RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2024, **14**, 37911

Comprehensive characterization of anthraquinones in *Damnacanthus indicus* using mass spectrometry molecular networking and metabolomics-based herb discrimination†

Lihua Zeng,[ab] Xing Yan,[ab] Ya Xu,[ab] Lulu Zheng,[ab] Wenwen Deng,[ab] Mengning Li,*[ab] Hui Li*[abc] and Zhixin Wang *[ab]

Damnacanthus indicus is a widely used folk medicine in China, renowned for its various bioactivities. The key active components, anthraquinones, have not been comprehensively profiled due to their complex chemical nature. Establishing a high-throughput strategy to systematically characterize these anthraquinones is essential. Additionally, the cultivation of *D. indicus* across various provinces results in significant quality differences in the harvested herbs. Thus, developing an effective strategy to distinguish herbs from different regions and identify characteristic chemical markers for quality evaluation and control is crucial. In this study, a strategy based on ultra-high performance liquid chromatography-mass spectrometry (UHPLC-MS) was employed to systematically characterize the chemical composition of *D. indicus*. Mass spectrometry molecular networking was utilized to rapidly recognize and identify anthraquinones. Principal component analysis (PCA) was applied to cluster the herbs from different habitats, while partial least square discriminant analysis (PLS-DA) was used to screen for chemical markers distinguishing herb origins. The result showed that a total of 112 anthraquinones and 66 non-anthraquinone compounds were identified in *D. indicus*. The biosynthetic pathways of anthraquinones in this herb were proposed. PCA grouped 15 batches of herbs from different origins into three clusters, corresponding to the climate types of their habitats. PLS-DA identified 27 significant chemical markers that could robustly distinguish the geographical origins of the herbs. This study provides a valuable reference for the quality evaluation and control of *D. indicus* and offers a scientific basis for the pharmacological research and rational utilization of these medicinal resources.

Received 18th September 2024
Accepted 17th November 2024

DOI: 10.1039/d4ra06732k

rsc.li/rsc-advances

1. Introduction

Damnacanthus indicus C. F. Gaertn. (Rubiaceae) is an evergreen shrub that grows throughout eastern Asia. The whole plant has been used as a folk medicine in China since ancient times, and is known for its ability to relieve rheumatism, activate blood circulation, and alleviate pain.[1] Modern pharmacological research has validated its anti-inflammatory,[2] antioxidant,[3] and antibacterial[3] properties. To date, more than 90 compounds have been isolated and identified from this herb.[4–8] Among these, anthraquinones are particularly significant due to their diverse bioactivities, including anticancer,[9] antioxidant,[10] antibacterial,[11] antiviral,[12] antimalarial,[13] antidiabetic,[14] and neuroprotective[15] effects. However, due to the complex chemical composition of *D. indicus*, few investigations have comprehensively profiled its anthraquinone composition. Thus, establishing a high-throughput strategy to systematically characterize the anthraquinones, along with other chemical components in this plant is both meaningful and challenging.

Ultra-high performance liquid chromatography coupled with mass spectrometry (UHPLC-MS) is a sophisticated analytical method extensively utilized for the characterization of plant metabolites.[16] This technique provides exceptional sensitivity and selectivity and crucial structural insights, while circumventing the arduous and lengthy procedures associated with the isolation and purification of individual compounds. Molecular networking facilitated by mass spectrometry is an invaluable approach for the identification of compounds, as it effectively identifies and groups those compounds with analogous chemical architectures.[17] For the current investigation, a comprehensive strategy was devised, leveraging UHPLC/QTOF-MS and

[a]Jiangxi Province Key Laboratory of Traditional Chinese Medicine Pharmacology, Institute of Traditional Chinese Medicine Health Industry, China Academy of Chinese Medical Sciences, Nanchang 330115, China
[b]Jiangxi Health Industry Institute of Traditional Chinese Medicine, Nanchang 330115, China
[c]Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4ra06732k

molecular networking, to expedite the annotation process of anthraquinones present in *D. indicus*.

Moreover, *D. indicus* is widely cultivated in many provinces in China, including Hunan, Jiangxi, Zhejiang, Yunnan, Guangxi, Guangdong, and Guizhou.[1] The differing natural conditions, such as sunlight, temperature, humidity, and soil, across these regions result in significant quality variations in the harvested herbs. Untargeted metabolomics, capable of high-throughput analysis of multicomponent and multisample compositions, enables the identification of chemical components.[16] Consequently, a metabolomics-based strategy was employed to discriminate herbs from different habitats and identify characteristic chemical markers for quality evaluation and control.

As a result, 112 anthraquinones and 66 non-anthraquinone compounds were identified or tentatively identified in *D. indicus*. The identified anthraquinones were classified into three types based on their structures: rubiadin-type, emodin-type, and other type. The biosynthetic pathway of anthraquinones was proposed based on the identified compounds. Regarding herb discrimination, 15 batches of herbs were unambiguously clustered into three groups by their chemical composition, correlating well with the three climate types of their original habitats. Additionally, 27 compounds were identified as significant chemical markers for distinguishing the internal quality of *D. indicus*. To our knowledge, this is the first study to systematically characterize anthraquinones and analyze the chemical composition of *D. indicus*. The findings provide crucial data for its pharmacological research and serve as a valuable reference for developing quality control and evaluation methods for this herb.

2. Materials and methods

2.1. Materials and reagents

Fifteen batches of *D. indicus* samples with whole plant were collected from seven provinces of China (Table 1). All the herb samples were authenticated with morphological and histological methods by Prof. Yuan Yuan (Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences). Voucher were preserved at the authors' laboratory.

The chemical standards alizarin, lucidin, rubiadin and physcion with purity ≥98% was purchased from Shanghai Standard Technology Co., Ltd (Shanghai, China); 2-hydroxymethyl-anthraquinone, rubiadin 1-methyl ether, anthragallol, kaempferol and genistein-d_4 with purity ≥98% was from Chengdu Chroma-Biotechnology Co., Ltd (Chengdu, China). Acetonitrile and methanol (all MS grade) were bought from Fisher Scientific (Fair Lawn, NJ, USA), formic acid (MS grade) was from Merck (Rahway, NJ, USA), and dimethyl sulphoxide (DMSO, AR grade) was from Solarbio (Beijing, China). Ultra-pure water was prepared with a Millipore-Q water purification system (Bedford, USA).

2.2. Sample preparation

Following the initial steps of weighing and grinding, 100 mg of air-dried *D. indicus* herb powder were carefully weighed and placed into 2 mL Eppendorf tubes. To each tube, 20 μL of a genistein-d_4 solution, serving as the internal standard at a concentration of 500 μg mL^{-1} and 980 μL of methanol were added. The mixture was then subjected to vigorous vortexing for a duration of 10 min to ensure thorough mixing. Subsequently, the extraction process involved sonication for 30 min, followed by centrifugation at a speed of 10 000 rpm for 10 min at a temperature of 4 °C. After centrifugation, the supernatant, which is the upper, less dense liquid layer, was carefully collected and transferred into fresh 2 mL tubes. To ensure the integrity and clarity of the samples, they were filtered using a 0.22 μm nylon membrane syringe filter (Sterlitech, Kent, USA). For the establishment of a quality control measure, a quality control (QC) sample was created by pooling equal volumes of the extracted solutions from all samples, allowing for a consistent benchmark throughout the analysis.

Table 1 Habitat information of 15 batches of *Damnacanthus indicus* herb

Sample	Source	Regional division	Climate type	Collection time
HN1	Hunan	Central-east China	Subtropical humid monsoon climate	2023/11/25
HN2	Hunan	Central-east China	Subtropical humid monsoon climate	2023/11/30
HH3	Hunan	Central-east China	Subtropical humid monsoon climate	2023/11/28
HN4	Hunan	Central-east China	Subtropical humid monsoon climate	2023/11/20
JX1	Jiangxi	Central-east China	Subtropical humid monsoon climate	2023/12/20
JX2	Jiangxi	Central-east China	Subtropical humid monsoon climate	2023/11/06
JX3	Jiangxi	Central-east China	Subtropical humid monsoon climate	2023/12/30
ZJ1	Zhejiang	Central-east China	Subtropical humid monsoon climate	2023/10/11
ZJ2	Zhejiang	Central-east China	Subtropical humid monsoon climate	2023/10/26
YN1	Yunnan	Southwest China	Subtropical plateau monsoon climate	2023/09/17
YN2	Yunnan	Southwest China	Subtropical plateau monsoon climate	2023/09/01
GZ1	Guizhou	Southwest China	Subtropical plateau monsoon climate	2023/12/12
GD1	Guangdong	South China	Subtropical monsoon climate	2023/10/23
GX1	Guangxi	South China	Subtropical monsoon climate	2023/11/16
GX2	Guangxi	South China	Subtropical monsoon climate	2023/11/18

2.3. UHPLC/QTOF-MS analysis

The UHPLC separation was performed on a XionLC platform (AB Sciex, Framingham, USA), using an ACQUITY UPLC BEH C_{18} column (100 × 2.1 mm, 1.7 μm; Waters, Milford, MA, USA) at a temperature of 35 °C. The mobile phase comprised a 0.1% formic acid aqueous solution (phase A) and 0.1% formic acid in acetonitrile (phase B), with a gradient elution program: 15–30% B from 0 to 8 min, 30–70% B from 8 to 23 min, and 70–100% B from 23 to 30 min. The flow rate was maintained at 0.3 mL min^{-1}, and the injection volume was 2 μL.

The X500R QTOF mass spectrometer (AB Sciex) was interfaced with the UHPLC *via* an electrospray ionization (ESI) source. The optimized source and gas parameters included: nebulizing gas and auxiliary gas at 55 psi, curtain gas at 35 psi, and source temperature at 550 °C, with a CAD gas setting of 7. The experiment utilized the information-dependent acquisition (IDA) method in positive ion mode. MS parameters were set as follows: spray voltage, 5500 V; declustering potential (DP), 60 V; DP spread, 0 V; collision energy (CE), 10 V; CE spread, 0 V; accumulation time, 0.2 s; mass range, 100–1300 Da. The MS/MS experiment applied dynamic background subtraction and excluded former candidate ions for 6 s after 2 occurrences, with a maximum of 12 candidate ions. MS/MS parameters were: DP, 60 V; DP spread, 0 V; CE, 30 V; CE spread, 15 V; accumulation time, 0.05 s; mass range, 50–1300 Da. Each sample was analyzed in triplicate.

2.4. Data processing and component identification

For the processing of nontargeted metabolomics data acquired with UHPLC-MS, a combined workflow utilizing MS-DIAL[18] and MS-FINDER[19] (Tokyo University of Agriculture and Technology, Tokyo, Japan) was developed. Initially, the raw data were deconvoluted using MS-DIAL v5.2. Data collection parameters included an MS^1 tolerance of 0.01 Da and an MS^2 tolerance of 0.025 Da. Peak detection was applied with a minimum peak height amplitude of 1000 and a mass slice width of 0.1 Da. Deconvolution settings were: sigma window value of 0.5 and an MS^2 abundance cutoff of 5 amplitudes. The retention time tolerance was set at 0.1 min. The resulting .mgf files were then exported to MS-FINDER v3.61 for compound identification. The chemical formula for each peak was determined based on the accurate mass of designated protonated molecular ions or adduct ions using the built-in formula predictor and their corresponding isobaric molecular ions. The element composition was restricted to C, H, O, and N. MS^1 and MS^2 tolerances were set to 5 ppm and 10 ppm, respectively. Potential candidates for each compound were identified by consulting a self-built plant-derived anthraquinones database (383 compounds)[20–23] and the chemical components database (80 compounds) of *D. indicus*. The results were ranked by similarity scores, calculated by comparing the experimental MS^2 spectra with those predicted *in silico*. Top-ranked candidates were chosen based on experience; however, if key fragments could not be well explained or if the scores were identical, the most plausible structure among the candidates was manually selected. For peaks with constitutional isomers, retention times were assigned based on *C* log *P* values, with compounds having higher *C* log *P* values showing longer retention times on the reverse-phase column.

2.5. Mass spectrometry molecular networking

The molecular network for anthraquinones was meticulously assembled using the data processed by MS-DIAL and the online Feature-Based Molecular Networking (FBMN) tool available on the Global Natural Products Social Molecular Networking (GNPS) platform.[24,25] The MS^2 spectra were refined by applying a filter that retained only the six most abundant fragment ions. Precision was ensured by setting the precursor ion mass tolerance to 0.005 Da and the MS^2 fragment ion tolerance to 0.01 Da. In constructing the network, connections, or edges, were established only if they had a cosine score exceeding 0.6 and matched more than five peaks. Additionally, an edge was maintained between two nodes only if each node was among the top 30 most similar nodes to the other. To manage the complexity of the network, the maximum size of a molecular family was capped at 200. The resulting spectra within the network were then cross-referenced against the GNPS spectral libraries. From the networks generated, only three were conclusively identified as anthraquinone-related through multiple library-matched nodes. These three networks were subsequently exported, consolidated, and visualized using Cytoscape v3.10.2 (National Human Genome Research Institute, Bethesda, USA).

2.6. Statistical analysis

After data processing by MS-DIAL, a list of retention times with corresponding peak areas for all the detected peaks from each sample was exported for statistical analysis. Principal component analysis (PCA) and partial least square discriminant analysis (PLS-DA) were performed by SIMCA v17.0 (Sartorius Stedim Data Analytics AB, Umeå, Sweden) to illuminate the variances in chemical composition among herb samples from different habitats. All variables were standardized before multivariate statistical analysis. The distinct cluster separation in PCA score plots means they are differentiated. The method repeatability was evaluated by investigating the clustering degree of the QC samples of multiple replicates. To identify the components which were significantly contributed to the discrimination, the variables with variable importance in projection (VIP) greater than 1 (p-value < 0.05) were kept as the potential chemical markers. Univariate statistics were completed by one-way analysis of variance (ANOVA) built in SIMCA, and statistical differences were considered significant at a p-value below 0.05. TBtools v2.106 was used to acquire the heatmap to visualize the peak areas of chemical markers and the dendrograms of cluster analysis of herbs from different habitats.

3. Results

3.1. Mass spectrometry molecular network of anthraquinones

As shown in Fig. 1, the mass spectrometry molecular network was analyzed to illustrate the structural similarity and

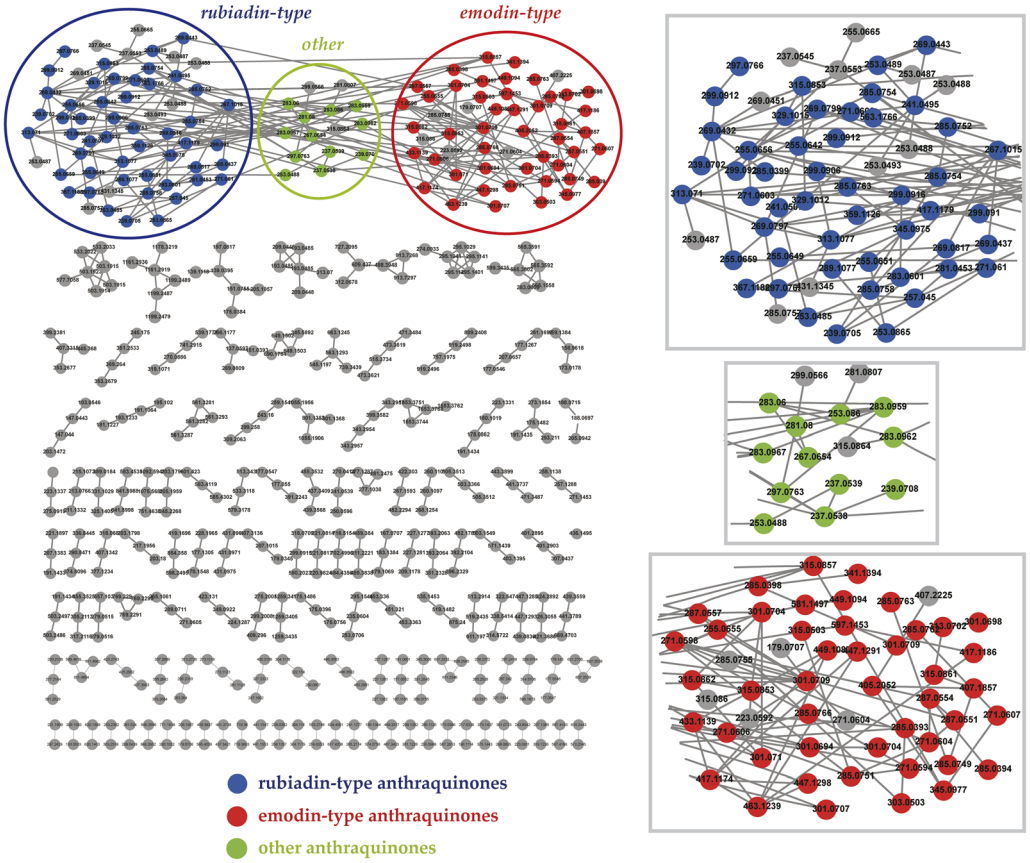


Fig. 1 Mass spectrometry molecular network of anthraquinones in *Damnacanthus indicus*. Each node represents a compound, and each connecting line between nodes denotes an edge.

derivation relationships of compounds in *D. indicus*. The molecular network incorporated a total of 20 760 nodes, forming 182 molecular clusters and 17 548 unconnected nodes. The recognized clusters could be categorized into three subnetworks of anthraquinones: rubiadin-type, emodin-type, and other type. Upon completion of the compound identification process, each of these nodes was visually enhanced by being highlighted in a variety of colors within the network. This color-coding approach serves to provide a clearer and more distinct representation of the different anthraquinone subtypes.

3.2 Identification of chemical components in *D. indicus*

The total ion current chromatogram (TIC) in positive mode with annotated peaks are shown in Fig. 2. Through the molecular network and high-resolution mass spectrometry (HR-MS)-based systematic identification strategy, a total of 112 anthraquinones were identified or tentatively identified. One of the representative rubiadin-type anthraquinones, lucidin, was selected as the reference compound to investigate the MS^2 fragmentation patterns of this type of compound (Fig. 3). In the positive ion mode, the $[M + H]^+$ quasi-molecular ion *m*/*z* 271.0611 of lucidin ($C_{15}H_{10}O_5$) could be readily formed. The cleavage of the middle ring of the parent ion led to the fragment ion *m*/*z* 105.0353, which could also undergo dehydration (losing H_2O) to produce the fragment ion *m*/*z* 253.0506. Continuous decarboxylation (losing CO) of the latter formed fragment ions *m*/*z* 225.0568, 197.0618, 169.0669, and 141.0718. Further dehydration of fragment *m*/*z* 225.0568 produced fragment *m*/*z* 207.0464. The homolytic cleavage of fragments *m*/*z* 197.0618 and 141.0718 led to the cleavage of methylene radicals, yielding fragments *m*/*z* 183.0452 and 127.0557, respectively. By elucidating the fragmentation pattern of lucidin, 50 other rubiadin-type anthraquinones were identified, with their chemical structures shown in Table 2 and related information listed in Table S1.†

Physcion was selected as the reference compound to investigate the MS^2 fragmentation patterns of emodin-type anthraquinones (Fig. 4). In the positive ion mode, the $[M + H]^+$ quasi-molecular ion *m*/*z* 285.0782 of physcion ($C_{16}H_{12}O_5$) was readily formed. Dehydration (losing H_2O) produced the fragment ion *m*/*z* 267.0680. Continuous decarboxylation (losing CO) of fragment *m*/*z* 267.0680 formed the fragment ions *m*/*z* 239.0722, 211.0772, 183.0837, and 139.0563. Furthermore, cleavage of the right ring and loss of a radical from the parent ion generated the fragment *m*/*z* 242.0593, and its subsequent decarboxylation and dehydration produced fragments *m*/*z* 214.0652 and 224.0503, resp. Further decarboxylation of the latter yielded the fragment *m*/*z* 196.0537. Based on the fragmentation pattern of physcion, a total of 50 other emodin-type anthraquinones were identified, with their chemical structures shown in Table 2 and related information listed in Table S1.†

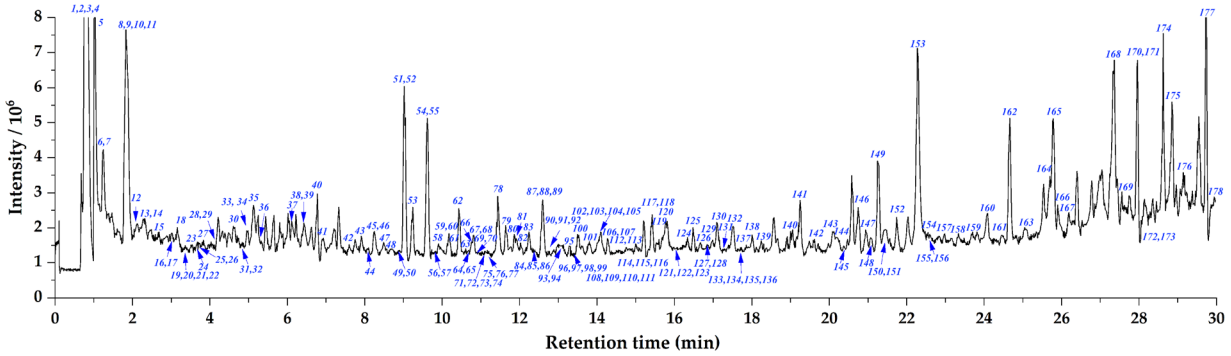


Fig. 2 Total ion current chromatogram (TIC) of *Damnacanthus indicus* extracts in positive mode with annotated peaks.

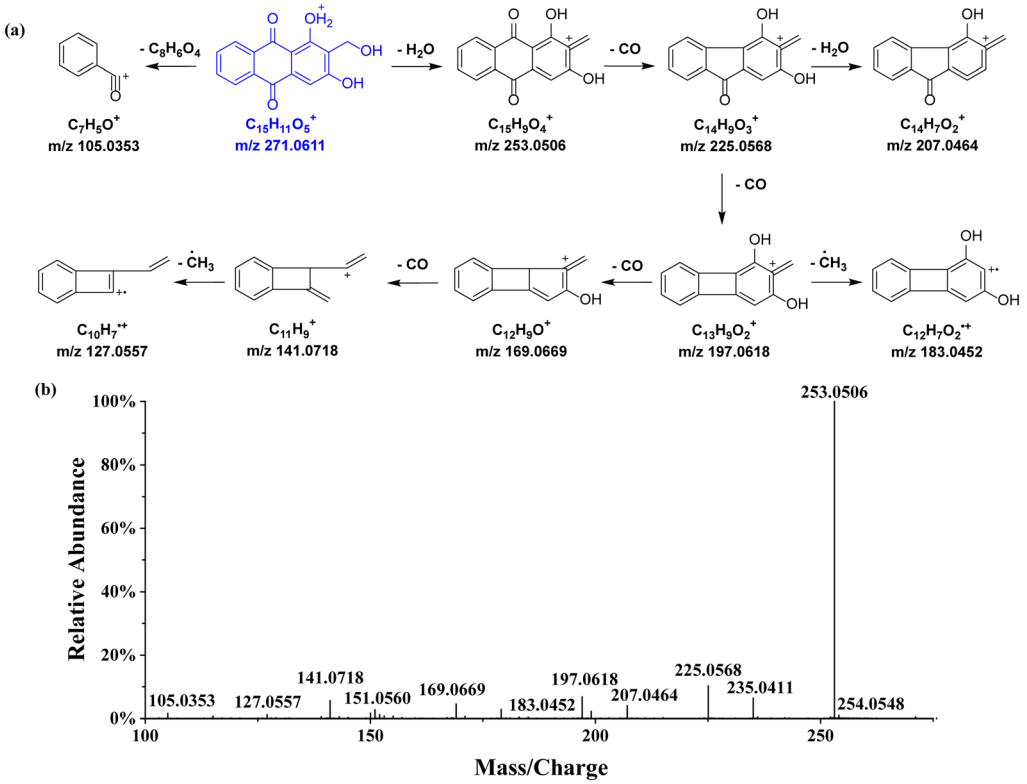


Fig. 3 Proposed fragmentation pattern in the positive-ion mode (a) and MS^2 spectrum (b) of lucidin. The $[M + H]^+$ quasi-molecular ion is shown in blue.

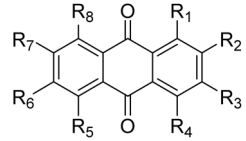
2-Hydroxymethyl-anthraquinone, an anthraquinone of other type, was selected to investigate the MS^2 fragmentation patterns (Fig. 5). The $[M + H]^+$ quasi-molecular ion *m/z* 239.0721 ($C_{15}H_{10}O_3$) was readily formed. Dehydration (losing H_2O) yielded the fragment ion *m/z* 221.0618, which was further decarboxylated to generate the fragment *m/z* 193.0666. The natural loss of one molecule of formaldehyde (HCHO) led to the production of the fragment ion *m/z* 209.0616, and further cleavage of the middle ring generated the same fragment *m/z* 153.0718. Similarly, 12 additional anthraquinones of other type were identified, with their chemical structures shown in Table 2 and related information listed in Table S1.†

In addition to anthraquinones, 66 non-anthraquinone compounds were also identified or tentatively identified. These include 12 phenylpropanoids, 9 triterpenoids, 9 terpenoids, 5 sesquiterpenoids, 4 alkaloids, 5 fatty acids, 5 organic acids, 3 steroids, 3 monoterpenoids, 2 flavonoids, 2 tetraterpenoids, 3 amino acids, and 4 esters. Their related information is listed in Table S1.†

3.3 Proposed biosynthetic pathway of anthraquinones in *D. indicus*

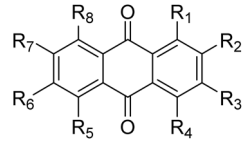
In plants, anthraquinones are primarily synthesized through two distinct pathways: the shikimic acid (SA) pathway and the

Table 2 Chemical structures of all the identified anthraquinones from *Damnacanthus indicus*



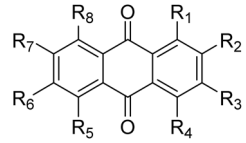
Type	Name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	References
Emodin-type	1-*O*-Glucoside-2,6,8-trihydroxyl-3-methyl-anthraquinone	O-glc[b]	OH	CH_3	H	H	OH	H	OH	26
	Chrysophanein	O-glc	H	CH_3	H	H	H	H	OH	27
	Physcion-1-*O*-glucoside	O-glc	H	CH_3	H	H	OCH_3	H	OH	28
	Aloe-emodin-7-hydroxyl-8-*O*-glucoside	OH	H	CH_2OH	H	H	H	OH	O-glc	23
	3-*O*-Glucoside-1,8-dihydroxyl-2-methoxymethyl-anthraquinone	OH	CH_2OCH_3	O-glc	H	H	H	H	OH	23
	Emodin-1-*O*-glucoside	O-glc	H	CH_3	H	H	OH	H	OH	29
	1-*O*-Primeveroside-3,6,8-trihydroxyl-2-methyl-anthraquinone	O-prim	CH_3	OH	H	H	OH	H	OH	23
	Glucoobtusifolin	OCH_3	O-glc	CH_3	H	H	H	H	OH	30
	3-*O*-Primeveroside-1,6,8-trihydroxyl-2-methyl-anthraquinone	OH	CH_3	O-prim	H	H	OH	H	OH	23
	Obtusin	OH	OCH_3	OCH_3	H	H	CH_3	OH	OCH_3	31
	1,5-Dihydroxy-2-methoxy-9,10-anthracenedione[a]	OH	OCH_3	H	H	OH	H	H	H	32
	Physcion-8-*O*-rutinoside	OH	H	CH_3	H	H	OCH_3	H	O-rut	33
	Franguloside	OH	H	CH_3	H	H	O-rha	H	OH	34
	Lucidin ω-methyl ether[a]	OH	CH_2OCH_3	OH	H	H	H	H	H	35
	Xanthorin	OH	H	CH_3	H	OH	OCH_3	H	OH	36
	1,7-Dihydroxyl-2-hydroxymethyl-anthraquinone	OH	CH_2OH	H	H	H	H	OH	H	23
	Madagascin	OH	H	CH_3	H	H	C_5H_9O	H	OH	37
	Rubiadin-6-hydroxyl	OH	CH_3	OH	H	H	OH	H	H	23
	1,6-Dihydroxyl-2,5-dimethoxyl-anthraquinone	OH	OCH_3	H	H	OCH_3	OH	H	H	23
	Lupinacidin A	OH	CH_3	OH	C_5H_{11}	OH	H	H	H	23
	Ophiohayatone A	H	CH_2OCH_3	OH	H	H	OH	H	H	38
	3,6-Dihydroxyl-1-methoxyl-2-hydroxymethyl-anthraquinone	OCH_3	CH_2OH	OH	H	H	OH	H	H	23
	Endocrocin	OH	COOH	CH_3	H	H	OH	H	OH	39
	1,3,6-Trihydroxy-2-methoxy-9,10-anthracenedione	OH	OCH_3	OH	H	H	OH	H	H	40
	Morindone 5-methyl ether	OH	CH_3	H	H	OCH_3	OH	H	H	41
	Emodin-6-geranyl	CH_3	$COOCH_3$	OH	H	H	H	H	OH	23
	Physcion-2-acety	OH	$COCH_3$	CH_3	H	H	OCH_3	H	OH	23
	1,6,7-Trihydroxyl-3-methoxyl-8-methoxymethyl-anthraquinone	OH	H	CH_3	H	H	OH	OH	CH_2OCH_3	23
	Norjuzunal[a]	OH	CHO	OH	H	OH	H	H	H	35
	1,7-Dihydroxyl-3,6-dimethoxyl-anthraquinone	OH	H	OCH_3	H	H	OCH_3	OH	H	23
	Aloe-emodin-2-hydroxyl	OH	OH	CH_2OH	H	H	H	H	OH	23
	Robustaquinone F	OH	OCH_3	OH	H	OH	H	H	OH	42
	Alatonal	OH	CHO	OH	H	H	H	H	OH	23
	1,3,5-Trihydroxy-2-(methoxymethyl)-9,10-anthracenedione[a]	OH	CH_2OCH_3	OH	H	OH	H	H	H	7
	2,7,8-Trihydroxyl-1-methoxyl-3-methyl-anthraquinone	OCH_3	OH	CH_3	H	H	H	OH	OH	23
	Obtusifolin-6-hydroxyl	H	OH	H	OH	OCH_3	OH	CH_3	H	43
	2,5-Dihydroxy-1-methoxy-9,10-anthracenedione	OCH_3	OH	H	H	OH	H	H	H	23
	1,6-Dihydroxy-2-methoxy-9,10-anthracenedione[a]	OH	OCH_3	H	H	H	OH	H	H	44
	1,5,8-Trihydroxyl-2-methoxyl-anthraquinone	OH	OCH_3	H	H	OH	H	H	OH	23

Table 2 (*Contd.*)



Type	Name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	References
	Morindone	OH	OH	H	H	OH	CH_3	H	H	7
	1,2,8-Trihydroxyl-3-methyl-6-hydroxymethyl-anthraquinone	OH	OH	CH_3	H	H	CH_2OH	H	OH	23
	Isorhodoptilometrin	OH	H	C_3H_7O	H	H	OH	H	OH	23
	1,5-Dihydroxyl-6-methoxyl-2-methyl-anthraquinone	OH	CH_3	H	H	OH	OCH_3	H	H	23
	Soranjidiol	OH	CH_3	H	H	H	OH	H	H	41
	1,5-Dihydroxyl-3-carboxy-anthraquinone	OH	H	COOH	H	OH	H	H	H	23
	Rhodoptilometrin	OH	H	OH	H	H	C_3H_7O	H	H	23
	Digitolutein-5-hydroxyl	OCH_3	OH	CH_3	H	OH	H	H	H	23
	Physcion[a]	OH	H	CH_3	H	H	OCH_3	H	OH	36
	Aurantio-obtusin-1-hydroxyl	OH	H	$C_{10}H_{17}$	H	H	CH_3	H	OH	23
	Emodin-6-geranyloxy	OH	H	CH_3	H	H	$C_{11}H_{19}$	H	OH	23
Rubiadin-type	Rubiadin-1-methoxyl-3-*O*-primeveroside	OCH_3	CH_3	O-prim	H	H	H	H	H	23
	8-Hydroxyl-2-methyl-3-carbomethoxy-anthraquinone	H	CH_3	$COOCH_3$	H	H	H	H	OH	23
	Rubiadin-3-*O*-glucoside	OH	CH_3	O-glc	H	H	H	H	H	23
	Damnacanthal[a]	OCH_3	CHO	OH	H	H	H	H	H	7
	Xanthopurpurin	OH	H	OH	H	H	H	H	H	45
	1,3-Dihydroxyl-2-methoxymethyl-anthraquinone	OH	CH_2OCH_3	OH	H	H	H	H	H	23
	Munjistin	OH	COOH	OH	H	H	H	H	H	35
	6-Hydroxyl-7-methoxyl-2-methoxymethyl-anthraquinone	H	CH_2OCH_3	H	H	H	OH	OCH_3	H	23
	Damnacanthol[a]	OCH_3	CH_2OH	OH	H	H	H	H	H	7
	1,3-Dihydroxyl-5,6-dimethoxyl-2-methyl-anthraquinone	OH	CH_3	OH	H	OCH_3	OCH_3	H	H	23
	2,3-Dihydroxyl-1-methoxyl-anthraquinone	H	OH	OH	OCH_3	H	H	H	H	23
	Alizarin 2-methyl ether[a]	OH	OCH_3	H	H	H	H	H	H	44
	Nordamnacanthal[a]	OH	CHO	OH	H	H	H	H	H	35
	Anthragallol	OH	OH	OH	H	H	H	H	H	7
	3-Hydroxyl-1-methyl-2-carbomethoxy-anthraquinone	CH_3	$COOCH_3$	OH	H	H	H	H	H	23
	3-Hydroxyl-2,4-dimethoxyl-1-methyl-anthraquinone	CH_3	OCH_3	OH	OCH_3	H	H	H	H	23
	1-Hydroxyl-5,6-dimethoxyl-2-methyl-anthraquinone	OH	CH_3	H	H	OCH_3	OCH_3	H	H	23
	Ethyl 9,10-dihydro-1,3-dihydroxy-9,10-dioxo-2-anthracenecarboxylate[a]	OH	$C_3H_5O_2$	OH	H	H	H	H	H	23
	Lucidin	OH	CH_2OH	OH	H	H	H	H	H	35
	Alizarin 1-methyl ether[a]	OCH_3	OH	H	H	H	H	H	H	43
	1-Hydroxyl-6,7-dimethoxyl-2-methyl-anthraquinone	OH	H	CH_3	H	H	OCH_3	OCH_3	H	23
	1-Hydroxyl-3,7-dialdehyde-anthraquinone	OH	H	CHO	H	H	H	CHO	H	23
	2-Hydroxyl-1,3-dimethoxyl-anthraquinone	H	OH	OCH_3	OCH_3	H	H	H	H	23
	Alizarin	OH	OH	H	H	H	H	H	H	35
	Fistulaquinones C	H	OH	OH	H	OCH_3	CH_2OCH_3	H	H	23
	Digiferruginol[a]	OH	CH_2OH	H	H	H	H	H	H	43
	1-Hydroxyl-8-carboxy-anthraquinone	OH	H	H	H	H	H	H	COOH	23

Table 2 (*Contd.*)



Type	Name	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	References
	1-Hydroxyl-2-carbaldehyde-anthraquinone	OH	CHO	H	H	H	H	H	H	23
	Damnacanthol-2-methoxymethyl	OCH_3	CH_2OCH_3	OH	H	H	H	H	H	23
	Isorhodoptilometrin-1-methoxyl	OCH_3	H	C_3H_7O	H	H	OH	H	OH	23
	Rubiadin 1-methyl ether[a]	OCH_3	CH_3	OH	H	H	H	H	H	44
	Emodin-1,6-dimethoxyl	OCH_3	H	CH_3	H	H	OCH_3	H	OH	23
	3-Hydroxyl-2-aldehyde-anthraquinone	H	CHO	OH	H	H	H	H	H	23
	1-Hydroxyl-2-ethoxy-anthraquinone	OH	OCH_2CH_3	H	H	H	H	H	H	23
	1-Hydroxy-2-methylanthraquinone[a]	OH	CH_3	H	H	H	H	H	H	44
	3-Hydroxyl-1,2-dimethoxyl-anthraquinone	H	OCH_3	OH	OCH_3	H	H	H	H	23
	1,2-Dihydroxy-6-methyl-9,10-anthracenedione	OH	OH	H	H	H	CH_3	H	H	46
	1,3-Dihydroxyl-6-methoxyl-2-ethoxymethyl-anthraquinone	OH	C_3H_7O	OH	H	H	OCH_3	H	H	23
	1,3-Dimethoxyl-2-hydroxymethyl-anthraquinone	OCH_3	CH_2OH	OCH_3	H	H	H	H	H	23
	7-Hydroxyanthra[1,2-*d*]-1,3-dioxole-6,11-dione[a]	OCH_2O		H	H	OH	H	H	H	7
	3-Hydroxyl-1,2-dimethyl-anthraquinone	CH_3	CH_3	OH	H	H	H	H	H	23
	Rubiadin[a]	OH	CH_3	OH	H	H	H	H	H	38
	Siameaquinone A	OH	H	OCH_3	H	H	CH_3	$C_5H_8O_2$		23
	3-Hydroxyl-1-methoxyl-2-carboxy-anthraquinone	OH	C_3H_7O	OCH_3	H	H	H	H	H	23
	6-Hydroxyl-2-methoxymethyl-anthraquinone	H	CH_2OCH_3	H	H	H	OH	H	H	23
	1,3-Hihydroxyl-5-methoxyl-2,6-dismethoxymethyl-anthraquinone	OH	CH_2OCH_3	OH	OCH_3	CH_2OCH_3	H	H	H	23
	2-Hydroxyl-3-methyl-anthraquinone	H	OH	CH_3	H	H	H	H	H	23
	2-Hydroxyl-3-methoxyl-7-hydroxymethyl-anthraquinone	H	OH	OCH_3	H	H	H	CH_2OH	H	23
	1-Hydroxyl-8-methoxyl-2-methyl-anthraquinone	OH	CH_3	H	H	H	H	H	CH_2CH_3	23
	Obtusin-6-hydroxyl-2-methoxyl	OCH_3	OCH_3	CH_3	H	H	OH	OCH_3	OH	23
Other type	2-Hydroxymethyl-anthraquinone	H	CH_2OH	H	H	H	H	H	H	47
	2-Acetoxymethyl-anthraquinone	H	$C_3H_5O_2$	H	H	H	H	H	H	23
	1-Formylanthraquinone	CHO	H	H	H	H	H	H	H	23
	1,5-Dimethoxyl-2-methyl-anthraquinone	OCH_3	H	H	H	OCH_3	CH_3	H	H	23
	2-(Methoxycarbonyl)anthraquinone	H	$COOCH_3$	H	H	H	H	H	H	48
	2-Formyl-1,3-dimethoxyanthraquinone	OCH_3	CHO	OCH_3	H	H	H	H	H	49
	2-Dimethoxymethyl-anthraquinone	H	$C_3H_7O_2$	H	H	H	H	H	H	23
	Morindaparvin A	OCH_2O		H	H	H	H	H	H	43
	2-Formylanthraquinone	H	CHO	H	H	H	H	H	H	47
	7-Methoxyanthra[1,2-*d*]-1,3-dioxole-6,11-dione[a]	OCH_2O		H	H	OCH_3	H	H	H	7
	Chrysophanol-1,8-dimethoxyl	OCH_3	H	CH_3	H	H	H	H	OCH_3	45
	1-Methoxyl-2-methyl-anthraquinone	OCH_3	CH_3	H	H	H	H	H	H	23

[a] Previously reported anthraquinones in *Damnacanthus indicus*. [b] glc = glucose ($C_6H_{11}O_5$), prim = primeverose ($C_{11}H_{19}O_9$), rut = rutinose ($C_{12}H_{21}O_9$), rha = rhamnose ($C_6H_{11}O_5$).

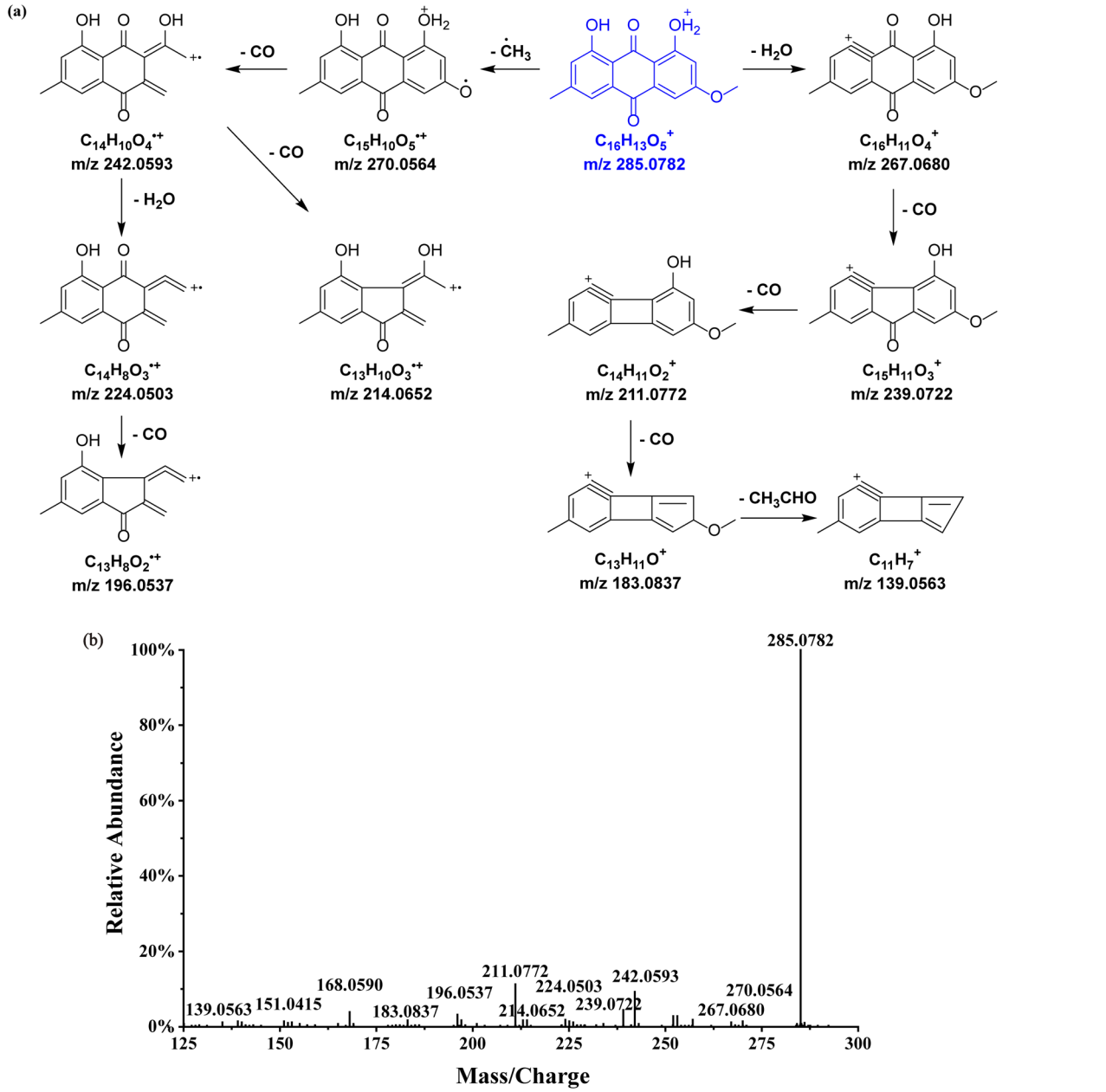


Fig. 4 Proposed fragmentation pattern in the positive-ion mode (a) and MS^2 spectrum (b) of physcion. The $[M + H]^+$ quasi-molecular ion is shown in blue.

polyketone pathway. The SA pathway is responsible for the production of rubiadin-type anthraquinones, whereas the polyketone pathway generates emodin-type anthraquinones.[23] Fig. 6 illustrates the proposed biosynthetic pathway of some typical dihydrochalcones identified in *D. indicus*.

The SA pathway involves several metabolic modules, including the tricarboxylic acid (TCA) cycle, the mevalonate (MVA) pathway, and the methyl erythritol phosphate (MEP) pathway. The first module of this pathway is the synthesis of 1,4-dihydroxy-2-naphthoic acid (DHNA). Starting with phosphoenol pyruvate (PEP) and erythrose-4-phosphate (E4P) as substrates, a series of enzyme-catalyzed reactions produce isochorismic acid (IA). IA, along with α-ketoglutarate from the TCA cycle, then participates in a three-step process to form DHNA, which constructs the A and B rings of the anthraquinone nucleus. The second module generates 3,3-dimethylallyl diphosphate (DMAPP). DMAPP can be synthesized from isoprenyl diphosphate (IPP) *via* the MVA pathway or from 4-hydroxy-3-methylbutenyl-1-diphosphate (HMBPP) *via* the MEP pathway, both involving multiple steps. DMAPP, a crucial precursor in anthraquinone biosynthesis, is primarily produced through the MVA pathway in the cytoplasm and the MEP pathway in the plastid. The complete anthraquinone three-ring structure is synthesized from DHNA and DMAPP through additional reactions. Subsequently, various enzymes catalyze the production of rubiadin-type anthraquinones. This includes alizarin,

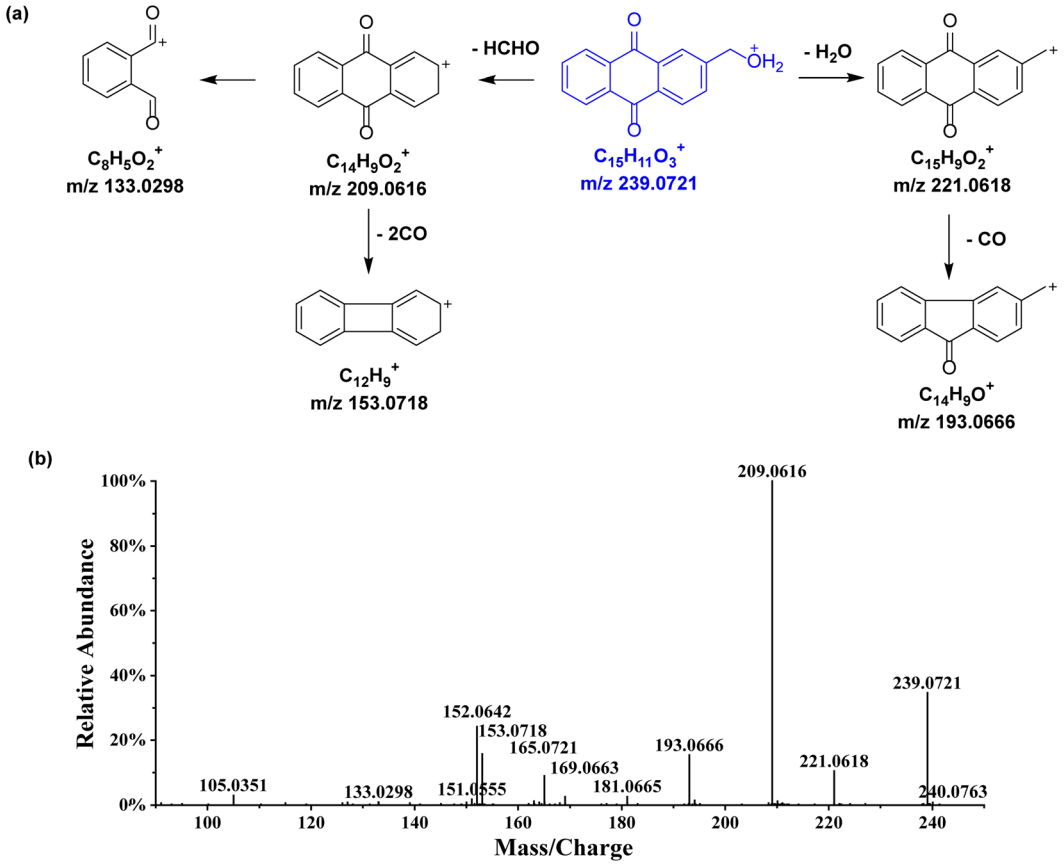


Fig. 5 Proposed fragmentation pattern in the positive-ion mode (a) and MS^2 spectrum (b) of 2-hydroxymethyl-anthraquinone. The $[M + H]^+$ quasi-molecular ion is shown in blue.

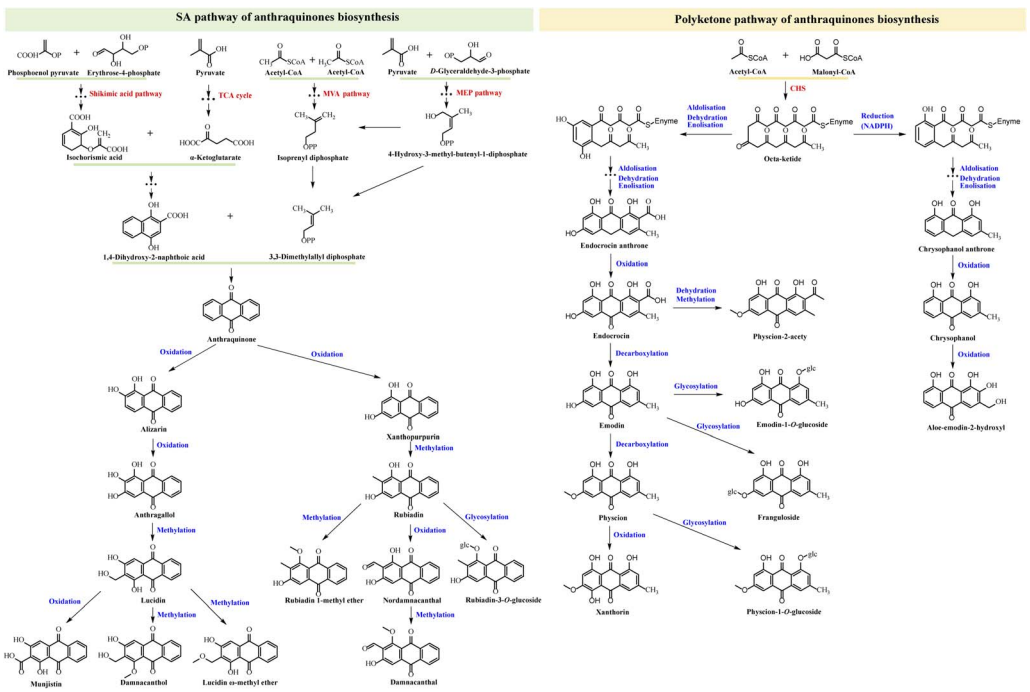


Fig. 6 Proposed biosynthetic pathway of several typical anthraquinones identified from *Damnacanthus indicus*. SA: shikimic acid; MVA: mevalonate; MEP: methyl erythritol phosphate; DHNA: 1,4-dihydroxy-2-naphthoic acid; PEP: phosphoenol pyruvate; E4P: erythrose-4.

anthragallol, and xanthopurpurin *via* oxidation; lucidin, damnacanthol, lucidin ω-methyl ether, rubiadin, and rubiadin 1-methyl ether *via* methylation; munjistin and nordamnacanthal through further oxidation; damnacanthal *via* additional methylation; and rubiadin-3-*O*-glucoside through glycosylation.

In the polyketone pathway, acetyl-CoA and malonyl-CoA serve as substrates for a series of condensation reactions mediated by chalcone synthase, resulting in the elongation of the carbon chain and ultimately forming an octa-ketide compound. This octa-ketide undergoes aldolization, dehydration, enolization, and/or reduction to yield intermediates, followed by multi-step reactions to produce endocrocin anthrone and chrysophanol anthrone. Subsequent oxidation leads to the formation of endocrocin, chrysophanol, and aloe-emodin-2-hydroxyl, while physcion-2-acety is produced through dehydration and methylation. Decarboxylation yields emodin and physcion, further oxidation produces xanthorin, and glycosylation results in emodin-1-*O*-glucoside, franguloside, and physcion-1-*O*-glucoside.

3.4 Herb discrimination and chemical marker discovery

The multivariate statistical analysis, based on the characterized chemical components, was utilized to elucidate the differences between the *D. indicus* herb samples. As illustrated in the PCA score plot (Fig. 7a), the three replicates of the QC samples clustered closely together near the origin, indicating the high repeatability of the analytical method and the reliability of the acquired data. The points representing the seven herb samples were distinctly grouped into three clusters. Notably, the clustering results corresponded well with the geographic regions of the herb origins, including Hunan–Jiangxi–Zhejiang, Guangdong–Guangxi, and Yunnan–Guizhou.

The PLS-DA score plot (Fig. 7b) further clarified the clustering of the 15 herb samples into three distinct groups according to their habitats, showing greater separation between the groups compared to the PCA results. The validity of the PLS-DA model was confirmed by a cross-validation with 200 permutations ($R^2 < 0.33$, $Q^2 < -0.537$) (Fig. 7c). Key chemical components contributing to the differentiation of the herb samples were identified based on their significant VIP values and substantial between-group variance (VIP > 1 and p-value < 0.05). These 27 robust components include 1-*O*-glucoside-2,6,8-trihydroxyl-3-methyl-anthraquinone (4), xanthopurpurin (37), norjuzunal (40), franguloside (45), 3,6-dihydroxyl-1-methoxyl-2-hydroxymethyl-anthraquinone (47), ophiohayatone A (51), digiferruginol (58), 1,6-dihydroxyl-2,5-dimethoxyl-anthraquinone (61), 1,5-dihydroxyl-3-carboxy-anthraquinone (79), 1-formylanthraquinone (86), 2,7,8-trihydroxyl-1-methoxyl-3-methyl-anthraquinone, obtusifolin-6-hydroxyl (95), rubiadin 1-methyl ether (97), 1-hydroxyl-2-carbaldehyde-anthraquinone (102), 1,5-dihydroxy-2-methoxy-9,10-anthracenedione (105), munjistin (119), 1,5-dimethoxyl-2-methyl-anthraquinone (123), 1,3-dihydroxyl-5,6-dimethoxyl-2-methyl-anthraquinone (126), rubiadin (128), 7-methoxyanthra[1,2-*d*]-1,3-dioxole-6,11-dione (138), 1-hydroxy-8-methoxyl-2-methyl-anthraquinone (151), deacetylasperuloside (3), *p*-coumaric acid (17), 4-oxo-β-ionone (21), methyl ferulate (34), 13-HOTE (152), and botulin (159). The majority of these compounds, specifically the first 21, are

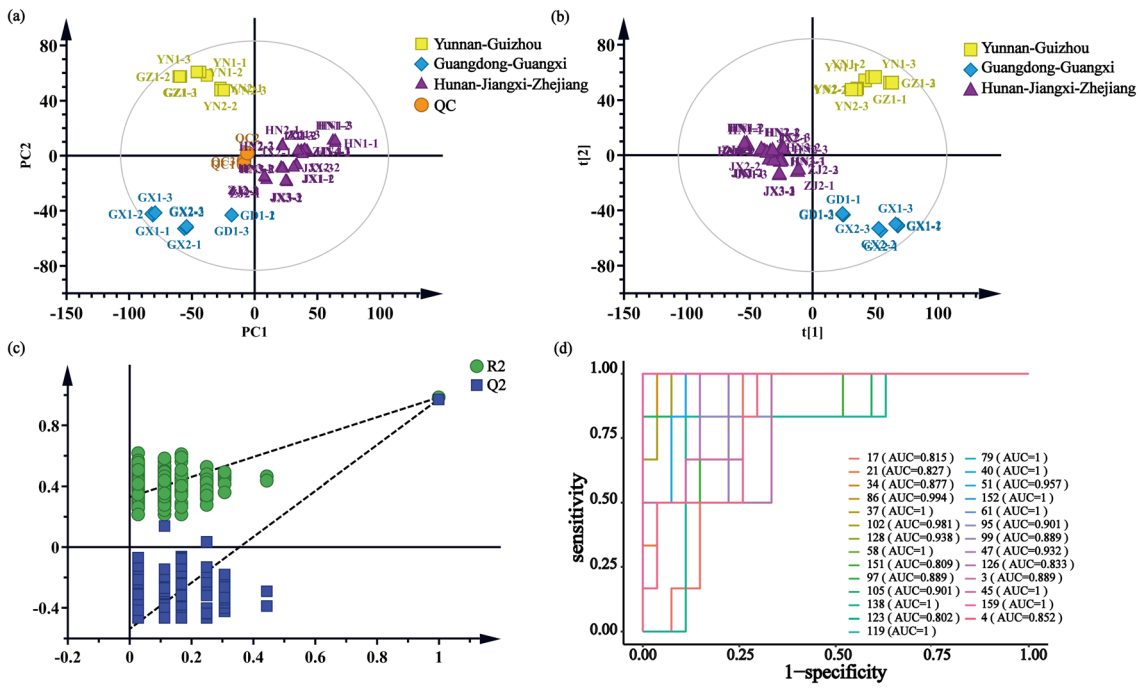


Fig. 7 Herb discrimination of *Damnacanthus indicus* from different habitats based on characterized chemical components and screening for chemical markers. PCA (a) and PLS-DA (b) score plots of herbs collected from 15 different habitats; (c) permutation test for PLS-DA model; (d) receiver operating characteristic (ROC) curves and area under the curves (AUCs) of chemical markers for validating effectiveness in distinguishing herbs from various habitats.

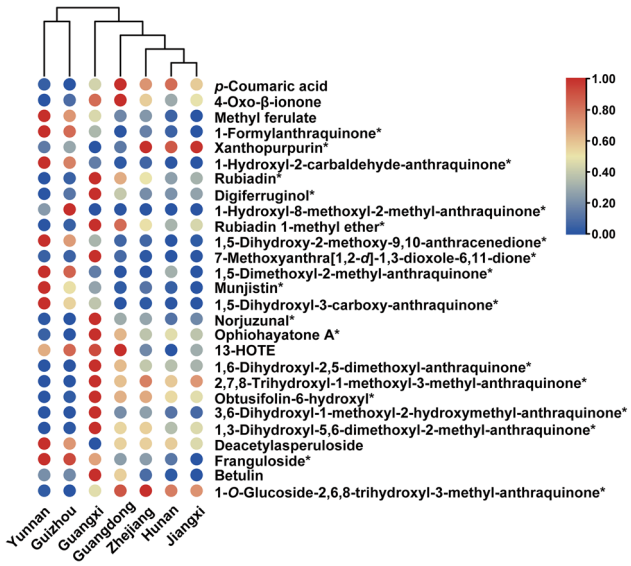


Fig. 8 Heatmaps visualizing comparison of contents of significant chemical markers with cluster dendrograms for *Damnacanthus indicus* from different habitats. Asterisks denote anthraquinones.

anthraquinones. The effectiveness of them in distinguishing samples from various habitats was further corroborated by receiver operating characteristic (ROC) curve analysis, with all compounds displaying an area under the curve (AUC) greater than 0.8 (Fig. 7d).

3.5 Content variance of chemical markers identified in *D. indicus* for herb discrimination

To further quantitatively evaluate the contribution of each potential chemical marker, a heatmap were generated based on the peak areas of each compound (Fig. 8). These metric colors revealed distinct differences in the content of these compounds among herbs derived from different habitats. Additionally, the hierarchical dendrograms from cluster analysis further corroborated the results of PCA and PLS-DA. Consequently, the above-mentioned 27 compounds could be regarded as the most critical chemical markers for discriminating the internal quality of *D. indicus*.

4 Discussion

Quality control of herbs is essential for ensuring their clinical efficacy and forms the bedrock of modern traditional Chinese medicine (TCM) research. This domain encompasses a multi-dimensional, multi-faceted, rigorous, and systematic analytical framework.[50,51] From the collection, processing, and storage of herbs to the precise identification and quantitative analysis of active ingredients, quality control guarantees the safety, efficacy, and stability of TCM.[52]

D. indicus, a historically significant and widely used folk medicine, requires meticulous quality control research. This plant comprises a variety of bioactive compounds, with anthraquinones being of particular interest due to their notable pharmacological properties. This class of anthraquinone compounds exhibits anti-inflammatory, antioxidant, anti-tumor, antiviral and other potential. Therefore, comprehensive chemical constituent analysis, particularly the qualitative and quantitative assessments of anthraquinones in *D. indicus*, is crucial for evaluating its efficacy and safety. As a modern analytical tool, UHPLC-MS, known for its high resolution, sensitivity, and rapid analysis capabilities, can provide robust tools for analyzing anthraquinones and accurately profile and identify various anthraquinones in *D. indicus*. Mass spectrometry molecular networking (MSMN), an advanced technique, enhances the efficiency of recognizing and classifying anthraquinones by clustering structurally similar compounds. Compared to traditional chemical analysis methods, MSMN offers superior data processing efficiency, enabling rapid identification numerous of compounds with a specific structural type.[53] This technique provides new perspectives and methodologies for studying TCM chemical constituents. By constructing a molecular network, in-depth comparative analyses of anthraquinones in *D. indicus* can be conducted, revealing their unique chemical characteristics and pharmacological properties.

Geographical origin discrimination, another critical aspect of quality control for medicinal materials, significantly impacts the quality and efficacy of TCM.[54] The climatic, soil, and ecological variations of different regions lead to differences in chemical constituents, influencing their pharmacological effects. Thus, geographical origin identification ensures the traceability of *D. indicus*, maintaining consistency and stability in its quality. Additionally, this identification aids in understanding the biodiversity of *D. indicus*, providing valuable scientific guidance for cultivation and resource conservation.

The investigation of the biosynthetic pathways of anthraquinones, the primary pharmacological components of *D. indicus*, is crucial for obtaining specific compounds and understanding the pharmacological mechanisms of the medicinal material. Research on biosynthetic pathways can elucidate the formation mechanisms of anthraquinones, providing theoretical support for optimizing cultivation and processing techniques. Additionally, this research may lead to the discovery of new active ingredients, expanding the application scope and clinical prospects of *D. indicus*.

This study is mainly focused on the characterization of chemical constituents, biosynthetic pathway deduction of anthraquinones, geographical origin discrimination and screening of chemical quality markers to establish a quality control system for *D. indicus*. However, due to the lack of preliminary chemical research, no reference substance was obtained through isolation and purification for active component evaluation. Consequently, the quality markers could not be determined based on activity data, and the gap would be addressed in subsequent research.

5 Conclusions

In this study, a UHPLC-MS-based strategy was utilized for the first time to comprehensively characterize the chemical composition of *D. indicus*. Through MSMN, 112 anthraquinones

and 66 non-anthraquinone compounds were identified, and their biosynthetic pathways in this herb were proposed. PCA grouped 15 batches of herbs from different origins into three clusters, corresponding with the climate types of their habitats. Additionally, PLS-DA identified 27 significant chemical markers that could robustly distinguish the geographical origins of the herbs. This study provides a valuable reference for the quality evaluation and control of *D. indicus* and offers a scientific basis for its pharmacological research and rational utilization of these medicinal resources.

Data availability

Data are contained within the article and ESI.†

Author contributions

Conceptualization, Hui Li and Zhixin Wang; funding acquisition, Lihua Zeng and Zhixin Wang; investigation, Ya Xu; methodology, Xing Yan and Wenwen Deng; project administration, Lihua Zeng and Zhixin Wang; resources, Lulu Zheng; validation, Lihua Zeng and Xing Yan; writing – original draft, Zhixin Wang; writing – review & editing, Lihua Zeng, Mengning Li, Hui Li and Zhixin Wang. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was supported by The General Project of Jiangxi Provincial Administration of Traditional Chinese Medicine (2023B1328), The Fundamental Research Funds for the Central Public Welfare Research Institutes (ZZ17-ND-12; KY-QD20230016) and National Natural Science Foundation of China (82104387; 82460854).

References

1 Editorial board of The Chinese Materia Medica of National Administration of Traditional Chinese Medicine, *The Chinese Materia Medica*, Shanghai Science and Technology Press, Shanghai, China, 1st edn, 1999.
2 Z. Y. Wu, Z. L. Ma, C. Q. Xie, J. J. Dong and K. H. Zhu, *Chin. J. Tradit. Med. Sci. Technol.*, 2019, **26**, 36–39.
3 R. L. Ma, Q. Q. Li, X. Q. Xu, M. Li, J. Tang and Z. Ouyang, *Chinese Wild Plant Res.*, 2014, **33**, 16–19.
4 Y. J. Yang, H. Y. Shu and Z. D. Min, *Yaoxue Xuebao*, 1992, **27**, 358–364.
5 S. Long, C. Yuan, Y. Wang, J. Zhang and G. Li, *J. Evidence-Based Complementary Altern. Med.*, 2019, **2019**, 1–9.
6 Y.-F. Cai and Q.-S. Huang, *Zhongyaocai*, 2012, **35**, 694–696.
7 K. Junko, O. Teruyo, T. Kiyoshi, K. Isao and I. Hiroshi, *Phytochemistry*, 1992, **31**, 709–710.
8 S.-W. Lee, S.-C. Kuo, Z.-T. Chen and Z.-S. Liu, *J. Nat. Prod.*, 1994, **57**, 1313–1315.
9 T. Nualsanit, P. Rojanapanthu, W. Gritsanapan, S.-H. Lee, D. Lawson and S. J. Baek, *J. Nutr. Biochem.*, 2012, **23**, 915–923.
10 G. C. Yen, P. D. Duh and D. Y. Chuang, *Food Chem.*, 2000, **70**, 437–441.
11 T. Qun, T. T. Zhou, J. K. Hao, C. M. Wang, K. Y. Zhang, J. Xu, *et al.*, *RSC Med. Chem.*, 2023, **14**, 1446–1471.
12 P. A. Cohen, J. B. Hudson and G. H. N. Towers, *Experientia*, 1996, **52**, 180–183.
13 L. A. Abu el Heiga, J. Katzhendler, K. F. Gean and U. Bachrach, *Biochem. Pharmacol.*, 1990, **39**, 1620–1623.
14 S.-C. Chien, Y.-C. Wu, Z.-W. Chen and W.-C. Yang, *J. Evidence-Based Complementary Altern. Med.*, 2015, **357357**, 1–13.
15 E. J. Yang, S. H. Kim, K. Y. Lee and K. S. Song, *J. Microbiol. Biotechnol.*, 2018, **28**, 12–21.
16 Z. Wang, Y. Niu, T. Vashisth, J. Li, R. Madden, T. S. Livingston, *et al.*, *Hortic. Res.*, 2022, **9**, uhac145.
17 Z. Wang, J. Li, A. Chambers, J. Crane and Y. Wang, *J. Agric. Food Chem.*, 2021, **69**, 555–567.
18 H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, *et al.*, *Nat. Methods*, 2015, **12**, 523–526.
19 Z. Lai, H. Tsugawa, G. Wohlgemuth, S. Mehta, M. Mueller, Y. Zheng, *et al.*, *Nat. Methods*, 2018, **15**, 53–56.
20 X. Yang, P. Zhang and J. Wu, *Biochem. Syst. Ecol.*, 2014, **52**, 49–52.
21 S. Nonomura, *Pharm. Bull.*, 1957, **5**, 366–368.
22 Y. Hirose, *Chem. Pharm. Bull.*, 1960, **8**, 417–426.
23 P. Wang, J. Wei, X. Hua, G. Dong, K. Dziedzic, A.-T. Wahab, *et al.*, *J. Cell. Physiol.*, 2024, **239**, 1–49.
24 M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, *et al.*, *Nat. Biotechnol.*, 2016, **34**, 828–837.
25 M. A. Farag, M. H. Baky, I. Morgan, M. R. Khalifa, R. Rennert, O. G. Mohamed, *et al.*, *RSC Adv.*, 2023, **13**, 21471–21493.
26 H. Guo, Z. Chang, R. Yang, D. Guo and J. Zheng, *Phytochemistry*, 1998, **49**, 1623–1625.
27 Y.-H. Kuo, P.-H. Lee and Y.-S. Wein, *J. Nat. Prod.*, 2002, **65**, 1165–1167.
28 S. Jan, A. N. Kamili, J. A. Parray and Y. S. Bedi, *Nat. Prod. Res.*, 2016, **30**, 608–612.
29 S. Yao, Y. Li and L. Kong, *J. Chromatogr. A*, 2006, **1115**, 64–71.
30 S. M. Wong, M. M. Wong, O. Seligmann and H. Wagner, *Phytochemistry*, 1988, **28**, 211–214.
31 L. Zhu, S. Yu, X. Zeng, X. Fu and M. Zhao, *Sep. Purif. Technol.*, 2008, **63**, 665–669.
32 N. Lebrasseur, G.-J. Fan, M. Oxoby, M. A. Looney and S. Quideau, *Tetrahedron*, 2005, **61**, 1551–1562.
33 C. Maksut, T. Satake, K. Hori, Y. Saiki and M. Tanker, *Phytochemistry*, 1990, **29**, 2018–2020.
34 G. Song, H. Liu, W. Zhang, M. Geng and Y. Li, *Bioorg. Med. Chem.*, 2010, **18**, 5183–5193.
35 I. Boldizsar, Z. Szucs, Zs. Füzfai and I. Molnár-Perl, *J. Chromatogr. A*, 2006, **1133**, 259–274.
36 W. Steglich, W. Loesel and W. Reininger, *Tetrahedron Lett.*, 1967, **47**, 4719–4721.

37 M. Politi, R. Sanogo, K. Ndjoko, D. Guilet, J.-L. Wolfender, K. Hostettmann, *et al.*, *Phytochem. Anal.*, 2004, **15**, 355–364.
38 H.-H. Chan, C.-Y. Li, A. G. Damu and T.-S. Wu, *Chem. Pharm. Bull.*, 2005, **53**, 1232–1235.
39 P. H. Hynninen, R. Raisanen, P. Elovaara and E. Nokelainen, *Z. Naturforsch., C: J. Biosci.*, 2000, **55**, 600–610.
40 A. D. Pawlus, B.-N. Su, W. J. Keller and A. D. Kinghorn, *J. Nat. Prod.*, 2005, **68**, 1720–1722.
41 H. Luo, W. Qin, H. Zhang, F.-C. Ren, W.-T. Fang, Q.-H. Kong, *et al.*, *Molecules*, 2022, **27**, 1730.
42 J. Schripsema, A. Ramos-Valdivia and R. Verpoorte, *Phytochemistry*, 1999, **51**, 55–60.
43 X.-I. Yang, P. Zhang and J.-Z. Wu, *Shizhen Guoyi Guoyao*, 2014, **25**, 20–22.
44 Y. J. Yang, H. Y. Shu and Z. D. Min, *Yaoxue Xuebao*, 1992, **27**, 358–364.
45 G. W. Van Eijk and H. J. Roeijmans, *J. Chromatogr.*, 1984, **295**, 497–502.
46 X.-H. Cai, X.-D. Luo, J. Zhou and X.-J. Hao, *J. Nat. Prod.*, 2005, **68**, 797–799.
47 P. Sumthong, R. R. Romero-Gonzalez and R. Verpoorte, *J. Wood Chem. Technol.*, 2008, **28**, 247–260.
48 M.-J. Liou, C.-M. Teng and T.-S. Wu, *J. Chin. Chem. Soc.*, 2002, **49**, 1025–1030.
49 K. Saha, K. W. Lam, F. Abas, H. A. Sazali, J. Stanslas, L. S. Hui, *et al.*, *Med. Chem. Res.*, 2013, **22**, 2093–2104.
50 J.-L. Ren, A.-H. Zhang, L. Kong, Y. Han, G.-L. Yan, H. Sun, *et al.*, *Phytomedicine*, 2020, **67**, 153165.
51 J. Xie, A.-H. Zhang, H. Sun, G.-L. Yan and X.-J. Wang, *RSC Adv.*, 2018, **8**, 812–824.
52 H. Zhang, Y. Zhang, T. Zhang and C. Liu, *J. Pharm. Biomed. Anal.*, 2022, **211**, 114588.
53 L.-F. Nothias, D. Petras, R. Schmid, K. Dührkop, J. Rainer, A. Sarvepalli, *et al.*, *Nat. Methods*, 2020, **17**, 905–908.
54 Z. Cai, Z. Huang, M. He, C. Li, H. Qi, J. Peng, *et al.*, *Food Chem.*, 2023, **422**, 136169.