Theor Appl Genet (1994) 89: 671–675

S. Matsumoto · A. Takeuchi · M. Hayatsu · S. Kondo

Molecular cloning of phenylalanine ammonia-lyase cDNA and classification of varieties and cultivars of tea plants (*Camellia sinensis*) using the tea PAL cDNA probe

Received: 25 January 1994 / Accepted 18 April 1994

Abstract Tea (*Camellia sinensis*) phenylalanine ammonia-lyase (PAL) cDNA was cloned using labelled rice PAL cDNA as a probe. The PAL genes of the tea plant were investigated by restriction fragment length polymorphism (RFLP) analysis using tea PAL cDNA. PAL genetic variation in tea plants was much larger than predicted due to the presence of various hybridized fragments in the Assam hybrids, which are hybrids between *C. sinensis* var 'assamica' and var 'sinensis'. On the other hand, hybridized band patterns of Japanese green tea cultivars belonging to var 'sinensis' could be divided into five groups. Furthermore, a short-length PAL probe, about 280 bp including the 3′ untranslated sequence, detected 3 DNA fragments of different lengths, which were named A, B and D. An experiment tracing the PAL gene heredity showed that A, B and D fragments were inherited according to the Mendelian monogenic ratio. Therefore, PAL genes identifiable by A, B and D fragments are multiple alleles, and the PAL gene is present as a single gene in the tea haploid genome. It was also clear that five groups of Japanese green tea cultivars were characterized by the composition of these PAL fragments. From RFLP analysis using tea PAL cDNA, we succeeded in distinguishing Assam hybrids and Japanese green tea cultivars with high and low catechin content, respectively, and in grouping Japanese green tea at the cultivar level.

Key words Phenylalanine ammonia-lyase
Camellia sinensis · RFLPs · Molecular marker

Introduction

Tea plants are distributed throughout India, China, Japan and other regions of Asia where the tea industry has been developed. Though the origin of the various types of tea plants is not clear, during the extension process tea plants differentiated not only by morphological traits but also by characteristic compounds contained in their leaves. Today, on the basis of these morphological characters, *Camellia sinensis* is botanically divided into four varieties (Chang and Bartholomew 1984). Two varieties, 'assamica' and 'sinensis', are the principle tea producers. The former was developed in India and the latter in China; they were subsequently introduced into Japan. However, since reproductive isolation has not been present between them, a number of intermediate types have been cultivated. These intermediate types are used for producing distinct tea products, which are dependent on the characteristic compounds contained in their leaves. For example, var 'assamica' and Assam hybrids are mainly utilized for black tea and var 'sinensis' for green tea and other tea products.

Many characteristic compounds related to tea quality have been detected in tea leaves (Nagata and Sakai 1984). One of these, catechin, which amounts to 10–30% of the dry weight of tea leaves, is a main component of astringency and a key determinant of tea quality (Iwasa 1968; Anan and Nakagawa 1974). Catechin content varies in varieties or cultivars; var 'assamica' and Assam hybrids contain a higher proportion than var 'sinensis'. Therefore, regulation of catechin content is very important in tea breeding and cultivation. This diversity in catechin content among tea cultivars may be caused by the different activity of the enzymes related to its biosynthesis. Of these enzymes, phenylalanine ammonia-lyase (PAL) catalyzes the deamination of L-phenylalanine to yield *trans*-cinnamic acid and NH_4^+, which is the first step in phenylpropanoid synthesis leading to catechin production. It has been reported that PAL activity in young tea leaves, which contain abundant catechin, is higher than in mature ones, which have a low catechin content (Iwasa 1977). Both the activity and content of catechin are controlled by light intensity (Saijo 1980). In addition, higher activity has also been observed in leaves of cultivars belonging to var 'assamica' and Assam hybrids than in Japanese green tea cultivars (Iwasa 1977). Therefore, PAL activity in

Communicated by G. Wenzel

S. Matsumoto (✉) · A. Takeuchi · M. Hayatsu · S. Kondo
National Research Institute of Vegetables,
Ornamental Plants and Tea, 2769 Kanaya, Shizuoka, Japan

tea leaves must have some connection with catechin content.

For tea breeding to control the catechin content, a technique must be established that is able to assess the catechin biosynthetic ability of various tea clones and to classify tea plants into useful genetic resources. It is a reasonable assumption to define and use the PAL gene as a genetic marker for the assessment of tea genetic resources. Therefore, a full length PAL cDNA was cloned and, PAL restriction fragment length polymorphisms (RFLPs) of the tea plants were detected. On the basis of these RFLPs, two tea leading varieties, var 'assamica' and var 'sinensis', could be distinguished. Furthermore, we showed that PAL of the green tea cultivars tested consisted of three multiple alleles on a single locus and that green tea plants could be classified at the cultivar level using these PAL gene compositions.

Materials and methods

Plant materials

Except for 'Karabeni' and 'Benitsukuba', the tea plants used in the present study were Japanese green tea cultivars and Assam hybrids. Japanese green tea cultivars belonging to var 'sinensis' are domestic varieties obtained from cultivated tea farms in Japan or hybrids of these varieties that are used as breeding materials: 'Yabukita', 'Kanayamidori', 'Tamamidori', 'Meiryoku', 'Kuritawase', 'Shunmei', 'Asatsuyu', 'Asagiri', 'Yaeho', 'Sayamakaori', 'Saemidori', 'Fushun', 'Natsumidori' and 'Z1'. Assam hybrid cultivars were 'Indo', 'Benikaori', 'Benitachiwase', 'Benifuji', 'Benihikari', 'Satsumabeni', 'Tadanishiki' and 'Akane'. 'Karabeni' is a variety introduced from China and may not be an Assam hybrid. 'Benitsukuba' is a black tea cultivar but a domestic variety in Japan. 'Assam' and 'Inzatsu 131' are Assam hybrid lines. These genetic materials were obtained from stocks maintained at the National Research Institute of Vegetables and Ornamental Plants and Tea, Kanaya, Shizuoka, Japan. Tea leaves used for DNA extraction were harvested in the field under normal agricultural conditions.

Construction of cDNA library, cloning and sequencing of PAL cDNA

For construction of the cDNA library, young tea (*C. sinensis* var 'sinensis' cv. 'Yabukita') leaves were harvested from bushes grown in the field, on a sunny afternoon in April, because the PAL activity of tea leaves is high under this condition. Total RNA was isolated from pulverized frozen leaves by a modification of the SDS-phenol method (Palmiter 1974). Polyadenylated RNA was extracted using the mRNA Purification Kit (Pharmacia). Double-stranded cDNA was synthesized using the cDNA Synthesis Kit (Boehringer Mannheim) and was inserted into the unique *Eco*RI site of the lambda gt11 vector using *Eco*RI linker addition. The cDNA library was screened using rice PAL cDNA (Minami et al. 1989) labelled with α-[^{32}P]-dCTP by the Random Primer DNA Labeling Kit (Takara Shuzo Co, Kyoto, Japan). The screening was carried out according to Sambrook et al. (1989). An insert of the positive clone was subcloned into pBluescript KS+ for further analysis. Sequencing reactions were carried out using the Taq Dye Primer Cycle Sequencing Kit (Applied Biosystems). A DNA Sequencer Model 373A (Applied Biosystems) was used for DNA sequence determination. Nucleotide and amino acid sequences were analyzed using the GENETYX software system (Software Development Co, Tokyo, Japan).

DNA preparation and RFLP analysis

Plant DNA was prepared from frozen leaf tissues according to Guillemanut et al. (1992) and some modified procedures. Fifteen microgram DNA extracted from tea cultivars was digested with *Hind*III or *Eco*RV and resolved in 0.8% agarose gel. DNA fragments were transferred to a nylon membrane and were fixed by exposure to UV light. The cloned tea PAL cDNA was labelled with α-[^{32}P]dCTP. The other two short probes that are partial PAL cDNAs which included the 5′ (N350) and 3′ untranslated regions (C280) were used. N350 and C280 contain about 200 bp of translated regions. Hybridization was carried out overnight at 65°C. After hybridization, the membranes were rinsed twice for 5 min in 2×SSC and washed at 65°C for 30 min in 0.2×SSC-0.1%SDS before being exposed to X-ray film with intensifying screens at –80°C for 3–5 days.

Results

Cloning of PAL cDNA and sequence determination

In the first screening using labelled rice PAL cDNA as a probe, 2 positive clones containing a short insert were obtained. The insert of 1 of these clone showed a higher DNA sequence homology with other plants cloned PAL sequence, and this was taken to indicate that the insert was part of the tea PAL cDNA. The insert DNA was used as a probe to obtain another clone with a longer insert. After hybridization and detection, 21 clones showed a positive signal on X-ray film. A clone with the largest insert was selected and designated as A21. The insert contained 2344 nucleotide sequences including a poly-A tail, and the open reading frame consisted of 2142 bp (data not shown; the nucleotide sequence will appear in the GSDB, DDBJ, EMBL and NCBI nucleotide sequence database with the accession number D26596). The molecular mass was calculated on the basis of the deduced amino acid sequence to be 77.750, which is similar in size to those of other plant PALs. A comparison of the deduced amino acid sequence of the insert with those of other plants showed high homology with dicotyledonous plants, i.e. 87.3%, 84.3%, and 82.2% of the amino acids of the A21 insert were homologous with those of pea (Kawamata et al. 1992), poplar (Subramaniam et al. 1993) and sweet potato (Tanaka et al. 1989), respectively. However, the homology was lower in monocotyledonous and gymnospermous plants, and the actual figures we found were 69.0% in rice (Minami et al. 1989) and 67.5% in pine (Whetten and Sederoff 1992). These results indicate that the insert DNA of A21 encoded the information of tea PAL cDNA.

Genetic variation of PAL in tea plants

To assess the genetic variation of PAL in tea plants, approximately 2.3 kbp of insert DNA of A21 was used as a probe. RFLP analysis detected 2 or 3 DNA fragments between about 5.3 and 7.3 kbp in Japanese green tea cultivars, and the 5.9 kbp fragment was found to be common to all of the tested cultivars (Fig. 1a). On the basis of the numbers and lengths of the detected fragments, we could

Table 1 Classification of Japanese green tea cultivars based on RFLP analysis. *Hin*dIII and *Eco*RV were used for DNA digestion (+ Detected hybridized bands)

*Hin*dIII					Cultivars	*Eco*RV					
5.3	5.9	6.7	7.3 kb[a]	Type		0.2	1.6	1.8	3.0	3.3	6.8 kb[a]
D	C	B	A								
	+		+	AA	Tamamidori, Kuritawase, Z1	+	+		+		+
	+	+	+	AB	Kanayamidori, Yaeho, Fushun	+	+	+	+	+	+
+	+		+	AD	Meiryoku, Shunmei, Asatsuyu	+	+	+	+		+
+	+	+		BD	Yabukita, Sayamakaori, Saemidori	+		+	+	+	+
+	+			DD	Asagiri, Natsumidori	+		+	+		+

[a] The lengths of the detected hybridized bands were estimated on the basis of the DNA molecular marker

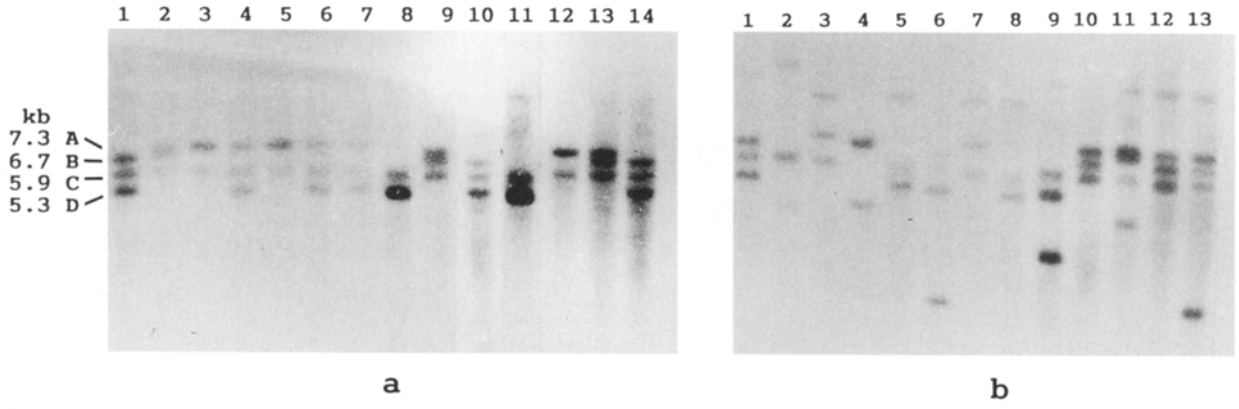

Fig. 1 RFLPs comparison of *C. sinensis* var 'sinensis' (**a**) with Assam hybrids (**b**) and 'Karabeni' and 'Benitsukuba'. The DNA was digested with *Hin*dIII, and the probe was 2.3-kb PAL cDNA. **a** *Lane 1* 'Yabukita', *2* 'Kanayamidori', *3* 'Tamamidori', *4* 'Meiryoku', *5* 'Kuritawase', *6* 'Shunmei', *7* 'Asatsuyu', *8* 'Asagiri', *9* 'Yaeho', *10* 'Sayamakaori', *11* 'Natsumidori', *12* 'Z1', *13* 'Fushun', *14* 'Saemidori'. **b** *Lane 1* 'Yabukita' (Cont.), *2* 'Assam', *3* 'Inzatsu131', *4* 'Indo', *5* 'Benitsukuba', *6* 'Benikaori', *7* 'Benitachiwase', *8* 'Benifuji', *9* 'Benihikari', *10* 'Satsumabeni', *11* 'Karabeni', *12* 'Tadanishiki', *13* 'Akane'

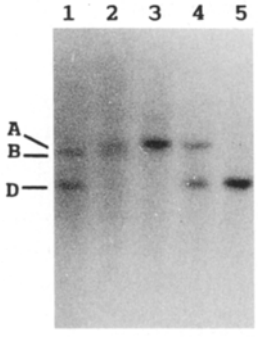

Fig. 2 Different hybridized fragment patterns (*A*, *B*, *D*) of 5 tea cultivars using C280 as a probe. *Lane 1* 'Yabukita', *2* 'Kanayamidori', *3* 'Tamamidori', *4* 'Meiryoku', *5* 'Asagiri'

classify Japanese green tea cultivars into five groups (Table 1). A further RFLP analysis using *Eco*RV divided these green tea cultivars into the same five groups consisting of the same cultivars as a result of *Hin*dIII digestion (data not shown but the result is summarized in Table 1). To distinguish between var 'sinensis' and var 'assamica' using the PAL gene, we compared hybridized band patterns between Japanese green tea cultivars and Assam hybrids and the black tea cultivars. On the other hand, Assam hybrids and 'Benitsukuba' and 'Karabeni' showed various fragment patterns on most cultivars (Fig. 1b); these patterns were distinct from those of Japanese green tea cultivars. The patterns of 'Satsumabeni' and 'Tadanishiki' were similar to those of Japanese green tea cultivars (Fig. 1b), but were different in another RFLP analysis using *Eco*RV as a restriction enzyme (data not shown). All of the Assam hybrids used in the present study could be distinguished from Japanese green tea cultivars by RFLPs.

Inheritance of PAL in tea plants

Two short probes, N350 and C280, were used for a subsequent analysis of Japanese green tea cultivars. Each fragment detected by RFLP analysis using *Hin*dIII was designated as A, B, C and D in order of longest (A) to shortest (D) (Fig. 1a). One green tea cultivar was selected from each of the classified groups based on RFLP analysis (Table 1) and in total, 5 cultivars were studied. From RFLP analysis using the N350 probe, the common fragment C, about 5.9 kbp, was the only one that emerged (data not shown). However, the three other kinds of fragments, A, B and D, were detected using the C280 probe and subsequently utilized for further classification (Fig. 2). These results suggest that the *Hin*dIII recognition site is different in the 3′ flanking region of the PAL gene but is the same in the 5′ flanking region. The *Hin*dIII recognition site, which separates the C fragment from the others, may be present in the intron sequence because it was not present in the cloned PAL cDNA.

Since the intensity of a single fragment detected in Fig. 2 was stronger than that of two others, it seemed that the former contained the PAL gene as a homozygote and

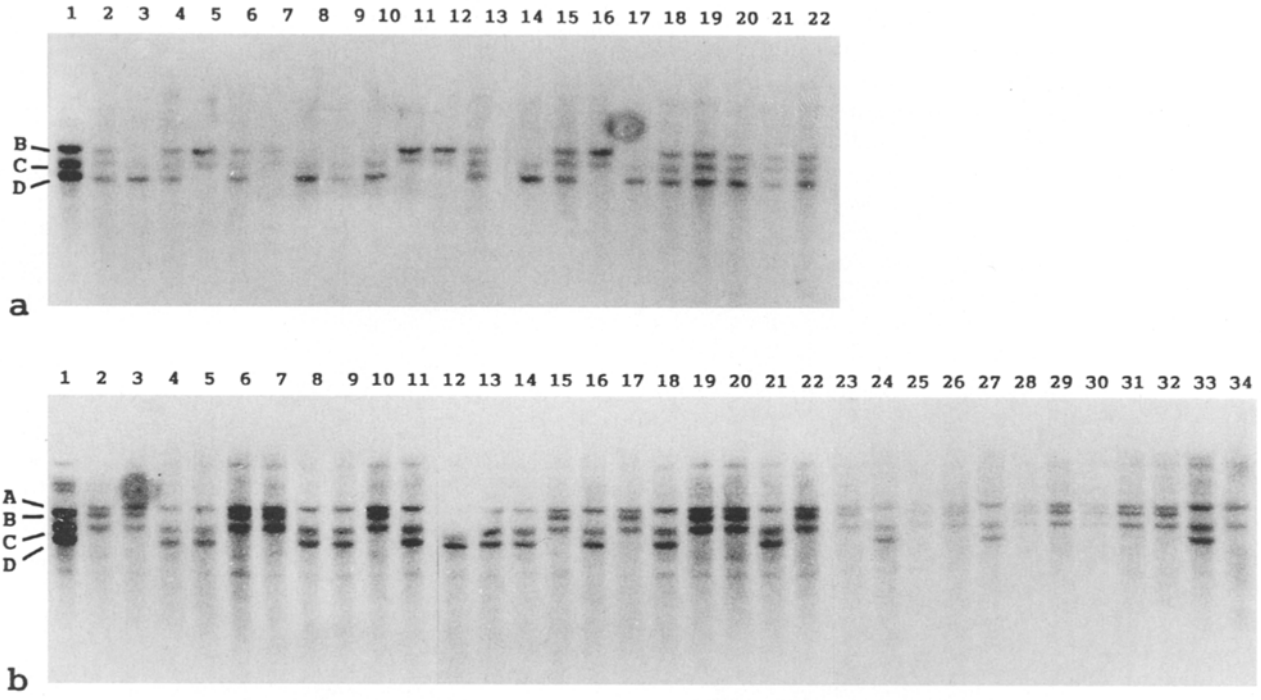

Fig. 3 a, b Segregation of PAL-specific RFLPs of 2 tea hybrid populations. The probe was the 2.3-kb PAL cDNA. **a** Hybrids (*lanes 2–21*) between 'Sayamakaori' (BD, *lane 1*) and 'Yabukita' (BD, *lane 22*). Twenty patterns of 22 hybrids assayed in this analysis are shown. **b** Hybrids (lanes 2–33) between 'Sayamakaori' (BD, *lane 1*) and 'Z1' (AA, *lane 34*)

Table 2 Segregation ratios of PAL cDNA hybridized fragments of tea plants

Parent cultivars	Number of hybrids			Total	χ^2	Value
	BB	BD	DD			
Sayamakaori(BD) ×Yabukita(BD)	5	12	5	22	0.18[a]	ns
	AB		AD			
Sayamakaori(BD) ×Z1(AA)	18		14	32	0.50[b]	ns

ns, not significant at P=0.05
[a] χ^2 for 1:2:1
[b] χ^2 for 1:1

the latter as a heterozygote. Therefore, we hypothesized that the A, B and D fragments are inherited as a single gene. To confirm this hypothesis, two progeny populations between 'Sayamakaori' (BD) and 'Yabukita' (BD), and between 'Sayamakaori' (BD) and 'Z1' (AA) were assayed. From the former population, each lane had, together with the C fragment, only the B, only the D or both the B and D fragments (Fig. 3a). Furthermore, only two fragment patterns, those of A and B or A and D, were observed from the latter population (Fig. 3b). These results are summarized in Table 2. A total of 22 hybrid plants of the former segregated 5 BB, 12 BD and 5 DD plants. Moreover, AB and AD plants eventually emerged from the latter. The A, B and D fragments were apparently multiple alleles distinguishable by RFLPs because these fragments were inherited as a single gene according to their Mendelian ratio. Since any additional hybridized fragments were not detected, the PAL gene of the tea plant was a single one per haploid genome. This finding indicates that the previously mentioned five groups are composed of the AA, AB, AD, BD and DD type (Table 1).

Discussion

Since PAL acts as a key enzyme in phenylpropanoid metabolism of plants, PAL cDNA or genes have been cloned from several plants, including the woody plants (Cramer et al. 1989; Joos and Hahlbrook 1992; Lois et al. 1989; Ohl et al. 1989, Gowri et al. 1991; Kawamata et al. 1992; Whetten and Sederoff 1992). A comparison of the amino acid sequence estimated from tea PAL cDNA with those of other plants showed that the tea plant has a higher homology to dicotyledons than to monocotyledons and gymnosperms. Although tea is the same woody type of plant as poplar and pine, this different homology matches the degree of evolutional distance of these plants. By RFLP analysis, we were able to reveal PAL genetic variation among tea plants whose catechin content was different. In Japanese green tea cultivars, five groups were determined (Table 1), whereas Assam hybrids could not be placed into any specific group because complex patterns were produced. Assam hybrids used in the present study were derived from crosses between var 'assamica' and Japanese domestic tea that have been made to breed black tea cultivars that can tolerate the cold climate of Japan. These hybrids contain

only one-half or one-quarter of the genes from var 'assamica'. In addition, the number of samples of the present study was not enough to detect the genetic PAL variation within all of var 'assamica'. In spite of these limitations, various PAL-hybridizing bands could be seen in most Assam hybrids, which indicates that PAL genetic variation is much wider than that we predicted. Moreover, 'Karabeni', which was derived from a Chinese wild-type tea belonging to var 'sinensis', showed different hybridizing bands from either the Assam hybrids or the Japanese green tea cultivars. An analysis extended to the Chinese wild-type tea plants would reveal the other different PALs. Therefore, it would be reasonable to speculate that a large number of PAL genes detectable by RFLPs might be distributed in tea plants.

Although in the Japanese green tea cultivars PAL is encoded by a single gene per haploid genome, RFLP analysis could distinguish at least three kinds of PAL genes, A, B and D. Three fragments were distinguished on the basis of the position of the *Hin*dIII recognition site in the 3′ flanking region of the PAL gene, but some other sequences also seemed to be present in the coding region of these genes. During the PAL cDNA cloning process, a clone with an incomplete length, which had a different nucleotide sequence from the insert of A21, was identified. Out of the 300 base pairs of nucleotide sequences of the clone, 37 base pairs were different, which caused 5 deduced amino acid conversions (data not shown). Since 'Yabukita' (BD) was used for the cDNA library construction, if A21 and the clone correspond to the B or D fragment, each of the PAL genes displayed by the B and D fragments has a different sequence with amino acid conversion. The diversity of the PAL gene in tea plants was shown in the present study. However, it is unknown whether this diversity arises not only in the flanking region but also in the coding region. The finding of two different sequences of PAL cDNA clones in 'Yabukita' suggests that the diversity could have arisen in the coding region of the PAL gene of the var 'assamica' or Assam hybrids.

We were able to classify Japanese green tea cultivars into five groups, AA, AB, AD, BD and DD. However, one question remains, why were no cultivars of the BB type found, which otherwise would have easily been obtained from a cross-hybridized population. If one of some inadequate phenotypes might be the cause of BB type formation, we could immediately use RFLP analysis for green tea breeding. Although the existence of many kinds of PAL genes distinguishable by RFLPs is predicted among the tea plants, only three PAL genes contribute to the development of Japanese green tea. We conclude that the key ancestries of Japanese green tea were derived from a few limited clones among all of the tea plants. Until now, many green tea cultivars were developed by crossing and selection from a genetically limited population.

Acknowledgements The authors thank Dr. E. Minami of National Institute of Agrobiological Resources, Tsukuba, Ibaragi, Japan, for the gift of the rice PAL cDNA probe. We also thank Prof. K. Arisumi of Kagoshima University and Prof. K. Takayanagi of Tsukuba University for careful reading of the manuscript.

References

Anan T, Nakagawa M (1974) Effect of light on chemical constituents in the tea leaves (in Japanese). Nippon Nogei Kagaku Kaishi 48:91–96

Chang HT, Bartholomew B (1984) Camellias. B.T.Batsford, London

Cramer CL, Edwards K, Dron M, Liang X, Dildine SL, Bolwell GP, Dixon RA, Lamb CJ, Schuch W (1989) Phenylalanine ammonia-lyase gene organization and structure. Plant Mol Biol 12:367–383

Gowri G, Paiva NL, Dixon RA (1991) Stress responses in alfalfa (*Medicago sativa* L.). 12. Sequence analysis of phenylalanine ammonia-lyase (PAL) cDNA clones and appearance of PAL transcripts in elicitor-treated cell cultures and developing plants. Plant Mol Biol 17:415–429

Guillemanut P, Marechal-Drouard L (1992) Isolation of plant DNA: a fast, inexpensive, and reliable method. Plant Mol Biol Rep 10:60–65

Iwasa K (1968) Influence of the shading culture on catechin composition in tea leaves (in Japanese). Study Tea 36:63–69

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

Joos HJ, Hahlbrock K (1992) Phenylalanine ammonia-lyase in potato (*Solanum tuberosum* L.)-genomic complexity, structural comparison of two selected genes and modes of expression. Eur J Biochem 204:621–629

Kawamata S, Yamada T, Tanaka Y, Sriprasertsak P, Kato H, Ichinose Y, Kato H, Shiraishi T, Oku H (1992) Molecular cloning of phenylalanine ammonia-lyase cDNA from *Pisum sativum*. Plant Mol Biol 20:167–170

Lois R, Dietrich A, Hahlbrock K, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive *cis*-acting elements. EMBO J 8:1641–1648

Minami E, Ozeki Y, Matsuoka M, Koizumi N, Tanaka Y (1989) Structure and some characterization of the gene for phenylalanine ammonia-lyase from rice plants. Eur J Biochem 185:19–25

Nagata T, Sakai S (1984) Differences in caffeine, flavanols and amino acids contents in leaves of cultivated species of *Camellia*. Jpn J Breed 34:459–467

Ohl S, Hedrick SA, Chory J, Lamb CJ (1990) Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. Plant Cell 2:837–848

Palmiter RD (1974) Magnesium precipitation of ribonucleoprotein complexes: expedient techniques for the isolation of undergraded polysomes and messenger ribonucleic acid. Biochemistry 13:3606–3615

Saijo R (1980) Effect on shade treatment on biosynthesis of catechins in tea plants. Plant Cell Physiol 21:989–998

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Subramaniam R, Reinold S, Molitor EK, Douglas CJ (1993) Structure, inheritance, and expression of hybrid poplar (*Populus trichocarpa* X *Populus deltoides*) phenylalanine ammonia-lyase genes. Plant Physiol 102:71–83

Tanaka Y, Matsuoka M, Yamamoto N, Ohashi Y, Kano-Murakami Y, Ozeki Y (1989) Structure and characterization of a cDNA clone for phenylalanine ammonia-lyase from cut-injured roots of sweet potato. Plant Physiol 90:1403–1407

Whetten RW, Sederoff RR (1992) Phenylalanine ammonia-lyase from loblolly pine. Plant Physiol 98:380–386