Biosci. Biotechnol. Biochem., **72** (4), 1038–1047, 2008

Identification of Diterpene Biosynthetic Gene Clusters and Functional Analysis of Labdane-Related Diterpene Cyclases in *Phomopsis amygdali*

Tomonobu TOYOMASU,[1,†] Rie NIIDA,[1,*] Hiromichi KENMOKU,[1,**] Yuri KANNO,[1,*] Shigeyoshi MIURA,[1] Chiaki NAKANO,[2] Yoshihito SHIONO,[1] Wataru MITSUHASHI,[1] Hiroaki TOSHIMA,[3] Hideaki OIKAWA,[4] Tsutomu HOSHINO,[2] Tohru DAIRI,[5] Nobuo KATO,[6] and Takeshi SASSA[1]

[1]*Department of Bioresource Engineering, Yamagata University, Tsuruoka, Yamagata 997-8555, Japan*
[2]*Graduate School of Science and Technology, Niigata University, Ikarashi, Niigata 950-2181, Japan*
[3]*Department of Bioresource Science, College of Agriculture, Ibaraki University, Inashiki, Ibaraki 300-0393, Japan*
[4]*Division of Chemistry, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan*
[5]*Biotechnology Research Center, Toyama Prefectural University, Imizu, Toyoma 939-0398, Japan*
[6]*The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567-0047, Japan*

Received December 5, 2007; Accepted December 26, 2007; Online Publication, April 7, 2008
[doi:10.1271/bbb.70790]

Two diterpene biosynthesis gene clusters in the fusicoccin-producing fungus, *Phomopsis amygdali*, were identified by genome walking from *PaGGS1* and *PaGGS4* which encode the geranylgeranyl diphosphate (GGDP) synthases. The diterpene cyclase-like genes, *PaDC1* and *PaDC2*, were respectively located proximal to *PaGGS1* and *PaGGS4*. The amino acid sequences of these two enzymes were similar to those of fungal labdane-related diterpene cyclases. Recombinant PaDC1 converted GGDP mainly into phyllocladan-16α-ol *via* (+)-copalyl diphosphate (CDP) and trace amounts of several labdane-related hydrocarbons which had been identified from the *P. amygdali* F6 mycelia. Since phyllocladan-16α-ol had not been identified in *P. amygdali* F6 mycelia, we isolated phyllocladan-16α-ol from the mycelia. Recombinant PaDC2 converted GGDP into (+)-CDP. Furthermore, we isolated the novel diterpenoid, phyllocladan-11α,16α,18-triol, which is a possible metabolite of phyllocladan-16α-ol in the mycelia. We propose that genome walking offers a useful strategy for the discovery of novel natural products in fungi.

Key words: biosynthesis; cyclase; diterpene; gene cluster; *Phomopsis amygdali*

The cyclic diterpenoids are derived from geranylgeranyl diphosphate (GGDP) through cyclization and chemical modification. Fungi produce a variety of diterpenoids with unique biological activities, *e.g.*, gibberellin is a phytohormone,[1)] aphidicolin is a specific inhibitor of DNA polymerase α,[2)] and fusicoccin A has not only potent plant plasma membrane H^+-ATPase-activating activity[3)] but also exerts novel effects on amphibian embryogenesis.[4)] The biosynthetic genes of gibberellins and aphidicolin, including a gene that encodes GGDP synthase (GGS), have recently been found to be clustered in the fungal genome.[5,6)] This suggested that the diterpene biosynthetic gene cluster could be obtained by isolating the *GGS* genes and subsequent genome walking. We have previously isolated six cDNA fragments that were derived from *GGS*-like genes (*PaGGS1* to *6*) from *Phomopsis amygdali* N2, and carried out genome walking from *GGS* to identify a cDNA that encodes fusicocca-2,10(14)-diene synthase and the fusicoccin biosynthetic gene cluster.[7)] As a result, we isolated a chimera diterpene synthase gene that encodes fusicocca-2,10(14)-diene synthase (PaFS), which possesses a terpene cyclase domain at the N-terminus and a prenyltransferase domain at the C-terminus.[7)] These results indicate that genome walking is a useful strategy for isolating cDNAs that encode unknown diterpene biosynthetic enzymes.

We describe in the present report the identification by

† To whom correspondence should be addressed. Fax: +81-235-28-2812; E-mail: toyomasu@tds1.tr.yamagata-u.ac.jp
* Present address: *Plant Science Center, RIKEN, Yokohama, Kanagawa 230-0045, Japan*
** Present address: *Tokushima Bunri University, Tokushima, Tokushima 770-8514, Japan*

Abbreviations: CDP, copalyl diphosphate; DMAPP, dimethylallyl diphosphate; FDP, farnesyl diphosphate; GC-MS, gas chromatography-mass spectrometry; GDP, geranyl diphosphate; GGDP, geranylgeranyl diphosphate; GGS, geranylgeranyl diphosphate synthase; GST, glutathione *S*-transferase; IPP, isopentenyl diphosphate; RACE, rapid amplification of cDNA ends; RT, reverse transcription; TLC, thin-layer chromatography

genome walking of two additional diterpene biosynthetic gene clusters, which include the labdane-related diterpene cyclase genes *PaDC1* and *PaDC2*, respectively located around *PaGGS1* and *PaGGS4*, in *P. amygdali*. In addition, we determined the structures of the two products converted from GGDP by recombinant PaDC1 and PaDC2, indicating that *PaDC1* and *PaDC2* encode phyllocladan-16α-ol synthase and (+)-copalyl diphosphate (CDP) synthase, respectively. Based on these evidences, we also identified from the *P. amygdali* mycelia the novel diterpenoid, phyllocladan-11α,16α,18-triol, which is a possible metabolite of phyllocladan-16α-ol. Thus, genome walking facilitated the discovery of two novel diterpene synthases and a novel diterpenoid.

Materials and Methods

Chemicals. Isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP), geranyl diphosphate (GDP), farnesyl diphosphate (FDP), and GGDP were purchased from Sigma-Aldrich. [1-^{14}C]IPP (CFA476, 2.15 GBq/mmol) was obtained from Amersham Bioscience. (+)-Copalol was synthesized by Toshima *et al.*,[8] and converted to the diphosphate ester, (+)-CDP, according to the method described previously.[9]

Fungal strains. Phomopsis amygdali strains N2 and F6 were respectively presented by Dr. Kanematsu (Apple Research Center, National Institute of Fruit Tree Science) and Professor A. Graniti (Bari University).

Genome walking from GGS. Genome walking was carried out by using the Universal Genome Walker kit (Clontech). Genomic DNA was extracted from the mycelia by using the Nucleon PhytoPure kit (Amersham Bioscience). The genome walker library was constructed with genomic DNA and was used for a nested polymerase chain reaction (PCR) according to the manufacturer's protocol. Primers for the first walking were designed from the cDNA fragment sequences of *PaGGS1* and *PaGGS4*.

Rapid amplification of cDNA ends (RACE). To determine the sequences of *PaGGS1* and *PaGGS4*, 5′-RACE and 3′-RACE were performed by using gene-specific primers based on the corresponding genomic DNA sequences according to the methods described previously.[10]

Functional analysis for prenyltransferase. To amplify the *PaGGS1* and *PaGGS4* ORF cDNAs, the following sets of primers were used: for *PaGGS1*, sense 5′-GGATCCATGACCTCCGGGTCAAAGC-3′ (*Bam*HI site underlined) and antisense 5′-CTGCAGATAAGAAATGCAGTACCAT-3′ (*Pst*I site underlined); and for *PaGGS4*, sense 5′-GGATCCATGGACTTCCCGATCCGCTC-3′ (*Bam*HI site underlined) and antisense 5′-CTGCAGATTGGTACTGCATTTCTTAT-3′ (*Pst*I site underlined). The PCR product was digested with *Bam*HI and *Pst*I and subcloned into the pQE30 vector (Qiagen), and the plasmid was introduced into *E. coli* BL21 (DE3). All the subsequent procedures have been described previously.[11] The allylic substrates used were DMAPP, GDP, and FDP. The products derived from the allylic substrates with [^{14}C]IPP were dephosphorylated and subjected to reversed-phase thin-layer chromatography (TLC), using LKC-18 and developing with acetone/H_2O (9:1).

Functional analysis for cyclase. To amplify the cDNAs for the ORFs of *PaDC1* and *PaDC2*, reverse transcription (RT)-PCR was carried out by using the following pairs of primers: for *PaDC1*, sense 5′-CCCGGGATGACTATTATGGACATCGA-3′ (*Sma*I site underlined) and antisense 5′-GCGGCCGCCTACGTGACCAGCATTGGAC-3′ (*Not*I site underlined); and for *PaDC2*, sense 5′-AGATCTATGGAGTTCGATGAACCACT-3′ (*Bgl*II site underlined) and antisense 5′-CCCGGGACACGACCCGACAGATCAGG-3′ (*Sma*I site underlined). The PCR products were digested with *Sma*I/*Not*I for *PaDC1* and *Bgl*II/*Sma*I for *PaDC2*, and ligated into the pGEX-4T-3 vector (Amersham Bioscience) to generate a glutathione *S*-transferase (GST) fusion protein. The plasmid was transformed into *E. coli* XL1-Blue. The procedures used for the growth of the *E. coli* cells, induction of gene expression, extraction and purification of recombinant enzymes, and measurement of cyclase enzyme activity have been described previously.[12] GDP, FDP and GGDP (20 μg) were used as the substrates. The hydrocarbon product and dephosphorylated derivative from the diphosphate product were analyzed as described previously.[12]

General procedures. The melting point (mp) data are uncorrected. Optical rotation values were measured with a Horiba SEPA-300 polarimeter, and the IR, UV and CD spectra were recorded with Jasco J-20A and Jasco J-20A spectrophotometers. Gas chromatography-mass spectrometry (GC-MS) was performed with an Agilent 6890N GC-5973N mass selective detector system. Mass spectra were obtained with a JMS-700 instrument, and ^{1}H- and ^{13}C-NMR spectra were acquired with a Jeol EX-400 or Brucker DPX 400 spectrometer.

Extraction and isolation of phyllocladan-16α-ol from the mycelia of P. amygdali *F6.* A hexane extract[13] (*ca.* 83.4 mg) from 95 g of *P. amygdali* F6 mycelia obtained from 10 culture flasks, each containing 100 ml of the culture medium (5.0% glucose, 1.0% Pharmamedia, 0.5% KH_2PO_4, and 0.1% $MgSO_4$ in deionized water), was separated by silica gel flash chromatography (Wakogel FC-40, 2 cm i.d. × 22 cm). From the fraction (7 mg) eluted with a mixture of hexane:EtOAc (10:1), a colorless oil (0.2 mg) with an Rf value of 0.48 was

obtained by preparative TLC on silica gel (Kieselgel 60 F254, Merck) by using a mixture of hexane:EtOAc (2:1); this compound was identified as phillocladan-16α-ol (see text).

Extraction and isolation of phyllocladan-11α, 16α, 18-triol from the mycelia of P. amygdali *F6.* For seed fermentation, *Phomopsis amygdali* F6 was grown in two 500-ml Sakaguchi flasks that contained 100 ml of a medium consisting of 3.0% glucose, 0.5% Pharmamedia, and 0.1% peptone. The inoculated flasks were incubated at 25 °C for 6 d on a rotary shaker. The culture broth from the two flasks was inoculated into 4500 ml of a medium containing 4% sugar, 0.5% soybean meal, 0.2% Phrmamedia, 0.5% KH_2PO_4, and 0.1% $MgSO_4 7H_2O$ in a jar fermenter (TS-A-7L; Takasaki Co., Tokyo, Japan) and cultured for 14 d. The culture broth was fermented with agitation at 25 °C, 180 rpm, and aeration at 2500 ml/min. After the incubation period, the culture supernatant was purified by filtration. The filtrate was basified with Na_2CO_3 and extracted with EtOAc. The organic layer was concentrated *in vacuo* to give an oily residue (8.8 g) which was subjected to silica gel column chromatography using stepwise elution with mixtures of $CHCl_3$–EtOH. The $CHCl_3$–EtOH (9:1) eluate (63.4 mg) was crystallized from MeOH to afford colorless needles (14.4 mg) which were identified as phyllocladan-11α,16α,18-triol (see the text).

Phyllocladan-11α,16α,18-triol: mp 295–300 °C; $[\alpha]_D^{20} = +3.5^\circ$ (*c* 0.1, MeOH); IR (KBr) ν_{max} cm^{-1}: 3292, 1382, 1357, 1057, 1018; HRFABMS *m/z* $[M + Na]^+$: calcd. for $C_{20}H_{34}O_3Na$, 345.2406; found, 345.2411; ^{1}H-NMR δ_H ($CDCl_3$): 0.89 (3H, s, Me-19), 1.29 (3H, s, Me-20), 1.25 (1H, dd, $J = 12.7$, 7.8 Hz, 6-H), 1.31 (1H, t, $J = 10.8$ Hz, H-14), 1.42 (overlapped signals, H-1 and H-3), 1.48 (1H, d, $J = 9.7$ Hz, H-9), 1.54 (1H, m, H-2), 1.57 (3H, s, Me-17), 1.64 (1H, m, 6-H), 1.70–1.74 (overlapped signals, H-5 and H-7), 1.72 (1H, d, $J = 14.7$ Hz, H-15), 1.78 (1H, m, H-2), 1.81 (1H, td, $J = 9.7$, 2.9 Hz, H-12), 1.93 (1H, td, $J = 13.7$, 3.9 Hz, H-3), 2.10 (1H, m, H-13), 2.32 (1H, d, $J =$ 14.7 Hz, H-15), 2.36 (1H, m, H-12), 2.46 (1H, ddd, $J = 10.8$, 6.8, 3.9 Hz, H-14), 3.08 (1H, br. d, $J =$ 13.7 Hz, H-1), 3.22 (1H, dd, $J = 10.8$, 5.4 Hz, H-18), 3.68 (1H, dd, $J = 10.8$, 5.4 Hz, H-18), 4.27 (1H, m, H-11), 5.30 (1H, d, $J = 7.3$ Hz, OH-11), 5.43 (1H, s, OH-16), 5.96 (1H, t, $J = 5.4$ Hz, OH-18); ^{13}C-NMR δ_C ($CDCl_3$, assigned by HSQC and HMBC data): 16.4 (q, C-20), 18.6 (q, C-19), 18.7 (t, C-2), 20.4 (t, C-6), 24.9 (q, C-17), 35.9 (t, C-3), 38.3 (s, C-4), 39.8 (s, C-10), 40.9 (t, C-12), 42.1 (t, C-1), 43.2 (t, C-7), 45.6 (s, C-8), 48.1 (d, C-13), 49.4 (t, C-14), 49.5 (d, C-5), 51.5 (t, C-15), 63.5 (d, C-9), 67.4 (d, C-11), 71.8 (t, C-18), 80.5 (s, C-16).

Large-scale preparation of the dephosphorylated derivative of the product converted from GGDP by recombinant PaDC2. The plasmid (pGEX-4T-3) that harbors *PaDC2* was transformed into *E. coli* BL21 (DE3). Transformants were grown at 37 °C in a 2× YT medium that contained ampicillin (50 μg/ml). When the optical density at 600 nm had reached 0.6, isopropyl-1-thio-β-D-galactoside was added to a final concentration of 0.2 mM. The bacteria, which were grown in 5000 ml of the 2× YT medium, were collected by centrifugation. To the pellets, 150 ml of a 50 mM Tris–HCl buffer (pH 7.5) that contained 0.1% Triton X-100 was added, and the suspension was subjected to ultrasonication and centrifugation to obtain a supernatant containing the enzyme. The reaction mixture contained 3 mg of GGDP (44 μM), a homogenous cell-free extract (140 ml), 1 mM $MgCl_2$, 0.5 mM EDTA, and 2 mM dithiothreitol in a total volume of 150 ml of a 50 mM Tris–HCl buffer solution (pH 7.5) plus 0.1% Triton X-100. After incubating at 25 °C for 15 h, and subjecting to a phosphatase treatment, the enzyme reaction was terminated by adding 100 ml of 15% KOH/MeOH and heating at 75 °C for 30 min, after which the lipophilic materials were extracted with hexane (150 ml × 3). The enzymatic product was purified by SiO_2 column chromatography, using gradient elution with hexane/EtOAc (100:0–100:5), which produced 1.2 mg of a pure compound. The structure of this was determined by mass and 2D-NMR spectral analyses to be that of copalol (oil)[8] with a specific rotation $[\alpha]_D^{24} = +30.6^\circ$ (*c* 0.12, $CHCl_3$). HRMS *m/z* (M^+): calcd. for $C_{20}H_{34}O$, 290.2610; found 290.2601; ^{1}H-NMR δ_H (400.13 MHz, C_6D_6): 0.855 (3H, s, H-20), 0.922 (3H, s, H-19), 0.963 (3H, s, H-18), 1.03 (m, H-1), 1.10 (dd, $J = 12.8$, 2.8 Hz, H-5), 1.23 (m, H-3), 1.38 (m, H-6), 1.48 (m, H-3), 1.53 (m, H-2), 1.60 (m, H-11), 1.65 (3H, s, H-16), 1.66 (m, H-2), 1.69 (m, H-9), 1.72 (m, H-11) 1.73 (m, H-6), 1.77 (m, H-1), 1.97 (m, H-12), 2.07 (m, H-7), 2.35 (m, H-12), 2.50 (m, H-7), 4.13 (2H, d, $J = 6.4$ Hz, H-15), 4.78 (bs, H-17), 5.07 (bs, H-17), 5.57 (bt, $J = 6.8$ Hz, H-14); ^{13}C-NMR δ_C (100.6 MHz, C_6D_6): 14.74 (C-20), 16.30 (C-16), 19.74 (C-2), 21.89 (C-19), 22.19 (C-11), 24.74 (C-6), 33.64 (C-4), 33.71 (C-18), 38.67 (C-7), 38.79 (C-12), 39.20 (C-1), 39.87 (C-10), 42.39 (C-3), 55.60 (C-5), 56.53 (C-9), 59.37 (C-15), 106.7 (C-17), 124.5 (C-14), 139.0 (C-13), 148.7 (C-8).

Results and Discussion

Identification of two diterpene biosynthesis gene clusters in P. amygdali

We identified by genome walking from *PaGGS1* and *PaGGS4* in *P. amygdali* N2 strain the genomic DNA sequences of about 26 kb and 14 kb, respectively (Fig. 1). These nucleotide sequences have been deposited in the databases with accession numbers AB254158 and AB254160, respectively. BLASTx searching identified a *diterpene cyclase*, five *P450 monooxygenases*, a *dehydrogenase*, and a *dioxygenase*-like gene proximal to *PaGGS1* (Fig. 1A), and a *diterpene cyclase*, two

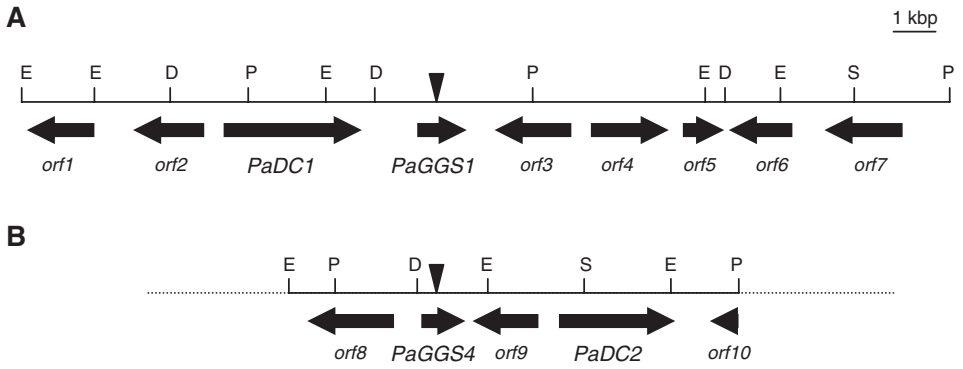


Fig. 1. Chromosome Map of Regions around *PaGGS1* (A) and *PaGGS4* (B).

The filled inverted triangle indicates the starting point for genome walking. These nucleotide sequences of A and B have been deposited in the databases with accession numbers AB254158 and AB254160, respectively. The restriction sites for *Dra*I, *Eco*RV, *Pvu*II and *Stu*I used in the genome walking process are represented by the letters D, E, P and S, respectively. The direction and deduced region of transcription are represented by arrows. Deduced *orf1*, *orf2*, *orf3*, *orf4*, *orf6*, *orf9*, and *orf10* represent *P450 monooxygenase*-like genes, and *orf5*, *orf7*, and *orf8* represent *dehydrogenase*, *dioxygenase*, and *transporter*-like genes, respectively. Homology searches were performed by using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/).

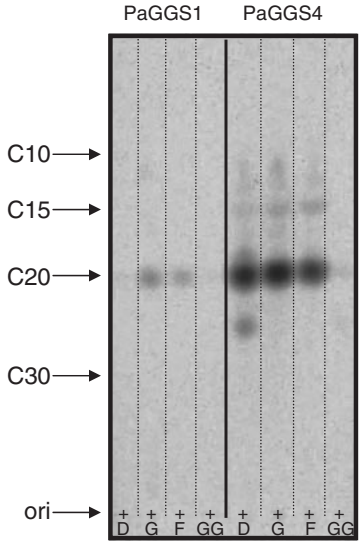


Fig. 2. TLC Autoradiography of the Alcohols Obtained by Hydrolysis of the Products Derived by Recombinant PaGGS1 and PaGGS4.

The '+' designation indicates the origin. DMADP (D), GDP (G), FDP(F), and GGDP (GG) were incubated with [^{14}C]IPP. Spots for authentic standard alcohols are indicated by arrows. C10, geraniol; C15, farnesol; C20, geranylgeraniol; C30, farnesylfarnesol.

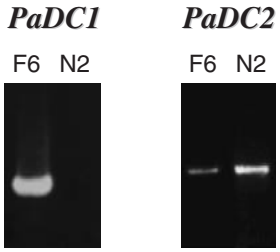


Fig. 3. Expression of *PaDC1* and *PaDC2* in Two Strains of *Phomopsis amygdali*.

RT-PCR was performed by using each set of gene-specific primers. F6 and N2 indicate template cDNAs derived from *P. amygdali* F6 and N2, respectively. Bands of approximately 3 kb are amplified.

P450 monooxygenases, and a *transporter*-like gene proximal to *PaGGS4* (Fig. 1B). We have named the two *diterpene cyclase*-like genes *PaDC1* and *PaDC2*, respectively. We determined the full-length cDNA sequences of *PaGGS1* and *PaGGS4* by RACE (nos. AB252830 and AB2528332), and carried out a functional analysis of each recombinant protein. Recombinant PaGGS1 catalyzed the conversion by IPP of GDP or FDP, but not DMAPP, into GGDP, and PaGGS4 catalyzed the conversion by IPP of DMAPP, GDP or FDP into GGDP (Fig. 2). It thus appears that both *PaGGS1* and *PaGGS4* encode GGS, which suggests that both of these genomic DNA regions contain diterpene biosynthesis gene clusters.

Isolation of ORF cDNAs of PaDC1 *and* PaDC2

To isolate the cDNA fragments that contain *PaDC1* and *PaDC2*, we performed RT-PCR using a cDNA pool derived from the mycelia of *P. amygdali* N2 as a template. The design of the 5′- and 3′-end primers for amplifying each ORF was based on the genomic DNA sequences. However, no band was amplified for *PaDC1*, whereas a band of the expected size was amplified for *PaDC2* (Fig. 3, lane N). Furthermore, we could not amplify the *PaDC1* band with certain combinations of other gene-specific primers (data not shown). In the mycelia of *P. amygdali* N2, fusicoccadiene-related hydrocarbons make up most of the hydrocarbon fraction.[14] However, in the hydrocarbon fraction of the extract from the mycelia of *P. amygdali* F6, several labdane-related hydrocarbons derived from (+)-CDP were detected, in addition to the foregoing hydrocarbons.[14] We therefore performed RT-PCR, using as

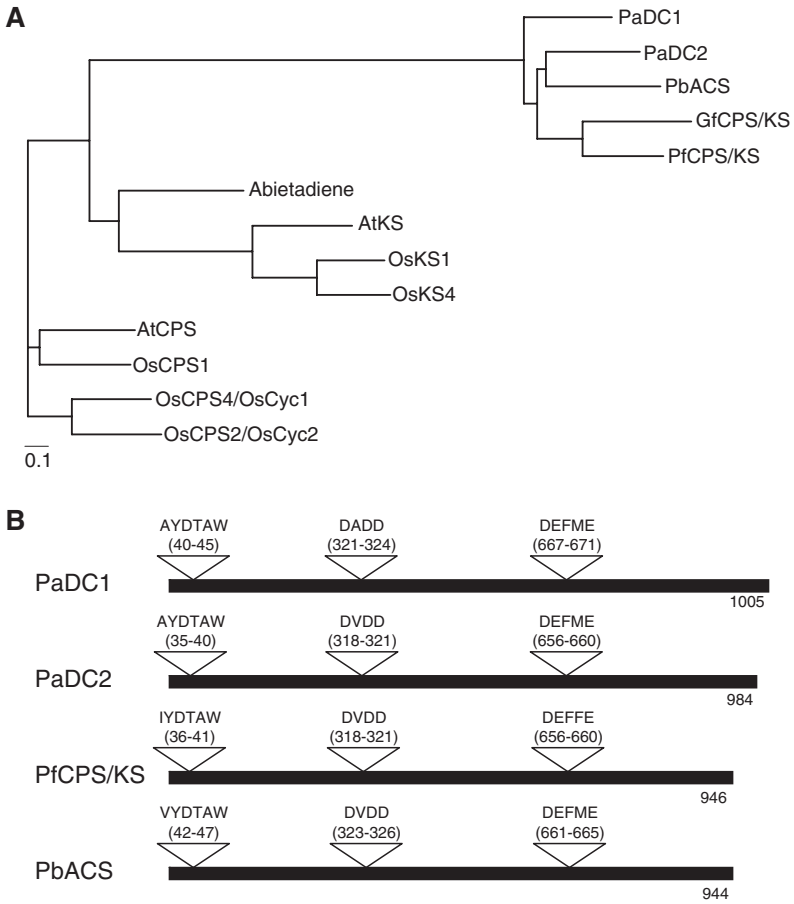


Fig. 4. Phylogenetic Tree and Schematic Primary Structures of PaDC1 and PaDC2.

A, The phylogenetic tree generated by CLUSTALW software (http://clustalw.ddbj.nig.ac.jp/top-j.html) includes the following: abeitadiene, grand fir abietadiene synthase (AAB05407); AtCPS, *Arabidopsis ent*-CDP synthase (NP_192187); AtKS, *Arabidopsis ent*-kaurene synthase (AAC39443); GfCPS/KS, *Gibberella fujikuroi ent*-kaurene synthase (AB013295); OsCPS1, rice *ent*-CDP synthase (BAF08464); OsCPS2/OsCyc2, rice *ent*-CDP synthase (AB066271); OsCPS4/OsCyc1, rice *syn*-CDP synthase (AB066270); OsKS1, rice *ent*-kaurene synthase (AN); OsKS4, rice 9b-pimara-7,15-diene synthase (AB126934); PbACS, *Phoma betae* aphidicola-16β-ol synthase (AB049075); PfCPS/KS, *Phaeosphaeria ent*-kaurene synthase (AB003395). B, Primary structures of PaDC1, PaDC2, PfCPS/KS, and PbACS. The inverted unfilled triangles indicate the AYDTAW, DDxxD and DExxE motifs. The AYDTAW motif is widely conserved among diterpene cyclases in plants and fungi. The DxDD and DExxE motifs in PfCPS/KS are suggested to be responsible for the cyclization of type B and type A, respectively, as described in the text.

template the cDNA from *P. amygdali* F6 mycelia. A band of the expected size for *PaDC1* was amplified, as well as a band for *PaDC2* (Fig. 3, lane F). Both bands were subcloned and sequenced. *PaDC1* (accession number AB252834) and *PaDC2* (AB252835) were found to encode 1005 amino acids (113 kDa) and 984 amino acids (110 kDa), respectively. The nucleotide sequence of the coding region of *PaDC2* from N2 was identical to that of *PaDC2* from F6.

Homology searches indicated that both PaDC1 and PaDC2 were homologous to the labdane-related diterpene cyclases, *ent*-kaurene synthases from *Phaeosphaeria* (PfCPS/KS)[15] and *Gibberella fujikuroi* (GfCPS/KS),[16] and aphidicolan-16β-ol synthase from *Phoma betae* (PbACS)[17] (Fig. 4A). These bifunctional diterpene cyclases convert GGDP sequentially into (−)-*ent*-CDP or (+)-*syn*-CDP, and into the specific cyclic hydrocarbons. The former reaction is initiated by protonation of the terminal double bond of GGDP (type B), and the latter is initiated by ionization of the diphosphate (type A).[18] A functional analysis with truncated mutants and site-directed mutagenesis of PfCPS/KS have demonstrated the aspartate/glutamate-rich motif DxDD in the N-terminal domain and DExxE in the C-terminal domain to be responsible for the formation of *ent*-CDP (type B cyclization) and *ent*-kaurene (type A cyclization), respectively.[19] Both PaDC1 and PaDC2 contained these two motifs, as well as PfCPS/KS, GfCPS/KS and PbACS; DADD (positions 321-324 in PaDC1) and DVDD (positions 318-321 in PaDC2) represent the DxDD motif, and DEFME (positions 667-671 in PaDC1) and DEFME (positions 656-660 in PaDC2) represent the DExxE motif (Fig. 4B). These results suggest that both *PaDC1* and *PaDC2* encode bifunctional labdane-related diterpene cyclases.

Functions of PaDC1 and PaDC2

To study the functions of PaDC1 and PaDC2, terpene cyclase assays using recombinant proteins were performed. The results of the GC-MS analysis to identify a product of GST-PaDC1 are shown in Fig. 5. A functional analysis with the GST-PaDC1 recombinant enzyme, using GDP, FDP or GGDP as a substrate, showed that GGDP was converted into peak 1 at a retention time 23 min on GC (Fig. 5A), whereas there was no conversion of either GDP or FDP (data not shown). The retention time and mass spectrum of peak 1 (Fig. 5A and D) were identical to those of phyllocladan-16α-ol (**11**) (Fig. 6) isolated from the hexane extract of *P. amygdali* F6 mycelia (Fig. 5C and F, peak 3). The identification of phyllocladan-16α-ol (**11**) in the mycelia is described next. Recombinant GST-PaDC1 also converted (+)-CDP into phyllocladan-16α-ol (**11**) (Fig. 5B and E, peak 2). These results indicate that PaDC1 converted GGDP into phyllocladan-16α-ol (**11**) *via* (+)-CDP (Fig. 6). The aspartate-rich motifs, DADD (positions 321-324) and DEFME (positions 667-671), in PaDC1 (Fig. 4B) should be responsible for the conversion of GGDP into (+)-CDP (type B cyclization) and for the conversion of (+)-CDP into phyllocladan-16α-ol (**11**) (type A cyclization), respectively. In the reaction mixture, trace amounts of (+)-isopimara-8,15-diene (**6**) (15.2 min), pimara-8,15-diene (**3**) (15.5 min), (+)-pimara-8(14),15-diene (**4**) (15.9 min), (−)-sandaracopimaradiene (**7**) (16.3 min), (+)-phyllocladene (**8**) (17.7 min), and (+)-kaurene (**5**) (18.3 min) were detected as by-products (data not shown). These labdane-related compounds have also been identified in the hydrocarbon fraction from *P. amygdali* F6 mycelia.[13] The general biogenetic pathways of these hydrocarbons are shown in Fig. 6A. (+)-CDP, which is converted from GGDP by type B cyclization, undergoes successive type A cyclization from *re*- or *si*-face attack of the C-13 olefin to generate tricyclic carbocations **1** and **2**. Deprotonation of 10-H and 14-H in **1** gave **3** and **4**, respectively, whereas sequential rearrangement of **1** provided **5**. While deprotonation of 10-H and 14-H in **2** gave **6** and **7**, respectively, sequential rearrangement of **2** afforded **8**. During the rearrangement of **2**, carbocation intermediate **10** was formed *via* **9**, and deprotonation of 17-H in **10** provided **8**, while stereoselective attack of the water molecules generated phyllocladan-16α-ol (**11**) (Fig. 6B). Therefore, *PaDC1* is responsible for the synthesis of the labdane-related hydrocarbons in *P. amygdali* F6. Although labdane-related diterpene cyclase *via* (+)-CDP as an intermediate has previously been isolated from a plant, *i.e.*, abietadiene synthase in grand fir,[20] such a diterpene cyclase has not been previously identified in fungi.

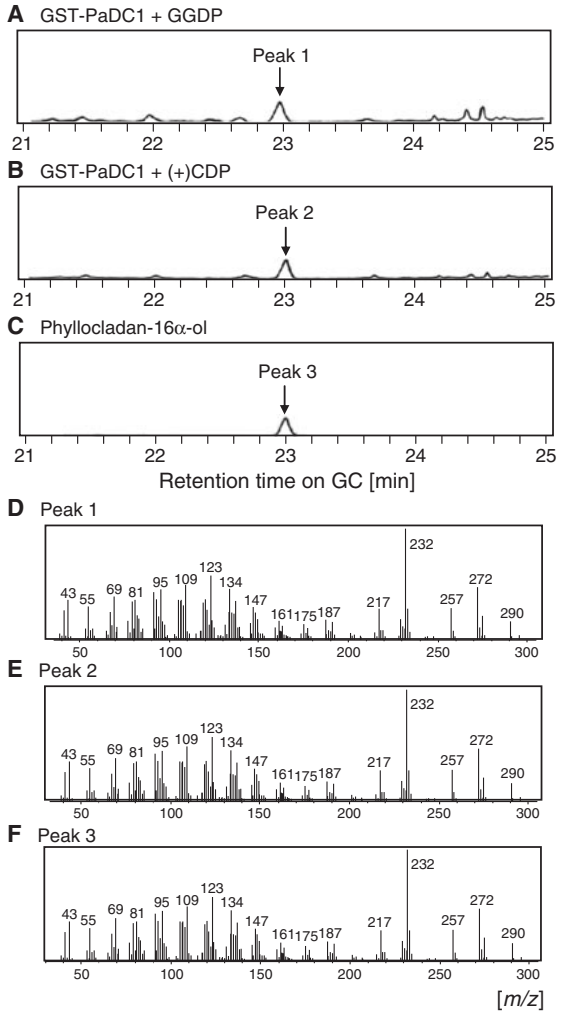


Fig. 5. GC-MS Results for Products Converted from GGDP by GST-PaDC1.

A, Total ion chromatograms of the products derived from the incubation of GGDP with recombinant GST-PaDC1. B, Total ion chromatograms of the products derived from the incubation of (+)-CDP with recombinant GST-PaDC1. C, Total ion chromatogram of authentic phyllocladan-16α-ol obtained from the mycelia of *P. amygdali* F6. D, Full-scan mass spectrum of peak 1. E, Full-scan mass spectrum of peak 2. F, Full-scan mass spectrum of phyllocladan-16α-ol (peak 3).

The functional analysis conducted with GST-PaDC2 showed that GDP, FDP and GGDP were not converted into hydrocarbon (data not shown). Therefore, the reaction products were subjected to GC-MS after dephosphorylating with alkaline phosphatase. Peak 4 with a retention time of 23.6 min was detected on GC from the reaction mixture by using GGDP as the substrate (Fig. 7A). The retention time and mass spectrum of peak 4 (Fig. 7A and C), representing the reaction products derived from incubating GGDP with GST-PaDC2, were identical to those of authentic (+)-copalol previously synthesized[8] (Fig. 7B and D, peak 5). It was not possible to distinguish between (+)-copalol and (−)-*ent*-copalol by our GC-MS system. We therefore determined the absolute stereostructure of the dephosphorylated derivative by measuring the specific rotation. The $[\alpha]_D$ value for copalol derived from the reaction product was +30.6, which is similar to the previously published value of +29.8,[21] demonstrating that PaDC2 converted GGDP into (+)-CDP, but not

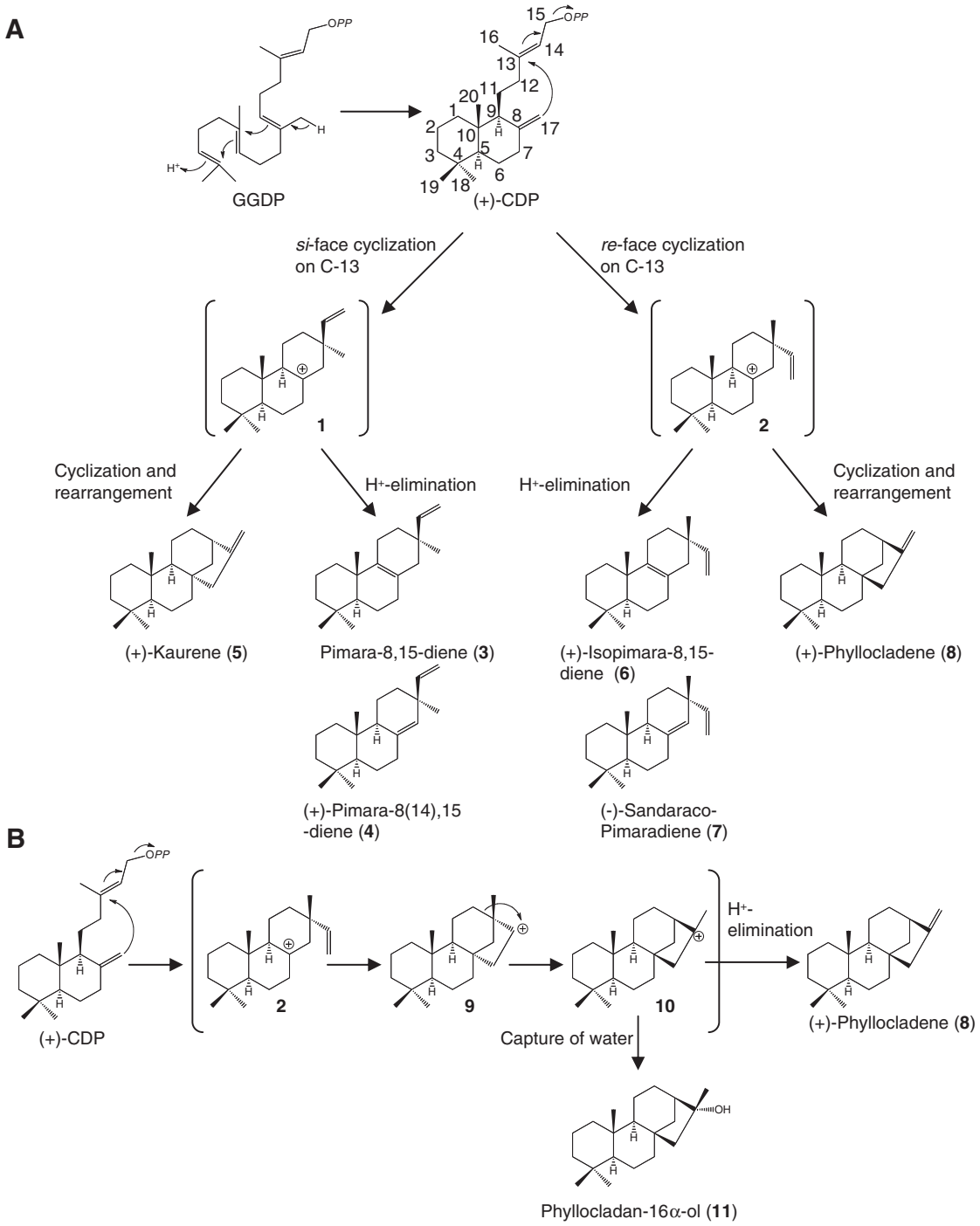


Fig. 6. The Biogenetic Pathways to Diterpene Hydrocarbons and Phyllocladan-16α-ol.
A, Formation of diterpene hydrocarbons from (+)-CDP by PaDC1. B, Formation of phyllocladan-16α-ol and phyllocladane by PaDC1.

into *ent*-CDP. Aspartate/glutamate-rich motif DVDD (positions 318-321; Fig. 4B) was expected to be responsible for the conversion of GGDP into (+)-CDP (type B cyclization). However, we did not detect any further cyclized product of (+)-CDP, even though PaDC2 contains another aspartate/glutamate-rich motif, DEFME (positions 656-660; Fig. 4B), which might be involved in type A cyclization. It is possible that a mutation occurred in the 3′-region, outside the coding region of the DEFME motif, in *PaDC2*. These results indicate that *PaDC2* encodes the (+)-CDP synthase. Labdane-related diterpenoids are biosynthesized from GGDP through specific diterpene hydrocarbon skeletal structures *via* a CDP intermediate, of which four common stereoisomers are known: *ent*-CDP, (+)-CDP, (+)-*syn*-CDP, and *ent*-*syn*-CDP.[18] Although the cDNAs that encode *ent*-CDP synthase and *syn*-CDP synthase have been isolated from plants and bacteria,[5,12,22,23] the cDNAs for (+)-CDP synthase and *ent*-*syn*-CDP synthase have not been identified in any organism. Thus,

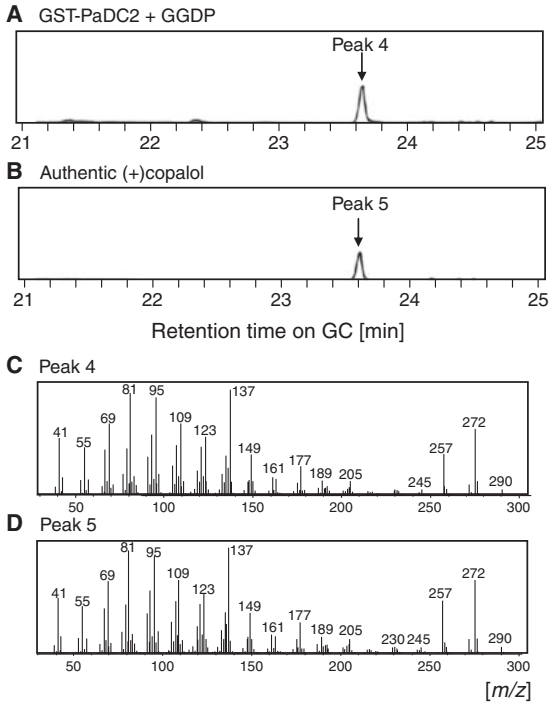


Fig. 7. GC-MS Results for the Dephosphorylated Derivatives of Products Converted from GGDP by GST-PaDC2.

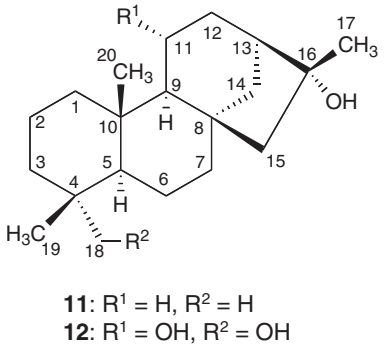
A, Total ion chromatograms of the products derived from the incubation of GGDP with recombinant GST-PaDC2. B, Total ion chromatogram of synthetic (+)-CDP. C, Full-scan mass spectrum of peak 4. D, Full-scan mass spectrum of peak 5.

this is the first identification of a natural enzyme which produces (+)-CDP as an end product.

These results suggest that *P. amygdali* N2 did not produce any labdane-related compounds,[14] probably owing to the suppression of *PaDC1* expression and loss of function of the C-terminal type A cyclization domain of PaDC2.

Identification of phyllocladan-16α-ol and its related compound in the mycelia of P. amygdali

Since we could not identify phyllocladan-16α-ol (**11**) in the hydrocarbon fraction of the *P. amygdali* F6 mycelial extract, as described in a previous report,[13] we purified compound **11** from the hexane extract of *P. amygdali* F6 mycelia and performed a GC-MS analysis (Fig. 5C and F). Peak 3 gave the M^+ ion at m/z 290, and characteristic fragment ions at m/z 272, 257, 217, 191, 147, 134 and 123. The mass spectrum of **11** was identical to that of phyllocladan-16α-ol listed in the mass spectral libraries (Vender Kit Wiley D.02.00; Agilent). Although **11** had very similar fragment ions to those of phyllocladan-16β-ol in the mass spectral libraries, the diagnostic ion at m/z 212 observed in phyllocladan-16β-ol was significantly different. Furthermore, the ^{1}H-NMR chemical shifts of the methyl groups of **11** were in good agreement with those of phyllocladan-16α-ol [**11**: δ_H (400 MHz, $CDCl_3$) 1.33 (3H, s, Me-17), 0.87 (3H, s, Me-18), 0.85 (3H, s, Me-19), 0.80 (3H, s, Me-20)].[24] These data strongly support the designation of **11** as phyllocladan-16α-ol (Fig. 8).



11: R^1 = H, R^2 = H
12: R^1 = OH, R^2 = OH

Fig. 8. Structures of Phyllocladan-16α-ol (**11**) and Phyllocladan-11α,16α,18-triol (**12**).

In general, the cyclic diterpene synthase enzymes convert a common acyclic biosynthetic precursor (GGDP) to diterpenes which are modified, initially by hydroxylation and subsequently by oxidation, to their corresponding alcohols, aldehydes, and acids. In our studies of phyllocladane-related diterpenes from this fungal strain, hydroxylated phyllocladane derivative **12** (Fig. 8) was isolated in the form of colorless needles. The molecular formula for **12**, $C_{20}H_{34}O_3$, which was determined by HRFABMS, has two additional oxygen atoms, as compared to **11**. The ^{1}H-NMR spectrum of **12** was similar to that of **11**, except for the presence of oxymethine [δ_H 4.27 (1H, m, H-11)] and oxymethylene signals [δ_H 3.22 (1H, dd, J = 10.8, 5.4 Hz, H-18) and 3.68 (1H, dd, J = 10.8, 5.4 Hz, H-18)]. A careful comparison of the ^{13}C-NMR spectra of **12** and phyllocladan-16β-ol[25] revealed the presence of two additional hydroxyl groups at C-11 and C-18, based on the substitution effects at C-4, C-9, C-11, C-12, and C-18. This was confirmed by HMBC correlations between OH-11 and C-9, between OH-18 and C-18, and between Me-19 and C-18. The relative stereochemistry of **12** was inferred from the nuclear Overhauser effects (NOEs) from H-11 to H-15β, Me-17 and Me-20, and from Me-20 to Me-19 and H-15β.

Although the absolute configurations of **11** and **12** have not been confirmed, these compounds are presumed to be (+)-phyllocladane (13β-kaurane)-type diterpenoids on the basis of biosynthetic considerations, as well as by the fact that all natural cyclic diterpene hydrocarbons isolated from this fungal strain to date have been assigned to the (+)-kaurane series (Fig. 7A).[13] We therefore conclude that **12** is phyllocladan-11α,16α,18-triol, which is reported here for the first time. Since the formation of **12** is thought to occur *via* **11**, this being followed by hydroxylation, it is biogenetically plausible that *cytochrome P450 oxygenase*-like genes near the *PaDC1* gene in the diterpene synthesis gene cluster (Fig. 1) are involved in the conversion of **11** into **12**. The biological activities of

11 and **12** have not been examined. However, **11** and/or **12** should possess unique biological activity, as it has been reported that a phyllocladanol-related compound from heartwood of *Cryptomeria japonica* exhibited termiticidal activity.[26)]

We identified in the present study two novel diterpene synthases, phyllocladan-16α-ol synthase and (+)-CDP synthase, and the novel diterpenoid metabolite, phyllocladan-11α,16α,18-triol. This is the first report of the identification of a novel terpenoid through the genome walking strategy. The identification of novel substances may indicate lead compounds for the creation of new medical and agricultural chemicals. We propose that genome walking offers a useful strategy for the discovery of novel natural products in fungi.

Acknowledgments

We are grateful to Dr. S. Kanematsu (Apple Research Center, National Institute of Fruit Tree Science) and Professor A. Graniti (Bari University) for providing the fungus, *P. amygdali* N2 and F6, respectively. We also thank K. Itoh, A. Kaneko (Yamagata University) and C. Ikeda (Toyama Prefectural University) for technical support. This study was supported in part by Grant-in-Aid for Scientific Research (A) (Nos. 13306009 and 16208012) from the Japanese Society for the Promotion of Science.

References

1) Yabuta, T., and Sumiki, Y., On the crystal of gibberellin, a substance to promote plant growth (in Japanese). *J. Agric. Chem. Soc. Japan*, **14**, 1526 (1938).
2) Ikegami, S., Taguchi, T., Ohashi, M., Oguro, M., Nagano, H., and Mano, Y., Aphidicolin prevents mitotic cell division by interfering with the activity of DNA polymerase α. *Nature*, **275**, 458–460 (1978).
3) Wurtele, M., Jelich-Ottmann, C., Wittinghofer, A., and Oecking, C., Structural view of a fungal toxin acting on a 14-3-3 regulatory complex. *EMBO J.*, **22**, 987–994 (2003).
4) Bunney, T. D., De Boer, A. H., and Levin, M., Fusicoccin signaling reveals 14-3-3 protein function as a novel step in left-right patterning during amphibian embryogenesis. *Development*, **130**, 4847–4858 (2003).
5) Hedden, P., Phillips, A. L., Cecilia Rojas, M., Carrera, E., and Tudzynski, B., Gibberellin biosynthesis in plants and fungi: a case of convergent evolution? *J. Plant Growth Regul.*, **20**, 319–331 (2002).
6) Toyomasu, T., Nakaminami, K., Toshima, H., Mie, T., Watanabe, K., Ito, H., Matsui, H., Mitsuhashi, W., Sassa, T., and Oikawa, H., Cloning of a gene cluster responsible for the biosynthesis of diterpene aphidicolin, a specific inhibitor of DNA polymerase α. *Biosci. Biotechnol. Biochem.*, **68**, 146–152 (2004).
7) Toyomasu, T., Tsukahara, M., Kaneko, A., Niide, R., Mitsuhashi, W., Dairi, T., Kato, N., and Sassa, T., Fusicoccins are biosynthesized by an unusual chimera diterpene synthase in fungi. *Proc. Natl. Acad. Sci. USA*, **104**, 3084–3088 (2007).
8) Toshima, H., Oikawa, H., Toyomasu, T., and Sassa, T., Total synthesis of both enantiomers of copalol *via* optical resolution of a general synthetic intermediate for drimarene sesquiterpenes and labdane diterpenes. *Tetrahedron*, **56**, 8443–8450 (2000).
9) Mohan, R. S., Yee, N. K. N., Coates, R. M., Ren, Y.-Y., Stamenkovic, P., Mendez, I., and West, C. A., Biosynthesis of cyclic diterpene hydrocarbons in rice cell suspension: conversion of 9,10-*syn*-labda-8(17),13-dienyl diposphate to 9β-pimara-7,15-diene and stemar-13-ene. *Arch. Biochem. Biophys.*, **330**, 33–47 (1996).
10) Toyomasu, T., Kawaide, H., Mitsuhashi, W., Inoue, Y., and Kamiya, Y., Phytochrome regulates gibberellin biosynthesis during germination of lettuce seeds. *Plant Physiol.*, **118**, 1517–1523 (1998).
11) Kawasaki, T., Hamano, Y., Kuzuyama, T., Itoh, N., Seto, H., and Dairi, T., Interconversion of product specificity of type I eubacterial farnesyl diphosphate synthase and geranylgeranyl diphosphate synthase by one amino acid substitution. *J. Biochem.*, **133**, 83–91 (2003).
12) Otomo, K., Kenmoku, H., Oikawa, H., Konig, W. A., Toshima, H., Mitsuhashi, W., Yamane, H., Sassa, T., and Toyomasu, T., Biological functions of *ent*- and *syn*-copalyl diphosphate synthases in rice: key enzymes for the branch point of gibberellin and phytoalexin biosynthesis. *Plant J.*, **39**, 886–893 (2004).
13) Kenmoku, H., Tanaka, M., Ogiyama, K., Kato, N., and Sassa, T., Identification of (+)-phyllocladen, (−)-sandaracopimaradiene, and (+)-kaurene as new fungal metabolites from fusicoccin-producing *Phomopsis amygdali* F6. *Biosci. Biotechnol. Biochem.*, **68**, 1574–1577 (2004).
14) Sassa, T., Kenmoku, H., Nakayama, K., and Kato, N., Fusicocca-3(16),10(14)-diene, and β- and δ-araneosenes, new fusicoccin biosynthesis-related diterpene hydrocarbons from *Phomopsis amygdali*. *Biosci. Biotechnol. Biochem.*, **68**, 1608–1610 (2004).
15) Kawaide, H., Imai, R., Sassa, T., and Kamiya, Y., *ent*-Kaurene synthase from the fungus *Phaeosphaeria* sp. L487. cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase in fungal gibberellin biosynthesis. *J. Biol. Chem.*, **272**, 21706–21712 (1997).
16) Toyomasu, T., Kawaide, H., Ishizaki, A., Shinoda, S., Otsuka, M., Mitsuhashi, W., and Sassa, T., Cloning of a full-length cDNA encoding *ent*-kaurene synthase from *Gibberella fujikuroi*: functional analysis of a bifunctional deterpene cyclase. *Biosci. Biotechnol. Biochem.*, **64**, 660–664 (2000).
17) Oikawa, H., Toyomasu, T., Toshima, H., Ohashi, S., Kawaide, H., Kamiya, Y., Ohtsuka, M., Shinoda, S., Mitsuhashi, W., and Sassa, T., Cloning and functional expression of cDNA encoding aphidicolan-16β-ol synthase: a key enzyme responsible for formation of an unusual diterpene skeleton in biosynthesis of aphidicolin. *J. Am. Chem. Soc.*, **123**, 5154–5155 (2001).
18) MacMillan, J., and Beale, M. H., Diterpene biosynthesis. In "Comprehensive Natural Products Chemistry No. 2, Isoprenoids Including Carotenoids and Steroids," eds. Barton, S. D., and Nakanishi, K., Elsevier Science, Oxford, pp. 217–243 (1999).
19) Kawaide, H., Sassa, T., and Kamiya, Y., Functional

analysis of the two interacting cyclase domains in *ent*-kaurene synthase from the fungus *Phaeosphaeria* sp. L487 and a comparison with cyclases from higher plants. *J. Biol. Chem.*, **275**, 2276–2280 (2000).

20) Vogel, B. S., Wildung, M. R., Vogel, G., and Croteau, R., Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *J. Biol. Chem.*, **271**, 23262–23268 (1996).

21) Yee, N. K. N., and Coates, R. M., Total synthesis of (+)-9,10-*syn*-copalol and (+)-9,10-*anti*-copalol *via* epoxy trienylsilane cyclization. *J. Org. Chem.*, **57**, 4598–4608 (1992).

22) Xu, M., Hillwig, M. L., Prisic, S., Coates, R. M., and Peters, R. J., Functional identification of rice *syn*-copalyl diphosphate synthase and its role in initiating biosynthesis of diterpenoid phytoalexin/allelopathic natural products. *Plant J.*, **39**, 309–318 (2004).

23) Ikeda, C., Hayashi, Y., Itoh, N., Seto, H., and Dairi, T., Functional analysis of eubacterial *ent*-copalyl diphosphate synthase and pimara-9(11),15-diene synthase with unique primary sequences. *J. Biochem.*, **141**, 37–45 (2007).

24) Carman, R. M., and Sutherland, M. D., Cupressene and other diterpenes of Cupressus species. *Aust. J. Chem.*, **32**, 1131–1142 (1979).

25) Carman, R. M., Phyllocladan-16β-ol (13β-kauran-16β-ol). *Aust. J. Chem.*, **34**, 923–926 (1981).

26) Sogabe, A., Kinjo, K., Abe, F., Yamauchi, T., and Yaga, S., Termiticidal substances from the heartwood of *Cryptomeria japonica* D. Don. *Mokuzai Gakkaishi*, **46**, 124–131 (2000).