Journal of Natural Medicines
https://doi.org/10.1007/s11418-020-01409-6

ORIGINAL PAPER

# Identifying the compounds that can distinguish between Saposhnikovia root and its substitute, *Peucedanum ledebourielloides* root, using LC-HR/MS metabolomics

**Taichi Yoshitomi[1] · Daigo Wakana[2] · Nahoko Uchiyama[1] · Takashi Tsujimoto[1] · Noriaki Kawano[3] · Tsuguo Yokokura[4] · Yutaka Yamamoto[5] · Hiroyuki Fuchino[3] · Takashi Hakamatsuka[1] · Katsuko Komatsu[6] · Nobuo Kawahara[3] · Takuro Maruyama[1]**

Received: 4 November 2019 / Accepted: 5 April 2020


## Abstract

Previously, we established a $^{1}$H NMR metabolomics method using reversed-phase solid-phase extraction column (RP-SPEC), and succeeded in distinguishing wild from cultivated samples of Saposhnikoviae radix (SR), and between SR and its substitute, *Peucedanum ledebourielloides* root (PR). Herein, we performed LC-HR/MS metabolomics using fractions obtained via RP-SPEC to identify characteristic components of SR and PR. One and three characteristic components were respectively found for SR and PR; these components were isolated with their *m/z* values and retention times as a guide. The characteristic component of SR was identified as 4′-*O*-β-D-glucosyl-5-*O*-methylvisamminol (**1**), an indicator component used to identify SR in the Japanese Pharmacopoeia. In contrast, the characteristic components of PR were identified as xanthalin (**2**), 4′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosyl-5-*O*-methylvisamminol (**3**), and 3′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosylhamaudol (**4**) based on spectroscopic data such as 1D- and 2D-NMR, MS, and specific optical rotation. Among them, **4** is a novel compound. For the correlation between the NMR metabolomics results in the present and our previous report, only **1** and **2** were found to correlate with the chemical shifts, and the other compounds had no correlation. As the chemical shifts for compounds **1**, **3**, and **4** were similar to each other, especially for the aglycone moiety, they could not be distinguished because of the sensitivity and resolution of $^{1}$H NMR. Accordingly, combining NMR and LC/MS metabolomics with their different

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s11418-020-01409-6) contains supplementary material, which is available to authorized users.

✉ Takuro Maruyama
t-maruya@nihs.go.jp

[1] Division of Pharmacognosy, Phytochemistry and Narcotics, National Institute of Health Sciences, 3-25-26 Tonomachi, Kawasaki-ku, Kawasaki, Kanagawa 210-9501, Japan

[2] Faculty of Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa-ku, Tokyo 142-8501, Japan

[3] Research Center for medicinal Plant Resources, National Institutes of Biomedical Innovation, Health and Nutrition, 1-2 Hachimandai, Tsukuba, Ibaraki 305-0843, Japan

[4] Nippon Funmatsu Yakuhin Co., Ltd, 71-3 Izumicho, Yao-shi, Osaka 581-0813, Japan

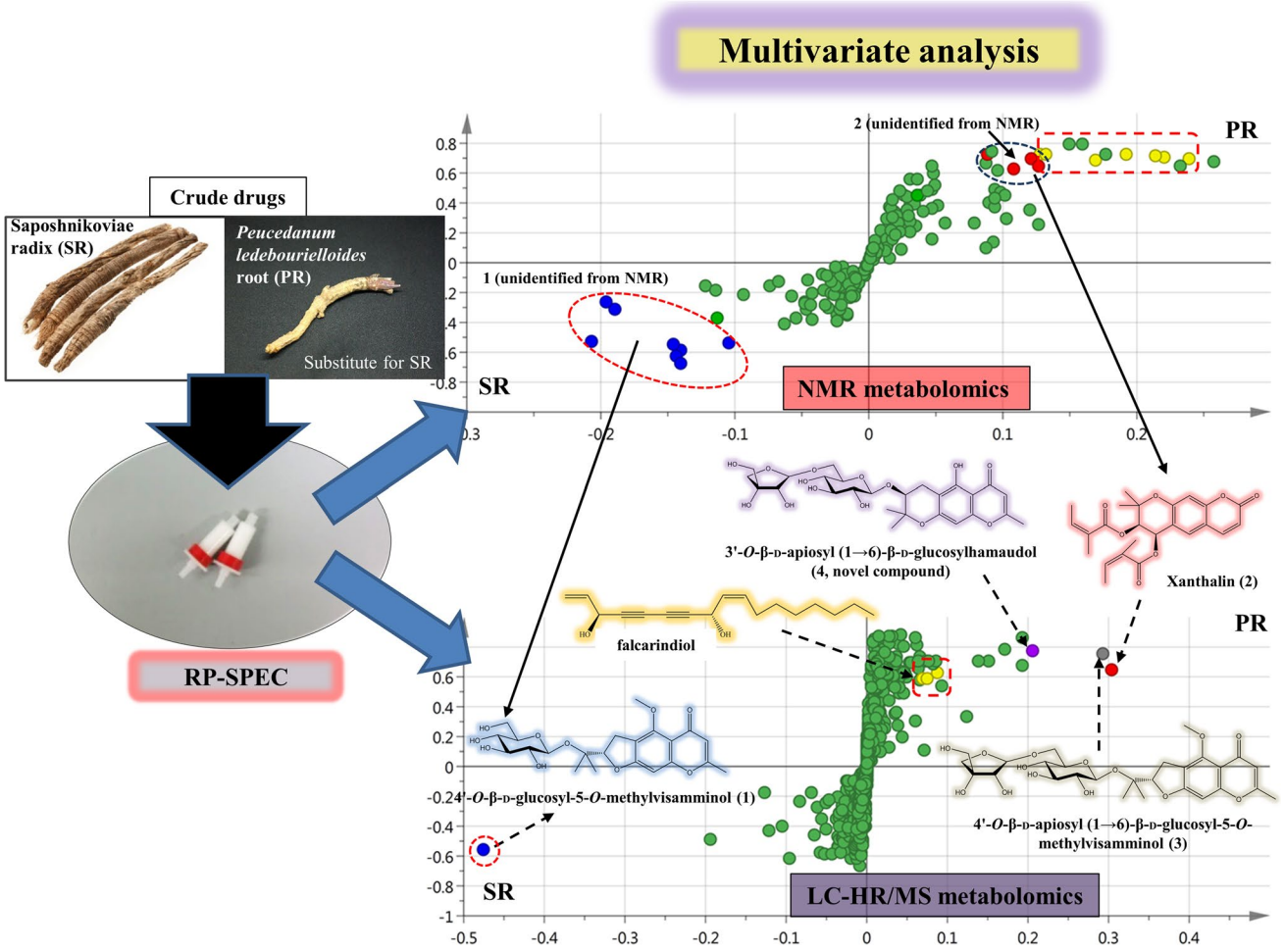
[5] Tochimoto Tenkaido Co., Ltd, Oniya Kaibara-Cho, Tamba, Hyogo 669-3315, Japan

[6] Department of Medicinal Resources Institute of Natural Medicine, University of Toyama, 2630 Sugitani, Toyama 930-0194, Japan

Published online: 24 April 2020

advantages is considered useful for metabolomics of natural products. The series of methods used in our reports could aid in quality evaluations of natural products and surveying of marker components.

**Graphic abstract**




**Keywords** $^{1}$H NMR-based metabolomics · LC-HR/MS metabolomics · Reversed-phase solid-phase extraction · Crude drugs · Saposhnikoviae radix (SR) · *Peucedanum ledebourielloides* root (PR)

## Introduction

In the 17th edition of the Japanese Pharmacopoeia (JP; The Ministry of Health, Labor and Welfare, Japan, 2016), Saposhnikoviae radix (SR) is defined as a crude drug derived from the roots and rhizomes of *Saposhnikovia divaricata* Schischkin (Umbelliferae; Apiaceae). SR is a crucial crude drug in oriental medicine and is utilized in many Kampo medicines (traditional Japanese medicines) owing to its perspiration and antipyretic actions [1]. Coumarins, chromones, essential oils, and polyacetylene compounds have been reported as ingredients in apiaceous plants [2] and are contained in SR. Besides, their pharmacological activities such as anti-inflammatory and antioxidant actions have been reported [2, 3]. On the other hand, SR is known to have a variety of components. In fact, Kim et al. and Batsukh et al. reported the chemical diversity in *S. divaricata* roots from Mongolia, China, and South Korea [4, 5]. Other apiaceous plants that naturally grow in China have been traditionally used as substitutes for SR in local regions [6–9]. Some of the SR available in Shaanxi and the surrounding regions were found to be derived from *Peucedanum ledebourielloides* root (PR) [9]. Owing to this background, constructing an objective method to clearly distinguish between SR and its substitutes is desirable.

In our previous paper, we applied $^{1}$H NMR metabolomics using a reversed-phase solid-phase extraction column (RP-SPEC) to SR and PR samples, and successfully discriminated between wild and cultivated samples of SR, and SR and PR samples [10]. In the report, saccharides, whose main content is sucrose, were identified as the components that aided in discriminating between wild and cultivated samples through comparison to authentic compounds. In contrast, the components contributing to the discrimination between SR and PR were assumed to be the secondary metabolites based on their chemical shifts in the s-plot; these, however, could not be identified. This is because components in apiaceous plants have high structural similarity, and obtaining the standard was difficult. Recently, Batsukh et al. investigated the chemical diversity of SR from different regions of Mongolia by multivariate analysis of LC/MS data. In the study, isolation of the characteristic components was needed to achieve unambiguous identification [5]. Furthermore, since the solvents used in metabolomics differ from those used in structural analysis, a direct comparison is difficult to perform. Thus, as we could not resolve these factors, even when $^{1}$H NMR metabolomics using RP-SPEC was adopted, their identification is still warranted.

In the present study, we performed LC-HR/MS metabolomics using RP-SPEC to obtain fractions to identify the compounds contributing to the discrimination between SR and PR. We succeeded in identifying 4 components including a novel compound. We also discussed the features of LC-HR/MS and $^{1}$H NMR metabolomics.

## Materials and methods

### General experimental procedures

Nuclear magnetic resonance (NMR) spectra were obtained using ECZ-600 and ECA-800 NMR Spectrometers (JEOL). Sep-Pak® plus C18 environmental cartridges (Waters) and vacuum manifold (GL Sciences) were used to pretreat the crude drug extracts. Automatic flash purification and recycling preparative HPLC were performed using Isolera one system (Biotage) equipped with a SNAP Ultra flash chromatography cartridge (25, 50, and 100 g), and the LaboAce LC-5060 system (Japan Analytical Industry) equipped with JAIGEL-ODS-AP-L (20 mm ID × 500 mm; Japan Analytical Industry), respectively. LC-HR/MS analysis was performed using the Ultimate 3000 HPLC system (Thermo Fisher Scientific) linked to an Orbitrap XL mass spectrometer with electron-transfer dissociation (Thermo Fisher Scientific) and equipped with Cortecs C18, 1.6 µm, 2.1 × 100 mm column (Waters). Specific optical rotation was measured by DIP-370 (JASCO). Methanol, *n*-hexane, chloroform, ethyl acetate, acetonitrile, and *n*-BuOH were purchased from Kanto Chemicals. Methanol-$d_4$, chloroform-*d*, and DMSO-$d_6$ were purchased from Taiyo Nippon Sanso. Ultrapure water was prepared by Integral-5 (Merck). 4′-*O*-β-D-glucosyl-5-*O*-methylvisamminol was purchased from Wako Pure Chemical, and 4′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosyl-5-*O*-methylvisamminol was purchased from Chem Scene. Falcarindiol was purchased from Wuhan ChemFaces. D- and L-Apiose were purchased from Carbosynth. L-Glucose was purchased from Sigma-Aldrich. D-Glucose was purchased from Kanto Chemical Co., Inc.

### Source and documentation

SR and PR samples were provided by the National Institutes of Biomedical Innovation, Health and Nutrition, Nippon Funmatsu Yakuhin Co. Ltd, and Tochimoto Tenkaido Co. Ltd. Further details are provided elsewhere [10].

### Sample preparation for LC-HR/MSS metabolomics

Solution C (fraction eluted with 100% methanol using solid-phase extraction) of SR and PR prepared in our previous study [10] was diluted fivefold with methanol for HPLC and filtered with a 0.45-µm filter to prepare the analytical samples (approximately 1 mg/ml; crude drug weight/solvent volume). In addition, reference solution for the alignment of chromatograms was prepared by mixing all sample solutions.

### Liquid chromatography/high-resolution mass spectrometry (LC-HR/MS) analysis

All sample solutions were analyzed using Ultimate 3000 HPLC with an Orbitrap high-resolution mass spectrometer and the Cortecs 1.6 µm, C18, 2.1 × 100 mm column maintained at 40 °C. Autosampler temperature was set to 4 °C and injection volume was 1 µl. Mobile phase consisting of 0.1% formic acid (solvent A) and acetonitrile containing 0.1% formic acid (solvent B) was run at a flow rate of 0.35 ml/min under the following gradient condition: 5% B, 3.5 min; 5–70% B, 10 min; 70–100% B, 4.5 min; and 100% B, 3.5 min. Electrospray ionization mass spectrometry (ESI–MS) parameters were: ion polarity, positive; scan range, *m/z* 100 to 1000; capillary temperature, 400 °C; sheath gas, 50 l/min; AUX gas, 25 l/min; capillary voltage, 15 V; and MS resolution, 60,000. All samples were analyzed in triplicate.

### Multivariate statistical analysis

Progenesis QI (Waters) was used for alignment and peak extraction from the chromatogram. For peak extraction, base peak mode (0.01%) was used. The obtained peaks were

converted into a data matrix by EZinfo statistical package (Umetrics). PCA and PLS analyses were carried out using SIMCA-14 software (Umetrics, Umeå, Sweden). Pareto scaling was applied as a preprocessing method and the number of latent variables was fixed as 3.

## Isolation of compound 1

SR powder (112.2 g) was extracted three times with methanol (500 ml) by sonication for 30 min. All extracts were combined and then evaporated under reduced pressure to yield the crude dry extract (12.8 g). Liquid–liquid partitioning of the water suspension was carried out with ethyl acetate, followed by *n*-BuOH. The *n*-BuOH extract (6.35 g) was then suspended in chloroform (3.0 ml), mixed with silica gel (8.0 g), added to a SNAP empty Samplet (10 g; Biotage), and used in flash chromatography. Mobile phase consisting of chloroform (A) and methanol (B) was used at a flow rate of 100 ml/min with A/B = 88/12 for 10 min. The fraction containing **1** was obtained at 5 min (276.4 mg), and was applied to recycling preparative HPLC with acetonitrile/water = 80/20 at a flow rate of 6.5 ml/min to isolate **1** (90.0 mg).

**4′-*O*-β-D-Glucosyl-5-*O*-methylvisamminol (1)**: white amorphous powder; $[\alpha]_D^{26}$ + 53.4, *c* 1.30, in methanol ($[\alpha]_D^{25}$ + 88.2, *c* 1.66, in methanol [19]), HR-ESI–MS *m/z* 453.1771 $[M+H]^+$ (Calcd. $C_{22}H_{29}O_{10}$, 453.1761). The chemical shifts from $^1$H and $^{13}$C NMR were identical to the published data [12].

## Isolation of compound 2

PR powder (8.5 g) was extracted three times with ethyl acetate (100 ml) by stirring for 60 min followed by evaporation under reduced pressure to yield the crude dry extract (1.27 g). The extract was then suspended in ethyl acetate (1 ml) and applied to a SNAP samplet HP-SIL (3 g; Biotage). After the removal of ethyl acetate under reduced pressure, the samplet was applied to flash chromatography. Mobile phase consisting of *n*-hexane and ethyl acetate mixture (4/1) was used at 75 ml/min for 5 min. The fraction collected during the first 3 min (23.0 mg) was again separated using ODS flash chromatography with water (A) and acetonitrile (B) at 25 ml/min under the following gradient condition: 40% B, 18 min; 40–60% B, 22 min; 60–75% B, 10 min. The peak at 36 min was collected and **2** was isolated (5.7 mg).

**Xanthalin (2)**: colorless amorphous powder; $[\alpha]_D^{24}$ − 91.9, *c* 0.73, in ethanol ($[\alpha]_D^{20}$ − 164.2, *c* 0.96, in ethanol [20]), HR-ESI–MS *m/z* 427.1757 $[M+H]^+$ (Calcd. $C_{24}H_{27}O_7$, 427.1757). The chemical shifts from $^1$H and $^{13}$C NMR were identical to the published data [9].

Isolation of compounds **3** and **4.**

PR powder (131.5 g) was extracted three times with chloroform/methanol = 1/1 (600 ml) by sonication for 30 min. All extracts were then combined and evaporated under reduced pressure to yield the crude dry extract (12.8 g). Liquid–liquid partitioning of the water suspension was carried out with ethyl acetate, followed by *n*-BuOH. The *n*-BuOH extract (2.63 g) was suspended in methanol (2.5 ml) and applied to a SNAP samplet HP-SIL (10 g; Biotage). After methanol removal under reduced pressure, the samplet was applied to SNAP Ultra 50 g (Biotage). Chloroform (A) and methanol (B) were used as the mobile phase at a flow rate of 100 ml/min under the following gradient condition: 12% B, 0–12 min; 12–15% B, 10 min; 20% B, 7 min. Fraction A (25 min) for **3** and fraction B (7 min) for **4** were obtained. Fraction A was again fractionated using SNAP Ultra 100 g (Biotage), with mobile phase, chloroform (A) and methanol (B), at 100 ml/min under the following gradient condition: 12% B, 0–10 min; 12–14% B, 17 min; 20% B, 5 min. Compound **3** was isolated in the 20% B fraction (15.6 mg) which was further fractionated using SNAP Ultra C18 30 g (Biotage). Mobile phase consisting of water (A) and acetonitrile (B) was used at 25 ml/min under the following gradient condition: 18% B, 0–24 min; 18–20% B, 15 min; **4** (12.4 mg) was obtained at 36 min.

**4′-*O*-β-D-Apiosyl (1 → 6)-β-D-glucosyl-5-*O*-methylvisamminol (3)**: white amorphous powder; $[\alpha]_D^{24}$ − 34.9, *c* 1.01, in methanol ($[\alpha]_D^{28}$ − 4.17, *c* 0.96, in methanol [7]), HR-ESI–MS *m/z* 585.2175 $[M+H]^+$ (Calcd. $C_{27}H_{37}O_{14}$, 585.2183), $^1$H NMR (DMSO-$d_6$, 600 MHz) $_H$: 6.67 (1H, s, H-8), 5.96 (1H, s, H-3), 4.83 (1H, t, *J* = 9.0 Hz, H-2′), 4.80 (1H, d, *J* = 3.0 Hz, H-1′′′), 4.40 (1H, d, *J* = 7.2 Hz, H-1′′), 3.82 (1H, d, *J* = 9.6 Hz, H-4′′′), 3.85 (3H, s, 5-$OCH_3$), 3.68 (1H, d, *J* = 3.0 Hz, H-2′′′), 3.66 (1H, br d, *J* = 9.6 Hz, H-6′′), 3.56 (1H, d, *J* = 9.6 Hz, H-4′′′), 3.38 (1H, m, H-3′), 3.32 (2H, m, H-5′′′), 3.29 (1H, m, H-6′′), 3.24 (2H, m, H-3′, H-5′′), 3.14 (1H, t, *J* = 9.0 Hz, H-3′′), 2.97 (1H, t, *J* = 9.0 Hz, H-4′′), 2.87 (1H, t, *J* = 8.1 Hz, H-2′′), 2.26 (3H, s, 2-$CH_3$), 1.26 (3H, s, 4′-*gem* $CH_3$), 1.23 (3H, s, 4′-*gem* $CH_3$); $^{13}$C NMR (DMSO-$d_6$, 150 MHz) $\delta_C$: 175.3 (C-4), 164.0 (C-7), 163.3 (C-2), 158.9 (C-8a), 155.0 (C-5), 117.1 (C-6), 110.9 (C-4a), 110.7 (C-3), 109.1 (C-1′′′), 97.1 (C-1′′), 93.1 (C-8), 89.9 (C-2′), 78.6 (C-3′′′), 76.8 (C-4′), 76.7 (C-3′′), 75.8 (C-2′′′), 75.0 (C-5′′), 73.3 (C-2′′), 73.1 (C-4′′′), 70.0 (C-4′′), 67.4 (C-6′′), 63.1 (C-5′′′), 60.1 (C-5-$OCH_3$), 27.0 (C-3′), 22.6 (C-4′-*gem* $CH_3$), 21.7 (C-4′-*gem* $CH_3$), 19.1 (2-$CH_3$).

**3′-*O*-β-D-Apiosyl (1 → 6)-β-D-glucosylhamaudol (4)**: white amorphous powder; HR-ESI–MS *m/z* 571.2018 $[M+H]^+$ (Calcd. $C_{26}H_{35}O_{14}$, 571.2027). $^1$H NMR (DMSO-$d_6$, 600 MHz), $\delta_H$: 13.1 (1H, s, 5-OH), 6.36 (1H, s, H-8), 6.17 (1H, s, H-3), 4.92 (1H, d, *J* = 3.0 Hz, H-1′′′), 4.34 (1H, d, *J* = 7.8 Hz, H-1′′), 3.91 (1H, t, *J* = 5.4 Hz, H-3′), 3.85 (1H, dd, *J* = 10.8, 1.2 Hz, H-6′′), 3.83 (1H, d, *J* = 9.6 Hz, H-4′′′),

3.75 (1H, dd, $J$ = 6.0, 2.4 Hz, H-2′′′), 3.60 (1H, d, $J$ = 9.0 Hz, H-4′′′), 3.44 (1H, dd, $J$ = 10.8, 7.2 Hz, H-6′′), 3.34 (2H, m, H-5′′′), 3.32 (1H, m, H-5′′), 3.18 (1H, ddd, $J$ = 8.7, 4.2, 4.2 Hz, H-3′′), 3.02 (1H, ddd, $J$ = 7.2, 5.4, 4.8 Hz, H-4′′), 2.94 (1H, ddd, $J$ = 8.4, 5.4, 5.4 Hz, H-2′′), 2.84 (1H, dd, $J$ = 17.2, 5.4 Hz, H-4′), 2.64 (1H, dd, $J$ = 17.2, 5.4 Hz, H-4′), 2.35 (3H, s, 2-$CH_3$), 1.33 (3H, s, 2′-*gem* $CH_3$), 1.31 (3H, s, 2′-*gem* $CH_3$); $^{13}$C NMR (DMSO-$d_6$, 150 MHz) $\delta_C$: 181.9 (C-4), 167.8 (C-2), 158.7 (C-5, C-8a), 155.4 (C-7), 109.2 (C-1′′′), 107.7 (C-3), 103.4 (C-4a, C-6), 100.9 (C-1′′), 94.3 (C-8), 78.8 (C-3′′′), 77.9 (C-2′), 76.8 (C-3′′), 75.8 (C-2′′′), 75.6 (C-5′′), 73.3 (C-4′′′), 73.2 (C-3′), 73.1 (C-2′′), 70.3 (C-4′′), 67.5 (C-6′′), 63.3 (C-5′′′), 25.2 (C-2′-*gem* $CH_3$), 22.0 (C-2′-*gem* $CH_3$), 21.7 (C-4′), 20.0 (2-$CH_3$).

### Hydrolysis of compound 4

We added 1,4-dioxane (1 ml) and 2 M HCl (1 ml) to **4** (19.2 mg), and refluxed the mixture for 3 h with a mantle heater. The reaction mixture was then extracted with water/ethyl acetate = 1/1 (2 ml), and the ethyl acetate layer was used in recycling preparative HPLC with acetonitrile at a flow rate of 6.25 mL/min to obtain compound **4a** (1.8 mg).

**Hamaudol (4a)**: white amorphous powder; $[\alpha]_D^{25}$ −39.8, $c$ 0.625, in ethanol ($[\alpha]_D^{24}$ −32.0, $c$ 0.13, in ethanol [8]), ESI–MS $m/z$ 277.1 $[M+H]^+$. $^1$H NMR (chloroform-$d$, 600 MHz) $\delta_H$: 13.0 (1H, s, 5-OH), 6.30 (1H, s, H-8), 5.97 (1H, s, H-3), 3.86 (1H, br-d, $J$ = 4.8 Hz, H-3′), 2.95 (1H, dd, $J$ = 17.4, 4.8 Hz, H-4′), 2.73 (1H, dd, $J$ = 17.4, 5.4 Hz, H-4′), 2.31 (3H, s, 2-$CH_3$), 1.37 (3H, s, 2′-*gem* $CH_3$), 1.33 (3H, s, 2′-*gem* $CH_3$); $^{13}$C NMR (chloroform-$d$, 150 MHz) (Table 1).

### Determination of D, L-sugar configuration for compounds 1, 3 and 4

The configuration of sugar moiety was determined according to the report of Tanaka et al.and Wang et al. [21, 22]. The compounds **1**, **3** and **4** were hydrolyzed by heating with 2 M HCl at 75 °C for 6 h and neutralized with 3.5 M ammonium hydroxide. After the evaporation of reaction mixture, the residue was dissolved in pyridine (2 mL) and added to L-cysteine methyl ester hydrochloride (15 mg), and heated at 80 °C for 1 h. Then, phenyl isothiocyanate (30 µL) was added to the reaction mixture, and was further heated at 80 °C for 1 h. The reaction mixture was directly analyzed under the following conditions by LC-HR/MS: Acquity UPLC BEH C18, 1.7 µm, 2.1 × 100 mm (Waters) was used as the analytical column at 40 °C. Mobile phase consisting of 0.1% formic acid (solvent A) and acetonitrile/methanol/isopropanol (50:25:25) containing 0.05% formic acid (solvent B) was run at a flow rate of 0.3 ml/min under the following gradient condition: 14%–16.5% B (0–23 min) and then 100% B (4 min hold). These derivatizing reactions and analyzes were also applied to the authentic D- and L-glucose and apiose. The retention times of the authentic standard derivatives were 15.78 min for D-glucose, 14.36 min for L-glucose, 20.97 min for D-apiose and 12.30 min for L-apiose.

## Results and discussion

### Quantitative range and reproducibility in LC-HR/MS analysis

First, we investigated the dynamic range and reproducibility of data acquisition for multivariate analysis. ESI was selected as the MS interface because of its wide ionization range from low to high polar compounds. The gradient program of 5% B to 100% B was also adopted as the mobile phase to cover the compounds' wide polarity range. Representative total ion current (TIC) chromatograms of SR and PR are shown in Fig. S1; the peak with the higher intensity (7.38 min; $m/z$ 453.1798 of the pseudo-molecular ion) and that with the relatively lower intensity (8.13 min; $m/z$ 291.1264 of the pseudo-molecular ion) were selected. Calibration curves were then prepared for peak areas and sample concentrations (Fig. S2). Good linearity was obtained at 0.001–10 mg/ml for SR and PR (Fig. S2). Therefore, 1.0 mg/ml was used as the sample concentration for multivariate analysis to ensure quantitative linearity in the concentration variation of the components from one-tenth to tenfold.

To verify the repeatability of the data, samples were repeatedly analyzed and variations in retention time were assessed (Fig. S2). A great change in retention time was not observed and the high reproducibility was also confirmed. In LC-HR/MS metabolomics, the data matrix consists of the qualitative parameter (retention time, $m/z$) and quantitative parameter (peak area). Data acquisition with the appropriate dynamic range and reproducibility is thus important to perform the correct analysis. Based on the above investigations, we deduced that the quality of the qualitative and quantitative parameters was sufficient for multivariate analysis.

### Peak extraction using parameters of PLS-DA

Subsequently, we determined the peak numbers in LC-HR/MS metabolomics via two extraction methods: (1) automatic extraction by software (normal mode and fewer mode; conditions A and B, respectively), (2) base peak-based extraction where peaks with intensities greater than 0.1% and 0.01% by the base peak, were extracted (conditions C and D). The respective peak numbers under conditions A to D were 9421, 7903, 230, and 1387. To estimate the optimum peak numbers, we determined the multiple correlation coefficient ($R^2$) and predictive ability parameter ($Q^2$) (Table S1). Values of $R^2$ and $Q^2$ were the same degree in any conditions,

**Table 1** Comparison of the $^{13}C$ NMR ( ppm in DMSO-$d_6$ or chloroform-*d*, 150 MHz) data for compounds **4** and **4a** to the literature values

| DMSO-$d_6$ | | | Chloroform-*d* | | |
|---|---|---|---|---|---|
| Position | Compound **4** | *sec*-*O*-glucosyl-hamaudol[1] | Hydrolyzate of Compound **4** (**4a**) | Hamaudol [2)] | Visamminol [2)] |
| 1 | – | – | – | – | – |
| 2 | 167.8 | 167.9 | 166.8 | 166.8 | 165.8 |
| 3 | 107.8 | 107.9 | 108.4 | 108.3 | 108.9 |
| 4 | 181.9 | 181.9 | 182.6 | 182.5 | 182.6 |
| 4a | 103.4 | 103.5 | 104.4 | 104.3 | 105.5 |
| 5 | 158.7 | 158.7 | 159.8 | 159.7 | 158.3 |
| 6 | 103.4 | 103.4 | 102.9 | 102.9 | 108.7 |
| 7 | 155.5 | 155.5 | 159.0 | 159.0 | 166.5 |
| 8 | 94.4 | 94.3 | 94.8 | 94.8 | 91.7 |
| 8a | 158.7 | 158.8 | 156.3 | 156.2 | 156.7 |
| 1’ | – | – | – | – | – |
| 2’ | 77.9 | Missing | 78.6 | 78.5 | 88.9 |
| 3’ | 73.2 | 72.7 | 68.8 | 68.7 | 20.4 |
| 4’ | 21.7 | 21.5 | 25.4 | 25.4 | 71.9 |
| Gem-$CH_3$ | 25.1 | 25.3 | 24.9 | 25.0 | 25.9 |
| Gem-$CH_3$ | 22.1 | 21.7 | 22.1 | 22.1 | 26.8 |
| 2-$CH_3$ | 20.0 | 20.0 | 20.6 | 20.6 | 23.9 |
| Glucose | | | | | |
| 1’’ | 100.9 | 100.6 | | | |
| 2’’ | 73.1 | 73.4 | | | |
| 3’’ | 76.8 | 76.9 | | | |
| 4’’ | 70.3 | 70.3 | | | |
| 5’’ | 75.7 | 76.9 | | | |
| 6’’ | 67.6 | 61.4 | | | |
| Apiose | | | | | |
| 1’’’ | 109.2 | | | | |
| 2’’’ | 75.8 | | | | |
| 3’’’ | 78.8 | | | | |
| 4’’’ | 73.2 | | | | |
| 5’’’ | 63.3 | | | | |

[1]Chem Pharm Bull, 2001, 49(2), 154–160

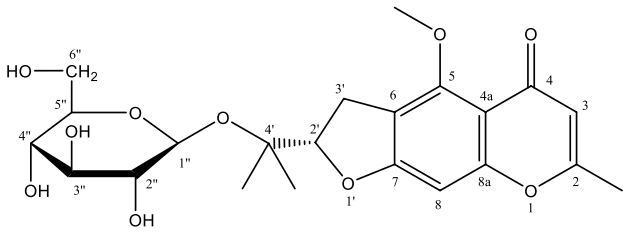
[2]Nat Pro Sci, 2017, 23(2), 97–102

and peak numbers greatly increased when conditions A and B were selected. Excessive peak extraction could induce the picking of the carry-over peaks derived from previous analyses. Therefore, condition D with a moderate number of peaks was selected as the peak extraction method.

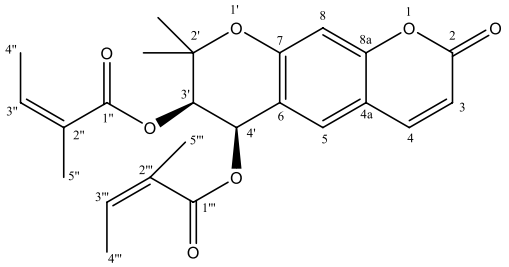
## PCA and PLS-DA in LC-HR/MS metabolomics

PCA was carried out using the resultant data matrix with the above-mentioned conditions, and the entire data were reviewed. SR (green circles) and PR (blue circles) samples were roughly separated on the PC1 vs PC2 plane (Fig. 1a); such discrimination was similar to that of $^1H$ NMR metabolomics [10]. For PC1, a positive correlation was confirmed between the sum of all peak areas obtained by LC/MS data and PC1 score (Fig. S3). Therefore, PC1 likely captured the total amount of the detected variables.

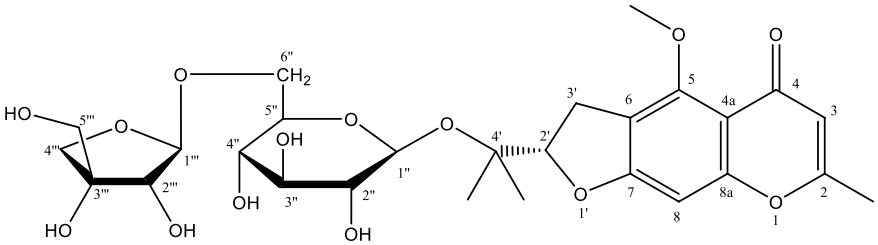
We then performed PLS-DA to search for components that could help to distinguish SR from PR. The PLS-DA score plot had a clearer separation of the SR and PR samples (Fig. 1b). Furthermore, the s-plot indicated the following four characteristic peaks: *m/z* 453.1750 ($C_{22}H_{28}O_{10}$; **1**) for SR, and *m/z* 427.1757 ($C_{24}H_{26}O_7$; **2**), 571.2071 ($C_{26}H_{34}O_{14}$; **4**), and 585.2175 ($C_{27}H_{36}O_{14}$; **3**) for PR (Fig. 2). These components were at the edge of the vertical and horizontal axes, and were explanatory variables that strongly contributed to the objective variable.



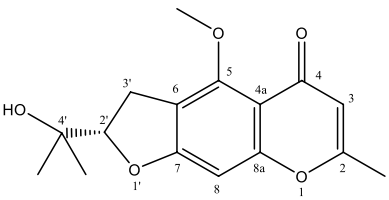
**1** (4'-*O*-β-D-glucosyl-5-*O*-methylvisamminol)



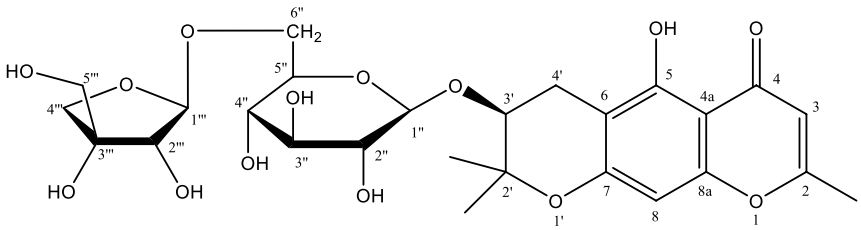
**2** (xanthalin)



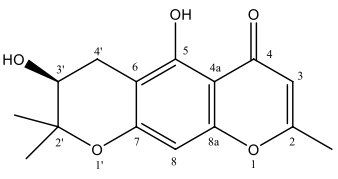
**3** (4'-*O*-β-D-apiosyl (1→6)-β-D-glucosyl-5-*O*-methylvisamminol)



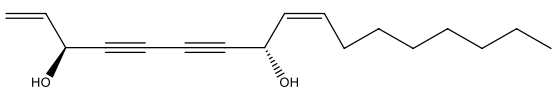
visamminol



**4** (3'-*O*-β-D-apiosyl (1→6)-β-D-glucosylhamaudol)



hamaudol



falcarindiol

**Fig. 1** Structures of the components used to distinguish between SR and PR in LC-HR/MS metabolomics

## Identification of the characteristic component in SR

As the characteristic compound **1** for SR had $C_{22}H_{28}O_{10}$ as the molecular weight due to its exact mass, we attempted a database search using the compositional formula obtained and the keyword "*Saposhnikovia*". With the dramatic advance made using mass spectrometers, the database constructed using exact mass was remarkably developed,

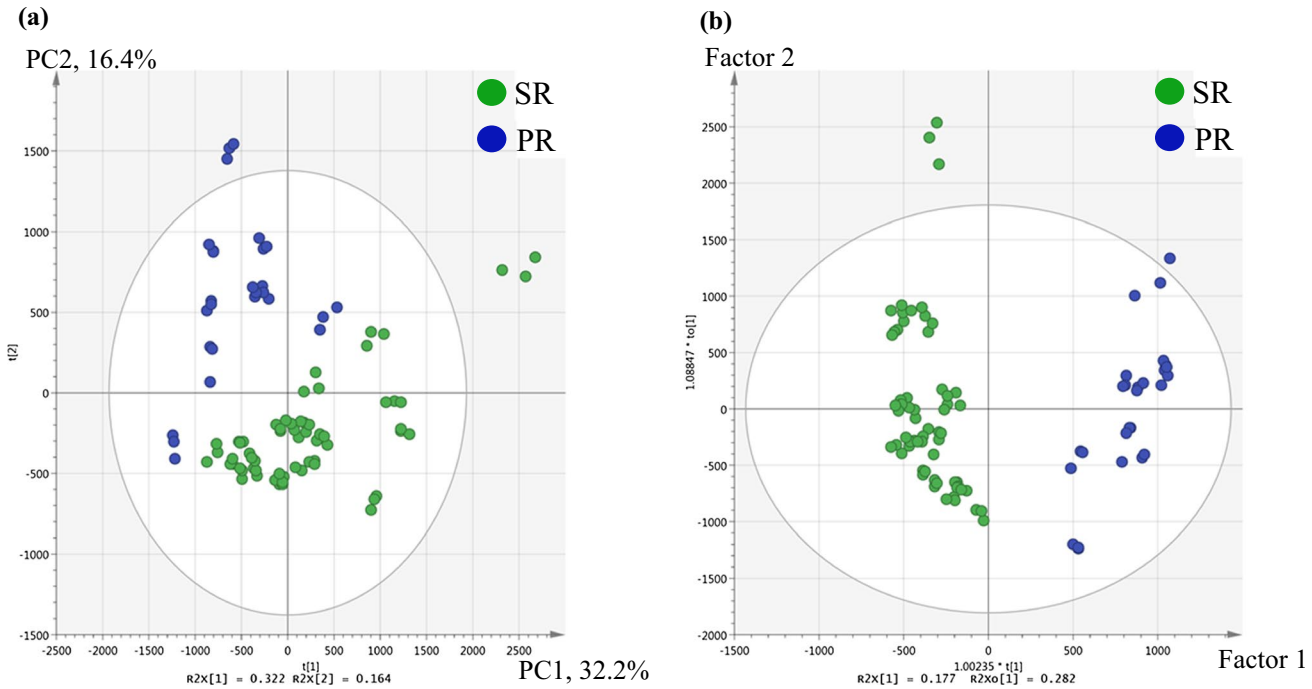



**Fig. 2** Score plots of PCA (a) and PLS-DA (b) for LC-HR/MS metabolomics. Green and blue circles indicate SR and PR samples, respectively

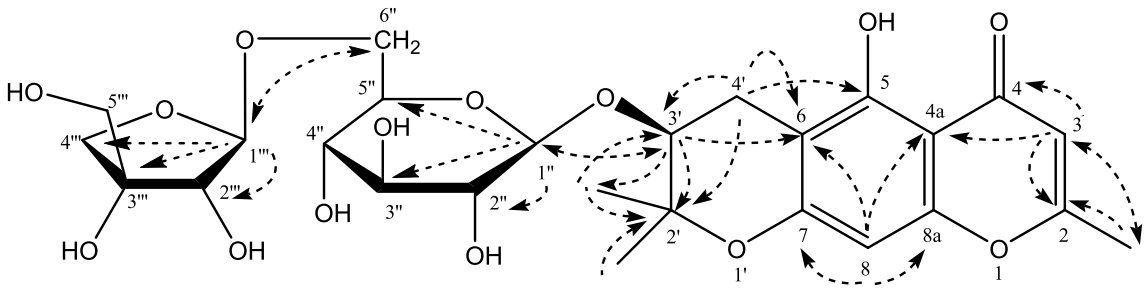


**Fig. 3** HMBC correlation of compound **4** (only correlations for the anomeric position are described for the sugar parts)

allowing a more efficient search [11]. By searching several databases, **1** was estimated as 4′-*O*-β-D-glucosyl-5-*O*-methylvisamminol (Fig. 3), and when we isolated **1** from the methanol extract of SR, this compound was confirmed by NMR and MS [12] (Fig. S4a). In addition, we also determined the D, L-configuration of the sugar moiety. The sugar generated from the hydrolyzation of **1** was converted to the diastereomeric derivative, and its configuration was determined by comparing the retention times of the obtained derivative in LC with those of the authentic D- and L-glucose derivatives [21, 22]. As the results, the retention time of sugar derivative from **1** was identical to the D-glucose. In the 17$^{th}$ edition of JP, **1** is used as an indicator component in the test to identify SR [13].

Compound **1** was detected in both SR and PR from the TICs (Fig. S4a). We calculated the average of all peak areas for **1** in the SR and PR samples, and verified that mean content of **1** in SR samples was much higher than in PR samples ($p < 0.005$ by $t$ test) (Fig. S4b). We think that the significant difference of **1** content between SR and PR is the reason for the picking of **1** as characteristic compound for SR.

## Identification of characteristic components in PR

Three characteristic compounds of PR were found using s-plot: *m/z* 427.1757 ($C_{24}H_{26}O_7$; **2**), 571.2071 ($C_{26}H_{34}O_{14}$; **4**), and 585.2175 ($C_{27}H_{36}O_{14}$; **3**). Thus, we performed a search for the compositional formulae and used the keyword, "*Peucedanum*". As a result, **2** was estimated to be a coumarin derivative, and was identified as xanthalin by MS and NMR (Fig. 3) [9]. Furthermore, we compared the retention time of **2** and the corresponding peak from sample solutions,

and observed a good alignment (Fig. S5). In supplement II of the 17th edition of JP, PR was defined as an impurity of SR and a purity test using xanthaline as the marker compound by TLC was designed [14].

For the identification of **3** and **4,** these compounds were isolated by normal and RP column chromatography. In the $^{1}$H NMR of **3**, the following signals were observed: (1) aromatic or olefinic protons at $\delta_H$ 5.96 (1H, s) and 6.67 (1H, s), (2) methoxy group at $\delta_H$ 3.85 (3H, s), and (3) methyl group on the $sp^2$ carbon at $\delta_H$ 2.26 (3H, s). These patterns were similar to the visamminol skeleton of **1** which was the characteristic component of SR. Additionally, two anomeric protons were found at $\delta_H$ 4.40 (1H, d, $J$ = 7.2 Hz) and 4.80 (1H, d, $J$ = 3.0 Hz). In 2D-NMR and the coupling pattern, glucose and apiose were identified as the sugars. Hence, we assumed that **3** was 4′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosyl-5-*O*-methylvisamminol. By comparing LC, MS pattern, and NMR data for standard **3** to those of isolated **3**, this compound was identified as 4′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosyl-5-*O*-methylvisamminol (Fig. 3). Furthermore, we also confirmed that each sugar has D-configuration in the same way as **1**. Compound **3** has already been isolated from *Libanotis laticalycina* (Umbelliferae/Apiaceae) and this plant has been previously used in the Henan and Shanxi provinces as a substitute plant for SR [7].

In the $^{1}$H NMR analysis of **4**, the characteristic signals were: 1) aromatic or olefinic protons at $\delta_H$ 6.17 (1H, s) and $\delta_H$ 6.36 (1H, s), 2) hydrogen-bonded phenolic hydroxyl group at $\delta_H$ 13.1 (1H, s), 3) methyl group on the $sp^2$ carbon at $\delta_H$ 2.35 (3H, s), and 4) anomeric protons at $\delta_H$ 4.34 (1H, d, $J$ = 7.8 Hz) and 4.92 (1H, d, $J$ = 3.0 Hz). The aglycon moiety of **4** was considered to be hamaudol or visamminol based on the literature values (Table 1) [8]. In addition, the structure of the sugar moiety was similar to that of **3** based on the $^{1}$H and $^{13}$C NMR spectra and thus, were presumed to also be glucose and apiose. To confirm the structure of the aglycone moiety, we hydrolyzed compound **4** and obtained aglycone, **4a**. $^{13}$C NMR data of **4a** were compared to the literature values of hamaudol and visamminol [15]; **4a** coincided with hamaudol (Table 1). To determine the configuration of the 3′ position, we measured the specific optical rotation of **4a**. This property showed a good agreement with the literature value [8]. Based on the above data, **4a** was identified as hamaudol. We also confirmed reasonable correlation in the aglycone moiety for the heteronuclear multiple bond coherence (HMBC) of **4** (Fig. 4). The correlation between the 3′ position of aglycone and anomeric position of glucose (1″), and the 6-position of glucose (6″) and anomeric position of apiose (1‴) were also confirmed (Fig. 4). We determined that each sugar has D-configuration in the same way of **1** and

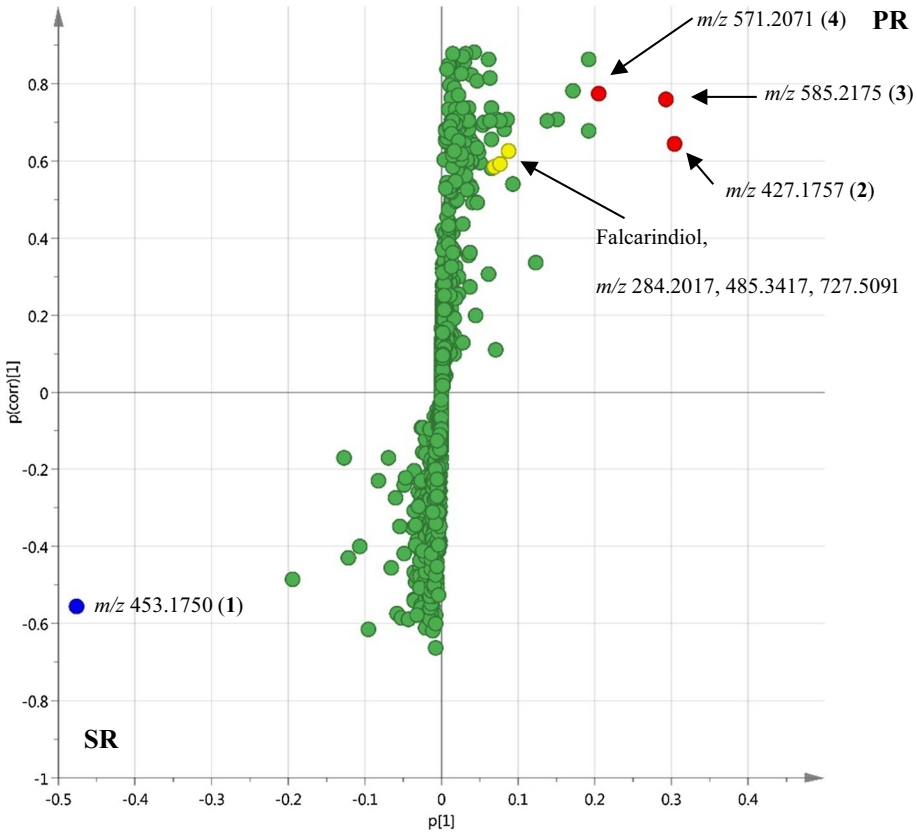



**Fig. 4** S-plots of PLS-DA for LC-HR/MS metabolomics. The characteristic variables are highlighted in blue, SR and red, PR; yellow, falcarindiol

**3**. Thus, **4** was identified as the novel compound, 3′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosylhamaudol (Fig. 3).

### Correlation between NMR and LC-HR/MS metabolomics with RP-SPEC

Finally, we reviewed the mass chromatograms of **2**–**4** from the SR and PR samples. Although only trace amounts of those compounds were detected in SR, the contents in PR were much higher than in SR (Fig. S6). Thus, **2**–**4** were indeed characteristic components of PR.

In the previous report [10], we were unable to adequately identify the variables contributing to the discrimination between SR and PR in $^{1}$H NMR metabolomics. Herein, we investigated the correlation between the components identified by LC-HR/MS metabolomics and the chemical shift obtained with $^{1}$H NMR metabolomics.

In the s-plot for $^{1}$H NMR metabolomics, the signals, 1.34, 2.30, 4.58, 6.02, and 6.57 ppm were confirmed at the edge on the SR side (blue circles, Fig. S7). These chemical shifts corresponded to those of **1** which is the characteristic component of SR in LC-HR/MS metabolomics (Table S2, Fig. S7). Thus, the variables obtained by the s-plot in $^{1}$H NMR metabolomics may have been derived from this compound.

For PR, the signals, 1.50–5.90 ppm, were confirmed at the edge of the s-plot (Fig. S7); some chemical shifts (1.50, 1.74, 1.86, 5.58 ppm) corresponded to those of **2** (red circles, Fig. S7), but most of the other shifts were identified as falcarindiol analogs (falcarindiol; yellow circles, Fig. S7) as described in our previous report [10]. Signals corresponding to **3** and **4** were not confirmed. We considered that this observation may be attributed to the structural similarity among **1**, **3**, and **4**. Most chemical shifts derived from these compounds were similar among each other, particularly the aglycone moiety which is indistinguishable due to the sensitivity and resolution of $^{1}$H NMR. Therefore, **1,** which is present in a high amount in SR, could be considered as the characteristic component of SR. Furthermore, **2,** which has a different skeleton, was recognized in $^{1}$H NMR metabolomics as the characteristic component of PR, supporting the speculation mentioned above.

The ions, *m/z* 284.2017, 485.3417, and 727.5091, in the s-plot of LC-HR/MS metabolomics may belong to falcarindiol (yellow circles, Fig. 2). Furthermore, *m/z* 243.1737 was assumed to be an ion where protons were added after the elimination of a water molecule $[M\text{-}H_2O + H]^+$; *m/z* 485.3417 and 727.5091 were expected to be the cluster ions (e.g., $[2(M\text{-}H_2O) + H]^+$, $[3(M\text{-}H_2O) + H]^+$) [16, 17]. Falcarindiol may not have been positioned at the edge of the s-plot for LC-HR/MS metabolomics because of the dispersed intensity of each ion by the generation of cluster ions. Falcarindiol is generally analyzed by APCI as the interface of MS, and cluster ions are not generated in this method [16, 17]. We analyzed the PR sample and standard falcarindiol with APCI and confirmed that cluster ions did not occur (Fig. S8). As APCI is suitable for analyzing hydrophobic compounds, it would be very effective in molecules such as falcarindiol [18]. Compared to NMR metabolomics, the need for optimized measurement conditions is one of the disadvantages of LC/MS metabolomics. Thus, combining NMR metabolomics which is a more comprehensive measurement method and LC/MS metabolomics with its high sensitivity and resolution, is considered to be a very useful approach for quality evaluation of natural products containing various components.

## Conclusions

We conducted LC-HR/MS metabolomics to identify the components used to distinguish between SR and PR which were found in the $^{1}$H NMR metabolomics carried out in our previous study [10]. As a result, **1** was identified as the characteristic component of SR, and its data correlated with the chemical shift obtained by $^{1}$H NMR metabolomics. In contrast, **2**–**4** were identified as characteristic components of PR, and **4** was identified as the novel compound, 3′-*O*-β-D-apiosyl (1 → 6)-β-D-glucosylhamaudol. Compounds **1** and **2** were found to correlate with the chemical shifts observed in the $^{1}$H NMR metabolomics, but a correlation was not found with the other compounds. The latter result could be attributed to the structural similarity of the components contained in both SR and PR, and the generation of cluster ions from falcarindiol in the ESI interface. It is important to note that compounds with high structural similarities such as coumarins and chromones are found in apiaceous plants. In this study, we succeeded in capturing the subtle component difference between SR and PR using a series of methods.

**Acknowledgements** This study was supported by a research grant from the Japan Agency for Medical Research and Development (the grant number: 17ak0101074j2001, 15ak0101017h0003).

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflicts of interest.

## References

1. Tai J, Cheung S (2017) Anti-proliferative and antioxidant activities of Saposhnikovia divaricata. Oncol Rep 18:227–234
2. Liao H, Li Q, Liu R, Liu J, Bi K (2014) Fingerprint analysis and multi-ingredient determination using a single reference standard for Saposhnikoviae Radix. Anal Sci 30:1157–1163
3. Kong X, Liu C, Zhang C, Zhao J, Wang J, Wan H, Zhu H, Zhang P, Chen W, Xiao Y, Lin N (2013) The suppressive effects of

*Saposhnikovia divaricata* (Fangfeng) chromone extract on rheumatoid arthritis via inhibition of nuclear factor-κB and mitogen activated proteinkinases activation on collagen-induced arthritis model. J Ethnopharmacol 148:842–850
4. Kim HS, Choi G, Lee AY (2018) Ultra-performance convergence chromatography method for the determination of four chromones and quality control of Saposhnikovia divaricata (Turcz.) Schischk. J Sep Sci 41:1682–1690
5. Batsukh Z, Toume K, Javzan B, Kazuma K, Cai S, Hayashi S, Kawahara N, Maruyama T, Komatsu K (2020) Metabolomic profiling of Saposhnikoviae Radix from Mongolia by LC–IT–TOF–MS/MS and multivariate statistical analysis. J Nat Med 74(1):170–188
6. Wang J, Materia LZ (1989) Medica research of Saposhunikoviae radix (provisional translation). Zhongguo Zhong Yao Za Zhi 14:579–581
7. Baba K, Qing X, Taniguchi M, Kozawa M, Fujita E (1991) Studies on Chiniese Medicine Fang-Feng (III) Constituents of Shui-Fang-Feng. Shoyakugaku Zasshi 45(2):167–173
8. Okuyama E, Hasegawa T, Matsushita T, Fujimoto H, Ishibashi Yamazaki M (2001) Analgesic Components of Saposhnikovia Root (*Saposhnikovia divaricata*). Chem Pharm Bull 49(2):154–160
9. Maruyama T, Ezaki M, Shiba M, Yamaji H, Yoshitomi T, Kawano N, Cheng SZX, Yokokura T, Yamamoto Y, Fuchino H, Sun H, Komatsu K, Kawahara N (2018) Botanical origin and chemical constituents of commercial SR and its related crude drugs available in Shaanxi and the surrounding regions. J Nat Med 72:267–273
10. Yoshitomi T, Wakana D, Uchiyama N, Tsujimoto T, Kawano N, Yokokura T, Yamamoto Y, Fuchino H, Hakamatsuka T, Komatsu K, Kawahara N, Maruyama T (2020) $^{1}$H NMR based metabolomic analysis coupled with reversed-phase solid-phase extraction for sample preparation of Saposhnikovia roots and related crude drugs. J Nat Med 74:65–75
11. Afendi F, Okada T, Yamazaki M, Hirai-Morita A, Nakamura Y, Nakamura K, Ikeda S, Takahashi H, Altaf-Ul-Amin M, Darusman L, Saito K, Kanaya S (2012) KNApSAcK family databases: integrated metabolite-plant species databases for multifaceted plant research. Plant Cell Physiol 53:e1–12
12. Sun A, Feng L, Liu R (2006) Preparative isolation and purification of prim-O-glucosyl-cinmifugin and 4'-O--D-glucosyl-5-O-methylvisamminol from Radix saposhnikoviae by high speed countercurrent chromatography. J Liq Chromatogr Rel Technol 29:751–759
13. The Ministry of Health, Labor and Welfare, Japan (2016) The Japanese pharmacopoeia seventeenth edition
14. The Ministry of Health, Labor and Welfare in Japan (2019) Supplement II to the Japanese Pharmacopoeia seventeenth edition. The MHLW Notification No. 49
15. Kwon Y, Kim H, Kim M, Chun W (2017) Acetylcholinesterase Inhibitors from *Angelica polymorpha* Stem. Nat Pro Sci 23(2):97–102
16. Kramer M, Mühleis A, Conrad J, Leitenberger M, Beifuss U, Carle R, Kammerer DR (2011) Quantification of polyacetylenes in apiaceous plants by high-performance liquid chromatography coupled with diode array detection. Z Naturforsch C 66(7–8):319–327
17. El-Houri R, Kotowska D, Christensen KB, Bhattacharya S, Oksbjerg N, Wolber G, Kristiansen K, Christensen LP (2015) Polyacetylenes from carrots (*Daucus carota)* improve glucose uptake in vitro in adipocytes and myotubes. Food Funct 6:2135–2144
18. Bruins AP (1994) Atmospheric-pressure-ionization mass spectrometry: II. Applications in pharmacy, biochemistry and general chemistry. Trends Anal Chem 13(2):81–90
19. Sasaki H, Taguchi H, Endo T, Yoshioka I (1982) The Constituents of Ledebouriella seseloides WOLFF. I. Structures of Three New Chromones. Chem Pharm Bull 30(10):3555–3562
20. Zheleva A, Soine TO, Bubeva-Ivanova L (1972) Natural Coumarins V: Isolation of Xanthalin and a New Pyranocoumarin, Peuarenine, from *Peucedanum arenarium*. W K J Pharm Sci 61(10):1643–1644
21. Tanaka T, Nakashima T, Ueda T, Tomii K, Kouno I (2007) Facile discrimination of aldose enantiomers by reversed-phase HPLC. Chem Pharm Bull 55(6):899–901
22. Wang YH, Avula B, Fu X, Wang M, Khan IA (2012) Simultaneous determination of the absolute configuration of twelve monosaccharide enantiomers from natural products in a single injection by a UPLC-UV/MS method. Planta Med 78:834–837