Journal of Pharmaceutical and Biomedical Analysis 236 (2023) 115722

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.journals.elsevier.com/journal-of-pharmaceutical-and-biomedical-analysis

Short communication

Phytochemical profiling of three *Amaranthus* species using LC-MS/MS metabolomic approach and chemometric tools

Ghada Abdel-Moez [a], Bharathi Avula [b], Hanaa Sayed [a], Azza Khalifa [a], Samir Ross [b,c,d,*], Kumar Katragunta [b], Ikhlas Khan [b,c], Shaymaa Mohamed [a]

[a] Department of Pharmacognosy, Faculty of Pharmacy, Assiut University, Assiut 71526, Egypt
[b] National Center for Natural Products Research, School of Pharmacy, University of Mississippi, University, MS 38677, USA
[c] Division of Pharmacognosy, Department of BioMolecular Sciences, School of Pharmacy, University of Mississippi, University, MS 38677, USA
[d] Asfendiyarov Kazakh National Medical University, Almaty, Kazakhstan

ARTICLE INFO

Keywords:
Amaranthus
Chemical profiling
Chemotaxonomic markers
Chemometrics
Principal component analysis
N-coumaroyl-L-tryptophan

ABSTRACT

Several *Amaranthus* vegetables (Amaranthaceae) have been recognized as valuable sources of minerals, vitamins, proteins, and phytonutrients, with health-promoting characteristics. In this study, three edible *Amaranthus* species, namely *A. hybridus* (AH), *A. blitum* (AB), and *A. caudatus* (AC), were chemically characterized using non-targeted liquid chromatography-tandem mass spectrometry (LC-MS/MS) technique. Further, multivariate chemometric analyses were conducted, including principal component analysis (PCA) and correlation-covariance plot (C-C plot). As a result, forty-one diverse compounds were identified, which varied in distribution and abundance across the investigated species. Amino acids and flavonoid glycosides were the most prevalent metabolites. Other identified compounds comprised nucleoside, chlorogenic acids, hydroxy cinnamoyl amides, and triterpenoid saponins. The most discriminant metabolites were flavonoid glycosides and hydroxy cinnamoyl amides, giving each species a chemotaxonomic identity. Advancing the chemotaxonomy of Amaranthaceae, adenosine nucleoside and *N*-coumaroyl-L-tryptophan were first reported from this family. Isorhamnetin and tricin glycosides were uniquely identified in AC, offering useful chemotaxonomic markers for this species. Notably, AB and AH profiles shared most metabolites, yet with varying abundance. These include adenosine, nicotiflorin, dicaffeoylquinic acids, and *N-trans*-feruloyl-4-*O*-methyldopamine. However, *N*-coumaroyl-L-tryptophan and kaempferol dirhamnoside were exclusively found in AB, separating it from AH. In conclusion, the applied analytical techniques established molecular fingerprints for the included species, identified specific biomarkers, and investigated their interconnections.

1. Introduction

Metabolomics is the study of "metabolomes" under a specified range of conditions. These metabolomes encompass all low molecular weight phytochemicals (up to 1800 Da). Given its high throughput, soft ionization, and extensive metabolome detection, LC-MS has gained prominence as a metabolomic tool [1].

Known for their tasty tender leaves and superior nutritive values, most *Amaranthus* species are consumed as leafy vegetables and grains worldwide [2]. Several amaranth species have been proven to possess various therapeutic properties such as analgesic, hypoglycemic, anti-inflammatory, anti-hyperlipidemic, and hepatoprotective [3–6]. These properties are mostly related to excellent antioxidant capacities that have been attributed to high polyphenols levels [7]. Nevertheless, only a few studies have investigated the chemical composition of these species. Additionally, previous studies were more concerned with the seeds of AC, but little is known about the leaves. Most studies have determined total phenolics, total flavonoids, or other phytochemicals in *Amaranthus* species, without analysis of their individual constituents [2,8,9]. Other studies have focused on their polyphenolic profiles; these studies were comprehensively summarized in a literature review [10]. Fourteen aliphatic compounds have recently been identified from AB through GC-MS analysis [5]; however, no data have been found on polar compounds.

To enrich chemical profiling of these species, this study investigated three *Amaranthus* species (AH, AB, and AC) using the hyphenated liquid

* Correspondence to: Department of Biomolecular Sciences, School of Pharmacy, University of Mississippi, University, MS 38677, USA.
E-mail address: sross@olemiss.edu (S. Ross).

https://doi.org/10.1016/j.jpba.2023.115722
Received 30 May 2023; Received in revised form 24 August 2023; Accepted 12 September 2023
Available online 14 September 2023

chromatography-tandem mass spectrometry (LC-MS/MS) technique. Besides, multivariate chemometric analyses were conducted, including principal component analysis (PCA) and correlation-covariance plot (C-C plot) to establish the relationship between the studied species. The applied analytical methods shed light on the heterogeneity of secondary metabolites classes and introduced respective chemotaxonomic markers for the studied species.

2. Experimental

2.1. Reference standards and solvents

Leucine, isoleucine, rutin, adenosine, phenylalanine, tyrosine, tryptophan and chlorogenic acid were purchased from Sigma (St. Louis, MO, USA). The identity and purity of these compounds were confirmed through spectral analysis, and they were established to be more than 95% pure. The solvents, including acetonitrile, ethanol, and formic acid, were of HPLC certified grade and obtained from Fisher Scientific (Santa Clara, CA, USA). Water used as a mobile phase was purified using a Milli-Q system from Millipore in Burlington, MA.

2.2. Plant materials

The aerial parts of AB were harvested in September 2020 from the Medicinal Plants Station, Pharmacognosy Department, Assiut University, Assiut, Egypt. Then, in October 2020, AH aerial parts were collected from Cairo-Alexandria agricultural road, while AC aerial parts were collected from Faculty of Agriculture, Assiut University, Assiut, Egypt. These species were identified by Dr. Mostafa Aboulela, Associate Professor of Plant Taxonomy, Botany and Microbiology Department, Faculty of Science, Assiut University. The collected plants were air-dried and powdered. The hydroalcoholic extract was prepared via macerating 50 g of powder in 0.5 L of 70% ethanol for 48 h and occasionally stirred. Then, the macerates were filtered and concentrated (40 °C, under reduced pressure) to give dried residues of approximately 5 g. Voucher specimens (No. 0002017, 0002029, and 0002030) were deposited in the Pharmacognosy Department Museum, Faculty of Pharmacy, Assiut University for AB, AH, and AC, respectively.

2.3. Liquid chromatography diode array detector-quadrupole time-of-flight mass spectrometry (LC-DAD-QToF)

The liquid chromatographic system was an Agilent Series 1290 comprised of a binary pump, vacuum solvent microdegasser, autosampler, thermostatically controlled column compartment, and photo diode array detector. Separation was achieved on an Agilent Poroshell 120 EC-C18 2.1 × 150 mm (2.7 μm) column. The mobile phase consisted of water with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B) at a flow rate of 0.2 mL/min. Analysis was performed using the following gradient elution: 99% A/1% B, isocratic for 3 min, in next 17 min to 60% A/40% B, in next 5 min to 20% A/80% B and in next 5 min to 100% B. Each run was followed by a 5 min wash with 100% B and an equilibration period of 5 min with 99% A/1% B. Two microliters of sample were injected. The column temperature was 35 °C.

The mass spectrometric analysis was performed with a QToF-MS-MS (Model #G6530A, Agilent Technologies, Santa Clara, CA, USA) equipped with an ESI source with Jet Stream technology using the following parameters: drying gas (N_2) flow rate, 11 L/min; drying gas temperature, 325 °C; nebulizer pressure, 30 psig, sheath gas temperature, 300 °C; sheath gas flow, 11 L/min; capillary voltage, 3500 V; nozzle voltage, 0 V; skimmer, 65 V; Oct RF V, 750 V; and fragmentor voltage, 135 V. All the operations, acquisition and analysis of data were controlled by Agilent MassHunter Acquisition Software Ver. A.05.01 and processed with MassHunter Qualitative Analysis Software Ver. B.06.00. Each sample was analyzed in positive mode over the range of m/z = 50–1300 and extended dynamic range (flight time to m/z 1700 at 2 GHz acquisition rate). A mass resolution of > 20,000 was applied to distinguish target analytes from interferences. Accurate mass measurements were obtained by means of reference ion correction using reference masses at m/z 121.0509 (protonated purine) and 922.0098 [protonated hexakis (1H, 1H, 3H-tetrafluoropropoxy) phosphazine or HP-921] in positive ion mode, while m/z 112.9856 (deprotonated trifluoroacetic acid-TFA) and 1033.9881 (TFA adducted HP-921) were used in negative ion mode. Samples were analyzed in all-ion MS-MS mode, where experiment 1 was carried out with collision energy of zero and experiment 2 with a fixed collision energy of 40 eV. For the ESI-MS-MS collision-induced dissociation (CID) experiments, precursor ions of interest were mass-selected by the quadrupole mass filter. The selected ions were then subjected to collision with nitrogen in a high-pressure collision cell. The compounds were confirmed in each spectrum. For this purpose, the reference solution was introduced into the ESI source via a T-junction using an Agilent Series 1200 isocratic pump (Agilent Technologies, Santa Clara, CA, USA) using a 100:1 splitter set at a flow rate of 20 μL/min.

2.4. Data processing and chemometric analysis

MassHunter Qualitative Analysis software (Agilent Technologies, Santa Clara, CA, USA) was used for the initial data processing of the raw LC-QToF data. The data were acquired from methanolic extracts in positive and negative ion modes, over the range of m/z 100–1300. The raw data were processed using the Find by Molecular Feature (MF) algorithm called Molecular Feature Extractor (MFE) within MassHunter Qualitative Analysis software. Extracted molecular features were processed to create a list of compounds. MFE allows for the unbiased grouping of all related ion signals into a single compound. Following feature extraction, the data were inspected and saved as a Compound Exchange File (.cef) for each sample file and imported into the Mass Profiler (Version B.03.00) was used to align mass and retention time data across the samples within the set to perform the statistical analyses required to profile the samples, including principal component analysis (PCA). In these steps, retention time and m/z indices alignment were carried out across the sample set using a tolerance window of 0.15 min and 10 mDa, respectively.

3. Results and discussion

3.1. LC-DAD-QToF method development and optimization

In this study, an analytical method, combining reversed-phase chromatography and tandem mass spectrometry, was developed for comparative chemical profiling of three *Amaranthus* species. For liquid chromatographic separation, the bonded stationary phase Poroshell 120 EC-C18 packing was selected, applying water (H_2O)/acetonitrile (CH_3CN) gradient solutions, modified with formic acid (FA), as mobile phases. The designed gradient elution provided fine-tuned separation and optimized resolution and sensitivity for various phytoconstituents of interest. Among acidic modifiers, FA at a concentration of 0.1% v/v was chosen relying on its lesser capacity to suppress fragment ion detection in QTOF analysis. This selection resulted in enhancement of the positive-ion electrospray ionization efficiency. Within 30 min, the gradient-customized elution systems, ranging from 1% CH_3CN in H_2O up to 100% CH_3CN, successfully detected diverse 41 phytochemicals. To assure satisfactory resolution, other separation parameters were adjusted, comprising column temperature of 35 °C and flow rates of 0.2 mL/min. Through proper internal calibration during acquisition and mass drift compensation (collected in single MS mode and applied to the tandem MS as a rolling average), the QToF instrument provided precise mass measurement with a limited error of ±5 ppm between measured and expected values. Most investigated constituents exhibited preference towards the positive-ion ionization mode. Representative total ion chromatograms (TICs) recorded for the three investigated *Amaranthus*

species in ESI positive ion mode are provided in (Fig. 1). In addition, some detected analytes showed responses in the negative-ion ionization mode. Representative total ion chromatograms (TICs) recorded for the studied extracts in ESI negative ion mode are included (Fig. 1). Extracted ion chromatograms of eight reference standards and *Amaranthus* species were shown as Fig. S1 of supplementary materials. To assist broad phytochemicals detection, four detection wavelengths, 254, 280, 330, and 370 nm were considered (Fig. S2 of supplementary materials). Taken together, our developed method surpassed other analytical methods in terms of several desirable criteria, including simplicity, fast analysis, broad dynamic range, and high accuracy.

3.2. Compounds characterization via their MS and MS/MS

Various constituents were identified at four different wavelengths using the positive ion mode with commonly detected precursor ions $[M+H]^+$ or/and $[M+Na]^+$ as well as the negative ion mode with frequent precursor ions $[M-H]^-$ or/and $[M+Cl]^-$, enabled precise interpretation of molecular formula. Besides, neutral losses of sugar moieties, such as ribose (Rib, 132 Da), glucose (Glc, 162 Da), and rhamnose (Rha, 146 Da), were commonly observed in the MS spectra to draw conclusion about the number of sugar substituents, their proper sequence, and type.

3.2.1. Amino acids

As listed in Table 1, six amino acids (peaks 1–6) were identified throughout the positive ESI, which exhibited an intense protonated precursor ion peak. Peak 1 was identified as valine, based on a protonated molecular ion peak $[M+H]^+$ at *m/z* 118.0863 (calcd 118.0863), produced *m/z* 72.0806 $[M+H-H_2O-CO]^+$ and *m/z* 55.0541 $[M+H-H_2O-CO-NH_3]^+$. Isoleucine/leucine (peaks 2 and 3) were identified based on the precursor ion adduct $[M+H]^+$ at *m/z* 132.1021 (calcd 132.1019). Moreover, they gave a fragment ion peak at *m/z* 118.0842 (calcd 118.0868), which was explained by the loss of a methyl group (CH_3, 14 Da) and produced two fragments by sequential losses of H_2O+CO (*m/z* 86.0968) and NH_3 (*m/z* 69.0701). Discrimination between these isomers, 2 and 3, was enabled based on their retention times. Peak 4 was established as tyrosine, based on a molecular ion peak $[M+H]^+$ at *m/z* 182.0820 (calcd 182.0812) and a fragment ion at *m/z* 165.0553 was formed by the loss of NH_3. Peak 5 was identified as phenylalanine, based on a molecular ion peak $[M+H]^+$ at *m/z* 166.0862 (calcd 166.0863). Further dissociation of this protonated molecular ion produced a characteristic fragment ion $[M+H-H_2O+CO]^+$ at *m/z* 120.0800 (calcd 120.0813), due to the loss of H_2O+CO moieties (46 Da) [11] and a dominant fragment ion at *m/z* 103.0532 (calcd 103.0547), formed by the further loss of NH_3 or corresponded to a deprotonated styrene $[styrene-H]^+$ [12]. Peak 6 was identified as tryptophan, based on a molecular ion peak $[M+H]^+$ at *m/z* 205.0981 (calcd 205.0972). Additionally, its MS/MS spectrum displayed a fragment ion $[M+H-NH_3]^+$ at *m/z* 188.0715 (calcd 188.0711). Notably, the preferential loss of ammonia (NH_3) from the protonated tryptophan has been previously proved as a fragmentation pattern, generated the diagnostic 2-carboxyspiro[cyclopropane-indolium] fragment ion [13]. As such, these findings highlight the anabolic benefit of the studied plants to supplement essential amino acids, offering a well-balanced, high--quality, and gluten-free protein source.

3.2.2. Nucleoside

In the positive ion mode, the MS/MS spectrum of peak 7 showed a parent protonated molecular ion $[M+H]^+$ at *m/z* 268.1045 (calcd 268.1040) and a product ion at *m/z* 136.0621 (136. 0623). This fragment ion was generated through cleavage of the glycosidic bond to give the protonated adenine aglycone, with the loss of a pentosyl D-ribose residue (Rib, 132 Da). Thus, this peak was identified as the adenosine nucleoside. Interestingly, this nucleoside is described for the first time from the family Amaranthaceae, providing new insights on the chemotaxonomic characteristics of the family. Defining the systematics of the studied species, it was detected in AB and AH. This indicates a close relationship between these two species and serves as a potential chemotaxonomic marker within the genus.

3.2.3. Chlorogenic and related caffeoylquinic acid derivatives

Four chlorogenic acid isomers (peaks 8–12), caffeoylquinic acids (CQA), were detected; their mass spectra shared the same deprotonated molecular ion peak $[M-H]^-$ at *m/z* 353.0869 and a fragment ion peak [quinic acid$-H]^-$ at *m/z* 191, owing to the caffeoyl moiety loss. As observed in Table 1, the elution orders, and distinct secondary ions, with variable relative abundances, allowed differentiation of these isomers. Specifically, peaks 8 and 9 showed similar fragment ions. Peak 8 was assigned as neochlorogenic acid [14] and yielded the deprotonated quinic acid (*m/z* at 191.0569) and *m/z* 135.0456 as the base peaks in the mass spectrum, and supplementary ions at *m/z* 173.0434 [quinic acid-$-H-H_2O]^-$ and 179.0338. Peak 9 was readily characterized as chlorogenic acid based on the diagnostic intense fragment ion at *m/z* 191.0563

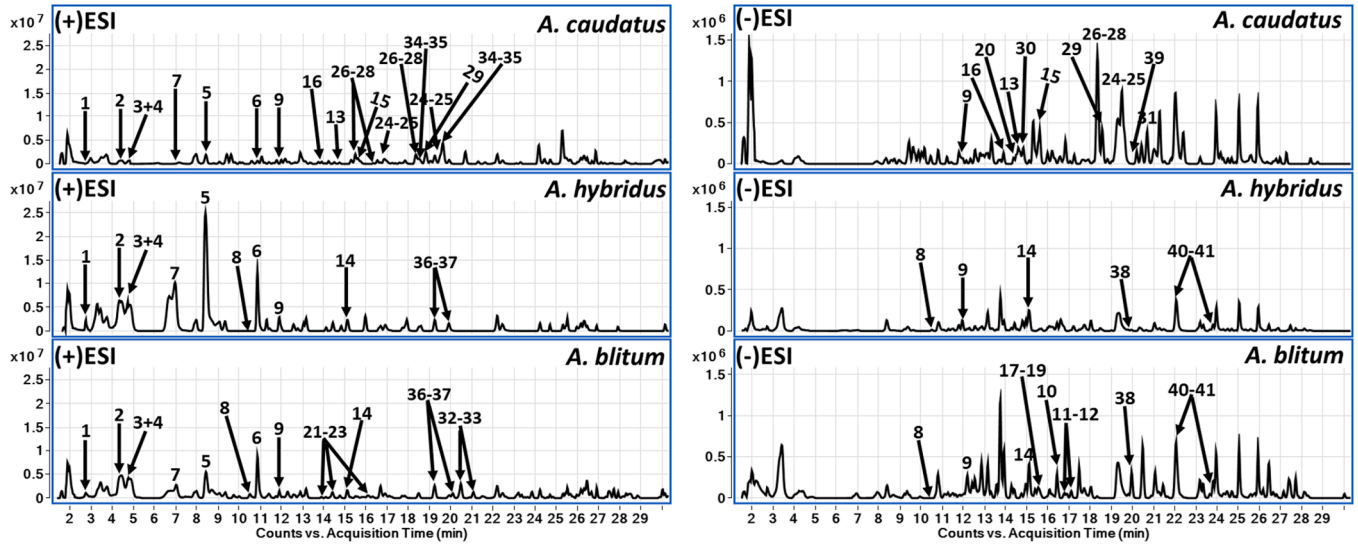

Fig. 1. Total Ion Chromatograms (TICs) recorded for three *Amaranthus* species (*A. caudatus*; *A. hybridus*; *A. blitum*) ethanolic extracts in ESI positive-ion and ESI negative-ion modes. Each sample was injected twice.

Table 1
Tentative identification of phytoconstituents in ethanolic extracts (10 mg/mL) of three *Amaranthus* species, *A. hybridus* (AH), *A. blitum* (AB), and *A. caudatus* (AC).

#	t_R (min)	Compound	Molecular formula	Positive mode adducts		Negative mode adducts		*Amaranthus species*		
				Prec. ion *m/z* $[M+H]^+$	Fragment ions	Prec. ion *m/z* $[M-H]^-$	Fragment ions	AH	AB	AC
1	2.8	Valine	$C_5H_{11}NO_2$	118.0863 (118.0863) *	72.0806 $[M+H-H_2O-CO]^+$, 55.0541 $[M+H-H_2O-CO-NH_3]^+$	-	-	+	+	+
2	4.5	Isoleucine	$C_6H_{13}NO_2$	132.1021	86.0968 $[M+H-H_2O+CO]^+$, 69.0701 $[M+H-H_2O+CO-NH_3]^+$	-	-	+	+	+
3	5.0	Leucine		(132.1019)						
4	4.8	Tyrosine	$C_9H_{11}NO_3$	182.0820 (182.0812)	165.0553 $[M+H-NH_3]^+$, 136.0758 $[M+H-CH_2O_2]^+$	-	-	+	+	+
5	8.9	Phenylalanine	$C_9H_{11}NO_2$	166.0862 (166.0863)	120.0800 $[M+H-H_2O+CO]^+$, 103.0532 $[M+H-H_2O+CO-NH_3]^+$	-	-	+	+	+
6	11.0	Tryptophan	$C_{11}H_{12}N_2O_2$	205.0981 (205.0972)	188.0715 $[M+H-NH_3]^+$	-	-	+	+	+
7	7.2	Adenosine	$C_{10}H_{13}N_5O_4$	268.1045 (268.1040)	136.0621 $[Adenine+H]^+$	-	-	+	+	+
8	10.6	Neochlorogenic acid	$C_{16}H_{18}O_9$	355.1031 (355.1024) 377.0846 (377.0843) $[M+Na]^+$	163.0390	353.0869 (353.0878)	191.0569 $[Quinic\ acid-H]^-$, 173.0434 $[C_7H_{11}O_6-H_2O]^-$, 179.0338 $[M-H-C_7H_{10}O_5]^-$, 161.0250 $[M-H-C_7H_{10}O_5-H_2O]^-$, 135.0457 $[M-H-C_7H_{10}O_5-CO_2]^-$	+	+	ND
9	12.1	Chlorogenic acid					179.0339 $[M-H-C_7H_{10}O_5]^-$, 135.0456 $[M-H-C_7H_{10}O_5-CO_2]^-$, 191.0563 $[Quinic\ acid-H]^-$,	+	+	+
10	16.3	Dicaffeoylquinic acids	$C_{25}H_{24}O_{12}$	-	-	515.1190 (515.1195)	353.0859 $[M-H-caffeoyl\ moiety]^-$, 191.0556 $[Quinic\ acid-H]^-$	t	t	t
11	16.6									
12	17.1									
13	14.6	Quercetin diglycoside	$C_{27}H_{30}O_{16}$	611.1609 (611.1607)	303.0481 $[Quercetin+H]^+$	-	301.0357 $[M-H-146-Glc]^-$	ND	t	+
14	15.1	Rutin		633.1417 (633.1426) $[M+Na]^+$	465.1044 $[M+H-146]^+$, 303.0514 $[M+H-146-162]^+$	609.1456 (609.1461)	301.0355 $[M-H-146-162]^-$	++	++	+
15	15.6	Isoquercetrin	$C_{21}H_{20}O_{12}$	465.1033 (465.1028)	303.0499 $[M+H-Glc]^+$	463.0877 (463.0882)	301.0351 $[M-H-146-Glc]^-$	+	+	++
16	13.9	Quercetin 3-rhamninoside	$C_{33}H_{40}O_{20}$	757.2188 (757.2186) 779.2002 (779.2005) $[M+Na]^+$	611.1603 $[M+H-146]^+$, 303.0511 $[M+H-Glc-146-146]^+$	755.2031 (755.2040)	301.0355 $[M-H-Glc-146-146]^-$	t	t	+
17	15.5	Kaempferol-dirhamnoside	$C_{27}H_{30}O_{14}$	579.1704 (579.1708)	433.1143 $[M+H-146]^+$, 287.0566 $[M+H-146-146]^+$	577.1555 (577.1563)	285.0399 $[M-H-146-146]^-$	t	+	t
18	15.3	Kaempferol 3-apiosyl-(1→4)-rhamnoside-7-rhamnoside	$C_{32}H_{38}O_{18}$	711.2130 (711.2131) 733.1945 (733.1950) $[M+Na]^+$	579.1704 $[M+H-132]^+$, 433.1143 $[M+H-132-146]^+$, 287.0565 $[M+H-132-146-146]^+$	709.1978 (709.1985)	565.1906 $[M-H-146]^-$, 285.0390 $[M-H-132-146-146]^+$	ND	+	ND
19	15.5									
20	14.3	Kaempferol 3-*O*-rutinoside 7-*O*-glucoside	$C_{33}H_{40}O_{20}$	757.2189 (757.2186) 779.2007 (779.2005)	595.1675 $[M+H-Glc]^+$ 449.1094 $[M+H-Glc-146]^+$ 287.0557 $[M+H-2Glc-146]^+$	755.2035 (755.2040)	285.0393 $[M-H-2Glc-146]^-$	ND	ND	+
21	13.9	Kaempferol-glucoside-rhamnoside	$C_{27}H_{30}O_{15}$	595.1660 (595.1657)	449.1090 $[M+H-146]^+$, 287.0557 $[M+H-146-Glc]^+$	593.1503 (593.1512)	285.0395 $[M-H-146-Glc]^-$	ND	+	ND
22	14.5							ND	+	ND
23	16.1	Nicotiflorin						+	+	ND
24	16.8	Isorhamnetin glucoside	$C_{22}H_{22}O_{12}$	479.1193 (479.1184)	317.0671 $[M+H-Glc]^+$	477.1031 477.1038	315.0529 $[M-H-Glc]^-$	ND	ND	+
25	19.5									
26	15.5	Isorhamnetin-glucoside-rhamnoside	$C_{28}H_{32}O_{16}$	625.1768 (625.1763) 647.1584 (647.1583) $[M+Na]^+$	479.1191 $[M+H-146]^+$, 317.0668 $[M+H-146-Glc]^+$	623.1610 (623.1618)	315.0517 $[M-H-146-Glc]^-$	ND	ND	+
27	16.3									
28	18.3									

(*continued on next page*)

Table 1 (*continued*)

#	t_R (min)	Compound	Molecular formula	Positive mode adducts		Negative mode adducts		*Amaranthus species*		
				Prec. ion *m/z* $[M+H]^+$	Fragment ions	Prec. ion *m/z* $[M-H]^-$	Fragment ions	AH	AB	AC
29	18.6	Isorhamnetin 3-vicianoside	$C_{27}H_{30}O_{16}$	611.1608 (611.1607)	317.0665 $[M+H-132-162]^+$	609.1466 (609.1461)	315.0521 $[M-H-132-162]^-$	ND	ND	+
30	14.7	Isorhamnetin-3-*O*-rhamnosylrutinoside	$C_{34}H_{42}O_{20}$	771.2339 (771.2342) 793.2155 (793.2162)	317.0670 $[M+H-146-146-162]^+$	769.2188 (769.2197)	315.0572 $[M-H-146-146-162]^-$	ND	ND	+
31	20.7	Tricin glucoside/Eupalitn glucoside	$C_{23}H_{24}O_{12}$	493.1346 (493.1341) 515.1168 (515.1165) $[M+Na]^+$	331.0811 $[M+H-Glc]^+$	491.1190 (491.1195)	329.0661 $[M-H-Glc]^-$	ND	ND	+
32	20.5	*N*-Coumaroyl-L-tryptophan	$C_{20}H_{18}N_2O_4$	351.1325 (351.1339)	147.0447 $[C_9H_7O_2+H]^+$	349.1178 (349.1183)	145.0288	ND	+	ND
33	21.1									
34	18.8	Sinapoyltyramine/	$C_{19}H_{21}NO_5$	344.1490 (344.1497)	145.0287	-	-	ND	ND	+
35	19.6	4-(2-Aminoethyl)phenol/ *N*-						ND	ND	+
36	19.4	(4-Hydroxy-3,5-dimethoxy-*E*-						+	+	ND
37	20.1	cinnamoyl)/ 3-Hydroxy-4-methoxyphenethylamine/ Feruloyl-4-*O*-methyldopamine/ *N*-(4-Hydroxy-3-methoxy-*E*-cinnamoyl)						+	+	ND
38	19.8	Dihydroxy-30-nor-oleanadien-23,28-dioic acid 3-*O*-D-glucurono-pyranoside, 28-*O*-D-glucopyranosyl ester	$C_{41}H_{60}O_{17}$	-	-	823.3751 (823.3758)	-	+	+	ND
39	20.0	3-*O*-β-D-glucopyranosyl-dihydroxy-30-nor-oleanadiene-23,28-dioic acid 28-*O*-β-D-glucopyranosyl ester	$C_{41}H_{62}O_{16}$	833.3919 (833.3935) $[M+Na]^+$	-	809.3952 (809.3960)	-	ND	+	+
40	22.2	Dihydroxy-30-nor-	$C_{35}H_{52}O_{11}$	671.3401	-	647.3431	-	+	t	ND
41	23.8	oleanadien-23,28-dioic acid 3-*O*-D-glucopyranoside/ Dihydroxy-30-nor-oleanadien-23,28-dioic acid, 28-*O*-D-glucopyranosyl ester/ 2,3,23-Trihydroxy-30-nor-12,20(29)-oleanadien-28-oic acid; (2β,3β)-form, 23-Carboxylic acid, 28-*O*-β- D-glucopyranosyl ester		(671.3402) $[M+Na]^+$		(647.3437)		+	+	ND

ND = Not detected; AH: A. hybridus; AB: A. blitum; AC: A. caudatus; t = trace; Glc = D-glucose. The parenthetical numbers (*) indicates the theoretical mass values.

and additional peaks at *m/z* 179.0339 and *m/z* 135.0456 [15].

Dicaffeoylquinic acids, peaks 10–12, were detected based on the negative ESI-MS spectrum, which exhibited $[M-H]^-$ at *m/z* 515.1190 (calcd 515.1195) and two fragment ions of $[M-H-caffeoyl]^-$ at *m/z* 353.0859 and $[quinic\ acid-H]^-$ at *m/z* 191.0556 (calcd 191.0555), due to the successive loss of the two caffeoyl moieties.

3.2.4. *Flavonoid glycosides*

Distinctive flavonoid profiles, with considerable qualitative and quantitative variation, were observed among the three species, presenting useful biomarkers. Four peaks (13–16) were assigned to be quercetin derivatives based on the ion at *m/z* $[quercetin+H]^+$ of 303, while peaks 17–23 were kaempferol derivatives defined with MS fragment ions of $[kaempferol+H]^+$ at *m/z* 287 or $[kaempferol-H]^-$ at *m/z* 285. Among quercetin glycosides, peak 14 represented a major constituent in both AB and AH but only a minor one in AC. It was identified as rutin as deduced from the protonated molecular ion peak $[M+H]^+$ at *m/z* 611.1609 (calcd 611.1607). A typical fragmentation pattern was observed in the MS/MS spectrum (Fig. S3 of supplementary materials), displaying fragment ions $[M+H-146]^+$ at *m/z* 465.1034 (calcd 465.1033) and $[M+H-146-162]^+$ at *m/z* 303.0514 (calcd 303.0504) through the neutral successive losses of the two sugar moieties. The negative ion mode further confirmed the structural identity based on a molecular ion peak $[M-H]^-$ 609.1456 (calcd 609.1461) and the characteristic fragment ion of 301.0355. Another quercetin diglycoside, peak 13, was detected at a different retention time, which showed the same molecular ion $[M+H]^+$ and a protonated fragment ion of the aglycone $[quercetin+H]^+$. Peak 15 was also detected in the three species, with the highest content being detected in AC, which identified as isoquercetrin, based on its MS spectrum which showed an ion peak $[M+H]^+$ at *m/z* 465.1033 (calcd 465.1028) and a fragment ion peak $[M+H-162]^+$ at *m/z* 303.0499 (calcd 303.0504). Peak 16 was detected in low amounts in AH and AB, identified as quercetin-rhamninoside at *m/z* 757.2188 and fragment ions at *m/z* 611.1603 $[M+H-146]^+$ and *m/z* 303.0511 $[M+H-Glc-146-146]^+$.

Peak 17 was assigned to kaempferol-dirhamnoside based on a parent ion peak $[M+H]^+$ at *m/z* 579.1704 (calcd 579.1708) and a fragment ion after the loss of a rhamnose moiety for $[M+H-146]^+$ at *m/z* 433.1143 (calcd 433.1134) and another fragment of the protonated aglycone $[kaempferol+H]^+$ at *m/z* 287.0566 (calcd 287.0555), after the loss of the second rhamnose unit. Peaks 18 and 19 have an accurate mass at *m/z* 711.2130 and fragment ions at *m/z* 579.1704, 433.1143 and 287.0565 (consecutive losses of two rhamnosyl and apiosyl units); these compounds were identified as kaempferol-apiosyl-dirhamnosides. The compound with a pseudomolecular ions $[M+H]^+$ at *m/z* 757.2189 and its fragment ions at *m/z* 595.1675, 449.1094 and 287.0557 (consecutive losses of rutinoside and hexosyl units) was tentatively identified as kaempferol-rutinoside-glucoside (peak 20). Peaks 21–23 were identified

as kaempferol-glucoside-rhamnoside based on a precursor ion peak of $[M+H]^+$ at *m/z* 595.1660 (calcd 595.1657) and a fragment ion of $[M+H-146]^+$ at *m/z* 449.1090 (calcd 449.1083), due to the loss of the rhamnose unit. Another fragment for the protonated kaempferol $[kaempferol+H]^+$ at *m/z* 287.0557 (calcd 287.0555) was also detected. The negative ion mode confirmed this identification through a molecular ion peak $[M-H]^-$ at *m/z* 593.1503 (calcd 593.1512) and a fragment ion of $[kaempferol-H]^-$ at *m/z* 285.0395 (calcd 285.0399). These quercetin and kaempferol glycosides have wide distribution in *Amaranthus* species [10].

Seven isorhamnetin glycosides (24−30) were identified only in AC. Peaks 24 and 25 were identified as isorhamnetin glucosides, deduced from a parent ion peak at *m/z* 479.1193 (calcd 479.1184). The mass spectrum also showed a fragment ion of $[M+H-162]^+$ at *m/z* 317.0671 (calcd 317.0661), due to the loss of the glucose or galactose unit. Peaks 26–28 were assigned as isorhamnetin-glucoside-rhamnoside, as deduced from a protonated ion peak $[M+H]^+$ at *m/z* 625.1768 (calcd 625.1763) and verified by another molecular ion $[M+Na]^+$ at *m/z* 647.1584 (calcd 647.1583). The MS/MS spectrum facilitated the identification through the neutral successive losses of the two sugar moieties. First, the loss of the rhamnose unit (Rha, 146 Da), resulted in $[M+H-146]^+$ fragment ion at *m/z* 479.1191 (calcd 479.1189), and subsequently the loss of the glucose unit (Glc, 162 Da) gave $[M+H-146-162]^+$ ion at *m/z* 317.0668 (calcd 317.0661). Peak 29 was identified as isorhamnetin -vicianoside based on a molecular ion peak $[M+H]^+$ at *m/z* 611.1608 (calcd 611.1607). The dominating fragment ion was $[M+H-132-162]^+$ at *m/z* 317.0665 (calcd 317.0661), corresponding to the loss of anhydro-arabinose (Ara, 132 Da) and glucose (Glc, 162 Da) residues. Peak 30 was characterized as isorhamnetin-*O*-rhamnosylrutinoside, based on a precursor ion peak $[M+H]^+$ at *m/z* 771.2339 (calcd 771.2342) in the positive mode, and confirmed through a molecular ion peak of $[M-H]^-$ at *m/z* 769.2188 (calcd 769.2197) in the negative mode. In addition, a secondary ion of the protonated aglycone $[isorhamnetin+H]^+$ at *m/z* 317.0670 (calcd 317.0661) was observed. Such *O*-methylated flavanol glycosides are rare in *Amaranthus*; nevertheless, they have been reported for *A. spinosis* [16].

One *O*-methylated flavone derivatives (31) were among the interesting, detected flavonoids, being only observed in AC, for the first time in this study. Peak 31 was annotated as either tricin glucoside or eupalitin glucoside, based on an intense molecular ion peak $[M+H]^+$ at *m/z* 493.1346 (calcd 493.1341) and $[M+Na]^+$ at *m/z* 515.1168 (calcd 515.1165). The fragmentation of the parent ion through cleavage of the glycosidic linkage and loss of the glucose unit provided a predominant fragment of the protonated aglycone at *m/z* 331.0811 (calcd 331.0817). Significantly, tricin derivatives have never been detected before in *Amaranthus* species.

3.2.5. Hydroxycinnamic acid amides

Six hydroxycinnamic acid amides, peaks 32–37, were detected. Peaks 32 and 33 were identified as *N*-coumaroyl-L-tryptophan (either *cis* or *trans* isomers at the double bond of the coumaroyl moiety), based on the positive ion mode MS spectrum, which showed a molecular ion peak of $[M+H]^+$ at *m/z* 351.1325 (calcd 351.1339). The negative ESI-MS-MS (Fig. S4 of supplementary materials) confirmed this identification through a deprotonated molecular ion peak at $[M-H]^-$ at *m/z* 349.1178 (calcd 349.1183) and diagnostic fragments at *m/z* 119.0497 (calcd 119.0496), 186.0555 (calcd 186.0555), and 229.0613 (calcd 229.0613), resulted from the reported fragmentation pathway [17], involving cleavages at two bonds.

Peaks 34–37 were established to be *N-trans*-sinapoyltyramines or derivatives, and provided a protonated precursor ion peak $[M+H]^+$ at *m/z* 344.1490 (calcd 344.1497). They were distinguished based on the predicted retention times. Considering the variable distribution among the three studied species, these amides could serve as key discriminant phytochemicals. Interestingly, the coumaroyl- and sinapoyl-amide conjugates, were first identified in Amaranthaceae family, whereas the feruloyl-amide, has been previously identified from AH [18]. In line with our findings, feruloyl amides have also been found in other *Amaranthus* species [19], and other Amaranthaceae members [20,21].

3.2.6. Triterpenoid saponins

Four nor-oleanane-type triterpenoid saponins (peaks 38–41) were detected. Peak 38 was proposed to be dihydroxy-30-nor-oleanadien-23,28-dioic acid, 3-*O*-D-glucurono-pyranoside, 28-*O*-D-glucopyranosyl ester based on the negative ion mode MS spectrum, which showed a molecular ion peak of $[M-H]^-$ at *m/z* 823.3751 (calcd 823.3758). Peak 39 was identified as 3-*O*-β-D-glucopyranosyl-2β,3β-dihydroxy-30-nor-oleane-12,20(29)-diene-23,28-dioic acid 28-*O*-β-D-glucopyranosyl ester based molecular ion peaks at *m/z* 833.3919 $[M+Na]^+$ (calcd 833.3935) and $[M-H]^-$ at *m/z* 809.3952 (calcd 809.3960). Peaks 40 and 41 were tentatively identified as two nor-oleanane monodesmosides based on the precursor ion adduct $[M-H]^-$ at *m/z* 647.3431 (calcd 647.3437). These four saponins shared the same nor-oleanane sapogenin, with glucuronic acid and/or glucose as the sugar substituents at C-3 and/or C-28. Reviewing literature data, these sugars have been reported to be predominant in numerous Amaranthaceae saponins [22]. Nor-oleanane saponins have been previously reported from several *Amaranthus* species, which is significant from a chemotaxonomic perspective. In specific, the same compound as peak 38 has been reported from *A. cruentus* seeds [23] and a similar compound as peak 39 has been reported from AC seeds [24].

3.3. Principle component analysis and correlation-covariance (C-C) plot

Several challenges are conjugated with LC-MS metabolomics in terms of data analysis, metabolite annotation, and identifying biological molecules from mass spectral data. To address such challenges, the acquired LC profiles of *Amaranthus* species were further analyzed via PCA and correlation-covariance (C-C) plot. PCA is a powerful approach in chemometrics that pinpoint the effects of technical variance in metabolic profile analysis and evaluate the quality of the data [25]. Following the pre-processing LC-MS data of the *Amaranthus* species metabolites, PCA was applied to explore sample clustering for both positive and negative ion modes data, and to evaluate the variation within the *Amaranthus* samples. Using PCA, the data's dimensionality is decreased while most data sets variation is preserved. This dimensionality reduction is achieved through finding the key directions—known as components—along which the data's maximum variance occurs. The remaining principal components, having minor effects on the model, were discarded. This analysis successfully differentiated AB, AH, and AC, establishing three reliable clustered groups, based on nature and the levels of metabolites. As shown in Table 2, the first principal component (PC1) mostly summarized the entire dataset since it captured most of the dataset variability in metabolite patterns. While the second principal component (PC2) revealed the second most variation in the dataset. The first two principal components (PC1 and PC2), representing 95.95% of the total variance, were extracted for analysis. The PCA scores plot (Fig. 2) demonstrated that differences among clusters along PC1 axis are larger than the similar-looking distances along PC2 axis. Additionally, AC was shown to be the most distant species compared with the other two species which were placed separately on the right side of the PC1, this demonstrating how the metabolites of AC were different from those

Table 2
Total variance explained for PCA of three *Amaranthus* species, *A. hybridus*, *A. blitum*, and *A. caudatus*.

Component	Variance (%)	Cumulative variance (%)
PC1	63.31	63.31
PC2	32.64	95.95
PC3	3.49	99.44
PC4	0.39	99.83

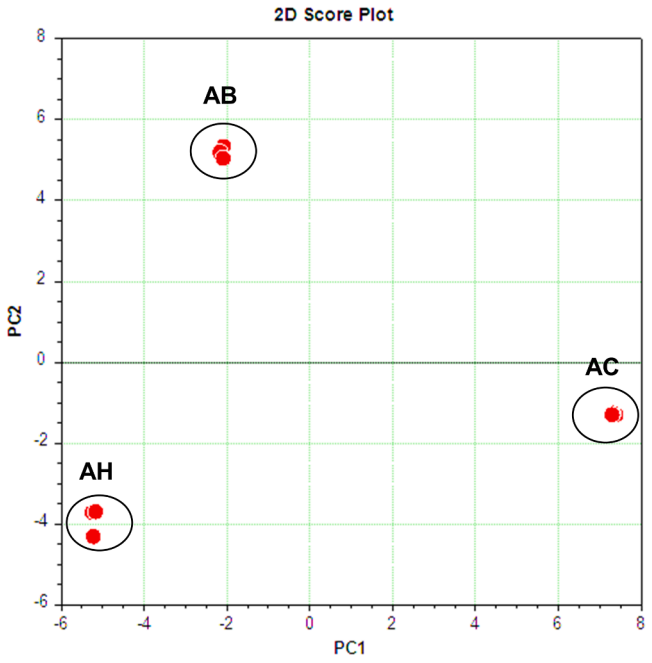

Fig. 2. The principal component analysis (PCA) score plot of the first two principal components for three Amaranthus species: *A. hybridus* (AH), *A. blitum* (AB), and *A. caudatus* (AC).

of the closely related species, AB and AH. Furthermore, PC2 can discriminate AB from AH by distinct clustering separation between the AB group in the positive side and AH group in the negative side, which maximizes the variance between these two samples.

The C-C plot is a helpful tool for understanding how data are separated depending on the compound variance. The C-C plot, generated from the PCA analysis of the metabolites of the *Amaranthus* species datasets, is shown in (Fig. 3). In this plot, each species is plotted based on the correlation score (the reliability of the relationship between the species metabolites) and the covariance score (the amount of change for each observed group across the dataset). Data points in each group were either relatively close to each other, indicating a close relationship among different compounds with same scattering and suggesting same sample variability.

4. Conclusions

The described hyphenated approach was successfully applied to identify diverse metabolites from *Amaranthus* species under investigation. Flavonoid glycosides and hydroxy cinnamoyl amides were the most discriminant metabolites, enabling discrimination between the studied species. Notably, isorhamnetin- and tricin-based glycosides were the most detected flavanol glycosides in AC. Under our experimental conditions, they were not detected in the other two species. Multivariate chemometric analyses clearly revealed the close relation between AH and AB, which shared most metabolites, barring a few variations. Evidently, AC was distinct from the other two species. In light of their phytochemical compositions and nutritional qualities, these *Amaranthus* species could serve as accessable adequate diet for food-insecure communities.

CRediT authorship contribution statement

Ghada Abdel-Moez: Investigation, Visualization, Writing – original draft. **Bharathi Avula:** Investigation, Visualization, Methodology, Validation, Data curation, Resources. **Hanaa Sayed:** Conceptualization, Supervision, Writing – review & editing. **Azza Khalifa:** Conceptualization, Supervision, Writing – review & editing. **Samir Ross:** Supervision, Resources, Project administration. **Kumar Katragunta:** Investigation, Visualization, Methodology, Data curation. **Ikhlas Khan:** Resources, Project administration. **Shaymaa Mohamed:** Investigation, Visualization, Supervision, Writing – review & editing.

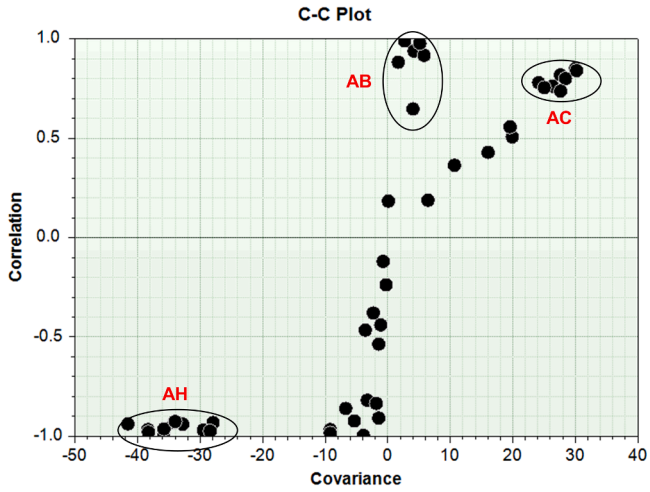

Fig. 3. Correlation-Covariance plot for LC-MS/MS datasets comparing three *Amaranthus* species: *A. hybridus* (AH), *A. blitum* (AB), and *A. caudatus* (AC).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors are grateful to the Department of Pharmacognosy, Faculty of Pharmacy, Assiut University, and The University of Mississippi, and the National Center for Natural Products Research, USA for financial support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jpba.2023.115722.

References

[1] D. González-Peña, L. Brennan, Recent advances in the application of metabolomics for nutrition and health, Annu. Rev. Food Sci. Technol. 10 (2019) 479–519.
[2] D.M. Jiménez-Aguilar, M.A. Grusak, Minerals, vitamin C, phenolics, flavonoids and antioxidant activity of *Amaranthus* leafy vegetables, J. Food Compos. Anal. 58 (2017) 33–39.
[3] K. Girija, K. Lakshman, Anti-hyperlipidemic activity of methanol extracts of three plants of *Amaranthus* in triton-*WR* 1339 induced hyperlipidemic rats, Asian Pac. J. Trop. Biomed. 1 (2011) S62–S65.
[4] E.U. Ejiofor, S.O. Oyedemi, S.O. Onoja, N.Y. Omeh, *Amaranthus hybridus* Linn. leaf extract ameliorates oxidative stress and hepatic damage abnormalities induced by thioacetamide in rats, S. Afr. J. Bot. 146 (2022) 213–221.
[5] S. Jahan, M. Nesa, M.E. Hossain, J.C. Rajbangshi, M.S. Hossain, In vivo and in silico evaluation of analgesic and hypoglycemic activities of *Amaranthus blitum* L, S. Afr. J. Bot. 150 (2022) 565–575.
[6] A. Kar, S. Bhattacharjee, Bioactive polyphenolic compounds, water-soluble vitamins, in vitro anti-inflammatory, anti-diabetic and free radical scavenging properties of underutilized alternate crop *Amaranthus spinosus* L. from Gangetic plain of West Bengal, Food Biosci. 50 (2022), 102072.

[7] P.I. Adegbola, A. Adetutu, T.D. Olaniyi, Antioxidant activity of *Amaranthus* species from the Amaranthaceae family – a review, S. Afr. J. Bot. 133 (2020) 111–117.
[8] U. Sarker, S. Oba, Nutrients, minerals, pigments, phytochemicals, and radical scavenging activity in *Amaranthus blitum* leafy vegetables, Sci. Rep. 10 (2020) 1–9.
[9] M.O. Jimoh, A.J. Afolayan, F.B. Lewu, Nutrients and antinutrient constituents of *Amaranthus caudatus* L. cultivated on different soils, Saudi J. Biol. Sci. 27 (2020) 3570–3580.
[10] P. Kongdang, N. Dukaew, D. Pruksakorn, N. Koonrungsesomboon, Biochemistry of *Amaranthus* polyphenols and their potential benefits on gut ecosystem: a comprehensive review of the literature, J. Ethnopharmacol. 281 (2021).
[11] P. Zhao, T. White, R. Graham Cooks, Q. Chen, Y. Liu, H. Chen, Detection of neutral CO lost during ionic dissociation using atmospheric pressure thermal dissociation mass spectrometry (APTD-MS), J. Am. Soc. Mass Spectrom. 29 (2018) 2317–2326.
[12] A.G. Harrison, Effect of phenylalanine on the fragmentation of deprotonated peptides, J. Am. Soc. Mass Spectrom. 13 (2002) 1242–1249.
[13] H. Lioe, R.A.J. O'Hair, G.E. Reid, Gas-phase reactions of protonated tryptophan, J. Am. Soc. Mass Spectrom. 15 (2004) 65–76.
[14] A.T. Ramabulana, P. Steenkamp, N. Madala, I.A. Dubery, Profiling of chlorogenic acids from Bidens pilosa and differentiation of closely related positional isomers with the aid of UHPLC-QTOF-MS/MS-based in-source collision-induced dissociation, Metabolites 10 (2020).
[15] S.J. Rowland, J.A.F. Rook. Analytical Methods, 1961.
[16] F.C. Stintzing, D. Kammerer, A. Schieber, H. Adama, O.G. Nacoulma, R. Carle, Betacyanins and phenolic compounds from *Amaranthus spinosus* L. and *Boerhavia erecta* L. Z. Naturforsch. Sect. C J. Biosci. 59 (2004) 1–8.
[17] B.J. Knollenberg, G.X. Li, J.D. Lambert, S.N. Maximova, M.J. Guiltinan, Clovamide, a hydroxycinnamic acid amide, is a resistance factor against *Phytophthora* spp. in *Theobroma cacao*, Front. Plant Sci. 11 (2020).
[18] S.K. Steffensen, H.A. Pedersen, R. Labouriau, A.G. Mortensen, B. Laursen, R.M. De Troiani, E.J. Noellemeyer, D. Janovska, H. Stavelikova, A. Taberner, C. Christophersen, I.S. Fomsgaard, Variation of polyphenols and betaines in aerial parts of young, field-grown *Amaranthus* genotypes, J. Agric. Food Chem. 59 (2011) 12073–12082.
[19] H.A. Pedersen, S.K. Steffensen, C. Christophersen, A.G. Mortensen, L.N. Jørgensen, S. Niveyro, R.M. De Troiani, R.J. Rodríguez-Enríquez, A.P. Barba-De La Rosa, I. S. Fomsgaard, Synthesis and quantitation of six phenolic amides in *Amaranthus* spp, J. Agric. Food Chem. 58 (2010) 6306–6311.
[20] A.H.M. Khurshid Alam, G. Sadik, H. Or-Rashid, C.M. Hasan, M.A. Rashid, *N-trans*-feruloyl-4-methyldopamine from *Achyranthes ferruginea*, Biochem. Syst. Ecol. 31 (2003) 1345–1346.
[21] J. Martin-Tanguy, F. Cabanne, E. Perdrizet, C. Martin, The distribution of hydroxycinnamic acid amides in flowering plants, Phytochemistry 17 (1978) 1927–1928.
[22] J. Zehring, V. Reim, D. Schröter, S. Neugart, M. Schreiner, S. Rohn, R. Maul, Identification of novel saponins in vegetable amaranth and characterization of their hemolytic activity, Food Res. Int. 78 (2015) 361–368.
[23] M. Junkuszew, W. Oleszek, M. Jurzysta, S. Piancente, C. Pizza, Triterpenoid saponins from the seeds of *Amaranthus cruentus*, Phytochemistry 49 (1998) 195–198.
[24] L. Rastrelli, R. Aquino, S. Abdo, M. Proto, F. De Simone, N. De Tommasi, Studies on the constituents of *Amaranthus caudatus* leaves: isolation and structure elucidation of new triterpenoid saponins and ionol-derived glycosides, J. Agric. Food Chem. 46 (1998) 1797–1804.
[25] E.R. Ziegel. Statistics and Chemometrics for Analytical Chemistry, 2004.