TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2018) 42: 1-11
© TÜBİTAK
doi:10.3906/biy-1606-54

Molecular and in silico cloning, identification, and preharvest period expression analysis of a putative cytochrome P450 monooxygenase gene from *Camellia sinensis* (L.) Kuntze (tea)

Ayşenur EMİNOĞLU[1,*], Yeşim AKTÜRK DİZMAN[1], Şule GÜZEL[2], Ali Osman BELDÜZ[3]

[1]Molecular Biology Research Laboratories, Department of Biology, Recep Tayyip Erdoğan University, Rize, Turkey
[2]Plant Ecology Research Laboratories, Department of Biology, Recep Tayyip Erdoğan University, Rize, Turkey
[3]Department of Biology, Faculty of Science, Karadeniz Technical University, Trabzon, Turkey

Received: 17.06.2016 • **Accepted/Published Online:** 01.11.2016 • **Final Version:** 15.02.2018

Abstract: Cytochrome P450 monooxygenases are one of the largest heme-containing protein groups, and the majority of them catalyze hydroxylation reactions dependent on nicotinamide adenine dinucleotide phosphate and oxygen. Cytochrome P450 (CYP) enzymes function in a wide range of monooxygenation reactions essential in primary and secondary metabolism in plants. *Camellia sinensis* (L.) Kuntze is a commercially and economically valuable plant due to its medicinally important secondary metabolites and as a beloved beverage. Cytochrome P450 monooxygenases play a significant role in the biosynthesis of a variety of secondary metabolites in tea. Although the biosynthesis of secondary metabolites has been investigated in detail, there have been limited studies conducted on identifying the genetic mechanisms of CYP-catalyzed secondary metabolic pathways in the *C. sinensis* (tea) plant. In our study, we characterized a putative *C. sinensis* (L.) Kuntze cytochrome P450 monooxygenase gene (Csp450), which has 1759 bp full-length cDNA with 49 bp of 5′ and 183 bp of 3′ untranslated regions. The CDS of the gene is 1527 bp and 508 amino acids in length. BLAST results of the deduced amino acid sequence revealed a high similarity with the CYP704C1-like superfamily. Preharvest period gene expression analysis from May, July, and September did not show any difference.

Key words: *Camellia sinensis* (L.) Kuntze (tea), cytochrome P450 monooxygenase, molecular and in silico cloning, seasonal gene expression

1. Introduction

Tea (*Camellia sinensis* (L.) Kuntze) is a commercially and medicinally important crop due to its secondary metabolites and as a beloved beverage with its aromatic, unique taste. *C. sinensis* leaves contain significant amounts of secondary metabolites (Chen et al., 2012), such as polyphenols (e.g., tannin), alkaloids (e.g., caffeine), vitamins (A, B1, B2, E, C), polysaccharides, and volatile oils. Tea polyphenols mainly consist of catechins, flavanones, and phenolic acid (Park et al., 2004). These polyphenolic compounds constitute approximately 30% of the dry weight of young tea leaves. They have significant effects on human health, besides having important functions in plant physiology (Singh et al., 2009a).

The shikimate pathway is the key metabolic pathway in the biochemical synthesis of aromatic metabolites in plants, which, in turn, function as signaling molecules, hormones, and structural compounds like suberin, lignin, condensed tannins, sporopollenin, and some defense compounds against pathogens and UV light (Werck-Reichhart, 1995; Weaver et al., 1997). The precursor substances of the aromatic compounds of phenylalanine, tyrosine, tryptophan, and chorismate, which are also the precursors of a variety of other secondary metabolites, are provided by the shikimate pathway (Weaver et al., 1997). In plants, the shikimate pathway is the starting point for the biosynthesis of phenylpropanoids.

Cytochrome P450 monooxygenases (CYPs) are heme-containing enzymes that catalyze hydroxylation reactions dependent on nicotinamide adenine dinucleotide phosphate and oxygen in all three domains of life (Denisova et al., 2005; Misra et al., 2012). Among a large variety of enzymes, CYPs play critical roles in the biosynthesis of a wide range of secondary metabolites such as phenylpropanoids (flavonoids, isoflavonoids, coumarins, etc.), terpenoids, lignin intermediates, alkaloids, fatty acids, hormones (gibberellins, jasmonic acid, brassinosteroids, etc.), pigments, and protective agents in the shikimic acid pathway (Schuler, 1996; Chapple, 1998; Schuler and Werck-Reichhart, 2003). They are also involved in critical processes like the metabolism of xenobiotics (Brazier et al., 2002; Gorinova et al., 2005), in-chain hydroxylation,

* Correspondence: aysenur.eminoglu@erdogan.edu.tr

omega-hydroxylation, and/or epoxidation of medium- and long-chain fatty acids. Protective layers such as cutins and suberins consist mainly of hydroxylated fatty acids (Salaün and Helvig, 1995). CYP86 and CYP94 are the major families involved in fatty acid oxidation in plants. Enzymes that belong to CYP703, CYP704, CYP709, CYP77, and CYP74 are also identified as catalysts of fatty acid oxidation reactions. It has been demonstrated that members of the CYP704 family participate in fatty acid metabolism in plants (Dobritsa et al., 2009; Li et al., 2010; Pinot and Beisson, 2011).

The ability of secondary metabolism, which is active in the production of more than 200,000 natural compounds in plants, to produce a variety of different products is based on the variety of genes involved in this metabolic pathway (Mizutani, 2012). Among these genes, P450 monooxygenases are the most important (Chapple, 1998). Diversification of P450s leads to the production of specific metabolites in different species and therefore most of the P450s that have been studied so far have produced a unique protein (Zhao, 2014). CYP genes comprise approximately 1% of the protein coding genes in several angiosperms (Nelson et al., 2008), and the determination of P450 genes and characterization of their proteins have become the primary focus of plant P450 studies (Chapple, 1998). Although an increasing number of studies have been carried out on the genetic mechanism of secondary metabolite biosynthesis in the tea plant (Singh et al., 2009b; Wang et al., 2012b; Zhao et al., 2014; Ono et al., 2016), there is still a need to explore new genes and identify the genetic mechanisms of CYP-catalyzed secondary metabolic pathways in the *C. sinensis* (tea) plant. Furthermore, the identification of the genes of *C. sinensis* would help the development and completion of the reference genome of this important crop. Studies on the discovery, investigation, and characterization of CYP genes, which encode the vast majority of the enzymes in secondary metabolism, have become increasingly important in *C. sinensis* (tea).

In the present study, cDNA of a putative cytochrome P450 monooxygenase (CYP704) gene (*Csp450*) from *C. sinensis* (L.) Kuntze was in silico and molecularly cloned, and its seasonal expression profile was investigated in young leaves.

2. Materials and methods

2.1. Plant materials and sampling

C. sinensis (L.) Kuntze plants were collected from the Çaykur Atatürk Tea and Garden Cultures Research Institution′s experimental tea plantation plots (Rize, Turkey). Among ten registered Turkish tea clones, high-yielding tea clone Fener-3, which has been intensively used for black tea production, was chosen as the RNA source in this study. Young shoots and the first two leaves were plucked from the plants during three consecutive harvesting periods (May, July, and September) and immediately frozen in liquid nitrogen and kept at –80 °C until RNA extraction.

2.2. In silico cloning

In silico cloning of *C. sinensis Csp450* gene was performed as described by Zhang et al. (2002) and Passier et al. (2004). In the first round, short sequence reads from the expressed sequence tag (EST) database (https://www.ncbi.nlm.nih.gov/nucest/) were investigated against the term "*Camellia. sinensis* P450 cytochrome monooxygenase" and the sequences found were used for the second-round search in the EST database. Obtained *C. sinensis* ESTs were assembled by contig assembly software, CAP3 (Huang and Madan, 1999). After several searches of the assembled contigs against the database, ESTs with the highest similarity to a putative *C. sinensis* P450 cytochrome monooxygenase gene were used to obtain the final sequence via contig assembly. Partial coding sequence (CDS) was predicted by ExPASY translate tool (http://web.expasy.org/translate/) as a standard genetic code and used as the template sequence for the RACE-PCR.

2.3. Total RNA extraction and cDNA synthesis

Total RNA was extracted from a 0.1 g leaf sample that had been homogenized in liquid nitrogen, with a G1-Total RNA Extraction Kit (Vivantis, Malaysia), according to the manufacturer's instructions. RNA concentrations were measured by a spectrophotometer (NanoDrop UV-Vis Spectrophotometer 2000, Thermo Scientific, MA, USA). Ratios of 260/280 and 260/230 nm were used to determine the quality of the RNA. cDNA was synthesized by two different methods. (1) In order to verify the in silico cloning, cDNA was obtained with a Transcriptor First Strand cDNA Synthesis Kit (Roche, Switzerland) from 1 µg of total RNA. (2) For the amplification of 5′ and 3′ ends of *Csp450*, a 5′/3′ RACE Kit, 2nd Generation (Roche) was used in cDNA synthesis according to the manufacturer′s protocol.

2.4. Verification of in silico cloning via sequencing

The partial CDS of *Csp450* obtained via in silico cloning was confirmed by sequencing. For this purpose, *Csp450* cDNA was used as a template for the amplification of the partial *Csp450* CDS with a pair of primers (Csp450_F and Csp450_R) (Table). The PCR reaction was carried out in a 50 µL reaction volume containing 1X Phusion HF buffer (Thermo Scientific), 200 µM of dNTP mix, 0.5 µM of each primer, and 0.01 U of Phusion High Fidelity DNA polymerase (Thermo Scientific). After purification (High Pure PCR Product Purification Kit, Roche) and 3′A tailing with GoTaq DNA polymerase (Promega, Madison, WI, USA), PCR fragments were cloned into the pCR2.1-TOPO vector (TOPO TA Cloning Kits for Subcloning; Invitrogen, Carlsbad, CA, USA), as described elsewhere. Positive clones were verified via colony PCR using the Csp450_F and the M13 reverse primer of the pCR 2.1-

Table. List of primers used in this study.

Csp450_F* ′	5′-GGAAAGAAGAAGTACCATCCAGTG-3′
Csp450_R*	5′-GGATCGGTTTCAGTCACTCCAAG-3′
Csp4505RACE1*	5′-GATTTGCAGAGAGAGGAGAGAGAGG-3′
Csp4505RACE2*	5′-CCTGAAAGGGCTTAACAACCTGTAAG-3′
Csp4505RACE3*	5′-CCACTGGATGGTACTTCTTCTTTCC-3′
3′Degenerate*	5′-CCRTCDAYRTGAAGAKTDATC-3′
Csp4503RACE*	5′-GTTTAGGAGCGGAAGGCCAG-3′
Oligo d(T)-anchor primer	5‘-GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV-3′
PCR anchor primer	5‘-GACCACGCGTATCGATGTCGAC-3′
Csp450cDNA_F*	5′-GATGATTTCCATGGATTCTCTC-3′
Csp450cDNA_R*	5′-GTTACACACTGGATCTGTGGAAG-3′
GAPDH_F (in expression)	5′-TTGGCATCGTTGAGGGTCT-3′
GAPDH_R (in expression)	5′CAGTGGGAACACGGAAAGC-3′
qCsp450_F (in expression)	5′-GGAACTAGCGAAGAAGGCACC-3′
qCsp450_R (in expression)	5′-GCCTTCCGCTCCTAAACTCAA-3′

* Primers were designed manually.

TOPO vector by 0.5 U of GoTaq DNA Polymerase with the following cycling conditions: initial denaturation steps at 95 °C for 3 min, denaturation at 95 °C for 45 s, followed by annealing at 58 °C for 1 min and primer extension at 72 °C for 40 s, followed by a step at 72 °C for 5 min, for 25 cycles. All PCR-verified vectors were sequenced by Macrogen Inc. (Seoul, Korea) using the M13 universal primer.

2.5. Rapid amplification of 5′ and 3′ ends of the *Csp450* cDNA

To determine the 5′ and 3′ end coding sequences of the *Csp450* cDNA, RACE-PCR was conducted using the 5′/3′ RACE Kit, 2nd Generation (Roche) without any modification according to the following steps. (1) 5′RACE was performed with three nested primers and one oligo d(T)-anchor primer (5′/3′ RACE Kit, 2nd Generation, Roche) by two-step PCR. Briefly, first, Csp4505RACE1 primer was used for the synthesis of the first strand *Csp450* cDNA by Transcriptor Reverse Transcriptase; then primer Csp4505RACE2, which is specific to the upstream region of the partial *Csp450* cDNA, and primer Csp4505RACE3 were used at the first and second rounds of 5′RACE-PCR for the identification of the 5′ end, respectively, following the manufacturer′s instructions. (2) The 3′ end of the *Csp450* cDNA was amplified by 3′RACE-PCR using the PCR anchor primer (5′/3′ RACE Kit, 2nd Generation, Roche) and a nested antisense primer (Csp4503RACE), after the synthesis of the first strand *Csp450* cDNA with oligo d(T)-anchor primer (5′/3′ RACE Kit, 2nd Generation, Roche, Switzerland), following the manufacturer′s procedure.

2.5.1. Determination of the 3′ end length

Determination of the 3′ end length was performed by a degenerate reverse primer with a single step of standard PCR using the 1X GoTaq Green Flexi Buffer, 1.5 mM $MgCl_2$, 0.6 µM 3′ degenerate primer, 0.5 µM Csp4503RACE primer, 200 µM dNTP mix, 2 µL cDNA, and 2 U GoTaq DNA polymerase (Promega). The 3′ terminal regions of five organisms′ amino acid sequences (GenBank accession numbers: XP_002529550, XP_007052008, XP_011093193, XP_011034689, AAZ39646.1, EXB52689.1), which were the sequence identities ranging from 75% to 76% to the corresponding amino acid sequence, were used in designing the degenerate primer.

2.6. Full-length *Csp450* cDNA cloning and sequencing

After assembly and analysis of the 5′ end, the 3′ end, and partial *Csp450* CDS sequences, two complete CDS specific primers (Csp450cDNA_F and Csp450cDNA_R) were designed to amplify the full-length cDNA sequence of *Csp450*. After amplification with a standard one-step PCR by Phusion High Fidelity DNA polymerase (Thermo Scientific), PCR products were cloned into the pCR2.1-TOPO, as described above, and then sequenced (Macrogen Inc.) using the M13 universal primer.

2.7. RT-qPCR and seasonal expression analysis in young leaves

RT-qPCR analysis of the full-length *Csp450* was carried out by CFX96 RT-qPCR system (Bio-Rad, Hercules, CA, USA). The iTaq Universal SYBR Green One-Step Kit was used to reverse transcript 300 ng of RNA extracted from the first two leaves of tea plants. Expressions level were measured with the iTaq Universal SYBR Green reaction mix (Bio-Rad) following the manufacturer′s instructions. An annealing temperature of 60 °C used to determine expression levels. Reactions were performed in a 10 µL reaction mixture containing 5 µL of iTaq Universal SYBR Green reaction mix (1X), 0.3 µM of each primer, 300 ng of RNA, and 0.125 µL of iScript reverse transcriptase (1X). The expression levels were assessed by the normalization of the *Ct* values to GAPDH housekeeping gene expression levels. To investigate the seasonal expression of the Csp450, plants were collected during the three consecutive harvesting periods (May, July, and September). All reactions were carried out in triplicate. All Rt-qPCR expression primers were designed by Primer3 v. 0.4.0 (Untergasser et al., 2012) and are given in the Table.

2.8. Bioinformatics analysis

Sequence analysis was conducted by SnapGene v.2.8 (www.snapgene.com) and CLC Main Workbench v.7.6.4 (https://www.qiagenbioinformatics.com/). Nucleotide and amino acid sequence resemblances were determined by NCBI blastn program (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch) and DELTA-BLAST server (Boratyn et al., 2012), respectively. Sequences were aligned by ClustalW (Thompson et al., 1994). Evolutionary relations were revealed using the maximum likelihood method based on the Poisson correction model (Zuckerkandl and Pauling, 1965). Initial tree(s) for the heuristic search were obtained automatically by applying the Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with a superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 16 amino acid sequences. All positions containing gaps and missing data were eliminated. There was a total of 416 positions in the final dataset. Phylogenetic familiarities were identified by MEGA7 (Kumar et al., 2015). Conserved domains were determined by the NCBI′s conserved domain database (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

3. Results

3.1. Molecular characterization of *Csp450*

When short sequence reads from EST databases were searched against the term "*C. sinensis* P450 cytochrome monooxygenase," three sequences were found, with GenBank accession numbers JK993199.1, ES323524.1, and ES323538.1. Each sequence was searched against the EST database with the default BLAST parameters using the blastn program again, and a total of nine reads were retrieved (GenBank accession numbers JZ486013.1, FS948329.1, FS949044.1, FS958246.1, FS946138.1, FS944460.1, JK342044.1, JK265118.1, CV699567.1). These sequences were assembled via CAP3 and a 917 bp sequence was obtained (Figure 1). BLAST of this sequence showed a significant similarity with *Nicotiana sylvestris* cytochrome P450 (XM_009778139.1) (76%) and *Citrus sinensis* cytochrome P450 (XM_006490889.1) (76%).

Based on this sequence, a 764 bp fragment was amplified from *Csp450* cDNA and cloned (Figure 2). In subsequent verification of this sequence, in order to determine the full-length cDNA of the *Csp450* gene with 5′ and 3′ untranslated regions (UTRs), we first aimed to identify the 5′ and 3′ ends of the *Csp450*.

```
  1 AAAAAAGCCA GTGGCCAACA CTGCACATCT CTCTCACTAC ACCAAGCTAT GATTTCCATG
 61 GATTCTCTCT CTAACTTCAT CTCTATACCA GCCTTGGTTC TCTCCATAAT CCTCTCTCTC
121 CTCTCTCTGC AAATCCTTAA AAGAAAACTC AATGGAAAGA AGAAGTACCA TCCAGTGGCA
181 GGGACCATGT TCAACCAGCT ACTCAATTTC AATAGGGTTC ATGATTACAT GACTGATCTT
241 GCTGCAAAGC ACAAGACTTA CAGGTTGTTA AGCCCTTTCA GGAATGAGAT TTACACTTCA
301 GATCCTGCAA ATGTTGAGTA CATTCTCAAA ACAAACTTTG AAAATTATGG CAAGGGGATG
361 TACAATTACT GCATCTTTAG GGATCTATTA GGTGATGGAA TTTTCACAGT CGATGGAGAC
421 AAACGGCGTC AACAACGAAA AGTATCAAGC CACGAATTCT CCACAAAGGT ATTGCGGGAT
481 TTTAGCAGCG TCATCTTCCA GAGAAGTGTC GCGAAACTTG CCAATATCAT CTCTAAAGCT
541 GCAACTTCCA ACCAAGTCAT GGATCTTCAA GATCTGTATA TGAGATCAAC CTTGGATTCA
601 ATATTCAAAG TTGGGTTTGG AGTTGAACTA GACAGTATGT GCGGAACTAG CGAAGAAGGC
661 ACCATATTTA CCAATGCCTT TGATGATTCA AGCGCAATAA CGCTTTTTCG ATTTGTTGAT
721 GTCTTTTGGA AGATAAAGAA AGCTTTGAGT TTAGGAGCGG AAGGCCAGTT AAAGAAAAAT
781 GTCAAAATCA TCAACGAGTT CGTATACAAG CTAATCCAAA GCAAGGCAGA GCAAATGAAA
841 AAGTCTCCAA ATGAATTTTC TGCAAAGAAA GAAGACATTT TATCAAGGTT TCTTGGAGTG
901 ACTGAAACCG ATCCGAC
```

Figure 1. 916 bp sequence of *Csp450* gene, generated by in silico cloning.



Figure 2. 750 bp sequence amplified from a 917 bp in silico cloned sequence (Binding regions of the primers are shown with orange bars. Arrows indicate the direction of the primers).

The 5′ end was identified by two nested primers. The first round of 5′ RACE gave a fragment about 250–300 bp in length, whereas the second round was approximately 200 bp. The sequence of this region was spliced, via overlapping the primer sequences, to the 916 bp sequence obtained by in silico cloning. Splicing revealed that the sequence obtained via in silico cloning already contained the start codon.

The 3′ RACE PCR that was carried out to define the 3′ end resulted in two fragments approximately 900 and 300 bp in length. To understand which of these fragments represented the real 3′ end of the *Csp450* cDNA, a standard PCR was performed with a degenerate primer that resulted in a distinct band of about 800 bp. Regarding this band, the 900 bp band generated by 3′ RACE PCR was considered the 3′ end fragment. Following the cloning and sequence analysis, the 3′ end sequence was aligned and assembled to the in silico cloned sequence with the 5′ end sequence. Cloning and sequencing of this provisional full cDNA demonstrated that the full-length *Csp450* cDNA was 1759 bp in length, with 49 bp of 5′ and 183 bp of 3′ untranslated regions (Figure 3). The CDS of the gene is 1527 bp (GenBank accession number KT007221) and 508 amino acids in length (Figure 4).

According to phylogenetic analysis, Csp450 is significantly similar to the CYP704C1-like superfamily (Figure 5). A conserved domain database search of the Csp450 amino acid sequence showed that the carboxyl-terminal region of Csp450 also contains the heme-binding domain (FXXGXXXCXG), which is highly conserved among CYPs (Umemoto et al., 1993) (Figure 6).

3.2. Expression analysis of *Csp450* in three consecutive harvesting periods

As a result of the gene expression analysis by RT-qPCR, similar expression patterns were observed in the first two leaves of the three preharvest period *C. sinensis* plants (Figure 7).

4. Discussion

In Turkey, the first tea garden was established in Rize in 1924, but agricultural *C. sinensis* (L.) Kuntze plant cultivation started after 1938 (Kafkas et al., 2009; Saklı, 2011) Since then, several tea clones, such as Derepazarı-7, Muradiye-10, Tuğlalı-10, Pazar-20, and Gündogdu-3, have been developed and used in commercial black tea production. In this study, due to its high yield, we used the first two leaves from *C. sinensis* var. *sinensis* clone Fener-3 (Zenginbal et al., 2014) for expression analysis. Because of their high phenolic content, a young bud and the first two or three leaves of the tea plant *C. sinensis* are used for high-quality Turkish tea production (Beris et al., 2005). It was suggested that the putative gene may play a role in catechin biosynthesis and thus its expression was analyzed in the first two leaves during three consecutive preharvest periods (De Vetten et al., 1999; Punyasiri et al., 2004; Liu et al., 2015). However, compared to the housekeeping gene GAPDH, normalized expression of the corresponding gene was found to be very low in each of the preharvest periods. Tea flowers at the end of August and at the beginning of September, and product accumulation occurs due to the high activity of the corresponding genes. It may be expected that CYP gene expression may decrease during the preharvest periods after the flowering season. Based on the phylogenetic analysis, Csp450 is significantly similar to a CYP704C1-like superfamily (Figure 5). There have been some studies reporting that members of the CYP704 family are involved in fatty acid metabolism in plants (Dobritsa et al., 2009; Li et al., 2010; Pinot and Beisson, 2011; Nelson and Werck-Reichhart, 2011); however, all these enzymes belong to the CYP704B family.



Figure 3. Schematic diagram of *Csp450* full cDNA nucleotide sequence. Highlighted sections show the 5′ and 3′ UTR sequences.

Studies on the seasonal variation of the phenolic compounds in *C. sinensis* have shown that there were some seasonal differences in the accumulation of phenolic compounds in different catechin classes; some of them increased in the warmer season while others in the cold season (Yao et al., 2005; Liu et al., 2015). It should be considered that sunlight and length of daytime have an effect on phenolic compound distribution in tea shoots (Harbowy and Balentine, 1997; Wang et al., 2012a, 2012b). In a study that addressed seasonal variation in phenolic compound distribution in Turkish tea from Rize (Erturk et al., 2010), it was demonstrated that total phenolic compounds were lower in May when compared to September, and increased from July to September. Moreover, cytochrome p450 genes are effective in responding to adverse environmental conditions like

```
  1 MISMDSLSNF ISIPALVLSI ILSLLSLQIL KRKLNGKKKY HPVAGTMFNQ LLNFNRVHDY
 61 MTDLAAKHKT YRLLSPFRNE IYTSDPANVE YILKTNFENY GKGMYNYCIF RDLLGDGIFT
121 VDGDKRRQQR KVSSHEFSTK VLRDFSSVIF QRSVAKLANI ISKAATSNQV MDLQDLYMRS
181 TLDSIFKVGF GVELDSMCGT SEEGTIFTNA FDDSSAITLF RFVDVFWKIK KALSLGAEGQ
241 LKKNVKIIND FVYKLIQSKA EQMKKSPNEF SAKKEDILSR FLGVTETDPT YLRDIILNFI
301 IAGKDTTATT LSWFTYALCK HPEIQEKIAR EIREATGIKE VTDYAEFAAG MREDVLEKMQ
361 YLHAALTETL RLYPAVPVDA KICFSDDTLP DGFSVKKGDM VAYQPYAMGR MKFLWGADAA
421 EFRPERWLDE NGIFRPESPF KFTAFQAGPR ICLGKEFAYR QMKIFSAVLL GCFVFKLSDP
481 TKPVNYRTMI NLHVDGGLQI RAFHRSSV*
```

Figure 4. Amino acid sequence of Csp450.

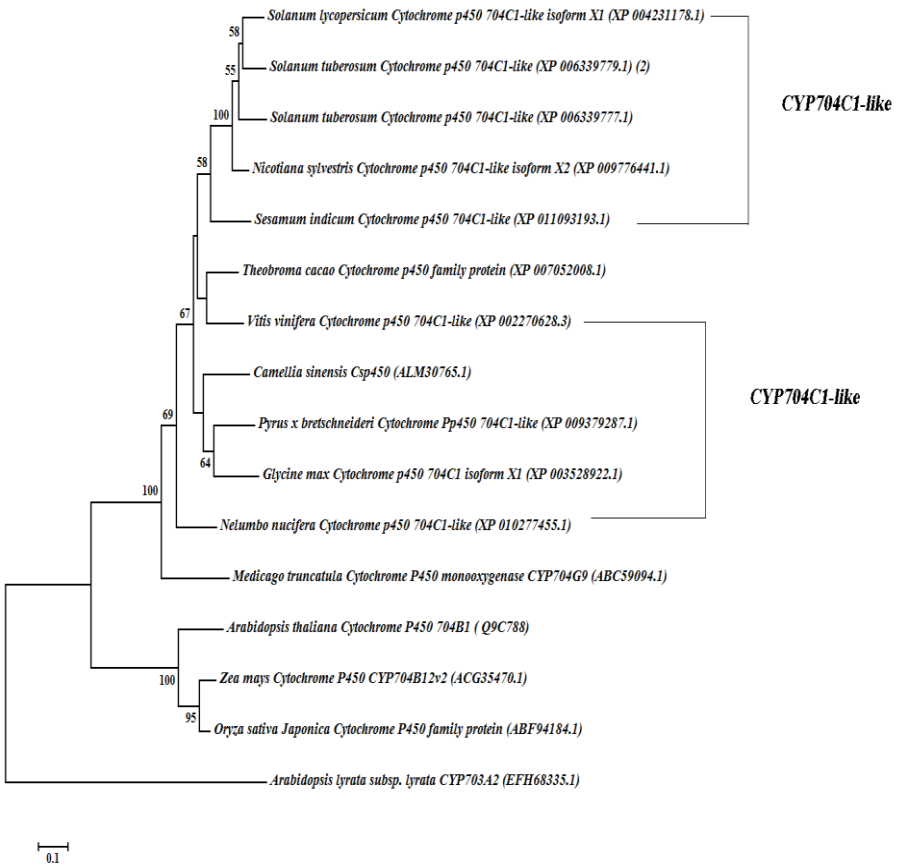


Figure 5. Evolutionary relationship analysis of certain plant CYP genes. The tree was drawn by maximum likelihood method based on the Poisson correction model of some plant cytochrome P450 monooxygenase amino acid sequences (accession numbers are given in parentheses). The phylogenetic tree with the highest log likelihood (–8275.8434) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. *A. lyrata* was selected as an outgroup member.

physiological stress, exogenous chemical treatments, and in xenobiotic metabolism in plants (Robineau et al., 1998; Morant et al., 2003; Mizutani, 2012; Tan et al., 2015). In the light of these studies, relatively high levels of expression were expected in September tea; however, the expression was not any different when compared to May and July. Due to poor or insufficient information on the activity of the CYP704C family in the literature, we were unable to estimate whether this result is correlated with catechin biosynthesis and therefore statistically significant.

Since the first discovery of plant CYPs in 1969 (Frear et al., 1969), CYPs have gained considerable attention because

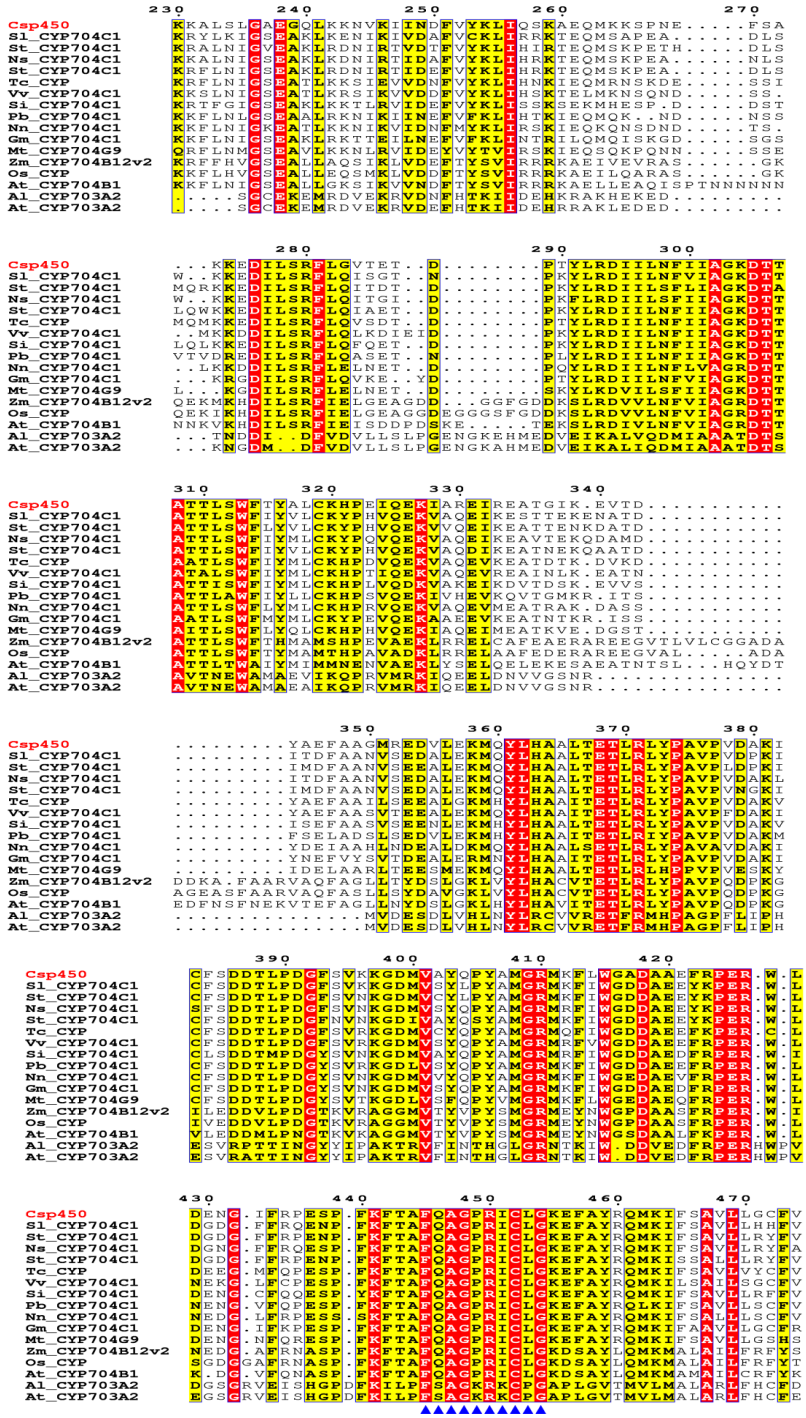


Figure 6. Multiple amino acid sequence alignment of certain CYP members; *Camellia sinensis Csp450* (Csp450, highlighted in red), *Solanum lycopersicum* cytochrome p450 704C1-like isoform X1 (Sl_CYP704C1), *Solanum tuberosum* cytochrome p450 704C1-like 1 (St_CYP704C1), *Nicotiana sylvestris* cytochrome p450 704C1-like isoform X2 (Ns_CYP704C1), *Solanum tuberosum* cytochrome p450 704C1-like 2 (St_CYP704C1), Theobroma cacao cytochrome p450 family protein (Tc_CYP), *Vitis vinifera* cytochrome p450 04C1-like (Vv_CYP704C1), *Sesamum indicum* cytochrome p450 704C1-like (Si_CYP704C1), *Pyrus* × *bretschneideri* cytochrome p450 704C1-like (Pb_CYP704C1), *Nelumbo nucifera* cytochrome p450 704C1-like (Nn_CYP704C1), *Glycine max* cytochrome p450 704C1 isoform X1 (Gm_CYP704C1), *Medicago truncatula* cytochrome P450 monooxygenase_CYP704G9 (Mt_CYP704G9), *Zea_mays_Cytochrome* P450 CYP704B12v2 (Zm_CYP704B12v2), *Oryza sativa* L. ssp *japonica* cytochrome P450 family protein (Os_CYP), *Arabidopsis thaliana* cytochrome P450 704B1 (At_CYP704B1), *Arabidopsis lyrata* ssp. *lyrata* CYP703A2 (Al_CYP703A2), and *Arabidopsis thaliana* cytochrome P450 703A2 (At_CYP703A2). Triangles indicate the heme-binding domain residues. The figure was drawn with ESPript (Robert and Gouet, 2014).

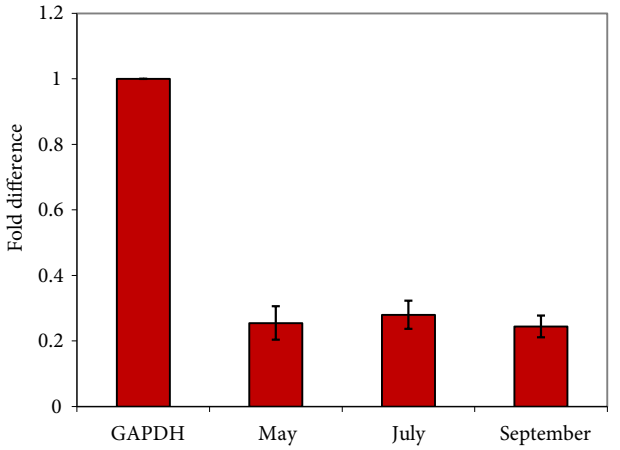


Figure 7. Normalized expression analysis of Csp450 relative to GAPDH during the preharvest periods.

of their essential functions in plants. Although the P450 superfamily enzymes have similar structures and features, they may have different effects on biological systems (Zhao et al., 2014). In this study, we identified a new CYP P450 monooxygenase gene that may have a particular role in the secondary metabolism of the agriculturally, medicinally, and industrially important crop *C. sinensis* (L.) Kuntze.

Further investigations and functional characterization of this gene may reveal the function of the gene in the tea plant metabolism.

To the best of our knowledge, this study is the first in silico cloning and molecular characterization of a putative CYP704 family gene in *C. sinensis* (L.) Kuntze.

Acknowledgments

This work was supported by a grant from the Scientific and Technological Research Council of Turkey (TÜBİTAK) with the Support No: 114Z598. Also, we would like to express our sincere thanks to the Çaykur Atatürk Tea and Garden Cultures Research Institution (Rize, Turkey) and to Ayhan Haznedar, who kindly helped us to identify and collect the tea samples.

References

Beris FS, Sandalli C, Canakci S, Demirbag Z, Belduz AO (2005). Phylogenetic analysis of tea clones (*Camelli sinensis*) using RAPD markers. Biologia 60: 451-456.

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL (2012). Domain enhanced lookup time accelerated BLAST. Biology Direct 7: 1.

Brazier M, Cole DJ, Edwards R (2002). O-Glucosyltransferase activities toward phenolic natural products and xenobiotics in wheat and herbicide-resistant and herbicide-susceptible black-grass (*Alopecurus myosuroides*). Phytochemistry 59: 149-156.

Chaple C (1998). Molecular-genetic analysis of plant cytochrome P450-dependent monooxygenase. Annu Rev Plant Physiol Plant Mol Biol 49: 311-343.

Chen BT, Li WX, He RR, Li YF, Tsoi B, Zhai, YJ, Kurihara H (2012). Anti-inflammatory effects of a polyphenols-rich extract from tea (*Camellia sinensis*) flowers in acute and chronic mice models. Oxid Med Cell Longev doi:10.1155/2012/537923.

Denisov IG, Makris TM, Sligar SG, Schlichting I (2005). Structure and chemistry of cytochrome P450. Chem Rev 105: 2253-2278.

De Vetten N, Ter Horst J, Van Schaik HP, De Boer A, Mol J, Koes R (1999). A cytochrome b5 is required for full activity of flavonoid 3′, 5′-hydroxylase, a cytochrome P450 involved in the formation of blue flower colors. Proc Natl Acad Sci USA 96: 778-783.

Dobritsa AA, Shrestha J, Morant M, Pinot F, Matsuno M, Swanson R, Møller BL, Preuss D (2009). CYP704B1 is a long-chain fatty acid ω-hydroxylase essential for sporopollenin synthesis in pollen of *Arabidopsis*. Plant Physiol 151: 574-589.

Erturk Y, Ercisli S, Sengul M, Eser Z, Haznedar A, Turan M (2010). Seasonal variation of total phenolic, antioxidant activity and minerals in fresh tea shoots (*Camellia sinensis* var. *sinensis*). Pak J Pharm Sci 23: 69-74.

Frear DS, Swanson HR, Tanaka FS (1969). N-demethylation of substituted 3-(phenyl)-1-methylureas: isolation and characterization of a microsomal mixed function oxidase from cotton. Phytochemistry 8: 2157-2169.

Gorinova N, Nedkovska M, Atanassov A (2005). Cytochrome P450 monooxygenases as a tool for metabolizing of herbicides in plants. Biotechnol Biotechnol Equip 19: 105-115.

Huang X, Madan A (1999). CAP3: A DNA sequence assembly program. Genome Res 9: 868-877.

Kafkas S, Ercişli S, Doğan Y, Ertürk Y, Haznedar A, Sekban R (2009). Polymorphism and genetic relationships among tea genotypes from Turkey revealed by amplified fragment length polymorphism markers. J Am Soc Hortic Sci 134: 428-434.

Kandel S, Sauveplane V, Olry A, Diss L, Benveniste I, Pinot F (2006). Cytochrome P450-dependent fatty acid hydroxylases in plants. Phytochem Rev 5: 359-372.

Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol 33: 1870-1874.

Li H, Pinot F, Sauveplane V, Werck-Reichhart D, Diehl P, Schreiber L, Franke R, Zhang P, Chen L, Gao Y et al. (2010). Cytochrome P450 family member CYP704B2 catalyzes the ω-hydroxylation of fatty acids and is required for anther cutin biosynthesis and pollen exine formation in rice. Plant Cell 22: 173-190.

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.). Hort Res 2: 15011.

Misra A, Chanotiya CS, Gupta MM, Dwivedi UN, Shasany AK (2012). Characterization of cytochrome P450 monooxygenases isolated from trichome enriched fraction of *Artemisia annua* L. leaf. Gene 510: 193-201.

Mizutani M (2012). Impacts of diversification of cytochrome P450 on plant metabolism. Biol Pharm Bull 35: 824-832.

Morant M, Bak S, Møller BL, Werck-Reichhart D (2003). Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr Opin Biotechnol 14: 151-162.

Nelson DR, Ming R, Alam M, Schuler MA (2008). Comparison of cytochrome P450 genes from six plant genomes. Trop Plant Biol 1: 216-235.

Nelson D, Werck-Reichhart D (2011). A P450-centric view of plant evolution. Plant J 66: 194-211.

Ono E, Handa T, Koeduka T, Toyonaga H, Tawfik MM, Shiraishi A, Murata J, Matsui K (2016). CYP74B24 is the 13-hydroperoxide lyase involved in biosynthesis of green leaf volatiles in tea (*Camellia sinensis*). Plant Physiol Biochem 98: 112-118.

Park JS, Kim JB, Hahn BS, Kim KH, Ha SH, Kim JB, Kim YH (2004). EST analysis of genes involved in secondary metabolism in *Camellia sinensis* (tea), using suppression subtractive hybridization. Plant Sci 166: 953-961.

Passier R, Doevendans PA (2004). Methods in molecular cardiology: in silico cloning. Neth Heart J 12: 400.

Pinot F, Beisson F (2011). Cytochrome P450 metabolizing fatty acids in plants: characterization and physiological roles. FEBS J 278: 195-205.

Punyasiri PAN, Abeysinghe ISB, Kumar V, Treutter D, Duy D, Gosch C, Martens S, Forkmann G, Fischer TC (2004). Flavonoid biosynthesis in the tea plant *Camellia sinensis*: properties of enzymes of the prominent epicatechin and catechin pathways. Arch Biochem Biophys 431: 22-30.

Robert X, Gouet P (2014). Deciphering key features in protein structures with the new ENDscript server. Nucl Acids Res 42: 320-324.

Robineau T, Batard Y, Nedelkina S, Cabello-Hurtado F, LeRet M, Sorokine O, Didierjean L, Werck-Reichhart D (1998). The chemically inducible plant cytochrome P450 CYP76B1 actively metabolizes phenylureas and other xenobiotics. Plant Physiol 118:1049-1056.

Sakli AR (2011). A critical rewiev of recent sectoral structure proposal for Turkish tea sector. Humanity & Social Science Journal 6: 1-7.

Salaün JP, Helvig C (1995). Cytochrome P450-dependent oxidation of fatty acids. Drug Metabol Drug Interact 12: 261-284.

Schuler MA (1996). Plant cytochrome P450 monooxygenases. Crit Rev Plant Sci 15: 235-284.

Schuler MA, Werck-Reichhart D (2003). Functional genomics of P450s. Annu Rev Plant Biol 54: 629-667.

Singh K, Rani A, Paul A, Dutt S, Joshi R, Gulati A, Ahuja PS, Kumar S (2009a). Differential display mediated cloning of anthocyanidin reductase gene from tea (*Camellia sinensis*) and its relationship with the concentration of epicatechins. Tree Physiol 29: 837-846.

Singh K, Kumar S, Yadav SK, Ahuja PS (2009b). Characterization of dihydroflavonol 4-reductase cDNA in tea [*Camellia sinensis* (L.) O. Kuntze]. Plant Biotechnol Rep 3: 95-101.

Tan LR, Lu YC, Zhang JJ, Luo F, Yang H (2015). A collection of cytochrome P450 monooxygenase genes involved in modification and detoxification of herbicide atrazine in rice (*Oryza sativa*) plants. Ecotoxicol Environ Saf 119: 25-34.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl Acids Res 22: 4673-4680.

Umemoto N, Kobayashi O, Ishizaki-Nishizawa O, Toguri T (1993). cDNAs sequences encoding cytochrome P450 (CYP71 family) from eggplant seedlings. FEBS Lett 330: 169-173.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012). Primer3 - New capabilities and interfaces. Nucleic Acids Res 40: 115.

Wang Y, Gao L, Shan Y, Liu Y, Tian Y, Xia T (2012a). Influence of shade on flavonoid biosynthesis in tea (*Camellia sinensis* (L.) O. Kuntze). Sci Hortic 141: 7-16.

Wang Y, Gao L, Wang Z, Liu Y, Sun M, Yang D, Weia C, Shana Y, Xia T (2012b). Light-induced expression of genes involved in phenylpropanoid biosynthetic pathways in callus of tea (*Camellia sinensis* (L.) O. Kuntze). Sci Hort 133:72-83.

Weaver LM, Herrmann KM (1997). Dynamics of the shikimate pathway in plants. Trends Plant Sci 2: 346-351.

Werck-Reichhart D (1995). Cytochromes P450 in phenylpropanoid metabolism. Drug Metabol Drug Interact 12: 221-244.

Winkel-Shirley B (2001). Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol 126: 485-493.

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

Zenginbal H, Haznedar A, Dolgun O (2014). Effects of indole-3-butyric acid (IBA) and cutting type on rooting of *Camellia sinensis* (L.) O. Kuntze. Am J Exp Agric 4: 1935-1943.

Zhang DL, Sun XJ, Ling LJ, Chen RS, Ma DL (2002). Molecular cloning, characterization, chromosomal assignment, genomic organization and verification of SFRS12 (SRrp508), a novel member of human SR protein superfamily and a human homolog of rat SRrp86. Acta Genet Sin 29: 377-383.

Zhao YJ, Cheng QQ, Su P, Chen X, Wang XJ, Gao W, Huang LQ (2014). Research progress relating to the role of cytochrome P450 in the biosynthesis of terpenoids in medicinal plants. Appl Microbiol Biotechnol 98: 2371-238.

Zuckerkandl E, Pauling L (1965). Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, editors. Evolving Genes and Proteins. New York, NY, USA: Academic Press, pp. 97-166.