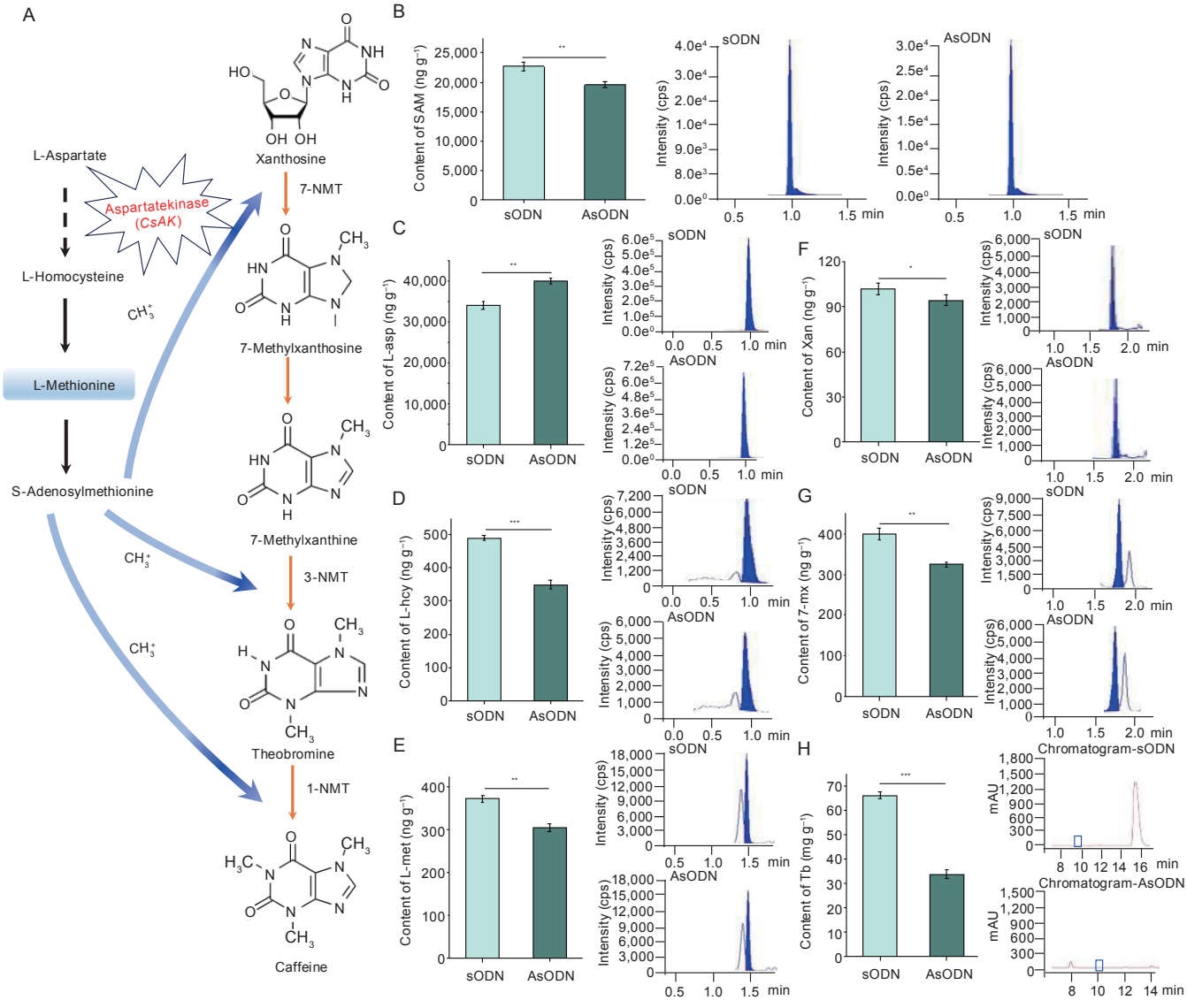


Fig. 7 Effects of *CsAK* silencing on dynamic changes of upstream and downstream substances in tea accessions. A, the *CsAK* gene metabolism pathway and the most important steps in the caffeine (CAF) synthesis pathway. B, the content and ion chromatogram peaks of S-adenosylmethionine (SAM). C, the content and ion chromatogram peaks of L-aspartate (L-asp). D, the content and ion chromatogram peaks of L-homocysteine (L-hcy). E, the content and ion chromatogram peaks of L-methionine (L-met). F, the content and ion chromatogram peaks of xanthosine (Xan). G, the content and ion chromatogram peaks of 7-methylxanthine (7-mx). H, the content and chromatogram of theobromine (Tb). *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$.

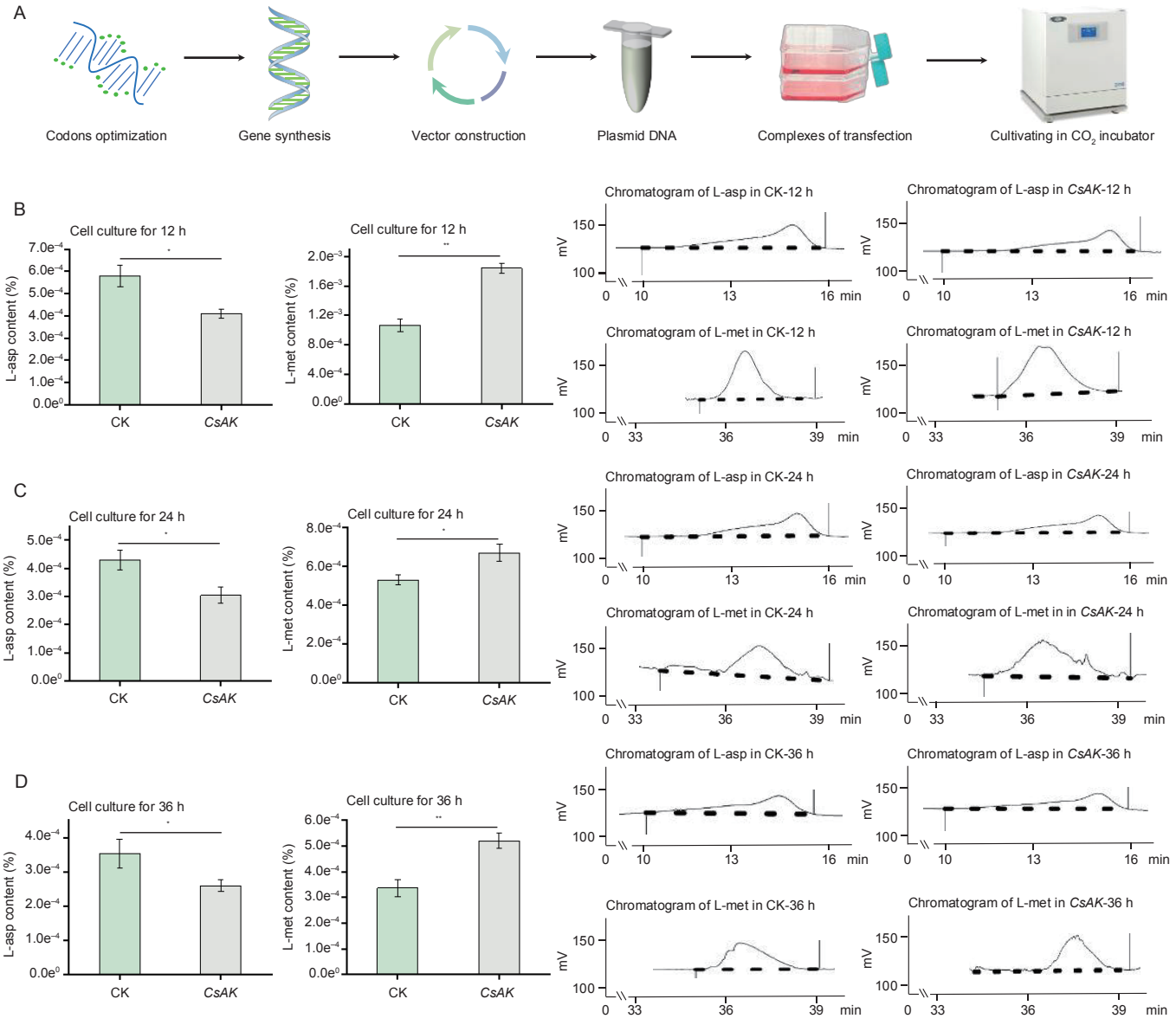


Fig. 8 Effects of overexpression of *CsAK* in cells on the accumulation of L-aspartate (L-asp) and L-methionine (L-met) contents. A, experiment flowchart. B–C, the contents of L-asp and L-met after 12 h (B), 24 h (C) and 36 h (D) of cultivation. *, $P<0.05$; **, $P<0.01$.

some accessions showed overlapping characteristics. The GWAS analysis was used as an effective method to detect the genetic variation locus of interested phenotypic traits. However, the application of GWAS in exploring the molecular mechanism of CAF biosynthesis in tea accessions is still not reported. Our study revealed that a total of 19 high-quality SNPs associated with CAF content ($-\log_{10}(P)\geq 4.0$) were identified based on the GWAS analysis. The LD decay rate was observed at 3.5 kb when r^2 decreased to half its maximum value, and this distance was used to identify candidate genes for specific traits. Although linkage genes are usually detected within the LD range during population genotype analysis, this approach may miss some important genes. To improve gene identification, we extended the search range to 200 kb, as recommended in previous studies (Lu *et al.* 2016). Within this expanded range, we identified two genes associated with the 19 SNPs linked to CAF content.

Aspartate kinase (AK) is the first key enzyme in the aspartate metabolism pathway. At present, researches on AK mainly focus on modifications in *Escherichia coli* to promote the accumulation of L-lysine (Kotaka *et al.* 2006; Xu *et al.* 2016), and the mutant of AK in *Corynebacterium pekinense* expressed strong binding affinity for aspartate and enhanced the synthesis of lysine (Gao *et al.* 2019). However, AK is rarely reported in plant research. Previous studies have shown that overexpression of aspartate kinase and cystathionine γ-synthase in tobacco leaves can enhance the accumulation of methionine and threonine (Hacham *et al.* 2008). Consequently, AK demonstrates critical importance for contributing to downstream metabolite accumulation. KEGG analysis and gene functional annotation further revealed that the *CsAK* gene encoding aspartate kinase was involved in catalyzing the L-methionine biosynthesis, which was further

generated into the methyl donor S-adenosylmethionine (SAM) (Ravanel *et al.* 1998; Jander and Joshi 2010; Wang *et al.* 2018). Previous studies had revealed that SAM was used as a methyl donor involved in CAF biosynthesis (Suzuki 1972; Jin *et al.* 2016; Xia *et al.* 2017; Gong *et al.* 2020; Mi *et al.* 2023). Our results suggest that the expression level of *CsAK* gene was consistent with the dynamic changes of CAF contents in three different tea accessions, indicating that *CsAK* gene was positively correlated with the CAF contents. The application of AsODN for the *CsAK* gene significantly suppressed the *CsAK* gene expression and decreased the CAF content in the apical buds treated with AsODN. Moreover, the content of intermediate metabolites during the CAF biosynthesis process was analyzed and found that six intermediates (L-homocysteine, L-methionine, S-adenosylmethionine, xanthosine, 7-methylxanthine, and theobromine) were significantly reduced in the AsODN-treated apical buds in comparison with the control groups, which were consistent with the findings of previous studies. The results further indicated that *CsAK* gene was involved in mediating the CAF biosynthesis. After overexpressing the *CsAK* gene in eukaryotic cells, we investigated the dynamic changes of two metabolites (L-aspartate and L-methionine) of the aspartate metabolic pathway in eukaryotic cells, found that the L-aspartate content was reduced in comparison with the control group, while the L-methionine content was increased. Above results indicated that *CsAK* gene was involved in catalyzing L-aspartate to generate the key intermediate-L-methionine, thereby providing methyl groups for CAF biosynthesis.

5. Conclusion

In this study, we identified the candidate gene *CsAK* related to CAF contents that is located with 200 kb upstream and downstream regions of significant SNPs. Functional annotation revealed that the *CsAK* gene may play a vital role in the high accumulation of CAF. The expression patterns of *CsAK* gene based on the RT-qPCR result were consistent with the dynamic changes of CAF contents in three tea accessions. Subcellular localization revealed that the protein encoded by *CsAK* gene was mainly located on cell membrane. AsODN interference in the buds and leaves significantly decreased the expression levels of the *CsAK* gene, which decreased the CAF content in the AsODN-treated apical buds. Additionally, overexpression of *CsAK* gene in eukaryotic cells resulted in the accumulation of key intermediate product, L-methionine during CAF biosynthesis process. These findings suggest that *CsAK* plays a key role in catalyzing L-aspartate to generate the key intermediate-L-methionine, thereby providing methyl groups for CAF biosynthesis. This study provides theoretical insights and practical applications for breeding tea accessions and essential groundwork for the cultivation of excellent germplasm with high or low levels of CAF in the future.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (32060700), the National Guidance Foundation for Local Science and Technology Development of China [(2023) 009], the Guizhou Provincial Basic Research Program (Natural Sciences) — Youth Initiative Project, China([2024] Youth 157), the Guiyang Science and Technology Plan Project (Construction Technology Contract, China [2023] 48-21), the Post-funding Project of the National Key R&D Program of China (2021YFD1100307), the Guizhou Provincial Science and Technology Program, China ([2023] General 481), the Qianxinan Prefecture Science and Technology Program, China (2022-1-52), the Youth Science and Technology Talent Project of Guizhou Provincial Department of Education, China (KY [2022]153), and the Basic Research Project of Guizhou University, China (2023-38).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2025.06.026

References

Ahammed G J, Li X. 2022. Hormonal regulation of health-promoting compounds in tea (*Camellia sinensis* L.). *Plant Physiology and Biochemistry*, **185**, 390–400.

Bradbury P J, Zhang Z, Kroon D E, Casstevens T M, Ramdoss Y, Buckler E S. 2007. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, **23**, 2633–2635.

Chen W, Hou L, Zhang Z Y, Pang X M, Li Y Y. 2017. Genetic diversity, population structure, and linkage disequilibrium of a core collection of *Ziziphus jujuba* assessed with genome-wide SNPs developed by genotyping-by-sequencing and SSR markers. *Frontiers in Plant Science*, **8**, 575.

Chen Y F, Jiang C J, Yin S X, Zhuang J H, Zhao Y, Zhang L J, Jiang X L, Liu Y J, Gao L P, Xia T. 2023. New insights into the function of plant tannase with promiscuous acyltransferase activity. *The Plant Journal*, **113**, 576–594.

Chen Y J, Niu S Z, Deng X Y, Song Q F, He L M, Bai D C, He Y Q. 2023. Genome-wide association study of leaf-related traits in tea plant in Guizhou based on genotyping-by-sequencing. *BMC Plant Biology*, **23**, 196.

Van Dam R M, Hu F B. 2022. Caffeine consumption and cardiovascular health. *Nature Reviews Cardiology*, **19**, 429–430.

Danecek P, Auton A, Abecasis G, Albers C A, Banks E, DePristo M A, Handsaker R E, Lunter G, Marth G T, Sherry S T, McVean G, Durbin R, 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Deng X L, Huang D J, Wang Y H, An H W, Bai D C, Wang X J, Niu S Z, Song X M. 2025. Genome-wide association study of salicylic acid provides genetic insights for tea plant selective breeding. *Horticulture Research*, **12**, 362.

Devitt G, Thomas M, Klibanov A M, Pfeiffer T, Bosch V. 2007. Optimized protocol for the large scale production of HIV pseudovirions by transient transfection of HEK293T cells with linear fully deacylated polyethylenimine. *Journal of Virological Methods*, **146**, 298–304.

Fang K X, Xia Z Q, Li H J, Jiang X H, Qin D D, Wang Q H, Wang Q, Pan C D, Li B, Wu H L. 2021. Genome-wide association analysis identified molecular markers associated with important tea flavor-related metabolites. *Horticulture Research*, **8**, 42.

Gao Y N, Han C J, Liu C L, Wang J, Zhao L, Fang L, Min W H. 2019. Enzymatic characterization and molecular mechanism of a novel

aspartokinase mutant M372I/T379W from *Corynebacterium pekinense*. *Royal Society of Chemistry Advances*, **9**, 21344–21354.
Gong A D, Lian S B, Wu N N, Zhou Y J, Zhao S Q, Zhang L M, Cheng L, Yuan H Y. 2020. Integrated transcriptomics and metabolomics analysis of catechins, caffeine and theanine biosynthesis in tea plant (*Camellia sinensis*) over the course of seasons. *BMC Plant Biology*, **20**, 294.
Grover A, Sharma P C. 2016. Development and use of molecular markers: Past and present. *Critical Reviews in Biotechnology*, **36**, 290–302.
Hacham Y, Matityahu I, Schuster G, Amir R. 2008. Overexpression of mutated forms of aspartate kinase and cystathionine gamma-synthase in tobacco leaves resulted in the high accumulation of methionine and threonine. *The Plant Journal*, **54**, 260–271.
He L M, Luo J, Niu S Z, Bai D C, Chen Y J. 2023. Population structure analysis to explore genetic diversity and geographical distribution characteristics of wild tea plant in Guizhou Plateau. *BMC Plant Biology*, **23**, 255.
Helyar S J, Hemmer-Hansen J, Bekkevold D, Taylor M I, Ogden R, Limborg M T, Cariani A, Maes G E, Diopere E, Carvalho G R, Nielsen E E. 2011. Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. *Molecular Ecology Resources*, **1**, 123–136.
Jander G, Joshi V. 2010. Recent progress in deciphering the biosynthesis of aspartate-derived amino acids in plants. *Molecular Plant*, **3**, 54–65.
Jiang P, Ren L J, Zhi L, Yu Z, Lv F X, Xu F L, Peng W, Bai X Y, Cheng K L, Quan L, Zhang X Q, Wang X H, Zhang Y, Yang D, Hu X L, Xiao R P. 2021. Negative regulation of AMPK signaling by high glucose via E3 ubiquitin ligase MG53. *Molecular Cell*, **81**, 629–637.
Jin J Q, Chai Y F, Liu Y F, Zhang J, Yao M Z, Chen L. 2018. Hongyacha, a naturally caffeine-free tea plant from Fujian, China. *Journal of Agricultural and Food Chemistry*, **66**, 11311–11319.
Jin J Q, Yao M Z, Ma C L, Ma J Q, Chen L. 2016. Association mapping of caffeine content with TCS1 in tea plant and its related species. *Plant Physiology and Biochemistry*, **105**, 251–259.
Jin Q F, Wang Z, Sandhu D, Chen L, Shao C Y, Shang F H Z, Xie S Y, Huang F Y, Chen Z Y, Zhang X Q, Hu J Y, Liu G Z, Su Q, Huang M D, Liu Z H, Huang J A, Tian N, Liu S Q. 2023. mRNA-miRNA analyses reveal the involvement of CsbHLH1 and miR1446a in the regulation of caffeine biosynthesis in *Camellia sinensis*. *Horticulture Research*, **11**, 282.
Kim H, Kang S H, Kim S H, Kim S H, Hwang J, Kim J G, Han K, Kim J B. 2021. Drinking coffee enhances neurocognitive function by reorganizing brain functional connectivity. *Scientific Reports*, **11**, 14381.
Koshiishi C, Kato A, Yama S, Crozier A, Ashihara H. 2001. A new caffeine biosynthetic pathway in tea leaves: Utilisation of adenosine released from the S-adenosyl-L-methionine cycle. *FEBS Letters*, **499**, 50–54.
Kotaka M, Ren J S, Lockyer M, Hawkins A R, Stammers D K. 2006. Structures of R- and T-state *Escherichia coli* aspartokinase III. Mechanisms of the allosteric transition and inhibition by lysine. *Journal of Biological Chemistry*, **281**, 31544–31552.
Li F, Deng X Y, Huang Z, Zhao Z F, Li C Y, Song Q F, He Y Q, Niu S N. 2023. Integrated transcriptome and metabolome provide insights into flavonoid biosynthesis in “P113”, a new purple tea of *Camellia tachangensis*. *Beverage Plant Research*, **3**, 3.
Li G F, Tan M, Ma J J, Cheng F, Li K, Liu X J, Zhao C P, Zhang D, Xing L B, Ren X L, Han M Y, An N. 2021. Molecular mechanism of MdWUS2-MdTCP12 interaction in mediating cytokinin signaling to control axillary bud outgrowth. *Journal of Experimental Botany*, **72**, 4822–4838.
Lu Q, Zhang M C, Niu X J, Wang C H, Xu Q, Feng Y, Wang S, Yuan X P, Yu H Y, Wang Y P, Wei X H. 2016. Uncovering novel loci for mesocotyl elongation and shoot length in *indica* rice through genome-wide association mapping. *Planta*, **243**, 645–657.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo M A. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297–1303.
Mi X Z, Yang C, Qiao D H, Tang M S, Guo Y, Liang S H, Li Y, Chen Z W, Chen J. 2023. De novo full length transcriptome analysis of a naturally caffeine-free tea plant reveals specificity in secondary metabolic regulation. *Scientific Reports*, **13**, 6015.
Misako K, Kouichi M. 2004. Caffeine synthase and related methyltransferases in plants. *Frontiers in Bioscience-Landmark*, **9**, 1833–1342.
Mohanpuria P, Kumar V, Yadav S K. 2010. Tea caffeine: Metabolism, functions, and reduction strategies. *Food Science and Biotechnology*, **19**, 275–287.
Niu S N, Koiwa H, Song Q F, Qiao D H, Chen J, Zhao D G, Chen Z W, Wang Y, Zhang T Y. 2020. Development of core-collections for Guizhou tea genetic resources and GWAS of leaf size using SNP developed by genotyping-by-sequencing. *PeerJ*, **8**, 8572.
Niu S N, Song Q F, Koiwa H, Qiao D H, Zhao D G, Chen Z W, Liu X, Wen X P. 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**, 328.
Qiao D H, Yang C, Chen J, Guo Y, Li Y, Niu S Z, Cao K M, Chen Z W. 2019. Comprehensive identification of the full-length transcripts and alternative splicing related to the secondary metabolism pathways in the tea plant (*Camellia sinensis*). *Scientific Reports*, **9**, 2709.
Pandey J, Scheuring D C, Koym J W, Coombs J, Novy R G, Thompson A L, Holm D G, Douches DS, Miller Jr J C, Vales M I. 2021. Genetic diversity and population structure of advanced clones selected over forty years by a potato breeding program in the USA. *Scientific Reports*, **11**, 8344.
Ravanel S, Gakière B, Job D, Douce R. 1998. The specific features of methionine biosynthesis and metabolism in plants. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 7805–7812.
Reddy V S, Shiva S, Manikantan S, Ramakrishna S. 2024. Pharmacology of caffeine and its effects on the human body. *European Journal of Medicinal Chemistry Reports*, **10**, 100138.
Shi R, Cao Y, Yang T H, Wang Y P, Xiang Y N, Chen F, Zhang W, Zhou X Y, Sun C M, Fu S X, Hu M L, Zhang J F, Wang X D. 2024. Genome-wide association study reveals the genetic basis of crude fiber components in *Brassica napus* L. shoots at stem elongation stage. *Journal of Agricultural and Food Chemistry*, **72**, 16530–16540.
Srdjenovic B, Djordjevic-Milic V, Grujic N, Injac R, Lepojevic Z. 2008. Simultaneous HPLC determination of caffeine, theobromine, and theophylline in food, drinks, and herbal products. *Journal of Chromatographic Science*, **46**, 144–149.
Stecher G, Tamura K, Kumar S. 2020. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. *Molecular Biology and Evolution*, **37**, 1237–1239.
Suzuki T. 1972. The participation of S-adenosylmethionine in the biosynthesis of caffeine in the tea plant. *FEBS Letters*, **24**, 18–20.
Tang W J, Wu T T, Ye J, Sun J, Jiang Y, Yu J, Tang J P, Chen G M, Wang C M, Wan J M. 2016. SNP-based analysis of genetic diversity reveals important alleles associated with seed size in rice. *BMC Plant Biology*, **16**, 93.
Teng J, Yan C Y, Zeng W, Zhang Y Q, Zeng Z, Huang Y H. 2020. Purification and characterization of theobromine synthase in a Theobromine-enriched wild tea plant (*Camellia gymnogyna* Chang) from Dayao Mountain, China. *Food Chemistry*, **311**, 125875.
Uffelmann E, Huang Q Q, Munung N S, Vries J D, Okada Y, Martin A R, Martin H C, Lappalainen T, Posthuma D. 2021. Genome-wide association studies. *Nature Reviews Methods Primers*, **1**, 1–21.
Wang J D, Zhang Z L, Gong Z L, Liang Y J, Ai X T, Sang Z W, Guo J P, Li X Y, Zheng J Y. 2022. Analysis of the genetic structure and diversity of upland cotton groups in different planting areas based on SNP markers. *Gene*, **809**, 146042.
Wang R J, Gao X F, Yang J, Kong X R. 2019. Genome-wide association study to identify favorable SNP allelic variations and candidate genes that control the timing of spring bud flush of tea (*Camellia sinensis*) using SLAF-seq. *Journal of Agricultural and Food Chemistry*, **67**, 10380–10391.
Wang W Y, Xu M Y, Wang G P, Galili G. 2018. New insights into the metabolism of aspartate-family amino acids in plant seeds. *Plant Reproduction*, **31**, 203–211.
Wang Z H, Huang R, Moon D G, Ercisli S, Chen L. 2023. Achievements and prospects of QTL mapping and beneficial genes and alleles

mining for important quality and agronomic traits in tea plant (*Camellia sinensis*). *Beverage Plant Research*, **3**, 22.

Xia E H, Tong W, Hou Y, An Y L, Chen L B, Wu Q, Liu Y L, Yu J, Li F D, Li R P, Li P H, Zhao H J, Ge R H, Huang J, Mallano A L, Zhang Y R, Liu S R, Deng W W, Song C K, Zhang Z L, *et al.* 2020a. The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation. *Molecular Plant*, **13**, 1013–1026.

Xia E H, Tong W, Wu Q, Wei S, Zhao J, Zhang Z Z, Wei C L, Wan X C. 2020b. Tea plant genomics: Achievements, challenges and perspectives. *Horticulture Research*, **7**, 7.

Xia E H, Zhang H B, Sheng J, Li K, Zhang Q J, Kim C, Zhang Y, Liu Y, Zhu T, Li W, Huang H, Tong Y, Nan H, Shi C, Shi C, Jiang J J, Mao S Y, Jiao J Y, Zhang D, Zhao Y, *et al.* 2017. The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. *Molecular Plant*, **10**, 866–877.

Xu J Z, Han M, Ren X D, Zhang W G. 2016. Modification of aspartokinase III and dihydrodipicolinate synthetase increases the production of L-lysine in *Escherichia coli*. *Biochemical Engineering Journal*, **114**, 79–86.

Yao X Z, Chen H F, Ai A T, Wang F, Lian S S, Tang H, Jiang Y H, Jiao Y J, He Y M, Li T, Lu L T. 2023. The transcription factor CsS40 negatively regulates *TCS*$_1$ expression and caffeine biosynthesis in connection to leaf senescence in *Camellia sinensis*. *Horticulture Research*, **10**, 162.

Yoshida H, Okada S, Wang F M, Shiota S, Mori M, Kawamura M, Zhao X, Wang Y Q, Nishigaki N, Kobayashi A, Miura K, Yoshida S, Ikegami M, Ito A, Huang L T, Caroline Hsing Y I, Yamagata Y, Morinaka Y, Yamasaki M, Kotake T, *et al.* 2023. Integrated genome-wide differentiation and association analyses identify causal genes underlying breeding-selected grain quality traits in *japonica* rice. *Molecular Plant*, **16**, 1460–1477.

Yu S W, Li P H, Zhao X C, Tan M M, Ahmad M Z, Xu Y J, Tadege M, Zhao J. 2021. CsTCPs regulate shoot tip development and catechin biosynthesis in tea plant (*Camellia sinensis*). *Horticulture Research*, **8**, 104.

Zhang C, Dong S S, Xu J Y, He W M, Yang T L. 2019. PopLDdecay: A fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*, **35**, 1786–1788.

Zhang S R, Jin J Q, Chen J D, Ercisli S, Chen L. 2022. Purine alkaloids in tea plants: Component, biosynthetic mechanism and genetic variation. *Beverage Plant Research*, **2**, 13.

Zhang W Y, Zhang Y J, Qiu H J, Guo Y F, Wan H L, Zhang X L, Scossa F, Alseekh S, Zhang Q H, Wang P, Xu L, Schmidt M H-W, Jia X X, Li D L, Zhu A T, Guo F, Chen W, Ni D J, Usadel B, Fernie A R, Wen W W. 2020. Genome assembly of wild tea tree DASZ reveals pedigree and selection history of tea varieties. *Nature Communications*, **11**, 3719.

Zhang Y R, Fu J M, Zhou Q Y, Li F D, Shen Y H, Ye Z L, Tang D K, Chi N, Li L Q, Ma S Y, Inayat M A, Guo T Y, Zhao J, Li P H. 2022. Metabolite profiling and transcriptome analysis revealed the conserved transcriptional regulation mechanism of caffeine biosynthesis in tea and coffee plants. *Journal of Agricultural and Food Chemistry*, **70**, 3239–3251.

Zhao M Y, Jin J Y, Wang J M, Gao T, Luo Y, Jing T T, Hu Y T, Pan Y T, Lu M Q, Schwab W, Song C K. 2022. Eugenol functions as a signal mediating cold and drought tolerance via UGT71A59-mediated glucosylation in tea plants. *The Plant Journal*, **109**, 1489–1506.

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

Zheng Y X, Yuan F, Huang Y Q, Zhao Y F, Jia X Y, Zhu L Y, Guo J J. 2021. Genome-wide association studies of grain quality traits in maize. *Scientific Reports*, **11**, 9797.

Zhou M Z, Yan C Y, Zeng Z, Luo L, Zeng W, Huang Y H. 2020. *N*-Methyltransferases of caffeine biosynthetic pathway in plants. *Journal of Agricultural and Food Chemistry*, **68**, 15359–15372.

Zou J N, Yu H, Yu Q, Jin X J, Cao L, Wang M Y, Wang M X, Ren C Y, Zhang Y X. 2021. Physiological and UPLC-MS/MS widely targeted metabolites mechanisms of alleviation of drought stress-induced soybean growth inhibition by melatonin. *Industrial Crops and Products*, **163**, 113323.

Executive Editor-in-Chief Sanwen Huang
Managing Editor Lingyun Weng