Holzforschung, Vol. 63, pp. 69-74, 2009 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/HF.2009.019

Cloning and characterization of α -pinene synthase from *Chamaecyparis formosensis* Matsum

Fang-Hua Chu¹, Pei-Min Kuo², Yu-Rong Chen¹ and Sheng-Yang Wang^{2,3,*}

- ¹ School of Forestry and Resource Conservation, National Taiwan University, Taipei, Taiwan
- ² Department of Forestry, National Chung-Hsing University, Taichung, Taiwan
- ³ Experimental Forest, National Chung-Hsing University, Taichung, Taiwan

*Corresponding author.

- Department of Forestry, National Chung-Hsing University, Taichung 402, Taiwan
- Phone: +886-4-22840345 ext 138

E-mail: taiwanfir@dragon.nchu.edu.tw

Abstract

Analyzing the gene sequences of terpene synthase (TPS) may contribute to a better understanding of terpenes biosynthesis and evolution of phylogenetic taxonomy. Chamaecyparis formosensis is an endemic and precious conifer of Taiwan. To understand the biosynthesis mechanism of terpenes in this tree, a full length of putative mono-TPS, named as Cf-Pin (GeneBank accession no. EU099434), was obtained by PCR method and RACE extension. The Cf-Pin has an 1887-bp open reading frame and encodes 628 amino acids. To identify the function of Cf-Pin, the recombinant protein from Escherichia coli was incubated with geranyl diphosphate, produced one major product, the structure of which was elucidated. GC/MS analysis and matching of retention time and mass spectrum with authentic standards revealed that this product is α -pinene. This is the first report of cloning of a mono-TPS and functionally expressed in E. coli and which could be identified as α -pinene synthase from a Cupressaceae conifer.

Keywords: α -pinene; *Chamaecyparis formosensis*; Cupressaceae; terpene synthase.

Introduction

Terpenes are important secondary metabolites, which are not only essential for plant growth and development, but also for various interactions of plants with environment (Trapp and Croteau 2001; Christianson 2006; Keeling and Bohlmann 2006). Many terpenes are commercially applied in foods and cosmetics because of their flavor and fragrance. In addition, some of the terpenes are raw materials for pharmaceuticals (Aharoni et al. 2005). Terpenes are synthesized by terpene synthase (TPS) from one of the three common prenyl diphosphate precursors in plant (Chappell 1995). Monoterpenes originate from geranyl diphosphate (GPP), sesquiterpenes are from farnesyl diphosphate (FPP), and diterpenes are from geranylgeranyl diphosphate (GGPP). Theses precursors are formed by the condensation of dimethylallyl diphosphate with one or more isopentenyl diphosphate (IPP), catalyzed by specific prenyltranferases (Keeling and Bohlmann 2006). A large number of various terpenes are generated in plants as a result of the combined action of monterpene synthase (mono-TPS), sesquiterpene synthase (sesqui-TPS), or diterpene synthase (di-TPS). TPSs belong to the first key enzymes leading to the diversity of terpenes.

Several mono-TPSs from conifers have been functionally characterized (Bohlmann et al. 1997, 1999; Faldt et al. 2003; Byun-McKay et al. 2003, 2006; Phillips et al. 2003; Martin et al. 2004; Huber et al. 2005). Most of them are from Pinaceae. Only one sequence of limonene/ borneol synthase (BAC92722) was found from Cupressaceae (Chamaecyparis obtusa), but it was not functionally approved to date. Chamaecyparis formosensis and C. obtusa are named hinoki wood in Taiwan and Japan. Both are precious and highly estimated for their excellent quality, fragrance, and durability (Wang et al. 2005). C. formosensis commonly known as the Taiwan red cypress, which is an endemic tree that grows at elevations of 1500 to 2150 m in Taiwan's central mountains (Wang et al. 2005, 2006). Investigation of the chemical ingredients from C. formosensis can be traced back to Kafuka and Ichikawa (1931) who studied the volatile constituents of its leaves in 1931. According to their findings, α -pinene is the major constituent (85%) of its leaf oil.

In this study, a "mono-TPS like" cDNA was cloned from *C. formosensis.* The full-length clone was further functionally expressed in *Escherichia coli* and could be identified as α -pinene synthase. The results of cloning and characterization of α -pinene synthase from *C. formosensis* will be reported.

Materials and methods

Plant material and RNA isolation

The inner bark sample of a 25-year-old *C. formosensis* was obtained from the Chi-Tou Tract of the Experimental Forest of National Taiwan University and was identified by Dr. Yen-Hsueh Tseng (National Chung-Hsing University). The total RNA extraction was isolated as described by Chen et al. (2004) and Lee et al. (2006). Briefly, tissue of inner bark was frozen in liquid nitrogen and was ground into powder, which was then suspended in an extraction buffer [2% hexadecyltrimethylammonium bromide (CTAB), 2% polyvinylpyrrolidinone K30, 100 mM Tris-HCI (pH 8.0), 25 mM EDTA, 2.0 M NaCl, and 0.5 g spermidine I⁻¹] plus 2% (v/v) β -mercaptoethanol. The homogenate was extracted twice with an equal volume of chloroform:isoamyl alcohol (24:1). The total RNA was precipitated overnight with 10 M LiCl at 4°C, and then was dissolved in RNase-free water and stored at -80°C.

Isolation of partial and full-length TPS cDNA

To isolate mono-TPS cDNAs, partial cDNA fragments were obtained using degenerate primers, TPSm-F [5'-GA(T/C)CT(G/C/A)AACTCAACTGC(C/G/T)TTGGG-3'] and TPSm-R [5'-CC(A/G)AAGGTGTCGTA(G/C)ATATC(G/A)TC(G/C/A)AG-3'], designed from nucleotide alignments of conifer TPS sequences. The full-length cDNA of putative monoterpene synthase was isolated using 5'-RACE (rapid amplification of 5' complementary DNA ends) and 3'-RACE (rapid amplification of 3' complementary DNA ends), performed by SMART RACE cDNA Amplification Kit (BD Biosciences) using primer TPSm-5'RACE (5'-CTCCCCGGGAAATGCTATGAGTGAACC-3') and TPSm-3'RACE (5'-GCTCGCCATCGCCATGTGGGAGTATTAC-3') for PCR amplification. The PCR products were cloned by means of pGEM-T Easy vector system (Promega) and sequenced using ABI 377 automatic sequencer (Perkin Elmer).

Sequence alignment and phylogenetic analyses

Multiple sequence alignment was performed with the ClustalW program at EMBL (http://www.ebi.ac.uk/) and GeneDoc. Homologs of monoterpene synthase were identified by BLAST algorithms at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/). The following accession numbers were examined for monoterpene synthase proteins: myrcene synthase of *Abies grandis* (AAB71084), β -phellandrene synthase of *A. grandis* (AAF61453), pinene synthase of *A. grandis* (AAB71085), (+)-3-carene synthase of *Picea abies* (AAO73863), (-)-limonene synthase of *P. abies* (AAS47694), (-)-linalool synthase of *P. abies* (AAS47696), and myrcene synthase of *P. abies* (AAS47696). Phylogenetic analysis was performed by ClustalW and the phylogenetic tree was generated by neighbor-joining algorithm from the phylip3.6 software. The program TreeView (V. 1.6.6) was used to visualize the format of phylogenetic tree. Distance analyses were completed on 1000 bootstrap replicated data sets (SEQBOOT) using PRODIST. The sequences analyzed are listed in Table 1. Homology analysis (similarity and identity,%) presented in Table 1 was performed by means of the BL2SEQ program at NCBI.

Functional identification

The truncated open reading frame (ORF) of the monoterpene synthase cDNA was removed out of plastidial signal peptide, which was then subcloned into the fusion protein expression vector pQE31 (Qiagen) with optional N-terminal $6 \times \text{HisTag}$ sequence, by *SphI* and *SalI* cloning sites, and the resultant construct was expressed in *E. coli* BL-21 (DE3). Soluble his-tagged fusion proteins were purified using His-tagged resin (Novagen). The transformed *E. coli* cells were cultured overnight at 37° C in Luria-Bertani medium supplemented with 100 μ g μ I⁻¹ carbenicillin and induced with 1 mM isopropyI- β -D-thiogalactoside

Table 1 Similarity and identity analysis of Cf-Pin with other known conifer terpene synthases.

Enzyme	Species	Accession no.	Similarity (identity) (%)
Monoterpene synthase			
(-)-Camphene synthase	Abies grandis	AAB70707	72(54)
(-)-(4S)-Limonene synthase	Abies grandis	AAB70907	71(53)
(-)-Limonene/(-)- α -pinene synthase	Abies grandis	AAF61455	71(53)
Myrcene synthase	Abies grandis	AAB71084	72(54)
β-Phellandrene synthase	Abies grandis	AAF61453	73(54)
(-)-Pinene synthase	Abies grandis	AAB71085	70(52)
Terpinolene synthase	Abies grandis	AAF61454	69(50)
Limonene/borneol synthase	Chamaecyparis obtusa	BAC92722	71(54)
(+)-3-Carene synthase	Picea abies	AAO73863	69(52)
(-)-Limonene synthase	Picea abies	AAS47694	75(55)
(-)-Linalool synthase	Picea abies	AAS47693	74(55)
Myrcene synthase	Picea abies	AAS47696	74(54)
(-)- α/β -Pinene synthase	Picea abies	AAS47692	74(53)
(-)-Pinene synthase	Picea sitchensis	AAP72020	70(50)
(-)-α-Pinene synthase	Pinus taeda	AAO61225	71(53)
(+)-α-Pinene synthase	Pinus taeda	AAO61228	72(52)
α -Terpineol synthase	Pinus taeda	AAO61227	70(50)
(-)-α-Pinene/(-)-camphene synthase	Pseudotsuga menziesii	AAX07267	72(53)
Terpinolene synthase	Pseudotsuga menziesii	AAX07264	70(51)
Sesquiterpene synthase			
(E) - α -Bisabolene synthase	Abies grandis	AAC24192	66(45)
(E) - α -Bisabolene synthase	Picea abies	AAS47689	66(45)
(\vec{E}) - γ -Bisabolene synthase	Pseudotsuga menziesii	AAX07266	65(45)
(E)-β-Farnesene synthase	Pseudotsuga menziesii	AAX07265	63(43)
Diterpene synthase			
Abietadiene synthase	Abies grandis	AAB05407	61(43)
Levopimaradiene synthase	Ginkgo biloba	AAL09965	61(41)
Isopimara-7,15-diene synthase	Picea abies	AAS47690	61(41)
Levopimaradiene/abietadiene synthase	Picea abies	AAS47691	61(41)
Levopimaradiene/abietadiene synthase	Pinus taeda	AAX07435	60(42)
Taxadiene synthase	Taxus baccata	CAC42773	66(46)
Taxadiene synthase	Taxus brevifolia	AAC49310	66(46)
Taxa-4(5),(11912)-diene synthase	Taxus canadensis	AAR13860	66(46)
Taxadiene synthase	Taxus chinensis	AAG02257	65(45)
Taxadiene synthase	Taxus chinensis var. mairei	AAR15329	66(46)
Taxadiene synthase	Taxus $ imes$ media	AAS18603	66(46)

 α -Pinene synthase from Chamaecyparis formosensis 71

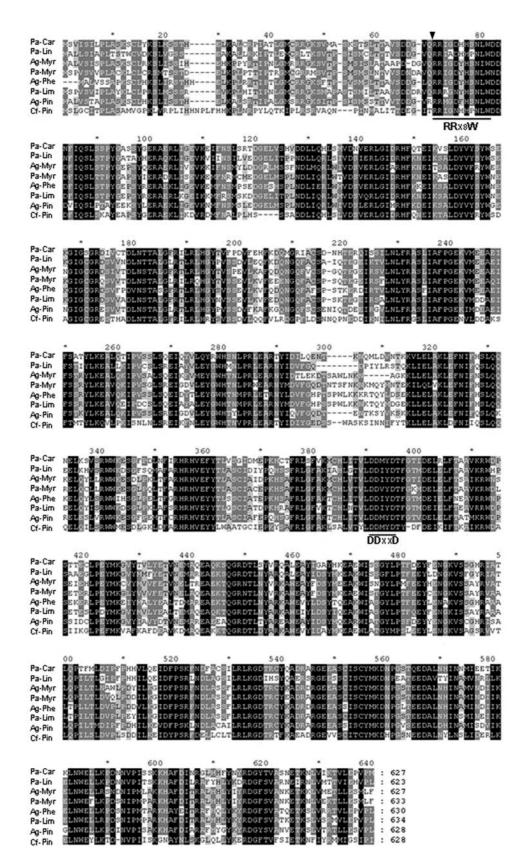


Figure 1 Amino acid alignments of Cf-Pin from *Chamaecyparis formosensis* with other reported coniferous monoterpene synthase by ClusterlX and GeneDoc. The RRX₈W and DDXXD are shown. Black triangle indicates the separate barrier with plastidial signal peptide. Conserved similarity shading is based on 100% (black), 70% (dark gray), and 60% (light gray). Pa-Car: *Picea abies* (+)-3-carene synthase (AAO73863); Pa-Lim: *Picea abies* (-)-limonene synthase (AAS47694); Pa-Lin: *Picea abies* (-)-linalool synthase (AAS47693); Pa-Myr: *Picea abies* myrcene synthase (AAS47696); Ag-Myr: *Abies grandis* myrcene synthase (AAB71084); Ag-Phe: *Abies grandis* β-phellandrene synthase (AAF61453); Ag-Pin: *Abies grandis* pinene synthase (AAB71085).

(IPTG) at 20°C overnight. The activity of TPS was measured in vitro by GC/MS analysis as described by Lewinsohn et al. (1991) and Martin et al. (2004) with slight modifications. Briefly, the purified protein (30 μ g) were added in mono-TPS buffer (25 mM HEPES, pH 7.2, 100 mM KCl, 10 mM MnCl₂, 10% (v/v) glycerol, 5 mM DTT) with 30 µM GPP, 1 mM ascorbic acid, and protease inhibitor. The reaction mixture was then covered with 1 ml of pentane to trap volatile products and incubated at 30°C for 3 h. The pentane extracts were combined and passed through a silica gel column (3 cm \times 5 mm i.d.) covered with MgSO₄ (1 cm). Monoterpene products were analyzed by a Trace GC PoLaris Q mass system, equipped with an RTx-5MS capillary column (30 m \times 0.25 mm i.d., film thickness 0.25 μ m). The oven temperature was held at 80°C for 1 min, then programmed to increase from 80°C to 200°C at a rate of 4°C min-1 and held for 5 min. The other parameters: injector temperature 250°C; ion source temperature 280°C; El 70 eV; carrier gas He at a flow rate of 1 ml min⁻¹; split ratio 1:50; mass range 45-425 m/z. The structure of the major compound synthesized by TPS of C. formosensis was confirmed by comparison with a standard, by spiking, and on the basis of its mass spectral fragmentation based on the Wiley/NBS Registry of Mass Spectral library and NIST MS Search.

Results and discussion

cDNA cloning

By combination of PCR method and RACE extension, the full length cDNA of putative mono-TPS, named as *Cf-Pin*, was obtained (GeneBank accession no. EU099434). The latter has an 1887-bp ORF, a 53-bp 5' non-coding region, and a 160-bp 3' non-coding region. The ORF encodes a 628-aa protein and the predicted molecular weight (M_w) of the polypeptides was 72.3 kDa and the theoretical isoelectric point (p/) was 5.76. The D³⁸¹D³⁸²XXD³⁸⁵ motif conserved in terpene synthase (Keeling and Bohlmann 2006) was present in the Cf-Pin (Figure 1). As in other mono-TPS, a transit peptide sequence upstream from R⁶⁷R⁶⁸X₈W⁷⁷ was found that may function in targeting the protein into plastids (Keeling and Bohlmann 2006) (Figure 1).

Homologous analysis

The TPS plant family has been classified into seven subfamilies, i.e., TPS-a through to TPS-d, based on sequence relatedness as well as functional assessment (Martin et al. 2004). Up to date, all known gymnosperm TPS clustered in the TPS-d subfamily which contains mono-TPS (TPS-d1), sesqui-TPS (TPS-d2), di-TPS (TPSd3) (Keeling and Bohlmann 2006). To the best of our knowledge, there is not any full length TPS from Cupressaceae reported until now. In this study, the first TPS, Cf-Pin, was obtained from a tree belonging to the Cupressaceae. Thus, the further understanding of the phylogeny of the TPS-d family can be substantiated with the inclusion of Cf-Pin together with other TPS gymnosperms available from the literature.

Sequence analysis showed that the deduced protein of Cf-Pin was close to TPS-d1 (mono-TPS) from other gymnosperms TPS-d subfamily (Figure 2 and Table 1). The Cf-Pin showed 69% to 75% amino acid similarity and 50% to 55% identity with other known monoterpene

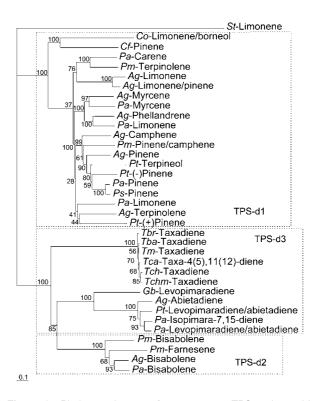


Figure 2 Phylogenetic tree of gymnosperm TPS amino acid sequences showing the TPS-d subfamily branching into three distinct groups. Limonene synthase from *Schizonepeta tenuifolia* shown as outgroup in this analysis. The neighbor-joining method was used for the protein sequences. Bootstrap values were labeled besides the branches. The scale bar corresponds to the branch length for an expected number of 0.1 substitutions per site. *Ag, Abies grandis; Co, Chamaecyparis obtusa; Gb, Ginkgo biloba; Pa, Picea abies; Ps, Picea sitchensis; Pt, Pinus taeda; Pm, Pseudotsuga menziesii; Tba, Taxus baccata; Tbr, Taxus brevifolia; Tca, Taxus canadensis; Tch, Taxus chinensis; Tchm, Taxus chinensis var. mairei; Tm, Taxus × media; St, Schizonepeta tenuifolia.*

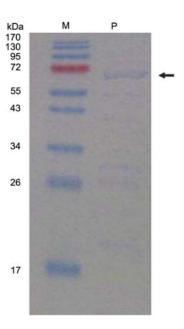


Figure 3 SDS/PAGE analysis of the proteins expressed in *E. coli* pQE31 system by IPTG induction. Lane M: protein molecular marker. Lane P: purified protein. The gel was shown in Coomassie blue staining.

α-Pinene synthase from Chamaecyparis formosensis 73

synthases. However, Cf-Pin has lower amino acid similarity with sesquiterpene synthase (63-68% similarity; 43-47% identity) and diterpene synthase (60-66% similarity; 41-46% identity) (Table 1). Phylogenetic analysis showed that Cf-Pin is closely clustered with other reported TPS-d1, although Pa-limonene synthase has the highest amino acid similarity to Cf-Pin (75%) and not to Co-limonene/borneol synthase (71%) (Table 1). It has, however, distinct protein sequences from Pinaceae conifer TPS-d1 and it is more closely related to the sequence of limonene/borneol synthase from C. obtusa (Figure 2). The reason for the differences observed in analysis of similarity, identity, and phylogeny might be due to TPS function variation in the TPS-d1 subfamily. It is also possible that the TPS is specified in gymnosperms (Keeling and Bohlmann 2006). The results not only account for the diversity of the pinene synthases and the evolutionary variety of conifers. The complete elucidation of the sequence of Cf-Pin might also reinforce the TPS investigation in conifers. Moreover, a functional diversification and the evolutionary context of plant TPS may be elucidated.

Functional identification

To determine the function of TPS gene obtained from C. formosensis, full-length and truncated recombinant proteins were tested for activity with the substrates GPP, FPP, and GGPP. However, only monterpene products were formed by the truncated recombinant protein. The recombinant plasmid pQE31/Cf-Pin, which transit peptide sequence, had been truncated and resulted in successful expression of fusion protein in E. coli after IPTG induction. The recombinant protein was purified by means of His-tagged resin. The molecular mass was approximately 63 kDa (Figure 3). GPP is considered to be the natural substrate for mono-TPS (Bohlmann et al. 1998). However, GPP cannot be cyclized directly due to the transdouble bond located at C2-C3; the reaction mechanism necessarily involves both isomerization and cyclization steps (Bohlmann et al. 1998). To identify the function of Cf-Pin, the recombinant protein from E. coli was incubated with GPP, produced a major product (Figure 4) according to GC analysis. MS analyses and retention time matching with authentic standards

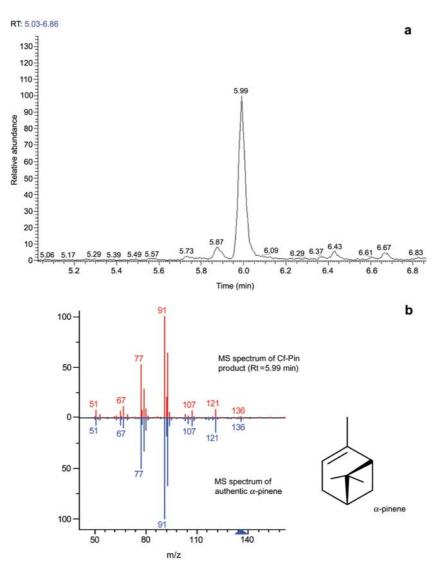


Figure 4 (a) GC/MS analysis of monoterpene products of recombinant protein product of Cf-Pin with geranyl diphosphate as the substrate. (b) MS of product are shown in the top of panel with the authentic α -pinene below. The major product of Cf-Pin is α -pinene.

revealed that the major product was α -pinene (RT: 5.99 min). According to the search from gene bank in NCBI (http://www.ncbi.nlm.nih.gov), Cf-Pin is the first mono-TPS identified from the Cupressaceae family, the function of which was approved. The cloning and identification of TPS cDNA may provide important information for understanding the molecular genetics and biosynthesis of terpenes in conifers. Further study of TPS in *C. formosensis*, including its physiological function and its interaction with the environment, is under investigation in our group.

Acknowledgements

This study was supported by a grant (96-2628-B-005-010-MY3) from the National Science Council (NSC) of Executive Yuan, Taiwan. The authors thank NSC for financial support. The authors also thank the reviewers and the Editorial Office of Holzforschung for their constructive and valuable comments.

References

- Aharoni, A., Jongsma, M.A., Bouwmeester, H.J. (2005) Volatile science? Metabolic engineering of terpenoids in plants. Trends Plant Sci. 10:594–602.
- Bohlmann, J., Steele, C.L., Croteau, R. (1997) Monoterpene synthases from grand fir (*Abies grandis*). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(*4S*)-limonene synthase, and (-)-(*1S*,*5S*)-pinene synthase. J. Biol. Chem. 272:21784–21792.
- Bohlmann, J., Meyer-Gauen, G., Croteau, R. (1998) Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. USA 95:4126–4133.
- Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S., Croteau, R. (1999) cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the *Tpsd* gene family from grand fir (*Abies grandis*). Arch. Biochem. Biophys. 368:232–243.
- Byun-McKay, A., Hunter, W., Goddard, K., Wang, S., Martin, D., Bohlmann, J., Plant, A. (2003) Insect attack and wounding induce traumatic resin duct development and gene expression of (α)-pinene synthase in Sitka spruce. Plant Physiol. 133:368–378.
- Byun-McKay, A., Godard, K.A., Toudefallah, M., Martin, D.M., Alfaro, R., King, J., Bohlmann, J., Plant, A.L. (2006) Woundinduced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. Plant Physiol. 140:1009–1021.

- Chappell, J. (1995) The biochemistry and molecular biology of isoprenoid metabolism. Plant Physiol. 107:1–6.
- Chen, Y.R., Lee, Y.R., Wang, S.Y., Chang, S.T., Shaw, J.F., Chu, F.H. (2004) Establishment of expressed sequence tags from Taiwania (*Taiwania cryptomerioides* Hayata) seedling cDNA. Plant Sci. 167:955–957.
- Christianson, D.W. (2006) Structural biology and chemistry of the terpenoid cyclases. Chem. Rev. 106:3412–3442.
- Faldt, J., Arimura, G., Gershenzon, J., Takabayashi, J., Bohlmann, J. (2003) Functional identification of AtTPS03 as (E)β-ocimene synthase: a monoterpene synthase catalyzing jasmonate- and wound-induced volatile formation in *Arabidopsis thaliana*. Planta 216:745–751.
- Huber, D.P., Philippe, R.N., Madilao, L.L., Sturrock, R.N., Bohlmann, J. (2005) Changes in anatomy and terpene chemistry in roots of Douglas-fir seedlings following treatment with methyl jasmonate. Tree Physiol. 25:1075–1083.
- Kafuka, K., Ichikawa, N. (1931) The volatile compounds from leaves of *Chamaecyparis formosensis*. Nippon Kagaku Kaishi 52:222–228.
- Keeling, C.I., Bohlmann, J. (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defense of conifers against insects and pathogens. New Phytologist 170:657–675.
- Lee, C.H., Chan, M.H., Wang, Y.N., Chu, F.H. (2006) Gene investigation into the inner bark of Taiwania (*Taiwania cryptomerioides*). Botan. Stud. 47:111–118.
- Lewinsohn, E., Gijzen, M., Savage, T.J., Croteau, R. (1991) Defense mechanisms of conifers – relationship of monoterpene cyclase activity to anatomical specialization and oleoresin monoterpene content. Plant Physiol. 96:38–43.
- Martin, D.M., Faldt, J., Bohlmann, J. (2004) Functional characterization of nine Norway spruce *TPS* genes and evolution of gymnosperm terpene synthases of the *TPS-d* subfamily. Plant Physiol. 135:1908–1927.
- Phillips, M.A., Wildung, M.R., Williams, D.C., Hyatt, D.C., Croteau, R. (2003) cDNA isolation, functional expression, and characterization of (+)-alpha-pinene synthase and (-)-alphapinene synthase from loblolly pine (*Pinus taeda*): stereocontrol in pinene biosynthesis. Arch. Biochem. Biophys. 411: 267–276.
- Trapp, S.C., Croteau, R.B. (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics 158:811–832.
- Wang, S.Y., Wu, C.L., Chu, F.H., Chien, S.C., Kuo, Y.H., Shyur, L.F., Chang, S.T. (2005) Chemical composition and antifungal activity of essential oil from *Chamaecyparis formosensis* wood. Holzforschung 59:295–299.
- Wang, S.Y., Wang, Y.S., Tseng, Y.H., Lin C.T., Liu, C.P. (2006) Analysis of fragrance compositions of precious coniferous. Holzforschung 60:528–532.

Received April 29, 2008. Accepted August 12, 2008. Previously published online September 15, 2008.