Development of CAPS markers to identify Indian tea (*Camellia sinensis*) clones with high catechin content

M. Elangbam and A.K. Misra

Department of Botany, North-Eastern Hill University, Shillong, Meghalaya, India

Corresponding author: M. Elangbam
E-mail: meena.elangbam@gmail.com

ABSTRACT. Tea leaves are rich in plant secondary phenolics, especially flavonoids. Catechins are considered to be the most valuable flavonoids, and the catechin content in tea is an important trait for determining its quality. We have developed cleaved amplified polymorphic sequence (CAPS)-based markers for evaluating total catechin content that target two important secondary metabolite pathway genes, *PAL* (phenylalanine ammonia-lyase) and *CHS* (chalcone synthase). Catechin content levels in the tea samples tested ranged from 9 to 33 mg/mg. The CAPS technique identified clones with the homozygous profile PRc1, which has relatively lower catechin content than clones with the heterozygous profile PRc2. A significant difference (t = 16.85) in the level of catechin content was also detected between heterozygotes and homozygotes in the tea seed stock TS379. We found a polynomial relationship between the marker developed for *CHS2* and catechin content in these tea samples with $R^2 = 0.9788$. Moreover, *PAL* has less of a relationship with catechin content. Therefore, we recommend tea clones with heterozygous CAPS profiles for the gene *CHS2+Rca*I for the further improvement in these clones.

Key words: Amplicon restriction pattern; Catechin; Chalcone synthase; Cleaved amplified polymorphic sequence; Molecular markers; Phenylalanine ammonia-lyase

INTRODUCTION

Tea [*Camellia sinensis* (L.) O. Kuntze] belongs to the family Theaceae and is one of the three main classes of cultivated species: *C. sinensis* (L.) O. Kuntze, *C. assamica* (Masters) Wight, and the intermediate *C. assamica* ssp *lasiocalyx* (Planchon ex Watt) Wight. These are also known as China, Assam, and Cambod varieties, respectively (Banerjee, 1992). Tea leaves are the source of the oldest known beverage. They are rich in flavonoids, which are one of the largest classes of plant phenolics occurring abundantly in vegetables, fruits, and green plants (Gershenzon, 2002). Catechins constitute the major group of flavonoids found in tea leaves and are responsible for the characteristic taste, aroma, and color of tea. The major catechins in green leaves are epicatechin (EC), epigallocatechin (EGC), epigallocatechin gallate (EGCG), and epicatechin gallate (ECG), all of which are water-soluble polyhydroxylated flavonoids. A high content of total catechin is a good indicator of high quality tea (Anan and Nakagawa, 1974; Obanda et al., 1997). Fresh tea shoot catechin can be quantified spectrophotometrically using the reagent diazotized sulphanilamide (Singh et al., 1999). Owing in part to the growing evidence supporting the health benefits of tea, it has become one of the most popular non-alcoholic drinks in many cultures worldwide. Consequently, the increase in global demand for tea leaves presents growers with the challenge of increasing its production. Furthermore, tea growers and breeders have also explored different approaches for selecting clones with higher catechin content with the aim of improving the production of high quality tea. With this in mind, we designed an experimental plan to examine tea quality at the molecular level.

DNA based markers have become a very popular tool in the selection of superior plants. In addition to being very useful in allelic selection and in the removal of undesirable traits in breeding programs (Bang et al., 2007), molecular markers can also be effective in the selection of superior genotypes (Verghese and Misra, 2000; Chauhan and Misra, 2002). Many molecular markers used in plants today are based on random DNA sequences. Such markers are often validated for agronomic traits by QTL analysis, which relies heavily on statistical tools. However, with increased availability of specific gene sequences in publically available global databases, markers targeting specific genes are becoming more reliable and popular. Markers such as cleaved amplified polymorphic sequence (CAPS) (Konieczny and Ausubel, 1993) or amplicon restriction pattern (ARP) (Yanthan and Misra, 2013) have been employed for marker-assisted selection in many plant species (Chauhan and Misra, 2002; Dillon et al., 2006; Gutierrez et al., 2006; Kuklev et al., 2009; Ben et al., 2015). This technique involves amplification of the DNA by polymerase chain reaction (PCR) followed by digestion of the PCR product or amplicon with different restriction enzymes. These fragments can then be separated by agarose gel electrophoresis to detect polymorphisms within the amplified genomic region. The greatest advantage of CAPS markers over other technologies first lies in its high level of specificity and repeatability. Second, since it is a PCR-based marker it is less laborious and time-consuming than analyzing alternative types of markers that require Southern hybridization. Third, as they have been developed from expressed genes (ESTs) CAPS markers are more useful as genetic markers than those derived from non-functional sequences, such as genomic microsatellite markers (Semagn et al., 2006). Finally, CAPS markers are inherited mainly in a co-dominant manner (Matsumoto and Tsumura, 2004).

Genetic diversity of tea in both '*assamica*' and '*sinensis*' varieties was first analyzed by Kaundun and Matsumoto (2003) using CAPS in the phenylalanine ammonia-lyase (*PAL*),

chalcone synthase 2 (*CHS2*) and dihydroflavonol 4-reductase (*DFR*) genes. The *PAL* gene in tea is present as a single-copy gene (Matsumoto et al., 1994), whereas *CHS* found to exist in three copies CHS1, CHS2 and CHS3 (Takeuchi et al., 1994). These genes are directly involved in the secondary metabolite pathway in plants leading to the biosynthesis of many secondary metabolites, including flavonoids. In the present study, we have tried to find out if there exist any link between total catechin content present in tea and CAPS marker using PAL and CHS2. Since these genes are integral part of the secondary metabolite pathway, the possibility of finding an association with catechin content and the CAPS marker increased many fold. Further, CAPS being a codominant marker could help us identify heterozygotes as well. Further it will also help to identify high quality Indian tea clones.

Since we chose genes that are integral part of the secondary metabolite pathway, the possibility of finding an association with between catechin content and the CAPS marker increases many fold. Further, CAPS, being a codominant marker, could help us identify heterozygotes as well.

MATERIAL AND METHODS

Plant material

Two biclonal seed stocks, TS379 and TS463, and three clonal seed stocks, TS656, TS657, and TS658, of tea maintained at the Tea Research Association (TRA) in Tocklai, Jorhat, India were selected for the study. Seed stock TS379 was obtained by crossing clones 14.12.16 (China type) and 14.5.35 (China type), and seed stock TS463 was obtained by crossing clones TV1 (Assam-China hybrid) and TV19 (Cambod type). Twenty F_2 progenies each of TS379 and TS463 seed stocks were taken. Twenty clones each of the clonal seed stocks all Assam-China types were also included in the analysis.

DNA isolation

Total DNA was extracted from frozen tissue (1gm) using the CTAB method (Porebski et al., 1997) with minor modifications. The collected DNA pellet was suspended in an appropriate quantity of nuclease-free Milli-Q water and stored at -20°C. Quality and concentration of the DNA samples were assessed with agarose gel electrophoresis and by using Lambda 35 UV/VIS Spectrophotometer, Perkin Elmer, MA, USA. The A260/A280 ratio for the DNA samples isolated ranged from 1.8 - 2.0.

Catechin estimation

Catechins were quantified as described by Singh et al. (1999). The apical bud and the adjoining first two young leaves were selected for catechin estimation. Fresh shoots stored at -80°C were weighed and finely ground in 100ml acetone and the macerate was filtered through sintered glass funnel (porosity G3). The residue was washed with 100 mL of 100% acetone, followed by 200 mL of 60% acetone to remove all the water-soluble polyphenols of the macerated tea shoots. Next, two volumes of petroleum benzene was added to this solution. The lower aqueous acetone layer containing the catechins and other polyphenolic compounds was removed, dried, and dissolved in water. Five µL of the prepared extract of tea shoots

were placed in triplicate sets of 25 mL volumetric flasks. One mL of diazotized sulfanilamide solution (1% w/v prepared in acetone) followed by 1mL of diluted hydrochloric acid (30:100 v/v concentrated hydrochloric acid:double-distilled water) were subsequently added to the extract and allowed to react at room temperature for 1 hour. The reagent blank minus tea extract was included. At the end of incubation period the volumes were brought up to the mark with water and mixed well, and the absorbance was measured at 425 nm (λ max of the colored complex) in UV visible spectrophotometer against blank reagent. A standard curve was prepared with d-($\pm$)-catechin.

Target DNA segments and primer design

Marker development was based on the two genes *PAL* and *CHS2*, for which nucleotide sequence information is available for tea in the global gene databases. The gene segments of exon 2 for both *PAL* and *CHS2* were targeted because exons constitute part of the mature mRNA transcript responsible for the synthesis of the corresponding proteins. Although intron segments tend to harbor greater diversity, we deliberately excluded them, as they are not part of the mature mRNA. For amplifying exon 2 of *PAL*, primers (PX2F 5'-AGGCTAACATACTCGCC-3' and PX2R 5'-TGCGATAAGAATTGCAC-3') were designed based on the *C. sinensis* cultivar Yabukita cDNA sequences of *PAL* in GenBank (Accession: D26596) using the DS GENE version 1.1 software. Primers reported by Kaundun and Matsumoto (2003) were used for amplifying exon 2 region of *CHS2*.

Polymerase Chain Reaction

The polymerase chain reaction was carried out in a 25-μL volume containing approx. 50-100 ng genomic DNA, 2 pM primers, 200 μM dNTP, 2 mM MgCl$_2$, 50 mM KCl, 500 ng BSA, 10 mM Tris-HCl, pH 8.3, and 0.5 U Taq polymerase. PCR was performed in a thermal cycler (Applied Biosystems 9700) programmed for an initial denaturation step of 94°C for 5 min followed by 35 cycles of 1 min at 94°C, 1 min at the annealing temperature, and 1 min at 72°C. A final elongation step of 10 min at 72°C was also included. The PCR products were separated on a 1% agarose gel.

Digestion of PCR products

A total volume of 50 μL PCR products was digested separately with restriction endonucleases including *Rsa*I, *Rca*I, and *Kpn*I following the respective manufacturer instructions (Roche Diagnostics and New England BioLabs). The digested products were separated on a 4% agarose gel.

Statistical analysis

The Student *t*-test was performed on the levels of catechin contents of the heterozygous and homozygous lines using GraphPad Prism 4.0 software to test for significant differences between the two groups. A simple regression analysis was carried out using MS Excel in Windows 7 to assess the relationship between the level of catechin content and allelic composition of *CHS2*+ *Rca*I.

RESULTS

This study was conducted with the aim of developing CAPS markers to the level of polymorphism in commercially important tea lines. We developed DNA based markers for a selection of tea clones with higher catechin content and determined the allelic composition of two important genes (*PAL* and *CHS*) in the biclonal population of *Camellia* spp. Although CAPS studies were carried out with a large number of restriction enzymes, data presented here represent only those that yielded applicable information for our experimental aims. Genomic DNA from all the accessions considered in the present study was amplified by PCR using primer pairs targeting exon 2 of both *CHS2* and *PAL* genes. This procedure generated an amplicon size of 801 and 1344 bp for *CHS2* and *PAL*, respectively (Figures 1 and 2).
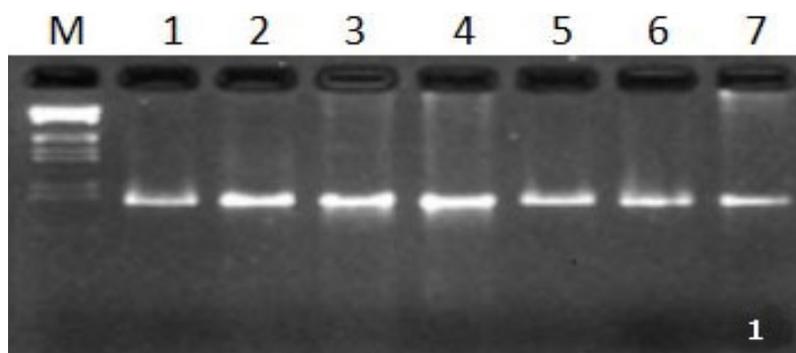
Figure 1. PCR profile of tea samples with *CHS2* exon 2 primers. *Lane M* = λ DNA/*Eco*RI + *Hin*dIII marker.
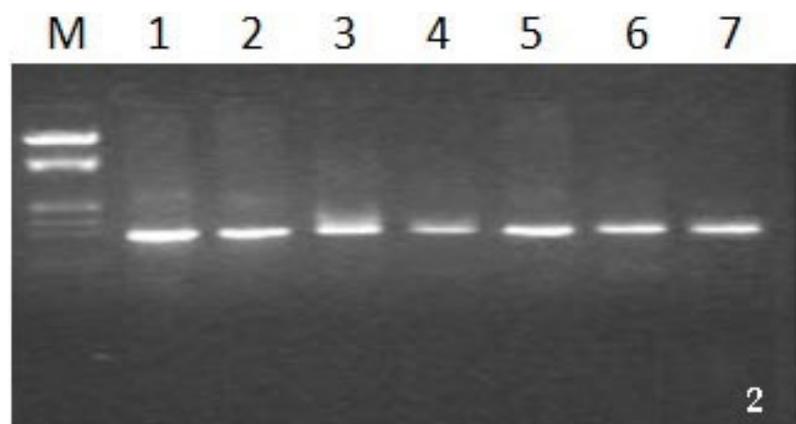
Figure 2. PCR profile of tea samples with *PAL* exon 2 primers. *Lane M* = λ DNA/*Eco*RI + *Hin*dIII marker.

Development of a CAPS marker in *CHS2* exon 2 digested with *Rsa*I

Two types of profiles were observed from the restriction profiles obtained for exon 2 of *CHS2* digested by *Rsa*I (Figure 3). One profile was found to correspond to a homozygous

condition for the restriction site (R1R1) and the other corresponded to a heterozygous condition for the restriction site (R1R2) (Figures 4 and 5). We designated the homozygous condition as profile PR1 and the heterozygous condition as profile PR2. In TS463, digestion of the *CHS2* exon 2 amplicon with *Rsa*I showed only the homozygous profile PR1. In TS379, both types of profiles could be observed: parent 14.12.16 and eleven F_2 progenies showed a homozygous profile PR1, while the other parent 14.5.35 and nine progenies showed a heterozygous profile PR2. The digestion of the *CHS2* exon 2 amplicon in TS656 and TS658 showed a heterozygous profile PR2, while TS657 showed a homozygous profile PR1 (Table 1). This polymorphism was observed owing to two different allelic forms at this locus.
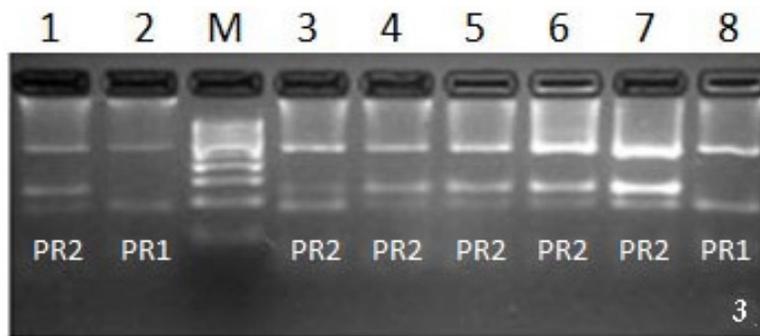
Figure 3. Amplicon restriction polymorphism observed in the *CHS2* exon 2 amplicon digested with *Rsa*I identifying the PR1 and PR2 profiles in TS379. *Lane M* = 100-bp ladder.
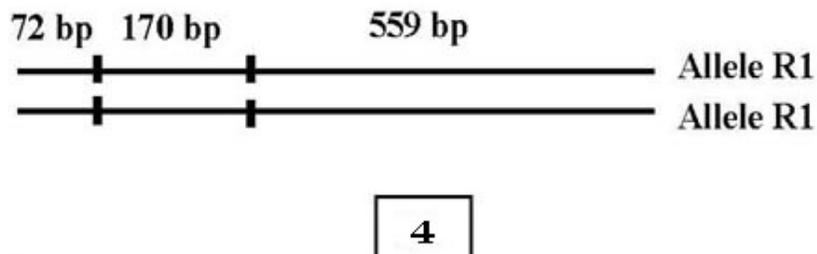
Figure 4. Restriction site map of the *CHS2* exon 2 digested with *Rsa*I in tea samples with the amplicon restriction profile PR1. (Distances not to scale).
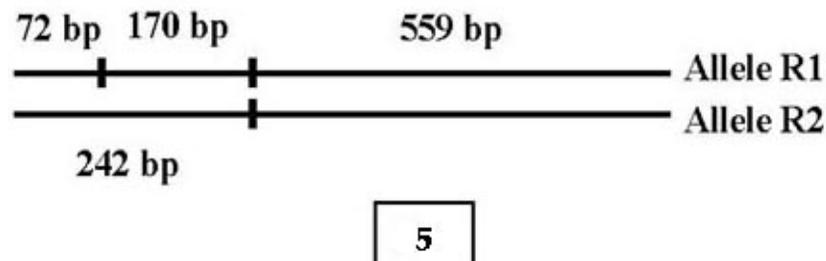
Figure 5. Restriction site map of the *CHS2* exon 2 digested with *Rsa*I in tea samples with the amplicon restriction profile PR2. (Distances not to scale).

Table 1. Relationship between catechin content and amplicon restriction profiles of DNA regions studied in TS463, TS656, TS657 and TS658. *Parent samples of TS463.

| Clone	Tea type	Catechin content (µg/mg)	CHS2 exon2		PAL exon2
			*Rsa*I profile	*Rca*I profile	*Kpn*I profile
TV19*	Cambod	29	PR1	PRc2	PK1
TS463 1M		29	PR1	PRc2	PK1
TS463 2M		30	PR1	PRc2	PK1
TS463 3M		28	PR1	PRc2	PK1
TS463 4M		31	PR1	PRc2	PK1
TS463 5M		30	PR1	PRc2	PK1
TS463 6M		29	PR1	PRc2	PK1
TS463 7M		28	PR1	PRc2	PK1
TS463 8M		30	PR1	PRc2	PK1
TS463 9M		30	PR1	PRc2	PK1
TS463 10M		31	PR1	PRc2	PK2
TS463 11M		30	PR1	PRc2	PK1
TS463 12M		29	PR1	PRc2	PK1
TS463 13M		29	PR1	PRc2	PK1
TS463 14M		30	PR1	PRc2	PK1
TS463 15M		31	PR1	PRc2	PK1
TS463 16M		29	PR1	PRc2	PK1
TS463 17M		30	PR1	PRc2	PK2
TS463 18M		31	PR1	PRc2	PK1
TS463 19M		30	PR1	PRc2	PK1
TS463 20M		31	PR1	PRc2	PK1
TV1*	Assam-China Hybrid	30	PR1	PRc2	PK2
TS656 C1		32	PR2	PRc2	PK1
TS656 C2		31	PR2	PRc2	PK1
TS656 C3		32	PR2	PRc2	PK1
TS656 C4		31	PR2	PRc2	PK1
TS656 C5		32	PR2	PRc2	PK1
TS656 C6		31	PR2	PRc2	PK1
TS656 C7		32	PR2	PRc2	PK1
TS656 C8		33	PR2	PRc2	PK1
TS656 C9		32	PR2	PRc2	PK1
TS656 C10		31	PR2	PRc2	PK1
TS656 C11		32	PR2	PRc2	PK1
TS656 C12		31	PR2	PRc2	PK1
TS656 C13		31	PR2	PRc2	PK1
TS656 C14		31	PR2	PRc2	PK1
TS656 C15		31	PR2	PRc2	PK1
TS656 C16		32	PR2	PRc2	PK1
TS656 C17		32	PR2	PRc2	PK1
TS656 C18		31	PR2	PRc2	PK1
TS656 C19		32	PR2	PRc2	PK1
TS656 C20		31	PR2	PRc2	PK1
TS657 C1		30	PR1	PRc2	PK1
TS657 C2		29	PR1	PRc2	PK1
TS657 C3		29	PR1	PRc2	PK1
TS657 C4	Assam-China Hybrid	30	PR1	PRc2	PK1
TS657 C5		31	PR1	PRc2	PK1
TS657 C6		29	PR1	PRc2	PK1
TS657 C7		29	PR1	PRc2	PK1
TS657 C8		30	PR1	PRc2	PK1
TS657 C9		30	PR1	PRc2	PK1
TS657 C10		29	PR1	PRc2	PK1
TS657 C11		30	PR1	PRc2	PK1
TS657 C12		30	PR1	PRc2	PK1
TS657 C13		29	PR1	PRc2	PK1
TS657 C14		30	PR1	PRc2	PK1
TS657 C15		30	PR1	PRc2	PK1
TS657 C16		30	PR1	PRc2	PK1
TS657 C17		31	PR1	PRc2	PK1
TS657 C18		29	PR1	PRc2	PK1
TS657 C19		29	PR1	PRc2	PK1
TS657 C20		29	PR1	PRc2	PK1
TS658 C1		33	PR2	PRc2	PK1
TS658 C2		32	PR2	PRc2	PK1

| Table 1. Continued.					
TS658 C3		32	PR2	PRc2	PK1
TS658 C4		33	PR2	PRc2	PK1
TS658 C5		32	PR2	PRc2	PK1
TS658 C6		32	PR2	PRc2	PK1
TS658 C7		33	PR2	PRc2	PK1
TS658 C8		31	PR2	PRc2	PK1
TS658 C9		32	PR2	PRc2	PK1
TS658 C10		32	PR2	PRc2	PK1
TS658 C11		32	PR2	PRc2	PK1
TS658 C12		33	PR2	PRc2	PK1
TS658 C13		31	PR2	PRc2	PK1
TS658 C14		32	PR2	PRc2	PK1
TS658 C15		32	PR2	PRc2	PK1
TS658 C16		31	PR2	PRc2	PK1
TS658 C17		31	PR2	PRc2	PK1
TS658 C18		32	PR2	PRc2	PK1
TS658 C19		32	PR2	PRc2	PK1
TS658 C20		32	PR2	PRc2	PK1

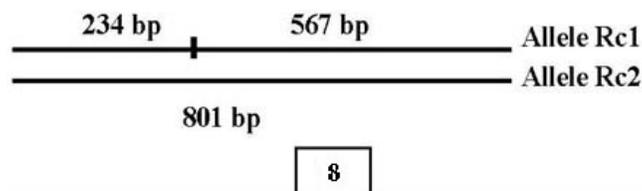## Development of CAPS markers in *CHS2* exon 2 digested with *Rca*I

Characteristic DNA patterns obtained by digesting the *CHS2* exon 2 amplicon with *Rca*I showed two different allelic forms because only one allele contained a restriction site for this enzyme (Figure 6). The homozygote for the restriction site was assigned Rc1Rc1 and the heterozygote with the absence of a restriction site was designated Rc1Rc2 (Figures 7 and 8). We named the profile for homozygosity at the *Rca*I restriction site as PRc1 and the profile for heterozygosity PRc2. In TS463, both parents and all twenty progenies showed a heterozygous profile PRc2. In TS379, the parent plant 14.5.35 produced two fragments while the other parent 14.12.16 revealed a heterozygous profile. Eleven progenies in TS379 showed the profile PRc1 while nine progenies showed the profile PRc2. The *CHS2* exon 2 amplicon restriction digest pattern in all three clonal seed stocks TS656, TS657, and TS658 had a heterozygous profile PRc2 (Table 1).
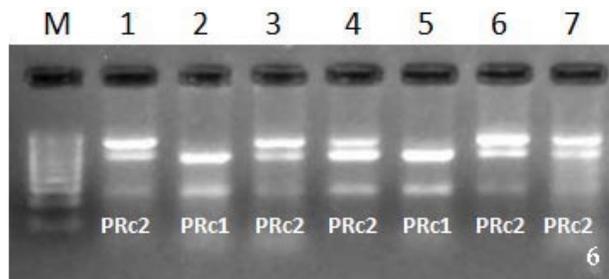
Figure 6. Amplicon restriction polymorphism observed in the *CHS2* exon 2 amplicon digested with *Rca*I identifying the PRc1 and PRc2 profiles in TS 379. *Lane M* = 100-bp ladder.
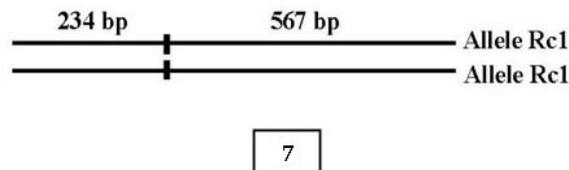
Figure 7. Restriction site map of the *CHS2* exon 2 digested with *Rca*I in tea samples with the PRc1amplicon restriction profile. (Distances not to scale).

Figure 8. Restriction site map of the *CHS2* exon 2 digested with *Rca*I in tea samples with the PRc2 amplicon restriction profile. (Distances not to scale).

Development of a CAPS marker in *PAL* exon 2 digested with *Kpn*I

A *Kpn*I restriction site in *PAL* was unevenly distributed across the samples studied, and found to be completely absent in most of the samples. In parent sample TV19 (TS463), the CAPS profile PK1 of the *PAL* exon 2 digested with *Kpn*I showed two restriction fragments of 385 and 959 bp (Figure 9). Parent sample TV1 (TS463) and two progenies, 10M and 17M, produced four fragments of 173, 208, 264, and 698 bp, respectively, when digested with *Kpn*I. This profile was designated PK2. The rest of the progenies showed profile PK1 (Figure 10). *Kpn*I digestion of the *PAL* exon 2 in samples belonging to seed stocks TS379 and TS656, and TS657 and TS658 (Table 1) showed two fragments consistent with the PK1 profile.
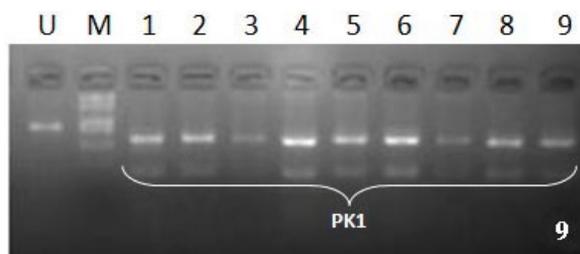
Figure 9. Amplicon restriction polymorphism observed in the *PAL* exon 2 amplicon digested with *Kpn*I identifying the PK1 profile in TS656. M=λ DNA/*Eco*RI+*Hin*dIII marker. U = Undigested amplicon.
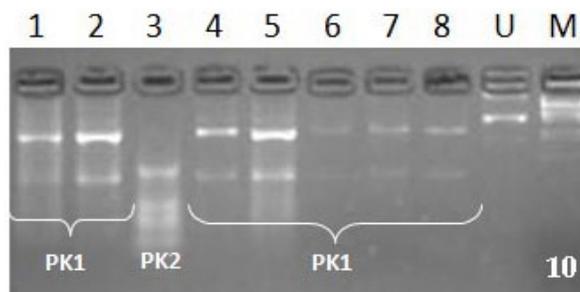
Figure 10. Amplicon restriction polymorphism observed in the *PAL* exon 2 amplicon digested with *Kpn*I with both the PKI and PK2 profiles in TS463. M = λ DNA/*Eco*RI+*Hind*III marker. U = Undigested amplicon.

Estimation of catechin content and its relation to the CAPS markers

Data on catechin content in the samples studied are presented in Tables 1 and 2. Our

goal was to link the catechin content in all the seed stocks with the restriction digest profiles characterized for *CHS2* and *PAL*. We observed that only profiles generated by *Rca*I digestion of the *CHS2* exon 2 showed a relationship with the level of catechin content. The samples with a homozygous digest pattern had relatively lower catechin content when compared to samples with a heterozygote digest pattern. The homozygous TS379 samples 14.5.35, 7A, 8A, 9A, 10A, 14A, 15A, 17A, 18A, and 19A had an average catechin content of 13.44 ± 1.119 mg/mg of fresh leaf tissue (Table 2). The heterozygote lines 14.12.16, 1A, 2A, 3A, 4A, 5A, 6A, 11A, 12A, 13A, 16A, and 20A had a relatively higher average catechin content of 30 ± 0.2599 mg/mg leaf tissue. Therefore, the difference of catechin content in homozygotes and that of heterozygotes was significantly different as confirmed by the Student *t*-test, ($t = 16.85$; P < 0.0001). Regression analysis revealed a polynomial relationship between the level of catechin and the allelic composition of *CHS2*+*Rca*I denoted by $y = 0.0027x^2 - 0.0593x + 2.294$; $R^2 = 0.9788$. All tea samples in TS463, TS656, TS657, and TS658 showed a heterozygous *CHS2* exon 2 amplicon *Rca*I digest profile with an average catechin content of 30, 31, 29, and 32 mg/mg, respectively.

Table 2. Relationship between catechin content and amplicon restriction profiles of DNA regions studied in TS379.

TS379	Tea type	Catechin content (μg/mg)	CHS2 exon2		PAL exon 2
			*Rsa*I Profile	*Rca*I Profile	*Kpn*I profile
14.5.35*	China	15	PR2	PRc1	PK1
14.12.16*	China	30	PR1	PRc2	
TS379 1A		30	PR1	PRc2	
TS379 2A		29	PR1	PRc2	
TS379 3A		31	PR1	PRc2	
TS379 4A		29	PR2	PRc2	
TS379 5A		30	PR2	PRc2	
TS379 6A		32	PR2	PRc2	
TS379 7A		15	PR1	PRc1	
TS379 8A		17	PR2	PRc1	
TS379 9A		9	PR1	PRc1	
TS379 10A		15	PR2	PRc1	
TS379 11A		31	PR1	PRc2	
TS379 12A		31	PR2	PRc2	
TS379 13A		31	PR2	PRc2	
TS379 14A		18	PR1	PRc1	
TS379 15A		14	PR1	PRc1	
TS379 16A		31	PR1	PRc2	
TS379 17A		14	PR1	PRc1	
TS379 18A		10	PR2	PRc1	
TS379 19A		9	PR1	PRc1	
TS379 20A		30	PR2	PRc2	
Average catechin content for samples					
Homozygous for *Rca*I		13.44 ± 1.119 μg/mg			
Heterozygous for *Rca*I		30.42 ± 0.2599 μg/mg			

The means for samples homozygous and heterozygous for *Rca*I digestion sites in CHS2 exon 2 were tested for an association with catechin content using the Student t-test (P < 0.0001). *Rsa*I and *Kpn*I profiles did not show any relationship with catechin content. *Parent samples of TS379.

DISCUSSION

Catechin content is used to determine tea quality and as a biochemical marker to study the diversity of tea germplasm. Accumulation of catechin has been found to differ the China, Assam, and Cambod cultivars. Catechin content has been reported to be lower in China types than in Assam and Cambod types (Takeda, 1994; Magoma et al., 2000; Gulati et al., 2009). Further, Iwasa (1977) reported that catechin content is higher in young leaves of var. *assamica* compared to older leaves and the var. *sinensis*. Variations in catechin content have also been observed among Japanese tea cultivars used in different tea products (Maeda-Yamamoto et al., 2001).

Catechin is one of the most important flavonoids synthesized via the phenyl propanoid pathway. *PAL* and *CHS* are two essential genes in the pathway, which plays a major role in the determination of tea taste and quality. Tea growers have always endeavored to develop superior clones for their tea gardens. In the present study, molecular markers for early screening of clones with higher catechin content were developed. The advantage of the CAPS approach is that the markers are developed from expressed genes (ESTs) that can be inherited by future clones. In this case, since catechin is produced through the secondary metabolite pathway, *PAL* and *CHS* were targeted for the development of CAPS markers. Both genes operate upstream of the final flavonoid production; we hypothesized that the marker(s) identified may have close association with catechin content. Further, CAPS markers would help us distinguish plants that were homozygous and heterozygous for the *CHS2* gene. We do not know whether catechin content would be affected by heterosis or not. An analysis of heterozygosity may also enable us to examine this aspect of catechin content.

No suitable markers have been thus far developed for screening tea clones for catechin content at the nursery level. To our knowledge, this is the first attempt to systematically develop such markers utilizing targeted genes involved in the secondary metabolite pathway. This approach reduces dependence on chance factors associated with random markers that need further statistical validation using QTL, making our approach novel and more reliable. CAPS can also be used for many other markers for quality traits associated with biochemical pathways. This approach is likely to succeed with quality traits that are under control of few genes with major effects, rather than quantitative traits controlled by polygenes with small individual effects.

CAPS analysis of the samples revealed that out of the two genes assessed, *CHS2* exhibited more useful polymorphisms than the *PAL* gene exon. For *PAL* exon 2 digestion with *Kpn*I, only TV1 (TS463) and its F_2 progenies 10M and 17M generated the profile PK2. TV19 (TS463) and the rest of the progenies, TS379, TS656, TS657, and TS658, generated the profile PK1. TV1, an 'Assam-China' hybrid, differed in the *Kpn*I restriction digest profile from TV19, which is a 'Cambod' type and also from 14.5.35 and 14.12.16, both 'China' types. TV1 is maintained as a quality clone in TRA and from the amplicon restriction profile of TV1 and is different from the rest of the samples studied. From this, we can conclude that the CAPS markers developed from the *PAL* exon 2 *Kpn*I digestion can serve as a marker for the TV1clone.

Comparison of the average catechin content with the CAPS markers did not reveal any meaningful association between the amplicon restriction profiles of *PAL* and catechin content. However, we found a codominant CAPS marker in the *Rca*I digestion of the *CHS2* exon 2 that showed a positive relationship between the heterozygosity for the enzyme site and the catechin content in tea. It was observed that the samples heterozygous for the *Rca*I site had relatively higher average catechin content than the samples that were homozygous for this restriction site. Further investigations will confirm whether this association is due to a heterotic effect for this

character. Despite screening 104 plants, we did not identify any plants homozygous for allele Rc2Rc2 and have not definitively shown the heterotic effect. However, we recommend that for breeding programs, lines showing heterozygosity for *CHS2* exon 2 alleles could be used for selecting clones with higher catechin content. Alternatively, the profiles with homozygosity for Rc1Rc1 may be used for rejecting clones with low catechin content. Tea clones TV1, TV19, 14.12.16, TS656, TS657, and TS658, which show heterozygosity for the *Rca*I restriction site, can also be used for the production of higher grade tea and can be included in future breeding programs tailored for higher catechin content. Our markers can be used for selecting clones in segregating populations and we provide evidence that the *CHS2* gene may be used as a genetic determinant for catechin content. Our CAPS markers, which distinguished plants with higher catechin content from those with low catechin content, will allow breeders to economically distinguish the plants at the seedling stage.

CONCLUSION

The present study utilized a targeted approach for developing molecular markers for marker-assisted selection of tea clones, wherein two important secondary metabolite pathway genes were identified for this purpose. A close association of the restriction patterns of the amplified genes obtained with an important quality trait revealed that heterozygosity for particular restriction sites was closely associated with catechin content. Thus, the CAPS markers developed herein can be effectively used by tea breeders for selection of tea clones with higher catechin content.

ACKNOWLEDGMENTS

REFERENCES

Anan T and Nakagawa M (1974). Effect of light on chemical constituents in the tea leaves. *J.Agr. Chem. Soc. Japan* 48: 91-96.

Banerjee B (1992). Botanical classification of tea. In: Tea cultivation to consumption (Wilson KC and Clifford MN, eds). Springer Netherlands, 25-51.

Bang H, Kim S, Leskovar D and King S (2007). Development of a codominant CAPS marker for allelic selection between canary yellow and red watermelon based on SNP in lycopene β-cyclase (LCYB) gene. *Mol. Breed.* 20: 63-72. http://dx.doi.org/10.1007/s11032-006-9076-4

Ben MS, Ben Tamarzizt H, Baraket G, Abdallah D, et al. (2015). Genetic diversity and differentiation in *Prunus* species (Rosaceae) using chloroplast and mitochondrial DNA CAPS markers. *Genet. Mol. Res.* 14: 4177-4188. http://dx.doi.org/10.4238/2015.April.27.33

Chauhan VS and Misra AK (2002). Development of molecular markers for screening of Alnus nepalensis (D. Don) genotypes for the nitrogenase activity of actinorhizal root nodules. *Mol. Genet. Genomics* 267: 303-312. http://dx.doi.org/10.1007/s00438-002-0645-x

Dillon S, Ramage C, Ashmore S and Drew RA (2006). Development of a codominant CAPS marker linked to PRSV-P resistance in highland papaya. *Theor. Appl. Genet.* 113: 1159-1169. http://dx.doi.org/10.1007/s00122-006-0375-2

Gershenzon J (2002). Secondary metabolites and plant defense. In: Plant physiology (Taiz L and Zeiger E, eds.) Sinauer Associates Inc., Sunderland, MA, 283-308.

Gulati A, Rajkumar S, Karthigeyan S, Sud RK, et al. (2009). Catechin and catechin fractions as biochemical markers to study the diversity of Indian tea (*Camellia sinensis* (L.) O. Kuntze) germplasm. *Chem. Biodivers.* 6: 1042-1052. http://dx.doi.org/10.1002/cbdv.200800122

Gutierrez N, Avila CM, Duc G, Marget P, et al. (2006). CAPs markers to assist selection for low vicine and convicine contents in faba bean (*Vicia faba* L.). *Theor. Appl. Genet.* 114: 59-66. http://dx.doi.org/10.1007/s00122-006-0410-3

Iwasa K (1977). Biosynthesis of catechins in tea plant. *Bull. Natl. Res. Inst. Tea.* 13: 101-126.

Kaundun SS and Matsumoto S (2003). Development of CAPS markers based on three key genes of the phenylpropanoid pathway in tea, *Camellia sinensis* (L.) O. Kuntze, and differentiation between assamica and sinensis varieties. *Theor. Appl. Genet.* 106: 375-383.

Konieczny A and Ausubel FM (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* 4: 403-410. http://dx.doi.org/10.1046/j.1365-313X.1993.04020403.x

Kuklev MY, Fesenko IA and Karlov GI (2009). [Development of a CAPS marker for the Verticillium wilt resistance in tomatoes]. *Russian J. Genet.* 45: 656-661.

Maeda-Yamamoto M, Sano M, Matsuda N, Miyase T, et al. (2001). The change of epigallocatechin-3-*O*-(3-*O*-methyl) gallate content in tea of different varieties, tea seasons of crop and processing method. *J. Japanese Soc. of Food Sci. Tech* 48: 64-68. http://dx.doi.org/10.3136/nskkk.48.64

Magoma GN, Wachira FN, Obanda M, Imbuga M, et al. (2000). The use of catechins as biochemical markers in diversity studies of tea (*Camellia sinensis*). *Genet. Resour. Crop Evol.* 47: 107-114. http://dx.doi.org/10.1023/A:1008772902917

Matsumoto A and Tsumura Y (2004). Evaluation of cleaved amplified polymorphic sequence markers for Chamaecyparis obtusa based on expressed sequence tag information from Cryptomeria japonica. *Theor. Appl. Genet.* 110: 80-91. http://dx.doi.org/10.1007/s00122-004-1754-1

Matsumoto S, Takeuchi A, Hayatsu M and Kondo S (1994). Molecular cloning of phenylalanine ammonia-lyase cDNA and classification of varieties and cultivars of tea plants (Camellia sinensis) using the tea PAL cDNA probe. *Theor. Appl. Genet.* 89: 671-675. http://dx.doi.org/10.1007/BF00223703

Obanda M, Owuor PO and Taylor SJ (1997). Flavanol composition and caffeine content of green leaf as quality potential indicators of Kenyan black teas. *J. Sci. Food Agric.* 74: 209-215. http://dx.doi.org/10.1002/(SICI)1097-0010(199706)74:2<209::AID-JSFA789>3.0.CO;2-4

Porebski S, Bailey LG and Baum BR (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol component. *Plant Mol. Biol. Report.* 15: 8-15. http://dx.doi.org/10.1007/BF02772108

Semagn K, Bjornstad A and Ndjiondjop MN (2006). An overview of molecular marker methods for plants. *Afr. J. Biotechnol.* 5: 2540-2568.

Singh HP, Ravindranath SD and Singh C (1999). Analysis of tea shoot catechins: spectrophotometric quantitation and selective visualization on two-dimensional paper chromatograms using diazotized sulfanilamide. *J. Agric. Food Chem.* 47: 1041-1045. http://dx.doi.org/10.1021/jf9807263

Takeda Y (1994). Differences in caffeine and tannin contents between tea cultivars and application to tea breeding. *Jpn. Agric. Res. Q.* 28: 117-123.

Takeuchi A, Matsumoto S and Hayatsu M (1994). Chalcone synthase from *Camellia sinensis:* isolation of the cDNAs and the organ-specific and sugar-responsive expression of the genes. *Plant Cell Physiol.* 35: 1011-1018.

Verghese SK and Misra AK (2000). PCR-RFLP based screening of *Frankia* in Alder nodules having different levels of nitrogenase activity. *Symbiosis* 28: 337-350.

Yanthan M and Misra AK (2013). Amplicon restriction patterns associated with nitrogenase activity of root nodules for selection of superior *Myrica* seedlings. *J. Biosci.* 38: 789-795. http://dx.doi.org/10.1007/s12038-013-9358-z