Chinese Journal of Natural Medicines

Discovery of fernane-type triterpenoids from *Diaporthe discoidispora* using genome mining and HSQC-based SMART technology

Yajing Wang, Yongfu Li, Yan Dong, Chunyan Yu, Chengwei Liu, Chang Li, Yi Sun, Yuehu Pei

Citation: Yajing Wang, Yongfu Li, Yan Dong, Chunyan Yu, Chengwei Liu, Chang Li, Yi Sun, Yuehu Pei, Discovery of fernane-type triterpenoids from *Diaporthe discoidispora* using genome mining and HSQC-based SMART technology, Chinese Journal of Natural Medicines, 2025, 23(3), 368-376. doi: 10.1016/S1875-5364(25)60837-5.

View online: https://doi.org/10.1016/S1875-5364(25)60837-5

Chinese Journal of Natural Medicines 23 (2025) 368-376

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Original article

Discovery of fernane-type triterpenoids from *Diaporthe discoidispora* using genome mining and HSQC-based SMART technology

Yajing Wang[a], Yongfu Li[a], Yan Dong[a], Chunyan Yu[b], Chengwei Liu[b], Chang Li[a, c,*], Yi Sun[d,*], Yuehu Pei[a,*]

[a] *Department of Medicinal Chemistry and Natural Medicine Chemistry, College of Pharmacy, Harbin Medical University, Harbin 150081, China*
[b] *Key Laboratory for Enzyme and Enzyme-like Material Engineering of Heilongjiang, College of Life Science, Northeast Forestry University, Harbin 150040, China*
[c] *Key Laboratory of Gut Microbiota and Pharmacogenomics of Heilongjiang Province, College of Pharmacy, Harbin Medical University, Harbin 150081, China*
[d] *Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China*

ARTICLE INFO

Article history:
Received 11 December 2023
Revised 26 March 2024
Accepted 8 April 2024
Available online 20 March 2025

Keywords:
Fernane
Triterpenoid
Genome mining
SMART technology
Endophytic fungi
Diaporthe discoidispora

ABSTRACT

In this study, we employed a combination of genome mining and heteronuclear single quantum coherence (HSQC)-based small molecule accurate recognition technology (SMART) technology to search for fernane-type triterpenoids. Initially, potential endophytic fungi were identified through genome mining. Subsequently, fine fractions containing various fernane-type triterpenoids were selected using HSQC data collection and SMART prediction. These triterpenoids were then obtained through targeted isolation and identification. Finally, their antifungal activity was evaluated. As a result, three fernane-type triterpenoids, including two novel compounds, along with two new sesquiterpenes and four known compounds were isolated from one potential strain, *Diaporthe discoidispora*. Their structures were elucidated through analysis of high-resolution electrospray ionization mass spectrometry (HR-ESI-MS) and nuclear magnetic resonance (NMR) spectroscopic data. The absolute configurations were determined using single-crystal X-ray diffraction analysis and electron capture detector (ECD) analysis. Compound **3** exhibited moderate antifungal activity against *Candida albicans* CMCC 98001 and *Aspergillus niger*.

1. Introduction

Microbial natural products have emerged as crucial resources for drug development due to their structural diversity and varied bioactivities [1, 2]. Among these, endophytic fungi have garnered significant attention, as compounds with substantial therapeutic potential, have been isolated from these predominantly microscopic organisms [3, 4]. The advancement of genome sequencing technology has positioned genome mining as an essential tool for identifying biosynthetic gene clusters and exploring their products [5-7].

Fernane-type triterpenoids represent a small category of triterpenoids characterized by a 6/6/6/6/5 pentacyclic hydrocarbon skeleton. They can be classified into two groups: group I with a cleaved E-ring and group II with an intact pentacyclic skeleton. Enfumafungin, one of the most renowned group I fernane-type triterpenoids, exhibits significant antifungal activity as a specific inhibitor of glucan synthesis [8]. To date, limited research has been conducted on the biosynthesis of these triterpenoids, with only two triterpene cyclases reported: *EfuA* and *PolA*. Studies have demonstrated that oxidosqualene can be converted to group I fernane-type triterpenoids by *EfuA* and to group II fernane-type triterpenoids by *PolA* in fungi [9, 10].

Small molecule accurate recognition technology (SMART) is a deep convolutional neural network-based tool employed for generating structural hypotheses from heteronuclear single quantum coherence (HSQC) data [11, 12]. While multiple bioactive compounds have been successfully isolated from various medicinal plants using the SMART strategy, to the best of our knowledge, there are limited reports on the application of SMART for isolating microbial natural products.

This study presents an effective approach for identifying fernane-type triterpenoids from endophytic fungi using genome mining and HSQC-based SMART strategy. The methodology comprises four stages: genome mining and large-scale fermentation, HSQC data collection and SMART prediction, targeted isolation and identification, and activity evaluation (Fig. 1). Initially, genome mining identified one strain, *Diaporthe discoidispora*, which possessed a triterpene cyclase similar to *EfuA*, named *DiapA* (54% identity). Subsequently, fraction 3.4 was further isolated using the SMART strategy. Targeted isolation yielded three fernane-type triterpenoids (including two novel compounds), two new sesquiterpenes, and four additional compounds. Structure elucidation was accomplished through spectroscopic analysis, calculated electron capture detector (ECD) method, and X-ray diffraction. Finally, the antifungal activities of these triterpenoids were assessed.

2. Results and Discussion

Genome mining of fernane-type triterpene cyclase. In our

* Corresponding author.
E-mail addresses: lichang661@126.com (C. Li); ysun@icmm.ac.cn (Y. Sun); peiyueh@vip.163.com (Y. Pei)

https://doi.org/10.1016/S1875-5364(25)60837-5

genome mining efforts focused on fernane-type terpene cyclases (TC), we identified a putative fernane-type terpene biosynthetic gene cluster in the fungus *D. discoidispora* (Fig. 2A). The cluster comprises fifteen genes: a terpene cyclase (*diapA*), three P450 monooxygenases (*diapB*, *diapC* and *diapJ*), an acetyltransferase (*diapD*), one glycosyl hydrolase (*diapG*), and one glycosyltransferase (GT; *diapO*). Additionally, the cluster includes an integral membrane protein (*diapI*), a fungal-specific transcription factor (*diapL*), a major facilitator superfamily domain-containing protein (*diapK*), an *α/β* hydrolase (*diapF*), a proline-specific peptidase (*diapE*), a putative aldehyde reductase (*diapM*), an adenylate-forming reductase (*diapN*), and a cutinase (*diapH*). Notably, *diap A* is a multifunctional enzyme consisting of a TC domain fused to a GT domain, exhibiting 54% identity to *EfuA* (Fig. 2).

SMART-guided isolation. The fermentation of *D. discoidispora* underwent ethyl acetate partitioning, yielding four fractions (Fr. 3.1–Fr. 3.4) through multiple chromatography columns. HSQC spectral data from Fr. 3.1–Fr. 3.4 were collected, converted to comma-separated value (CSV) files, and uploaded to the SMART system (http://smart.ucsd.edu/classic) to assess potential structure types within the fractions. Analysis of Fr. 3.4 revealed that the top 10 structures with the highest similarity were all triterpenoids, including fernane-type terpenoids. Consequently, Fr. 3.4 was selected for further isolation (Fig. 3).

2.1. Structural elucidation

Diaporpiol A (**1**) was isolated as colorless crystals. Its molecular formula was determined to be $C_{32}H_{50}O_6$ based on high-resolution electrospray ionization mass spectrometry (HR-ESI-MS) data at *m/z* 553.3501 $[M + Na]^+$ (Calcd. for $C_{32}H_{50}O_6Na$, 553.3499), indicating 8 degrees of unsaturation. The ^1H nuclear magnetic resonance (NMR) spectral data of **1** revealed ten methyl signals, including seven characteristic singlet methyl signals at δ_H 1.26, 1.10, 0.99, 0.98, 0.96, 0.75 and 1.99, as well as three doublet methyl signals at δ_H 0.76 (3H, d, *J* = 6.8 Hz), 0.86 (3H, d, *J* = 7.6 Hz) and 0.87 (3H, d, *J* = 7.6 Hz), indicating the presence of ten methyl groups in the structure. Additionally, there was one hydroxyl proton from the carboxylic group at δ_H 12.22, one exchangeable proton at δ_H 4.86, and two oxygenated methine protons at δ_H 4.81 (1H, ddd, *J* = 11.9, 10.0, 4.2 Hz) and 3.06 (1H, dd, *J* = 10.0, 4.9 Hz). Analysis of ^{13}C NMR and HSQC data revealed the presence of ten methyls at δ_C 28.4, 25.7, 23.8, 21.3, 21.3, 20.6, 18.5, 17.0, 16.2 and 10.2, two olefinic carbons at δ_C 133.9 and 131.3, seventeen sp^3 carbons (including two oxygenated at δ_C 78.0 and 72.3), and three carbonyl carbons at δ_C 213.9, 173.4 and 170.2. The ^1H NMR and ^{13}C NMR data of **1** (Tables 1 and 2) were closely related to those of co-isolated known compound WF11605 (**3**) [13], except for the absence of one glucoside moiety, suggesting that **1** was the aglycone of **3**, also a fernane-type tetracyclic triterpenoid. Its planar structure was further confirmed by the ^1H-^1H correlation spectroscopy (COSY) and heteronuclear multiple bond correlations (HMBCs) (Fig. 4).

The relative configuration of compound **1** was determined through nuclear overhauser effect spectroscopy (NOESY) correlations (Fig. 5) and coupling constants. The large coupling constant between H-2 and H-3 (*J* = 10.0 Hz) confirmed their *trans* configuration. The methyl group at C-28 was assigned a *β*-orientation based on the NOESY spectrum, which revealed correlations between H-2 (δ_H 4.81) and H-28 (δ_H 0.99). Furthermore, the NOESY correlation of H-3 (δ_H 3.06) with H-5 (δ_H 1.12) led to the assignment of an *α*-orientation for H-3 and H-5. The observed NOESY correlations of H-7*α* (δ_H 1.85) with H-5 (δ_H 1.12) indicated their shared *α* orientation. Additionally, the NOESY correlations of H-15*β* (δ_H 2.18) with H-7*β* (δ_H 2.09) and H-28 (δ_H 0.99) positioned these protons on the same side. Moreover, the NOESY cross-peaks of H-15*α* (δ_H 2.69) with H-26 (δ_H 1.10) and H-24 (δ_H 1.26) placed H-26 and H-24 on the opposite side. The NOESY correlations between H-18 (δ_H 2.88) and H-27 (δ_H 0.98) indicated

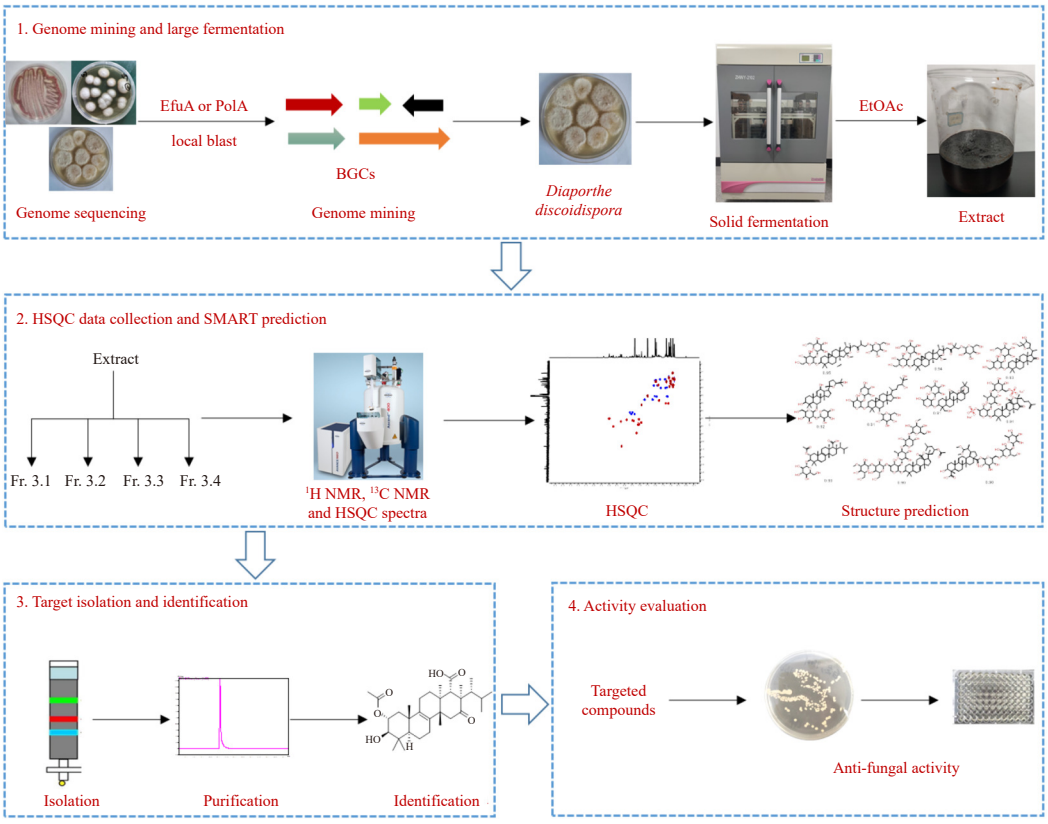

Fig. 1 Strategy for the discovery of fernane-type compounds by genome mining and SMART prediction.

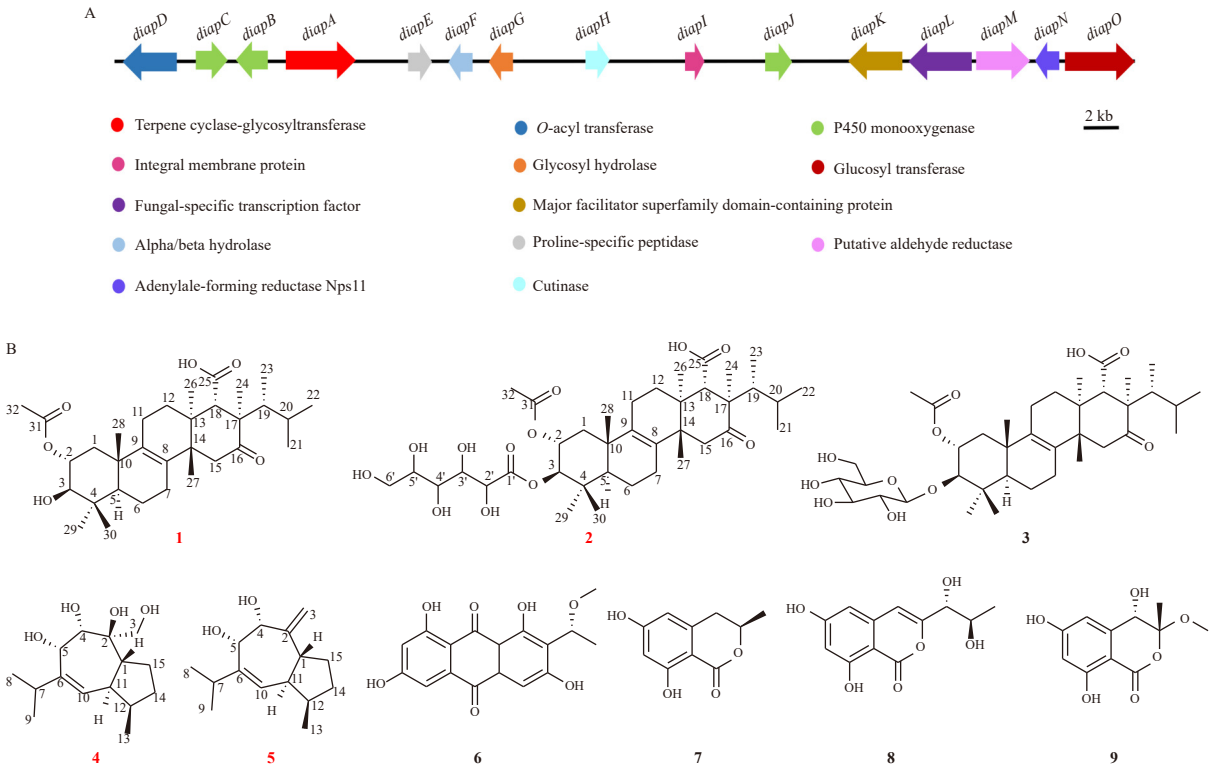

Fig. 2 The proposed biosynthetic gene cluster (A) and chemical structures of isolated compounds (red are new) (B).

their β configuration. Single crystals of **1** were obtained from MeOH); UV (MeOH)–CH_2Cl_2 (1∶1) suitable for X-ray diffraction (Fig. 6), unambiguously determining its planar structure and absolute configuration. The absolute configuration of **1** was established as 2*R*, 3*R*, 5*R*, 10*S*, 13*S*, 14*S*, 17*R*, 18*S*, 19*R*.

Diaporpiol B (**2**) was isolated as white powders, and its molecular formula was determined to be $C_{38}H_{60}O_{12}$ based on HR-ESI-MS data at *m/z* 731.3981 [M + Na]$^+$ (Calcd. for $C_{38}H_{60}O_{12}Na$, 731.3982), indicating 9 degrees of unsaturation. The ^{1}H NMR spectral data of **2** exhibited similar seven singlet methyl signals at δ_H 1.93, 1.27, 1.10, 1.02, 0.99, 0.84 and 0.82, as well as three doublet methyl signals at δ_H 0.77 (3H, d, *J* = 6.7 Hz), 0.85 (3H, d, *J* = 6.8 Hz) and 0.87 (3H, d, *J* = 7.1 Hz). Additionally, five characteristic exchangeable protons were observed at δ_H 5.35 (1H, d, *J* = 6.0 Hz), 4.53 (1H, d, *J* = 2.4 Hz), 4.37 (1H, br s), 4.35 (1H, br s) and 4.25 (1H, br s), along with six oxygenated methine signals at

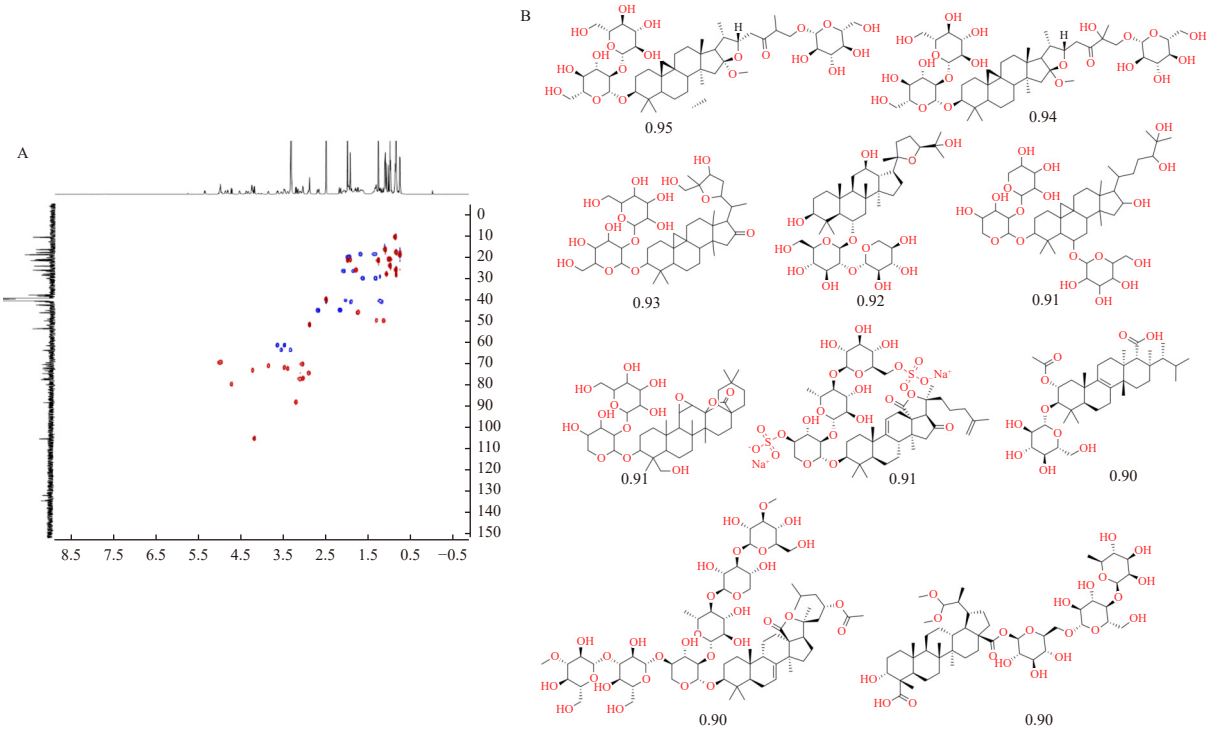

Fig. 3 HSQC spectrum of Fr. 3.4 (A) and SMART results (top 10 structures based on cosine similarity score) of Fr. 3.4 (B).

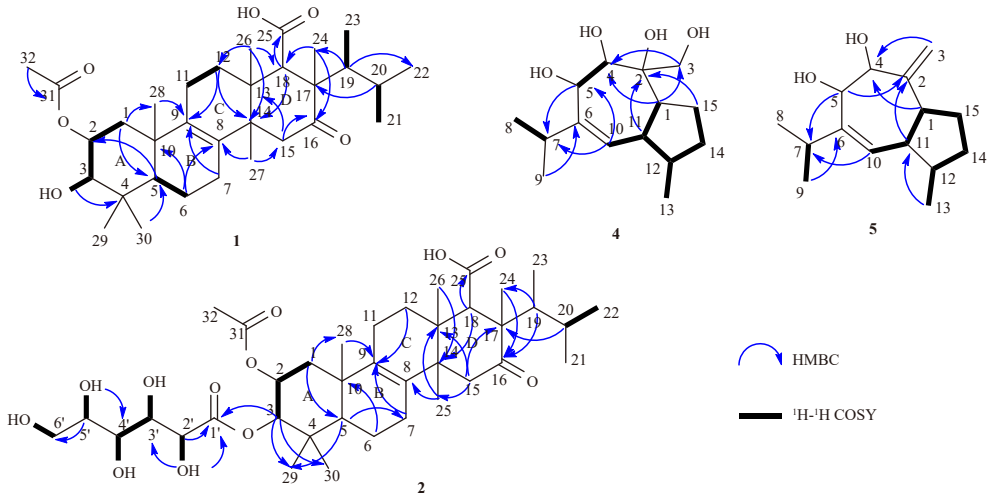

Fig. 4 Key ^{1}H-^{1}H COSY and HMBCs of compounds **1**–**2** and **4**–**5**.

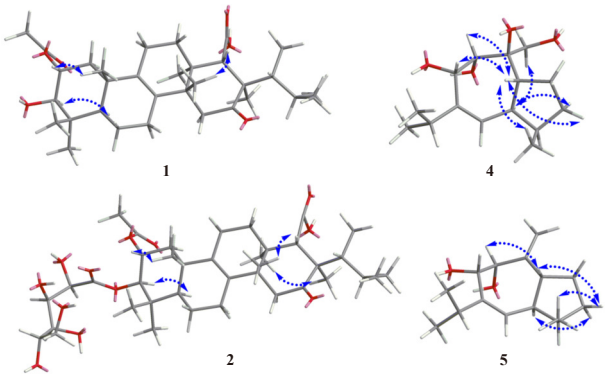

Fig. 5 Key NOESY correlations of compounds **1**–**2** and **4**–**5**.

δ_H 5.01 (1H, m), 4.72 (1H, d, J = 10.4 Hz), 4.23 (1H, br s), 3.86 (1H, brs), 3.49 (1H, br s) and 3.41 (1H, br s). The ^{13}C NMR and HSQC spectrum revealed 38 carbon resonances, including four carbonyl group signals at δ_C 213.9, 173.4, 172.1 and 170.1, two olefinic carbons at δ_C 133.6 and 131.6, seven oxygenate carbons at δ_C 79.4, 72.9, 72.1, 71.6, 70.9, 69.1 and 63.3, ten methyl groups at δ_C 28.0, 25.7, 23.9, 21.3, 20.9, 20.7, 18.3, 17.5, 16.2 and 10.4, as well as fifteen sp^3 carbons.

Comparison of the NMR data of **2** with those of **1** indicated that compound **2** was an analog of **1**, differentiated by an additional monosaccharide moiety. The ^{1}H-^{1}H COSY correlations of HO-2' (δ_H 5.35)/H-2' (δ_H 4.23), HO-3' (δ_H 4.25)/H-3' (δ_H 3.86)/H-4' (δ_H 3.41)/HO-4' (δ_H 4.37), HO-5' (δ_H 4.53)/H-5' (δ_H 3.49), HO-6' (δ_H 4.35)/H-6' (δ_H 3.56, 3.33), as well as the observed HMBCs from HO-2' (δ_H 5.35) to C-1' (δ_C 172.1) and C-3' (δ_C 70.9), from H-2' (δ_H 4.23) to C-1' (δ_C 172.1), from HO-5' (δ_H 4.53) to C-4' (δ_C 72.1), and from H-5' (δ_H 3.49) to C-6' (δ_C 63.3) demonstrated the presence of a monosaccharide moiety. The connection point was determined to be at C-3, evidenced by the upfield shift of H-3 from δ_H 3.06 in **1** to δ_H 4.72 in **2**. This was further corroborated by an HMBC between H-3 (δ_H 4.72) and the carbonyl carbon C-1' (δ_C 172.1). The monosaccharide was identified as gluconic acid through comparison with published NMR data [14]. To the best of our knowledge, this represents the first report of a natural product containing a gluconic acid moiety.

The relative configuration of compound **2** was determined through NOESY correlations (Fig. 4). NOESY correlations between H-2 (δ_H 5.01) and H-28 (δ_H 1.02) indicated their β configuration. Correlations of H-3 (δ_H 4.72) with H-5 (δ_H 1.31) placed these protons in the α configuration. Observed NOESY correlations of H-5 (δ_H 1.31)/H-7α (δ_H 1.87)/H-15α (δ_H 2.70) indicated α-configuration. Additionally, NOESY correlations of H-24 (δ_H 1.27)/H-26 (δ_H 1.10), and both protons with H-15α (δ_H 2.70) led to the assignment of an α-orientation for H-24 and H-26. In the NOESY spectrum of **2**, nuclear overhauser effect (NOE) correlations of H-28 (δ_H 1.02)/H-7β (δ_H 2.09)/H-27 (δ_H 0.99)/H-15β (δ_H 2.19) positioned these protons on the same side. Additionally, NOESY correlations between H-18 (δ_H 2.89) and H-27 (δ_H 0.99) suggested

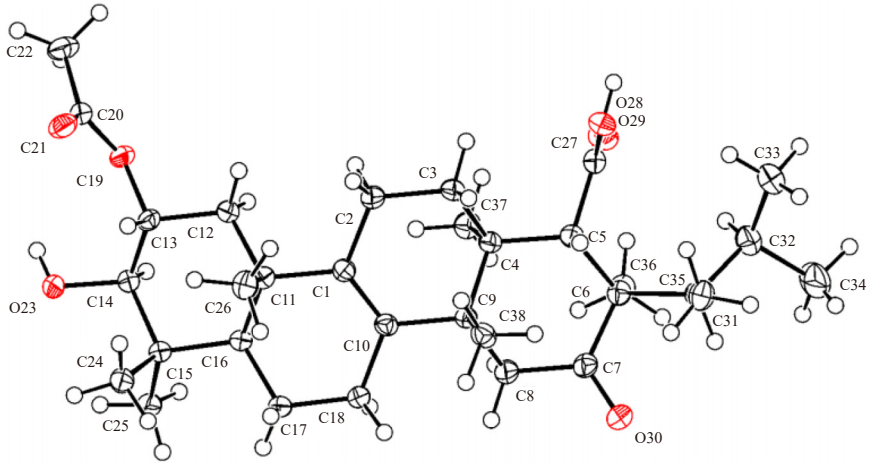

Fig. 6 The ORTEP drawing of compound **1**.

their β-oriented. To determine the stereochemistry of **2**, ECD analysis was conducted by comparison with **1**. The ECD spectrum of **2** exhibited similar Cotton effects to those of **1**, indicating that the parent nuclei of compound **2** possessed the same absolute configuration as **1**, except for its flexible side chain (Fig. 7).

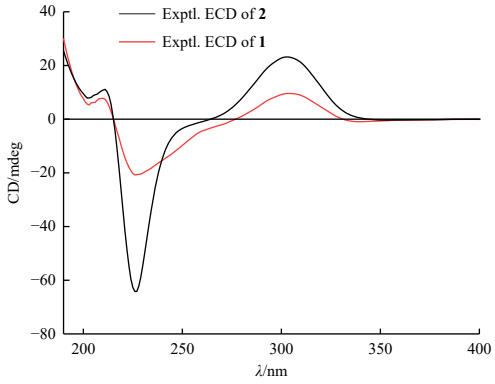

Fig. 7 Experimental ECD spectra of compounds **1**–**2** in MeOH.

Oxytropiol K (**4**) was isolated as white powders, and its molecular formula was determined to be $C_{15}H_{26}O_4$ based on HR-ESI-MS data at m/z 293.1722 $[M + Na]^+$ (Calcd. for $C_{15}H_{26}O_4Na$, 293.1723), indicating 3 degrees of unsaturation. The ^1H NMR spectrum of **4** recorded in CD_3OD revealed three methyl signals at δ_H 0.89 (3H, d, J = 7.1 Hz), 1.03 (3H, d, J = 6.8 Hz), and 1.04 (3H, d, J = 6.8 Hz), one pair of oxygenated methylene signals at δ_H 3.82 (1H, d, J = 11.2 Hz) and 4.19 (1H, d, J = 11.2 Hz), two oxygenated methine signals at δ_H 3.47 (1H, d, J = 1.7 Hz) and 4.29 (1H, d, J = 1.7 Hz), and one olefinic proton signal at δ_H 5.71 (1H, d, J = 3.5 Hz). The ^{13}C NMR and HSQC results indicated the presence of 15 carbon resonances attributable to two olefinic carbons (δ_C 147.2 and 129.1), four oxygenated carbons (δ_C 82.3, 78.5, 75.0 and 63.6), two methylenes (δ_C 33.7 and 25.6), four methines (δ_C 51.4, 41.8, 38.7 and 37.5), and three methyl groups (δ_C 22.2, 22.0 and 15.3). The ^1H-^1H COSY correlations of H-10 (δ_H 5.71)/H-11 (δ_H 2.91)/H-1 (δ_H 1.92), H-4 (δ_H 3.47)/H-5 (δ_H 4.29), combined with the HMBCs from H-10 (δ_H 5.71) to C-5 (δ_C 75.0), from H-1 (δ_H 1.92) to C-2 (δ_C 78.5) and C-4 (δ_C 82.3), from H-11 (δ_H 2.91) to C-2 (δ_C 78.5) confirmed a seven-membered ring. The ^1H-^1H COSY correlations of H-7 (δ_H 2.35)/H-8 (δ_H 1.04) and HMBCs from H-9 (δ_H 1.03) to C-6 (δ_C 147.2), from H-10 (δ_H 5.71) to C-7 (δ_C 37.5), from H-5 (δ_H 4.29) to C-7 (δ_C 37.5) indicated that the isopropyl group linkage was at the C-6 position. The ^1H-^1H COSY correlations of H-11 (δ_H 2.91)/H-12 (δ_H 2.17)/H-14 (δ_H 1.72, 1.34)/H-15 (δ_H 1.81), H-12 (δ_H 2.17)/H-13 (δ_H 0.89), combined with the HMBCs from H-11 (δ_H 2.91) and H-15 (δ_H 1.81) to C-2 (δ_C 78.5) revealed the fused five-membered ring, suggesting compound **4** possessed a guaiane-type sesquiterpene skeleton [15]. Additionally, HMBCs from H-1 (δ_H 1.92) to C-3 (δ_C 63.6), from H-3 (δ_H 4.19, 3.82) to C-4 (δ_C 82.3) demonstrated that the carbinol group was linked to the C-2 position. Thus, the planar structure of **4** was elucidated. The relative configuration of **4** was established based on its NOESY correlations and coupling constant analysis. The NOESY correlations from H_α-14 (δ_H 1.72) to H-11 (δ_H 2.91) and H-3 (δ_H 4.19, 3.82) indicated that H-3 and H-11 were in the α orientation. The NOESY cross-peaks of H-4 (δ_H 3.47)/H-1 (δ_H 1.92)/H-5 (δ_H 4.29), H-13 (δ_H 0.89)/H-1 (δ_H 1.92)/H_β-14 (δ_H 1.34) suggested that these groups were β configured. Furthermore, the small coupling constant between H-4 and H-5 (J = 1.7 Hz) confirmed that these two protons were in the *cis* configuration. The absolute configuration of **4** was determined by theoretical calculations of its ECD (Fig. 8). The calculated ECD spectrum of (1*R*,2*R*,4*S*,5*S*,11*S*,12*R*)-**4** showed good agreement with the measured spectrum. Therefore, the absolute configuration of **4** was established as 1*R*, 2*R*, 4*S*, 5*S*, 11*S*, 12*R*.

Oxytropiol L (**5**) was isolated as white powders, and its molecular formula was determined to be $C_{15}H_{24}O_2$ based on HR-ESI-MS data at m/z 259.1667 $[M + Na]^+$ (Calcd. for $C_{15}H_{24}O_2Na$, 259.1668), indicating 4 degrees of unsaturation. The ^1H NMR spectrum of **5** recorded in CD_3OD revealed three methyl signals at δ_H 0.94 (3H, d, J = 7.1 Hz) and 1.05 (6H, d, J = 6.8 Hz), one *exo*-methylene at δ_H 5.22 (1H, d, J = 1.5 Hz) and 5.11 (1H, d, J = 1.5 Hz), one olefinic proton signal at δ_H 5.74 (1H, d, J = 3.7 Hz), and two oxygenated methine signals at δ_H 4.06 (1H, d, J = 0.8 Hz) and 4.17 (1H, brs). The ^{13}C NMR and HSQC results indicated the presence of 15 carbon resonances attributable to four olefinic carbons (δ_C 151.6, 147.4, 128.5 and 107.3), two oxygen-bearing carbons (δ_C 77.8 and 75.9), two methylenes (δ_C 33.7 and 28.5), four methines (δ_C 49.0, 47.6, 38.3 and 38.0), and three methyl groups (δ_C 21.9, 21.9 and 17.4). These ^{13}C NMR data were similar to those of **4**, except that the two signals in **4** at 78.5, 63.6 were replaced by 151.6, 107.3 in **5**, suggesting that **5** was also a guaiane-type sesquiterpenoid. Its structure was further elucidated by its key ^1H-^1H COSY and HMBCs (Fig. 4). The relative configuration of **5** was established based on its NOESY correlations and coupling constant analysis. The NOESY correlations from H-1 (δ_H 2.21) to H-4 (δ_H 4.06), from H_β-14 (δ_H 1.31) to H-1 (δ_H 2.21) and H-13 (δ_H 0.94) indicated that H-1, H-4, H-13 and H_β-14 were in the β orientation. Additionally, the small coupling constant between H-4 and H-5 (J = 0.8 Hz) confirmed that the dihedral angle of the two protons is close to 90°, and these two protons were in the *cis* configuration. The NOESY cross-peaks of H_α-14 (δ_H 1.94)/H-11 (δ_H 2.55) implied that these groups were α configured. Comparison of the calculated ECD spectra data with experimental data of compound **5** indicated that the calculated curve matched well with the experimental one (Fig. 8). The absolute configuration of compound **5** was assigned to be 1*R*, 4*R*, 5*S*, 11*S*, 12*R*.

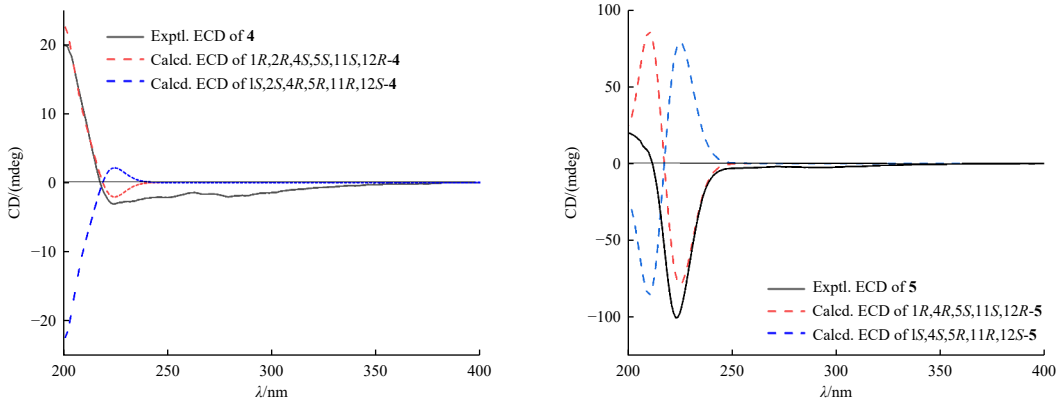

Fig. 8 Experimental CD spectra and calculated ECD spectra of compounds **4**–**5** in MeOH.

The five previously identified compounds were characterized as WF11605 (**3**) [13], 1,3,6,8-tetrahydroxy-anthraquinone (**6**) [16], (*R*)-6-hydroxymellein (**7**) [17], 6,8-dihydroxy-3-(10*R*,20*R*-dihydroxypropyl)-isocoumarin (**8**) [18], and penicoffrazin B (**9**) [19] through analysis of their spectroscopic data and comparison with previously reported literature values.

2.2. Proposed biosynthesis pathway

A local BLAST search using EfuA as a query in the genome of the fungus *D. discoidispora* identified a cluster potentially related to fernane-type triterpenes. Eight genes within this cluster likely contribute to the biosynthesis of compounds **1–3**. These include the TC-GT fusion protein *DiapA*, acetyltransferase *DiapD*, three P450 monooxygenases (*DiapB*, *DiapC*, *DiapJ*), glycosyl hydrolase *DiapG*, glycosyl transferase *DiapO*, and *α*/*β* hydrolase *DiapF*. The proposed biosynthetic pathway begins with *DiapA* catalyzing cyclization and glycosylation using 2,3-oxidosqualene as a substrate. Subsequently, hydroxylation at C-2 likely occurs *via* *DiapB/C/J*, concurrent with acetylation by *DiapD*. The structure then undergoes oxidation at C-25 and lactonization through a Baeyer-Villiger reaction, catalyzed by monooxygenases *DiapB/C/J*, followed by ring cleavage to synthesize **3**. Compound **3** may then be converted to **1** by *DiapG* and further to **2** by *DiapO*. Alternatively, **3** could be converted to **2** by *DiapF* and other oxidative proteins (Fig. 9).

2.3. Evaluation of antifungal bioactivity

Antifungal assays of compounds **1**–**7** were conducted against *Candida albicans* CMCC 98001 and *Aspergillus niger*. The results demonstrated that only compound **3** exhibited moderate activity, with minimum inhibitory concentrations (MICs) of 7.81 and 31.25 μg·mL^{-1} (Table 3), respectively.

This study presents an integrated approach combining genome mining and HSQC-based SMART technology for the identification of fernane-type triterpenoids from endophytic fungi. The investigation led to the selection of the strain *D. discoidispora* for further analysis. From Fraction 3.4, three fernane-type triterpenoids (including two novel compounds), two new sesquiterpenes, and four known compounds were isolated and characterized. Antifungal assays revealed that compound **3** exhibited moderate activity against *Candida albicans* CMCC 98001 and *A. niger*, suggesting its potential as a lead compound for antifungal agent development. This research offers a novel perspective on targeted isolation procedures compared to conventional methods. Furthermore, the findings contribute to expanding the structural diversity of fernane-type triterpenoids.

3. Experimental

Optical rotation was measured using a SAC-i (Atago, Japan) spectrometer. The UV spectrum was recorded on a UV-2550 (Shimadzu, Japan) UV/Vis spectrometer. IR spectrum was obtained using an IR Tracer-100 (Shimadzu, Japan). 1D and 2D NMR spectra were acquired on Bruker AV-600 and Bruker AV-400 spectrometers. HR-ESI-MS data were collected using a Waters G2-S QTOF mass spectrometer. ECD spectra were measured on a Bio-logic MOS-450 circular dichrometer. Open CC was performed using silica gel (200–300 mesh, Qingdao Haiyang Chemical Group Corp. Qingdao, China) and octadecylsilane (ODS) (50 μm, YMC, Japan). The samples were purified by semipreparative high-performance liquid chromatography (HPLC) equipped with an LC-20AR system, a UV detector, and an HPLC column (5 μm, 10 mm × 250 mm, COSMOSIL AR II-C_{18} column).

3.1. Strain isolation and cultivation

The fungal strain was isolated from the inner tissue of the surface-sterilized fruit of *Rauwolfia yunnanensis* Tsiang, collected in Yunnan Province, China, in 2017. It was identified as *D. discoidispora* through rRNA gene sequence analysis and assigned the Genbank accession number OQ552614. A voucher specimen (No. luo2-twig-2i-f) was deposited at the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing. The strain underwent fermentation on autoclaved rice solid-substrate medium (one hundred and fifty 500 mL Erlenmeyer flasks, each containing 40 g rice and 60 mL water) for 30 d at 28 °C.

3.2. Genome sequencing, genome mining, and purification

The genome sequencing of *D. discoidispora* was conducted utilizing the Illumina HiSeq × 10 Platform at Shanghai Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China), resulting in 1386 scaffolds encompassing approximately 56.1Mb. Subsequently, gene prediction was executed using Maker 2. The predicted gene products were then employed to construct a database for local BLAST search purposes.

Following incubation, the solid rice medium was extracted thrice with EtOAc to yield a crude extract (67.0 g), which was suspended in 95% MeOH and partitioned with petroleum ether to obtain a 95% MeOH total extract (61.4 g). The extract was fractionated by ODS flash column chromatography (5 cm × 30 cm) eluting with 4 L of each of 40%, 60%, 80%, and 100% (*V*/*V*)

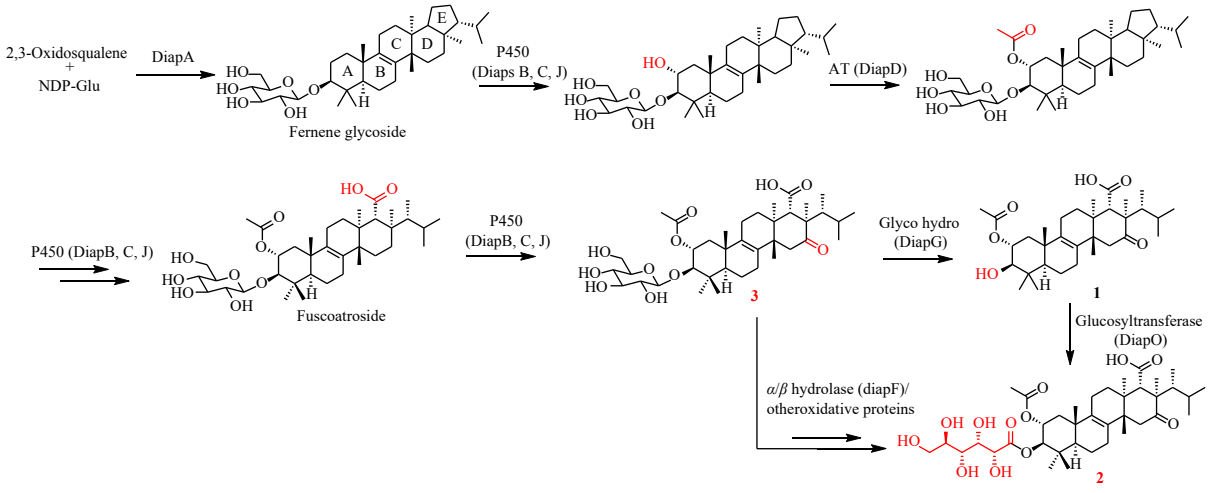

Fig. 9 Proposed biosynthetic pathway of compounds **1**–**3**.

MeOH/H_2O to yield four fractions (Fr. 1–Fr. 4). Fraction Fr. 3 (3.3 g) was chromatographed on silica gel using gradient elution with CH_2Cl_2/Acetone (100∶1 to 0∶100) to yield four subfractions Frs. 3.1–3.4. HSQC spectral data from Frs.3.1–3.4 were collected, converted to CSV files, and uploaded into the SMART system to identify possible structure types from these four fractions. Based on the SMART analysis, Fr. 3.4 was purified by a reversed-phase column using ODS, eluting gradually with H_2O/MeOH (4∶6 to 2∶8) to give three subfractions (Frs. 3.4.1–3.4.3). Subfraction Fr. 3.4.1 was applied to a Sephadex LH-20 column with MeOH/CH_2Cl_2 (1∶1) as eluent to obtain compound **6** (2.0 mg). Subfraction Fr. 3.4.2 was purified by silica-gel column chromatography (CC) (CH_2Cl_2/MeOH, 1∶0 to 0∶1) to obtain compound **2** (25.0 mg) and compound **3** (50.0 mg). Subfraction Fr. 3.4.3 recrystallized with CH_2Cl_2/MeOH (1∶1) to yield compound **1** (5.0 mg). Fr 2 (10.4 g) was applied to a reversed-phase column using ODS, eluting gradually with H_2O/MeOH (5∶5 to 3∶7) to give five subfractions (Frs. 2.1–2.5). Subfraction Fr. 2.2 was purified by silica-gel CC, eluted with a PE/Acetone (10∶1 to 1∶1) gradient to give compound **4** (30.0 mg). Compound **5** (10.2 mg, t_R = 31.5 min) was obtained from Fr. 2.3 using semipreparative HPLC with MeOH/H_2O (65/35, *V*/*V*) as eluent. Fr. 1 was subjected to silica-gel CC, eluted with a CH_2Cl_2/MeOH (1∶0 to 0∶1) gradient, to yield four subfractions (Frs. 1.1–1.4). Fraction Fr. 1.1 recrystallized with methanol to yield compound **7** (200.0 mg). Fraction Fr. 1.4 was purified by a Sephadex LH-20 column with MeOH as eluent to give three subfractions (Frs.1.4.1–1.4.5). Compounds **8** (5.0 mg) and **9** (4.0 mg) were obtained from Fr. 1.4.2 using silica-gel CC with a CH_2Cl_2/MeOH (10∶1 to 0∶1) gradient.

Diaporpiol A (**1**): colorless crystals; $[\alpha]_D^{20}$ −57 (*c* 0.2, MeOH λ_{max} (log ε): 203 (2.74) nm; IR (KBr) ν_{max}: 3456, 1700 cm^{-1}. CD (MeOH) 226 ($\Delta\varepsilon$ −1.89) nm, 304 (+0.90) nm;^{1}H NMR (600 MHz, DMSO-d_6) and ^{13}C NMR (100 MHz, DMSO-d_6) data (Table 1; HR-ESI-MS *m*/*z* [M + Na]$^+$ 553.3501 (Calculated for $C_{32}H_{50}O_6Na$, 553.3499).

Diaporpiol B (**2**): white powder; $[\alpha]_D^{20}$ −72 (*c* 0.2, MeOH λ_{max}

Table 1 ^{1}H NMR (600 MHz) and ^{13}C NMR (100 MHz) data for compounds **1**–**2**

No.	1		2	
	δ_H (*J* Hz)	δ_C, type	δ_H (*J* Hz)	δ_C, type
1	1.94, m 1.09, d (11.9)	40.0, CH_2	2.04, dd (11.9, 4.7) 1.24, d (11.9)	40.1, CH_2
2	4.81, ddd (11.9, 10.0, 4.2)	72.3, CH	5.01, m	69.1, CH
3	3.06, dd (10.0, 4.9)	78.0, CH	4.72, d (10.4)	79.4, CH
4	-	39.2, C	-	39.1, C
5	1.12, d (12.4)	49.6, CH	1.31 (1H, br s)	49.3, CH
6	1.68, m 1.37, m	18.6, CH_2	1.71, dd (12.7, 7.2) 1.41, m	18.5, CH_2
7	2.09, m 1.85, m	26.3, CH_2	2.09, m 1.87, m	26.2, CH_2
8	-	131.3, C	-	131.6, C
9	-	133.9, C	-	133.6, C
10	-	38.2, C	-	38.0, C
11	1.96, m	19.8, CH_2	1.97, m	19.9, CH_2
12	1.65, m 1.34, m	29.6, CH_2	1.64, m 1.33, m	29.5, CH_2
13	-	42.1, C	-	42.1, C
14	-	37.4, C	-	37.3, C
15	2.69, d (16.8) 2.18, d (16.8)	44.7, CH_2	2.70, d (16.8) 2.19, d (16.8)	44.7, CH_2
16	-	213.9, C=O	-	213.9, C=O
17	-	53.2, C	-	53.1, C
18	2.88, s	51.4, CH	2.89, s	51.5, CH
19	1.75, m	45.7, CH	1.71, m	45.6, CH
20	1.81, m	25.9, CH	1.79, m	26.0, CH
21	0.76, d (6.8)	18.5, CH_3	0.77, d (6.7)	17.5, CH_3
22	0.87, d (7.6)	10.2, CH_3	0.87, d (7.1)	10.4, CH_3
23	0.86, d (7.6)	25.7, CH_3	0.85, d (6.8)	25.7, CH_3
24	1.26, s	21.3, CH_3	1.27, s	21.3, CH_3
25	-	173.4, C=O	-	173.4, C=O
26	1.10, s	16.2, CH_3	1.10, s	16.2, CH_3
27	0.98, s	23.8, CH_3	0.99, s	23.9, CH_3
28	0.99, s	20.6, CH_3	1.02, s	20.7, CH_3
29	0.96, s	28.4, CH_3	0.82, s	28.0, CH_3

Continued

No.	1		2	
	δ_H (*J* Hz)	δ_C, type	δ_H (*J* Hz)	δ_C, type
30	0.75, s	17.0, CH_3	0.84, s	18.3, CH_3
31	-	170.2, C=O	-	170.1, C=O
32	1.99, s	21.3, CH_3	1.93, s	20.9, CH_3
3-OH	4.86, d (4.9)	-		
32-COOH	12.22, br s	-	12.21, br s	-
1'			-	172.1, C=O
2'			4.23, br s	72.9, CH
3'			3.86, br s	70.9, CH
4'			3.41, br s	72.1, CH
5'			3.49, br s	71.6, CH
6'			3.56, d (11.0) 3.33, overlap	63.3, CH_2
2'-OH			5.35, d (6.0)	-
3'-OH			4.25, br s	-
4'-OH			4.37, br s	-
5'-OH			4.53, d (2.4)	-
6'-OH			4.35, br s	-

Proton coupling constants (*J*) in Hz are given in parentheses. The assignments were based on ^{1}H-^{1}H COSY, HSQC, and HMBC experiments.

(log ε): 200 (4.41) nm; IR (KBr) ν_{max}: 3439, 2963, 2941, 1734 cm^{-1}. CD (MeOH) 226 ($\Delta\varepsilon$ −1.83) nm, 304 (+0.66) nm;^{1}H NMR (600 MHz, DMSO-d_6) and ^{13}C NMR (100 MHz, DMSO-d_6) data (Table 1; HR-ESI-MS *m/z* [M + Na]$^+$ 731.3981 (Calculated for $C_{38}H_{60}O_{12}Na$, 731.3982).

Oxytropiol K (**4**): white powder; $[\alpha]_D^{20}$ +73 (*c* 0.1, MeOH λ_{max} (log ε): 200 (3.80) nm, 222 (2.58) nm; IR (KBr) ν_{max}: 3380, 2963 cm^{-1}. CD (MeOH) 224 ($\Delta\varepsilon$ −0.40) nm; ^{1}H NMR (600 MHz, CD_3OD) and ^{13}C NMR (100 MHz, CD_3OD) data (Table 2); HR-ESI-MS *m/z* [M + Na]$^+$ 293.1722 (Calculated for $C_{15}H_{26}O_4Na$, 293.1723).

Oxytropiol L (**5**): white powder; $[\alpha]_D^{20}$ +64 (*c* 0.1, MeOH λ_{max} (log ε): 200 (4.08) nm; IR (KBr) ν_{max}: 3382, 2958, 2870 cm^{-1}. CD

Table 2 ^{1}H NMR (600 MHz) and ^{13}C NMR (100 MHz) data for compounds **4**–**5**

No.	4		5	
	δ_H (*J* Hz)	δ_C, type	δ_H (*J* Hz)	δ_C, type
1	1.92, m	51.4, CH	2.21, m	49.0, CH
2	-	78.5, C	-	151.6, C
3	4.19, d (11.2)	63.6, CH_2	5.22, d (1.5)	107.3, CH_2
	3.82, d (11.2)		5.11, d (1.5)	
4	3.47, d (1.7)	82.3, CH	4.06, d (0.8)	77.8, CH
5	4.29, d (1.7)	75.0, CH	4.17, br s	75.9, CH
6	-	147.2, C	-	147.4, C
7	2.35, m	37.5, CH	2.35, m	38.3, CH
8	1.04, d (6.8)	22.2, CH_3	1.05, d (6.8)	21.9, CH_3
9	1.03, d (6.8)	22.0, CH_3	1.05, d (6.8)	21.9, CH_3
10	5.71, d (3.5)	129.1, CH	5.74, d (3.7)	128.5, CH
11	2.91, ddd (10.5, 6.9, 3.5)	41.8, CH	2.55, ddd (11.4, 7.9, 3.7)	47.6, CH
12	2.17, qd (6.9, 2.3)	38.7, CH	2.23, m	38.0, CH
13	0.89, d (7.1)	15.3, CH_3	0.94, d (7.1)	17.4, CH_3
14	1.72, m	33.7, CH_2	1.94, m	33.7, CH_2
	1.34, m		1.31, m	
15	1.81, m	25.6, CH_2	1.90, m	28.5, CH_2
			1.76, m	

Proton coupling constants (*J*) in Hz are given in parentheses. The assignments were based on ^{1}H-^{1}H COSY, HSQC, and HMBC experiments.

Table 3 Antifungal activities of compounds **1**–**7**

Compound	*C. albicans* (MIC/μg·mL^{-1})	*A. niger* (MIC/μg·mL^{-1})
1	> 500	> 500
2	> 500	500
3	7.81	31.25
4	500	500
5	500	> 500
6	500	> 500
7	500	> 500
Caspofungin	0.98	0.12

(MeOH) 223 (Δε −0.59) nm;^{1}H NMR (600 MHz, CD_3OD) and ^{13}C NMR (100 MHz, CD_3OD) data (Table 2); HR-ESI-MS *m/z* [M + Na]$^+$ 259.1667 (Calculated for $C_{15}H_{24}O_2Na$, 259.1668).

X-ray crystallography of diaporpiol A (**1**). Single crystals were obtained from dichloromethane and MeOH. A suitable crystal was selected and analyzed on an XtaLAB Synergy Custom diffractometer: Kappa single diffractometer utilizing Cu Kα radiation. The crystal was maintained at 100.01(10) K during data collection. Using Olex2 [20], the structure was elucidated with the ShelXT [21] structure solution program employing Direct Methods and refined with the ShelXL [22] refinement package using Least Squares minimization. Crystal data for $C_{32}H_{50}O_6$ (*Mr*: 530.72): orthorhombic, space group $P2_12_12_1$, a = 6.74265(10) Å, b = 20.1715(3) Å, c = 21.9665(3) Å, α = 90°, β = 90°, γ = 90°, V = 2987.65(7) Å3, Z = 4, T = 100.01(10) K, μ(CuKα) = 0.634 mm^{-1}, *Dcalc* = 1.180 g·cm^{-3}, 28899 reflections measured (5.948° ≤ 2θ ≤ 153.92°), 6108 unique R_{int} = 0.0366, R_{sigma} = 0.0259 which were used in all calculations. The final R_1 was 0.0318 (I > 2σ (I)), and wR_2 was 0.0825. The goodness of fit on F^2 was 1.070. Flack parameter = −0.06(6). The crystallographic data for compound **1** (CCDC 2268746) have been deposited in the Cambridge Crystallographic Data Centre.

The conformational search was conducted using the random search method with MMFF94s force field as implemented in the Sybyl 8.0 software package. The obtained conformers (> 95%) were selected and further optimized at the B3LYP/6-31G(d) level using the Gaussian 09 software. The resulting stable conformers underwent ECD calculation at the CAM-B3LYP/6-311+G(d,p) level with 10 excited states in methanol. The overall predicted CD curves of **4** and **5** were simulated according to Boltzmann weighting with a half-width of 0.2 eV using the SpecDic 1.70.1 program, demonstrating good agreement with the measured CD curves.

Antifungal bioactivity was assessed using a 2-fold dilution assay in 96-well microtiter plates. The fungal strains *C. albicans* CMCC 98001 and *A. niger* were cultivated on sabouraud dextrose agar medium (2% glucose, 1% peptone, and 1.8% agar) at 28 °C for 4–7 d. Spores were collected and adjusted to a concentration of 10^7–10^9/mL with 0.9% saline. Subsequently, 100 μL of this mixture was added to 200 mL sabouraud dextrose broth to create a spore solution. Compounds were dissolved in DMSO to a concentration of 10 mg·mL^{-1}. 10 μL of sample solution was mixed with 200 μL of spore solution, which was then serially diluted with spore solution until reaching a concentration of 0.03 μg·mL^{-1}. Caspofungin served as the positive control, while DMSO acted as the negative control. The 96-well microtiter plates were incubated at 28 °C for 48 h. The MIC was determined as the lowest concentration at which fungal growth was inhibited.

Funding

This work was supported by the Outstanding Youth Foundation of Heilongjiang Province (No. YQ2021H009).

Supporting Information

HRMS and NMR spectra of compounds **1**–**2**, **4**–**5**, and X-ray files for compound **1** (CIF) can be obtained by contacting the corresponding authors *via* E-mail.

Declaration of Competing Interest

The authors reported no potential conflicts of interest.

References

1 Zhang HW, Song YC, Tan RX. Biology and chemistry of endophytes. *Nat Prod Rep*. 2006;23(5):753-771. https://doi.org/10.1039/b609472b.
2 Singh A, Singh DK, Kharwar RN, et al. Fungal endophytes as efficient sources of plant-derived bioactive compounds and their prospective applications in natural product drug discovery: insights, avenues, and challenges. *Microorganisms*. 2021;9(1):197. https://doi.org/10.3390/microorganisms9010197.
3 Liu ZG, Sun Y, Tang MY, et al. Trichodestruxins A–D: cytotoxic cyclodepsipeptides from the endophytic fungus *Trichoderma harzianum*. *J Nat Prod*. 2020;83(12):3635-3641. https://doi.org/10.1021/acs.jnatprod.0c00808.
4 Gakuubi MM, Munusamy M, Liang ZX, et al. Fungal endophytes: a promising frontier for discovery of novel bioactive compounds. *J Fungi*. 2021;7(10):786. https://doi.org/10.3390/jof7100786.
5 Cheng JT, Wang HM, Yu JH, et al. Discovery of a potential liver fibrosis inhibitor from a mushroom endophytic fungus by genome mining of a silent biosynthetic gene cluster. *J Agric Food Chem*. 2021;69(38):11303-11310. https://doi.org/10.1021/acs.jafc.1c03639.
6 Morishita Y, Tsukada K, Murakami K, et al. Synthetic biology-based discovery of diterpenoid pyrones from the genome of *Eupenicillium shearii*. *J Nat Prod*. 2022;85(2):384-390. https://doi.org/10.1021/acs.jnatprod.1c00973.
7 Shen QY, Dai GZ, Li AY, et al. Genome-guided discovery of highly oxygenated aromatic polyketides, saccharothrixins D–M, from the rare marine Actinomycete *Saccharothrix* sp. D09. *J Nat Prod*. 2021;84(11):2875-2884. https://doi.org/10.1021/acs.jnatprod.1c00617.
8 Schwartz RE, Smith SK, Onishi JC, et al. The isolation and structure determination of enfumafungin, a triterpene glycoside antifungal agent from the fermenation of a *Hormonema* sp.. *J Am Chem Soc*. 2000;122:4882-4886.
9 Kuhnert E, Li Y, Lan N, et al. Enfumafungin synthase represents a novel lineage of fungal triterpene cyclases. *Environ Microbiol*. 2018;20(9):3325-3342. https://doi.org/10.1111/1462-2920.14333.
10 Li XY, Lv JM, Cao ZQ, et al. Biosynthetic characterization of the antifungal fernane-type triterpenoid polytolypin for generation of new analogues *via* combinatorial biosynthesis. *Org Biomol Chem*. 2023;21(4):851-857. https://doi.org/10.1039/d2ob02158g.
11 Li SF, Yu XQ, Li YL, et al. Vibsane-type diterpenoids from *Viburnum odoratissimum* and their cytotoxic activities. *J Bioorg Chem*. 2021;106:104498. https://doi.org/10.1016/j.bioorg.2020.104498.
12 Bai M, Xu W, Zhang X, et al. HSQC-based small molecule accurate recognition technology discovery of diverse cytotoxic sesquiterpenoids from *Elephantopus tomentosus* L. and structural revision of molephantins A and B. *Phytochemistry*. 2023;206:113562. https://doi.org/10.1016/j.phytochem.2022.113562.
13 Shigematsu N, Tsujii E, Kayakiri N, et al. WF11605, an antagonist of leukotriene B4 produced by a fungus II. Structure determination. *J Antibiot*. 1992;45(5):704-708. https://doi.org/10.7164/antibiotics.45.704.
14 Kaur R, Macleod J, Foley W, et al. Gluconic acid: and antifungal agent produced by *Pseudomonas* species in biological control of take-all. *Phytochemistry*. 2006;67(6):595-604. https://doi.org/10.1016/j.phytochem.2005.12.011.
15 Tan X, Zhang X, Yu M, et al. Sesquiterpenoids and mycotoxin swainsonine from the locoweed endophytic fungus *Alternaria oxytropis*. *Phytochemistry*. 2019;164:154-161. https://doi.org/10.1016/j.phytochem.2019.05.012.
16 Brauers G, Edrada RA, Ebel R, et al. Anthraquinones and betaenone derivatives from the sponge-associated fungus *Microsphaeropsis* species: novel inhibitors of protein kinases. *J Nat Prod*. 2000;63(6):739-745. https://doi.org/10.3168/jds.S0022-0302(01)74580-8.
17 Schlingmann G, Milne L, Carter GT. Isolation and identification of antifungal polyesters from the marine fungus *Hypoxylon oceanicum* LL-15G256. *Tetrahedron*. 2002;58(34):6825-6835. https://doi.org/10.1016/S0040-4020(02)00746-9.
18 Xiao Y, Liang W, Zhang Z, et al. Polyketide derivatives from the endophytic fungus *Phaeosphaeria* sp. LF5 isolated from *Huperzia serrata* and their acetylcholinesterase inhibitory activities. *J Fungi*. 2022;8(3):232. https://doi.org/10.3390/jof8030232.
19 Gao J, Li XM, Li X, et al. New lactone and isocoumarin derivatives from the marine mangrove-derived endophytic fungus *Penicillium coffeae* MA-314. *Phytochem Lett*. 2019;32:1-5. https://doi.org/10.1016/j.phytol.2019.04.018.
20 Dolomanov OV, Bourhis LJ, Gildea RJ, et al. OLEX2: a complete structure solution, refinement and analysis program. *J Appl Cryst*. 2009;42(2):339-341. https://doi.org/10.1515/ncrs-2023-0461.
21 Sheldrick GM. SHELXT-Integrated space-group and crystal-structure determination. *Acta Cryst*. 2015;A71:3-8. https://doi.org/10.1107/S2053273314026370.
22 Sheldrick GM. Crystal structure refinement with SHELXL. *Acta Cryst*. 2015;C71:3-8. https://doi.org/10.1107/S2053229614024218.