Associations Between SNPs and Vegetation Indices: Unraveling Molecular Insights for Enhanced Cultivation of Tea Plant (Camellia sinensis (L.) O. Kuntze)

Daria Kuzmina [1, 2], **Lyudmila S Malyukova** [1], **Karina Manakhova** [1, 2], **Tatyana Kovalenko** [1, 2], **Jaroslava Fedorina** [1, 2], **Aleksandra O Matskiv** [1], **Alexey V Ryndin** [1], **Maya V Gvasaliya** [1], **Yuriy L Orlov** [Corresp., 3], **Lidiia S Samarina** [Corresp. 1, 2]

[1] Federal Research Centre the Subtropical Scientific Centre of the Russian Academy of Sciences, Sochi, Russia

[2] Sirius University of Science and Technology, Sochi, Russia

[3] Agrarian and Technological Institute, Patrice Lumumba Peoples' Friendship University of Russia, Moscow, Russia

Corresponding Authors: Yuriy L Orlov, Lidiia S Samarina
Email address: orlov@bionet.nsc.ru, q11111w2006@yandex.ru

Background. Breeding programs for nutrient-efficient tea plant varieties could be advanced by the combination of genotyping and phenotyping technologies. This study was aimed to search functional SNPs in key genes related to the nitrogen-assimilation in the collection of tea plant *Camellia sinensis* (L.) Kuntze. To gain a better knowledge of the mechanisms underlying nitrogen uptake efficiency, an additional objective of the study was to identify potential associations between SNPs and phenotypic traits.

Methods. The study was conducted on the tea plant collection of *Camellia sinensis* (L.) Kuntze of Western Caucasus grown without nitrogen fertilizers. Phenotypic data was collected by measuring the spectral reflectance of leaves in the 350–1100 nm range calculated as vegetation indices by the portable hyperspectral spectrometer. Coefficients of determination and prediction analysis were used to identify the most effective vegetation indices for phenotypic nutrient efficiency in tea plants. Single nucleotide polymorphisms were identified by pooled amplicon sequencing, and SNPs were annotated with the variant annotation and effect prediction SnpEFF tool. To reveal associations between the functional SNPs and the efficient vegetation indices a linear regression model was applied.

Results. PCA and regression analysis revealed significant vegetation indices with high R2 values (more than 0.5) that correspond to nitrogen-deficiency tolerant (ZMI, CNDVI, RENDVI, VREI1, GM2, GM1, PRI, PSRI, PSRI, ARI2, ARI1, WBI, NDVI, SIPI, Lic1, WBI) and susceptible (MDATT, Ctr2, TCARI, MCARI1, VREI3, VREI2) genotypes. The following are the most reliable indices that, based on prediction analysis, have the greatest difference between tolerant and susceptible genotypes: ZMI, CNDVI, RENDVI, VREI1, GM2, GM1, PRI, and Ctr2, VREI3, VREI2. The largest SNPs frequency was observed in several flavonoid biosynthesis and N assimilation genes, namely *F3'5'Hb*, *UFGTa*, *UFGTb*, *4Cl*, and *AMT1.2*. Mutations in *NRT2.4*, *PIP*, *AlaDC*, *DFRa*, and *GS1.2* are inherent in genotypes with significant alterations in nitrogen and theanine concentrations, which are primarily susceptible to N deficiency. Also, only susceptible tea types are known to harbor mutations in *AlaAT1*, *MYB4*, and *WRKY57*, which lead to alterations in protein structure. Associations were discovered between SNPs that change the amino acids in *ASNb*, *PIP*, and *FRI* as well as between two mutations in *4Cl* and *WBI*.

Conclusions. The results will be useful to identify tolerant and susceptible tea plant genotypes under nitrogen deficiency . Revealed associations between mutations and vegetation indices improve our understanding of nitrogen-efficiency processes in future investigations on tea plant selection.

Associations Between SNPs and Vegetation Indices: Unraveling Molecular Insights for Enhanced Cultivation of Tea Plant (Camellia sinensis (L.) O. Kuntze)

Daria Kuzmina[1,2], Lyudmila S Malyukova[1], Karina Manakhova[1,2], Tatyana Kovalenko[1,2], Jaroslava Fedorina[1,2], Aleksandra O Matskiv[1], Alexey V Ryndin[1], Maya V Gvasaliya[1], Yuriy L Orlov[3], Lidiia S Samarina [1,2]

[1] Federal Research Centre the Subtropical Scientific Centre of the Russian Academy of Sciences, Sochi, Russia
[2] Sirius University of Science and Technology, Sochi, Russia
[3] Agrarian and Technological Institute, Patrice Lumumba Peoples' Friendship University of Russia, 117198 Moscow, Russia

Corresponding Author:
Lidiia Samarina[1]
Yana Fabritsiusa str. 2/28, Sochi, 354002, Russia
Email address: q11111w2006@yandex.ru

Abstract

Background. Breeding programs for nutrient-efficient tea plant varieties could be advanced by the combination of genotyping and phenotyping technologies. This study was aimed to search functional SNPs in key genes related to the nitrogen-assimilation in the collection of tea plant *Camellia sinensis* (L.) Kuntze. To gain a better knowledge of the mechanisms underlying nitrogen uptake efficiency, an additional objective of the study was to identify potential associations between SNPs and phenotypic traits.

Methods. The study was conducted on the tea plant collection of *Camellia sinensis* (L.) Kuntze of Western Caucasus grown without nitrogen fertilizers. Phenotypic data was collected by measuring the spectral reflectance of leaves in the 350–1100 nm range calculated as vegetation indices by the portable hyperspectral spectrometer. Coefficients of determination and prediction analysis were used to identify the most effective vegetation indices for phenotypic nutrient efficiency in tea plants. Single nucleotide polymorphisms were identified by pooled amplicon sequencing, and SNPs were annotated with the variant annotation and effect prediction SnpEFF tool. To reveal associations between the functional SNPs and the efficient vegetation indices a linear regression model was applied.

Results. PCA and regression analysis revealed significant vegetation indices with high R2 values (more than 0.5) that correspond to nitrogen-deficiency tolerant (ZMI, CNDVI, RENDVI, VREI1, GM2, GM1, PRI, PSRI, PSRI, ARI2, ARI1, WBI, NDVI, SIPI, Lic1, WBI) and susceptible (MDATT, Ctr2, TCARI, MCARI1, VREI3, VREI2) genotypes. The following are the most reliable indices that, based on prediction analysis, have the greatest difference between tolerant and susceptible genotypes: ZMI, CNDVI, RENDVI, VREI1, GM2, GM1, PRI, and Ctr2, VREI3, VREI2. The largest SNPs frequency was observed in several flavonoid biosynthesis and N assimilation genes, namely *F3'5'Hb*, *UFGTa*, *UFGTb*, *4Cl*, and *AMT1.2*. Mutations in *NRT2.4*, *PIP*, *AlaDC*, *DFRa*, and *GS1.2* are inherent in genotypes with significant alterations in nitrogen and theanine concentrations, which are primarily susceptible to N deficiency. Also, only susceptible tea types are known to harbor mutations in *AlaAT1*, *MYB4*, and *WRKY57*, which lead to alterations in protein structure. Associations were discovered between SNPs that change the amino acids in *ASNb*, *PIP*, and *FRI* as well as between two mutations in *4Cl* and *WBI*.

Conclusions. The results will be useful to identify tolerant and susceptible tea plant genotypes under nitrogen deficiency. Revealed associations between mutations and vegetation indices improve our understanding of nitrogen-efficiency processes in future investigations on tea plant selection.

Keywords: tea plant, SNP, vegetation, Camellia sinensis, flavonoid biosynthesis, nitrogen uptake efficiency, tea plant genotypes

Introduction

Tea, derived from the perennial evergreen woody plant *Camellia sinensis* (L.) O. Kuntze, stands as one of the world's most consumed beverages, prized for its aromatic flavor and potential health benefits (Samanta, 2020; Sánchez et al., 2020; Malyukova et al., 2022)□. Tea has demonstrated numerous pharmacological properties, including antioxidant and anticancer effects, as well as the ability to reduce metabolic issues and prevent cardiovascular diseases (Chan et al., 2011; Filippini et al., 2020; Brimson et al., 2023)□. The secondary metabolites that determine the tea quality, such as theanine, caffeine, flavonoids, and amino acids, enhance the beneficial biological activities and taste of tea plants (Gai et al., 2019)□. The metabolism of these plant compounds, and hence the tea quality, is dependent on a variety of factors, including nitrogen supply (Yang et al., 2018)□.

Nitrogen (N), a crucial component for plant development, is frequently supplied via fertilizers to guarantee optimal growth. However, excess N inhibits the formation of flavonol glycosides, whereas decreasing N availability reduces amino acid and caffeine concentrations in mature tea leaves (Li et al., 2016; Dong et al., 2019). In addition, long-term nitrogen fertilization is not only expensive, but it also causes an array of environmental issues, including greenhouse gas emissions, soil pH changes, eutrophication, and microbial community disruption (Gao & Cabrera Serrenho, 2023; Kamran et al., 2023; Liu et al., 2023; Tang et al., 2023). Use of tea cultivars with high NUE (nitrogen uptake efficiency) and high quality is necessary to preserve environmental pollution and promote productivity.

Nitrogen-efficient varieties are likely to have polymorphisms in the genes that control nitrogen metabolism and determine the tea quality (Li et al., 2017; Yang et al., 2020; Xie et al., 2023). Genes involved in N uptake (*PIP, LHT1)*, transport (*AMT1.2, NRT2.4*) and assimilation (*AlaAT1, GDHa, GDH2*, *GS1.2*) were identified as well as genes that regulate secondary metabolites which expression changes depending on the nitrogen level (Wang et al., 2021b; Li et al., 2021; Xie et al., 2023; Wang et al., 2021a; Tang et al., 2021; Chen et al., 2023; Zhang et al., 2023). Genes *MYB7, MYB4, WD40, HLH35, HLH36, UFGTa, UFGTb, DFRa, F3'5'Hb* and *F3'5'Ha* are involved in the flavonoid pathway, whereas *AlaDC* controls the theanine synthesis (Huang et al., 2018; Liu et al., 2018; Dong et al., 2019; Guo et al., 2019; Wang et al., 2021b; Ye et al., 2021; Li et al., 2023). Gene *4Cl* mediates phenylpropanoid metabolism, *bG* is critical for tea aroma generation, *ANRb*, *ANSa*, *ANSb*, and *LAR* regulate the catechin pathway, and *WRKY57* modulates stress responses (Chen et al., 2009; Liu et al., 2015; Wani et al., 2021; Li et al., 2022; Zhao et al., 2022a). In a previous work, we described 20 tea genotypes from Northwest Caucasia that are susceptible or tolerant to nitrogen deficit. A number of polymorphisms in the tea quality genes and their relationships with certain phenotypic traits as biochemical measurements were revealed in the tea collection (Samarina et al., 2023).

SNPs markers and numerous metabolic profile approaches could be utilized for identifying nitrogen-efficient cultivars (Hazra et al., 2018)□. The remote sensing technology provides a non-destructive and rapid approach to gauge plant health and development, offering insights into the metabolism change of tea plant response to nitrogen deficiency (Cao et al., 2022)□. The changes in plant chemical composition could be described by reflectance light-based indices or vegetation indices (VIs) developed based on the reflectance data (Kior et al., 2021)□. Combinations of spectral bands could be utilized for generating vegetation indices because pigments have the ability to absorb light in certain bands. Vegetation indices such as WBI, PRI, NDVI, TCARI, TVI, and ZMI, FRI, ARI1, and ARI2 provide information regarding plant water status, photosynthetic factors, and secondary metabolism, respectively (Frels et al., 2018; Prey, Hu & Schmidhalter, 2020)□. The use of vegetation indices to determine insect, cold, drought and nitrogen shortage stress enable the selection of the best growing conditions for tea plants (Chen et al., 2021; Zhao et al., 2022b; Mao et al., 2023)□. Few research using unidentified aerial vehicles (UAVs) were conducted on the quality of tea and nitrogen deficiency (Luo et al., 2022)□. However, handled spectrometry was not tested to reveal the most efficient VIs for tea phenotyping, while experiments with potted plants rather than field studies are relevant for QTL and association mapping (Hazra et al., 2018)□.

In this study, we evaluate the efficiency of 31 VIs collected by a handheld spectrometer to reveal their efficiency for distinguishing ND-tolerant and ND-susceptible tea genotypes. We analyzed SNPs in 30 key genes related to N-assimilation and quality in the collection of 34 genotypes of tea plants in Western Caucasus. We aimed to identify relationships between genotype and phenotype traits in ND-tolerant and ND-susceptible tea cultivars. The findings of the study could be used as markers for screening ND-tolerant tea genotypes. This research may advance precise breeding strategies aimed to enhance yield quality of *Camellia sinensis* (L.) O. Kuntze under ND by defining the genetic determinants and chemical composition linked to ND-response.

Materials & Methods

Plant material

The plant materials were obtained from the field gene bank of the Russian Academy of Sciences' Federal Research Center's Subtropical Scientific Center (FRC SSC RAS) (Samarina et al., 2020). This study comprised mutant forms obtained between 1970 and 1980 from seeds (mostly cultivars "Kolkhida" and "Qimen") exposed to γ-irradiation. Each genotype of plants was clonally reproduced using 30–60 repetitions, and they were cultivated on acid soil from a brown forest (pH 5.5) with 30 mg kg−1 of nitrogen (as opposed to the ideal 80 mg kg−1 N for tea plantations). For the past 27 years, no fertilizers have been added to the experimental plot.

Library preparation and amplicon sequencing

The library preparation and sequencing procedure for the following 14 genotypes of tea plants is explained; our earlier research on gene selection and primer design, long-range polymerase chain reaction, and sequencing for the remaining 20 variations can be reviewed in (Samarina et al., 2023).

Using the NEBNext Ultra II DNA Reagent Kit Library Prep Kit for Illumina and following the manufacturer's instructions, fragment DNA libraries were created equimolarly from the mixed PCR results. The libraries were subjected to a qualitative assessment with High Sensitivity D5000 ScreenTape and High kits Sensitivity D5000 Reagents (Agilent, Santa Clara, CA, USA) on an Agilent bioanalyzer TapeStation 4150. Using the KAPA Library Quantification Kit (KAPA Biosystems, Wilmington, MA, USA), a real-time PCR was used to provide a quantitative assessment of the products.

The DNA library fragments were mixed equimolarly into a pool and sequenced on the Illumina MiSeq using pair-end reads 76+76 bp and single-end reads 151 bp. Using the default settings of the bcl2fastq v2.20.0.422 software, sequencing data were demultiplexed by index sequences. For each DNA library, a total of 184,000–392,000 pairs of reads were collected. The FastQC v0.11.2 program was used to carry out the first quality evaluation of the deep sequencing data. Low-quality sequences and adapters were eliminated using AdapterRemoval v2 programs (with settings --trimqualities, --minquality 20, --minlength 50). Following filtering, 94.34% of the read pairs were retained.

Data that had been filtered were mapped against the tea plant's reference genome (GCF_004153795.1). The BWA programs package's bwa mem function was utilized for mapping. Duplicates were eliminated using the MarkDuplicates function of the Picard tools v2.22.2 (Picard toolkit) software package. Samtools v1.9, a software application, was used to assess the alignments' quality. Using the COVERAGE_CAP = 10,000 option, the CollectWgsMetrics function of the Picard-tools software package (https://broadinstitute.github.io/picard/, accessed on March 2, 2024) was used to measure the depth coverage of the target genomic regions. 96.44% of reads on average were mapped to the genome of tea. On average, we were able to get 261-fold coverage of the target genomic areas of tea for each sample.

The raw data are deposited in the NCBI SRA database under accession numbers PRJNA1015448 (https://www.ncbi.nlm.nih.gov/sra/SRX21783698) and PRJNA977584.

Genotype analysis

Using BWA-MEM (version 0.7.12), the clean reads were aligned to the reference genome "Shuchazao" (Xia et al., 2020), and SAMtools (version 1.16.1) was used for sorting and combination of paired-end and single-end reads into a single-bam file. The GATK software (version 4.2) was used to add read groups. Variant calling was done using the GATK-

HaplotypeCaller method, with default parameters for diploid/unknown ploidy varieties and --sample-ploidy 3 and --sample-ploidy 4 for tetraploids and known triploids, respectively. The following parameters were utilized by the GATK software to select and filter SNPs/InDels: 'QD < 2.0||FS > 60.0||MQ < 40.0||SOR_filter||SOR > 4.0||DP < 261' and 'QD < 2.0||FS > 200.0||SOR > 10.0||DP < 261', respectively.

SnpEFF (version 5.0) was used to build the database for the reference genome "Shuchazao," and it then served to annotate the remaining variants. High, moderate, low, or modifier effect impact classifications were obtained via the SnpEff tool variation annotation. These genetic differences known as impact variations are expected to have an indirect, mild, moderate, or severe effect on the protein.

In order to facilitate further study, the discovered SNP data of the 14 tea varieties were combined with published data on 20 tea sorts (Supplementary Data S1). The formula for SNP density was mean SNP per gene divided by the gene's fragment length in kb. We normalized the SNP frequency in each gene to get a summary of the SNP distribution and potential SNP enrichments for the genes. Each SNP gene frequency was determined using the following formula: SNP_freq = (SNP_count/per_gene)/gene length × 10^3, where gene_length is the length of the gene and SNP_count/per_gene is the number of SNPs found in a particular gene. To make a fair comparison more straightforward, the SNP_Freq values were leveraged by applying factor 10^3 to the denominator.

Phenotypic analysis

In this work the efficiency of 31 different VIs was evaluated to phenotype ND-response in tea collection. Using a Ci-710s Miniature Leaf Spectrometer (CID Bio-Science, USA), the leaf spectral reflectance in the 350–1100 nm region was measured and 31 VIs were calculated. Five technical replications of each 33 genotypes were used to measure the reflectance in the middle of each leaf, next to the primary vehicle between 11:00 and 14:00. Data were statistically analyzed using the XLSTAT program (free trial version). To identify significant changes between the genotypes, one-way ANOVA, Fisher's and Tukey tests were performed. In addition, the study employed Pearson (n) PCA. The measured values of each VI as well as the results of statistical testing can be found in Supplementary Data S2.

Genotype and phenotype association analysis

For the association analysis, we combined SNP data from 20 and 14 different tea varieties. Locations of SNPs with moderate and high effect were mapped based on the alternative homozygous and heterozygous states of each allele (Supplementary Data S3). Only 26 types were subjected to a further study since phenotype data were available for a portion of the genotypes that had been sequenced. To determine the relationships between SNPs and the

phenotypes, a linear regression model was combined with a statistical test adjusted for multiple comparisons (Bonferroni and False Discovery Rate). Significant associations were identified at Bonferroni- and FDR-corrected p-values < 0.05. Statistical analysis and visualization were performed using R package (version 4.2.3).

Results

Phenotypic characterization

To reveal efficient VIs for phenotyping of ND- response, tea genotypes were classified as tolerant or susceptible to ND based on their leaf quality and N content. Eight genotypes were determined as tolerant of nitrogen deficit, whereas ten varieties were susceptible to it. Of these, fourteen genotypes did not exhibit any clear response to nitrogen deprivation and were classified as non-responsive.

We used PCA biplot to illustrate the related traits and correlations between vegetation indices and genotypes (Figure 1). The first two PCs exhibited a cumulative variation of approximately 73.03%. Both the ND-susceptible and ND-tolerant genotypes were clearly separated in the biplot. Most of the vectors of VIs were distributed with high loading on the positive side of PC1 and the negative side of PC2. The highest loading was observed in the following VIs: ZMI, VRE1, RENDVI, CNDVI, PRI, PSRI, GM2, GM1, and CRI2, NDVI, SIPI, and CRI1. The majority of the ND- tolerant genotypes were distributed closely around the positive side of the PC1 vectors, suggesting a significant correlation between those vectors and the nearest accessions. In contrast, genotypes with no clear response to ND were placed on the PC2 negative side. The vectors of TCARI, Ctr2, VRE2, VRE3, and MDATT were positioned on the negative side of PC1, whileLic2, SRPI, and MRESP1 - on the positive side of PC2. The most part of ND-susceptible genotypes were placed closely to them. Finally, few ND-tolerant accessions were spread out along all PCA plot sides.

Vegetation indices with a coefficient of determination R2 larger than 0.5 and a p value < 0.0001 were selected as efficient for distinguishing ND-tolerant and ND-susceptible tea accessions (Table 1). Tukey's multiple comparison analysis resulted in efficient indices, significantly different between tolerant and susceptible groups (P value 0.0001). The following VIs were corresponded to tolerant genotypes with the highest R2: ZMI, CNDVI, RENDVI, VREI1, GM2, GM1, PRI, PSRI, PRI, ARI2, ARI1, WBI, NDVI, SIPI Lic1, and WBI. On the other hand, ND-susceptible genotypes displayed higher values for MDATT, Ctr2, TCARI, MCARI1, VREI3, and VREI2.According to the prediction analysis, the greatest distance between susceptible and tolerant genotypes was observed by ZMI (Figure 2). ZMI values were found to be above 1.9 in most data points for tolerant genotypes and below 1.7 - for susceptible genotypes. RENDI and CNDVI were also demonstrated to be reliable; susceptible genotypes showed values below 1.35 while tolerant - above 1.40. Additionally, the remarkable differences were observed by PRI,

GM1, GM2 and VREI1, and tolerant genotypes displayed greater values. In contrast, ND-susceptible genotypes showed larger values of VREI3, VREI2, and Ctr2 above -0.07, -0.07 and 0.25, respectively.

Identification of SNPs in tea quality genes and their associations with phenotypes

Among 34 tea accessions, genes *4CL, AMT1,2,* and *F3'5'Hb* had the highest SNP densities (1.0-2.0) in exon regions, while genes *AlaAT1, GDH2, LAR, WD40, bG,* and *bHLH35* had the lowest densities (Table 2). The highest SNP densities in introns (5.0–6.0) was found in *4CL* and *GS1,2*. There were no SNPs found in *MYB7* or *bHLH36*. The largest percentage of polymorphisms in exon per gene (more than 45%) were detected for *bG, F3'5'Hb*, and *DFRa*.

The high-effect SNPs were observed in the following accessions: #619, #2697, #536, #1385 and #3986 (Figure 3). Each genotype showed SNPs with different degrees of effect: low, moderate, and modifier. Low-effect SNPs were found to have the highest percentages in #582, #157, and cv. Karatum, ranging from 4.0 to 25.0% across all genotypes. In cv. Sochi, #35, and #1292, moderate-effect SNPs have the highest rate, varying from 5.0 to 15.9% for all accessions. The highest percentages of SNPs with modifying effects were detected in #321, #619, and #3509, and ranged from 63.0 to 86.7% across all genotypes.

According to the annotation tool, the intron variant (8.9-57.14%) is the most common variant across all genotypes; #321 and #35 have the lowest proportion, while susceptible genotypes #551, #507, and #1467 have the highest percentage (Figure 3, Supplementary Data S4). Genotypes #321, #35 have the highest percentage of SNPs within a gene (intragenic variations) at roughly 56–67%, whereas #619, #1385 have the lowest percentage at 1.5–4.5%. Intergenic areas were detected with the highest value of 9–11% in #321 and #3823 and the lowest value of 0.5–0.8 % in #Sochi and #837. The 3′-UTR (0.6-8.0%), 5′-UTR (0.5-6.38%), and splice region or accepter variations (0.2-2.13%) were shown to be the least prevalent variants. 5′-UTR SNP rates were lowest in #Sochi and #4605, and greatest in #582, cv. Karatum, and #551. On the other hand, the percentage of 3′-UTR SNPs was lowest in cv. Karatum, #1476, #837, and highest in #1292, #1385. Splice areas and splice acceptor variations were rare observed in #3986 and #619 and are predominantly occurring in cv. Karatum, #582, and #3823. The frequency of polymorphisms that were classified as downstream and upstream gene variations was 5.6-21.38% and 0.5-18%, respectively. The greatest frequencies were found in #619, #3180, #855 and #257, #501, while the lowest - in #551 and #3823, #527.

The highest SNPs frequency was observed in exons of *UFGTa*, *4Cl*, *UFGTb,* and *AMT1.2*, while the lowest - in *GDH2*, *WD40, bHLH35*, *AlaAT1*, *LAR*, *GDHa*. The clustering method indicated 2 distinct branches (Figure 4). The first branch consisted of four tea varieties with the highest SNP frequencies in *UFGTa*: #507, #1476, #1484 and #Sochi. Among them, ND-susceptible #507 displayed the lowest leaf N-content, #1476 – ND-tolerant with high leaf N-content, and #1484

and #Sochi showed an uncertain reaction to ND. The second branch consisted of the two sub-branches. The first sub-branch combined the ND-susceptible genotypes with the low leaf nitrogen content, namely #1385, #3986, #1467, #582, #527, #1877, and #536. Besides, ND-tolerant genotype #619 and the high nitrogen-content genotypes #316, #212, and #1405 were joined to this sub-branch. All these tea plant genotypes displayed significant SNP frequencies in *NRT2.4*, *PIP*, *AlaDC*, *DFRa*, *GS1.2*, *F3'5'Hb*, *UFGTa*, *UFGTb*, *4Cl*, and *AMT1.2*. The second sub-branch combined ND-susceptible genotypes (#501, #551), ND-tolerant genotypes (#157, #2697, #3609, #4605) and non-responsive to ND.

Amino acid changes were caused by 109 SNPs that were categorized as missense variations with a moderate effect (Supplementary Data S5). A single SNP in *WRKY57* with a significant effect was identified as a splice acceptor and intron variant in ND-susceptible genotypes #3986 and #1385, as well as ND-tolerant genotypes #619, #2697, and #536. The most frequent amino acid alterations were discovered in genes *4CL, F3'5'Hb, F3'5'Ha* and *ANRb-ANR1*, while the rest were detected in *UFGTa, ANSb, ANSa, WRKY57, AlaDC, AlaAT1*, *GDHa, bG, MYB4, NRT2.4, PIP*, *UFGTb, WD40, GS1.2, AMT1.2, GDH2* and *DFRa.* A number of mutations were discovered to be specific for ND-susceptible genotypes and genotypes with low N content #855, #3574, and #536. These mutations lead to amino acid changes in the genes *AlaAT1*, *MYB4*, and *WRKY57.*

Four significant associations (p value < 0.05) were revealed between the SNPs and vegetation indices (Table 3). Two SNPs in gene *4Cl* are associated with the Water Band Index (WBI), which has a significant coefficient of determination (R2=0.624). Both SNPs of the *4Cl* gene were occurred in #1292 and the ND-susceptible genotype #507. Previously, we showed that one SNP in *4Cl* was substantially associated with Theaflavin content (Samarina et al., 2023). Additionally, associations between the Flavonols Reflectance Index or FRI (R2=0.211) and SNPs that alter the amino acid composition of *PIP* and *ANSb* were found. While the SNP in *ANSb* was observed in ND-tolerant genotypes #619, #157, ND-susceptible #582, #1385 and #536 with low tea quality. Besides, the mutation in the *PIP* gene was found in ND-susceptible and ND-tolerant genotypes #157 and #212, which are characterized by high leaf N and caffeine levels.

Discussion

This study was the first to employ portable spectrometry to reveal efficient vegetation indices for phenotyping of ND-tolerant tea plants. Totally, 20 of 31 VIs showed to be efficient for ND-response phenotyping. Also, prediction analysis indicated the greatest gap for ZMI, RENDI, CNDVI, PRI, GM1, GM2, VREI1 (tolerant genotypes have higher values) and VREI3, VREI2, Ctr2 (susceptible genotypes have higher values), suggesting that these are the most reliable vegetation indices for ND-response phenotyping. Tolerant and susceptible genotypes are represented by VIs with the following designations: ZMI, GM1, GM2, VRE1, CNDVI, RENDVI, SIPI, Lic1, NDVI, and VRE2, VRE3, Ctr2, TCARI, MDATT, and MCARI1,

respectively. These VIs are sensitive to chlorophyll concentration and nitrogen stress (Penuelas, Baret & Filella, 1995; Lichtenthaler et al., 1996; Haboudane et al., 2004; Jain et al., 2007; Sun et al., 2013; Burns et al., 2022; Vogelmann, Rock & Moss, 1993)□. One of the main traits of tea plant adaptability is the amount of chlorophyll in the leaves, which rises proportional with the amount of nitrogen applied (Qiu et al., 2024)□. Chlorophyll prevention strategies could be developed by genotypes that are tolerant ND. Nitrate levels correspond to vegetation indices that are sensitive to chlorophyll concentration, such as Ctr2, NDVI, RENDVI, and TCARI (Katsoulas et al., 2016; Ihuoma & Madramootoo, 2020)□. The chlorophyll-sensitive VIs used in our study could serve as markers for genotypes that react differentially to nitrogen limitation. Vegetation index PRI, which is higher in tolerant genotypes, describes the intensity of photosynthesis based on the amount of chlorophyll (Xiao et al., 2018)□. Additionally linked to resistant genotypes are the carotenoid pigment-sensitive indicator PSRI and the anthocyanin detectors ARI1/ARI2, which indicate plant senescence or active growth (Merzlyak et al., 1999; Gitelson, Merzlyak & Chivkunova, 2001; Foster et al., 2012; Tayade et al., 2022)□.

Additionally, carotenoid pigment-sensitive indicator PSRI and the anthocyanin detectors ARI1/ARI2, which indicate plant senescence or active growth, were higher in ND-tolerant genotypes. This corresponds with our suggestion that genotypes exhibiting elevated N content also exhibit elevated levels of polyphenols, specifically flavonols, which can be detected using PSRI and ARI1/ARI2. Long-term N treatment was known to increase carotenoid concentration in tea leaves, while ND promotes oxidative stress in plants (Chen et al., 2021b)□. Furthermore, it was shown that anthocyanins and carotenoids are present in stressed vegetation and promote the antioxidant process (Stahl & Sies, 2003; Xiang et al., 2022)□. Thus, it can be suggested that oxidative stress-protective mechanisms are triggered in these genotypes. Some researchers demonstrated that ND enhances stomatal resistance and decreases transpiration, which may have an impact on the leaf water content (Nagarajah, 1981)□. Further evidence showed a significant increase in water use efficiency with increasing leaf N content (Katsoulas et al., 2016)□. This is in line with our findings that NDVI, which correlates with photosynthetic efficiency, biomass, nitrogen and water content, as well as WBI, assigns the cluster of ND-tolerant genotypes, including the VIs mentioned above (Peñuelas et al., 1994; Badzmierowski, McCall & Evanylo, 2019)□. Thus, the genotypes of tea plants that are susceptible or tolerant to ND can be identified using these reflectance light-based indices. A portion of the genotypes are categorized as non-responsive, while some of them exhibit remarkable variations in leaf N-content, and the remainder have an unknown chemical composition. As a result, according to certain vegetation indices, there may not be a substantial difference between genotypes that are susceptible and tolerant. To ascertain how the remaining tea plant varieties would react to a nitrogen deficit, more research is needed to analyze their N content and leaf quality.

Using association analysis between SNPs and VIs, four mutations causing amino changes in the N metabolism and tea quality genes were found. The water content indicator is associated with two mutations in the *4Cl* gene, which codes for 4-coumarate:CoA ligase and is involved in the phenylpropanoid biosynthesis pathway (Li et al., 2022)□. As we previously demonstrated, the

antioxidant polyphenol theaflavin is similarly linked to the identical mutation in *4Cl*. Flavonoids and polyphenols are known for their role in defense against biotic and abiotic stressors including water stress. Water stress has been shown to be a cause of phenolic compound formation, and a decrease in soil water content lowers the phenols content in tea (Cheruiyot et al., 2007; Hodaei et al., 2018). Consequently, WBI has the potential to be utilized as an indirect indicator of phenylpropanoid leaf content. Water and phenol content in leaves could be better understood by investigating how genotypes of nitrogen-efficient tea plants react to water stress. However, the two mutations change amino acids with similar properties (Thr to Ser and Ile to Val), which could have a minor impact on the enzyme structure and functions. Mutations in the *ANSb* and *PIP* genes appeared to be linked to the flavonol content indicator (Merzlyak et al., 2005). Phenolic compounds found in tea, anthocyanins, are synthesized and accumulated with the help of anthocyanidin synthase, which is encoded by *ANSb (*Anggraini et al., 2019; Huang et al., 2022*)*. Moreover, the coloration of anthocyanins is influenced by the increased production of ROS by plasma membrane intrinsic proteins (PIPs), which also control N absorption (Li et al., 2017a; Zhang et al., 2020a; Maritim et al., 2021). However, the vegetation index showed a poor R2, yet the association between both mutations and FRI is evident. The phenotypic data is only known for a portion of the genotypes, whereas the genotype data was obtained from two combined studies. This could have caused some gaps in the data and affected the findings of the association study.

According to the findings of the genotyping study, the genes controlling ammonium transport (*AMT1.2*) and flavonoid pathways (*UFGTa, UFGTb, 4Cl, F3'5'Hb, ANRb-ANR1*) had the highest SNP frequencies across all the varieties, which is in accordance with our previous research. That investigation showed no relationship between the amount of N and the quantity of catechins, whose synthesis is controlled by the genes *ANR* and *4Cl* in tea leaves (Zhang et al., 2020b). On the other hand, a strong positive association was found between the concentration of N leaves and flavanols, specifically L-theanine, theaflavins, and thearubigins, as well as tannins like gallic acid. Earlier research demonstrated that the accumulation of a variety of flavonoids by nitrogen-deficient tea plants correlates with increased expression of genes such as *F3H*, *FNS*, *UFGT*, *bHLH35*, and *bHLH36 (*Huang et al., 2018*)*. Additionally, there is an increase in the expression of dihydroflavonol 4-reductase (*DFR*), anthocyanidin synthase (*ANS*), anthocyanidin reductase 1 (*ANR1*), and 3′,5′-hydroxylase (*F3'5'H)* under conditions of N excess compared to deficiency (Dong et al., 2019). Further research using high-performance liquid chromatography is required to demonstrate the leaf content of proanthocyanidins and how it relates to N content. Hierarchical clustering showed that a group of ND-susceptible genotypes and several ND-tolerant genotypes have more SNPs in *NRT2.4*, *PIP*, *AlaDC*, *DFRa*, and *GS1.2*. The nitrate transporter gene (*NRT2.4*) and aquaporin gene (*PIP*) are responsible for effective N uptake, while the theanine synthesis enzyme (Alanine decarboxylase; *AlaDC*) is crucial for nitrogen storage (Wang et al., 2021b; Xie et al., 2023; Bai et al., 2019, 2021). As well, the theanine pathway and ammonium assimilation are facilitated by glutamine synthetase (*GS1.2*) (Zhang et al., 2023). SNPs in the previously stated genes may account for the notable variations in the

chemical contents in leaves; #316 has the highest theanine and nitrogen content, whereas #1467, #1877, #527, #536, and #507 have the lowest. Mutations that change amino acids in the *AlaAT1* and *MYB4* genes were determined to be specific to susceptible tea genotypes and those characterized by the low leaf N-content. Alanine aminotransferase *AlaAT* plays a role in the biosynthesis and accumulation of L-theanine as well as the efficiency of nitrogen use (Wang et al., 2021a; Zhang et al., 2022). Based on our prior findings, we suggest that this alteration in the structure of the enzyme could cause a decrease in theanine in genotypes *#3986* and *#1467*. The leaf N-content, which is low in susceptible genotypes, is known to be positively linked with flavan-3-ols and other phenolic compounds whose accumulation is inhibited by *MYB4* (Li et al., 2017b; Ye et al., 2021). One mutation in the transcription factor *WRKY57*, which participates in hormone signaling during stressor activation, is also present in accessions with low nitrogen content and insusceptible genotypes (Jiang et al., 2014; Chen et al., 2019, 2021c). However, the role of *WRKY57* in nitrogen stress has yet to be investigated. Combining datasets under different experimental settings presents data integration challenges that could impair accuracy and result in missing values in SNPs positions (Dergilev et al., 2021; Chao et al., 2023). Further phenotype studies and Sanger sequencing will be used to validate the polymorphisms that have been discovered.

Conclusions

We identified efficient vegetation indices to distinguish ND-tolerant and ND-susceptible tea genotypes: ZMI, RENDI, CNDVI, PRI, GM1, GM2, VRI1, VRE3, VRE2, Ctr2. Numerous SNPs that could be exploited for genotyping were discovered. In particular, mutations in *NRT2.4, PIP, AlaDC, DFRa, GS1.2*, *AlaAT1, MYB4*, and *WRKY57* are specific for ND-susceptible tea genotypes. Four associations were detected between the SNPs and vegetation indices. Particularly, water band index (WBI) and far red index (FRI) were associated with SNPs in the flavonoid regulators *4Cl*, *ANSb*, and *PIP*. Further characterization of tea varieties cultivated under ND-conditions, as well as the validation using sequencing and metabolic techniques, could improve the accuracy of detecting genotypes that are tolerant or susceptible to ND. The phenotypic and genetic data obtained in this study could be used in breeding programs aimed at developing nitrogen-efficient tea plants.

Acknowledgements

This study was funded by Russian Science Foundation grant # 22-16-00058. All phenotypic analyses were conducted at Subtropical scientific Center. All genetic analyses were conducted at Sirius University of Science and Technology. The plant material for this study was provided by the program # FGRW-2024-0003.

References

Anggraini T, Wilma S, Syukri D, Azima F. 2019. Total phenolic, anthocyanin, catechins, DPPH radical scavenging activity, and toxicity of *lepisanthes alata (Blume) leenh*. *International Journal of Food Science* **2019**:9703176 DOI: 10.1155/2019/9703176.

Badzmierowski MJ, McCall DS, Evanylo G. 2019. Using hyperspectral and multispectral indices to detect water stress for an urban turfgrass system. *Agronomy* **9**(8):439. DOI: 10.3390/agronomy9080439.

Bai P, Wang L, Wei K, Ruan L, Wu L, He M, Ni D, Cheng H. 2021. Biochemical characterization of specific Alanine Decarboxylase (AlaDC) and its ancestral enzyme Serine Decarboxylase (SDC) in tea plants (Camellia sinensis). *BMC Biotechnology* **21**:17. DOI: 10.1186/s12896-021-00674-x.

Bai P, Wei K, Wang L, Zhang F, Ruan L, Li H, Wu L, Cheng H. 2019. Identification of a novel gene encoding the specialized alanine decarboxylase in tea (*Camellia sinensis*) plants. *Molecules* 24(3):540. DOI: 10.3390/molecules24030540.

Brimson JM, Prasanth MI, Kumaree KK, Thitilertdecha P, Malar DS, Tencomnao T, Prasansuklab A. 2023. Tea Plant (*Camellia sinensis*): A Current Update on Use in Diabetes, Obesity, and Cardiovascular Disease. *Nutrients* **15**(1):37. DOI: 10.3390/nu15010037.

Burns BW, Green VS, Hashem AA, Massey JH, Shew AM, Adviento-Borbe MAA, Milad M. 2022. Determining nitrogen deficiencies for maize using various remote sensing indices. *Precision Agriculture* **23**:791–811. DOI: 10.1007/s11119-021-09861-4.

Cao Q, Yang G, Duan D, Chen L, Wang F, Xu B, Zhao C, Niu F. 2022. Combining multispectral and hyperspectral data to estimate nitrogen status of tea plants (Camellia sinensis (L.) O. Kuntze) under field conditions. *Computers and Electronics in Agriculture* 198:107084. DOI: 10.1016/j.compag.2022.107084.

Chan EWC, Soh EY, Tie PP, Law YP. 2011. Antioxidant and antibacterial properties of green, black, and herbal teas of Camellia sinensis. *Pharmacognosy Research* **33**(4):266-272. DOI: 10.4103/0974-8490.89748.

Chao H, Zhang S, Hu Y, Ni Q, Xin S, Zhao L, Ivanisenko V, Orlov Y, Chen M. 2023. Integrating omics databases for enhanced crop breeding. *Journal of Integrative Bioinformatics* 20(4):20230012. DOI: 10.1515/jib-2023-0012.

Chen S, Gao Y, Fan K, Shi Y, Luo D, Shen J, Ding Z, Wang Y. 2021a. Prediction of Drought-Induced Components and Evaluation of Drought Damage of Tea Plants Based on Hyperspectral Imaging. *Frontiers in Plant Science* **12**:695102. DOI: 10.3389/fpls.2021.695102.

Chen W, Hao WJ, Xu YX, Zheng C, Ni DJ, Yao MZ, Chen L. 2019. Isolation and characterization of CsWRKY7, a subgroup IId WRKY transcription factor from Camellia sinensis, linked to development in Arabidopsis. *International Journal of Molecular Sciences* **20**(11):2815. DOI: 10.3390/ijms20112815.

Chen T, Ma J, Li H, Lin S, Dong C, Xie Y, Yan X, Zhang S, Yang T, Wan X, Zhang Z. 2023. CsGDH2.1 negatively regulates theanine accumulation in late-spring tea plants (Camellia sinensis var. sinensis). *Horticulture Research* **10**(1):uhac245, DOI: 10.1093/hr/uhac245.

Chen Y, Wang F, Wu Z, Jiang F, Yu W, Yang J, Chen J, Jian G, You Z, Zeng L. 2021b. Effects of long-term nitrogen fertilization on the formation of metabolites related to tea quality in subtropical China. *Metabolites* **11**(3):146. DOI: 10.3390/metabo11030146.

Chen L, Zhao L, Ma C, Wang X, Yao M. 2009. Differential expression analysis of β-glucosidase and β-primeverosidase genes in tea plant (Camellia sinensis). *Acta Horticulturae Sinica* **36**(1):87–92.

Chen W, Zheng C, Yao M, Chen L. 2021c. The tea plant CsWRKY26 promotes drought tolerance in transgenic Arabidopsis plants. *Beverage Plant Research* **1**:3. DOI: 10.48130/BPR-2021-0003.

Cheruiyot EK, Mumera LM, Ngetich WK, Hassanali A, Wachira F. 2007. Polyphenols as potential indicators for drought tolerance in tea (Camellia sinensis L.). *Bioscience, Biotechnology and Biochemistry* **71**(9):2190–2197. DOI: 10.1271/bbb.70156.

Dergilev AI, Orlova NG, Dobrovolskaya OB, Orlov YL. 2021. Statistical estimates of multiple transcription factors binding in the model plant genomes based on ChIP-seq data. *Journal of Integrative Bioinformatics* **19**(1):20200036. DOI:10.1515/jib-2020-0036

Dong F, Hu J, Shi Y, Liu M, Zhang Q, Ruan J. 2019. Effects of nitrogen supply on flavonol glycoside biosynthesis and accumulation in tea leaves (Camellia sinensis). *Plant Physiology and Biochemistry* **138**:48-57. DOI: 10.1016/j.plaphy.2019.02.017.

Filippini T, Malavolti M, Borrelli F, Izzo AA, Fairweather-Tait SJ, Horneber M, Vinceti M. 2020. Green tea (Camellia sinensis) for the prevention of cancer. *Cochrane Database of Systematic Reviews* 3:CD005004. DOI: 10.1002/14651858.CD005004.pub3.

Foster AJ, Kakani VG, Ge J, Mosali J. 2012. Discrimination of switchgrass cultivars and nitrogen treatments using pigment profiles and hyperspectral leaf reflectance data. *Remote Sensing* **4**(9):2576-2594. DOI: 10.3390/rs4092576.

Frels K, Guttieri M, Joyce B, Leavitt B, Baenziger PS. 2018. Evaluating canopy spectral reflectance vegetation indices to estimate nitrogen use traits in hard winter wheat. *Field Crops Research* **217**:82-92. DOI: 10.1016/j.fcr.2017.12.004.

Gai Z, Wang Y, Jiang J, Xie H, Ding Z, Ding S, Wang H. 2019. The quality evaluation of tea (camellia sinensis) varieties based on the metabolomics. *HortScience horts* **54**(3):409-415. DOI: 10.21273/HORTSCI13713-18.

Gao Y, Cabrera Serrenho A. 2023. Greenhouse gas emissions from nitrogen fertilizers could be reduced by up to one-fifth of current levels by 2050 with combined interventions. *Nature Food* **4**(2):170-178. DOI: 10.1038/s43016-023-00698-w.

Gitelson AA, Merzlyak MN, Chivkunova OB. 2001. Optical Properties and Nondestructive Estimation of Anthocyanin Content in Plant Leaves. *Photochemistry and Photobiology* **74**(1):38-45. DOI: 10.1562/0031-8655(2001)074<0038:opaneo>2.0.co;2.

Guo L, Gao L, Ma X, Guo F, Ruan H, Bao Y, Xia T, Wang Y. 2019. Functional analysis of flavonoid 3′-hydroxylase and flavonoid 3′,5′-hydroxylases from tea plant (Camellia sinensis), involved in the B-ring hydroxylation of flavonoids. *Gene* **717**:144046. DOI: 10.1016/j.gene.2019.144046.

Haboudane D, Miller JR, Pattey E, Zarco-Tejada PJ, Strachan IB. 2004. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment* **90**(3):337-352. DOI: 10.1016/j.rse.2003.12.013.

Hazra A, Dasgupta N, Sengupta C, Das S. 2018. Next generation crop improvement program: Progress and prospect in tea (*Camellia sinensis* (L.) O. Kuntze). *Annals of Agrarian Science* **16**(2):128-135. DOI: 10.1016/j.aasci.2018.02.002.

Hodaei M, Rahimmalek M, Arzani A, Talebi M. 2018. The effect of water stress on phytochemical accumulation, bioactive compounds and expression of key genes involved in flavonoid biosynthesis in *Chrysanthemum morifolium* L. *Industrial Crops and Products* **120**:295-304. DOI: 10.1016/j.indcrop.2018.04.073.

Huang F, Duan J, Lei Y, Kang Y, Luo Y, Chen Y, Ding D, Li S. 2022. Metabolomic and transcriptomic analyses reveal a MYB gene, CsAN1, involved in anthocyanins accumulation separation in F1 between ‘Zijuan’ (Camellia sinensis var. assamica) and ‘Fudingdabaicha’ (C. sinensis var. sinensis) tea plants. *Frontiers in Plant Science* 13. DOI: 10.3389/fpls.2022.1008588.

Huang H, Yao Q, Xia E, Gao L. 2018. Metabolomics and Transcriptomics Analyses Reveal Nitrogen Influences on the Accumulation of Flavonoids and Amino Acids in Young Shoots

of Tea Plant (Camellia sinensis L.) Associated with Tea Flavor. *Journal of Agricultural and Food Chemistry* **66**(37):9828–9838. DOI: 10.1021/acs.jafc.8b01995.

Ihuoma SO, Madramootoo CA. 2020. Narrow-band reflectance indices for mapping the combined effects of water and nitrogen stress in field grown tomato crops. *Biosystems Engineering* 192. DOI: 10.1016/j.biosystemseng.2020.01.017.

Jain N, Ray SS, Singh JP, Panigrahy S. 2007. Use of hyperspectral data to assess the effects of different nitrogen applications on a potato crop. *Precision Agriculture* 8. DOI: 10.1007/s11119-007-9042-0.

Jiang Y, Liang G, Yang S, Yu D. 2014. Arabidopsis WRKY57 functions as a node of convergence for jasmonic acid- and auxin-mediated signaling in jasmonic acid-induced leaf senescence. *Plant Cell* 26. DOI: 10.1105/tpc.113.117838.

Kamran M, Yan Z, Chang S, Ning J, Lou S, Ahmad I, Ghani MU, Arif M, El Sabagh A, Hou F. 2023. Interactive effects of reduced irrigation and nitrogen fertilization on resource use efficiency, forage nutritive quality, yield, and economic benefits of spring wheat in the arid region of Northwest China. *Agricultural Water Management* 275. DOI: 10.1016/j.agwat.2022.108000.

Katsoulas N, Elvanidi A, Ferentinos KP, Kacira M, Bartzanas T, Kittas C. 2016. Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review. *Biosystems Engineering* 151. DOI: 10.1016/j.biosystemseng.2016.10.003.

Kior A, Sukhov V, Sukhova E. 2021. Application of reflectance indices for remote sensing of plants and revealing actions of stressors. *Photonics* 8. DOI: 10.3390/photonics8120582.

Li F, Dong C, Yang T, Bao S, Fang W, Lucas WJ, Zhang Z. 2021. The tea plant CsLHT1 and CsLHT6 transporters take up amino acids, as a nitrogen source, from the soil of organic tea plantations. *Horticulture Research* 8. DOI: 10.1038/s41438-021-00615-x.

Li M, Guo L, Wang Y, Li Y, Jiang X, Liu Y, Xie DY, Gao L, Xia T. 2022. Molecular and biochemical characterization of two 4-coumarate: CoA ligase genes in tea plant (Camellia sinensis). *Plant Molecular Biology* 109. DOI: 10.1007/s11103-022-01269-6.

Li Y, Huang J, Song X, Zhang Z, Jiang Y, Zhu Y, Zhao H, Ni D. 2017a. An RNA-Seq transcriptome analysis revealing novel insights into aluminum tolerance and accumulation in tea plant. *Planta* 246. DOI: 10.1007/s00425-017-2688-6.

Li M, Li Y, Guo L, Gong N, Pang Y, Jiang W, Liu Y, Jiang X, Zhao L, Wang Y, Xie DY, Gao L, Xia T. 2017b. Functional characterization of tea (Camellia sinensis) MYB4a

transcription factor using an integrative approach. *Frontiers in Plant Science* 8. DOI: 10.3389/fpls.2017.00943.

Li J, Liu X, Xu L, Li W, Yao Q, Yin X, Wang Q, Tan W, Xing W, Liu D. 2023. Low nitrogen stress-induced transcriptome changes revealed the molecular response and tolerance characteristics in maintaining the C/N balance of sugar beet (Beta vulgaris L.). *Frontiers in Plant Science* 14. DOI: 10.3389/fpls.2023.1164151.

Li W, Xiang F, Zhong M, Zhou L, Liu H, Li S, Wang X. 2017c. Transcriptome and metabolite analysis identifies nitrogen utilization genes in tea plant (Camellia sinensis). *Scientific Reports* 7. DOI: 10.1038/s41598-017-01949-0.

Li ZX, Yang WJ, Ahammed GJ, Shen C, Yan P, Li X, Han WY. 2016. Developmental changes in carbon and nitrogen metabolism affect tea quality in different leaf position. *Plant Physiology and Biochemistry* 106. DOI: 10.1016/j.plaphy.2016.06.027.

Lichtenthaler HK, Lang M, Sowinska M, Heisel F, Miehé JA. 1996. Detection of vegetation stress via a new high resolution fluorescence imaging system. *Journal of Plant Physiology* 148. DOI: 10.1016/S0176-1617(96)80081-2.

Liu Y, Hou H, Jiang X, Wang P, Dai X, Chen W, Gao L, Xia T. 2018. A WD40 repeat protein from camellia sinensis regulates anthocyanin and proanthocyanidin accumulation through the formation of MYB–bHLH–WD40 ternary complexes. *International Journal of Molecular Sciences* 19. DOI: 10.3390/ijms19061686.

Liu M, Tian HL, Wu JH, Cang RR, Wang RX, Qi XH, Xu Q, Chen XH. 2015. Relationship between gene expression and the accumulation of catechin during spring and autumn in tea plants (Camellia sinensis L.). *Horticulture Research* 2. DOI: 10.1038/hortres.2015.11.

Liu Y, Zhang M, Li Y, Zhang Y, Huang X, Yang Y, Zhu H, Xiong H, Jiang T. 2023. Influence of Nitrogen Fertilizer Application on Soil Acidification Characteristics of Tea Plantations in Karst Areas of Southwest China. *Agriculture (Switzerland)* 13. DOI: 10.3390/agriculture13040849.

Luo D, Gao Y, Wang Y, Shi Y, Chen S, Ding Z, Fan K. 2022. Using UAV image data to monitor the effects of different nitrogen application rates on tea quality. *Journal of the Science of Food and Agriculture* 102. DOI: 10.1002/jsfa.11489.

Malyukova LS, Koninskaya NG, Orlov YL, Samarina LS. 2022. Effects of exogenous calcium on the drought response of the tea plant (*Camellia sinensis* (L.) Kuntze). *PeerJ*. 10:e13997. DOI:10.7717/peerj.13997

Mao Y, Li H, Wang Y, Wang H, Shen J, Xu Y, Ding S, Wang H, Ding Z, Fan K. 2023. Rapid monitoring of tea plants under cold stress based on UAV multi-sensor data. *Computers and Electronics in Agriculture* 213. DOI: 10.1016/j.compag.2023.108176.

Maritim TK, Korir RK, Nyabundi KW, Wachira FN, Kamunya SM, Muoki RC. 2021. Molecular regulation of anthocyanin discoloration under water stress and high solar irradiance in pluckable shoots of purple tea cultivar. *Planta* 254. DOI: 10.1007/s00425-021-03736-8.

Merzlyak MN, Gitelson AA, Chivkunova OB, Rakitin VY. 1999. Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum* 106. DOI: 10.1034/j.1399-3054.1999.106119.x.

Merzlyak MN, Solovchenko AE, Smagin AI, Gitelson AA. 2005. Apple flavonols during fruit adaptation to solar radiation: Spectral features and technique for non-destructive assessment. *Journal of Plant Physiology* 162. DOI: 10.1016/j.jplph.2004.07.002.

Nagarajah S. 1981. The effect of nitrogen on plant water relations in tea (Camellia sinensis). *Physiologia Plantarum* 51. DOI: 10.1111/j.1399-3054.1981.tb04482.x.

Penuelas J, Baret F, Filella I. 1995. Semi-empirical indices to assess carotenoids/chlorophyll a ratio from leaf spectral reflectance. *Photosynthetica* 31.

Peñuelas J, Gamon JA, Fredeen AL, Merino J, Field CB. 1994. Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sensing of Environment* 48. DOI: 10.1016/0034-4257(94)90136-8.

Prey L, Hu Y, Schmidhalter U. 2020. High-Throughput Field Phenotyping Traits of Grain Yield Formation and Nitrogen Use Efficiency: Optimizing the Selection of Vegetation Indices and Growth Stages. *Frontiers in Plant Science* 10. DOI: 10.3389/fpls.2019.01672.

Qiu Z, Liao J, Chen J, Li A, Lin M, Liu H, Huang W, Sun B, Liu J, Liu S, Zheng P. 2024. Comprehensive analysis of fresh tea (Camellia sinensis cv. Lingtou Dancong) leaf quality under different nitrogen fertilization regimes. *Food Chemistry* 439:138127. DOI: https://doi.org/10.1016/j.foodchem.2023.138127.

Samanta S. 2020. Potential Bioactive Components and Health Promotional Benefits of Tea (Camellia sinensis). *Journal of the American College of Nutrition*. DOI: 10.1080/07315724.2020.1827082.

Samarina LS, Bobrovskikh AV, Doroshkov AV, Malyukova LS, Matskiv AO, Rakhmangulov RS, Koninskaya NG, Malyarovskaya VI, Tong W, Xia E, Manakhova KA, Ryndin AV, Orlov YL. 2020. Comparative Expression Analysis of Stress-Inducible

643 Candidate Genes in Response to Cold and Drought in Tea Plant [Camellia sinensis (L.)
644 Kuntze]. *Frontiers in Genetics* **11**:611283. DOI: 10.3389/fgene.2020.611283

645 Samarina L, Fedorina J, Kuzmina D, Malyukova L, Manakhova K, Kovalenko T, Matskiv
646 A, Xia E, Tong W, Zhang Z, Ryndin A, Orlov YL, Khlestkina EK. 2023. Analysis of
647 Functional Single-Nucleotide Polymorphisms (SNPs) and Leaf Quality in Tea Collection
648 under Nitrogen-Deficient Conditions. *International Journal of Molecular Sciences* 24. DOI:
649 10.3390/ijms241914538.

650 Sánchez M, González-Burgos E, Iglesias I, Lozano R, Gómez-Serranillos MP. 2020. The
651 pharmacological activity of camellia sinensis (L.) kuntze on metabolic and endocrine
652 disorders: A systematic review. *Biomolecules* 10. DOI: 10.3390/biom10040603.

653 Stahl W, Sies H. 2003. Antioxidant activity of carotenoids. *Molecular Aspects of Medicine*
654 24. DOI: 10.1016/S0098-2997(03)00030-X.

655 Sun H, Li M, Zhang Y, Zheng L, Zhang Y. 2013. Correlation between chlorophyll content
656 and vegetation index of maize plants under different fertilizer treatments with multi-spectral
657 imaging. *Sensor Letters* 11. DOI: 10.1166/sl.2013.2888.

658 Tang D, Jiao Z, Zhang Q, Liu MY, Ruan J. 2021. Glutamate dehydrogenase isogenes
659 CsGDHs cooperate with glutamine synthetase isogenes CsGSs to assimilate ammonium in
660 tea plant (Camellia sinensis L.). *Plant Science* 312. DOI: 10.1016/j.plantsci.2021.111031.

661 Tang S, Zhou J, Pan W, Sun T, Liu M, Tang R, Li Z, Ma Q, Wu L. 2023. Effects of
662 combined application of nitrogen, phosphorus, and potassium fertilizers on tea (Camellia
663 sinensis) growth and fungal community. *Applied Soil Ecology* 181. DOI:
664 10.1016/j.apsoil.2022.104661.

665 Tayade R, Yoon J, Lay L, Khan AL, Yoon Y, Kim Y. 2022. Utilization of Spectral Indices
666 for High-Throughput Phenotyping. *Plants* 11. DOI: 10.3390/plants11131712.

667 Vogelmann J, Rock B, Moss DM. 1993. Red edge spectral measurements from sugar maple
668 leaves. *Int. J. Remote Sens.* **14**(8):1563-1575. DOI: 10.1080/01431169308953986.

669 Wang J, Chen W, Li Y, Zhao X, Xiang Q, Li M. 2021a. Identification and Partial
670 Characterization of an Alanine Aminotransferase Gene CsAlaAT2 in Response to Shade in
671 Tea (Camellia sinensis). *Journal of Plant Biology* 64. DOI: 10.1007/s12374-020-09273-3.

672 Wang Y, Cheng X, Yang T, Su Y, Lin S, Zhang S, Zhang Z. 2021b. Nitrogen-Regulated
673 Theanine and Flavonoid Biosynthesis in Tea Plant Roots: Protein-Level Regulation
674 Revealed by Multiomics Analyses. *Journal of Agricultural and Food Chemistry* 69. DOI:
675 10.1021/acs.jafc.1c02589.

Wani SH, Anand S, Singh B, Bohra A, Joshi R. 2021. WRKY transcription factors and plant defense responses: latest discoveries and future prospects. *Plant Cell Reports* **40**(7):1071-1085. DOI: 10.1007/s00299-021-02691-8.

Xia E, Tong W, Hou Y, An Y, Chen L, Wu Q, et al. 2020. The Reference Genome of Tea Plant and Resequencing of 81 Diverse Accessions Provide Insights into Its Genome Evolution and Adaptation. *Molecular Plant* **13**(7):1013-1026. DOI: 10.1016/j.molp.2020.04.010.

Xiang F, Zhou L, Liu H, Li W. 2022. Improving Tea Quality by Balancing ROS and Antioxidant System through Appropriate Ammonium Nitrogen Application. *Sustainability (Switzerland)* **14**(15):9354. DOI: 10.3390/su14159354.

Xiao F, Yang Z, Huang H, Yang F, Zhu L, Han D. 2018. Nitrogen fertilization in soil affects physiological characteristics and quality of Green Tea Leaves. *HortScience* **53**(5): 715–722. DOI: 10.21273/HORTSCI12897-18.

Xie X, Sun Z, Zhang X, Han X. 2023. Novel Aspects of Regulation of Nitrogen Responses in the Tea Plant (Camellia sinensis (L.)). *Agronomy* 13. DOI: 10.3390/agronomy13010144.

Yang T, Li H, Tai Y, Dong C, Cheng X, Xia E, Chen Z, Li F, Wan X, Zhang Z. 2020. Transcriptional regulation of amino acid metabolism in response to nitrogen deficiency and nitrogen forms in tea plant root (Camellia sinensis L.). *Scientific Reports* **10**(1):6868. DOI: 10.1038/s41598-020-63835-6.

Yang Y, Wang F, Wan Q, Ruan J. 2018. Transcriptome analysis using RNA-Seq revealed the effects of nitrogen form on major secondary metabolite biosynthesis in tea (Camellia sinensis) plants. *Acta Physiologiae Plantarum* **40**:127. DOI: 10.1007/s11738-018-2701-0.

Ye JH, Lv YQ, Liu SR, Jin J, Wang YF, Wei CL, Zhao SQ. 2021. Effects of light intensity and spectral composition on the transcriptome profiles of leaves in shade grown tea plants (Camellia sinensis l.) and regulatory network of flavonoid biosynthesis. *Molecules* **26**(19):5836. DOI: 10.3390/molecules26195836.

Zhang X, Liu H, Pilon-Smits E, Huang W, Wang P, Wang M, Guo F, Wang Y, Li R, Zhao H, Ni D. 2020a. Transcriptome-wide analysis of nitrogen-regulated genes in tea plant (Camellia sinensis l. o. kuntze) and characterization of amino acid transporter cscat9.1. *Plants* **9**(9):1218. DOI: 10.3390/plants9091218.

Zhang W, Ni K, Long L, Ruan J. 2023. Nitrogen transport and assimilation in tea plant (Camellia sinensis): a review. *Frontiers in Plant Science* **14**:1249202. DOI: 10.3389/fpls.2023.1249202.

Zhang Y, Wang L, Wei K, Ruan L, Wu L, He M, Tong H, Cheng H. 2020b. Differential regulatory mechanisms of secondary metabolites revealed at different leaf positions in two related tea cultivars. *Scientia Horticulturae* **272**:109579. DOI: 10.1016/j.scienta.2020.109579.

Zhang X, Wen B, Zhang Y, Li Y, Yu C, Peng Z, Wang K, Liu Z, Huang J an, Xiong L, Li J. 2022. Transcriptomic and biochemical analysis reveal differential regulatory mechanisms of photosynthetic pigment and characteristic secondary metabolites between high amino acids green-leaf and albino tea cultivars. *Scientia Horticulturae* **295**:110823. DOI: 10.1016/j.scienta.2021.110823.

Zhao LQ, Shan CM, Shan TY, Li QL, Ma KL, Deng WW, Wu JW. 2022a. Comparative transcriptomic analysis reveals the regulatory mechanisms of catechins synthesis in different cultivars of Camellia sinensis. *Food Research International* **157**:111375. DOI: 10.1016/j.foodres.2022.111375.

Zhao X, Zhang J, Huang Y, Tian Y, Yuan L. 2022b. Detection and discrimination of disease and insect stress of tea plants using hyperspectral imaging combined with wavelet analysis. *Computers and Electronics in Agriculture* **193**:106717. DOI: 10.1016/j.compag.2022.106717.

Figure 1

Figure 1. Principal component analysis of 31 vegetation indices in 33 tea genotypes with different responses to nitrogen deficit.

PCA biplot illustrates the related traits and correlations between vegetation indices and genotypes. The first two PCs exhibited a cumulative variation of approximately 73.03%. Both the ND-susceptible and ND-tolerant genotypes were clearly separated in the biplot. Most of the vectors of VIs were distributed with high loading on the positive side of PC1 and the negative side of PC2. The highest loading was observed in the following VIs: ZMI, VRE1, RENDVI, CNDVI, PRI, PSRI, GM2, GM1, and CRI2, NDVI, SIPI, and CRI1.

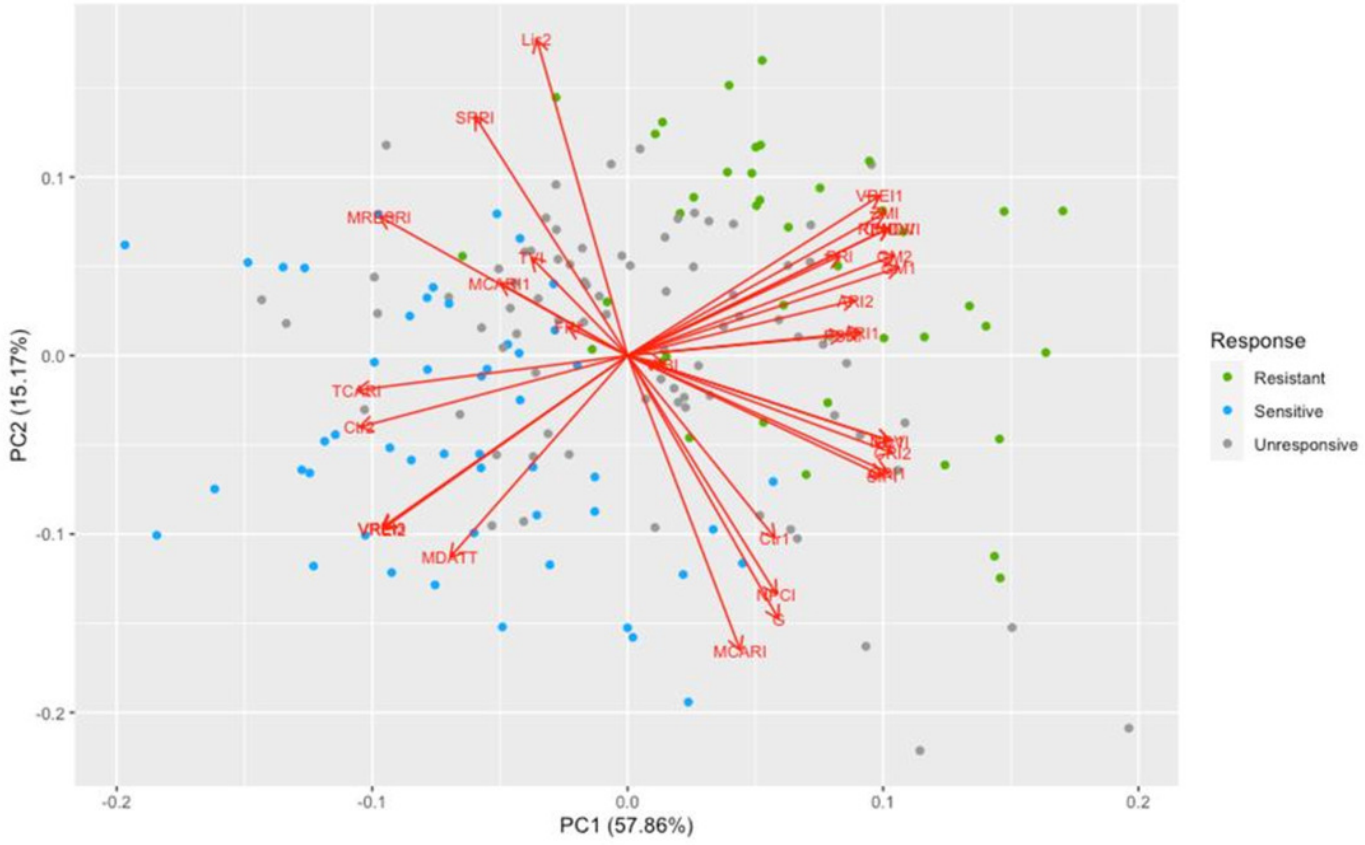

Figure 2

Figure 2. Data points distributions and prediction in the tea genotypes with different responses to nitrogen deficit.

Green - tolerant genotypes; blue - susceptible genotypes; gray - non-responsive

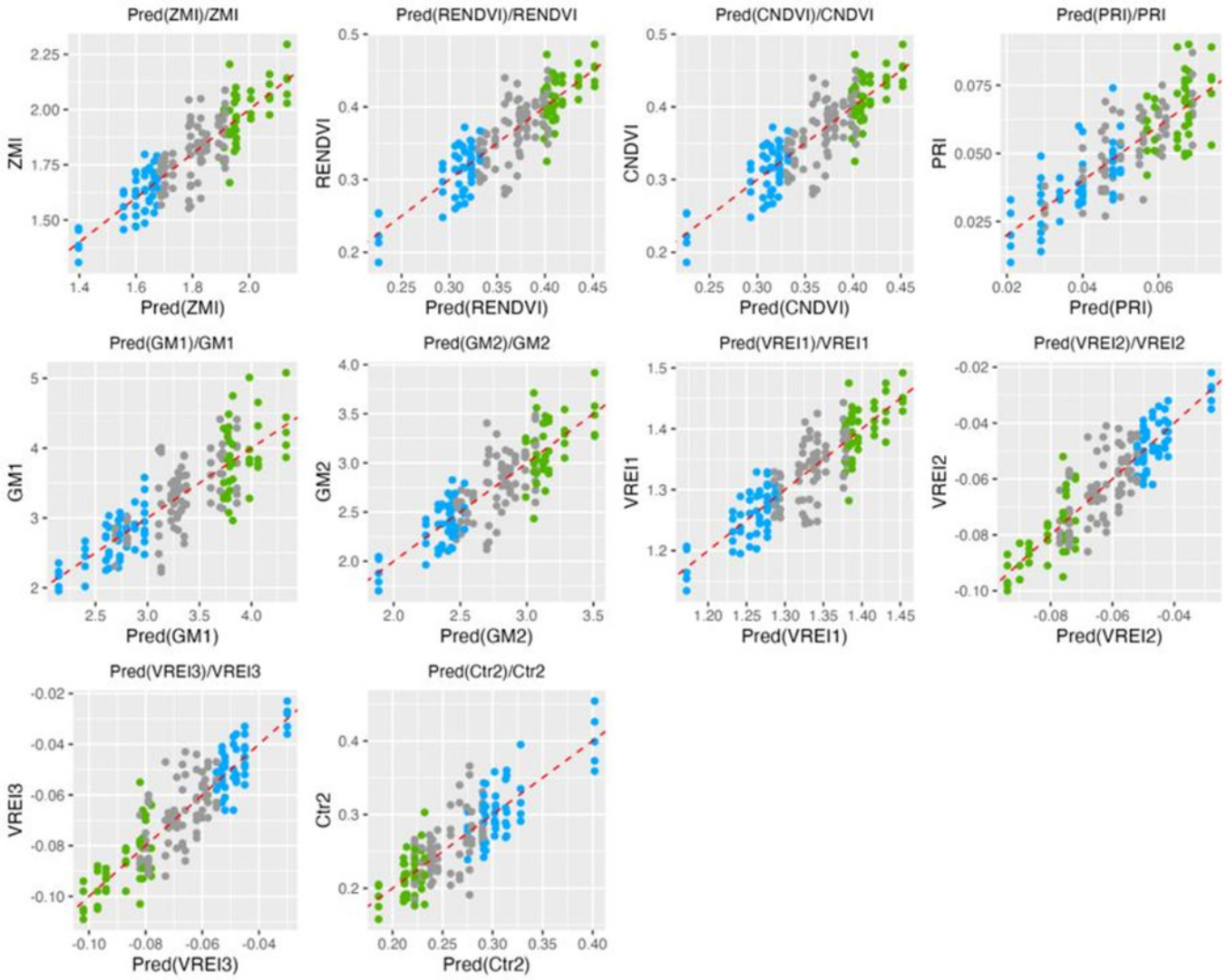

Figure 3

Figure 3. SNPs effect (A) and impact (B) on the phenotypes of 34 tea genotypes.

SNPs effect (A) and impact (B) on the phenotypes of 34 tea genotypes.

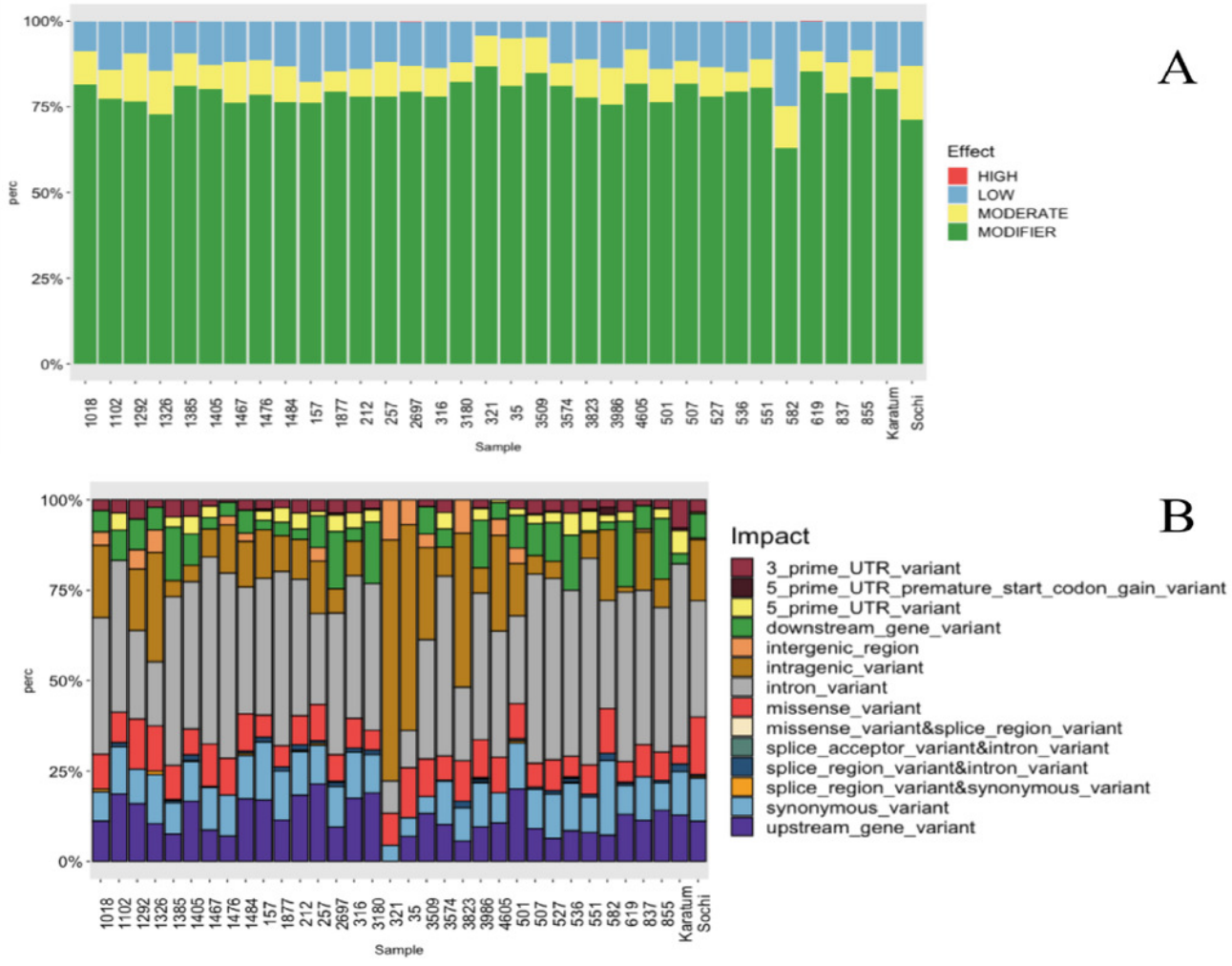

Figure 4

Figure 4. Heatmap of exon SNPs frequencies.

The columns represent the tea genotypes, and the rows represent the different genes. Green frames indicate ND-tolerant genotypes, blue frames – ND-susceptible genotypes.

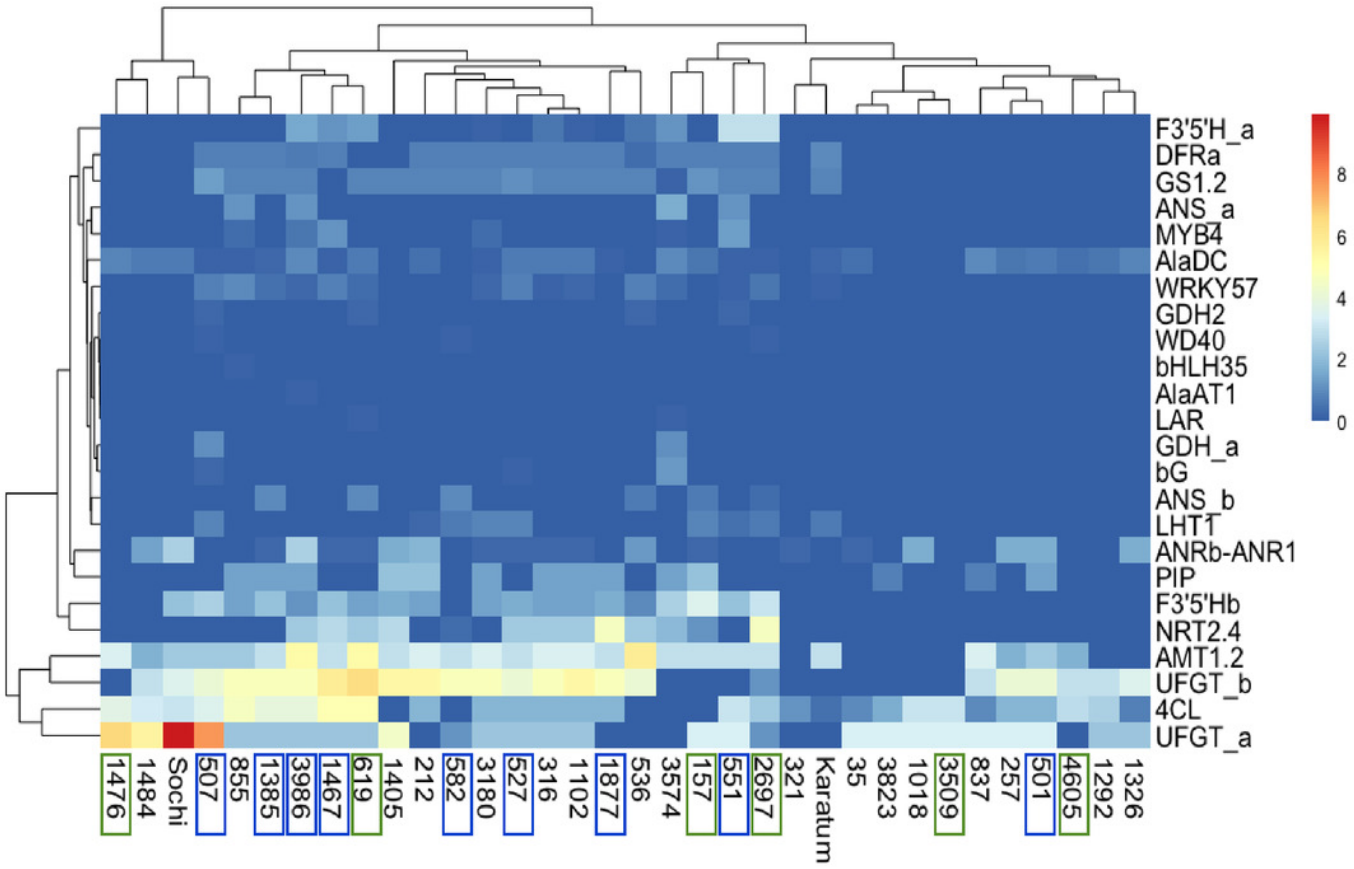

Table 1(on next page)

Table 1. Determination coefficients of vegetation indices at *p* value < 0.05 and determination coefficient R2 > 0.5.

Determination coefficients of vegetation indices at *p* value < 0.05 and determination coefficient R2 > 0.5.

VI	R2	F	Pr > F
CNDVI	0.728	11.487	< 0.0001
RENDVI	0.728	11.487	< 0.0001
SIPI	0.479	3.928	< 0.0001
NDVI	0.528	4.792	< 0.0001
MDATT	0.624	7.102	< 0.0001
Lic1	0.528	4.792	< 0.0001
Ctr2	0.683	9.224	< 0.0001
ARI2	0.652	8.022	< 0.0001
TCARI	0.635	7.451	< 0.0001
MCARI1	0.548	5.198	< 0.0001
WBI	0.624	7.115	< 0.0001
ZMI	0.744	12.437	< 0.0001
VREI1	0.759	13.501	< 0.0001
GM2	0.695	9.757	< 0.0001
GM1	0.674	8.849	< 0.0001
PSRI	0.603	6.508	< 0.0001
PRI	0.674	8.864	< 0.0001
VREI3	0.770	14.357	< 0.0001
VREI2	0.771	14.333	< 0.0001
ARI2	0.625	7.136	< 0.0001

Table 1. Determination coefficients of vegetation indices at *p* value < 0.05 and determination coefficient R2 > 0.5.

Table 2(on next page)

Table 2. The distribution of SNPs in 34 distinct varieties of tea in the exon and intron regions of the target genes (N = 20).

The distribution of SNPs in 34 distinct varieties of tea in the exon and intron regions of the target genes (N = 20).

Gene	Fragment Length, bp	Mean SNPs Number in Introns	Mean SNPs Number in Exons	SNP Density in Introns	SNP Density in Exons	SNP % in Exons
4CL	5264	32,94	9,17	6,26	1,74	21,78
AMT1,2	2643	5,40	4,29	2,04	1,62	44,25
AlaAT1	8058	1,06	0,03	0,13	0,00	2,63
AlaDC	7227	6,83	2,97	0,94	0,41	30,32
DFRa	6600	2,54	2,49	0,39	0,38	49,43
F3'5'H_a	5118	6,57	1,86	1,28	0,36	22,03
F3'5'Hb	4435	3,66	4,54	0,82	1,02	55,40
GDH2	4915	2,57	0,11	0,52	0,02	4,26
GS1,2	6202	33,11	2,77	5,34	0,45	7,72
LAR	8600	0,31	0,06	0,04	0,01	15,38
LHT1	5107	2,00	0,80	0,39	0,16	28,57
MYB4	5342	5,00	0,60	0,94	0,11	10,71
MYB7	3376	0,03	0,00	0,01	0,00	0,00
NRT2,4	3060	6,03	2,37	1,97	0,77	28,23
PIP	2006	1,37	0,83	0,68	0,41	37,66
WD40	3844	2,34	0,34	0,61	0,09	12,77
WRKY57	11214	23,91	2,46	2,13	0,22	9,32
bG	8605	0,11	0,37	0,01	0,04	76,47
bHLH35	5743	1,91	0,03	0,33	0,00	1,47
bHLH36	2953	24,23	0,00	8,20	0,00	0,00

Table 2:
The distribution of SNPs in 34 distinct varieties of tea in the exon and intron regions of the target genes (N = 20).

Table 3(on next page)

Table 3. Significant associations between SNPs and the phenotypes (at *p* value < 0.05).

N (REF/ALT) — nucleotide change (reference/alternative),VI — vegetation indice, DF — Degrees of Freedom, Adj.P-value — Bonferroni-corrected P-value

Gene	Position	N (REF/ ALT)	VI	Amino (REF/ ALT)	Property (REF/ALT)	Test Statistics	DF	Exact P-value	Adj. P-value
4CL	2130421	A/T	WBI	p.Thr16Ser	Neutral/Neutral	4.188505	23	0.0003519712	0.03554909
4CL	2132938	A/G	WBI	p.Ile417Val	Hydrophobic/Hydrophobic	4.188505	23	0.0003519712	0.03554909
PIP	220018	A/T	FRI	p.Val182Glu	Hydrophobic/Charged_acidic	-7.629861	14	2.366524e-06	0.0002390189
ANSb	805981	A/G	FRI	p.Phe39Leu	Aromatic/Hydrophobic	-7.32134	13	5.820357e-06	0.0005878 56

Table 3:

Significant associations between SNPs and the phenotypes at *p* value < 0.05.

N (REF/ALT) — nucleotide change (reference/alternative),VI — vegetation indice, DF — Degrees of Freedom, Adj.P-value — Bonferroni-corrected P-value