Beverage Plant Research

ARTICLE

Open Access

https://doi.org/10.48130/bpr-0024-0035

Beverage Plant Research **2025**, 5: e004

Bioactive oriented discovery of diterpenoids in *Coffea arabica* basing on 1D NMR and LC-MS/MS molecular network

Guilin Hu[1], Chenxi Quan[1,2], Abdulbaset Al-Romaima[1,2], Haopeng Dai[1,2] and Minghua Qiu[1,2*]

[1] *State Key Laboratory of Phytochemistry and Natural Medicines, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, Yunnan, China*
[2] *University of Chinese Academy of Sciences, Beijing 100049, China*
* Corresponding author, E-mail: mhchiu@mail.kib.ac.cn

Abstract

Coffee diterpenes are a class of characteristic components in coffee, which have potential biological activities including the prevention of cancer, obesity, diabetes, and other diseases. Due to the complex chemical composition of roasted coffee beans, analyzing the composition and potential activity of coffee diterpenes has always been a challenge. In the current research, based on activity-oriented research strategies, three novel coffee diterpene esters (**1**–**3**) with moderate α-glucosidase inhibitory activity were separated from roasted Arabica coffee beans. The structures of the three new compounds were determined through comprehensive spectral analysis. To explore trace active diterpene esters of the same type in coffee, a molecular network based on LC-MS/MS was constructed, and three novel coffee diterpene esters (**4**–**6**) were identified.

Citation: Hu G, Quan C, Al-Romaima A, Dai H, Qiu M. 2025. Bioactive oriented discovery of diterpenoids in *Coffea arabica* basing on 1D NMR and LC-MS/MS molecular network. *Beverage Plant Research* 5: e004 https://doi.org/10.48130/bpr-0024-0035

Introduction

In addition to providing humans with essential nutrients, functional foods also bring many biologically active ingredients to humans. These ingredients will exert a wide range of biological activities including anti-oxidation, hypoglycemic, neuroprotection, and lipid-lowering[1]. It has always been a hot and difficult area in food chemistry research to quickly find these potential functional ingredients from the complex extracts of food. Generally, research on functional ingredients in food mostly use phytochemical research procedures, including extraction, separation, structure analysis, and activity evaluation[2]. However, this traditional research method is usually blind, will bring a lot of extra work, and ultimately may not be able to determine active ingredients. In recent years, to improve the research efficiency, state-of-the-art instruments like NMR[3,4], HPLC-SPE-NMR[5], and HPLC-MS^n[6–13] have been used alone or combined with activity detection or metabolomics analysis to quickly carry out dereplication.

Since the method by which the target bioactive metabolites can be quickly obtained at a low cost, activity tracking is an important concept in the dereplication of metabolites. In activity-oriented research, NMR, and LC-MS^n are still the two most frequently adopted detection methods[14]. In NMR detection, through simple sample pretreatment, the comprehensive information of the protons and/or carbons of the mixture can be quickly captured in a non-destructive manner without chromatographic separation. Then, by comparing with the NMR database, part of known compounds can be dereplicated[3,4]. Compared with NMR, the sensitivity advantage of LC-MS^n allows it to realize the detection of trace compounds present in crude mixture. It is worth mentioning that, in recent years, LC-MS/MS-based molecular networking has been widely used in the tracking of active compounds[12,15]. Although several cases have shown that LC-MS^n can perform well in the mining of active compounds, but the complexity of fragment interpretation limits it as an independent detection method for structural identification[14]. Therefore, the identification scheme of active compounds by integrating NMR and LC-MS/MS is worth exploring.

Coffee, as the most popular beverage in the world, is grown in over 70 countries and regions, intimately connected to the lives of billions of people worldwide, and stands as one of the most economically valuable plants. Currently, the globally cultivated coffee consists primarily of two species: Arabica (*Coffea arabica*) and Robusta (*Coffea canephora*). Arabica coffee excels in taste and aroma compared to Robusta coffee, making it the most widely planted, produced, and traded coffee globally. Coffee diterpenoids are a class of key characteristic components found in coffee, with the majority (about 99.6%) existing in the form of diterpene esters in coffee oil[15]. Approximately 24 fatty acids participate in the formation of these esters, with palmitic acid and linoleic acid diterpene esters being the most abundant. The content of coffee diterpenoids varies significantly among different coffee species, with Arabica coffee generally containing higher levels of coffee alcohols compared to Robusta coffee[15]. Reports on the structure of coffee diterpenoids can be traced back to the 1930s. In 1932, Bengis & Anderson and others first detected Kahweol in coffee beans[16]. So far, over 70 types of coffee diterpenes, including multiple sub-types, have been reported[1]. Research has shown that C&K can exhibit anti-cancer effects through various pathways[17]. Additionally, C&K may potentially lower blood sugar levels by stimulating insulin secretion and increasing glucose uptake[18,19]. Because coffee roasted beans are a very complex mixture system, whether there are other novel active coffee diterpenes in it is a question worth exploring in depth.

In the current research, an activity-oriented scheme combining NMR and LC-MS/MS technology are proposed. Basing on the strategy, three new coffee diterpene esters exhibiting moderate α-glucosidase inhibitory activity were isolated from roasted Arabica coffee beans. The structures of these novel compounds were elucidated through thorough spectral analysis. To investigate similar trace active diterpenoids in coffee, a molecular network was established using LC-MS/MS, leading to the identification of three additional novel coffee diterpene esters.

Materials and methods

Instrument, chemicals and coffee material

The ^{1}H NMR, ^{13}C NMR, DEPT, and 2D NMR data were collected using Bruker DRX-600 instruments (Bruker, Zurich, Switzerland). The deuterated chloroform for NMR experiments was purchased from Beijing Yinuokai Technology Co., Ltd (Beijing, China). The purification of the compounds was performed on an Agilent HP1100 series instrument equipped with a UV L-2400 detector. The acetonitrile (99.9%) for semi-preparation was produced by Oceanpak Alexative Chemical, Ltd (Sweden). The optical rotation and UV spectra were obtained with a Jasco P-1020 polarimeter (Jasco, Japan) and a UV-2401 PC spectrophotometer (Shimadzu, Japan), respectively. The LC-MS/MS and HRESIMS experiments were performed on a UPLC system (1260, Agilent) coupled to a quadrupole time-of-flight mass spectrometer (Agilent 6540 Q-TOF, Agilent Technologies).

Acarbose and p-nitrophenyl-α-$_{D}$-glucopyranoside were obtained from Aladdin Co., Ltd. (Shanghai, China). The α-glucosidase was purchased from Sigma-Aldrich Trading Co., Ltd (Lot # 089M4087V) (Shanghai, China). The green coffee beans of *Coffea* arabica were purchased in Yunnan province, China. Green coffee beans were roasted to a moderate degree according to the color value by a professional barista[20]. The diterpene ester extract (80.7 g) used as demonstration example in the current study was from the PE/EtOAc = 8:1 fraction of the petroleum ether extraction.

D NMR detection and α-glucosidase activity screening

Silica gel column chromatography was applied to separate the diterpene extract into Fr.1–Fr.19 (Fr.1, 1.7 g; Fr.2, 2.2 g; Fr.3, 4.7 g; Fr.4, 3.8 g; Fr.5, 5.4 g; Fr.6 3.9 g; Fr.7, 2.4 g; Fr.8, 0.9 g; Fr.9, 1.6 g; Fr.10, 2.4 g; Fr.11, 1.2 g; Fr.12, 3.0 g; Fr.13, 1.9 g; Fr.14, 3.4 g; Fr.15, 4.6 g; Fr.16, 4.9 g; Fr.17, 3.4 g; Fr.18, 0.7 g; Fr.19, 2.6 g) with PE/EA (5:2–1:1) system. The right amount of extract from each fraction were transferred into a 1.5 mL EP tube. A 25 mg/mL solution was prepared with deuterated chloroform, and 600 μL was pipetted into a nuclear magnetic tube, and ^{1}H NMR data collected on a 600 MHz nuclear magnetic resonance instrument. Refer to previous literature for ^{1}H NMR acquisition parameters[21]. The method described in a previous article was adopted to detect α-glucosidase (50 μg/mL) inhibitory activity for Fr.1–Fr.19[20].

MestReNova (version: 6.1.0–6224) was adopted to align and standardize the^{1}H NMR data of Fr.1–Fr.19, solvent peaks (7.250–7.280 ppm) and areas without characteristic signals (8.396–9.000 ppm) were removed. The standardization parameter was set as 100, the integration interval was 5.0–11.0 ppm. The interval was 0.004 ppm, and the integrated data was stored as a file with the suffix '.csv' to obtain an integrated area matrix of 1,167 × 19. Further, a clustering heat map containing the ^{1}H NMR data matrix and α-glucosidase inhibitory activity data was plotted by using the R program package 'pheatmap' in RStudio (version: 1.3.1056). Representative samples Fr.9 were selected from the groups showing α-glucosidase inhibitory activity for ^{13}C DEPT NMR detection.

Purification, identification, and activity verification of target compounds

The representative Fr.9 was further purified by semi-preparative HPLC (ZORBAX, RX-C8, 5 μm, 9.4 mm × 250 mm, $CNCH_3 : H_2O$ = 95:5) to obtain compounds **1** (t_R = 16 min, 8.6 mg), **2** (t_R = 24 min, 3.4 mg) and **3** (t_R = 31 min, 3.1 mg)**.** Then, ^{1}H NMR and ^{13}C NMR were performed for the establishment of exact structures of compounds **1**–**3**. To further confirm the structures of compounds **1**–**3**, **2** was selected for 2D NMR (HSQC, ^{1}H-^{1}H COSY, HMBC and ROESY) detection. The α-glucosidase inhibitory activities of **1**–**3** were screened according to the methodology reported in a previous study[22].

Caffaldehyde A (***1***) colorless oily liquid, $[\alpha]^{20}_{D}$ –82.19 (c = 0.05, MeOH), UV (MeOH) λ_{max} (log ε) 218.5 (3.74), 263.5 (3.50). *m/z* 591.4022 $[M+Na]^+$ (calcd for $C_{36}H_{56}O_5Na^+$, 591.4020. ^{1}H and ^{13}C NMR spectral data are shown in Supplemental Table S1, Supplemental Figs S3 & S4.

Caffaldehyde B (***2***) colorless oily liquid, $[\alpha]^{21}_{D}$ –65.16 (c=0.05, MeOH), UV (MeOH) λ_{max} (log ε) 218.5 (3.66), 268.0 (3.45). HRESIMS *m/z* 597.4517 $[M+H]^+$ (calcd for $C_{38}H_{61}O_5^+$, 597.4514). ^{1}H and ^{13}C NMR spectral data are shown in Supplemental Table S1 and Supplemental Figs S5 & S6. 2D NMR spectra are shown in Supplemental Figs S7–S10.

Caffaldehyde C (***3***) colorless oily liquid, $[\alpha]^{21}_{D}$ –60.67 (c = 0.05, MeOH), UV (MeOH) λ_{max} (log ε) 221 (3.75), 267.5 (3.59). HRESIMS *m/z* 625.4831 $[M+H]^+$ (calcd for $C_{40}H_{65}O_5^+$, 625.4827). ^{1}H and ^{13}C NMR spectral data are shown in Supplemental Table S1, Supplemental Figs S11 & S12

LC-MS/MS analysis and molecular network construction

UPLC-MS/MS analysis was performed using an UPLC system (1260, Agilent) coupled to a quadrupole time-of-flight mass spectrometer (Agilent 6540 Q-TOF, Agilent Technologies) equipped with an ESI source operating with a positive polarity. An RX-C8 column (150 mm × 4.6 mm × 5 μm) was used for separation. Water (A) and acetonitrile (B) were selected as eluents, and the optimal condition was set as: acetonitrile from 70% to 100% for 30 min, then eluting with 100% acetonitrile for 30 min. The flow rate was 0.5 mL/min. The detection wavelengths of the DAD detector were 210, 254, 270, and 330 nm. The MS acquisition conditions and the construction method of the molecular network were the same as previous reports[22].

Results and discussion

General strategy for discovery of bioactive ingredients in roasted coffee

In the experimental design of this study, it is hoped that the scheme will take into account several aspects: (1) Assist in the discovery and separation of active compounds; (2) Direct identification of trace novel compounds without chromatographic purification; (3) Avoid chromatographic separation as much as possible to reduce the use of solvents and consumables; (4) Reduce the difficulty of spectrum analysis and shorten the research time.

After weighing the pros and cons of various dereplication methods, a three-step scheme, as shown in Fig. 1, was proposed. Firstly, the sub-fraction samples are subjected to ^{1}H-NMR and activity detection. The ^{1}H NMR and activity data obtained are analyzed by cluster-heatmap to group and screen out the active groups. Secondly, fractions of interest were screened out for ^{13}C-DEPT NMR detection. Fractions likely to contain active novel compounds are purified by semi-preparative HPLC. The isolated novel compounds were confirmed by 2D NMR and HRESIMS and verified for activity. Finally, LC-MS/MS based molecular network is adopted to explore potential trace novel active compounds.

Bioactive compounds from *Coffea arabica*

In natural product research, valuable compounds account for only a small portion, most of the extracts are worthless. Therefore, the earlier to locate the valuable part of the extract, the less the workload will be. In the current study, the crude diterpene extract of roasted *coffea Arabica* was chosen to to discover active novel compounds in natural products.

^{1}H NMR integrates several advantages such as high sensitivity, fast detection, good repeatability, and non-destructive testing,

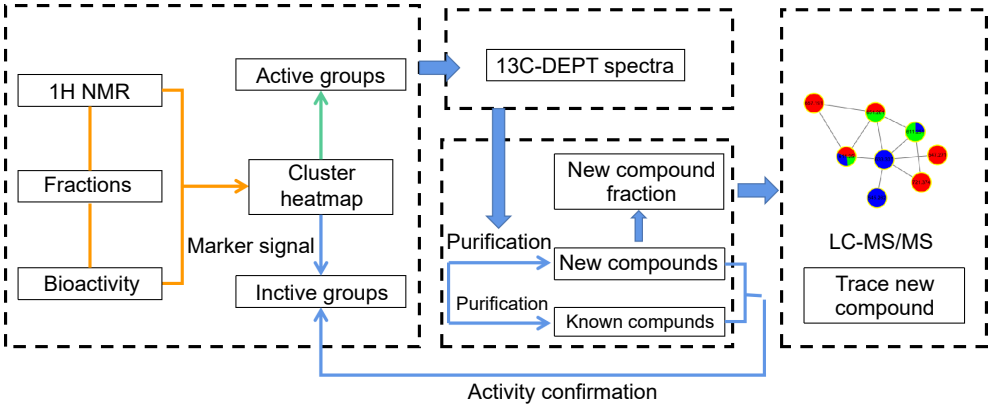

Fig. 1 General strategy for discovery of bioactive ingredients in roasted coffee.

making it a very convenient qualitative method to carry out rapid testing on a large number of samples to obtain high-quality spectra. In previous studies, several teams achieved rapid identification of the main components in coffee D_2O extracts by ^{1}H NMR, indicating the possibility of identifying compounds in the mixture[23,24]. In the current study, the diterpene extract of roasted coffee was divided into 19 fractions using silica gel column chromatography. To achieve the purpose of shortening the research time and reducing the use of solvents, ^{1}H-NMR combining activity detection was selected as the first step of repetition for the 19 fractions to make an early stop-or-go decision (Supplementary Fig. S1). The ^{1}H NMR spectra of the fractions were processed with MestReNova to obtain the integral value of the characteristic region (5.0–11.0 ppm). At the same time, solvent peak areas and the areas without characteristic signals were manually deleted. Therefore, a matrix with 19 rows, 1,167 columns was generated for the construction of cluster heatmap (Fig. 2). Meanwhile, the α-glucosidase inhibitory activity of the fractions at 50 μg/mL were marked at the top of the heatmap according to the shade of the color.

As shown in Fig. 2, according to the cluster result basing on ^{1}H NMR data, fractions were divided into five groups (Fr.1–Fr.3, Fr.4–Fr.5, Fr.6–Fr.8, Fr.9–Fr.13, and Fr.14–Fr.19). The Fr.1–Fr.3 group did not show any α-glucosidase inhibitory activity, therefore, no follow-up analysis was performed on this group. The fractions from the remaining four groups exhibited varying degrees of inhibitory activity. Interestingly, the fractions with significant activity were mainly concentrated in Fr.9–Fr.13. Combining clustering heatmap and ^{1}H NMR spectra found that Fr.9–Fr.13 possessed common characteristic signals at δ_H 10.19 (d, 7.2 Hz), δ_H 7.25 (d, 10.01 Hz), δ_H 6.00 (d, 10.01 Hz), and δ_H 5.97 (d, 7.2 Hz), indicating that the compounds with these characteristic signals may play a key role in Fr.9–Fr.13 exerting α-glucosidase inhibitory activity (Fig. 2). Since fractions within group have high ^{1}H NMR similarity, it is only necessary to randomly select a representative fraction in each group of interest for subsequent analysis, which can effectively reduce the workload.

After the first step, the location and characteristic signal of the bio-active compounds have been found. However, it was hard to infer more information based on the limited ^{1}H NMR signals. Therefore, in the second step of dereplication, the representative sample Fr.9 from the active group (Fr.9–Fr.13) was tested for ^{13}C DEPT NMR data (Supplementary Fig. S2). The results indicate that there is a characteristic carbon signal of aldehyde groups (δ_C 192.9) in the mixture, which is consistent with its characteristic proton signal (δ_H 10.19 d, 7.2 Hz). Furthermore, the representative sample Fr.9 was

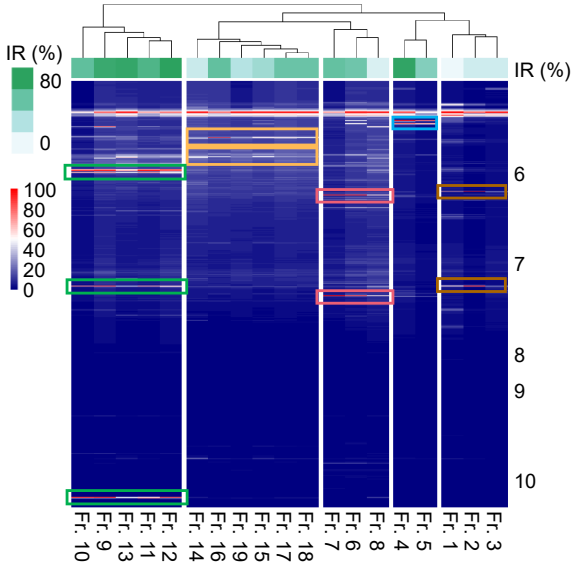

Fig. 2 Cluster heatmap of coffee diterpene fractions basing on ^{1}H NMR data. The α-glucosidase inhibitory rate (IR, 50 μg/mL, n = 3) of each fraction are marked on the top. The darker the color indicates the stronger the activity.

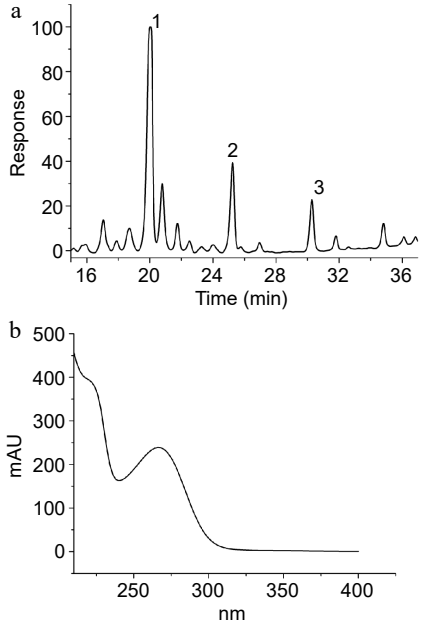

Fig. 3 (a) HPLC chromatogram of Fr.9. (b) UV spectrum of 1 (UV_{max} = 218 ± 5 nm, 265 ± 5 nm).

separated by HPLC (Fig. 3). The three main chromatographic peaks (t_R = 16 min, t_R = 24 min, t_R = 31 min) with UV_{max} at 218 ± 5 and 265 ± 5 nm were purified by semi-preparative HPLC to obtain compounds **1**–**3** (Figs 3, 4).

Compound **1** was obtained as colorless oily liquid. ^{1}H NMR spectrum showed two methyl at δ_H 1.17 (3H, s) and δ_H 0.88 (3H, t, 7.0 Hz), one oxidation methylene signal at δ_H 4.26 (2H, s), three olefinic protons δ_H 7.25 (1H, d, 10.1 Hz), δ_H 5.97 (1H, dd, 7.2 Hz, 1,8 Hz), and δ_H 6.00 (1H, d, 10.1 Hz), one aldehyde proton signal δ_H 10.18 (1H, d, 7.2 Hz). The ^{13}C-DEPT NMR spectrum showed that **1** possessed a total of 36 carbon atoms, which were two methyl groups, 21 methylene groups (including one oxymethylene group), and seven methine groups (including three olefinic carbons, one aldehyde carbon), six quaternary carbons (including one oxygenated tertiary carbon, one alkene carbon, one ketone carbonyl, one ester carbonyl). Through literature comparison, it was found that **1** contained exactly the same diterpene skeleton reported by De Lucia et al. in 2009[25]. In addition, a palmitic acid signal appeared in ^{13}C NMR, indicating that **1** is a diterpene ester with sixteen carbon saturated fatty acids. The ^{1}H NMR and ^{13}C NMR data of compounds **2** and **3** are exactly similar to those of **1**, except the ^{13}C NMR and HRESIMS indicated that **2** has two more methylenes and **3** has four more methylenes in the long chain of fatty acids compared to **1**. Therefore, the fatty acid moieties of compounds **2** and **3** were determined to be stearic acid (18:0) and arachidic acid (20:0), respectively. To further confirm the structures of these three compounds, **2** was selected for 2D NMR detection The esterification location was proved to be at C-17 based on the HMBC correlation from δ_H 4.26 (H-17) to δ_C 174.0 (C-1') (Fig. 4). The above inference was consistent with all 2D NMR correlations and HRESIMS data. The ROESY correlations of H-13/H-17 assigned OH-16 as α-orientation. Consequently, the structures of **1**–**3** were determined as shown and named as caffaldehydes A-C, respectively.

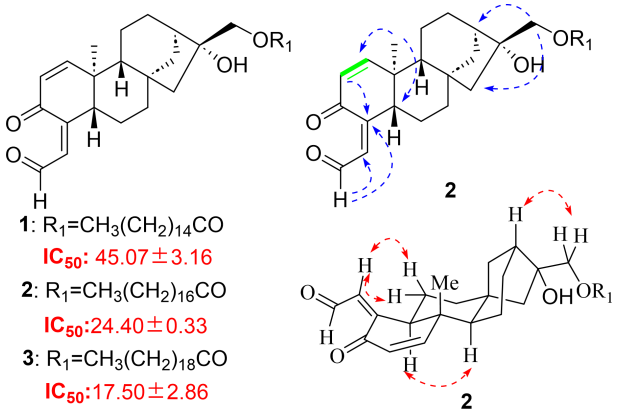

Fig. 4 Structures of compounds **1**–**3** and their IC_{50} to α-glucosidase (μM, data expressed as means ± SD, n = 3); Key 2D NMR correlations of compound **2**.

Furthermore, the α-glucosidase inhibitory activities of compounds **1**–**3** were verified, and their IC_{50} values were calculated as 45.07 ± 3.16, 24.40 ± 0.33, and 17.50 ± 2.86 μM respectively (Fig. 4), compared to the positive control (acarbose, IC_{50}: 60.71 ± 16.45 μM). This proves that **1**–**3** are the components of Fr.9 that exert α-glucosidase inhibitory activity.

In the previous two steps, some known compounds have been identified without chromatographic separation and the active compounds with a high content (**1**–**3**) have been identified and separated. However, due to the sensitivity limitation of NMR or HPLC, it is hard to identify and separate compounds with lower content. To achieve this purpose, Fr.1–Fr.19 were combined to form 4 groups (G1: Fr.1–Fr.5; G2: Fr.6–Fr.8; G3: Fr.9–Fr.13; G4: Fr.14–Fr.19) and LC-HRMS/MS was performed. Based on the MS/MS data from G1–G4, a molecular network was constructed using GNPS (http://gnps.ucsd.edu) and visualized with cytoscape 3.8.0 (Fig. 5a, Supplementary Fig. S13). The specific parameters and results of the molecular network can be accessed through the link https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=e5f985a8148c4ee58fa5c8e4f90ed50c. In the molecular network, a node represents

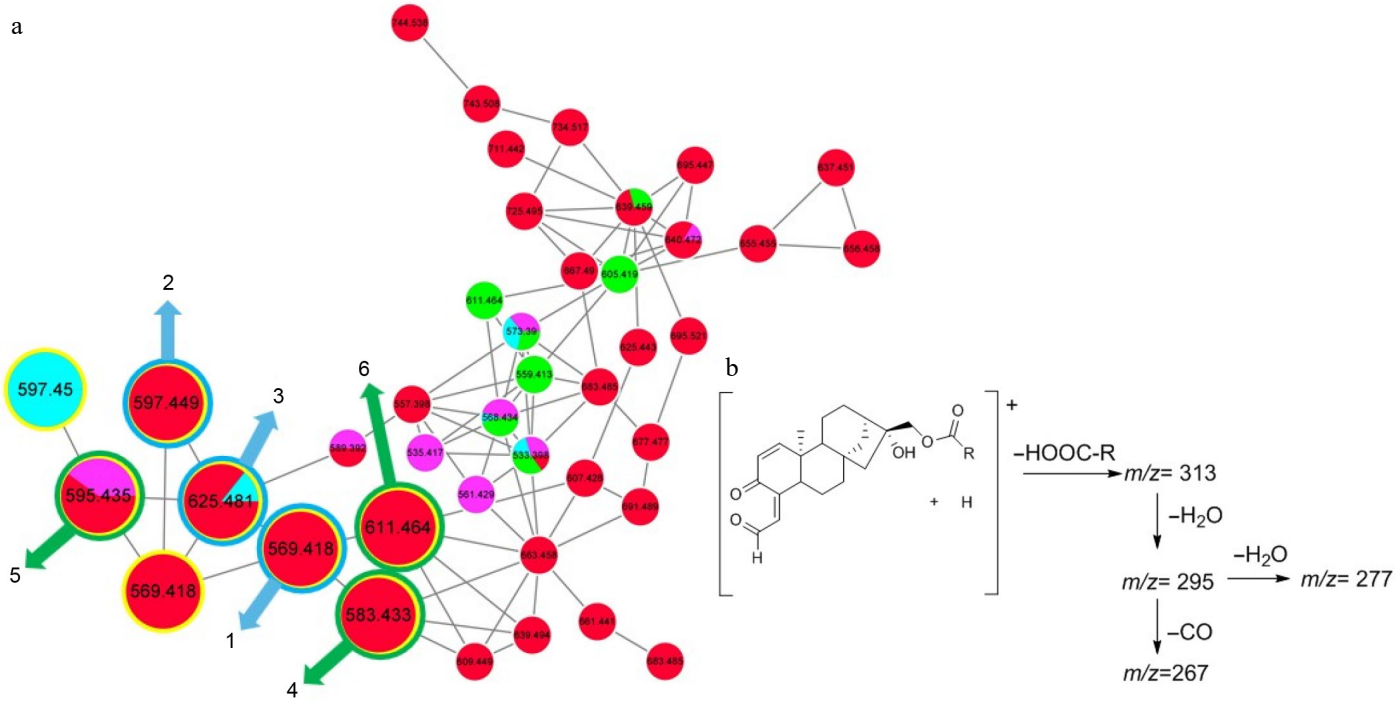

Fig. 5 (a) MN1 containing diterpenes (G1: Fr.1–Fr.5, green; G2: Fr.6-–Fr.8, blue; G3: Fr.9–Fr.13, red; G4: Fr.14–Fr.19, purple). (b) Possible formation pathways of characteristic fragments of compounds **1**–**6**.

Table 1. The HRESIMS data of compounds **1**–**6**.

No.	HRMS	Formula	Category	Δ ppm
1	569.4178	$C_{36}H_{56}O_5$	$(M+H)^+$	4.13
2	597.4492	$C_{38}H_{60}O_5$	$(M+H)^+$	3.89
3	625.4799	$C_{40}H_{64}O_5$	$(M+H)^+$	4.04
4	583.4329	$C_{37}H_{58}O_5$	$(M+H)^+$	4.52
5	595.4339	$C_{38}H_{58}O_5$	$(M+H)^+$	3.22
6	611.4650	$C_{39}H_{62}O_5$	$(M+H)^+$	3.71

the quasi-molecular ion peak of a compound, and the color represents the group in which the compound exists.

Compounds **1**–**3** were identified in the MN-1 based on the HRESIMS data of their $[M+H]^+$ peak and characteristic fragment ion peak comparison. As shown in Fig. 5a, the three nodes representing **1**–**3** were in adjacent positions in MN1. MS/MS analysis found that **1**–**3** have common fragment ion peaks *m/z* 313, 295, 277, and 267, which were calculated as $C_{20}H_{25}O_3^+$ (313.1813, −4.64 ppm), $C_{20}H_{23}O_2^+$ (295.1682, 3.41 ppm), $C_{20}H_{21}O^+$ (277.1577, 2.94 ppm) and $C_{19}H_{23}O^+$ (267.1737, 5.92 ppm), respectively. Interestingly, near the nodes of **1**–**3**, there were three nodes 583.433 (**4**, 583.4329, $C_{37}H_{59}O_5^+$, 4.52 ppm), 595.435 (**5**, 595.4339, $C_{38}H_{59}O_5^+$, 3.22 ppm) and 611.464 (**6**, 611.4650, $C_{39}H_{63}O_5^+$, 3.71 ppm) with the same characteristic fragments (Table 1), indicating **4**–**6** were also diterpene esters whose diterpene moieties were the same as **1**–**3**. Through the molecular weight search of DATAanalyte, it was found that **4**–**6** did not exist in the database, indicating that they were also novel coffee diterpene esters. Combined with the molecular formula from the HRESIMS, it could be inferred that the fatty acids esterified with **4**–**6** were magaric acid (17:0), octadecene acid (18:1), and nonadecanoic acid (19:0), respectively. The possible formation pathways of the characteristic fragments of **1**–**6** were deduced in Fig. 5b and Supplementary Fig. S14–S19.

Conclusions

An active compound-oriented separation strategy was proposed. The proof-of-principle of the strategy was successfully achieved in the diterpene extract of *Coffea arabica*. Three novel diterpene ester compounds (**1**–**3**) with significant α-glucosidase inhibitory activity were rapidly achieved targeted separation with very small amounts of organic solvent. Three other trace coffee diterpene esters (**4**–**6**) with potential α-glucosidase inhibitory activity were directly identified without enrichment with the aid of an LC-MS/MS-based molecular network. The results indicated that the multi-spectral technology dereplication strategy proposed in the current research can greatly improve the research efficiency of novel active compounds. In the future, this method is worthy of being applied to the targeted separation of active metabolites of other food.

Author contributions

The authors confirm contribution to the paper as follows: methodology, validation, formal analysis: Hu G, Quan C, Al-Romaima A, Dai H; funding aquisition: Hu G, Qiu M; data curation, visualization, writing - original draft, review & editing: Hu G; supervision, project administration: Qiu M. All authors reviewed the results and approved the final version of the manuscript.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This study was supported financially by the National Natural Science Foundation of China, China (U1902206) , Key Research and Development Project of Yunnan Province (202003AD15 0006), China Postdoctoral Science Foundation (2023M743591), Postdoctoral Fellowship Program of CPSF (GZC20232766), Special Research Assistant of Chinese Academy of Sciences, Yunnan Caiyun Postdoctoral Program, Coffee Industry Science and Technology Mission in Zhenkang County, Yunnan (202204BI090009), Project of Yunnan International Joint Innovation platform (202203AP140106), Project of Expert workstation of Yunnan Science and Technology Association (2024).

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary information accompanies this paper at (https://www.maxapress.com/article/doi/10.48130/bpr-0024-0035)

Dates

Received 8 June 2024; Revised 20 August 2024; Accepted 23 September 2024; Published online 18 February 2025

References

1. Hu GL, Wang X, Zhang L, Qiu MH. 2019. The sources and mechanisms of bioactive ingredients in coffee. *Food & Function* 10:3113–26
2. Wang X, Meng Q, Peng X, Hu G, Qiu M. 2018. Identification of new diterpene esters from green Arabica coffee beans, and their platelet aggregation accelerating activities. *Food Chemistry* 263:251–57
3. Hubert J, Nuzillard JM, Purson S, Hamzaoui M, Borie N, et al. 2014. Identification of natural metabolites in mixture: A pattern recognition strategy based on ^{13}C NMR. *Analytical Chemistry* 86:2955–62
4. Clendinen CS, Lee-McMullen B, Williams CM, Stupp GS, Vandenborne K, Hahn DA, Walter GA, Edison AS. 2014. ^{13}C NMR metabolomics: Applications at natural abundance. *Analytical Chemistry* 86:9242–50
5. Johansen KT, Wubshet SG, Nyberg NT. 2013. HPLC-NMR revisited: using time-slice high-performance liquid chromatography-solid-phase extraction-nuclear magnetic resonance with database-assisted dereplication. *Analytical Chemistry* 85:3183–89
6. Tabudravu JN, Pellissier L, Smith AJ, Subko K, Autréau C, et al. 2019. LC-HRMS-Database screening metrics for rapid prioritization of samples to accelerate the discovery of structurally new natural products. *Journal of Natural Products* 82:211–20
7. Li P, Anandhi Senthilkumar H, Figueroa M, Wu SB, Fata JE, et al. 2016. UPLC-QTOFMS(E)-guided dereplication of the endangered chinese species Garcinia paucinervis to identify additional benzophenone derivatives. *Journal of Natural Products* 79:1619–27
8. Wei W, Hou J, Yao C, Bi Q, Wang X, et al. 2019. A high-efficiency strategy integrating offline two-dimensional separation and data post-processing with dereplication: characterization of bufadienolides in *Venenum Bufonis* as a case study. *Journal of Chromatography A* 1603:179–89
9. Nguyen DD, Wu CH, Moree WJ, Lamsa A, Medema MH, et al. 2013. MS/MS networking guided analysis of molecule and gene cluster families. *Proceedings of the National Academy of Sciences of the United States of America* 110:E2611–E2620
10. Allard PM, Péresse T, Bisson J, Gindro K, Marcourt L, et al. 2016. Integration of molecular networking and *in-silico* MS/MS fragmentation for natural products dereplication. *Analytical Chemistry* 88:3317–23
11. Wu C, van der Heul HU, Melnik AV, Lübben J, Dorrestein PC, et al. 2019. Lugdunomycin, an angucycline-derived molecule with unprecedented chemical architecture. *Angewandte Chemie - International Edition* 58:2809–14

12. Abbas-Mohammadi M, Farimani MM, Salehi P, Ebrahimi SN, Sonboli A, et al. 2022. Molecular networking based dereplication of AChE inhibitory compounds from the medicinal plant *Vincetoxicum funebre* (Boiss & Kotschy). *Journal Of Biomolecular Structure & Dynamics* 40:1942–51
13. Wolfender JL, Litaudon M, Touboul D, Queiroz EF. 2019. Innovative omics-based approaches for prioritisation and targeted isolation of natural products - new strategies for drug discovery. *Natural Product Reports* 36:855–68
14. Hubert J, Nuzillard JM, Renault JH. 2017. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochemistry Reviews* 16:55–95
15. Moeenfard M, Alves A. 2020. New trends in coffee diterpenes research from technological to health aspects. *Food Research International* 134:109207
16. Bengis RO, Anderson RJ. 1932. The chemistry of the coffee-bean I. Concerning the unsaponifiable matter of the coffee-bean oil. Preparation and properties of kahweol. *Journal of Biological Chemistry* 97(1):99–113
17. Iwamoto H, Izumi K, Kadono Y, Mizokami A. 2022. Anticancer effects of coffee diterpenes kahweol acetate and cafestol on prostate cancer. *Cancer Science* 113:1733–33
18. Mellbye FB, Jeppesen PB, Shokouh P, Laustsen C, Hermansen K, et al. 2017. Cafestol a bioactive substance in coffee, has antidiabetic properties in KKAy mice. *Journal of Natural Products* 80:2353–59
19. Mellbye FB, Jeppesen PB, Hermansen K, Gregersen S. 2015. Cafestol, a bioactive substance in coffee, stimulates insulin secretion and increases glucose uptake in muscle cells: Studies in Vitro. *Journal of Natural Products* 78:2447–51
20. Hu G, Peng X, Dong D, Nian Y, Gao Y, et al. 2020. New *ent*-kaurane diterpenes from the roasted arabica coffee beans and molecular docking to α-glucosidase. *Food Chemistry* 345:128823
21. Hu G, Peng X, Gao Y, Huang Y, Li X, et al. 2020. Effect of roasting degree of coffee beans on sensory evaluation: Research from the perspective of major chemical ingredients. *Food Chemistry* 331:127329
22. Hu GL, Gao Y, Peng XR, Liu JH, Su HG, et al. 2020. Lactam *ent*-Kaurane Diterpene: a new class of diterpenoids present in roasted beans of *Coffea arabica*. *Journal of Agricultural and Food Chemistry* 68:6112–21
23. Wei F, Furihata K, Hu F, Miyakawa T, Tanokura M. 2011. Two-dimensional $^{1}H^{13}C$ nuclear magnetic resonance (NMR)-based comprehensive analysis of roasted coffee bean extract. *Journal of Agricultural and Food Chemistry* 59:9065–73
24. Ciaramelli C, Palmioli A, Airoldi C. 2019. Coffee variety, origin and extraction procedure: implications for coffee beneficial effects on human health. *Food Chemistry* 278:47–55
25. De Lucia M, Panzella L, Melck D, Giudicianni I, Motta A, et al. 2009. Differential reactivity of purified bioactive coffee furans, cafestol and kahweol, with acidic nitrite: Product characterization and factors controlling nitrosation versus ring-opening pathways. *Chemical Research In Toxicology* 22:1922–28