JOURNAL OF NATURAL PRODUCTS

pubs.acs.org/jnp

Article

Species Prioritization Based on Spectral Dissimilarity: A Case Study of Polyporoid Fungal Species

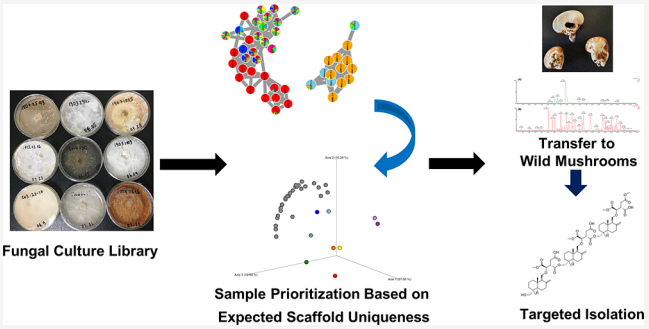
Huong T. Pham, Kwang Ho Lee, Eunah Jeong, Sunmin Woo, Jinsuh Yu, Woo-Young Kim, Young Woon Lim, Ki Hyun Kim,* and Kyo Bin Kang*

Cite This: https://dx.doi.org/10.1021/acs.jnatprod.0c00977

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

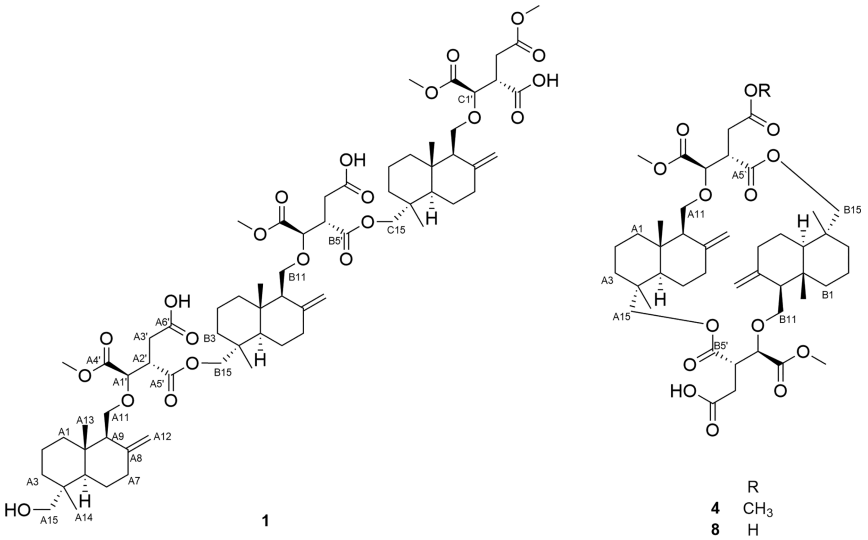
ABSTRACT: Biological species collections are critical for natural product drug discovery programs. However, prioritization of target species in massive collections remains difficult. Here, we introduce an untargeted metabolomics-based prioritization workflow that uses MS/MS molecular networking to estimate scaffold-level distribution. As a demonstration, we applied the workflow to 40 polyporoid fungal species. Nine species were prioritized as candidates based on the chemical structural and compositional similarity (CSCS) metric. Most of the selected species showed relatively higher richness and uniqueness of metabolites than those of the others. *Cryptoporus volvatus*, one of the prioritized species, was investigated further. The chemical profiles of the extracts of *C. volvatus* culture and fruiting bodies were compared, and it was shown that derivative-level diversity was higher in the fruiting bodies; meanwhile, scaffold-level diversity was similar. This showed that the compounds found from a cultured fungus can also be isolated in wild mushrooms. Targeted isolation of the fruiting body extract yielded three unknown (**1**−**3**) and six known (**4**−**9**) cryptoporic acid derivatives, which are drimane-type sesquiterpenes with isocitric acid moieties that have been reported in this species. Cryptoporic acid T (**1**) is a trimeric cryptoporic acid reported for the first time. Compounds **2** and **5** exhibited cytotoxicity against HCT-116 cell lines with IC_{50} values of 4.3 and 3.6 μM, respectively.



Fungi are talented producers of specialized metabolites. Structurally diverse metabolites have been isolated from fungi,[1] and many have been important for natural product-derived drug discovery. Fungal fruiting bodies, commonly known as mushrooms, have been consumed as nutritional and medicinal products.[2] Harvesting wild mushrooms is a challenging process. In many cases, fungal fruiting bodies are short-lived, appearing overnight after rain and then disappearing within a few days. Such life cycles often make collecting and further chemical investigation or bioactivity screening of fruiting bodies difficult. The possibly devastating effects of collection on biodiversity is also an important and controversial concern.[3] Thus, mycelial culture has been suggested as an alternative solution for biological and chemical studies of macro-fungi, although the chemical profiles of fruiting bodies and mycelial cultures are often different.[4,5]

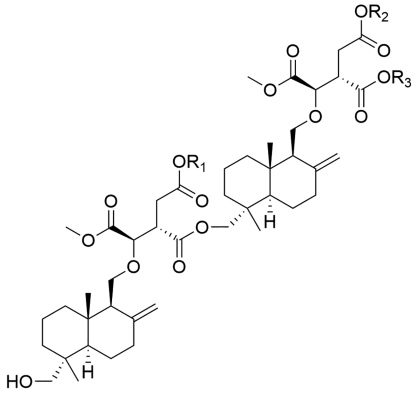
As structure-based discovery has risen as an alternative approach for natural product discovery,[6] collections of source organisms are becoming increasingly important. In structure-based workflows, the targeting of organisms for isolation is prioritized based on data related to their metabolic phenotypes and potential, which are typically provided by untargeted metabolomics or genome mining efforts. Thus, the larger collections generally allow for richer information gathering. However, a massive data set also creates a dilemma; time and effort for data analysis will be dramatically increased, which will reduce the efficiency of data-driven prioritization. Thus, the need for automated and comprehensive data analysis platforms in the dereplication of natural products to solve this problem is increasing. MS/MS molecular networking,[7] which is publicly available via the GNPS web platform,[8] has become one of the most popular data analysis workflows for MS-based untargeted metabolomics approaches in natural product discovery. MS/MS molecular networking clusters similar spectra into molecular families, of which spectral node members are assumed to have structural similarities; for natural products, structural similarities often occur between metabolites sharing the same scaffolds. Many natural product studies have utilized MS/MS molecular networking for target prioritization,[9] but their selection criteria were different from each other. Nevertheless, the prioritization step in most of these studies has a common fact: They primarily relied on manual

Received: September 8, 2020

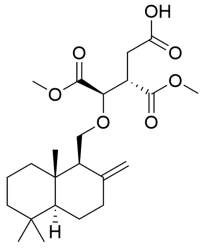


1

	R
4	CH_3
8	H



	R_1	R_2	R_3
2	H	CH_3	CH_3
3	H	H	$CH(CH_3)_2$
5	CH_3	H	CH_3
6	H	H	CH_3
7	CH_3	CH_3	CH_3



9

inspection of each molecular family and spectral annotation, which requires considerable time for the analysis of large data sets built from large collections.

Here, we introduce an untargeted metabolomics-based prioritization pipeline focusing on scaffold-level uniqueness of compounds. The Chemical Structural and Compositional Similarity (CSCS) metric is a distance metric developed by Sedio et al. as an attempt to take account of the chemical structural similarity between metabolites.[10] However, the exact chemical similarity can only be calculated with accurate identification of metabolites, so the metric estimates structural similarity based on spectral similarity measured during molecular networking. Previously, it was shown that the CSCS metric could discriminate samples based on their different chemical compositions, mainly focusing on distributions of molecular families.[11] We applied this statistical method to an LC−MS/MS data set acquired from multiple fungal isolates to highlight species that produce metabolites clustered in species-specific molecular families. As an example, 40 species of the family Polyporaceae, a family of poroid fungi containing well-known medicinal mushrooms, such as *Wolfiporia cocos*, *Polyporus umbellatus*, and *Trametes versicolor*, were selected as candidate taxa. The chemical profiles of extracts from mycelial cultures and from fruiting bodies of *Cryptoporus volvatus* were compared to determine whether the chemical uniqueness retrieved from a cultured fungus can also be found in wild mushrooms. Further results for validation of the workflow, such as chemical dereplication, targeted isolation, and biological evaluation, are also described here.

RESULTS AND DISCUSSION

Extracts from 40 fungi cultured on PDA plates were analyzed using LC−MS/MS, and then the data were analyzed by feature-based molecular networking workflow.[12] As a result, 5100 MS features were extracted from the entire data set, and 2240 MS features were grouped into 751 molecular families; the remaining 2800 MS features were singletons (nodes not having any molecular relatives) (Figure S1, Supporting Information). In the principal coordinate analysis (PCoA) plot based on the weighted (by ion intensities of MS features) CSCS metric, nine species, *Cryptoporus volvatus* (KMRB16041505), *Daedaleopsis tricolor* (KMRB15061217), *Lopharia cinerascens* (KMRB15090836), *Microporus vernicipes* (KMRB15071608), *Phlebiopsis gigantea* (KMRB17083132), *Polyporus brumalis* (KMRB15102011), *Skeletocutis nivea* (KMRB15090322), *Trametes versicolor* (KMRB15032003), and *Tyromyces chioneus* (KMRB14070116), showed relatively high dissimilarity from the other strains (Figure 1A). This finding suggests that the chemical compositions of these species are expected to be unique within the set of analyzed

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

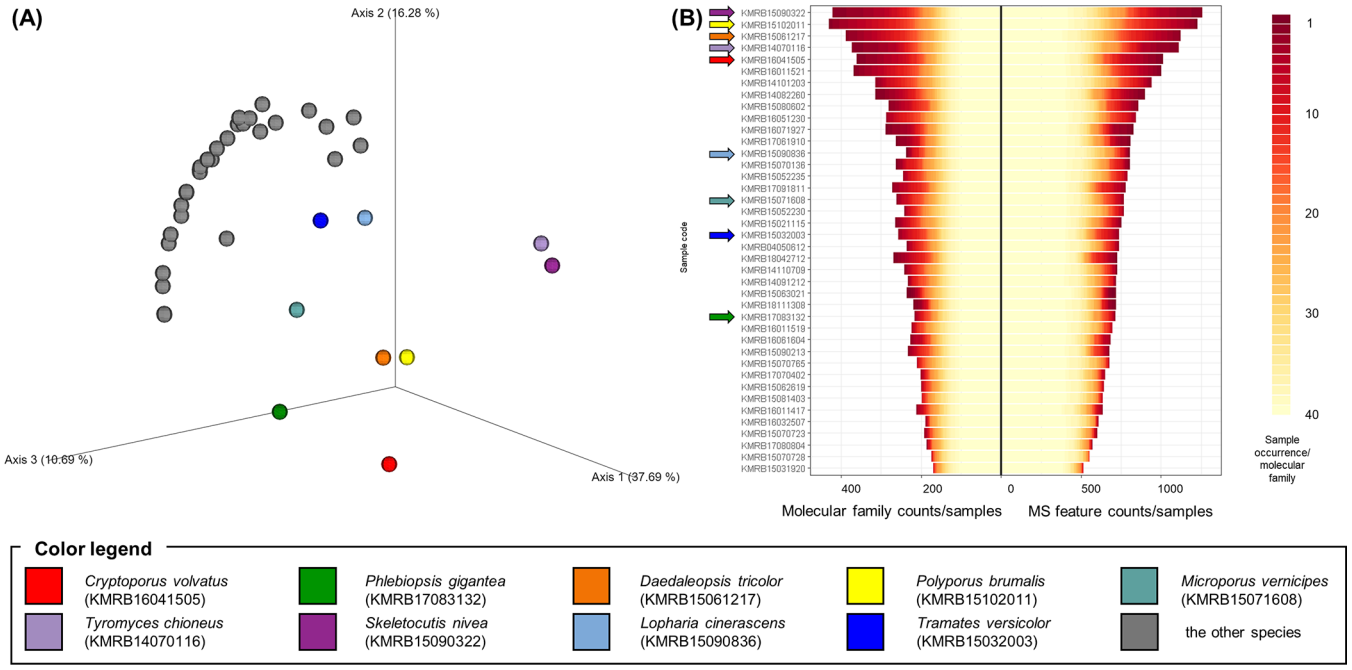


Figure 1. Chemistry-based species prioritization using 40 polyporoid culture extracts. (A) The PCoA plot based on the weighted CSCS metric highlights the chemical dissimilarity between samples. Nine selected samples showing high dissimilarity from the others are visualized with colors. (B) The numbers of molecular families and MS features in each sample are visualized. The molecular network organizes MS features into molecular families, and the color of the bar illustrates the number of samples each molecular family occurs in. The selected samples captioned in part A are highlighted with colored arrows representing each species.

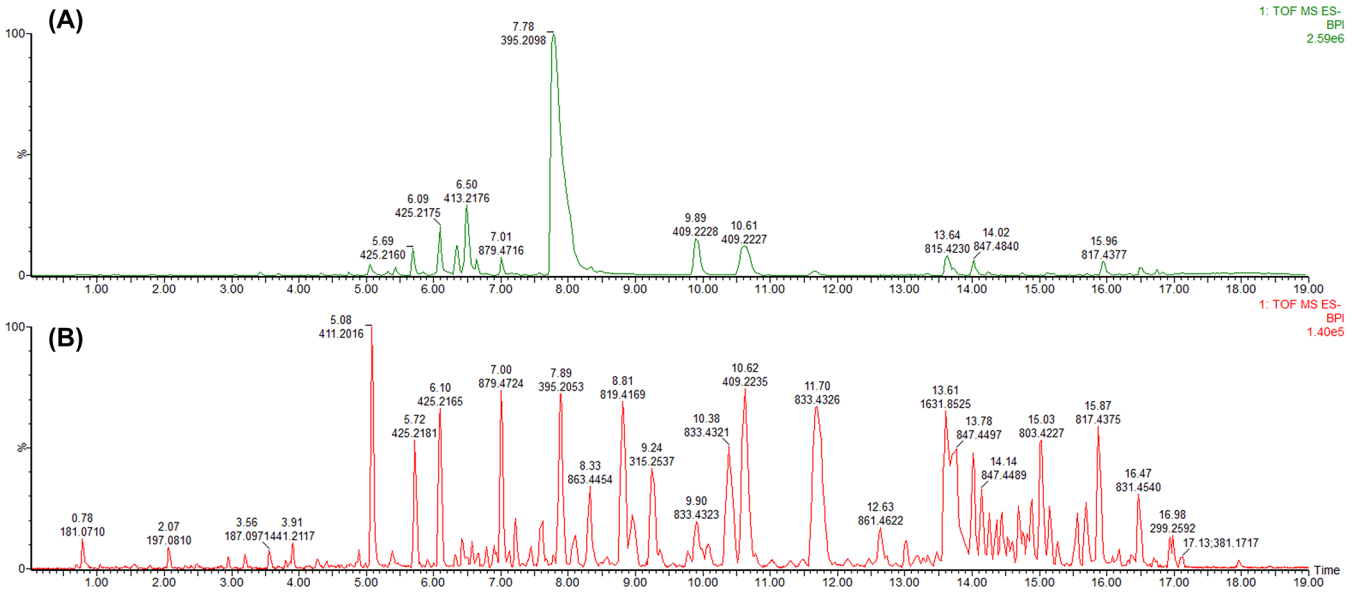


Figure 2. LC−MS base peak ion (BPI) chromatograms of (A) *C. volvatus* culture (KMRB16041505) extract and (B) *C. volvatus* fruiting body extract. Numbers above chromatographic peaks indicate retention times (min) and *m*/*z* values of the base peak ions.

samples, especially at the molecular family level, because the distance between samples were weighted by spectral dissimilarity. To further investigate the chemical uniqueness of each species, we visualized MS feature counts for each sample (Figure 1B). Additionally, we added a color gradient to each MS feature to illustrate the number of samples each molecular family occurs in (the counting method is exemplified in Figure S2, Supporting Information). This plot revealed that most of the candidate species were high in both the richness (number of different chemicals) and uniqueness of metabolites. However, there are exceptional cases such as *P. gigantea* (KMRB17083132) and *T. versicolor* (KMRB15032003), and the molecular network was manually inspected to determine the reason. *T. versicolor*, *M. vernicipes*, and *L. cinerascens* contained unique molecular families, but the size of these molecular families was small. *P. gigantea* metabolites were presented in other species, but their relative ion intensities were much higher in *P. gigantea* (Figure S1, Supporting Information). Thus, it could be supposed that *P. gigantea* was different from other species quantitatively, rather than

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

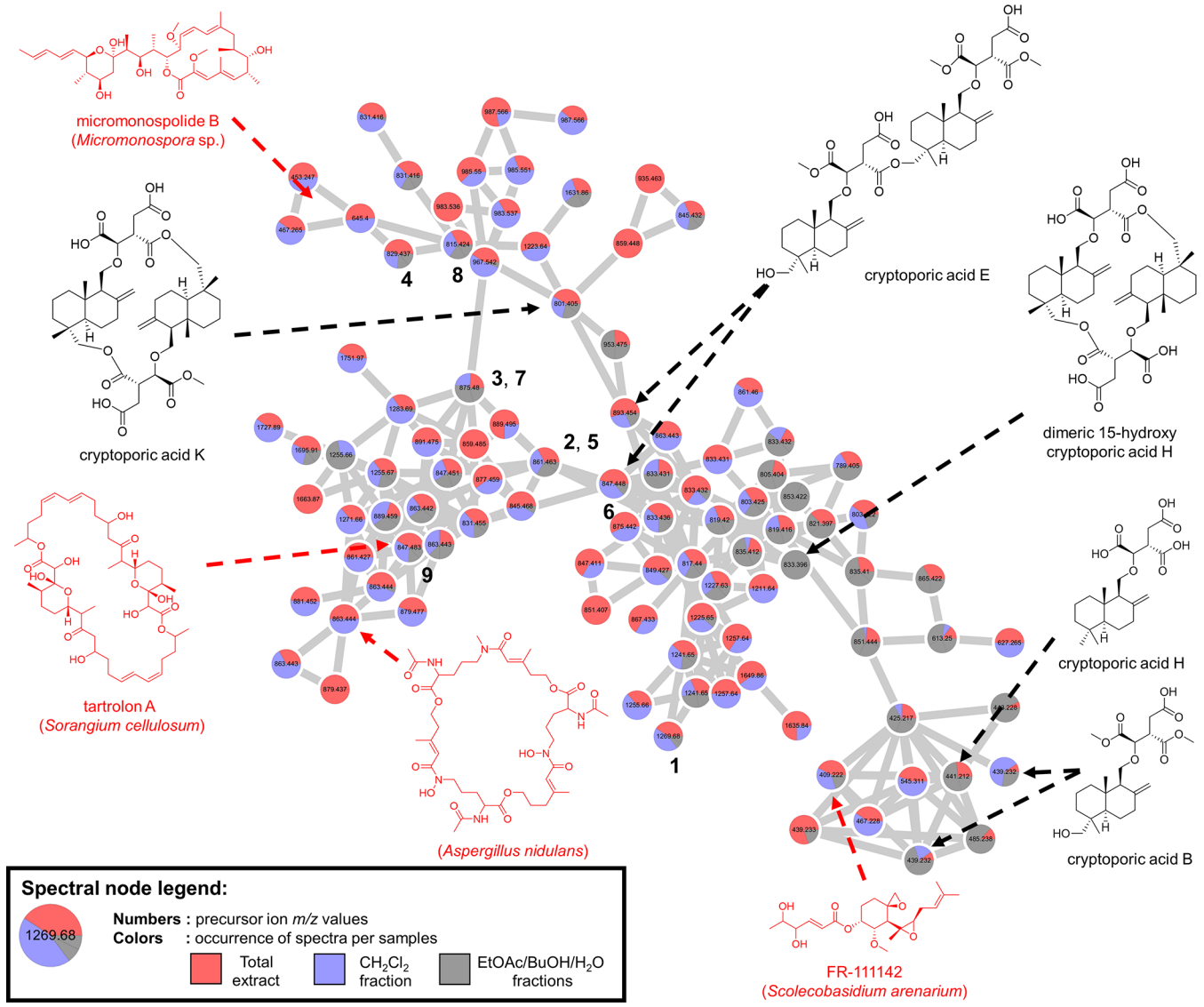


Figure 3. Dereplication and targeted isolation of cryptoporic acid derivatives from the *C. volvatus* fruiting body extract using MS/MS molecular networking. The structures shown here denote the candidate structures suggested by the NAP *in silico* annotation. Red-colored candidates were discarded, mainly because the producing organisms (shown with structures) are phylogenetically distant from *C. volvatus*. From these annotations, it was concluded that the spectra in this molecular family represent cryptoporic acid derivatives.

qualitatively. The weighted CSCS metric was used here, so ion intensity of each feature also largely affects the distances between samples. These showed that MS/MS molecular networking and further chemical dissimilarity-based statistical analysis can provide a comparative overview on multiple metabolomes, which could be further facilitated for structure-based sample prioritization.

Among the nine fungi prioritized, only two species, *C. volvatus* and *T. versicolor*, have been investigated for their pharmacological activity.[13−15] We selected *C. volvatus* as a subject for further chemical investigation focusing on its antiviral and chemopreventive effect against carcinogenesis, both of which were observed *in vivo*.[14,15] Comparison of the LC−MS data revealed that most of the major peaks observed in the culture extract were also found in the fruiting body extract, and the fruiting bodies had a more complex chemical profile than those of the cultured fungi (Figure 2). The molecular networking analysis of these two extracts revealed that most chromatographic peaks observed in the fruiting body extract were clustered together with the major peaks of the culture extract, which implies structural similarity (Figure S3, Supporting Information). Thus, it was assumed that the fruiting body extract would yield more diverse derivatives with scaffolds similar to the metabolites in the culture extract.

The fruiting body extract was further fractionated by liquid−liquid extraction, and the chemical compositions of the crude extract and fractions were analyzed using MS/MS molecular networking. Most spectral nodes were clustered into a sizable molecular family, which suggested their structural similarity (Figure S4, Supporting Information). None of the spectral nodes in this molecular family were annotated through GNPS spectral library matching, so they were given *in silico* annotations by applying Network Annotation Propagation (NAP)[16] with the structural library of microbial natural products available from the NPAtlas.[17] The NAP annotation provided candidate structures with different scaffolds of macrolides and sesquiterpenoids. We manually inspected which species were previously reported as producers of

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

Table 1. ^{1}H and ^{13}C NMR Spectroscopic Data of Compounds 1−4 (500/125 MHz)

	1[a]		2[a]		3[a]		4[b]	
position	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
A1	38.5	1.62, m	38.5	1.62, m	38.8	1.68, m	40.4	1.75, m
		1.14, m				1.17, m		1.26, m
A2	18.5	1.53, m	18.6	1.62, m	18.7	1.56, m	19.5	1.66, m
				1.27, m				1.60, m
A3	35.3	1.56, m	35.4	1.66, m	35.4	1.45, m	36.7	1.74, m
		1.26, m		1.26, m		1.28, m		1.24, m
A4	38.0		38.8		38.0		38.0	
A5	48.0	1.45, m	48.2	1.44, m	48.2	1.46, m	48.0	1.44, m
A6	23.5	1.61, m	23.7	1.43, m	23.7	1.60, m	24.3	1.54, m
		1.28, m		1.33, m		1.34, m		1.25, m
A7	37.1	2.35, m	37.4	2.36, m	37.5	2.36, m	38.9	2.19, m
		1.93, m		2.04, m		2.06, m		2.05, m
A8	146.5		146.6		146.7		148.6	
A9	54.9	1.94, m	55.7	1.99, m	55.7	2.01, d (4.3)	58.1	2.02, m
A10	38.4		38.0		146.7		39.8	
A11	68.5	3.91, m	68.4	3.90, m	68.6	3.89, dd (9.5, 8.1)	70.5	4.08, ddd (9.5, 6.8, 2.9)
		3.56, m		3.54, m		3.58, m		3.43, ddd (9.5, 5.8, 4.0)
A12	108.3	4.85, s	108.3	4.85, d (1.1)	108.2	4.85, d (1.1)	109.0	4.90, s
		4.71, s		4.73, s		4.72, s		4.80, s
A13	15.9	0.71, s	16.0	0.76, s	16.0	0.77, s	16.0	0.76, s
A14	17.8	0.74, s	17.8	0.75, s	17.8	0.76, s	18.3	0.79, s
A15	72.0	3.41, m	72.1	3.40, m	72.2	3.42, m	72.1	4.37, d (5.0)
		3.08, d (11.0)		3.09, d (11.0)		3.11, d (10.9)		3.05, m
A1′	78.8	4.12, d (4.9)	78.8	4.10, d (5.2)	79.0	4.13, d (5.4)	80.2	4.14, d (3.6)
A2′	44.4	3.48, m	44.6	3.45, m	44.3	3.49, m	46.2	3.50, m
A3′	32.0	2.87, m	31.8	2.83, dd (17.6, 10)	32.4	2.84, m	33.7	3.00, m
		2.62, dd (17.7, 3.5)		2.59, m		2.63, dd (17.8, 2.1)		2.71, dd (17.0, 5.6)
A4′	171.0		170.9		170.9		172.8	
A5′	170.9		171.0		171.0		171.4	
A6′	176.4		174.4		178.9		173.9	
A4′-OMe	52.5	3.68, s	52.2	3.76, s	52.3	3.76, s	52.3	3.78, s
B1	38.6	1.61, m	38.5	1.61, m	38.3	1.58, m	40.4	1.75, m
		1.14, m		1.52, m		1.10, m		1.26, m
B2	18.7	1.57, m	18.5	1.50, m	18.4	1.52, m	19.5	1.66, m
				1.16, m				1.60, m
B3	35.5	1.53, m	35.5	1.53, m	35.2	1.51, m	36.6	1.72, m
		1.26, m		1.27, m		1.21, m		1.66, m
B4	37.5		38.7		38.0		37.9	
B5	47.5	1.28, m	47.5	1.37, m	46.6	1.26, m	48.0	1.44, m
B6	23.6	1.61, m	23.7	1.60, m	23.4	1.70, m	24.3	1.54, m
		1.31, m		1.33, m		1.26, m		1.25, m
B7	37.4	2.35, m	37.1	2.30, m	37.1	2.37, m	38.9	2.17, m
		2.06, m		1.97, m		1.89, m		2.05, m
B8	146.2		146.6		146.3		148.6	
B9	55.0	1.88, m	54.8	2.03, m	55.1	1.82, d (7.6)	58.1	2.02, m
B10	38.7		37.3		38.4		39.8	
B11	67.9	3.93, m	68.1	3.90, m	67.7	3.96, t (9.6)	70.5	4.08, ddd (9.5, 6.8, 2.9)
		3.58, m		3.54, m		3.60, m		3.43, ddd (9.5, 5.8, 4.0)
B12	108.6	4.88, s	108.2	4.83, s	108.5	4.88, s	109.0	4.80, s
		4.75, s		4.70, s		4.75, s		4.90, s
B13	16.0	0.76, s	17.8	0.77, s	16.0	0.71, s	16.0	0.76, s
B14	14.8	0.76, s	15.9	0.72, s	18.0	0.75, s	18.3	0.79, s
B15	72.1	4.03, d (2.8)	72.7	4.03, d (11)	71.0	4.09, d (11.3)	72.1	4.39, d (5.0)
		3.58, m		3.53, m		3.66, d (11.3)		3.02, m
B1′	77.4	4.27, d (4.3)	78.3	4.10, d (5.2)	76.9	4.36, d (4.3)	80.3	4.16, d (3.6)
B2′	43.9	3.45, m	44.6	3.39, m	43.6	3.42, m	46.2	3.50, m
B3′	31.6	2.82,m	32.1	2.78, m	31.9	2.86, m	33.7	3.00, m
		2.74, m		2.59, m		2.84, m		2.71, dd (16.7, 5.6)
B4′	171.3		171.3		171.2		172.9	

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

Table 1. continued

	1[a]		2[a]		3[a]		4[b]	
position	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
B5′	171.1		172.0		169.8		171.8	
B6′	178.6		172.6		179.6		173.9	
B4′-OMe	52.3	3.76, s	52.2	3.76, s	52.1	3.75, s	52.4	3.78, s
B5′-OMe			52.3	3.68, s				
B6′-OMe			52.6	3.68, s			52.4	3.73, s
B7′					69.2	5.01, m		
B7′-Me					21.7	1.20, d (6.2)		
B7′-Me					21.9	1.22, d (6.2)		
C1	38.8	1.68, m						
		1.17, m						
C2	18.5	1.52, m						
C3	35.5	1.45, m						
		1.26, m						
C4	37.5							
C5	47.3	1.32, m						
C6	23.7	1.61, m						
		1.33, m						
C7	37.4	2.35, m						
		2.06, m						
C8	146.6							
C9	55.8	1.99, m						
C10	38.4							
C11	68.0	3.93, m						
		3.58, m						
C12	108.2	4.85, s						
		4.74, s						
C13	15.9	0.72, s						
C14	17.9	0.77, s						
C15	72.3	4.05, d (2.7)						
		3.55, m						
C1′	77.9	4.20, d (5.1)						
C2′	44.1	3.48, m						
C3′	32.4	2.85, m						
		2.72, m						
C4′	171.3							
C5′	177.4							
C6′	171.1							
C4′-OMe	52.4	3.77, s						
C6′-OMe	52.3	3.76, s						

[a]In $CDCl_3$. [b]In CD_3OD.

candidate structures through NPAtlas and discarded annotations when source species were taxonomically far from *C. volvatus*, considering them as false positives (red colored structures in Figure 3). After filtering the annotated structures based on phylogeny of source organisms, cryptoporic acid derivatives were the major scaffolds left, which suggested that the spectra in the molecular family represent cryptoporic acid derivatives (black colored structures in Figure 3). Although 42 cryptoporic acid derivatives including dimers have been reported[18−24] since the initial isolation of cryptoporic acids A and B,[25] this metabolite family have been reported from only *C. volvatus*, *Fomitella fraxinea*,[22] and *Polyporus ciliatus*.[26] Thus, the *in silico* annotation result demonstrated that the untargeted metabolomics workflow described here can be used for identifying strains producing scaffolds unique to individual species. From the dereplication results, some spectral nodes from unknown derivatives did not provide any *in silico* candidates, and these were prioritized for isolation. Many spectra showed relatively high ion intensities in the CH_2Cl_2 fraction, so targeted isolation was performed with this fraction to yield compounds **1**−**9**. Among these, compounds **5**−**9** were identified as 6′,6‴-cryptoporic acid G dimethyl ester (**5**), cryptoporic acid E (**6**), cryptoporic acid E pentamethyl ester (**7**), crytoporic acid D (**8**), and cryptoporic acid A (**9**) by comparing their NMR and optional rotation data with those reported in the literature.[18,24]

Compound **1** was prioritized because of its deprotonated molecular ion peak at HRESIMS m/z 1255.6630 $[M - H]^-$ (calcd for $C_{67}H_{99}O_{22}$, 1255.6628), which suggested that compound **1** is a trimeric derivative of cryptoporic acid. The ^{1}H NMR data of **1** (Table 1) showed the presence of six methyl (δ_H 0.71, 0.72, 0.74, 0.76, 0.76, and 0.77), four methyl ester (δ_H 3.68, 3.76, 3.76, and 3.77), and three pairs of exomethylene (δ_H 4.71 and 4.85, 4.75 and 4.88, and 4.74 and 7.85) groups. The ^{13}C NMR spectrum showed nine carbonyl groups (δ_C 170.9, 171.0, 171.1, 171.1, 171.3, 171.3, 171.4,

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

176.4, and 178.6), indicating that compound **1** is a trimeric cryptoporic acid derivative. The $^{1}H-^{1}H$ COSY spectrum confirmed the spin systems of the three isocitric acid moieties (Figure 4). The HMBC correlations from H-A1′ (δ_H 4.12) to

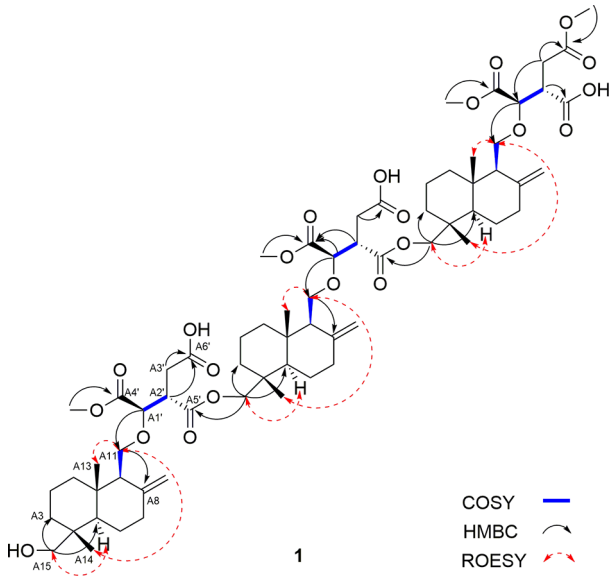


Figure 4. Key $^{1}H-^{1}H$ COSY, HMBC, and ROESY correlations of compound **1**.

C-A11 (δ_C 68.5), from H-B1′ (δ_H 4.27) to C-B11 (δ_C 67.9), and from H-C1′ (δ_H 4.20) to C-C11 (δ_C 68.0) established the ether linkages between the sesquiterpene and isocitric moieties of each monomer, while the HMBC correlations from H-B15 (δ_H 4.03) to C-A5′ (δ_C 170.9) and from H-C15 (δ_H 4.05) to C-B5′ (δ_C 171.1) demonstrated the connections between three monomeric moieties through the ester bonds (Figure 4). The positions of the methyl ester groups were determined to be C-A4′ (δ_C 171.0), C-B4′ (δ_C 171.3), C-C4′ (δ_C 171.3), and C-C6′ (δ_C 171.1) by the HMBC experiment. The relative configurations of the sesquiterpene moieties were established based on the ROESY correlations between H-A11 and H-A13/H-A14, H-B11 and H-B13/H-B14, H-C11a and H-C13/H-C14, H-A5 and H-A15, H-B5 and H-B15, and H-C5 and H-C15 (Figure 4). The absolute configurations of the isocitric acid moieties were proposed to be either (1′*R*, 2′*S*) or (1′*S*, 2′*R*) for all the three moieties by comparing ^{1}H NMR chemical shifts of H-12, H-1′, and H-2′ with four synthesized diastereomers of cryptoporic acid A methyl ester.[27] The absolute configurations of the isocitric acids were finally confirmed as (1′*R*, 2′*S*) by comparing the optical rotation with previously reported values of **6**.[18] Based on the above data, the structure of **1** was determined and named cryptoporic acid T.

Compound **2** was inferred to be a dimeric cryptoporic acid with the molecular formula $C_{46}H_{70}O_{15}$, which was suggested by its deprotonated molecular ion at HRESIMS *m*/*z* 861.4651 $[M - H]^-$ (calcd for $C_{46}H_{69}O_{15}$, 861.4642). The ^{1}H and ^{13}C NMR spectra of **2** were similar to those of **5** (Table 1).[24] Molecular formulas of **2** and **5** were assigned to be identical based on their MS data, but the difference in their HPLC retention times suggested that the structures of **2** and **5** are different. The HMBC spectrum of **2** revealed the position of the methyl ester groups as C-A4′ (δ_C 170.9), C-B4′ (δ_C 171.3), C-B5′ (δ_C 172.0), and C-B6′ (δ_C 172.6). The relative configurations of sesquiterpene moieties were identified by the ROESY spectrum. The absolute configuration of the isocitric moieties of **2** was determined to be (A1′*R*, A2′*S*, B1′*R*, B2′*S*), based on the similarities of the ^{1}H NMR and optical rotation to those of **5**. Thus, the structure of compound **2** was established to be 5‴,6‴-cryptoporic acid G dimethyl ester.

Compound **3** was also suggested to be a dimer with a molecular formula of $C_{47}H_{72}O_{15}$, according to its deprotonated molecular ion at HRESIMS *m*/*z* 875.4782 $[M - H]^-$ (calcd for $C_{47}H_{71}O_{15}$, 875.4799). The NMR data (Table 1) of **3** were similar to those of **5**,[24] except for one more isopropyl group at the C-B5′ position. The isopropyl group was established based on the $^{1}H-^{1}H$ COSY correlation between a methine proton (δ_H 5.0) and two doublet methyl signals (δ_H 1.20 and 1.22). HMBC confirmed the connection between the isopropyl group H-B7′ (δ_H 5.01) and C-B5′ (δ_C 169.8). The relative configuration was defined based on the ROESY spectrum, as in **2**. The absolute configuration of **3** in the isocitric was

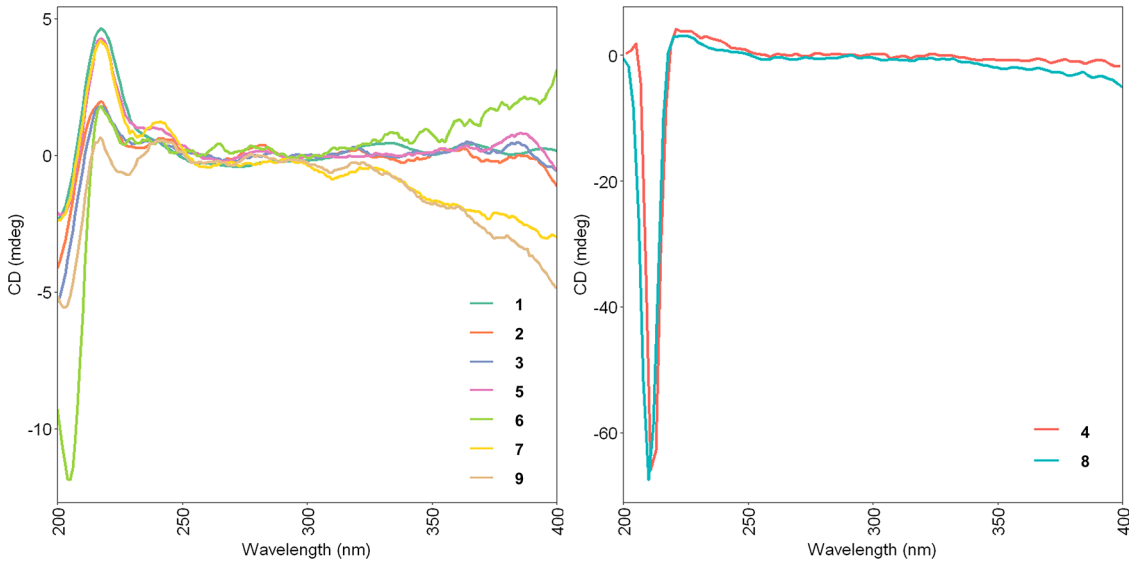


Figure 5. Experimental ECD spectra of compounds **1**–**9**.

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

determined to be (A1′*R*, A2′*S*, B1′*R*, B2′*S*) in the same manner as compound **2**. Accordingly, compound **3** was identified as 5″-cryptoporic acid E isopropyl ester. Although this compound was detected in the LC−MS analysis of the crude extract, we propose that compound **3** is an experimental artifact derived from residual isopropyl alcohol during the extraction process.[28]

Compound **4** was a dimeric cryptoporic acid analogue with a molecular formula of $C_{45}H_{66}O_{14}$, which was suggested by the deprotonated ion at *m*/*z* 829.4370 $[M - H]^-$ (calcd for $C_{45}H_{65}O_{14}$, 829.4380). The NMR spectroscopic data (Table 1) suggested that compound **4** is a derivative of cryptoporic acid D,[18] which is a symmetric dimer. However, some resonances such as H-3, H-7, and H-15 did not fully overlap, which suggests that compound **4** is not a fully symmetric analogue. The relative intensities of the ^{1}H NMR peaks at δ_H 3.78 (s, 6H) and 3.73 (s, 3H) suggested that compound **4** possesses three methyl ester groups. The HMBC data confirmed that compound **4** is a cyclic dimer showing correlations from H-A15 to C-B5′ and from H-B15 to C-A5′. The methyl ester groups were located at C-A4′, C-A6′, and C-B4′ based on the HMBC data. The absolute configuration of **4** was identified by the same method used for **1**−**3**. Consequently, the structure of **4** was established, as shown. This structure was reported previously;[29] however, the reported spectral data are not identical to our data. We suppose that the compound reported by Meng and others had a different structure, but unfortunately, raw NMR spectra were not provided, so we could not revise their structural identification. Meng et al. named this compound crypotoporic acid J, but this name was also used for another monomeric compound in an article published at approximately the same time.[23] To avoid confusion, we suggest that compound **4** be named cryptoporic acid D trimethyl ester.

Electronic circular dichroism (ECD) analysis was performed to provide further evidence for absolute configurations of isolated compounds. Compounds **1**−**3**, **5**−**7**, and **9** showed similar ECD spectra with a weak negative Cotton effect at 204 nm and weak positive Cotton effects at 218 and 240 nm (Figure 5). This result supported that all of these compounds have identical absolute configurations. On the other hand, two cyclic dimeric derivatives, **4** and **8**, showed strong negative Cotton effects at 212 nm (Figure 5). 3D structural models of cyclic dimers were different from those of acyclic dimers (Figure S5, Supporting Information), and this spatial difference causes the difference in the ECD spectra. The results described here should facilitate the use of ECD analysis for absolute configuration determination in further studies on cryptoporic acid derivatives.

C. volvatus is known in Korean folk medicine to be useful for cancer treatment, and cryptoporic acids D and E are reported to exhibit tumor-suppressive effects *in vivo*.[15,30] Compounds **1**−**9** were evaluated for their cytotoxicity against the human colon cancer cell line, HCT-116. Compounds **2** and **5** exhibited moderate cytotoxicity with IC_{50} values of 4.3 and 3.6 μM, respectively, while the other compounds showed weak or no cytotoxicity (paclitaxel, the positive control, showed an IC_{50} value of 7.7 nM; Table S3, Supporting Information).

Here, we introduced an untargeted metabolomics-based prioritization pipeline for selecting target strains from sample collections with a focus on the uniqueness of scaffold structures, taking advantage of structural similarities predicted by MS/MS molecular networking. In this study, the workflow analyzed a fungal extract collection, but this pipeline can be used for other types of natural product libraries, such as plant or bacterial extract collections. Although the case study on *C. volvatus* did not yield compounds with a novel scaffold, it validated that this workflow can highlight species-specific metabolite producers while minimizing efforts during spectral investigations. Moreover, the case study also showed that the workflow could be applied to the transition from fungal isolate cultures to wild mushrooms, which will be beneficial for future fungal natural product studies. Freely available bioinformatics platforms that do not require advanced computational skills, such as GNPS,[8] AntiSMASH,[31] and BiG-SCAPE,[32] are encouraging the application of computational tools in natural product discovery studies. Among these, GNPS and related advances such as bioactive molecular networking or *in silico* annotations have contributed to the progress of chemistry-based dereplication and prioritization.[33−35] The workflow described in this study will facilitate the earliest analysis of large collection-based natural product drug discovery programs and the selection of target samples.

EXPERIMENTAL SECTION

General Experimental Procedures. Optical rotations were measured using a Jasco P-2000 digital polarimeter (Jasco, Tokyo, Japan) with a 10 cm cell at 20 °C and a sodium lamp (589 nm). UV absorption and electronic circular dichroism (ECD) spectra were measured on a Jasco J-1500 spectropolarimeter (Jasco, Tokyo, Japan). NMR spectra were obtained using a Bruker Avance III HD 500 MHz spectrometer (Bruker, Billerica, MA, USA) equipped with a 5 mm BBFO smart probe. MPLC separations were performed with a CombiFlash NextGen 300 system (Teledyne Isco Inc., Lincoln, NE, USA) equipped with a silica gel column cartridge. Preparative HPLC separations were performed with a Waters 600 HPLC system equipped with a Hector C_{18} column (250 × 21.2 mm or 250 × 10.0 mm, 5 μm, RS Tech, Daejeon, Korea) or a Spursil C_{18} EP column (250 × 10.0 mm, 5 μm, Dikma Technologies, Foothill Ranch, CA, USA). Extra-pure grade solvents for extraction, fractionation, and isolation were purchased from Daejung Chemical & Metal Co., Ltd. (Siheung, Korea). HPLC-grade water and MeCN for LC−MS were purchased from J. T. Baker (Avantor, Phillipsburg, NJ, USA), and formic acid was acquired from Sigma-Aldrich (St. Louis, MO, USA). Deuterated solvents for NMR analysis were purchased from Cambridge Isotope Laboratories (Cambridge, MA, USA).

Fungal Materials. Forty Polyporaceae fungal isolates were obtained from the Korea Mushroom Resource Bank (KMRB) at Seoul National University. The full list of these species is given in Table 2. The fruiting bodies of *C. volvatus* were collected from the Nambu Research Forest of Seoul National University, Baegwoon Mountain, Gwangyang, Korea (GPS N35°06′, E127°37′) in July 2017. The samples were maintained at −72 °C until extraction. These samples were authenticated by the late Prof. Sang Hyun Sung (Seoul National University, Seoul, Korea). The voucher specimen (SMU-19-001) was deposited at the College of Pharmacy, Sookmyung Women's University, Seoul, Korea.

MS/MS Molecular Networking-Based Prioritization. Each species was cultured on potato dextrose agar (PDA) medium for 4 weeks. Two Petri plates for each species were extracted with 80% aqueous MeOH (2 × 150 mL) for 2 days at rt; the quantities of extracts are summarized in Table 2. Dried extracts were stored at −72 °C until LC−MS/MS analyses.

Samples were dissolved in MeOH at a concentration of 5 mg/mL and filtered through a PTFE syringe filter (Woongki Science Co., Ltd., Seoul, Korea). LC−MS/MS analyses were performed on a Waters Acquity UPLC system (Waters Co., Milford, MA, USA) coupled to a Waters Xevo G2 Q/TOF mass spectrometer (Waters MS Technologies, Manchester, UK) equipped with an electrospray ionization interface (ESI). Chromatographic separations were

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

Table 2. General Information for the 40 Polyporoid Fungal Species Analyzed in This Study

fungal species	strain accession ID	extract amount (mg)
Cerrena aurantiopora	KMRB14110709	192.1
Cerrena consors	KMRB15062619	399.5
Cerrena unicolor	KMRB16011521	107.9
Cinereomyces lindbladii	KMRB17061910	377.2
Coriolopsis strumosa	KMRB15070136	84.0
Cryptoporus volvatus	KMRB16041505	72.7
Daedaleopsis confragosa	KMRB15081403	283.1
Daedaleopsis styracina	KMRB16032507	357.3
Daedaleopsis tricolor	KMRB15061217	126.5
Datronia mollis	KMRB16011417	131.1
Fomes fomentarius	KMRB14091212	232.1
Lenzites betulina	KMRB14082260	159.1
Lenzites styracina	KMRB17091811	74.0
Lopharia cinerascens	KMRB15090836	189.0
Microporus affinis	KMRB17080804	284.3
Microporus vernicipes	KMRB15071608	184.8
Neofavolus alveolaris	KMRB17070402	244.1
Neolentinus lepideus	KMRB15063021	199.7
Nigroporus vinosus	KMRB16071927	32.4
Perenniporia fraxinea	KMRB15070723	304.8
Perenniporia koreana	KMRB18042712	98.6
Perenniporia minutissima	KMRB15080602	106.2
Perenniporia subacida	KMRB15031920	315.1
Phlebiopsis castanea	KMRB18111308	218.8
Phlebiopsis gigantea	KMRB17083132	52.5
Polyporus arcularius	KMRB16061604	297.5
Polyporus brumalis	KMRB15102011	105.3
Polyporus tuberaster	KMRB15052230	313.6
Pycnoporus coccineus	KMRB15070728	370.2
Skeletocutis nivea	KMRB15090322	210.6
Trametes gibbosa	KMRB15090213	195.2
Trametes hirsuta	KMRB04050612	181.0
Trametes orientalis	KMRB15070765	218.6
Trametes suaveolens	KMRB14101203	175.6
Trametes trogii	KMRB15052235	272.2
Trametes versicolor	KMRB15032003	83.4
Trametopsis cervina	KMRB16011515	245.0
Trichaptum abietinum	KMRB15021115	158.2
Trichaptum biforme	KMRB16051230	104.3
Tyromyces chioneus	KMRB14070116	262.9

performed on a Waters Acquity UPLC BEH C_{18} (100 × 2.1 mm, 1.7 μm) column, and the temperature was maintained at 40 °C. Mixtures of H_2O (A) and MeCN (B) were eluted at a flow rate of 0.3 mL/min with an optimized gradient as follows: 10−54% B (0−6 min); isocratic at 54% B (6−11 min); 54−95% B (11−16 min); 95−100% B (16−19 min); isocratic B at 100% (19−22 min); 100−10% B (22−22.1 min); isocratic B at 10% (22.1−25 min). The samples (2.0 μL injection volume) were analyzed in data-dependent acquisition (DDA) mode consisting of a full MS survey scan in the m/z 100−2000 Da range (scan time = 150 ms) followed by MS/MS scans for the three most intense ions (m/z 50−2000 Da; scan time = 100 ms). The gradient of collision energy was set as 20−80 V. Raw MS/MS data were deposited in the MassIVE Public GNPS data sets (https://massive.ucsd.edu) with accession no. MSV000085974.

LC−MS/MS data were analyzed via classical or feature-based molecular networking workflows,[35] both of which are available in the GNPS web platform (https://gnps.ucsd.edu).[8] Data preprocessing for feature-based molecular networking was performed using MZmine 2.53 (detailed parameter for MZmine preprocessing is described in Table S1, Supporting Information).[36]

The feature-based MS/MS molecular network of 40 Polyporaceae species is accessible at the GNPS Web site via the following link: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f0688ba26f20404d979c3da1e53946f1.

The classical MS/MS molecular network of the *C. volvatus* culture extract and fruiting body extract is accessible at the GNPS Web site via the following link: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=842dbd2352514ea7b79ad30ca113829d.

The classical MS/MS molecular network and NAP annotation for the crude extract and fractions of *C. volvatus* fruiting bodies are accessible at the GNPS Web site with the following links: https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=2f25e613db5a4130beaafafbaac792f8 and https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=a736c296-f9844434b4324a65b964ecb8.

The weighted CSCS distance metric was calculated using the R package rCSCS (https://github.com/askerdb/rCSCS),[37] and the distance metric was visualized by performing principal coordinates analysis (PCoA) in the Qiime 2 environment.[38] All scripts, data, and results from data analyses are publicly accessible at: https://github.com/KyobinKang/supplementary-Polyporaceae.

Extraction and Isolation of Metabolites from the Fruiting Bodies of *C. volvatus*. Dried fruiting bodies of *C. volvatus* (76.80 g) were extracted with 95% EtOH (2 × 3 L) to yield the crude extract (11.66 g). The dried crude extract was dissolved in H_2O and then successively extracted with CH_2Cl_2, EtOAc, and BuOH. The CH_2Cl_2 fraction (10.32 g) was applied to MPLC eluted with CH_2Cl_2−MeOH 1:0 → 0:1 to obtain eight fractions, C1−C8. C5 was subjected to MPLC again to give eight subfractions C5a−C5h. C5e was further separated into six subfractions (C5e1−C5e6) by preparative HPLC (Hector C_{18} column, 4 mL/min, MeCN−H_2O 65:35 → 100:0). Nine subfractions C5e4a−C5e4i were derived from C5e4 by preparative HPLC (Hector C_{18} column, 4 mL/min, MeCN−H_2O 7:3 → 10:0). Compound **4** (3.6 mg, t_R = 15.5 min) was purified from C5e4f by preparative HPLC (Spursil C_{18} EP column, 4 mL/min, MeCN−H_2O 75:25 → 90:10). Preparative HPLC separation of C5E4e (Spursil C_{18} EP column, 4 mL/min, MeCN−H_2O 75:25 → 95:5) afforded compounds **6** (52.0 mg, t_R = 12.0 min) and **8** (22.6 mg, t_R = 12.4 min). Subfraction C5e5 was further fractionated using preparative HPLC (Spursil C_{18} EP column, 4 mL/min, MeCN−H_2O 75:25 → 100:0) to obtain eight subfractions (C5e5a−C5e5h); among them, subfraction C5e6 yielded compound **7** (22.6 mg, t_R = 15.3 min). Subfraction C5e5d was separated into two subfractions C5e5d1 and C5e5d2 by preparative HPLC (Hector C_{18} column, 4 mL/min, MeCN−H_2O 75:25 → 100:0). Compounds **9** (5.1 mg, t_R = 8.6 min) and **3** (4.8 mg, t_R = 12.7 min) were purified from C5e5d1 by preparative HPLC (Spursil C_{18} EP column, 4 mL/min, MeCN−H_2O 80:20 → 90:10), while compounds **2** (18.7 mg, t_R = 8.3 min) and **5** (7.5 mg, t_R = 9.1 min) were obtained from C5e5d2 by preparative HPLC (Hector C_{18} column, 4 mL/min, MeCN−H_2O 85:15 → 90:10). Subfraction C5e6 was fractionated by preparative HPLC (Spursil C_{18} EP column, 4 mL/min, MeCN−H_2O 75:25 → 100:0) into seven subfractions (C5e6a−C5e6g). Subfraction C5e6c was further purified by semipreparative HPLC (Spursil C_{18} EP column, 4 mL/min, isocratic MeCN−H_2O 80:20) to yield compound **1** (15.0 mg, t_R = 17.2 min).

*Cryptoporic Acid T (**1**).* This compound is a yellowish solid; $[\alpha]_D^{20}$= +49.2 (c 0.025, MeOH); λ_{max} (log ε) 208 (3.8) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 218 (4.5) nm; ^{1}H NMR and ^{13}C NMR data, see Table 1; HRESIMS m/z 1255.6630 $[M - H]^-$ (calcd for $C_{67}H_{99}O_{22}$, 1255.6628); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724260#%7B%7D.

*5‴,6‴-Cryptoporic Acid G Dimethyl Ester (**2**).* This compound is a yellowish solid; $[\alpha]_D^{20}$= +36.4 (c 0.025; MeOH); λ_{max} (log ε) 214 (3.7) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 218 (4.2), 242 (2.2) nm; ^{1}H NMR and ^{13}C NMR data, see Table 1; HRESIMS m/z 861.4651 $[M - H]^-$ (calcd for $C_{46}H_{69}O_{15}$, 861.4642); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724261#%7B%7D.

*5″-Cryptoporic Acid E Isopropyl Ester (**3**).* This compound is an amorphous powder; $[\alpha]_D^{20}$= +48.4 (*c* 0.025, MeOH); λ_{max} (log ε) 210 (3.7) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 217 (4.1), 242 (2.0) nm; ^{1}H NMR and ^{13}C NMR data, see Table 1; HRESIMS *m*/*z* 875.4782 [M − H]$^-$ (calcd for $C_{47}H_{71}O_{15}$, 875.4799); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724263#%7B%7D.

*Cryptoporic Acid D Trimethyl Ester (**4**).* This compound is an amorphous powder; $[\alpha]_D^{20}$= +46.4 (*c* 0.025, MeOH); λ_{max} (log ε) 218 (3.8) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 212 (−45.8) nm; ^{1}H NMR and ^{13}C NMR data, see Table 1; HRESIMS *m*/*z* 829.4370 [M − H]$^-$ (calcd for $C_{45}H_{65}O_{14}$, 829.4380); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724266#%7B%7D.

*6′,6‴-Cryptoporic Acid G Dimethyl Ester (**5**).* This compound is an amorphous powder; $[\alpha]_D^{20}$= +40.6 (*c* 0.025, MeOH); λ_{max} (log ε) 220 (3.8) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 218 (4.8), 239 (1.8) nm; ^{1}H NMR and ^{13}C NMR data, see Table S2 (Supporting Information); HRESIMS *m*/*z* 861.4656 [M − H]$^-$ (calcd for $C_{46}H_{69}O_{15}$, 861.4642); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724262#%7B%7D.

*Cryptoporic Acid E (**6**).* This compound is a yellowish solid; $[\alpha]_D^{20}$= +49.7 (*c* 0.025, MeOH); λ_{max} (log ε) 210 (3.7) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 204 (−4.5), 217 (1.0) nm; ^{1}H NMR and ^{13}C NMR data, see Table S2 (Supporting Information); HRESIMS *m*/*z* 847.4473 [M − H]$^-$ (calcd for $C_{45}H_{67}O_{15}$, 847.4486); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724265#%7B%7D.

*Cryptoporic Acid E Pentamethyl Ester (**7**).* This compound is a yellow solid; $[\alpha]_D^{20}$= +52.8 (*c* 0.025, MeOH); λ_{max} (log ε) 213 (3.8) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 217 (4.6), 241 (2.0) nm; ^{1}H NMR and ^{13}C NMR data, see Table S2 (Supporting Information); HRESIMS *m*/*z* 875.4787 [M − H]$^-$ (calcd for $C_{47}H_{71}O_{15}$, 875.4799); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724264#%7B%7D.

*Cryptoporic Acid D (**8**).* This compound is a yellow solid; $[\alpha]_D^{20}$= +60.2 (*c* 0.025, MeOH); λ_{max} (log ε) 217 (3.7) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 210 (−48.5) nm; ^{1}H NMR and ^{13}C NMR data, see Table S2 (Supporting Information); HRESIMS *m*/*z* 815.4206 [M − H]$^-$ (calcd for $C_{44}H_{63}O_{14}$, 815.4223); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724267#%7B%7D.

*Cryptoporic Acid A (**9**).* This compound is a yellowish solid; $[\alpha]_D^{20}$= +48.8 (*c* 0.025, MeOH); λ_{max} (log ε) 208 (3.8) nm; ECD (MeOH) λ_{max} ($\Delta\varepsilon$) 203 (−4.5), 217 (0.9), 228(−1.0) nm; ^{1}H NMR and ^{13}C NMR data, see Table S2 (Supporting Information); HRESIMS *m*/*z* 423.2372 [M − H]$^-$ (calcd for $C_{23}H_{35}O_7$, 423.2388); the MS/MS spectrum is deposited in the GNPS spectral library, https://gnps.ucsd.edu/ProteoSAFe/gnpslibraryspectrum.jsp?SpectrumID=CCMSLIB00005724268#%7B%7D.

Original NMR FIDs are available at: 10.5281/zenodo.4262237.

Cytotoxicity Assay Using HCT-116 Cells. Human colon cancer HCT-116 cells were obtained from the American Type Culture Collection. Cells were seeded into 96-well plates at a density of 3 × 10^2 cells/well with 100 μL of RPMI medium. Cells were incubated at 37 °C for 24 h in a 5% CO_2 incubator. Cells were then exposed to different concentrations of the compounds. After 48 h of incubation in 5% CO_2 at 37 °C, cell viability was determined using the MTS [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium] assay. Cells were treated with 10 μL of MTS solution (Promega, Madison, WI, USA) and 40 μL of PBS and then incubated for 1 h. The absorbance at 450 nm was measured using a SpectraMAX MS multiplate reader (Molecular Devices, Sunnyvale, CA, USA). Paclitaxel (Selleckchem, Houston, TX, USA) was used as a positive control.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jnatprod.0c00977.

Three molecular networks generated, detailed parameters for LC−MS preprocessing, NMR data of compounds **5**−**9**, and raw NMR spectra of compounds **1**−**4** (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Ki Hyun Kim − *School of Pharmacy, Sungkyunkwan University, Suwon 16419, Korea;* orcid.org/0000-0002-5285-9138; Email: khkim83@skku.edu

Kyo Bin Kang − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea;* orcid.org/0000-0003-3290-1017; Email: kbkang@sookmyung.ac.kr

Authors

Huong T. Pham − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea*

Kwang Ho Lee − *School of Pharmacy, Sungkyunkwan University, Suwon 16419, Korea*

Eunah Jeong − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea*

Sunmin Woo − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea; College of Pharmacy and Research Institute of Pharmaceutical Sciences, Seoul National University, Seoul 08826, Korea;* orcid.org/0000-0001-7561-2333

Jinsuh Yu − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea*

Woo-Young Kim − *Research Institute of Pharmaceutical Sciences, College of Pharmacy, Sookmyung Women's University, Seoul 04310, Korea*

Young Woon Lim − *School of Biological Sciences and Institute of Microbiology, Seoul National University, Seoul 08826, Korea*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jnatprod.0c00977

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was supported by the National Research Foundation of Korea (NRF) grants funded by the Ministry of Science, ICT, and Future Planning (NRF-2019R1F1A1058068 and NRF-2020R1C1C1004046).

■ REFERENCES

(1) González-Medina, M.; Owen, J. R.; El-Elimat, T.; Pearce, C. J.; Oberlies, N. H.; Figueroa, M.; Medina-Franco, J. L. *Front. Pharmacol.* **2017**, *8*, 180.

(2) Ha, J. W.; Kim, J.; Kim, H.; Jang, W.; Kim, K. H. *Nat. Prod. Sci.* **2020**, *26*, 118−131.
(3) Ruiz-Almenara, C.; Gándara, E.; Gómez-Hernández, M. *PeerJ* **2019**, *7*, e8325.
(4) Kabbaj, W.; Breheret, S.; Guimberteau, J.; Talou, T.; Olivier, J. M.; Bensoussan, M.; Sobal, M.; Roussos, S. *Appl. Biochem. Biotechnol.* **2002**, *102*, 463−469.
(5) Lu, R. L.; Bao, G. H.; Hu, F. L.; Huang, B.; Li, C. R.; Li, Z. Z. *Food Chem.* **2014**, *145*, 1066−1071.
(6) Henke, M. T.; Kelleher, N. L. *Nat. Prod. Rep.* **2016**, *33*, 942−950.
(7) Watrous, J.; Roach, P.; Alexandrov, T.; Heath, B. S.; Yang, J. Y.; Kersten, R. D.; Van Der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B. S.; Laskin, J.; Bandeira, N.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 1743−1752.
(8) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, C. A. P.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotechnol.* **2016**, *34*, 828−837.
(9) Fox Ramos, A. E.; Evanno, L.; Poupon, E.; Champy, P.; Beniddir, M. A. *Nat. Prod. Rep.* **2019**, *36*, 960−980.
(10) Sedio, B. E.; Echeverri, J. C. R.; Boya, C. A.; Wright, S. J. *Ecology* **2017**, *98*, 616−623.
(11) Kang, K. B.; Woo, S.; Ernst, M.; van der Hooft, J. J. J.; Nothias, L. F.; da Silva, R. R.; Dorrestein, P. C.; Sung, S. H.; Lee, M. *Phytochemistry* **2020**, *173*, 112292.
(12) Nothias, L.-F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H.; Fleischauer, M.; Aicheler, F.; Aksenov, A. A.; Alka, O.; Allard, P.-M.; Barsch, A.; Cachet, X.; Caraballo-Rodriguez, A. M.; Da Silva, R. R.; Dang, T.; Garg, N.; Gauglitz, J. M.; Gurevich, A.; Isaac, G.; Jarmusch, A. K.; Kameník, Z.; Kang, K. B.; Kessler, N.; Koester, I.; Korf, A.; Le Gouellec, A.; Ludwig, M.; Martin, H. C.; McCall, L.-I.; McSayles, J.; Meyer, S. W.; Mohimani, H.; Morsy, M.; Moyne, O.; Neumann, S.; Neuweger, H.; Nguyen, N. H.; Nothias-Esposito, M.; Paolini, J.; Phelan, V. V.; Pluskal, T.; Quinn, R. A.; Rogers, S.; Shrestha, B.; Tripathi, A.; van der Hooft, J. J. J.; Vargas, F.; Weldon, K. C.; Witting, M.; Yang, H.; Zhang, Z.; Zubeil, F.; Kohlbacher, O.; Böcker, S.; Alexandrov, T.; Bandeira, N.; Wang, M.; Dorrestein, P. C. *Nat. Methods* **2020**, *17* (9), 905−908.
(13) Habtemariam. *Biomedicines* **2020**, *8*, 135.
(14) Gao, L.; Sun, Y.; Si, J.; Liu, J.; Sun, Gl.; Sun, Z.; Cao, L. *PLoS One* **2014**, *9*, e113604.
(15) Narisawa, T.; Fukaura, Y.; Kotanagi, H.; Asakawa, Y. *Jpn. J. Cancer Res.* **1992**, *83*, 830−834.
(16) da Silva, R. R.; Wang, M.; Nothias, L. F.; van der Hooft, J. J. J.; Caraballo-Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. *PLoS Comput. Biol.* **2018**, *14*, e1006089.
(17) Van Santen, J. A.; Jacob, G.; Singh, A. L.; Aniebok, V.; Balunas, M. J.; Bunsko, D.; Neto, F. C.; Castaño-Espriu, L.; Chang, C.; Clark, T. N.; Cleary Little, J. L.; Delgadillo, D. A.; Dorrestein, P. C.; Duncan, K. R.; Egan, J. M.; Galey, M. M.; Haeckl, F. P. J.; Hua, A.; Hughes, A. H.; Iskakova, D.; Khadilkar, A.; Lee, J. H.; Lee, S.; Legrow, N.; Liu, D. Y.; Macho, J. M.; McCaughey, C. S.; Medema, M. H.; Neupane, R. P.; O'Donnell, T. J.; Paula, J. S.; Sanchez, L. M.; Shaikh, A. F.; Soldatou, S.; Terlouw, B. R.; Tran, T. A.; Valentine, M.; Van Der Hooft, J. J. J.; Vo, D. A.; Wang, M.; Wilson, D.; Zink, K. E.; Linington, R. G. *ACS Cent. Sci.* **2019**, *5*, 1824−1833.
(18) Asakawa, Y.; Hashimoto, T.; Mizuno, Y.; Tori, M.; Fukazawa, Y. *Phytochemistry* **1992**, *31*, 579−592.
(19) Isaka, M.; Chinthanom, P.; Danwisetkanjana, K.; Choeyklin, R. *Phytochem. Lett.* **2014**, *7*, 97−100.
(20) Wang, J. C.; Li, G. Z.; Lv, N.; Shen, L. G.; Shi, L. L.; Si, J. Y. *J. Asian Nat. Prod. Res.* **2017**, *19*, 719−724.
(21) Hirotani, M.; Furuya, T.; Shiro, M. *Phytochemistry* **1991**, *30*, 1555−1559.
(22) Yoshikawa, K.; Koso, K.; Shimomura, M.; Tanaka, M.; Yamamoto, H.; Imagawa, H.; Arihara, S.; Hashimoto, T. *Molecules* **2013**, *18*, 4181−4191.
(23) Wu, W.; Zhao, F.; Ding, R.; Bao, L.; Gao, H.; Lu, J. C.; Yao, X. S.; Zhang, X. Q.; Liu, H. W. *Chem. Biodiversity* **2011**, *8*, 1529−1538.
(24) Wang, J.; Li, G.; Gao, L.; Cao, L.; Lv, N.; Shen, L.; Si, J. *Phytochem. Lett.* **2015**, *14*, 63−66.
(25) Hashimoto, T.; Tori, M.; Mizuno, Y.; Asakawa, Y. *Tetrahedron Lett.* **1987**, *28*, 6303−6304.
(26) Cabrera, G. M.; Julia Roberti, M.; Wright, J. E.; Seldes, A. M. *Phytochemistry* **2002**, *61*, 189−193.
(27) Tori, M.; Hamada, N.; Sono, M.; Sono, Y.; Ishikawa, M.; Nakashima, K.; Hashimoto, T.; Asakawa, Y. *Tetrahedron Lett.* **2000**, *41*, 3099−3102.
(28) Maltese, F.; van der Kooy, F.; Verpoorte, R. *Nat. Prod. Commun.* **2009**, *4*, 447−454.
(29) Meng, J.; Li, Y. Y.; Ou, Y. X.; Song, L. F.; Lu, C. H.; Shen, Y. M. *Mycology* **2011**, *2*, 30−36.
(30) Matsunaga, S.; Furuya-suguri, H. *Carcinogenesis* **1991**, *12*, 1129−1131.
(31) Medema, M. H.; Blin, K.; Cimermancic, P.; De Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res.* **2011**, *39*, W339−W346.
(32) Navarro-Muñoz, J. C.; Selem-Mojica, N.; Mullowney, M. W.; Kautsar, S. A.; Tryon, J. H.; Parkinson, E. I.; De Los Santos, E. L. C.; Yeong, M.; Cruz-Morales, P.; Abubucker, S.; Roeters, A.; Lokhorst, W.; Fernandez-Guerra, A.; Cappelini, L. T. C.; Goering, A. W.; Thomson, R. J.; Metcalf, W. W.; Kelleher, N. L.; Francisco Barona-Gomez, F.; Medema, M. H. *Nat. Chem. Biol.* **2020**, *16*, 60−68.
(33) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; De Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonsi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J. Nat. Prod.* **2013**, *76*, 1686−1699.
(34) Nothias, L. F.; Nothias-Esposito, M.; Da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. *J. Nat. Prod.* **2018**, *81*, 758−767.
(35) Kang, K. B.; Park, E. J.; Da Silva, R. R.; Kim, H. W.; Dorrestein, P. C.; Sung, S. H. *J. Nat. Prod.* **2018**, *81*, 1819−1828.
(36) Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. *BMC Bioinf.* **2010**, *11*, 395.
(37) Brejnrod, A.; Ernst, M.; Dworzynski, P.; Rasmussen, L. B.; Dorrestein, P. C.; van der Hooft, J. J. J.; Arumugam, M. *BioRxiv* **2019**, 546150.
(38) Bolyen, E.; Rideout, J. R.; Dillon, M. R.; Bokulich, N. A.; Abnet, C. C.; Al-Ghalith, G. A.; Alexander, H.; Alm, E. J.; Arumugam, M.; Asnicar, F.; Bai, Y.; Bisanz, J. E.; Bittinger, K.; Brejnrod, A.; Brislawn,

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX

C. J.; Brown, C. T.; Callahan, B. J.; Caraballo-Rodríguez, A. M.; Chase, J.; Cope, E. K.; Da Silva, R.; Diener, C.; Dorrestein, P. C.; Douglas, G. M.; Durall, D. M.; Duvallet, C.; Edwardson, C. F.; Ernst, M.; Estaki, M.; Fouquier, J.; Gauglitz, J. M.; Gibbons, S. M.; Gibson, D. L.; Gonzalez, A.; Gorlick, K.; Guo, J.; Hillmann, B.; Holmes, S.; Holste, H.; Huttenhower, C.; Huttley, G. A.; Janssen, S.; Jarmusch, A. K.; Jiang, L.; Kaehler, B. D.; Kang, K. B.; Keefe, C. R.; Keim, P.; Kelley, S. T.; Knights, D.; Koester, I.; Kosciolek, T.; Kreps, J.; Langille, M. G. I.; Lee, J.; Ley, R.; Liu, Y. X.; Loftfield, E.; Lozupone, C.; Maher, M.; Marotz, C.; Martin, B. D.; McDonald, D.; McIver, L. J.; Melnik, A. V.; Metcalf, J. L.; Morgan, S. C.; Morton, J. T.; Naimey, A. T.; Navas-Molina, J. A.; Nothias, L. F.; Orchanian, S. B.; Pearson, T.; Peoples, S. L.; Petras, D.; Preuss, M. L.; Pruesse, E.; Rasmussen, L. B.; Rivers, A.; Robeson, M. S.; Rosenthal, P.; Segata, N.; Shaffer, M.; Shiffer, A.; Sinha, R.; Song, S. J.; Spear, J. R.; Swafford, A. D.; Thompson, L. R.; Torres, P. J.; Trinh, P.; Tripathi, A.; Turnbaugh, P. J.; Ul-Hasan, S.; van der Hooft, J. J. J.; Vargas, F.; Vázquez-Baeza, Y.; Vogtmann, E.; von Hippel, M.; Walters, W.; Wan, Y.; Wang, M.; Warren, J.; Weber, K. C.; Williamson, C. H. D.; Willis, A. D.; Xu, Z. Z.; Zaneveld, J. R.; Zhang, Y.; Zhu, Q.; Knight, R.; Caporaso, J. G. *Nat. Biotechnol.* **2019**, *37*, 852−857.

https://dx.doi.org/10.1021/acs.jnatprod.0c00977
J. Nat. Prod. XXXX, XXX, XXX−XXX