JOC Article

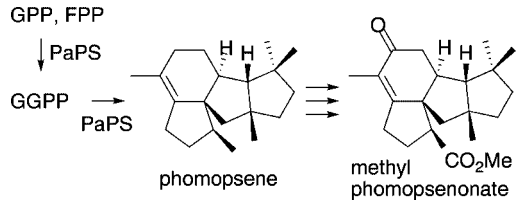
Biosynthetic Gene-Based Secondary Metabolite Screening: A New Diterpene, Methyl Phomopsenonate, from the Fungus *Phomopsis amygdali*

Tomonobu Toyomasu,[†] Akane Kaneko,[†] Tetsuo Tokiwano,[‡,¶] Yuya Kanno,[†] Yuri Kanno,[†,||] Rie Niida,[†] Shigeyoshi Miura,[†] Taiki Nishioka,[†] Chiho Ikeda,[§] Wataru Mitsuhashi,[†] Tohru Dairi,[§] Tomikazu Kawano,[⊥,#] Hideaki Oikawa,*[,‡] Nobuo Kato,[⊥] and Takeshi Sassa[†]

Department of Bioresource Engineering, Yamagata University, Tsuruoka, Yamagata 997-8555, Japan, Division of Chemistry, Graduate School of Science, Hokkaido University, Sapporo 060-0810, Japan, Biotechnology Research Center, Toyama Prefectural University, Imizu, Toyoma 939-0398, Japan, and The Institute of Scientific and Industrial Research, Osaka University, Ibaraki, Osaka 567-0047, Japan

hoik@sci.hokudai.ac.jp

Received October 16, 2008



The presence of the geranylgeranyl diphosphate synthase (GGS) gene is a common feature of gene clusters for diterpene biosynthesis. We demonstrated identification of a diterpene gene cluster using homology-based PCR of GGS genes and the subsequent genome walking in the fungus *Phomopsis amygdali* N2. Structure determination of a novel diterpene hydrocarbon phomopsene provided by enzymatic synthesis with the recombinant terpene synthase PaPS and screening of fungal broth extracts with reference to characteristic NMR signals of phomopsene allowed us to isolate a new diterpene, methyl phomopsenonate. The versatility of the gene-based screening of unidentified diterpenes is discussed in regard to fungal genomic data.

Introduction

Terpenoids are a rich source of bioactive compounds and have a large structural diversity with a molecular skeleton constructed by the cyclization of oligoprenyl diphosphate with terpene cyclase.[1] Thus, terpene cyclase play a key role in the construction of characteristic skeletons of terpenoids. Clustering of the genes responsible for biosynthesis of secondary metabolites is a common feature in most microorganisms including streptomycetes[2] and fungi.[3] Therefore, homology-based PCR screening of the terpene cyclase gene should be useful for identifying gene cluster of terpenoid biosynthesis. However, this approach is less frequently used, although a similar approach is generally applied for searching for gene clusters of polyketides[4] and non-ribosomal polypeptides[5] whose biosyntheses involve highly homologous modular enzymes for construction of molecular backbones. The major reason for less frequent use of the PCR screening is that microbial terpene cyclases show a low homology[6] compared with the modular enzymes described above. This is a clear difference between plant and microbial terpene cyclases since

[†] Yamagata University.
[‡] Hokkaido University.
[§] Toyama Prefectural University
[⊥] Osaka University.
[¶] Current address: Department of Biotechnology, Faculty of Bioresource Sciences, Akita Prefectural University, Akita 010-0195, Japan.
[||] Current address: Plant Science Center, RIKEN, Yokohama, Kanagawa 230-0045, Japan.
[#] Current address: School of Pharmacy, Iwate Medical University, Yahaba, Shiwa-gun, Iwate 028-3694, Japan.

(1) Cane, D. In *Isoprenoids Including Carotenoids and Steroids*; Vol. 2 in *Comprehensive Natural Products Chemistry*; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Elsevier: Amsterdam, 1999.

(2) (a) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531. (b) Oliynyk, M.; Samborskyy, M.; Lester, J. B.; Mironenko, T.; Scott, N.; Dickens, S.; Haydock, S. F.; Leadlay, P. F. *Nat. Biotechnol.* **2007**, *25*, 447–453.

(3) Keller, N. P.; Turner, G.; Bennett, J. W. *Nat. Rev. Microbiol.* **2005**, *3*, 937–947.

(4) Rascher, A.; Hu, Z. H.; Viswanathan, N.; Schirmer, A.; Reid, R.; Nierman, W. C.; Lewis, M.; Hutchinson, C. R. *FEMS Microbiol. Lett.* **2003**, *218*, 223–230.

(5) Martens, T.; Gram, L.; Grossart, H. P.; Kessler, D.; Muller, R.; Simon, M.; Wenzel, S. C.; Brinkhoff, T. *Microb. Ecol.* **2007**, *54*, 31–42.

10.1021/jo802319e CCC: $40.75 © 2009 American Chemical Society
Published on Web 01/22/2009

plant terpene cyclases have a relatively high homology and homology-based PCR is useful for cloning of the corresponding genes.[7]

During our search for novel diterpene syntheses from fungi, we and others found that the occurrence of geranylgeranyl diphosphate synthase (GGS) on the gene clusters is a common feature in several biosynthetic diterpene gene clusters such as gibberellin[8] and aphidicolin.[9] In addition, *GGS* was also found in the gene cluster for the biosynthesis of the indole diterpene, paxilline.[10] Fungal GGS possessed relatively high homology to design degenerated primers, suggesting that diterpene gene clusters would be securely cloned by identification of both constitutive and cluster-specific GGS genes and subsequent genome walking. Recently, homology-based PCR cloning was successfully applied to clone diterpene gene clusters[11−13] and was also used to identify a gene cluster for terpenoids with an isoprene unit more than C20.[14]

The plant pathogenic fungus *Phomopsis amygdali* produces diterpene phytotoxin fusicoccins[15] whose structurally related derivative cotylenin is a promising drug lead as a differentiation-inducing agent on cancer cells.[16] We recently found the partial biosynthetic gene cluster of fusicoccins in *P. amygdali*, using homology-based PCR cloning and genome walking.[11] This study allowed us to find a *P. amygdali* fusicoccadiene synthase (PaFS), which is responsible for construction of the fusicoccane skeleton. PaPS constitutes two domains, a terpene cyclase domain at the N-terminus and a prenyltransferase domain at the C-terminus.[11] During our genome walking, we unexpectedly found another chimera cyclase, PaFS homologue. In this paper, we describe a typical example of biosynthetic gene-based screening of novel diterpenes phomopsene and methyl phomopsenonate.

Results

Cloning and Sequencing of the Partial Diterpene Gene Cluster Including PaFS Homologue. In the study identifying the diterpene gene cluster in *P. amygdali*, degenerate primers designed to amplify GGS genes essential for the biosynthesis of the diterpene metabolites were used to amplify the corresponding homologues.[11] Previously, we reported six unique GGS gene fragments (*PaGGS1, 2, 3, 4, 5,* and *6*) among the amplimers cloned and sequenced.[11] When *PaGGS3* was used for genome walking, we found the partial sequence of a terpene cyclase-like gene.

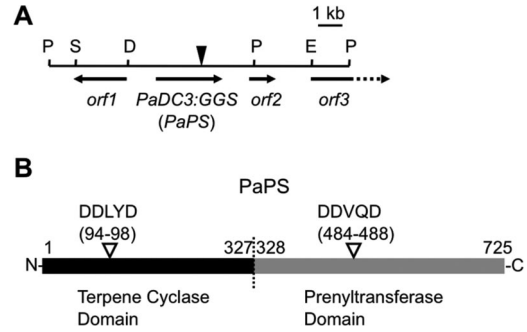


FIGURE 1. Chromosome map around *PaPS* and the primary structure of PaPS. (A) Chromosome map. A filled reverse triangle indicates the start point of genome walking. The restriction sites for *Dra* I, *Eco* RV, *Pvu* II, and *Stu* I used in genome walking are represented by the letters D, E, P, and S, respectively. The direction and deduced region of transcription are represented by arrows. The deduced *orf1*, *orf2*, and *orf3* represent *P450 monooxygenase*, *peroxiredoxin*, and *transporter* like genes, respectively. Homology searches were performed using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). (B) Schematic of two putative domains of PaPS. Black and gray bars indicate the terpene cyclase domain and the prenyltransferase domain, respectively. Reverse open triangles indicate the DDxxD motives of both domains. The borders between the terpene cyclase domain and the prenyltransferase domain were deduced by the homology with other fungal aristolochene synthases or GGSs.

The nucleotide sequence of the full-length cDNA was determined by a rapid amplification of cDNA ends (RACE) using gene-specific primers. This contained the predicted 2178-bp open reading frame (ORF), encoding a product of 725 amino acids, indicating that *PaGGS3* is fused to the terpene cyclase-like gene at the 3′ end, like *PaFS*.[11] This gene contained five intron insertion sites, on the basis of comparison with the corresponding genomic DNA sequence. According to the results shown later, we named this fused-type gene as a *P. amygdali* phomopsene synthase gene (*PaPS*). The predicted peptide constitutes two possible domains, the putative terpene cyclase domain (position 1−327) at the N-terminus and the putative prenyltransferase domain (position 328−725) at the C-terminus (Figure 1B). The enzyme encoded by *PaPS* exhibited homology with PaFS (35% identitiy and 53% similarity). The DDxxD motives, proposed to coordinate with the Mg^{2+} ion,[6a,7] was conserved at both domains: DDLYD (position 94−98) and DDVQD (position 484−488) (Figure 1B).

At the 5′ flanking region of *PaPS*, we found a putative cytochrome P-450 monooxygenase gene (*orf1*: *Aspergillus clavatus* NRRL 1, identitiy 40%, similarity 58%) and a putative peroxiredoxin (*orf2*: identity *Taiwanofungus camphoratus*61%, similarity 75%). At the 3′ flanking region, partial ORF (*orf3Emericella nidulans*) that encodes a putative self-resistance protein was also found. Although sequence data for the whole gene cluster is not currently available, the information shown above strongly indicated that the gene cluster is responsible for the biosynthesis of diterpene metabolites.

(6) (a) Hohn, T. M. In *Isoprenoids Including Carotenoids and Steroids*; Vol. 2 in *Comprehensive Natural Products Chemistry*; Barton, D., Nakanishi, K., Meth-Cohn, O., Eds.; Elsevier: Amsterdam, 1999; pp 201−215. (b) Kawaide, H.; Imai, R.; Sassa, T.; Kamiya, Y. *J. Biol. Chem.* **1997**, *272*, 21706–21712. (c) Toyomasu, T.; Kawaide, H.; Ishizaki, A.; Shinoda, S.; Otsuka, M.; Mitsuhashi, W.; Sassa, T. *Biosci. Biotechnol. Biochem.* **2000**, *64*, 660–664. (d) Oikawa, H.; Toyomasu, T.; Toshima, H.; Ohashi, S.; Kawaide, H.; Kamiya, H.; Ohtsuka, M.; Shinoda, S.; Mitsuhashi, W.; Sassa, T. *J. Am. Chem. Soc.* **2001**, *123*, 5154–5155.

(7) Bohlmann, J.; Meyer-Gauen, G.; Croteau, R. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4126–4133.

(8) Tudzynski, B. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 597–611.

(9) Toyomasu, T.; Nakaminami, K.; Toshima, H.; Mie, T.; Watanabe, K.; Ito, H.; Matsui, H.; Mitsuhashi, W.; Sassa, T.; Oikawa, H. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 146–152.

(10) (a) Young, C.; McMillan, L.; Telfer, E.; Scott, B. *Mol. Microbiol.* **2001**, *39*, 754–764. (b) Saikia, S.; Nicholson, M. J.; Young, C.; Parker, E. J.; Scott, B. *Mycol. Res.* **2008**, *112*, 184–199.

(11) Toyomasu, T.; Tsukahara, M.; Kaneko, A.; Niida, R.; Mitsuhashi, W.; Dairi, T.; Kato, N.; Sassa, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 3084–3088.

(12) Toyomasu, T.; Niida, R.; Kanno, Y.; Kenmoku, H.; Miura, S.; Nakano, C.; Shiono, Y.; Mitsuhashi, W.; Toshima, H.; Oikawa, H.; Hoshino, T.; Dairi, T.; Kato, N.; Sassa, T. *Biosci. Biotechnol. Biochem.* **2008**, *72*, 1038–1047.

(13) (a) Zhang, S.; Monahan, B. J.; Tkacz, J. S.; Scott, B. *Appl. Environ. Microbiol.* **2004**, *70*, 6875–6883. (b) Young, C. A.; Bryant, M. K.; Christensen, M. J.; Tapper, B. A.; Bryan, G. T.; Scott, B. *Mol. Genet. Genomics* **2005**, *274*, 13–29.

(14) Hayashi, Y.; Onaka, H.; Itoh, N.; Seto, H.; Dairi, T. *Biosci. Biotechnol. Biochem.* **2007**, *71*, 3072–3081.

(15) (a) Ballio, A.; Brufani, M.; Casinovi, C. G.; Cerrini, S.; Fedeli, W.; Pellicciari, R.; Santurbano, B.; Vaciago, A. *Experientia* **1968**, *24*, 631–635. (b) Barrow, K. D.; Barton, D. H.; Chain, E.; Ohnsorge, U. F. W.; Thomas, R. *Chem. Commun.* **1968**, 1198–1200.

(16) (a) Honma, Y.; Ishii, Y.; Yamamoto-Yamaguchi, Y.; Sassa, T.; Asahi, K. *Cancer Res.* **2003**, *63*, 3659–3666. (b) Honma, Y.; Akimoto, M. *Cancer Sci.* **2007**, *98*, 1643–1651.

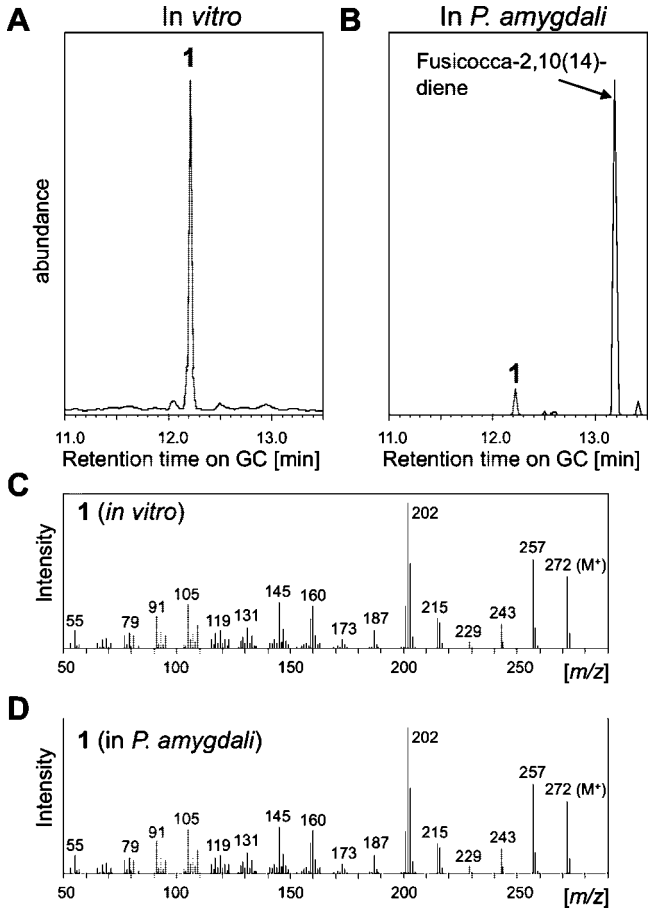
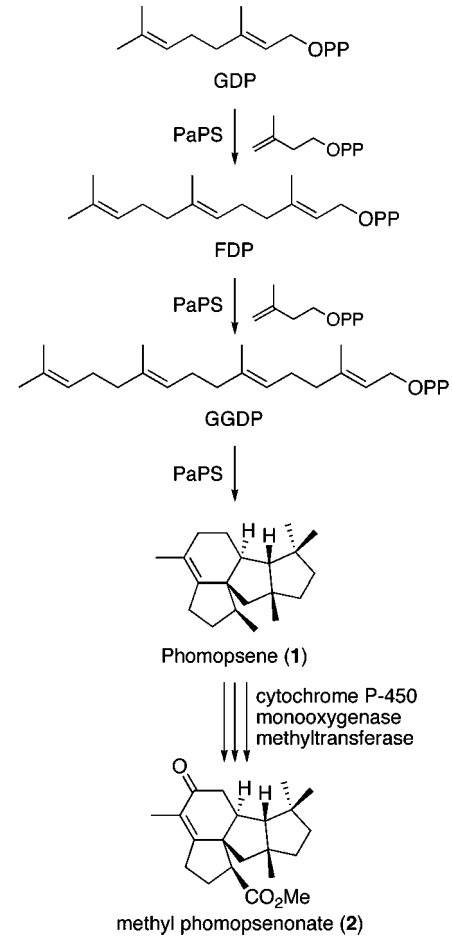


FIGURE 2. GC−MS results for products converted from GGPP by GST-PaPS. (A) Total ion chromatograms of the products after incubation of GGPP with the recombinant GST-PaPS. (B) Total ion chromatogram of the hydrocarbon fraction obtained from *P. amygdali* N2. (C) Full-scan mass spectrum of an enzymatic reaction product **1**. (D) Full-scan mass spectrum of phomopsene (**1**) from *P. amygdali* N2.

Functional Analysis of PaPS. To study the function of the PaPS, a full-length cDNA was ligated into a pGEX 4T-3 vector and was expressed in *Escherichia coli* BL21trx as the glutathione *S*-transferase (GST) fusion protein. In hexane extracts of the reaction mixture obtained by incubation of purified GST-PaPS with geranylgeranyl diphosphate (GGPP), an unidentified hydrocarbon **1** was observed at a retention time of 12.2 min on GC−MS analysis (Figure 2A and C), whereas the product was not detected from either geranyl diphosphate (GPP) or farnesyl diphosphate (FPP) (data not shown). The same product was obtained in the incubation of the enzyme with isopentenyl diphosphate (IPP) and FPP, indicating that the enzyme has both GGS and diterpene cyclase activities. In the mycelial extracts of *P. amygdali*, the diterpene hydrocarbon **1** was also found (Figure 2B and D). The structure of metabolite **1** was determined as a novel 5/6/5/5-tetracyclic diterpene hydrocarbon, named phomopsene (Scheme 1). Details of the structure determination of phomopsene are described later.

To demonstrate the function of both putative domains, N- and C-terminal truncated mutants were prepared (Figure S2A, Supporting Information). The N347 (position 1−347) mutant exhibited cyclase activity, which was estimated by the production of **1** from GGPP (Figure 1B and Figure S2B, Supporting Information), and the C398 mutant (position 328−725) catalyzed conversion of GPP or FPP, not dimethylallyl diphosphate (DMAPP), into GGPP (Figure 3). These results demonstrated that the N-terminal 40-kDa peptide catalyzed the cyclization of GGPP to **1**, and the C-terminal 44-kDa peptide catalyzed formation of GGPP.

SCHEME 1. Proposed Biosynthetic Pathway of Phomopsene (1) and Methyl Phomopsenonate (2)[a]



[a] PaPS: phomopsene synthase.

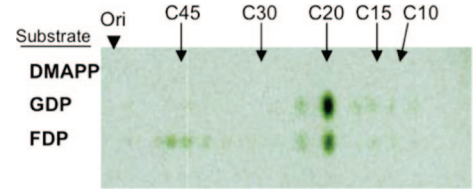


FIGURE 3. TLC autoradiography of the alcohols obtained by hydrolysis of the products derived by the recombinant C398. Ori indicates origin. DMAPP, GPP, and FPP were incubated with [^{14}C]IPP. Spots of authentic standard alcohols are indicated by arrows: C10, geraniol; C15, farnesol; C20, geranylgeraniol; C30, farnesylfarnesol; C45.

Structure of Novel Diterpene Hydrocarbon, Phomopsene. To determine the structure of the diterpene hydrocarbon phomopsene (**1**) detected in the enzymatic reaction of PaPS with GGPP, the mycelial acetone extract of *P. amygdali* N2 was partitioned with hexane−acetonitrile. The resultant hexane extract was purified by silica gel column chromatography and by reverse-phase HPLC to give **1**, which was also obtained by fermentation of the *E. coli* harboring pGEX-PaPS. The HR-EIMS of the isolated hydrocarbon showed the molecular formula of $C_{20}H_{32}$. Its ^{13}C NMR and DEPT spectra indicated five

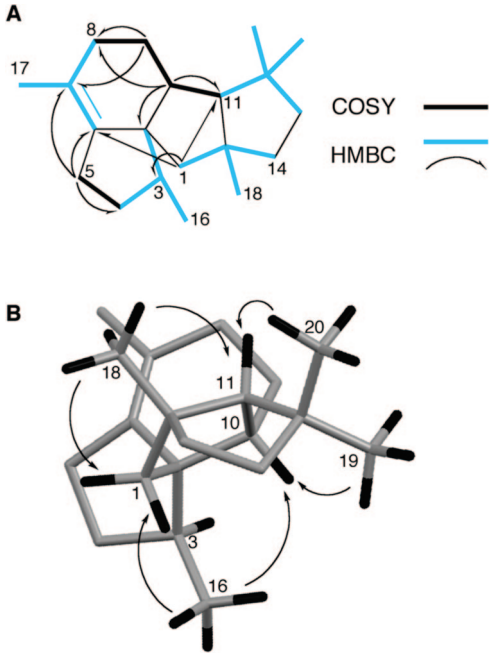


FIGURE 4. (A) Planar structure of phomopsene (**1**) deduced by $^{1}H-^{1}H$ COSY and HMBC data. (B) Stereostructure of **1** proposed by differential NOE data.

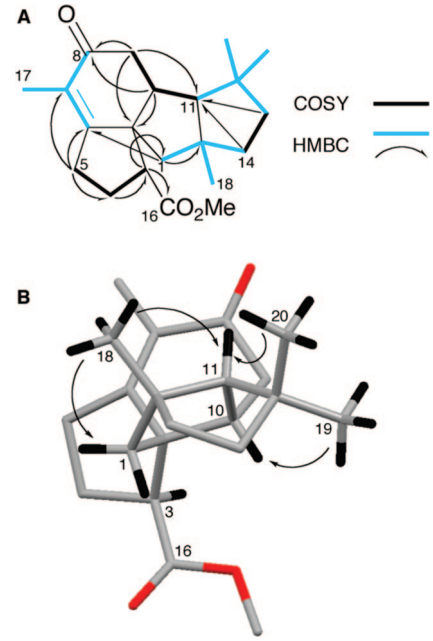


FIGURE 5. (A) Planar structure of methyl phomopsenonate (**2**) deduced by $^{1}H-^{1}H$ COSY and HMBC data. (B) Stereostructure of methyl phomopsenonate proposed by differential NOE data.

quaternary carbons including the olefin (140.7 and 123.1 ppm), three CH, seven CH_2, and five CH_3. The ^{1}H NMR spectrum showed the presence of the doublet methyl group (0.88 ppm) and the allylic methyl group (1.58 ppm). The ^{1}H and ^{13}C NMR signals in $CDCl_3$ were fully assigned by COSY, HMQC, HMBC, and NOE experiments (Figure 4). In addition to observation of COSY correlations of two spin systems H4-H5 and H8-H9-H10-H11, strong HMBC correlations between methyl protons and neighboring carbons provided three substructures as shown in Figure 4A. Additional HMBC correlations of H1, H5, and H10 unambiguously determined the 5/6/5/5-membered tetracyclic fused structure of **1** with the cyclohexene component. The apparent NOEs supported this planar structure and revealed the relative configuration of **1** (Figure 4B). The C11/C15 ring junction was assigned as cis on the basis of strong NOE correlations between H-11 and 18-Me/20-Me. The NOE correlations between 16-Me and H-1β/H-10 and between H-10 and 19-Me showed special proximity of these protons and allowed us to propose relative configurations at C2, C3, and C10.

Next, we turned our attention to the isolation of the unidentified diterpene metabolite biosynthesized by this gene cluster. The presence of modification enzymes such as a monooxygenase on the gene cluster strongly indicated that a target molecule should have polar oxygen functionality and essentially the same molecular skeleton as that of **1**, which has three singlet methyl and a doublet methyl group in its NMR spectrum. Using these characteristic NMR signals as markers, we screened oxidatively modified phomopsene diterpene in the culture broth. Although *P. amygdali* produces fusicoccin diterpene-glycosides as major metabolites,[17] characteristic signals originated from OMe and a sugar moiety allowed us to eliminate fusicoccin derivatives. Thus, unidentified diterpene **2**, structurally related to phomopsene (**1**), was isolated from the ethyl acetate extract of the fermentation broth of the phomopsene-producing strain by silica gel column chromatography and recrystallization.

The HR-EIMS and NMR data of **2** showed the molecular formula $C_{21}H_{30}O_3$. The ^{13}C NMR and DEPT spectra indicated seven quaternary carbons including the α,β-unsaturated ketone (199.22 ppm), the ester carbonyl (172.85 ppm), the olefin (162.84 and 129.56 ppm), three CH, six CH_2, and five CH_3. The ^{1}H NMR spectrum showed the methyl ester (OMe, 3.74 ppm) and the allylic methyl group (1.74 ppm). The ^{1}H and ^{13}C NMR signals in $CDCl_3$ were fully assigned by extensive analysis (COSY, HMQC, and HMBC). In addition to COSY correlations observed on three spin systems, H3-H4-H5, H9-H10-H11, and H13-H14, strong HMBC correlations between methyl protons and neighboring carbons allowed us to propose three substructures as shown in Figure 5A. Additional HMBC correlations of H1, H3, H5, and H9 unambiguously determined the tetracyclic structure of **2** as essentially identical to that of phomopsene. The observed NOEs established the configuration of **2**, which was the same as that of **1** (Figure 5B). Further confirmation of the structure of **2** was obtained by the X-ray crystallographic analysis shown in Figure 6. We named this diterpene metabolite methyl phomopsenonate.

To assign the absolute configuration of **2**, the modified Mosher's method[18] was applied to the secondary allylic alcohol **3** prepared by reduction of the C-8 ketone (Scheme 2). Treatment of **2** with $NaBH_4$ in the presence of $CeCl_3$ provided the single diastereomer **3**, which was converted to the corresponding *R*- and *S*-α-methoxy-α-trifluoromethylphenylacetic acid (MTPA) esters, **4a** and **4b**, respectively. The NOE analysis of Mosher ester **4a** allowed us to determine the stereochemistry at C-8 as shown in Figure 7. Calculated $\Delta\delta$ values were negative for signals of H-5 and H-17 and nearly zero for H-4 signals whereas $\Delta\delta$ values were positive for signals of H-9, H-10, and

(17) (a) Muromtsev, G. S.; Voblikova, V. D.; Kobrina, N. S.; Koreneva, V. M.; Krasnopolskaya, L. M.; Sadovskaya, V. L. *J. Plant Growth Regul.* **1994**, *13*, 39–49. (b) Tajima, N.; Kume, H.; Kanematsu, S.; Kato, N.; Sassa, T. *J. Pesticide Sci.* **2002**, *27*, 64–67. (c) Tajima, N.; Nukina, M.; Kato, N.; Sassa, T. *Biosci. Biotechnol. Biochem.* **2004**, *68*, 1125–1130.

(18) Ohtani, I.; Kusumi, T.; Kashman, Y.; Kakisawa, H. *J. Am. Chem. Soc.* **1991**, *113*, 4092–4096.

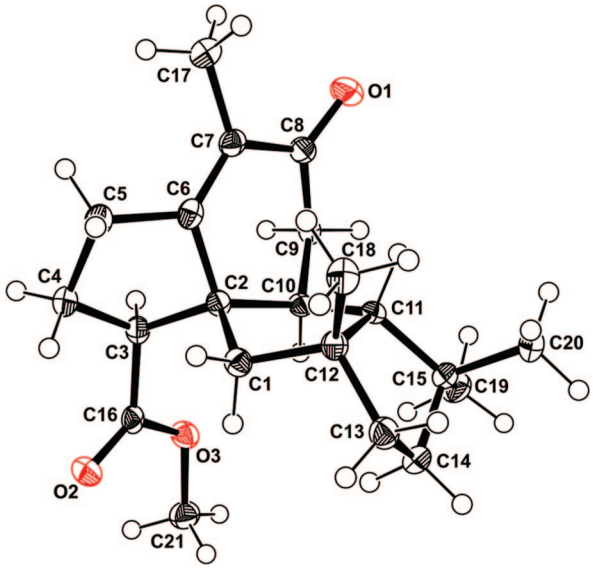
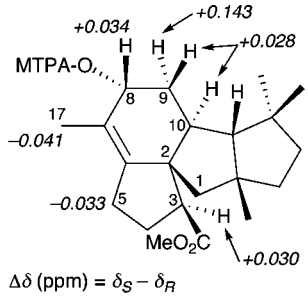
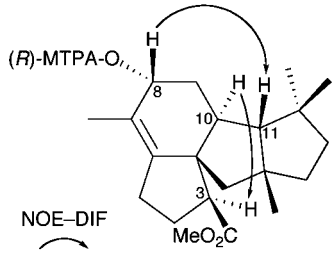


FIGURE 6. ORTEP drawing of methyl phomopsenonate (**2**).





4a

FIGURE 7. Δδ values obtained for (*R*)- and (*S*)-MTPA esters (**4a** and **4b**) and NOE correlations of **4a**.

H-3 (Figure 7). Other Δδ values were nearly zero within the experimental error for NMR measurements. On the basis of these results, the absolute configuration of **2** was proposed to be 2*R*, 3*S*, 10*S*, 11*R*, and 15*R*. Since phomopsene (**1**) is considered to be the putative biosynthetic precursor of methyl phomopsenonate (**2**), it is reasonably assumed that both metabolites have the same absolute configuration.

Discussion

On the basis of information regarding the gene cluster, the biosynthetic pathway of methyl phomopsenonate (**2**) is proposed in Scheme 1. At first, the universal precursor of diterpene GGPP is provided and is cyclized by the unusual bifunctional enzyme PaPS to give phomopsene (**1**). Since the oxidation of a methyl group to a carboxyl group is frequently catalyzed by cytochrome P-450,[8] the C-16 methyl group would be oxidized by a putative P-450 ORF3. Subsequently, oxidation of the allylic position and methylation of the carboxyl group may give **2**. Although further study is necessary to identify genes such as a monooxygenase and a methyltransferase, the predicted functions of genes on the cluster are correlated with the structure of **2**.

Two fungal metabolites, diterpene conidiogenone[19] and sesterterpene mangicol A,[20] have molecular skeletons closely related to methyl phomopsenonate (**2**) as shown in Figure 8. The biosynthetic pathway of mangicol A was proposed by feeding experiments with ^{13}C-labeled precursors.[20] However, its detailed cyclization mechanism was not elucidated. To explore a versatile biosynthetic pathway of these metabolites, we are currently investigating an enzymatic cyclization of the chimeric diterpene cyclase PaPS.

Gene-based screening of diterpenes relies on two essential factors: (1) versatility of homology-based PCR cloning and (2) functional analysis of diterpene synthase. Recently, genomic information on various fungi available from the database (http://www.broad.mit.edu/annotation/genome/aspergillus_terreus/Blast.html) allowed us to find a diterpene gene cluster based on bioinfomatics. The first factor is relatively easy to examine. For example, we can find more than five candidates, i.e., (1) AO090023000070, AO090023000073; (2) AO090012000686, AO090012000688; (3) AO090113000170, AO090113000171; (4) AO090701000221, AO090701000222; and (5) AO090038 000495 (chimera)), in *Aspergillus oryzae* in which putative diterpene synthase and GGS genes are located close to each other. Identification of six GGS genes and four diterpene gene clusters[11,12] on *P. amygdali* provided strong support for the reliability of a gene-based search of diterpene metabolites. For the second factor, diterpene synthases are relatively small proteins compared with modular enzymes such as polyketides synthases and non-ribosomal peptide synthetases. From our experiences with the handling of plant and fungal diterpene synthases, most can be successfully expressed in *E. coli*, simply using a GST-fusion protein expression system.[21] This allowed us to determine a structure of an enzymatic cyclization product.

Another factor, which we have to consider for the reverse genetic approach, is a screening marker for the diterpene metabolite. In this study, we used characteristic ^{1}H NMR signals as markers. Alternatively, analytically sensitive ^{2}H NMR signals can be used for screening of diterpenes. Since fungal diterpenes are biosynthesized via the mevalonate pathway, multiple methyl groups of diterpenes would be labeled with ^{2}H by feeding of [C^{2}H$_3$]-acetate.[22] In the ^{2}H NMR spectra of fungal crude extracts, these diagnostic methyl signals appear at the high field region (0.8−1.0 ppm) and usually do not overlap with other signals such as those of the polyketide metabolites. Thus, isolation of a target diterpene should be achieved with these NMR methods.

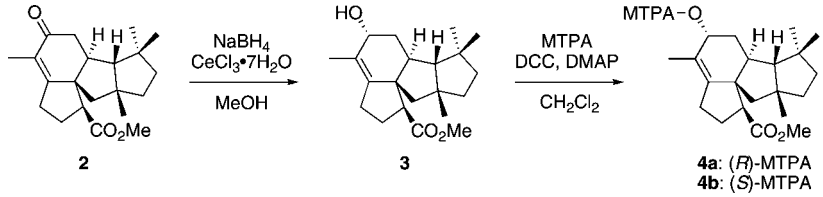
In conclusion, we have demonstrated biosynthetic gene-based screening to identify novel diterpenes. Our approach is proven to be effective for cloning the gene cluster of diterpene biosynthesis from microorganisms whose genomic information is not available. Conventionally, novel bioactive metabolites are found by bioassay-guided screening. Gene-guided screening as

(19) Roncal, T.; Cordobes, S.; Ugalde, U.; He, Y. H.; Sterner, O. *Tetrahedron Lett.* **2002**, *43*, 6799–6802.
(20) Renner, M. K.; Jensen, P. R.; Fenical, W. *J. Org. Chem.* **2000**, *65*, 4843–4852.
(21) Toyomasu, T. *Biosci. Biotechnol. Biochem.* **2008**, *72*, 1168–1175.
(22) Seo, S.; Yoshimura, Y.; Uomori, A.; Takeda, K.; Seto, H.; Ebizuka, Y.; Sankawa, U. *J. Am. Chem. Soc.* **1988**, *110*, 1740–1745.

TABLE 1. ^{1}H and ^{13}C NMR Assignments of Phomopsene (1) and Methyl Phomopsenonate (2)

	phomopsene (**1**)		methyl phomopsenonate (**2**)	
	δ_C	δ_H (multiplicity, J in Hz)	δ_C	δ_H (multiplicity, J in Hz)
1	43.08	1.51−1.42 (α-H) 1.17 (d, 12.8, β-H)	49.66	1.77 (d, 13.6, β-H) 1.51 (d, 13.6, α-H)
2	54.63		56.37	
3	39.45	1.51−1.42	46.50	2.65 (dd, 8.8, 10.6)
4	30.14	1.72 (m) 1.14 (m)	23.46	2.30 (dddd, 5.6, 10.6, 12.5, 13.4, β-H) 1.95 (dddd, 4.2, 8.8, 9.2, 13.4, α-H)
5	24.28	2.25 (ddd, 6.0, 9.5, 13.8) 2.14−2.08	25.01	2.54 (br dd, 12.5, 16.4, β-H) 2.45 (ddd, 5.6, 9.2, 16.4, α-H)
6	140.73		162.84	
7	123.13		129.56	
8	26.19	2.14−2.08 1.60 (br)	199.22	
9	24.50	1.36−1.32 1.27 (br tt, 4.4, 12.6)	39.57	2.39 (dd, 5.0, 15.2, α-H) 2.35 (dd, 2.7, 15.2, β-H)
10	39.61	2.03 (dt, 6.7, 3.4)	37.25	2.74 (ddd, 2.7, 5.0, 8.3)
11	64.28	1.51−1.42	66.17	1.33 (d, 8.3)
12	42.18		42.23	
13	40.81	1.51−1.42 1.36−1.32	39.52	1.41−1.35 1.33−1.30
14	40.63	1.51−1.42	40.23	1.46 (ddd, 6.2, 13.8, 14.2, β-H) 1.41−1.35 (α-H)
15	48.61		48.23	
16			172.85	
16-Me	14.21	0.88 (d, 6.7)		
17-Me	20.13	1.58 (s)	12.01	1.74 (d, 1.4)
18-Me	30.37	1.02 (s)	30.37	0.97 (s)
19-Me	24.98	1.01 (s)	24.49	1.03 (s)
20-Me	30.83	0.96 (s)	29.64	0.93 (s)
OMe			51.70	3.74 (s)

SCHEME 2



described in this report could become an alternative approach to identify novel secondary metabolites. When combined with a suitable heterologous expression system, this approach provides a powerful method to produce natural bioactive diterpenes that are obtained in small quantities or even not produced under normal fermentation conditions. According to a database search of fungal genomes, we found many GGS homologues in the *Aspergillus* fungi. Therefore, our reverse genetic approach would also be useful for genome mining of novel diterpenes and sesterterpenes.

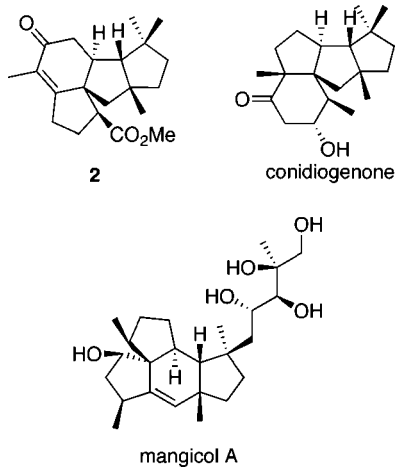


FIGURE 8. Biosynthetically related metabolites of methyl phomopsenonate (**2**).

Experimental Section

General. All commercially supplied reagents were used as received. ^{1}H and ^{13}C chemical shifts were referenced to the solvent signals ($CDCl_3$): 7.26 and 77.0 ppm. IPP, DMAPP, GPP, FPP, and GGPP were purchased from Sigma-Aldrich. [1-^{14}C]IPP (CFA476, 2.15 GBq/mmol) was obtained from Amersham Pharmacia.

Genome Walking from GGS. Genome walking was carried out using a Universal Genome Walker kit (Clontech). Genomic DNA was extracted from the 3-day-cultured mycelia of *P. amygdali* N2 by a Nucleon PhytoPure Kit (Amersham Pharmacia). A genome walker library was constructed with the genomic DNA and was used for nested PCR according to the manufacture's protocol. Primers were first designed on the basis of the cDNA fragment sequences of *PaGGS3*.[11] Gene prediction was carried out based on homology search using BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). The sequence of the identified biosynthetic gene cluster reported in this paper has been deposited in the DDBJ database with the accession code AB254159.

3′ RACE. To determine the 3′ sequences of the terpene cyclase-like gene, 3′ RACE was done using gene specific primers based on the corresponding genome DNA sequences, according to

methods described previously.[23] A full-length cDNA sequence of *PaPS* determined using RACE has been deposited in the DDBJ database with accession code AB252833.

Functional Analysis for Cyclase. To amplify the ORF cDNAs of *PaPS*, reverse transcriptase-polymerase chain reaction (RT-PCR) was carried out using a set of primers, 5′-GGATCCGTCCATCAACATTCATTCATCATG-3′ (sense, *Bam*HI site underlined) and 5′-GCGGCCGCACACGACCCGACAGATCAGG-3′ (antisense, *Not* I site underlined). For amplification of a cDNA fragment encoding N347 mutant, another reverse primer was used: 5′-GCGGCCGC**CTA**GGTAAGCTTGGTCACTTCG-3′ (antisense, *Not* I site underlined and added stop codon in boldface type). Each PCR product was digested with *Bam*HI and *Not* I and ligated into pGEX-4T-3 vector (Amersham Pharmacia) for expression as fusion protein with GST at the N-terminus. The plasmid was transformed into *E. coli* BL21trx. Procedures for growth of the *E. coli* cells, induction of the gene expression, extraction and purification of recombinant enzymes, and measurement of cyclase enzyme activity were identical with those described previously.[24] Twenty micrograms of GPP, FPP, or GGPP was used as the substrate with or without IPP. Hydrocarbon products and dephosphorylated derivatives from allylic diphosphates remained in the reaction mixture were analyzed by GC−MS.[24]

Functional Analysis for Prenyltransferase. For amplification of a cDNA fragment encoding C398 mutant, a set of primers were used: 5′-GGATCCATGATGGTTGCTCGCATGAA-3′ (sense, *Bam*HI site underlined) and 5′-GTCGACCTAAACTTTGAGTAAGCTGA (antisense, *Sal*I site underlined). PCR product was digested with *Bam*HI and *Sal*I and subcloned into pQE30 vector (QIAGEN), and the plasmid was transformed into *E. coli* BL21 (DE3). All procedures following were the same as those described previously.[25] Allylic substrates used were DMAPP, GPP, or FPP. The products derived from allylic substrates with [^{14}C]IPP were dephosphorylated and subjected to reverse-phase TLC analysis using LKC-18 developed with acetone/H_2O (9:1).

Extraction and Partial Purification of the Hydrocarbon Fraction. The *n*-hexane extract was prepared in the usual way from the samples of *P. amygdali* mycelia or *E. coli* cells. This extract was loaded into a column packed with silica gel (BW-820MH, Fuji Silysia Chemical). The column was eluted by *n*-hexane, and the eluate was evaporated and subjected to GC−MS analysis.

Isolation of Phomopsene. Fermentation of *P. amygdali* N2 was carried out at 25 °C on a reciprocal shaker for 3 days in two hundred 500-mL Sakaguchi flasks each containing 120 mL of medium of 5.0% glucose, 1.0% Pharmamedia (Southern Cotton Oil Inc.), 0.5% KH_2PO_4, and 0.1% $MgSO_4$. The mycelia were collected by filtration and extracted with acetone. The acetone layer was concentrated in vacuo, and the residual aqueous layer was adjusted to pH 9.5 with 5% aqueous Na_2CO_3. The resulting aqueous layer was extracted with ethyl acetate, and the organic layers were concentrated in vacuo. The ethyl acetate extract (11 g) was partitioned between hexane and acetonitrile. The hexane layer was concentrated in vacuo, and the residue (8 g) was purified by silica gel column chromatography (hexane) twice giving the mixture of hydrocarbon products (8 mg). This mixture was further separated by reverse-phase HPLC (Develosil ODS UG-5, ϕ 10 × 250 mm, acetonitrile) affording **1** (4 mg): colorless oil; $[\alpha]^{25}_D$ −97° (*c* 0.27, $CHCl_3$). See Table 1 for ^{1}H and ^{13}C NMR data. EI-HR-MS (positive): calcd for $C_{20}H_{32}$ (M) 272.2504, found *m*/*z* 272.2505.

Isolation of Methyl Phomopsenonate (2). The mycelial ethyl acetate extract (1.6 g) from 3.5 L of 6-day cultural broth of *P. amygdali* N2 was separated by silica gel column chromatography (hexane/ethyl acetate = 10:1). The fraction showing characteristic ^{1}H NMR methyl signals like those of phomopsene was purified by silica gel column chromatography (hexane/acetone = 20:1), and recrystallization from hexane/ethyl acetate afforded methyl phomopsenonate (5.0 mg): colorless crystal; mp 151 °C; $[\alpha]^{25}_D$ 12° (*c* 0.50, $CHCl_3$); UV λ_{max} nm (ϵ): 318 (370), 248 (17000). See Table 1 for ^{1}H and ^{13}C NMR data. IR ν_{max} cm^{-1}: 1731, 1718 (w), 1666; EI-HR-MS (positive): calcd for $C_{21}H_{30}O_3$ (M) 330.2195, found *m*/*z* 330.2199.

X-ray Analysis of Methyl Phomopsenonate (2).[26] The 0.33 × 0.33 × 0.26 mm^3 crystal was orthorhombic ($P2_12_12_1$), a = 8.9881(7), b = 13.7561(10), c = 14.6616(11) Å, V = 1812.8(2) Å^3, Z = 4, and D_c = 1.211 g/cm^3. All diffraction intensities with 2θ < 136.5° were collected at −180 °C in the ω-2θ scan mode by a Rigaku RAXIS RAPID imaging plate area detector with graphite monochromated Cu Kα radiation (λ = 1.54187 Å). Of the 8191 reflections collected, 1117 were unique, and the structure was solved by direct methods (*SIR92*).[27] The non-hydrogen atoms were refined anisotropically. Hydrogen atoms were refined isotropically. Full-matrix least-squares refinement on F converged to a final R_w factor of 0.0280 and an R factor of 0.0245 for the 18,222 observed reflections ($I > 2.00\sigma(I)$).

(*R*)-α-Methoxy-α-trifluoromethylphenylacetate 4a. To a solution of methyl phomopsenonate (**2**, 20 mg, 0.061 mmol) and $CeCl_3 \cdot 7H_2O$ (38 mg, 0.10 mmol) in methanol (0.2 mL) was added $NaBH_4$ (100 mg, 2.64 mmol) by portions at 0 °C. Acetone (0.5 mL) was added, and the mixture was extracted with ethyl acetate. The organic layers were dried over anhydrous Na_2SO_4 and concentrated in vacuo. The residue was purified by preparative TLC (hexane/ethyl acetate = 8:1) giving the allylic alcohol **3**. The product was unstable and was immediately used for the next reaction. To a solution of **3** and (*R*)-MTPA (18 mg, 0.077 mmol) in CH_2Cl_2 (0.5 mL) were added dicyclohexylcarbodiimide (62 mg, 0.30 mmol) and 4-(dimethylamino)pyridine (22 mg, 0.18 mmol). After standing at 25 °C for 12 h, the reaction mixture was filtered through Celite and concentrated in vacuo. The residue was purified by preparative TLC (hexane/ethyl acetate = 6:1) affording the (*R*)-MTPA ester **4a** (5 mg, 15% in 2 steps): colorless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.55 (2H, m), 7.42 (3H, m), 5.66 (1H, m, H-8), 3.69 (3H, s, COOMe), 3.65 (3H, s, OMe), 2.695 (1H, m, H-10), 2.443 (1H, dd, J = 8.7, 10.7 Hz, H-3), 2.288 (2H, broad, t, J = 7.4 Hz, H-5), 2.10 (1H, m, H-4), 1.992 (1H, ddd, J = 3.2, 4.9, 12.2 Hz, H-9b), 1.82 (1H, m, H-4), 1.594 (3H, s, H-17) 1.55 (1H, d, J = 7.2 Hz, H-11), 1.47−1.35 (7H, m), 1.024 (3H, s), 1.016 (3H, s), 1.010 (3H, s); ESI-HR-MS (positive): calcd for $C_{31}H_{39}F_3NaO_5$ (M + Na), 571.2647, found *m*/*z* 571.2691.

(*S*)-α-Methoxy-α-trifluoromethylphenylacetate (4b). The (*S*)-MTPA ester **4b** was obtained with (*S*)-MTPA by using the procedure just described for **4a**. **4b**: colorless oil; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.58 (2H, m), 7.41 (3H, m), 5.69 (1H, m, H-8), 3.70 (3H, s, COOMe), 3.61 (3H, s, OMe), 2.723 (1H, m, H-10), 2.473 (1H, dd, J = 8.7, 10.5 Hz, H-3), 2.255 (2H, broad, t, J = 7.2 Hz, H-5), 2.10 (1H, m, H-4), 2.020 (1H, ddd, J = 3.3, 4.9, 12.1 Hz, H-9b), 1.82 (1H, m, H-4), 1.553 (3H, s, H-17) 1.55 (1H, d, J = 7.2 Hz, H-11), 1.47−1.33 (7H, m), 1.018 (3H, s), 1.013 (6H, s); ESI-HR-MS (positive): calcd for $C_{31}H_{39}F_3NaO_5$ (M + Na), 571.2647, found *m*/*z* 571.2661.

(23) Toyomasu, T.; Kawaide, H.; Mitsuhashi, W.; Inoue, Y.; Kamiya, Y. *Plant Physiol.* **1998**, *118*, 1517–1523.

(24) Otomo, K.; Kenmoku, H.; Oikawa, H.; Konig, W. A.; Toshima, H.; Mitsuhashi, W.; Yamane, H.; Sassa, T.; Toyomasu, T. *Plant J.* **2004**, *39*, 886–893.

(25) Kawasaki, T.; Hamano, Y.; Kuzuyama, T.; Itoh, N.; Seto, H.; Dairi, T. *J. Biochem.* **2003**, *133*, 83–91.

(26) CCDC 703562 contains the supplementary crystallographic data for this paper. These data can be obtained free of charge via www.ccdc.cam.ac.uk/data_request/cif, by emailing, or by contacting The Cambridge Crystallographic Data Centre, 12, Union Road, Cambridge CB2 1EZ, U.K.; fax +44 1223 336033.

Acknowledgment. We are grateful to Dr. S. Kanematsu (National Institute of Fruit Tree Science, Japan) for providing the fungus, *P. amygdali* N2, and to Prof. M. Yotsu-Yamashita (Tohoku University, Japan) for HR-EIMS measurements. This work was supported in part by a grant-in-aid for Science Research (A) (nos. 13306009 and 16208012) from the Japanese Society for the Promotion of Science.

(27) Altomare, A.; Cascarano, G.; Giacovazzo, C.; Guagliardi, A. *J. Appl. Crystallogr.* **1993**, *26*, 343–350.

Supporting Information Available: Data on PaPS, ^{1}H and ^{13}C NMR spectra for compounds, and CIF file for methyl phomopsenonate. This material is available free of charge via the Internet at http://pubs.acs.org.

JO802319E