



# Article Identification of an Oxidosqualene Cyclase Gene Involved in Steroidal Triterpenoid Biosynthesis in *Cordyceps farinosa*

Gi-Hong An <sup>1,†</sup>, Jae-Gu Han <sup>1,†</sup>, Hye-Sung Park <sup>1</sup>, Gi-Ho Sung <sup>2</sup> and Ok-Tae Kim <sup>1,\*</sup>

- <sup>1</sup> Department of Herbal Crop Research, National Institute of Horticultural and Herbal Science, Rural Development Administration, Eumseong 27709, Korea; agiho@korea.kr (G.-H.A.); jaeguhan@gmail.com (J.-G.H.); hyesung2@korea.kr (H.-S.P.)
- <sup>2</sup> Catholic Kwandong University International St. Mary's Hospital, Incheon 22711, Korea; sung97330@gmail.com
- \* Correspondence: kimot99@hanmail.net; Tel.: +82-43-871-5730
- † These authors contributed equally to this study.

**Abstract:** Various fungi including *Cordyceps farinosa*, an entomopathogenic fungus, can produce steroidal triterpenoids. Protostadienol (protosta-17(20)Z,24-dien-3β-ol) is a precursor of steroidal triterpenoid compounds. To identify oxidosqualene cyclase (OSC) gene candidates involved in triterpenoid biosynthesis, genome mining was performed using Illumina sequencing platform. In the sequence database, two OSC genes, *CfaOSC1* and *CfaOSC2*, in the genome of *C. farinosa* were identified. Predicted amino-acid sequences of CfaOSC2 shared 66% similarities with protostadienol synthase (OSPC) of *Aspergillus fumigatus*. Phylogenetic analysis showed a clear grouping of CfaOSC2 in the OSPC clade. Function of CfaOSC2 was examined using a yeast INVSc1 heterologous expression system to endogenously synthesize 2,3-oxidosqualene. GC–MS analysis indicated that CfaOSC2 produced protosta-13(17),24-dien-3β-ol and protostadienol at a 5:95 ratio. Our results demonstrate that CfaOSC2 is a multifunctional triterpene synthase yielding a predominant protostadienol together with a minor triterpenoid. These results will facilitate a greater understanding of biosynthetic mechanisms underlying steroidal triterpenoid biosynthesis in *C. farinosa* and other fungi.

**Keywords:** *Cordyceps farinosa;* fusidic acid; oxidosqualene cyclase; protostadienol synthase; triterpenoid

## 1. Introduction

Steroidal triterpenoids such as ganoderic acid, pachymic acid, fusidic acid, and helvolic acid are produced by fungal cells. These compounds can protect fungal cells against other pathogens. Ganoderic acid and pachymic acid have anti-cancer, immuneenhancing, and anti-inflammation effects [1–3]. Fusidic acid and helvolic acid are widely used as antibiotics against Gram-positive bacteria. The biosynthesis mechanism of fusidic acid and helvolic acid has been recently elucidated using a heterologous expression system [4,5]. However, the pathway of other steroidal triterpenoid biosynthesis is not clear yet.

A number of studies have suggested that lanosterol and protostadienol (protosta-17(20)Z,24-dien-3 $\beta$ -ol) are precursors of steroidal triterpenoid compounds [6,7]. In the case of triterpene biosynthesis in fungal cells, lanosterol synthase (OSLC) and protostadienol synthase (OSPC) belonging to a class of oxisqualene cyclases (OSCs) are involved in the biosynthesis of lanostane-type and fusidane-type triterpenoid, respectively (Figure 1). It has been reported that these OSCs are key enzymes that regulate biosynthesis [8]. In both fungi and animals, OSLC is responsible for producing sterol precursors for cholesterol and ergosterol. OSLC genes from several fungal species, animals, plants, and microorganisms have been characterized [9–13]. Interestingly, genomes of some fungi encode more than one OSC. Other OSCs can also produce triterpenoid secondary metabolites via their OSC



Citation: An, G.-H.; Han, J.-G.; Park, H.-S.; Sung, G.-H.; Kim, O.-T. Identification of an Oxidosqualene Cyclase Gene Involved in Steroidal Triterpenoid Biosynthesis in *Cordyceps farinosa. Genes* **2021**, *12*, 848. https:// doi.org/10.3390/genes12060848

Academic Editor: Quan Zou

Received: 23 April 2021 Accepted: 28 May 2021 Published: 31 May 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). function. So far, only one OSPC gene has been characterized in *A. fumigatus* through a heterologous expression system in yeast cells [6]. OSPC gene from other organisms has not been characterized yet.



**Figure 1.** Putative biosynthesis pathway of steroidal triterpenoids in fungi. Dot arrows mean multiple steps.

The species *Cordyceps farinosa* belongs to the Cordycipitaceae family in the Hypocreales order of Ascomycota and is well-known as entomopathogenic fungi with a worldwide distribution in temporate and tropical zones [14]. These entomopathogenic fungi can also produce a variety of chemicals that cause severe adverse reactions in other organisms, including bacteria and other animals. *Sarocladium oryzae, Acremonium fusidioides, Metarhizium anisopliae*, and *Aspergillus fumigatus* are fungi that can produce fusidanes as steroidal antibiotics [15]. In particular, *Cordyceps kogane* can produce fusidic acid that is structurally similar to fusidane [16].

To discover genes related to triterpenoid biosynthesis, *Cordyceps farinosa* mycelium was chosen as a biosynthesis model of steroidal triterpenoid. We chose *C. farinosa* because, in the preliminary comparative genome analysis of Hypocrealean species, we recognized that a gene cluster of *C. farinosa* was similar to a well-known gene cluster related to helvolic acid biosynthesis in *A. fumigatus*. In the present study, whole genome sequences of *C. farinosa* KMCC47486 were analyzed and two OSC genes, lanosterol synthase and protostadienol synthase, were found. We isolated a gene encoding a putative protostadienol synthase from *C. farinosa* mycelium. Its function was then determined using a heterologous expression system in yeast along with product analysis using GC-MS. In addition, a gene cluster including OSPC gene was identified. It consisted of seven genes highly conserved in *A. fumigauts* for helvolic acid biosynthesis. We also discussed triterpenoids produced by discovering genes that matched with the gene cluster related to helvolic acid biosynthesis.

#### 2. Materials and Methods

# 2.1. Strain of C. farinosa

A strain of *C. farinosa* (KACC47486) was obtained from Korean Agricultural Culture Collection (KACC), Rural Development Administration (RDA), Republic of Korea. The

mycelium of *C. farinosa* was cultured on PDA (Potato Dextrose Agar) medium at 25 °C in darkness for 2 weeks. To identify the strain through sequence comparison of its ITS rDNA region, phylogenetic tree analysis was performed using the Neighbor-Joining method.

## 2.2. Genome Sequencing and Assembly

Genomic DNA was extracted from freeze-dried mycelium using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions. Concentration and purity of the extracted genomic DNA were assessed using a NanoDrop ND-1000 spectrophotometer (Thermo–Fisher Scientific, Waltham, MA, USA). Next generation sequencing (NGS) library construction and sequencing were carried out by Macrogen Inc. (Seoul, Korea). Paired-end and mate-pair libraries with insert sizes of 300 bp and 5 kbp, respectively, were prepared. Subsequently, NGS was carried out using an Illumina HiSeq platform (KAPA Biosystems, Roche, San Francisco, CA, USA) with a coverage of approximately 350×. Generated sequence reads with low-quality scores or short lengths were filtered out using a FASTX Toolkit 0.0.13 (http://hannonlab.cshl.edu/fastx\_toolkit (accessed on 28 May 2021)). Trimmed paired-end and mate-pair reads were then assembled using a Velvet 1.2.01 assembler [17]. All reads were deposited at NCBI, and can be accessed in the BioProject data base under project accession number PRJNA314175.

## 2.3. Phylogenetic Analysis

Amino acid sequences of OSCs in C. farinosa and its related species were extracted using tblastn of NCBI BLAST 2.2.28+ package [18] with a query sequence of A. fumigatus OSLC (XP\_747936) against selected fungal genomes belonging to the order Hypocreales: Beauveria bassiana ARSEF2860 (PRJNA38719), Claviceps purpurea 20.1 (PRJEA76493), Cordyceps militaris CM01 (PRJNA41129), Epichloë festucae FI1 (PRJNA51625), Epichloë typhina E8 (PRJNA174036), Fusarium graminearum PH-1 (PRJNA243), Fusarium oxysporum Fo5176 (PRJNA68027), Fusarium verticillioides 7600 (PRJNA15553), Metarhizium acridum CQMa 102 (PRJNA38715), Metarhizium anisopliae ARSEF23 (PRJNA38717), Tolypocladium inflatum NRRL8044 (PRJNA73163), Trichoderma atroviride IMI206040 (PRJNA19867), Trichoderma reesei QM6a (PRJNA15571), Trichoderma virens Gv29-8 (PRJNA19983), Beauveria pseudobassiana KACC47484 (PRJNA314175), Beauveria sungii KACC47481 (PRJNA314175), Cordyceps pruinosa KACC44470 (PRJNA314175), Cordyceps farinosa KACC47486 (PRJNA314175), and Cordyceps tenuipes KACC47485 (PRJNA314175). Predicted amino acid sequences were aligned using Clustal Omega [19] with default settings and then manually adjusted. Phylogenies were inferred using neighbor joining (NJ) and maximum likelihood (ML) methods of a MEGA6 software [20]. Relative robustness of individual branches was estimated with 500 replicates using bootstrapping (BS).

#### 2.4. Generation of Plasmid Vectors

Sequences for two *C. farinosa* genes were deposited in GenBank with accession numbers of MF972281 for CfaOSC1 and MF972287 for CfaOSC2. To amplify the CfaOSC2 gene, forward (5'-TGATGCCTGTCGCCGATATTGAC-3') and reverse (5'-TTATCCTTTGTA AAAGGACGCTC-3') primers were designed. PCR was then performed with an Ex-taq DNA polymerase (Takara, Tokyo, Japan) using the following cycling conditions: 30 cycles at 95 °C for 1 min, 58 °C for 1 min, and 72 °C for 2 min; followed by a final extension step at 72 °C for 10 min. The obtained fragment was purified and cloned into a pYES2.1 vector using a TOPO TA-expression kit (Invitrogen, Carlsbad, CA, USA), enabling the construction of an expression plasmid pYES-CfaOSC2 in yeast cells. After sequence (pYES-CfaOSC2) confirmation, the plasmid was used for yeast transformation.

## 2.5. Functional Expression of CfaOSC2 in Yeast

Functional characterization was carried out in yeast strain INVSc1 purchased from Invitrogen (Carlsbad, CA, USA). Yeast transformation and insert-DNA overexpression were carried out as described by Kushiro et al. [21]. Single clones including pYES–CfaOSC2 were incubated in 15 mL of a synthetic complete medium containing 2% glucose without uracil at 30 °C with shaking at 220 rpm for two days. After three-day induction with 2% galactose, cells from two 50 mL conical tubes were collected into one tube, refluxed with 8 mL of 20% potassium hydroxide (KOH) and 50% ethanol (EtOH), and extracted three times with hexane at the same volume. Extracts were concentrated under a stream of nitrogen gas (N<sub>2</sub>) and resuspended in 1 mL of chloroform (CHCl<sub>3</sub>). Subsequently, 100 µL of the extract was transferred into a new vial and concentrated under a stream of N<sub>2</sub> at 70 °C. For GC–MS analysis, the extract was silylated with 50 µL of *N*,*O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) and 50 µL of pyridine for 30 min at 70 °C.

## 2.6. GC-MS Analysis

GC–MS analyses were run under the same conditions as described by Wang et al. [22]. Briefly, a 1- $\mu$ L aliquot of the solution was analyzed using a 7890N gas chromatography (Agilent Technologies, Santa Clara, CA, USA) equipped with a 5973-inert mass spectrometer (MS) detector (Agilent Technologies, Santa Clara, CA, USA) and an Agilent HP-5 capillary column (length of 30 m, i.d. of 250  $\mu$ m, film thickness of 0.25  $\mu$ m). The injection temperature was set at 50 °C for 2 min. The column-temperature program was as follows: 40 °C/min ramp to 200 °C, a hold at 200 °C for 2 min, followed by an increase to 320 °C at a rate of 3 °C/min; and lastly a hold at 320 °C for 30 min. Triterpenoids were identified by comparison with authentic extracts from yeast cells co-expressing Afu-OSC3 (*A. fumigatus* protostadienol synthase) and CYP5081A1 (*A. fumigatus* cytochrome P450 oxidase) as reported by Mitsuguchi et al. [7]. The constructed vector of pESC(-Ura)-AfuOSC3/CYP5081A1 was kindly provided by Dr. Tetsuo Kushiro (Meiji University, Tokyo, Japan). As previously described, the same protocol was applied for the transformation and culturing of yeast strain INVSc1 with the vector.

## 2.7. Synteny Analysis

The OSC gene cluster of *A. fumigatus* Af293 proposed by Lodeiro et al. [6] was retrieved from GenBank (XP\_751348-XP\_751356), including a series of genes catalyzing monooxy-genation, dehydrogenation, and acyl transfer to convert protostadienol into helvolic acid. Based on the result of our phylogenetic analysis, the synteny of OSC gene cluster of *A. fu-migatus* was searched within the three selected hypocrealean species., viz. *M. anisopliae, B. sungii*, and *C. farinosa* (Group III in Figure 4). Orthologous regions were identified using the reciprocal best hit (RBH) approach [23]. First, local databases were created from genome sequences of *C. farinosa*. After BLAST-searching *AfuOSC3* gene clusters in databases, nucleotide sequences highly matched with protein sequences of AfuOSC3 were extracted from each genome. They were then conversely queried against the *AfuOSC3* gene cluster. In each pair-wise comparison, reciprocally best-matched genes (or regions) were regarded as orthologs.

#### 3. Results and Discussion

## 3.1. Identification of C. farinosa

The mycelium of putative *C. farinosa*, which was provdied by the KACC (Korean Agriculture Culture Collection), was cultured on PDA medium for 2 weeks (Figure 2a). To identify the strain, the fungal source was analyzed by ITS gene sequencing. The BLAST search of the ITS sequence of the strain showed the highest similarity to *Cordyceps* and the phylogenetic tree was constructed using MEGA 5.2 software. As shown in Figure 2b, this strain clearly was grouped with *C. farinosa*. In addition, *Isaria farinosa* as the scientific name has long been used for this fungus. However, the genetic position of *Isaria* was changed to the *Cordyceps* by the principal of priority because of the discontinuance of dual nomenclature for pleomorphic fungi in 2011 [24]. After confirmation of the strain, we performed genome sequencing.



**Figure 2.** Identification of *Cordyceps farinosa* (KACC47486). (a) Strain of *C. farinosa* (KACC47486) cultured on PDA medium for two weeks; (b) Phylogenetic analysis using Neighbor-Joining method based on partial sequences of the ITS rDNA region.

## 3.2. Sequence Features of Two OSC Genes

Mass DNA-sequence data were produced using an Illumina sequencing platform to identify OSC genes involved in triterpenoid biosynthesis. In the sequence database, two *OSC* genes, *CfaOSC1* (*C. farinosa* lanosterol synthase) and *CfaOSC2* (*C. farinosa* protostadienol synthase), were identified from the genome of *C. farinosa*. Both genes contained full-length cDNAs. Open reading frames (ORFs) of *CfaOSC1* and *CfaOSC2* were 2223 bp and 2205 bp in length, respectively. These ORFs were predicted to encode CfaOSC1 and CfaOSC2 proteins having 741 and 735 amino acids with masses of 84.407 kDa and 82.615 kDa, respectively. Deduced amino acid sequences of CfaOSC1 (70%) and CfaOSC2 (66%) were orthologous to lanosterol synthase (Afu5g04080) and protostadienol synthase (AfuOSC3) of *A. fumigatus*, respectively. As shown in Figure 3, CfaOSC1 and CfaOSC2 contained repeats of the QW motif [25], a typical feature of the triterpene-synthase superfamily. CfaOSC2 consisted of <sup>699</sup>ACPGGMR<sup>705</sup> motif in the C-terminal region. This motif is known to play a role in protostadienol formation [26].

To determine the relationship of these two CfaOSCs with other fungal OSCs, phylogenetic analysis was performed as to determine its relationships with other fungi (Figure 4). As expected, CfaOSC1 and CfaOSC2 were clearly grouped in OSLC and OSPC clades, respectively. Therefore, CfaOSC2 is likely to have function in protostadienol production.

		20		40		60		80		100	
Afu5g04080	M TGGP I A	SWR TAAQG	HLTPDENGDL	KTDYSRWRLL	DDRGRQTWHY	LESDEENEKW	PQSVADKHFL	GLPTVTPLAP	AATNRMFAIH	ELPELPKAKT	95
CfaOSC1	MATATGRETA	SARNRTRDSN	HGTQKPKIEG	KTDYSRWRLR	DDDSRHTWHY	LTDDTELAKW	PQSLAEKYFM	NLPL		GLPDLPKATT	84
CfaOSC2	MPVADID	MVTGSGG	HVNRKAASEY	STDLTRWRLK	VHEGRHMWEY	VEA-DVAATQ	HQTFAEKYWL	GQD	Y	ELPKMSTPTC	77
ATUOSC3	M-AIDSS	MPGTVTGKAE	FSDIKAASEF	GIDLSRWRLN	V DNGRHMWE Y	LESEDEARKR	POSFLEKYWL	GLP	Y	ELPARPRATC	80
		120		140		100		180		200	
Afu5g04080	PLQCAENGLE	FFSKLQLPPG	NWACEYGGPM	FLLPGLIITY	YVTNTPIPPE	YATEIKRYLF	ARQHPEDGGW	GLHIEAHSSV	FGTCMNYVAL	RLIGVSEDDP	195
CfaOSC1	ALDAANNGID	LFEKLQMPSG	HWACEYGGPM	FLLPCVVCAW	YVTKTPIPEA	YATETKNYLT	ARAHPEDGGW	GLHIEGESTV	FGTSLNYTVL	RLVGVDPEDP	184
460803	ALEAVENGWE	FEKRIOTADO	HWGCNDDGPL	EVISCIVIAS	VIVGIPIDSH	MKOEMCRYLL	NVVN-EDGGW	GLEIOSPSTV	FGTVMNYCML	RILGIOPEHP	170
A00003	ALEATENONE	220	in ponoboli L	240	11101110011	260	NUM EDOON	280	1 OT VINITY OWE	300	
		1		1		1		I I		EL VEODVOOL	005
Atu5g04080	MIKARGELH	KEGGATYGPH	WAKFWLSVLG	VMEWECVNPV	PPELWLLPDW	VPFIPWRWWI	HIRQVELPMS	TLWSKKFTHP	EDDLIBCIRG	ELFTOPTOST	295
CfaOSC2	VIIDARAALC	RMGSARALPT	WGKEWMCALG	AYEWSGMIPI	IPEPILAPGE	L P I NPGNWWV	HTRNVYVSMS	YLEGHRETAO	ITPI VOFI RE	ELYDMPYHEI	276
AfuOSC3	AMAKARNTLH	RLGSARATPT	WGKFWLCVLG	VYEWEGMVPL	PPEPLLVPAS	LPFNPGKWWV	HTRNVYISMS	YLYGHRFSMP	PNKLVQALRD	ELYDIPYQQI	279
		320		340		360		380		400	
Afu5c04080	SEANHRNSTH	AADNYYPKTW		SVWNPYERIP	AL VKRAFEWT	WELLBMEDEN	TDYAGI GPVS	NPMNMVACYL	HDGPDSYSVR		395
CfaOSC1	NWASHRNSIA	PEDNYHPKSW	LLNTVNWALV	NVYNPYIRPN	FIKNKAEAWA	SQLVDMEDAN	TDYADLAPVN	APLNTVVCYI	RDGPAAFSVK	RHIERLEEFL	384
CfaOSC2	DWFAQRTNIS	DNDRLFPPTL	LQKGLVNALC	SYEYIKVP	FLRRKALDEA	LFQVEMEVRN	TSYLCIAPVN	FASNMLVMLH	AHGSNSHWVK	GMADRIIDPM	374
AfuOSC3	NWPAQRTNVS	AADRLTDPTW	IQRSFTSALT	MYETFKIP	FLRRRALNEA	LFQIETETRN	THYLCIAPVS	FASNMLALYH	AHGRDSHWIR	GMRDRFIDPM	377
		420		440		460		480		500	
Afu5g04080	WMKNEGMLMN	GTNGVQVWDT	AFITQAIVVA	GFADDPKWRP	MLTKALEFLE	DHQLRENVPD	QEKCYRQHRK	GAWPFSNKTQ	GYTVSDCTAE	GLRSTIQLQE	495
CfaOSC1	WVKDEGMLVN	GTNGVQCWDT	AFLIQAVVEA	GLHTDPRWRP	MLLKALGYLD	RQQIRDNCVD	QEKCYRQPRK	GGWPFSNRDQ	GYGVSDCISE	ALKATILLQK	484
ClaUSC2	WICREGMAAS	GTNGTS WDT	ALTVOATIDA	GLAERPENHD	INCOALKFIN	NSOLPEDRIC	VUUVVDOPTR	CAMPESTROQ	ATAVSDITAE SVAVSDITAE	AVRAVIQUA	4/4
A00303	WECKEGEARS	620	ALIVEATION	SEARCE ENGA	TERRALEFID	NO GTINE DF EG	VIIIVIII	580	STAVSDITAL	800	4//
		T. T		I		T		T. T. T.		T	
Atu5g04080	MHNYPRLISV	ORLEDAVDTL	LLMQNPSGGF	TEYETTRGSE	KLEWLNAAEV	FGGIMIGYDY	PECITASVIA	LSLFSRFYPD	VRADEIKAAK	DRAVKYIKRV	595
CfaOSC2		ERLIEAVDIL	LGMENKCGGY	SAFEPIRAPK	ALELINVIEL	VENVMTESLY	PECTSSVILC	LETETKAYPH	HRPMEIOATI	GROVRYLIKA	574
AfuOSC3	IEGFPSRISD	ERLQQAIDLI	LGMENAGGGF	SAYEPVRGPK	FLELLNITEL	YENVMTDNLY	PECTSSVIMC	LTTFAREYPT	YRPRDIQACL	SRSIDYLLRS	577
		620		640		660		680		700	
Afu5q04080	QRPDGSWYGS	WGICFTYAAM	FALESLASVG	ETYETSEYAR	RGCEFLLSKQ	KEDGGWGESY	LSSEKHVYVQ	HEK-SQVVQT	AWACLALMEA	EYPHKEPLQK	694
CfaOSC1	QRIDGSWYGS	WGICFTYATM	FALESMASVG	ETYGNSYSAS	KGCEFLLSKQ	RADGGWSESY	KACETMTYVE	HPMGSLVVQT	AWALIGLMEA	EYPHVEPLRK	684
CfaOSC2	QFPCGGW/AS	WGVCFTYATM	FALQGLETVG	LRESNSETCR	DACSFLLQYQ	NEDGGWGEDL	ISIREKHYVQ	DPAGSQVTCT	AYALTGLIAA	HCSNRDALRR	674
AfuOSC3	QYPNGGWFAS	WGVCFTYATM	FALQGLACMG	WNESNCAACQ	RACSFLLQHQ	NPDGGWGESL	DTVRFKQYLP	HPDGSQVTNT	AYAVIGLLAA	RCGNHEAIRR	677
		720		740		760					
Afu5g04080	AMKLLMSRQQ	PNGEW_QES I	EGVFNQSCMI	SYPNYKFYWP	IRALGLYSRK	FGNEELM	751				
CfaOSC1	GVQLIMSRQQ	ANGEWLQEAI	EGVFNKSCMI	SYPNYKFTFT	IKALGMFAKR	FPDEQF	740				
CtaOSC2	GTHWLMSHQQ	ATGEWLPGSL	EGIFACPGGM	PRENE	LAAIGKYIER	HONECLAL	734				

**Figure 3.** Alignment of amino-acid sequences of OSC genes in *Cordyceps farinosa* and *Aspergillus fumigatus*. Blue and red boxes indicate the QW(QXXXGXW) motif and the specific motif that plays a role in protostadienol production, respectively.





## 3.3. Functional Characterization of CfaOSC2 in Yeast

Full-length cDNA of *CfaOSC2* was amplified using polymerase chain reaction (PCR) to elucidate the function of *CfaOSC2* gene in steroidal triterpenoid biosynthesis. The full-length cDNA was cloned into a yeast expression vector pYES2.1 under the control of a

GAL1 promotor. The CfaOSC2 gene functionally was characterized using a heterologous expression system in yeast INVSc1 that could endogenously synthesize 2,3-oxidosqualene. After three days of galactose induction, yeast cells were harvested and then extracted with hexane. To check whether or not alcohol triterpenes were extracted from yeast cells, we added lupeol as a internal standard before the extraction. Results of gas chromatographymass spectrometry (GC-MS) analysis indicated that CfaOSC2-overexpressing yeast cells produced two products not present in control yeast cells carrying an empty vector pYES2 (Figure 5a). Peaks of yeast cells co-expressing AfuOSC3/CYP5081A1 as described by Mitsuguchi et al. [7] were used as authentic standards. Results of retention-time comparison indicated that these two products at a ratio of 5:95 were protosta-13(17),24-dien-3 $\beta$ -ol (1) and protosta-17(20)Z,24-dien-3β-ol (2, protostadienol), respectively. In addition, the pattern of mass fragments of these identified compounds in yeast cells overexpressing AfuOSC3/CYP5081A1 was compared. Results are shown in Figure 5b. Both peaks were alcohol triterpenes as they comprised the m/z 498 molecular ion which was trimethylsilylated (TMS). In addition, the characteristic fragment ions in MS spectrum of protostadienol were m/z = 191, 339 and 429, and these values were consistent with those reported by Ledro et al. [6] (Supplementary Figure S1). Therefore, CfaOSC2 is a multifunctional triterpene synthase producing compounds 1 and 2.



Figure 5. Cont.



**Figure 5.** Identification of triterpenoids in CfaOSC2-overexpressing yeast cells by gas chromatography–mass spectrometry (GC–MS) analysis. (**a**) GC chromatograms from the strain expressing the *CfaOSC2* gene, the control yeast strain harboring the pYES2 vector, and the authentic standards from the yeast strain co-expressing *AfuOSC3/CYP5081A1* genes. IS means internal standard (Lupeol). (**b**) MS spectra of trimethylsilylated triterpenoids produced from the CfaOSC2-expressing strain (left) and the standard (right).

## 3.4. Synteny Comparison of a Gene Cluster for Helvolic Acid Biosynthesis

Whole fungi-genome sequences are available from public databases. Lodeiro et al. [6] and Mitsuguchi et al. [7] have identified a gene cluster related to helvolic acid biosynthesis that consists of OSPC, four cytochrome P450 (CYP) monooxygenases, two transferasefamily proteins, and two dehydrogenase genes. Nine of these gene-cluster genes have been recently characterized using a heterologous-expression system in Aspergillus oryzae NSAR1, resulting in the detection of helvolic acid together with 21 protostadienol derivatives [4]. We compared the gene cluster of A. *fumigatus* for helvolic acid biosynthesis with that of C. farinosa based on DNA sequences (Figure 6). A gene cluster of C. farinosa consists of seven genes, encoding one protostadienol synthase, three cytochrome P450 (CYP), one acyltransferase, one short-chain dehydrogenase/reductase and one ketosteroid dehydrogenase. Among seven genes, amino acids of five genes are highly similar to those of HelA (65%), HelB1 (64%), HelB2 (60%), HelC (61%) and HelB4 (66%). However, amino acids of two genes encoding acyltransferase and ketosteroid dehydrogenase showed the low level of similarity to those of HelD2 (38%) and HelE (26%). The gene cluster of C. farinosa showed deletion of a CYP gene (Gene No. 5) and an acyltransferase gene (Gene No. 6) compared to the helvolic acid cluster. We performed HPLC analysis to determine helvolic acid in C. farinosa (Supplementary Figure S2). The result indicated that a peak which was not consistent with that of helvolic acid was detected. Therefore, the result might indicate that helvolic acid is not biosynthesized because of deletion of the two genes in C. farinosa. Based on these results, we depicted a putative pathway of triterpenoid biosynthesis in C. farinosa (Figure 7). We inferred that the final product of operating this gene cluster might be 16β-acetyloxy-29-norprotosta-1,17(20)Z,24-trien-3-one-21-oic acid considering that those two genes (Gene No. 5 and Gene No. 6) were deleted in the gene cluster. Future studies are needed to identify all genes except OSPC gene using a heterologous expression system.



No.	Gene name	Description (Accession No.)
1	HelA	Protostadienol synthase (XP_751356.1)
2	HelB1	Cytochrome P450 monooxygenase (XP_751355.1)
3	HelB2	Cytochrome P450 monooxygenase (XP_751354.1)
4	HelC	Short chain dehydrogenase (XP_751353.1)
5	HelB3	Cytochrome P450 monooxygenase (XP_751352.2)
6	HelD1	Transferase family protein (XP_751351.1)
7	HelB4	Cytochrome P450 monooxygenase (XP_751350.1)
8	HelD2	Transferase family protein (XP_751349.1)
9	HelE	Extracellular 3-ketosteroid 1-dehydrogenase (XP_751348.1)

Figure 6. Synteny comparison of two gene clusters involved in triterpenoid biosynthesis.



Figure 7. Putative pathways of triterpenoid biosynthesis in C. farinosa.

## 4. Conclusions

This study revealed that the *CfaOSC2* gene encoded an enzyme that could catalyze the transformation of oxidosqualene to a predominant protostadienol along with a minor triterpenoid and that *CfaOSC2* was associated with protostane-type triterpenoid biosynthesis using a heterologous expression system in yeast cells. However, the next catalytic step of the cycling triterpenoid remains unclear. Based on our DNA-sequence database, future studies need to elucidate the function of candidate genes in the gene cluster associated with *CfaOSC2*. In addition, it will be necessary to identify 16 $\beta$ -acetyloxy-29-norprotosta-1,17(20)Z,24-trien-3-one-21-oic acid or any triterpene compound in *C. farinosa* cells. Results of this study provide a greater understanding about biosynthetic mechanisms underlying steroidal triterpenoid biosynthesis in *C. farinosa* and other fungi.

**Supplementary Materials:** The following are available online at https://www.mdpi.com/article/10 .3390/genes12060848/s1, Figure S1: Putative fragmentation mechanism of protostadienol leading to the characteristic fragment ions observed in the GC-MS spectrum, Figure S2: HPLC analysis of helvolic acid. Two chromatograms of HPLC analysis, (a) an authentic helvolic acid and (b) the extract from C. farinosa mycelium.

**Author Contributions:** G.-H.A. carried out yeast experiments and GC-MS analysis. J.-G.H., H.-S.P. and G.-H.S. participated in DNA-sequencing analysis. O.-T.K. planned the work and wrote the manuscript. All authors have read and approved the final version of the manuscript.

**Funding:** This work was supported by a grant (PJ01194901) from the Next-Generation BioGreen 21 Program funded by Rural Development Administration, Republic of Korea.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

#### References

- Radwan, F.F.Y.; Hossain, A.; God, J.M.; Leaphart, N.; Elvington, M.; Nagarkatti, M.; Tomlinson, S.; Haque, A. Reduction of myeloid-derived suppressor cells and lymphoma growth by a natural triterpenoid. *J. Cell. Biochem.* 2015, 116, 102–114. [CrossRef] [PubMed]
- Sun, K.X.; Xia, H.W. Pachymic acid inhibits growth and induces cell cycle arrest and apoptosis in gastric cancer SGC-7901 cells. Oncol. Lett. 2018, 16, 2517–2524. [CrossRef]
- 3. Zhao, S.; Gao, Q.; Rong, C.; Wang, S.; Zhao, Z.; Liu, Y.; Xu, J. Immunomodulatory effects of edible and medicinal mushrooms and their bioactive immunoregulatory products. *J. Fungi* **2020**, *6*, 269. [CrossRef] [PubMed]
- 4. Lv, J.M.; Hu, D.; Gao, H.; Kushiro, T.; Awakawa, T.; Chen, G.D.; Wang, C.X.; Abe, I.; Yao, X.S. Biosynthesis of helvolic acid and identification of an unusual C-4-demethylation process distinct from sterol biosynthesis. *Nat. Commun.* **2017**, *8*, 1644. [CrossRef]
- Cao, Z.; Li, S.; Lv, J.; Gao, H.; Chen, G.; Awakawa, T.; Abe, I.; Yao, X.; Hu, D. Biosynthesis of clinically used antibiotic fusidic acid and identification of two short-chain dehydrogenase/reductases with converse stereoselectivity. *Acta Pharm. Sin. B* 2019, *9*, 433–442. [CrossRef] [PubMed]
- 6. Lodro, S.; Xiong, Q.; Wilson, W.K.; Ivanova, Y.; Smith, M.L.; May, G.S.; Matsuda, S.P. Protostadienol biosynthesis and metabolism in the pathogenic fungus *Aspergillus fumigatus*. *Org. Lett.* **2009**, *11*, 1241–1244. [CrossRef] [PubMed]
- Mitsuguchi, H.; Seshime, Y.; Fujii, I.; Shibuya, M.; Ebizuka, Y.; Kushiro, T. Biosynthesis of steroidal antibiotic fusidanes: Functional analysis of oxidosqualene cyclase and subsequent tailoring enzymes from *Aspergillus fumigatus*. J. Am. Chem. Soc. 2009, 131, 6402–6411. [CrossRef]
- 8. Fei, Y.; Zhang, D.H.; Xu, J.W. Increased production of ganoderic acids by overexpression of homologous farnesyl diphosphate synthase and kinetic modeling of ganoderic acid production in *Ganoderma lucidum*. *Microb. Cell Factories* **2019**, *18*, 115. [CrossRef]
- Abe, I.; Prestwich, G.D. Molecular cloning, characterization, and functional expression of rat oxidosqualene cyclase cDNA. *Proc. Natl. Acad. Sci. USA* 1995, 92, 9274–9278. [CrossRef]
- 10. Joffrion, T.M.; Collins, M.S.; Sesterhenn, T.; Cushion, M.T. Functional characterization and localization of *Pneumocystis carinii* lanosterol synthase. *Eukaryot. Cell* **2010**, *9*, 107–115. [CrossRef] [PubMed]
- Shang, C.H.; Shi, L.; Ren, A.; Qin, L.; Zhao, M.W. Molecular cloning, characterization, and differential expression of a lanosterol synthase gene from *Ganoderma lucidum*. *Biosci. Biotechnol. Biochem.* 2010, 74, 974–978. [CrossRef] [PubMed]
- 12. Shi, Z.; Buntel, C.J.; Griffin, J.H. Isolation and characterization of the gene encoding 2,3-oxidosqualene-lanosterol cyclase from *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7370–7374. [CrossRef]
- 13. Suzuki, M.; Xiang, T.; Ohyama, K.; Seki, H.; Saito, K.; Muranaka, T.; Hayashi, H.; Katsube, Y.; Kushiro, T.; Shibuya, M.; et al. Lanosterol synthase in dicotyledonous plants. *Plant Cell Physiol.* **2006**, *47*, 565–571. [CrossRef] [PubMed]
- 14. Du, C.; Wu, J.; Cuthbertson, A.G.; Bashir, M.H.; Sun, T.; Ali, S. Morphological, molecular and virulence characterisation of six *Cordyceps* spp. isolates infecting the diamondback moth, *Pluttela xylostella*. *Biocontrol Sci. Technol.* **2021**, *31*, 373–386. [CrossRef]
- 15. Zhao, M.; Godecke, T.; Gunn, J.; Duan, J.A.; Che, C.T. Protostane and fusidane triterpenes: A Mini-Review. *Molecules* **2013**, *18*, 4054–4080. [CrossRef]
- 16. Hikino, H.; Asada, Y.; Arihara, S. Fusidic acid, a steroidal antibiotic from *Isaria kogane*. *Chem. Pharm. Bull.* **1972**, *20*, 1067–1069. [CrossRef]
- 17. Zerbino, D.R.; Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 2008, *18*, 821–829. [CrossRef]

- Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. J. Mol. Biol. 1990, 215, 403–410.
  [CrossRef]
- Sievers, F.; Wilm, A.; Dineen, D.G.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011, 7, 539. [CrossRef]
- 20. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef] [PubMed]
- 21. Kushiro, T.; Shibuya, M.; Ebizuka, Y. β-amyrin synthase-Cloning of oxidosqualene cyclase that catalyzes the formation of the most popular triterpene among higher plants. *Eur. J. Biochem.* **1998**, 256, 238–244. [CrossRef] [PubMed]
- Wang, Z.; Yeats, T.; Han, H.; Jetter, R. Cloning and characterization of oxidosqualene cyclases from *Kalanchoe daigremontiana*: Enzymes catalyzing up to 10 rearrangement steps yielding friedelin and other triterpenoids. *J. Biol. Chem.* 2010, 285, 29703–29712. [CrossRef] [PubMed]
- 23. Moreno-Hagelsieb, G.; Latimer, K. Choosing BLAST options for better detection of orthologs as reciprocal best hits. *Bioinformatics* **2008**, *24*, 319–324. [CrossRef]
- 24. Kepler, R.M.; Luangsa-Ard, J.J.; Hywel-Jones, N.L.; Quandt, C.A.; Sung, G.H.; Rehner, S.A.; Aime, M.C.; Henkel, T.W.; Sanjuan, T.; Zare, R.; et al. A phylogenetically-based nomenclature for *Cordycipitaceae (Hypocreales)*. *IMA Fungus* 2017, *8*, 335–353. [CrossRef]
- 25. Poralla, K.; Hewelt, A.; Prestwich, G.D.; Abe, I.; Reipen, I.; Sprenger, G. A specific amino acid repeat in squalene and oxidosqualene cyclases. *Trend Biochem. Sci.* **1994**, *19*, 157–158. [CrossRef]
- Kimura, M.; Kushiro, T.; Shibuya, M.; Ebizuka, Y.; Abe, I. Protostadienol synthase from *Aspergillus fumigatus*: Functional conversion into lanosterol synthase. *Biochem. Biophys. Res. Commun.* 2010, 391, 899–902. [CrossRef]