Journal of Plant Biology, June 2006, 49(3) : 200-204

Cloning, Characterization, and Expression of Two cDNA Clones for a Rice Ferulate-5-Hydroxylase Gene, a Cytochrome P450-Dependent Monooxygenase

Jin-Hong Kim[1], Dae Hwa Yang[1], Jae-Sung Kim[1], Myung-Hwa Baek[1], Young Mi Park[1], Seung Gon Wi[1], Jae-Young Cho[2], and Byung Yeoup Chung[1]*

[1]Division of Radiation Application Research, Korea Atomic Energy Research Institute, Jeongeup 580-185, Korea
[2]Division of Biological Resources Sciences, Chonbuk National University, Jeonju 561-756, Korea

Ferulate-5-hydroxylase (*F5H*) is a cytochrome P450-dependent monooxygenase that catalyses the hydroxylation of ferulic acid, coniferaldehyde, and coniferyl alcohol in the pathways leading to sinapic acid and the syringyl unit of lignin. We have isolated two *F5H* genes, *OsF5HL* and *OsF5HL2*, from a japonica-type rice (*Oryza sativa* L. cv. Ilpoombyeo). They are the first *F5H* genes experimentally identified in monocotyledonous plants. Phylogenetic analysis indicated that both genes are closely related to dicot *F5H* genes from *Arabidopsis thaliana* and *Brassica napus*. Southern blot analysis showed that these genes exist as single copies in the rice genome. Alignments of the *OsF5HL* and *OsF5HL2* cDNAs to their genomic DNAs revealed that *OsF5HL* has an open reading frame (ORF) of 1590 b from four exons, while *OsF5HL2* has an ORF of 1560 b from two exons. Expression of *OsF5HL* is highest in young leaves, whereas that of *OsF5HL2* is greatest in mature leaves. In the roots and stems, transcription levels for both genes are markedly low. These data suggest that the *OsF5HL* and *OsF5HL2* genes belong to the *CYP84* subfamily and that their expressions are tissue-specific.

Keywords: F5H, lignin, *OsF5HL*, *OsF5HL2*, rice

Plant cytochrome P450-dependent monooxygenases (Cyt P450s) participate in many biochemical pathways, such as for phenylpropanoids, alkaloids, terpenoids, lipids, cyanogenic glycosides, glucosinolates, and brassinosteroids (Gravot et al., 1993; Frank et al., 1996; Urban et al., 1997; Ro et al., 2001). Among these, the phenylpropanoid pathway is one of the most critical processes because its general products are important to plant survival. Those products include UV-absorptive plant secondary metabolites, e.g., flavonoids, hydroxycinnamic acid esters, and lignin. Several Cyt P450 genes involved in the phenylpropanoid pathway have now been isolated and characterized: *phenylalanine ammonia lyase* (*PAL*), *cinnamate-4-hydroxylase* (*C4H*), and *4-coumarate CoA ligase* (*4CL*) (Hahlbrock and Scheel, 1989; Gravot et al., 1993; Mizutani et al., 1997; Chapple, 1998; Yang et al., 2005). Unfortunately, *ferulate-5-hydroxylase* (*F5H*) is a more difficult target for identification because its protein is relatively unstable, of low abundance, and membrane-bound (Chapple, 1998; Meyer et al., 1998). Despite these difficulties, *F5H* cDNAs and genomic clones have been reported from many dicot species, e.g., *Arabidopsis thaliana* (AF068574), *Brassica napus* (AF214008), *Broussonetia papyrifera* (AY850934), *Camptotheca acuminata* (AY621153), *Lycopersicon esculentum* x *L. peruvianum* (AF150881), and *Populus balsamifera* (AJ010324). F5H has been extensively studied because it gives rise to a wide array of important metabolites, including guaiacyl lignin, sinapic esters (sinapoylglucose, sinapoylcholine, and sinapoylmalate), and a host of other secondary metabolites. This enzyme catalyzes an irreversible hydroxylation step in the phenylpropanoid pathway that diverts ferulic acid away from guaiacyl lignin biosynthesis and toward sinapic acid and syringyl lignin (Chapple, 1998; Meyer et al., 1998) (Fig. 1). In angiosperms, sinapic acid is an intermediate in syringyl lignin biosynthesis and, in some taxa, it serves as a precursor for soluble secondary metabolites. *F5H* defines the *CYP84* subfamily in the Cyt P450 superfamily (Nelson et al., 1993; Chapple, 1998; Ruegger et al., 1999). Although genetic information for dicot *F5H* is available from the GenBank database, no reports have been made of this gene in monocotyledonous plants.

Rice was used in the current study because it is not only an important food crop but also a scientific model plant for monocots due to its small genome size, high synteny to other monocots, and efficiency in its transformation (Ahn and Tanksley, 1993; Hiei et al., 1994; Havukkala, 1996; Izawa and Shimamoto, 1996). Here, we describe the isolation and characterization of two cDNA clones for rice F5H, as well as their expression analysis in different tissues.

MATERIALS AND METHODS

Plant Material

Rice (*Oryza sativa* L. cv. Ilpoombyeo) plants were

*Corresponding author; fax +82-63-570-3339
e-mail bychung@kaeri.re.kr

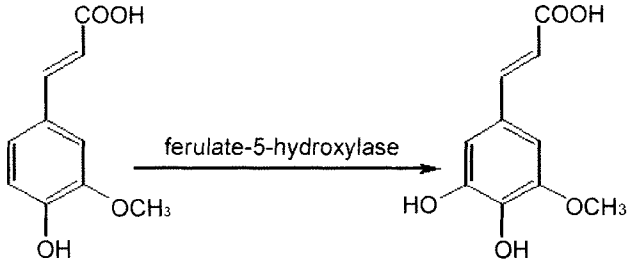

Figure 1. The conversion of ferulic acid to 5-hydroxyferulic acid is catalyzed by ferulate-5-hydroxylase.

hydroponically cultivated in half-strength Murashige and Skoog (MS) nutrient solution. The growth chamber was maintained at 28/20°C (D/N), under a 14-h photoperiod provided by two natrium lamps and six fluorescence lamps, with a photosynthetic photon flux density (PPFD) at pot level of 330 μmol m^{-2} s^{-1}. To extract total RNA and genomic DNA, three-week-old seedlings were harvested and excised into root, stem, and leaf segments. They were immediately frozen in liquid nitrogen and stored at -70°C.

Cloning of Rice *F5H* cDNA

Total RNA from the rice stems was extracted with TRIzol reagent (Invitrogen, USA), according to the manufacturer's instructions. Single-strand cDNA was synthesized from 5 μg of total RNA, using the RevertAid™ H Minus First Strand cDNA Synthesis Kit (MBI Fermentas, USA). The synthesized cDNAs were amplified with *F5H*-specific primers -- forward (5'-CTTCTCCCACTACGGCCACT-3') and reverse (5'-AGCTCGTACATCCCGAGCAC-3') – that had been designed by comparing the sequences of other *F5H* genes, e.g., *A. thaliana* (AF068574), *C. acuminata* (AY621153), and *P. balsamifera* (AJ010324). The amplified fragments were then cloned into the pGEM-T-easy vector (Promega, USA) and sequenced. To determine the full-length *F5H* cDNA sequences, 5'/3' rapid amplification of cDNA ends (5'/3' RACE) was performed using a 5'/3' RACE kit (Roche Diagnostics, USA), and the partial cDNAs were isolated and sequenced. Finally, the full-length cDNA sequences of our two rice *F5H* genes, *OsF5HL* (AB207252) and *OsF5HL2* (AB207253), were submitted to the DNA Data Bank of Japan (DDBJ, http://www.ddbj.nig.ac.jp)

Sequence and Phylogenetic Analyses

Sequence analyses were performed online (http://www.ncbi.nlm.nih.gov, http://www.shigen.nig.ac.jp, http://www.gramene.org, and http://www.softberry.com). The amino acid sequences were aligned by ClustalX (Thompson et al., 1997) and phylogenetic trees were generated by the neighbor-joining method, with 1000 repeats, using MEGA2 (Kumar et al., 2001). All neighbor-joining trees were drawn by NJplot (Perrière and Gouy, 1996). The GenBank accession numbers for the genes used in this sequence comparison are: *F5H* (AF068574, *A. thaliana*), *BNF5H2* (AF214008, *B. napus*), *F5H* (AY850934, *B. papyrifera*), *F5H* (AY621153, *C. acuminata*), *CYP84* (AF150881, *L. esculentum* x *L. peruvianum*), and *F5H* (AJ010324, *P. balsamifera*).

Southern Blot Analysis

Genomic DNA was extracted from young rice leaves using the DNeasy Plant Mini kit (Qiagen, USA). Aliquots (10 μg) of genomic DNA were digested with *Eco*RV, *Bam*HI, and *Hin*dIII, and then fractionated on a 0.8% (w/v) agarose gel. The gel was blotted onto a nylon membrane (Hybond N^+; Amersham, UK) and the membrane was cross-linked and hybridized with a ^{32}P-labeled cDNA fragment that included either the *OsF5HL* or the *OsF5HL2* open reading frame (0.6 kb) as a probe. Hybridization was carried out overnight at 65°C in a modified Church buffer (Church and Gilbert, 1984) containing 7% (w/v) sodium dodecyl sulfate (SDS), 0.5 M EDTA, 0.5 M sodium phosphate, and 1% (w/v) bovine serum albumin. The blots were washed twice with 2X SSC, 0.1% SDS for 20 min at 65°C; then twice with 1X SSC, 0.1% SDS for 10 min at 65°C, before being exposed for 1 d using the PERSONAL MOLECULAR IMAGER FX System (Bio-Rad, USA).

Expression Analysis

To investigate the transcription levels of our two rice *F5H* genes in different tissues, total RNA was extracted as described above from the roots, stems, and leaves for Reverse Transcription (RT)-PCR analysis. PCR was carried out as follows: denaturation at 94°C for 5 min, 26 cycles of 95°C (30 s), 58.2°C (30 s) and 72°C (30 s), and an extension at 72°C for 7 min. *OsF5HL*-specific primers forward (5'-AAAATTACTATGGCCTGGTGAAGTC-3') and reverse (5'-GACGAAGTAGGCATCTTGACTGTAG-3'); as well as *OsF5HL2* forward (5'-CGACCTGGAGAAGCTCCCCTTC-3') and reverse (5'-CTCCACACGTTGACCACCACGCG-3'); were designed by comparing the sequences of the respective *OsF5HL* and *OsF5HL2* genes. As a control reaction, amplification of *OsAct1* (Yamanouchi et al., 2002) was performed with forward (5'-TCCATCTTGGCATCTCTCAG-3') and reverse (5'-GTACCCTCATCAGGCATCTG-3') primers to ensure that equal amounts of cDNA were added to each PCR reaction.

RESULTS AND DISCUSSION

Isolation and Sequence Analysis of Rice *F5H* cDNA Clones

We used RACE (Frohman et al., 1988) and a cDNA pool made from rice stems to isolate the *OsF5HL* and

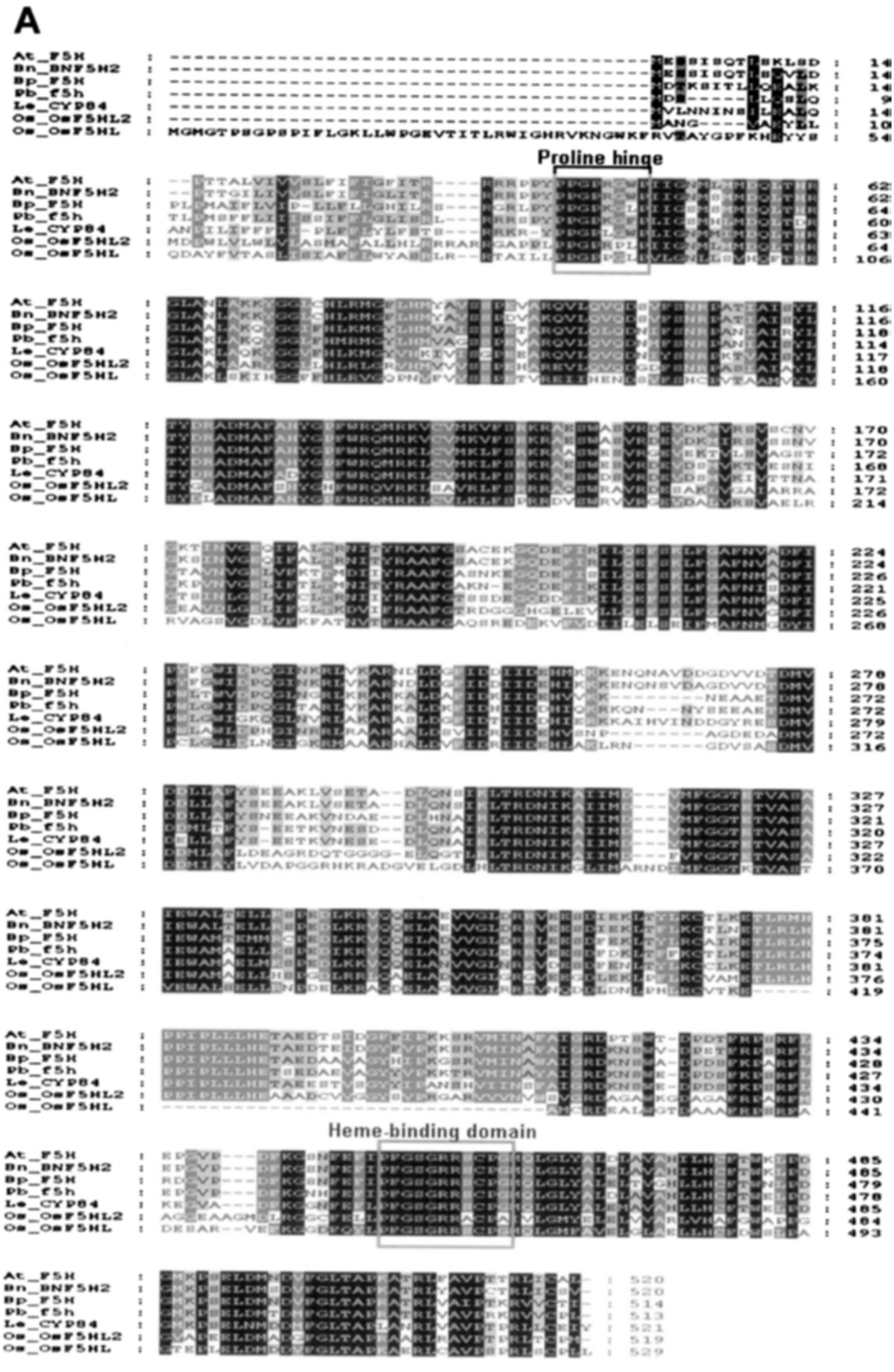

Figure 2. Sequence alignment and phylogenetic relationship of rice *F5H* genes. **A**, Comparison of deduced amino-acid sequences for *OsF5HL* and *OsF5HL2* with those previously reported for *At_F5H* from *A. thaliana*; *Bn_BNF5H2*, *B. napus*; *Bp_F5H*, *B. papyrifera*; *Ca_F5H*, *C. acuminata*; *Le_CYP84*, *L. esculentum* x *L. peruvianum*; and *Pb_f5h*, *P. balsamifera*. **B**, Phylogenetic relationship of *OsF5HL* and *OsF5HL2* from *CYP85* (*F5H*) subfamily. Neighbor-joining tree was generated by Clustal X. Numbers next to nodes are bootstrap values from 1000 replicates.

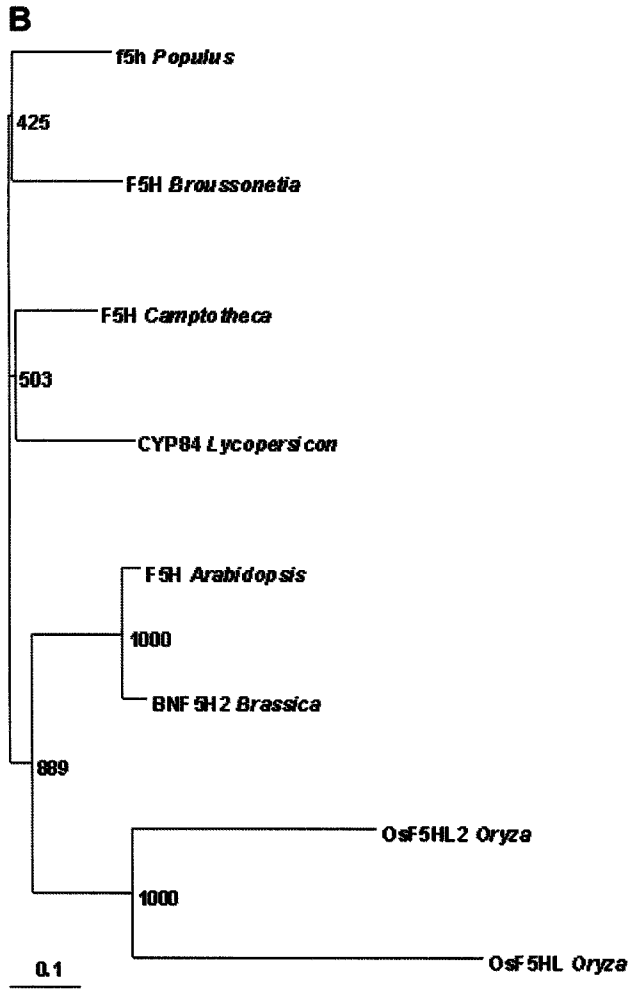

Figure 2. Continued.

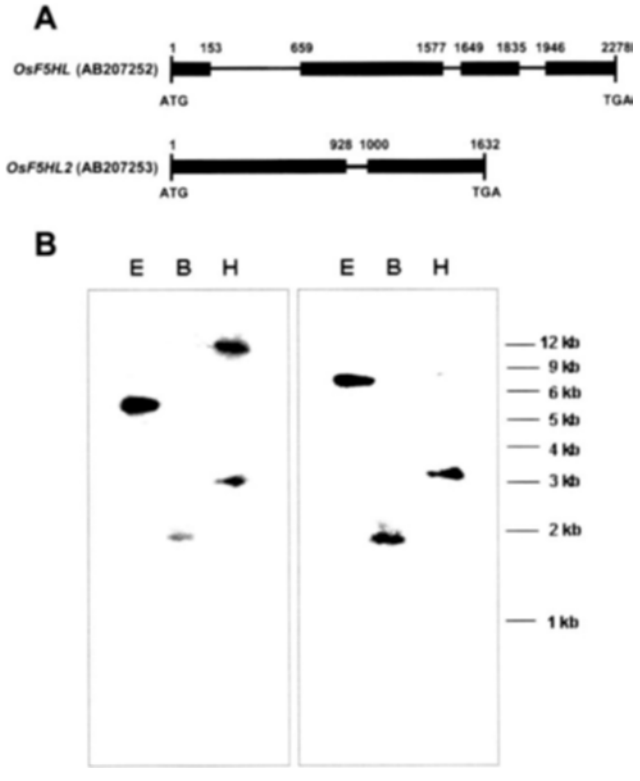

Figure 3. Schematic diagrams of rice *F5H* genes, and Southern blot analysis of *OsF5HL* and *OsF5HL2* loci in rice genome. **A**, Schematic diagrams of *OsF5HL* and *OsF5HL2* genes. Exons and introns are presented as solid bar and line, respectively. Parentheses indicate accession numbers. **B**, For Southern blot analysis, genomic DNA (10 μg per lane) was digested with enzymes and followed by hybridization, with about 0.6-kb cDNA fragment of respective gene as probe. E, *Eco*RV; B, *Bam*HI; H, *Hin*dIII.

OsF5HL2 genes. Sequence analyses showed open reading frames of 1590 b (*OsF5HL*) and 1560 b (*OsF5HL2*), encoding putative proteins of 530 and 520 amino acids, respectively. BLAST searches revealed that our rice *OsF5HL* shares 49% identity with *F5H* from *A. thaliana*, while the sequence homology is 53% between *OsF5HL2* and *F5H* of *B. papyrifera*. The predicted amino acid sequences of both genes showed the primary characteristics of Cyt P450 gene family members, i.e., a heme-binding region (FxxGxxxCxG) near the C-terminus and a consensus sequence (PPGPx(G/P)xP) of the proline hinge in the N-terminal region (Fig. 2A). Subfamilies in that superfamily are grouped by their amino acid sequence identity; within a Cyt P450 subfamily, homology should be greater than 40%. Only when a novel gene has 40% or less identity can a new gene family be defined (Frank et al., 1996; Chapple, 1998). Therefore, we can suggest that our isolated cDNA clones, *OsF5HL* and *OsF5HL2*, represent an *OsF5H*-like gene from *O. sativa* L., and that they belong to the *F5H* gene family.

To investigate the evolutionary relationships among various *F5H* (*CYP84*) genes, we conducted a phylogenetic analysis using genes in the *CYP84* subfamily (Nelson et al., 1993). Our two *OsF5H*-like genes from a monocot plant fell into that subfamily, showing high affinity to *F5H* genes from two dicot species, *A. thaliana* and *B. napus* (Fig. 2B)

The genomic sequences of our rice genes were estimated based on data from various websites: http://www.gramene.org, http://www.ncbi.nlm.nih.gov, http://www.shigen.nig.ac.jp, and http://www.softberry.com. *OsF5HL* and *OsF5HL2* were located to Chromosomes 6 and 3, respectively, in the rice genome, where their coding sequences are interrupted by either three introns – at 506 b, 72 b, and 112 b (Chromosome 6) – or one intron -- 74 b (Chromosome 3) (Fig. 3A). Southern blots confirmed that both genes exist as single copies in the genome (Fig. 3B).

OsF5HL and *OsF5HL2* Expression in Rice

Expression patterns were investigated using total RNA isolated from roots, young and mature leaves, and stems (Fig. 4). RT-PCR analysis revealed that *OsF5HL* and *OsF5HL2* transcript levels are highest in the young leaves and the mature leaves, respectively. However, in

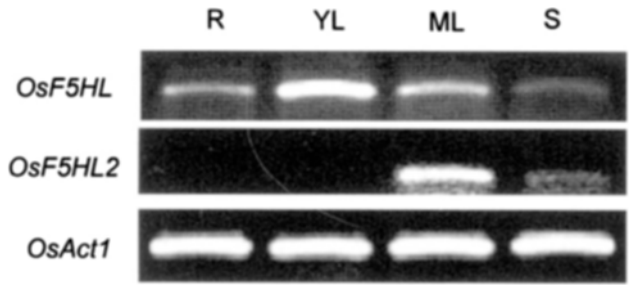

Figure 4. Expression profiles of *OsF5HL* and *OsF5HL2*. RT-PCR was performed with total RNA prepared from different tissues: S, stems; R, roots; YL, young leaves; ML, mature leaves. All tissues, except mature leaves (from 14-week-old plants) were excised from 3-week-old seedlings. Amplification of *OsAct1* was performed as control to ensure that equal amounts of cDNA were added to each PCR reaction.

the roots and stems, transcription of both genes is markedly low. This may mean that their expressions are tissue-specific.

Here, we have demonstrated the isolation and sequencing of two *F5H* genes from rice. Although we can speculate that their expressions vary according to the organ in which they are induced, we must still use enzyme activity assays of their products to confirm that *OsF5HL* and *OsF5HL2* are members of the *CYP84* subfamily. A detailed expression analysis of their functioning also is necessary.

ACKNOWLEDGEMENT

This project was supported by the Nuclear R&D Program of the Ministry of Science and Technology, Korea.

Received January 2, 2006; accepted April 4, 2006.

LITERATURE CITED

Ahn S, Tanksley SD (1993) Comparative linkage maps of the rice and maize genomes. Proc Natl Acad Sci USA 90: 7980-7984

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

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81: 1991-1995

Frank MR, Deyneka JM, Schuler MA (1996) Cloning of wound-induced cytochrome P450 monooxygenases expressed in pea. Plant Physiol 110: 1035-1046

Frohman MA, Dush MK, Martin GR (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene specific oligonucleotide primer. Proc Natl Acad Sci USA 85: 8998-9002

Gravot A, Larbat R, Hehn A, Lièvre K, Gontier E, Goergen JL, Bourgaud F (1993) Cinnamic acid 4-hydroxylase mechanism-based inactivation by psoralen derivatives: Cloning and characterization of a *C4H* from a psoralen producing plant - *Ruta graveolens* - exhibiting low sensitivity to psoralen inactivation. Arch Biochem Biophys 305: 509-515

Hahlbrock K, Scheel D (1989) Physiology and molecular biology of phenylpropanoid metabolism. Annu Rev Plant Physiol Plant Mol Biol 40: 347-369

Havukkala IJ (1996) Cereal genome analysis using rice as a model. Curr Opin Genet Dev 6: 711-714

Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. Plant J 6: 271-282

Izawa T, Shimamoto K (1996) Becoming a model plant: The importance of rice to plant science. Trends Plant Sci 1: 95-99

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software. Bioinformatics 17: 1244-1245

Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci USA 95: 6619-6623

Mizutani M, Ohta D, Sato R (1997) Isolation of a cDNA and a genomic clone encoding *cinnamate-4-hydroxylase* from *Arabidopsis* and its expression manner in planta. Plant Physiol 113: 755-763

Nelson DR, Kamataki T, Waxman DJ, Guengerich FP, Estabrook RW, Feyereisen R, Gonzalez FJ, Coon MJ, Gunsalus IC, Gotoh O, Okuda K, Nebert DW (1993) The P450 superfamily: Update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol 12: 1-51

Perrière G, Gouy M (1996) www-Query: An on-line retrieval system for biological sequence banks. Biochemie 78: 364-369

Ro DK, Mah N, Ellis BE, Douglas CJ (2001) Functional characterization and subcellular localization of poplar (*Populus trichocarpa* X *Populus deltoides*) cinnamate-4-hydroxylase. Plant Physiol 126: 317-329

Ruegger M, Meyer K, Cusumano JC, Chapple C (1999) Regulation of ferulate-5 hydroxylase expression in *Arabidopsis* in the context of sinapate ester biosynthesis. Plant Physiol 119: 101-110

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 24: 4876-4882

Urban P, Mignotte C, Kazmaier M, Delorme F, Pompon D (1997) Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 *CYP73A5*. J Biol Chem 272: 19176-19186

Yamanouchi U, Yano M, Lin H, Ashikari M, Yamada K (2002) A rice spotted leaf gene, *Spl7*, encodes a heat stress transcription factor protein. Proc Natl Acad Sci USA 99: 7530-7535

Yang DH, Chung BY, Kim JS, Kim JH, Yun PY, Lee YK, Lim YP, Lee MC (2005) cDNA Cloning and sequence analysis of rice cinnamate-4-hydroxylase gene, a cytochrome p450-dependent monooxygenase, involving in the general phenylpropanoid pathway. J Plant Biol 48: 311-318