AJ ACADEMIC JOURNALS
expand your knowledge

Journal of Medicinal Plants Research

Full Length Research Paper

Comprehensive liquid chromatography-mass spectrometry-based metabolomic analysis of *Moringa oleifera* seeds

Famurewa Oluwayemisi Juliannah*, Istifanus Yarkasuwa Chindo and Auwal Adamu Mahmoud

Department of Chemistry, Faculty of Sciences, Abubakar Tafawa Balewa University Bauchi, Nigeria.

Moringa oleifera seeds which are less explored and nutriment-rich have attracted scientific interest as the seed kernels contain numerous bioactive components with a variety of traditional uses. Besides its medicinal uses, *Moringa oleifera* biodiesel has shown remarkable potentiality in conducing to the decrease of greenhouse gases and guaranteeing sustainable supply of energy. In this study, the comprehensive analysis of the *M. oleifera* seeds metabolome was carried out by generating a Molecular Network (MN) from Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) data to profile the ethyl acetate extract. The dereplication information was then collected by the MN, which then compares the MS/MS spectra of the investigated compounds and groups them into clusters based on their fragmentation route similarities. Therefore, identification of the compounds was conducted based on their full MS and MS/MS spectra obtained in positive ion mode. Through mass spectrometry-based molecular networking a total of 54 metabolites were putatively identified encompassing different classes including coumarins, alkaloids, amino acids, flavonoids, terpenoids, fatty acids, steroids and lipids among others. Thus, the identification highlights that *M. oleifera* seeds could serve as potential biomarker for new drug discovery and can have a wide variety of applications in food industry. Also, these fatty acids (saturated and unsaturated) suggest that the seed is a good candidate for biodiesel production, since they are fundamental to whether *M. oleifera* seeds can be used as a biofuel feedstock.

Key words: *Moringa oleifera* seeds, biodiesel, sustainable energy, LC-MS/MS, molecular networking.

INTRODUCTION

Recent technological developments and methodological advances of both liquid chromatography (LC) and mass spectrometry (MS) have allowed LC-MS-based plant metabolomics to become a common tool for investigating quantity, quality, and chemical diversity of plant metabolites. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) is highly sensitive, selective,

and enables extensive detection of metabolites within a sample (Sawada and Hirai, 2013; Shimizu et al., 2018). Metabolomics generate a huge amount of metabolic data using wide range of analytical instruments. Although several analytical techniques can be used for metabolomics analysis (Wolfender et al., 2018), the application of LC–MS/MS-based metabolic profiling of biological systems

*Corresponding author. E-mail: famurewaoluwayemisi4@gmail.com.

has gained more extensive use in identifying drug metabolite, developing metabolite maps and lending clues to the mechanism of bioactivation (Goulitquer et al., 2012). Using LC-MS/MS based plant metabolomics approach, a few hundreds to thousands of metabolites with high molecular weight (>500 kDa), heat-labile functional groups, chemically unstable functional groups, and high-vapor-point can be detected in a plant extract. It does not require volatilization of the metabolites (Zeki et al., 2020; Piasecka et al., 2019).

LC-MS/MS paired with the computational technique of molecular networking is a cutting-edge data visualization approach that has most notably been used in discovering new drugs from natural sources. The chemical structure of a molecule dictates how it will fragment during the MS/MS procedure (Matt, 2022). Molecular Networking (MN) is a computational strategy used to visualize the structural link between molecules belonging to the same molecular family and interpret complex data arising from MS analysis making it easier to identify unknown metabolites (Messaili et al., 2020). MN is able to identify potential similarities among all MS/MS spectra within the dataset and to propagate annotation to unknown but related molecules (Wang et al., 2016).This approach exploits the assumption that structurally related molecules produce similar fragmentation patterns, and therefore they should be related within a network (Quinn et al., 2017).

In MN, MS/MS data are represented in a graphical form, where each node represents an ion with an associated fragmentation spectrum; the links among the nodes indicate similarities of the spectra. By propagation of the structural information within the network, unknown but structurally related molecules can be highlighted and successful dereplication can be obtained which are useful for metabolite identification (Vincenti et al., 2020; Yang et al., 2013). Metabolite extraction is a critical step prior to metabolomic experiments. The choice of solvents used for extraction is a key factor in determining the metabolites of interest to be extracted, since the main aim of the step is to extract as wide a spectrum of chemical compounds as possible from the sample in consideration (Lu et al., 2017). The ethyl acetate which is a medium polar solvent has been reported to be the best extraction solvent in terms of number of metabolites with large chemical and structural diversity detected by MS (Colnaghi et al., 2007; Beaulieu et al., 2013; Lindow et al., 2014; Di Masi et al., 2022).

Moringa oleifera Lam. belongs to a single genus family Moringaceae; a highly valued plant, distributed in many countries of the tropics and subtropics. It has an impressive range of medicinal uses with high nutritional value. The seed has continued to gain a wider acceptance in various global ethnomedicines for managing several communicable and lifestyle diseases aside its vital nutritional application as emerging food additives. *M. oleifera* seeds have been shown to elicit a myriad of pharmacological potential and health benefits, including:

antimicrobial, anticancer, antidiabetic, antioxidant, antihypertensive, anti-inflammatory and cardioprotective properties. The health benefits of bioactive components in the seeds are promising and demonstrate enough potential to facilitate the development of functional foods (Galuppo et al., 2013; Elsayed et al., 2015; Christian et al., 2022).

Studies on seeds are much accentuated on the purification of water and oil extraction. *M. oleifera* seeds are used as nature-based solutions for the problem of water purification in developing countries, using them as an alternative to Western methods (El-Haddad et al., 2019). The oil is not only free of toxicant, but it also exhibited high biological value as compared to commercial oil (Saa et al., 2019). Biodiesel produced from *M. oleifera* seed oil exhibit enhanced oxidative ability, high cloud point and a higher cetane number of approximately 67 which is higher than most biodiesels (Rashid et al., 2008).

However, fewer metabolomics studies have been conducted for large-scale detection of low molecular weight metabolite in *M. oleifera* seeds which is ideal for incorporation into diets and its metabolite composition contribute to its biological effect. Thus, as liquid chromatography tandem mass spectroscopy (LC-MS/MS) facilitates metabolite identification, quantification and identify patterns in chemical diversity in a complex mixture of molecules, we will gain a better understanding of the properties of the ethyl acetate extract by mapping the chemical profile to the nutritional and pharmacological effects of the *Moringa oleifera* seeds.

Therefore, this study will employ ultra-high performance liquid chromatography (UHPLC-MS/MS) metabolomics approach to comprehensively profile *M. oleifera* seeds metabolome to detect the bioactive metabolites present in the ethyl acetate extract. In addition, LC-MS/MS data will be subjected to a molecular networking analysis.

MATERIALS AND METHODS

Moringa oleifera seeds preparation

The mature seeds of *Moringa oleifera* were collected locally from the open market, Bauchi, Bauchi State, Nigeria and authenticated by a taxonomist. The seeds were de-husked manually, air dried and milled into fine powder with the aid of laboratory mortar and pestle. The fine powder was stored at room temperature before extraction.

Solvents and chemicals

Formic acid and acetonitrile of High Performance Liquid Chromatography grade were purchased from Baker (The Netherlands). All other solvents, standards, and chemicals were procured from Sigma Aldrich (St. Louis, MO, USA).

Moringa Oleifera seeds extraction procedure

100g of the powdered seeds was soaked in 300 mL of ethyl acetate

(medium polar solvent) for three days at room temperature.

The supernatant was then collected, filtered, and the solvent evaporated using a vacuum rotary evaporator. This step was done twice as according to the methodology described by Jeyaseelan et al. (2012) with little modification. The crude extract was stored at sterile laboratory conditions until further analysis. The ethyl acetate fraction was then used for LC-MS analysis.

Sample preparation for UHPLC-MS/MS analysis

The ethyl acetate extract (2 mg) was dissolved in LCMS-grade methanol (1 mL). Dissolved extract was vortexed for 10 min, centrifuged for 10 min and filtered through a nylon filter (0.22 µm) into a glass vial for LC-MS/MS analysis following method as described by De Oliveira et al. (2017).

Ultra-high performance liquid chromatography–MS/MS (UHPLC-MS/MS) analysis

Chromatography

Separation was performed using Thermo Scientific C18 column (AcclaimTM Polar Advantage II, 3 x 150 mm, 3 µm particle size) on an UltiMate 3000 UHPLC system (Dionex). Gradient elution was performed at flow rate of 0.4 mL/min and 40°C column temperature using H_2O + 0.1% Formic Acid (A) and 100% Acetonitrile (B) with 22 min total run time. The injection volume of sample was 3 µL. The gradient started at 5% B (0-3 min); 80% B (3-10 min); 80% B (10-15 min) and 5% B (15-22 min).

Mass spectrometry

High resolution mass spectrometry was carried out using a MicroTOF QIII Bruker Daltonic using an ESI positive ionization with the following settings: capillary voltage, 4500 V; nebulizer pressure, 2.0 bar; drying gas, 8 L/min at 300°C. The mass range was at 50-1500 m/z.

Data processing

The accurate mass data of the molecular ions, provided by the TOF analyzer, were processed by Compass Data Analysis software (Bruker Daltonik GmbH). The metabolites characterization was performed using Thermo Xcalibur 2.2.0 (Thermo Fisher Scientific Inc., Waltham, MA, USA) and their comparison was carried out using literature data and standard online databases (freely available), such as PubChem, Human Metabolome Database (HMDB), Chemspider, LIPID MAPS, Metanetx and Swisslipids. In addition, the MS^2 in positive ionization mode and relatively low mass error supported the confirmation of newly identified compounds.

Molecular networking (MN)

The molecular networks based on MS/MS data of *M. oleifera* seeds were generated using the online workflow Global Natural Products Social Molecular Networking (GNPS) platform (http://gnps.ucsd.edu, accessed on 5th April, 2023) with a registered account. The raw MS data including blank were first converted into mzXML format using MSConvert software downloaded from Proteowizard website (http://proteowizard.sourceforge.net/tools.shtml, accessed on 5th April, 2023) before uploading the data into GNPS. Then, the converted data files were uploaded to GNPS using FileZilla 3.64.0 software (https://filezilla-project.org/, accessed on 5th April, 2023). In the GNPS data analysis workflow, sample and blank data were selected as G1 and G2, respectively, with precursor ion mass tolerance set to 0.02 Da and a fragment ion mass tolerance of 0.02 Da. A network was processed with edges that were filtered to have a cosine score above 0.7 and a minimum 6 matched peaks in line with the procedure of Wu et al. (2015). After processing, the spectral networks were imported using the Cytoscape 3.9.1 software, and visualized using a force-directed layout (Institute of Systems Biology, Seattle, WA, USA).

RESULTS AND DISCUSSION

Metabolite profiling of *Moringa Oleifera* seeds extract Via-UHPLC-ESI MS/MS analysis

The ultra- high performance liquid chromatography (UHPLC) coupled with electrospray ionization (ESI)-micrOTOF-Q III which is a more advanced system known for its high resolution, sensitivity and excellent mass accuracy, was utilized to analyze ethyl acetate extract of *M. oleifera* seeds. The identification of the compounds was conducted based on their full MS and MS/MS spectra obtained in positive ion mode. All the compounds in the sample were readily ionized in the positive ion mode. The total ion chromatogram (TIC) of the extract is shown in Figure 1, in which a total of 41 chromatographic peaks were annotated. The identities with retention time in minutes (t_R) and fragment ion(s) for each metabolite are presented in Table 1, and the MS/MS spectra with the structures for every molecular ion detected are available in the Supplementary Material (Supplementary Figures S1 and S2).

MS/MS-based molecular networking

Molecular networking (MN) is also known as mass spectral networking. The MN is a graph-based workflow that organizes massive MS datasets by mining spectral similarity between different MS/MS fragmentation patterns, but structurally-related precursor ions. The basic principle underlying MN is to compare the MS/MS spectra of different ions in a sample and to organize those spectra based on similarities. The outcome is a network or graph, in which nodes represent precursor ions and edges represent spectral similarity between the MS/MS spectra of those ions (Nothias et al., 2020). In the present study, the metabolomics mass profile of *M. oleifera* seeds ethyl acetate extract was analyzed more comprehensively and accurately using the Global Natural Product Social Molecular Networking (GNPS) based on UHPLC-MS/MS analysis data. Figure 6 shows the generated MN with the different clusters in the network, whereby each cluster shares some distinct fragments and fragmentation patterns. The results demonstrated a total of 97 nodes assigned for the parent ions of *M. oleifera*
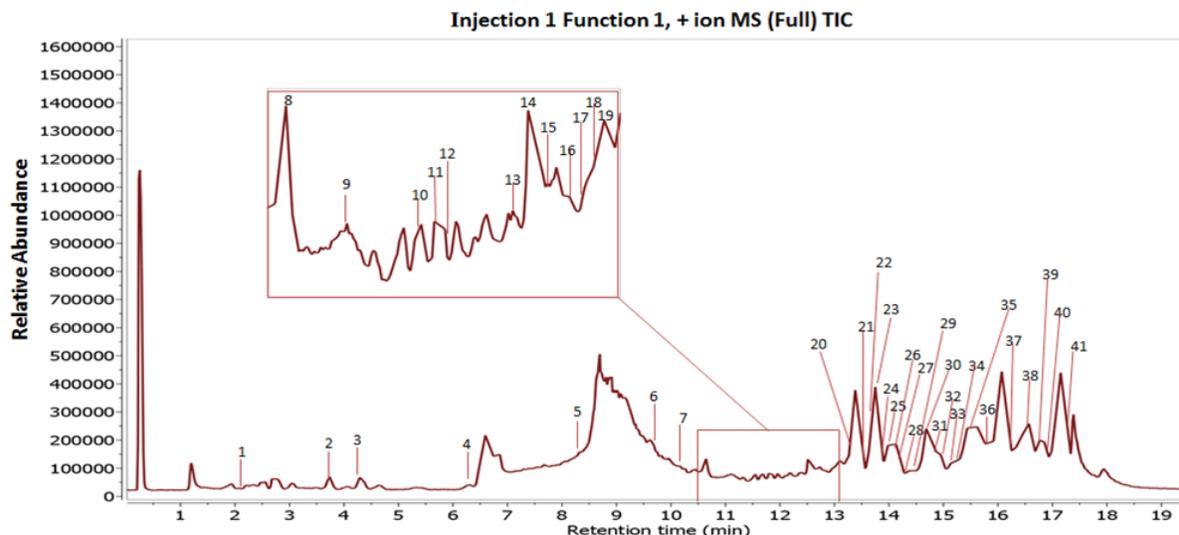
Figure 1. Total ion chromatogram (TIC) in positive ionization mode of *Moringa oleifera* seeds. The number above each peak represents peak numbers, corresponding to the peak numbers in Table 1.

seeds. The parent ions spectra in the network was matched with GNPS' spectral libraries and were also identified using different mass spectroscopic databases such as PubChem, HMDB and Chemspider platforms resulting in the annotation of 41 metabolites comprising coumarins, alkaloids, amino acids, flavonoids, terpenoids, fatty acids, and steroids among others Table 1.

Further identification using MN platform on MS/MS data discovered 12 glycerophospholipids that are categorized as (LysoPC, LysoPE, LysoPI, LysoPG, PG) with one glycerolipid as (DG) Table 2. Overall, UHPLC-ESI MS/MS analysis and molecular networking analysis resulted in the tentative identification of 54 compounds in *M. oleifera* seeds with all of them visualized in different clusters due to their slightly different structures.

Identification of coumarins

Coumarins containing the unique 2*H*-chromen-2-one motif are secondary metabolites beneficial to human health. They are known for their pharmacological properties such as anti-inflammatory, anticoagulant, antibacterial, antifungal, antiviral, anticancer, antihypertensive, antitubercular, anticonvulsant, antiadipogenic, antihyperglycemic, antioxidant, and neuroprotective properties (Venugopala et al., 2013). Three coumarins were present in the ethyl acetate extract of *M. oleifera* seeds. Peaks 33 and 41 with t_R 15.14 min and 17.34 min produced [M+Na]$^+$ molecular adduct ions of m/z 384 and 419 respectively. Both exhibited [M+H]$^+$ ions at m/z 362 and 397. At peak 33, the [M+H]$^+$ ion at m/z 362 produced a prominent ion at m/z 344 which was attributed to the loss of a H_2O molecule. The ion at m/z 344 was further fragmented by losses of CH_4O and CH_3COOH leading to

the formation of ions at m/z 330, and 302 respectively. Peak 41, the [M+H]$^+$ ion at m/z 397 yielded ions at m/z 379, 369 and 355 signaling the losses of H_2O, CO and C_2H_2O in a relative manner.

Another coumarin was putatively identified as 7-Hydroxy-2H-chromen-2-one, (peak 1), also known as umbelliferone at t_R 2.15 min. This peak showed a parent ion at m/z 161 [M-H]$^-$ ion and characteristics MS/MS fragments at m/z 133 due to loss of CO moiety, m/z 117 corresponding to the loss of CO_2 and m/z 105 produced as a result of loss of 2CO. The proposed mass fragments resulting from the fragmentation of 7-Hydroxy-2H-chromen-2-one is shown in (Figure 2) and is in agreement with previously published result (Zhou et al., 2018).

Identification of amino acids

Amino acids are the fundamental units of proteins, which are also the important components of active peptidases and other bioactive molecules (Liang et al., 2019; Duan et al., 2020). Five amino acid peaks were identified in the UHPLC chromatogram. At t_R 4.29 min, peak 3 with m/z 227 was observed to fragment into 209, 199 and 171. These signals result in the loss of a H_2O molecule, CO and C_4H_8 respectively. Peak 3 was tentatively identified as Cyclo(L-Leu-trans-4-hydroxy-L-Pro). Peak 4 with m/z 211was identified as Cyclo(L-Leu-L-Pro) by fragment ions m/z 183 and 155 at t_R 6.41 min by losses of CO and C_4H_8.

Interestingly, peak 11 and 12 with t_R 11.79 min and 11.87 min showed similar characteristics MS/MS fragmentation patterns. Peak 11 produced sodiumated molecular ion of m/z 467 and displayed protonated molecular ion of m/z 445 while peak 12 only give

Table 1. Putative metabolites identified based on LC-MS/MS and MN in ethyl acetate extract of *M. Oleifera* seeds.

Peak no.	RT (min)	Putative metabolite	Molecular formula	Adduct	SpecMZ	LibMZ	Mass error (ppm)	MS/MS fragments (ESI+)	Chemical class
1.	2.15	7-Hydroxy-2H-chromen-2-one	$C_9H_6O_3$	[M-H]-	159.066	161.024	-1.958	133, 117, 105	Coumarin
2.	3.73	Dextrorphan	$C_{17}H_{23}NO$	[M+H]+	259.078	258.190	0.892	240, 227, 201	Alkaloid
3.	4.29	Cyclo(L-Leu-trans-4-hydroxy-L-Pro)	$C_{11}H_{18}N_2O_3$	[M+H]+	229.068	227.140	1.929	209, 199, 171	Amino acid
4.	6.41	Cyclo(L-Leu-L-Pro)	$C_{11}H_{18}N_2O_2$	[M+H]+	209.154	211.145	-1.991	183,155	Amino acid
5.	8.47	1-[4-hydroxy-3-(3-methylbut-2-enyl)phenyl]ethanone	$C_{13}H_{16}O_2$	[M+H]+	205.086	205.122	-0.036	187, 163, 149	Organooxygen compound
6.	9.84	Tricetin	$C_{15}H_{10}O_7$	[M+H]+	304.298	303.050	1.248	301, 285, 126	Flavonoid
7.	10.12	Didymic acid	$C_{22}H_{26}O_5$	[M-H]-	368.352	369.170	-0.818	351, 325, 337	Dibenzofuran
8.	10.49	Benzyldimethyltetradecylammonium	$C_{23}H_{42}N$	[M+H]+	332.330	333.339	-1.009	304, 290, 278	Benzenoid
9.	11.12	3-(1,2,3,6-tetrahydropyridin-4-yl)-1H-pyrrolo[2,3-b]pyridine	$C_{12}H_{13}N_3$	[M+Na]+	223.060	222.100	0.959	183, 171, 157	Alkaloid
10.	11.74	Nonactic acid	$C_{10}H_{18}O_4$	[M+Na]+	225.194	225.278	-0.084	185, 157, 129	Polyketide
11.	11.79	Benzenepropanamide, N-[2-(acetyloxy)-1-(phenylmethyl)ethyl]-alpha-(benzoylamino)-	$C_{27}H_{28}N_2O_4$	[M+Na]+	467.193	467.194	-0.001	427, 403, 385, 353	Amino acid
12.	11.87	Aurantiamide acetate	$C_{27}H_{28}N_2O_4$	[M+H]+	445.210	445.212	-0.002	427, 403, 385, 353	Amino acid
13.	12.44	2H,9H-Naphtho[2,1-b]pyrano[3,2-e]pyran-1,11(5H,10H)-dione, 3-(acetyloxy)-3,4,4a,6,6a,12,12a,12b-octahydro-4,4,6a,9,12b-pentamethyl-	$C_{23}H_{32}O_6$	[M+Na]+	425.229	427.209	-1.980	387, 363, 345, 277	Terpenoid
14.	12.51	Esterastin	$C_{28}H_{46}N_2O_6$	[2M+Na]+	1035.420	1035.658	-0.238	489, 434, 377	Steroid
15.	12.63	(S,S)-asperphenamate	$C_{32}H_{30}N_2O_4$	[M+Na]+	529.206	529.210	-0.004	489, 429, 403	Amino acid
16.	12.78	3-Deoxycaryoptinol	$C_{24}H_{34}O_7$	[M+Na]+	455.244	457.220	-1.976	417, 375, 315	Terpenoid
17.	12.95	Monoelaidin	$C_{21}H_{40}O_4$	[M+H]+	357.287	357.299	-0.012	339, 237	Glycerolipid
18.	13.01	3-(2-methoxy-3-(((2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)tetrahydro-2H-pyran-2-yl)oxy)phenyl)propanoic acid	$C_{16}H_{22}O_9$	[M+Na]+	379.266	381.116	-1.850	341, 315, 299	Terpenoid
19.	13.10	(E)-5-(1,2,4a,5-tetramethyl-7-oxo-3,4,8,8a-tetrahydro-2H-naphthalen-1-yl)-3-methylpent-2-enoic acid	$C_{20}H_{30}O_3$	[M+Na]+	339.250	341.209	-1.959	301, 291, 259	Terpenoid
20.	13.27	(1S,4aS,5R)-5-[2-(furan-3-yl)ethyl]-1,4a-dimethyl-6-methylidene-3,4,5,7,8,8a-hexahydro-2H-naphthalene-1-carboxylic acid	$C_{20}H_{28}O_3$	[M-H]-	313.233	315.197	-1.964	297, 287, 271	Terpenoid
21.	13.56	Echinulin	$C_{29}H_{39}N_3O_2$	[M+H]+	462.307	462.313	-0.006	420, 406, 392	Alkaloid
22.	13.63	Celecoxib	$C_{17}H_{14}F_3N_3O_2S$	[M+H]+	381.260	382.083	-0.823	380, 316, 247	Alkaloid
23.	13.75	Lauroyl L-carnitine	$C_{19}H_{37}NO_4$	[M+H]+	342.334	344.278	-1.944	285	Fatty acid ester
24.	13.90	Ethyl 3-(2,4-dihydroxy-6-methylpyridin-3-yl)-3-(3-hydroxyphenyl)propanoate	$C_{17}H_{19}NO_5$	[M+Na]+	341.264	340.120	1.144	300, 290, 274	Pyridine
25.	13.93	5-(4-carboxy-3-methylbutyl)-5,6,8a-trimethyl-3-oxo-4a,6,7,8-tetrahydro-4H-naphthalene-1-carboxylic acid	$C_{20}H_{30}O_5$	[M+Na]+	371.274	373.199	-1.925	333, 323, 305, 291	Terpenoid
26.	14.13	4-(4-Hydroxy-2,6,6-trimethyl-1-cyclohexen-1-yl)-2-butanyl beta-D-glucopyranoside	$C_{19}H_{34}O_7$	[M+Na]+	395.274	397.220	-1.946	357, 345, 333	Terpenoid

Table 1. Contd.

27.	14.35	Stearamide	$C_{18}H_{37}NO$	$[M+H]^+$	283.259	284.294	-1.035	266, 239, 225	Fatty amide
28.	14.45	Methyl (1S,4aS,7aS)-1-(beta-D-glucopyranosyloxy)-7-(hydroxymethyl)-1,4a,5,7a-tetrahydrocyclopenta[c]pyran-4-carboxylate	$C_{17}H_{24}O_{10}$	$[M+Na]^+$	409.288	411.126	-1.838	371, 319	Terpenoid
29.	14.52	Pseudo-anisatin	$C_{15}H_{22}O_6$	$[M+Na]^+$	321.238	321.345	-0.107	281, 271, 269, 253	Terpenoid
30.	14.68	Pseudo-anisatin	$C_{15}H_{22}O_6$	$[M+Na]^+$	319.222	321.131	-1.909	281, 271, 269, 253	Terpenoid
31.	14.87	9-(octanoyloxy)octadecanoic acid	$C_{26}H_{50}O_4$	$[M+H]^+$	429.298	427.380	1.920	409, 367	Fatty acid ester
32.	14.95	Beta-D-Glucopyranoside, 2-phenylethyl 6-O-beta-D-xylopyranosyl-	$C_{19}H_{28}O_{10}$	$[M+Na]^+$	437.283	439.157	-1.874	399. 387, 373	Organooxygen compound
33.	15.14	(2R)-methyl 3-methyl-2-(2-((4-methyl-2-oxo-2H-chromen-7-yl)oxy)acetamido)pentanoate	$C_{19}H_{23}NO_6$	$[M+Na]^+$	383.198	384.140	-0.942	344, 330, 302	Coumarin
34.	15.22	Lysionotin	$C_{18}H_{16}O_7$	$[M+Na]^+$	367.278	367.476	-0.198	327, 313, 301	Flavonoid
35.	15.44	Docosatetra-7Z,10Z,13Z,16Z-enoyl dopamine	$C_{30}H_{45}NO_3$	$[M+H]^+$	469.309	468.350	-0.962	450, 424, 398	Fatty amide
36.	15.77	Caudatin	$C_{28}H_{42}O_7$	$[M+Na]^+$	511.319	513.280	-1.961	489, 361, 343, 325, 307, 283, 259, 243	Steroid
37.	16.25	Conjugated linoleic Acid (10E,12Z)	$C_{18}H_{32}O_2$	$[M+H]^+$	281.246	281.247	-0.001	263, 221	Fatty acid
38.	16.57	Alpha-hydroxydeoxycholic acid	$C_{24}H_{40}O_4$	$[M-H]^-$	393.257	391.290	1.972	373, 347	Steroid
39.	16.68	2-[3,4-bis[[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy]phenyl]-5,7-dihydroxychromen-4-one	$C_{27}H_{30}O_{16}$	$[M+NH4]^+$	628.182	628.187	-0.005	593, 449	Flavonoid
40.	16.93	Emetine	$C_{29}H_{40}N_2O_4$	$[M+H]^+$	481.309	481.311	-0.002	453, 449, 438	Alkaloid
41.	17.34	2-(3,5-dimethyl-7-oxo-7H-furo[3,2-g]chromen-6-yl)-N-(3-(2-oxopyrrolidin-1-yl)propyl)acetamide	$C_{22}H_{24}N_2O_5$	$[M+Na]^+$	419.272	419.384	-0.112	379, 369, 355	Coumarin

protonated molecular ion of m/z 445. Peak 11 and 12 were fragmented and firstly produced ion at m/z 427 by loss of one water molecule. This was followed by the loss of C_2H_2O which resulted in the ion at m/z 403. The next fragmentation was due to the losses of CH_3COOH and C_7H_8 giving ions at m/z 385 and 353 respectively. Initially, it was thought that this could be attributed to two compounds with stereoisomer structures. However, when viewed in the MN, the m/z 467 and 445 masses appeared in two different clusters. From the MN, it can be concluded that peak 11 and peak 12 are two different compounds with different skeletons, although they both share the same MS/MS fragmentation patterns. Peak 15 at t_R 12.63 min was identified as (S,S)-asperphenamate with sodiumated adduct at m/z 529 and protonated adduct at m/z 507. The fragment ions at m/z 489, 429 and 403 were observed due to elimination of H_2O molecule, C_6H_6 and C_7H_4O correspondingly.

Identification of terpenoids

Terpenoids are known to display a wide range of biological activities which include cancer chemo-preventive effects, antimicrobial, antifungal, antiviral, anti-hyperglycemic, anti-inflammatory, anti-parasitic activities and memory enhancers (Kuma et al., 2022). In this study, a total of ten terpenoids were tentatively characterized.

Peak 13 was detected at t_R 12.44 min with sodiumated adduct at m/z 427 displaying protonated adduct at m/z 405. The mass 405 showed MS^2 fragment ions at m/z 387, 363, 345 and 277 corresponding to the losses of H_2O, C_2H_2O, CH_3COOH and $C_6H_8O_3$ accordingly. Peak 16 with t_R 12.78 min was tentatively assigned as 3-deoxycaryoptinol with sodiumated molecular ion at m/z 457 and protonated molecular ion at m/z 435. Besides, MS^2 analysis revealed fragment ions at m/z 417, 375 and 315 suggesting losses of H_2O, CH_3COOH and $C_4H_8O_4$ respectively. Peak18 at t_R 13.01 min with $[M+H]^+$ ion m/z 359 and

Table 2. Putative annotation of lipids identified in ethyl acetate extract of *M. Oleifera* seeds based on LC-MS/MS and MN.

No.	t_R (min)	Lipids	Molecular formula	Adduct	Parent mass (m/z)	Category	Main class	Sub class
1.	11.20	LysoPC(0:0/18:0)	$C_{26}H_{54}NO_7P$	$[M+H]^+$	524.334	Glycerophospholipid	Glycerophosphocholine	Monoacylglycerophospho-choline
2.	11.69	LysoPE(12:0/0:0)	$C_{17}H_{36}NO_7P$	$[M+H]^+$	398.239	Glycerophospholipid	Glycerophosphoethanolamine	Monoacylglycer0phospho-ethanolamine
3.	13.18	LysoPE(0:0/18:3 (6Z,9Z,12Z))	$C_{23}H_{42}NO_7P$	$[M+H]^+$	476.289	Glycerophospholipid	Glycerophosphoethanolamine	Monoacylglycerophospho-ethanolamine
4.	14.28	LysoPG(18:2 (9Z,12Z)/0:0)	$C_{24}H_{45}O_9P$	$[M-H]^-$	507.341	Glycerophospholipid	Glycerophosphoglycerol	MonoacylglyceropHospho-glycerol
5.	14.31	LysoPE(0:0/20:1(11Z))	$C_{25}H_{50}NO_7P$	$[M-H]^-$	506.325	Glycerophospholipid	Glycerophosphoethanolamine	Monoacylglycerophospho-ethanolamine
6.	14.60	PG(22:1(11Z)/12:0)	$C_{40}H_{77}O_{10}P$	$[M-H]^-$	747.508	Glycerophospholipid	Glycerophosphoglycerol	Diacylglycerophospho-glycerol
7.	15.35	LysoPE(0:0/18:1(9Z))	$C_{23}H_{46}NO_7P$	$[M+H]^+$	480.304	Glycerophospholipid	Glycerophosphoethanolamine	Monoacylglycerophospho-choline
8.	15.92	LysoPG(15:0/0:0)	$C_{21}H_{43}O_9P$	$[M+H]^+$	471.326	Glycerophospholipid	Glycerophosphoglycerol	MonoacylglyceroPhosphor-glycerol
9.	16.07	LysoPI(18:1(9Z)/0:0)	$C_{27}H_{51}O_{12}P$	$[M-H]^-$	597.503	Glycerophospholipid	Glycerophosphoinositol	Monoacylglycerophospho-inositol
10.	16.45	LysoPE(0:0/22:2 (13Z,16Z))	$C_{27}H_{52}NO_7P$	$[M+H]^+$	534.374	Glycerophospholipid	Glycerophosphoethanolamine	Monoacylglycerophospho-ethanolamine
11.	16.77	LysoPG(13:0/0:0)	$C_{19}H_{39}O_9P$	$[M-H]^-$	441.317	Glycerophospholipid	Glycerophosphoglycerol	Monoacylglycerophospho-glycerol
12.	16.98	DG(18:3(6Z,9Z,12Z)/18:1(9Z)/0:0)[iso2]	$C_{39}H_{68}O_5$	$[M+NH4]^+$	634.139	Glycerolipid	Diradylglycerol	Diacylglycerol
13.	17.04	LysoPI(18:0/0:0)	$C_{27}H_{53}O_{12}P$	$[M-H]^-$	599.428	Glycerophospholipid	Glycerophosphoinositol	Monoacylglycerophospho-inositol

$[M+Na]^+$ ion m/z 381 was identified based on their fragment ions m/z 341, 315 and 299 due to the losses of H_2O molecule, C_2H_4O and CH_3COOH subsequently. At t_R 13.10 min, peak 19 with $[M+Na]^+$ adduct ion m/z 341 and $[M+H]^+$ adduct ion m/z 319 was observed to lose H_2O, C_2H_4 and CH_3COOH to form fragments m/z 301, 291 and 259 respectively. Peak 20 at t_R 13.27 min shows deprotonated adduct ion at m/z 315 with subsequent losses of H_2O, CO and C_2H_4O to form fragment ions at m/z 297, 287 and 271 correspondingly. Peak 25 at t_R 13.93min with parent ion m/z $[M+Na]^+$ 373 displayed m/z $[M+H]^+$ 351, showing characteristic losses of H_2O, CO, CH_2O_2 and CH_3COOH to produce product ions at m/z 333, 323, 305 and 291 accordingly.

The peak at t_R 14.13 min (peak 26) showed m/z $[M+H]^+$ 375 and m/z $[M+Na]^+$ 397. The m/z at 375 dissociated to give fragment ions at m/z 357, 345 and 333 corresponding to losses of H_2O, CH_2O and C_2H_2O respectively. Peak 28 at t_R 14.45 min

having m/z $[M+Na]^+$ 411 and m/z $[M+H]^+$ 389 reveals MS/MS fragment ions at m/z 371 and 319 due to the losses of H_2O molecule and C_5H_{1O}. Peak 29 and 30 with t_R 14.52 min and 14.68 min were attributed to pseudo-anisatin. Both peaks showed similar parent ions at m/z 321 $[M+Na]^+$ and m/z 299 $[M+H]^+$ displaying fragment ions corresponding to successive losses of H_2O,CO, CH_2O and CH_2O_2 moieties at m/z 281, 271, 269 and 253 respectively. In addition, the m/z 321 and 319 masses appeared in the same cluster in the MN. Thus, these two peaks suggest they are isomers.

Identification of alkaloids

Alkaloids have shown broad-spectrum antimicrobial activities, and several studies have suggested that these compounds could play an important role in tackling pathogenesis of a variety

of infection agents (Cushnie et al., 2008; Casciaro et al., 2020). They are important chemical compounds that serve as a rich reservoir for drug discovery (Lu et al., 2012). UHPLC-MS/MS analysis in positive ionization mode identified five alkaloids, most of which belongs to the isoquinoline type. Dextrorphan with protonated parent ion at m/z 258 was assigned to peak 2 at t_R 3.73 min and displaying fragments ions corresponding to successive losses of H_2O, NH_2CH_3 and C_4H_9 at m/z 240, 227 and 201 respectively. Echinulin having an MS parent ion m/z 462 $[M+H]^+$ at t_R 13.56 min (peak 21) produces ions at m/z 420, 406 and 392 suggesting the losses of C_3H_6, C_4H_8 and C_5H_{10} in a relative manner.

The peak with t_R 13.63 min (peak 22) gives a mass peak of m/z 382 $[M+H]^+$ is attributed to celecoxib. Some characteristics MS fragments were observed (Figure 3). The first fragment occurred from the loss of one H to produce m/z
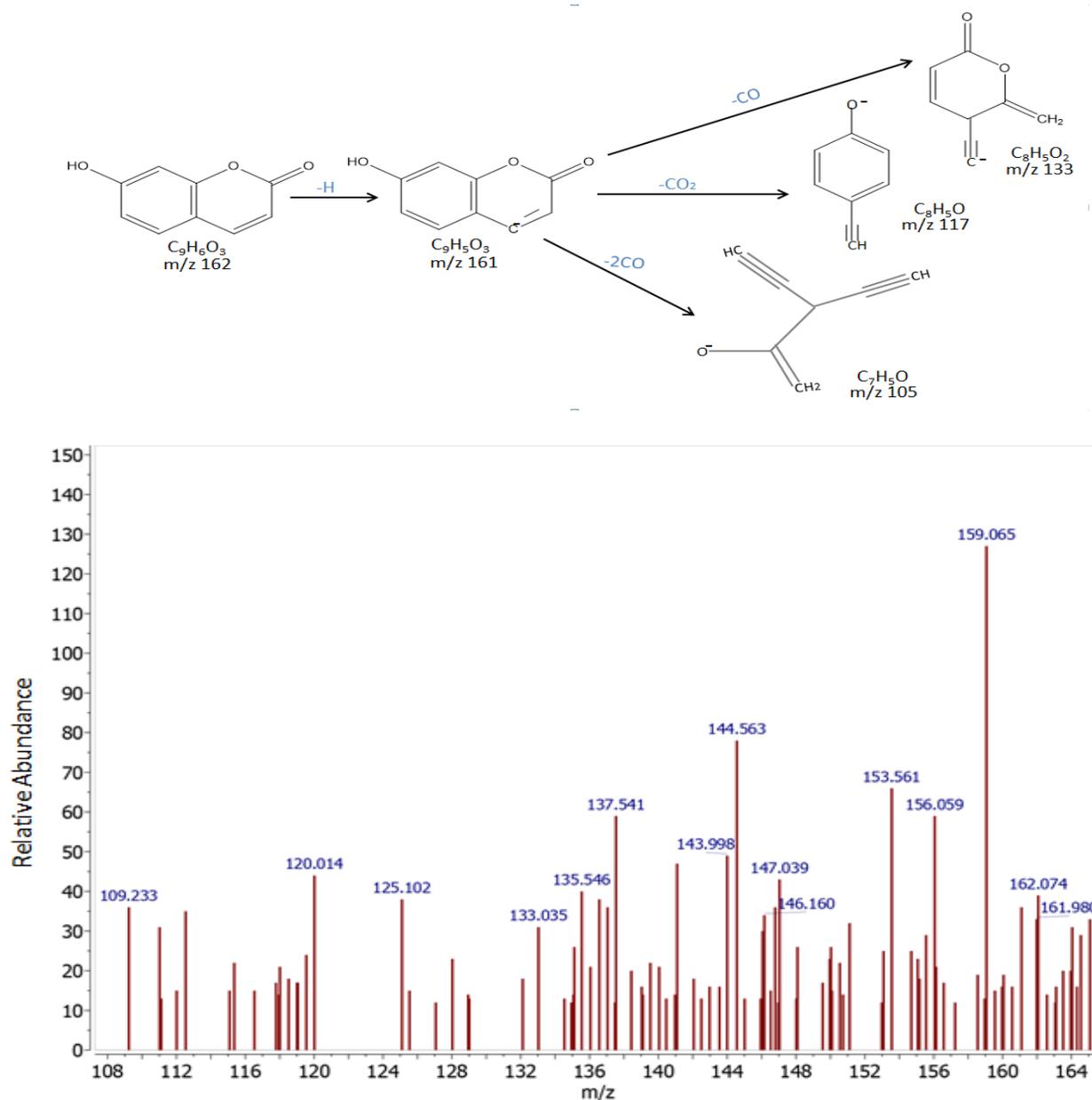
Figure 2. MS/MS spectrum of 7-Hydroxy-2H-chromen-2-one and proposed fragmentation patterns.

380. A second fragment was derived after the loss of SO_2 yielding ion m/z 316. The next fragmentation was due to the loss of CF_3 which resulted in m/z 247. Emetine (peak 40) was identified based on the characteristics fragmentation pattern with parent ion at m/z 481 [M+H]$^+$ which corresponds to its MS base peak at t_R 16.93 min. The ion at m/z 481 was fragmented leading to the formation of ions at m/z 453,449 and 438 by the losses of C_2H_4, CH_3OH and C_3H_7 successively.

Peak 9 at t_R 11.12 min gave [M+H]$^+$ ion at m/z 200 and [M+Na]$^+$ ion at m/z 222. The [M+H]$^+$ ion at m/z 200 was introduced to MS/MS experiment and yielded ions at m/z

183, 171 and 157 signaling losses of NH_3, H_3CN^+ and C_2H_5N respectively. This result agrees with previously published data (Taheri et al., 2016).

Identification of fatty acid/amides/esters and glycerolipid

These are regarded as important components of lipids necessary for cellular processes in humans and have wide range of commercial applications. As depicted in the UHPLC chromatograms of *M. oleifera* seeds ethyl acetate extract, one fatty acid, two fatty amide, two fatty
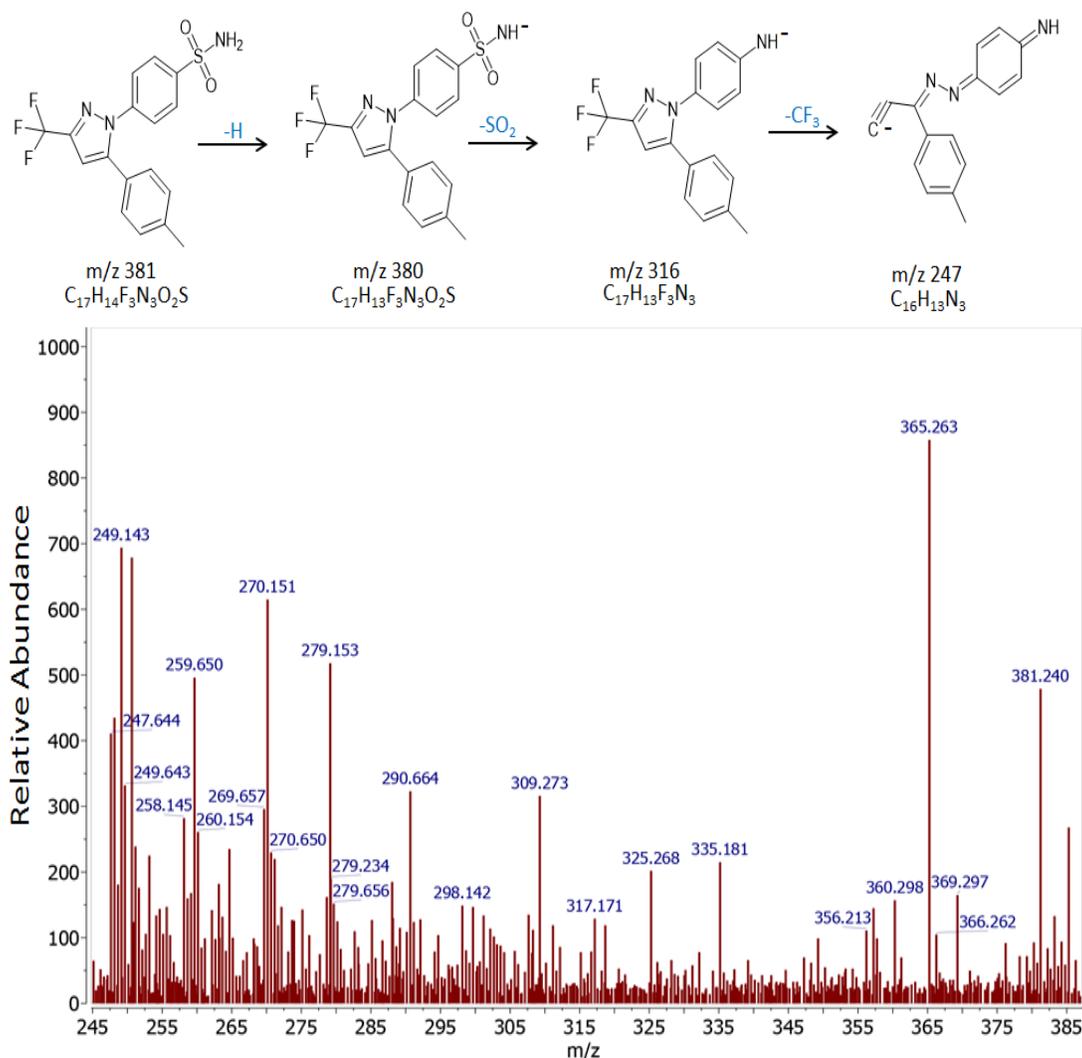
Figure 3. MS/MS spectrum of celecoxib and proposed fragmentation patterns.

acid esters and one glycerolipid were identified.

Fatty acid: Peak 37 with t_R 16.25 min was attributed to conjugated linoleic acid (10E,12Z). This peak showed a parent ion at m/z 281[M+H]$^+$ and characteristic MS/MS fragments ion at m/z 263 due to loss of water molecule from the protonated parent ion and m/z 221 was produced from the loss of acetic acid from the acylated side chain of long-chain fatty acid.

Fatty amides: Peak 27 at t_R 14.35 min with m/z [M+H]$^+$ 284 was tentatively identified as stearamide. The fragment ions at m/z 266, 239 and 225 were due to the losses of H_2O molecule, C_2H_5O and C_3H_6OH respectively. Peak 35 at t_R 15.44 min, having parent ion m/z 468 [M+H]+ exhibited MS/MS product ions at m/z 450, 424 and 398 suggesting losses of H_2O molecule, C_3H_8 and C_5H_{10} successively.

Fatty acid esters: Lauroyl L-carnitine with parent ion m/z [M+H]$^+$ 344 at t_R 13.75 min was assigned to peak 23. The parent ion dissociated to give fragment ion at m/z 285 as a result of loss of C_3H_9N. The characteristics MS/MS fragmentation pattern for peak 31 at t_R 14.87 min with fragment ions m/z 409 and 367 corresponding to the losses of H_2O molecule and CH_3COOH respectively was attributable to 9-(octanoyloxy) octadecanoic acid having a parent ion m/z [M+H]$^+$ 429.

Glycerolipid: Peak 17 at t_R 13.75 min with m/z 357 [M+H]$^+$ was identified as monoeladin based on the fragment ions m/z 339 and 237 due to consecutive losses of H_2O molecule and $C_4H_8O_4$ in a relative manner.

Identification of Organooxygen compounds

Only two organooxygen compounds were identified in the

ethyl acetate extract of *M. oleifera* seeds. Peak 5 was detected at t_R 8.47 min with protonated adduct $[M+H]^+$ at m/z 205. The MS^2 spectrum was characterized by losses of H_2O, C_3H_6 and C_4H_8 moieties to exhibit fragments at m/z 187, 163 and 149 correspondingly. Peak 32 at t_R 14.95 min gave an $[M+Na]^+$ ion at m/z 439 and exhibited an $[M+H]^+$ ion at m/z 417. The m/z 417 ion yielded product ions at m/z 399, 387 and 373 suggesting losses of H_2O, CH_2O and C_2H_4O accordingly.

Identification of Flavonoids

Flavonoids possess a number of medicinal benefits, including anticancer, antioxidant, anti-inflammatory, antiviral properties, neuroprotective and cardio-protective effects (Ullah et al.*,* 2020). Three flavonoids were obtained in the UHPLCMS/MS analysis of *M. oleifera* seeds extract. Peak 6 with t_R 9.84 min showed an $[M+H]^+$ ion at m/z 303. The m/z 303 ion was subjected to MS/MS analysis and gave product ions at m/z 301, 285 and 126 related to the losses of one H, OH and $C_9H_3O_3$ respectively as shown in Figure 4. By referring to mass spectral libraries, peak 6 was tentatively identified as tricetin. Lysionotin at t_R 15.22 min (peak 34) gave an $[M+H]^+$ ion at m/z 345 and an $[M+Na]^+$ ion at m/z 367. The $[M+H]^+$ ion at m/z 345 revealed fragment ion peaks at m/z 327, 313 and 3011 due to the elimination of H_2O, CH_4O and C_2H_4O accordingly. Peak 39 with t_R 16.68 min produced an $[M+H]^+$ ion m/z 611 and an $[M+NH_4]^+$ ion m/z 628. The $[M+H]^+$ ion at m/z 611 displayed fragment ions at m/z 593 and 449 due to loss of H_2O molecule and consequent loss of $C_6H_{10}O_5$.

Identification of benzenoid, dibenzofuran, polyketide and pyridine compounds

These different classes of compounds have various medicinal, agricultural, and industrial applications. Each of benzenoid, dibenzofuran, polyketide and pyridine compounds was produced by *M. oleifera* seeds ethyl acetate extract.

Benzenoid: Peak 8 at t_R 10.49 min with characteristic fragmentation pattern m/z 304, 290 and 278 is attributed to benzyldimethyltetradecylammonium affirming the losses of C_2H_5, C_3H_7 and C_4H_7 accordingly from the protonated precursor ion m/z 333.

Dibenzofuran: Peak 7 at t_R 10.12 min with m/z $[M-H]^-$ 369 was putatively identified as didymic acid. The fragment ions at m/z 351, 337 and 325 corresponding to the losses of H_2O, CH_4O and CO_2 respectively. Didymic acid is a member of dibenzofuran endowed with antimicrobial activity (Dieu et al., 2012).

Polyketide: Nonatic acid (peak 10) at t_R 10.49 min produced $[M+Na]^+$ ion at m/z 225 displaying $[M+H]^+$ ion at

m/z 203. The ion m/z 203 was subjected to MS^2 analysis yielding product ions m/z 185,157 and 129 due to the losses of H_2O molecule, CH_2O_2 and $C_3H_6O_2$ correspondingly.

Pyridine: Peak 24 gave m/z $[M+Na]^+$ 340 and m/z $[M+H]^+$ 318 at t_R 13.90 min with MS^2 spectrum characterized by losses of H_2O molecule, CO and C_2H_4O moieties to exhibit fragments at m/z 300, 290 and 274 successively.

Identification of steroids

Plant steroids are unique class of chemical compound that possess many interesting medicinal, pharmaceutical and agrochemical activities like anti-tumor, immunosuppressive, hepatoprotective, antibacterial, plant growth hormone regulator, sex hormone, antihelminthic, cytotoxic and cardiotonic activity (Patel et al.*,* 2015). In the seed extract, three steroids were putatively identified. Peak 14 with m/z $[2M+Na]^+$ 1035 and m/z $[M+H]^+$ 507 at t_R 12.51 min was annotated as esterastin. The MS/MS spectra of esterastin displayed product ion signals at m/z 489 (loss of H_2O), 434 (loss of $C_3H_5O_2$) and 377 (loss of $C_6H_{10}O_3$).

Caudatin (peak 36) at t_R 15.77 min produced $[M+Na]^+$ ion at m/z 513 displaying $[M+H]^+$ ion at m/z 491. The m/z 489 (A) observed from the loss of one H fragment was selected as precursor ion to perform MS^2 analysis from which the m/z 361 (B) was acquired due to the loss of $C_7H_{12}O_2$. Three additional fragment ions m/z 343 (C), 325 (D) and 307 (E) were also detected which may be ascribed to the sequential elimination of H_2O molecules from (B). To further the investigations, fragmentation patterns of ions B, C and D were chosen as precursor ions in MS^3 analyses to generate ions F, G and H. Ion F at m/z 259 was derived from ion B due to the loss of $C_5H_{10}O_2$. Ion G m/z 243 was generated from ion C suggesting the loss of $C_5H_8O_2$ and ion H m/z 283 was produced through elimination of C_2H_2O from ion D. Ions B-H were identified as key fragment ions for caudatin (Figure 5).

Another steroid was identified as alpha-hydroxydeoxycholic acid (peak 38) at t_R 16.57 min. The MS/MS spectrum of the peak exhibited fragment ions at m/z 373 and 347 affirming the losses of water and carbon dioxide respectively from the deprotonated precursor ion m/z 391. The fragmentation results agree with previously published data (Chen et al.*,* 2015)

Identification of lipids

Lipids are crucial components of cellular membranes and lipid particles such as lipoproteins, which play many essential roles in cellular functions, including cellular barriers, membrane matrices, signaling, and energy depots (Yang and Han, 2016). Lipids are also valuable energy rich compounds that have the potential to replace
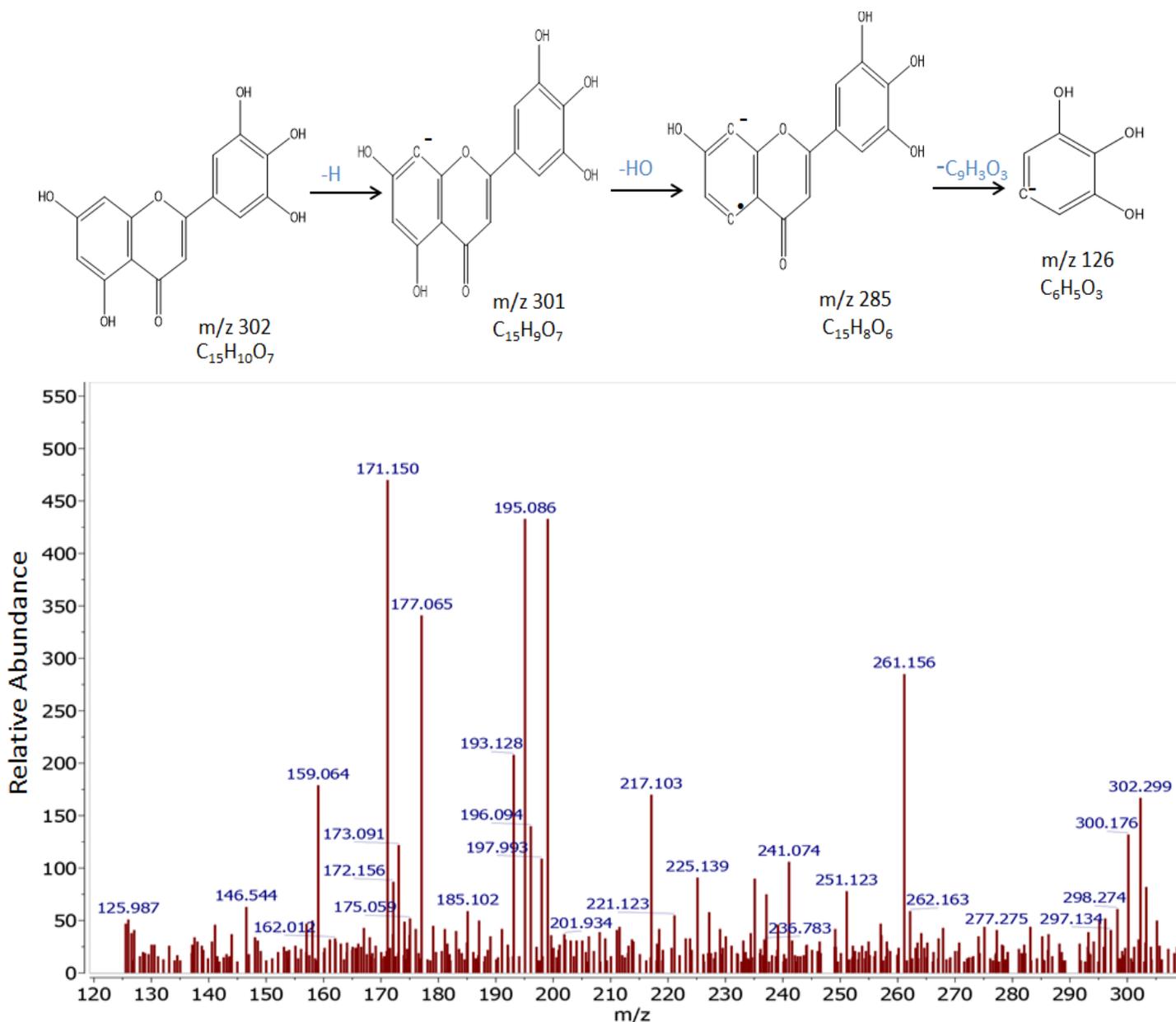
Figure 4. MS/MS spectrum of tricetin and proposed fragmentation patterns

conventional fossil fuels through the production of biofuels (Chew et al., 2018). Studies by Takase et al. (2022) indicate that biodiesel made from *Moringa oleifera* seed oil has a stronger oxidative ability, high cloud point, and a cetane number of around 67, which is higher than most biodiesels. In the current study, the typical nature of the lipid content that makes up the important property of *M. oleifera* seeds as a suitable candidate for nutrition and biofuel production was revealed through putative annotation of different lipids using MN as shown in Figure 6. The putative identified metabolites matched with the GNPS' spectral libraries and were also identified using different external data bases, namely LIPID MAPS,

different external data bases, namely LIPID MAPS, HMDB, Metanetx and Swisslipids.

The major families identified in the network were the glycerophospholipids, including five glycerophosphoethanolamines (LysoPE), three glycerophosphoglycerols (LysoPG) with one (PG), two glycerophosphoinositols (LysoPI) and one glycerophosphocholine (LysoPC), whereas the only glycerolipid identified is the diradylglycerol (DG) as referred to in Table 2. The glycerophospholipids are derivatives of sn-glycero-3-phosphoric acid. They contain an O-acyl or O-alkyl or O-alk-1′-enyl residue at the sn-1 position and an O-acyl residue at the sn-2 position of the
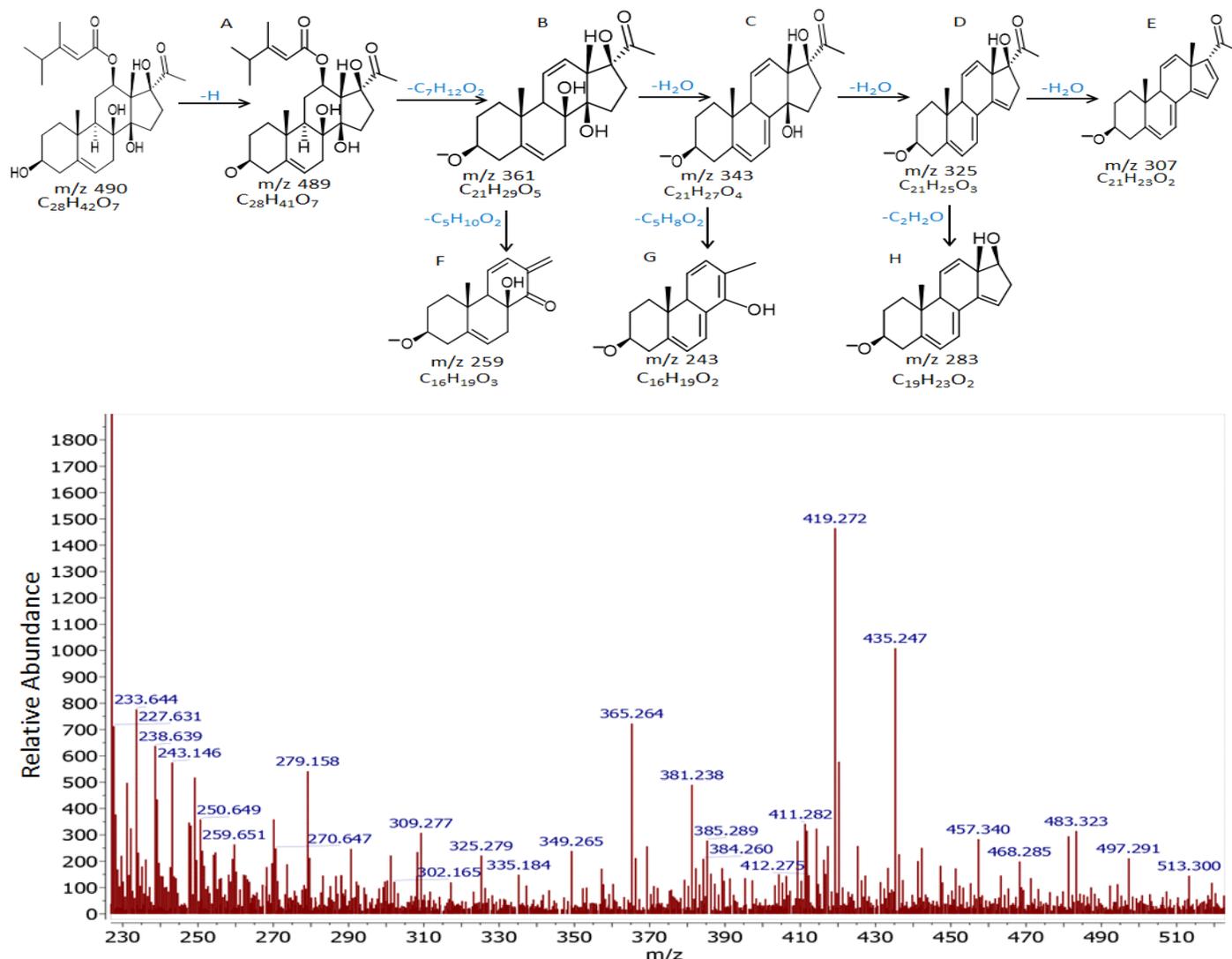
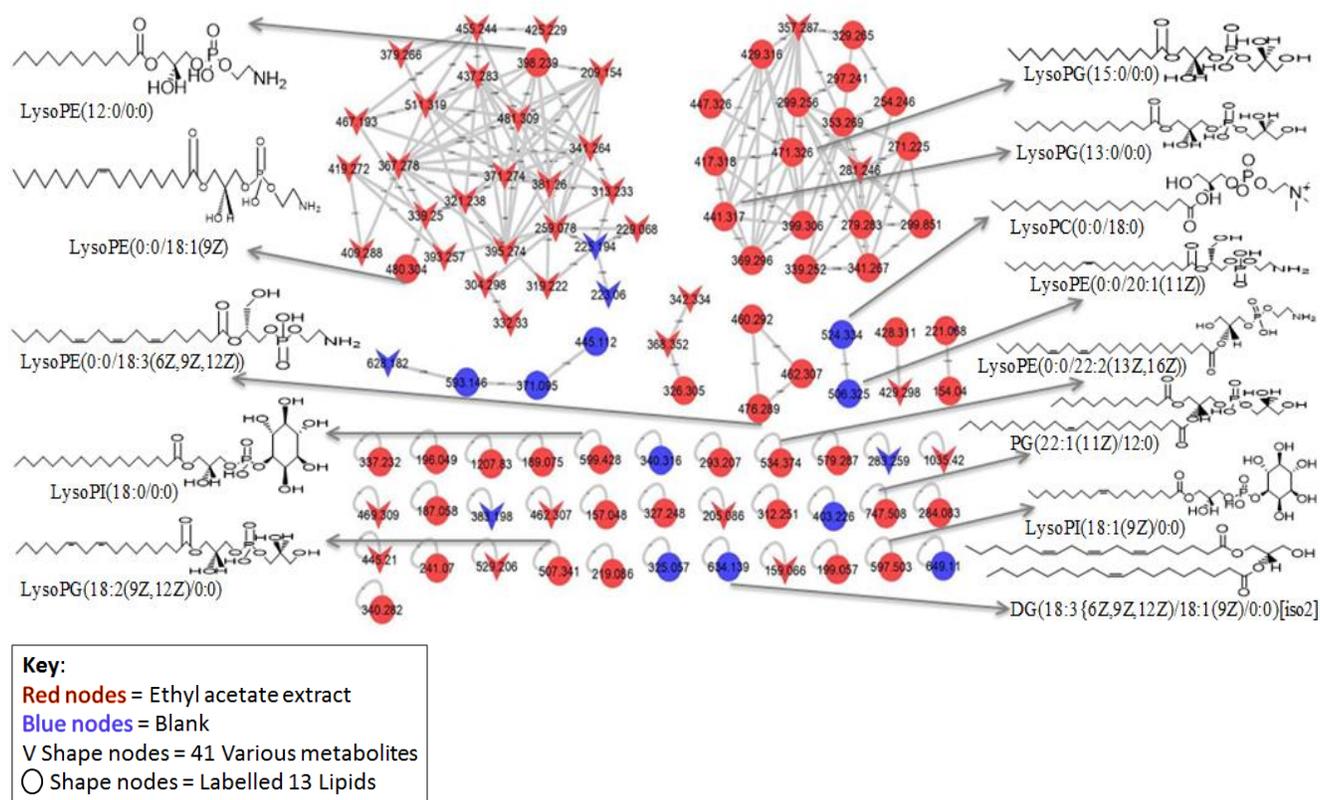**Figure 5.** MS/MS spectrum of caudatin and proposed fragmentation patterns.

glycerol moiety and are defined on the basis of the substituents on the phosphoric acid at the sn-3 position (Farooqui et al., 2000). Glycerophosphocholines (LysoPC) is a monoglycerophospholipid in which the glycerol is esterified with a fatty acid at O-1 position, and linked at position 3 to a phosphocholine. It consists of one chain of stearic acid at the C-1 position.

Glycerophosphoethanolamines (LysoPE) are glycerophospholipids where the glycerol is esterified with a fatty acid at O-2 position, and linked at position 3 to a phosphoethanolamine. They are known as lysophospholipid, which refers to any phospholipid that is lacking one of its two O-acyl chains. Glycerophosphoglycerol (LysoPG) are glycerophospholipids, at which only one fatty acid is bonded to the 1-glycerol moiety through an ester linkage. They consist of one chain of saturated fatty acid at the C-

1(sn-1) position. Unlike (LysoPG), the (PG) contained two fatty acids bonded to the 1-glycerol moiety through ester linkages. They can have many different combinations offatty acids of varying lengths and saturation attached to the C-1 and C-2 positions.

Glycerophosphoinositols (LysoPI) are glycerophospholipids where the glycerol is acylated only at position O-1 with a fatty acid. They can have different combinations of fatty acids of varying lengths and saturation attached at the C-1 (sn-1) or C-2 (sn-2) position. The glycerolipids are a class of lipids containing glycerol to which long-chain hydrocarbons are attached to the hydroxyl groups via carboxylic acid ester linkages (Dowhan et al., 2016).

Diradylglycerol (DG), consisting of two fatty acid chains one chain of linolenic acid at the C-1 position and one chain of oleic acid at the C-2 position covalently bonded

Key:
Red nodes = Ethyl acetate extract
Blue nodes = Blank
V Shape nodes = 41 Various metabolites
◯ Shape nodes = Labelled 13 Lipids

Figure 6. Full molecular networking of *M. oleifera* seeds ethyl acetate extract based on tandem mass (MS/MS) spectrometry data in the positive ionization mode identifying 54 various metabolites. Structures shown are representative examples of the lipids identified. Nodes are labeled with parent m/z values and edges are labeled with cosine scores from 0 to 1.

to a glycerol molecule through ester linkages without phosphorylethanolamine moiety. This finding has proven the existence of several monoacylglycerols and diacylglycerols in the lipophilic extracts of *M. oleifera* seeds obtained from Gas chromatography–mass spectrometry (GC-MS) fatty acid analysis.

Moreover, the seeds harvested in the rainy season, favor the higher content of mono- and diacylglycerides when compared with the seeds harvested during the dry season (Flávia et al., 2022). In another study, fourier transform infrared spectroscopy (FT-IR) and GC-MS analyses were used to characterize the biodiesel in order to investigate the quality and corresponding fatty acid methyl ester (FAMEs) composition in *Moringa* seed oil, to explore the biodiesel potential (Ruslan et al., 2021).

Molecular networking (MN) facilitates data mining via the clustering of the MS/MS spectra based on fragmentation cosine similarities (Esposito et al., 2017). Figure 7a presents the MS/MS spectrum of LysoPG (15:0/0:0) depicting common fragments found in MS/MS spectrum of LysoPG(13:0/0:0) which appeared in the same cluster. Both shared several fragments such as m/z 229 derived from glycerol moiety and phosphate components and m/z 214 established as cleavage of saturated hydrocarbon chain at its glycerol ester oxygen.

Meanwhile, Figure 7b displayed the represented MS/MS spectrum of LysoPE(0:0/18:1(9Z)) sharing other common fragments with MS/MS spectrum of LysoPE (12:0/0:0) located in the same cluster like m/z 242 forming as glycerol and phosphate backbone free from its ethanolamine moiety and another fragment of unsaturated hydrocarbon chain-forming at m/z 209.

Lastly, some glycerophospholipid which were totally absent in their class cluster due to their slightly different structures were seen in the resulting MN sharing common fragments. The MS/MS spectrums for all metabolites are provided as Supplementary Figure S2).

Conclusion

A novel strategy using UHPLC-QTOF/MS data acquisition combined with the MN was adopted to characterize a large set of metabolites with a wide range of classes in *Moringa oleifera* seeds. In this research, an efficient exploitation of datasets was employed for automated data treatment and access to dedicated fragmentation databases during MN. The MS/MS-based molecular networking approach, which had never been done on *M oleifera* seeds has succeeded in the discovery of a

(7a)

(7b)
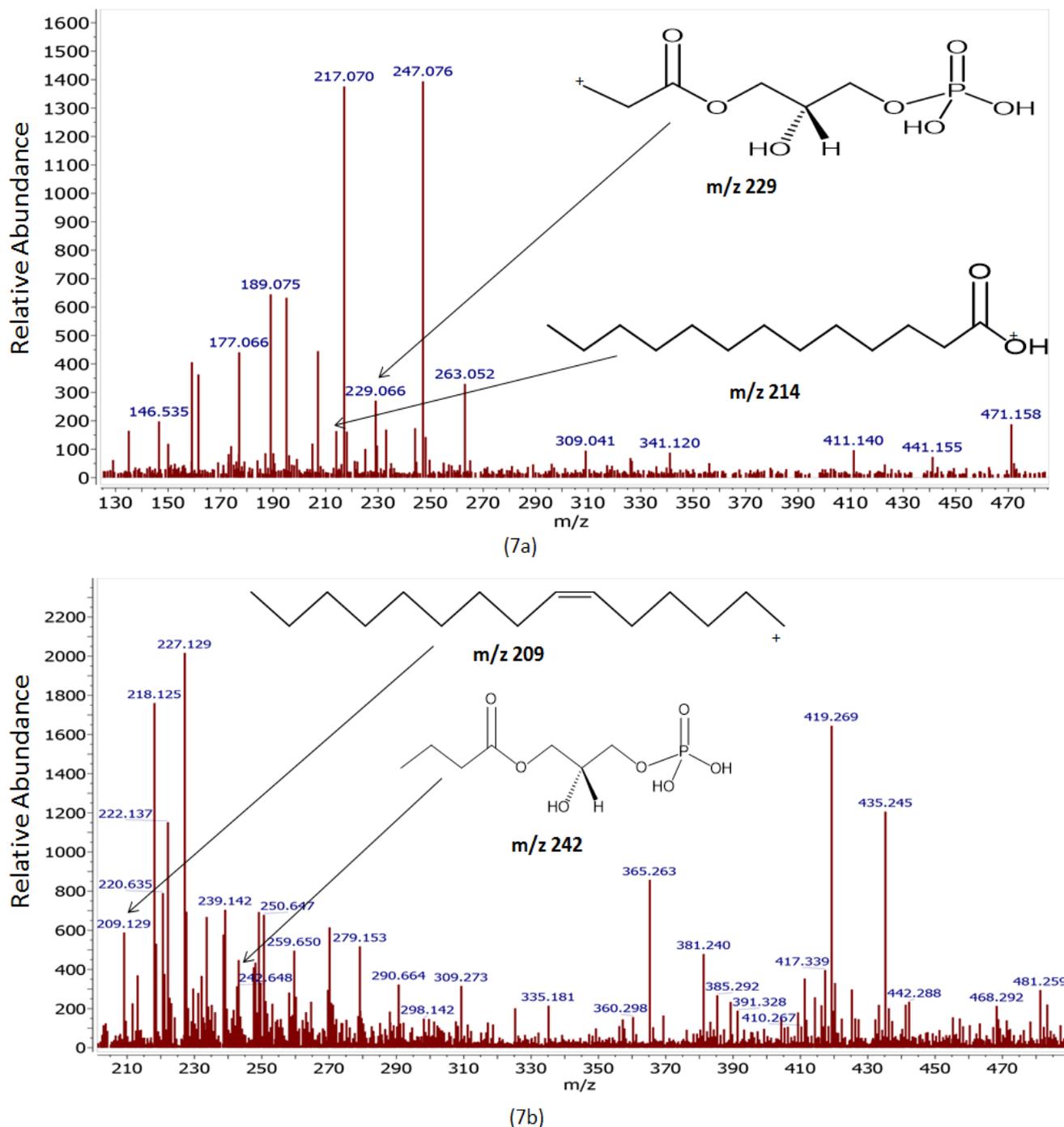
Figure 7. Represent MS/MS spectra showing common fragments shared among (a) LysoPG(15:0/0:0) and LysoPG(13:0/0:0); (b) LysoPE(0:0/18:1(9Z)) and LysoPE(12:0/0:0) in the same cluster in the MN acquired by UHPLC–MS/MS in positive mode.

total of 54 metabolites belonging to different bioactive phytochemical classes including coumarins, alkaloids, amino acids, flavonoids, terpenoids, fatty acids, steroids and lipids among others. The results of the study confirms the therapeutic potency of the *M. oleifera* seeds which could serve as potential biomarker for new drug discovery, and also has a wide application in food

industry.

In addition, the presence of saturated and unsaturated fatty acids suggests its benefit for biodiesel production. However, for further studies, fatty acids can be converted into (FAMEs) which could be produced more efficiently by ex-situ transesterification of lipid extracted from *M. oleifera* seeds. By optimizing the transesterification

reaction involving studying the effect of different reaction conditions, catalysts, and process parameters on the biodiesel production the yield and quality of biodiesel can be improved.

CONFLICT OF INTERESTS

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

AUTHOR'S CONTRIBUTIONS

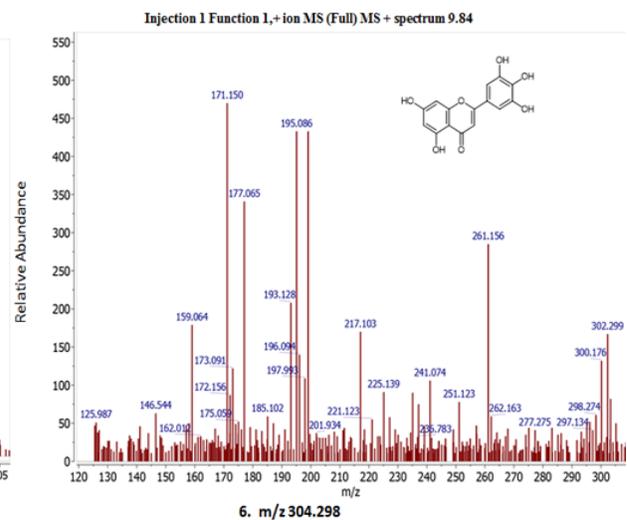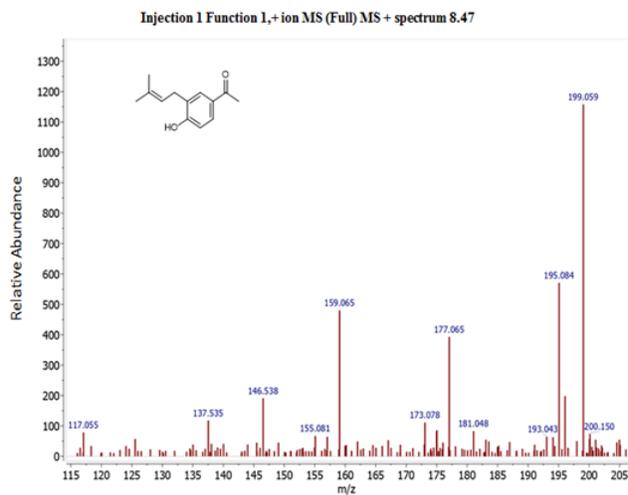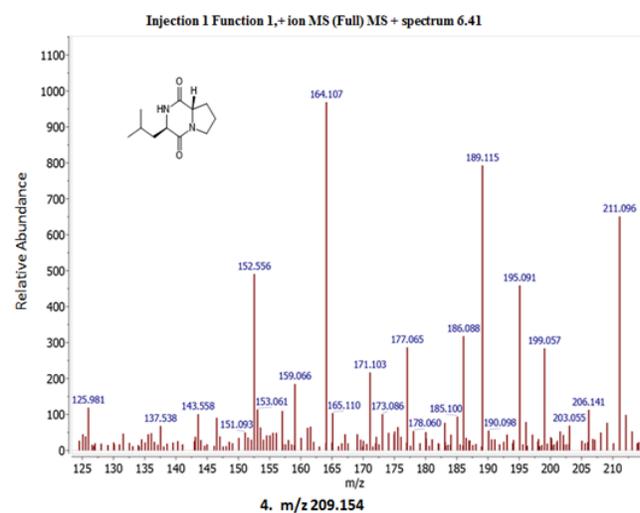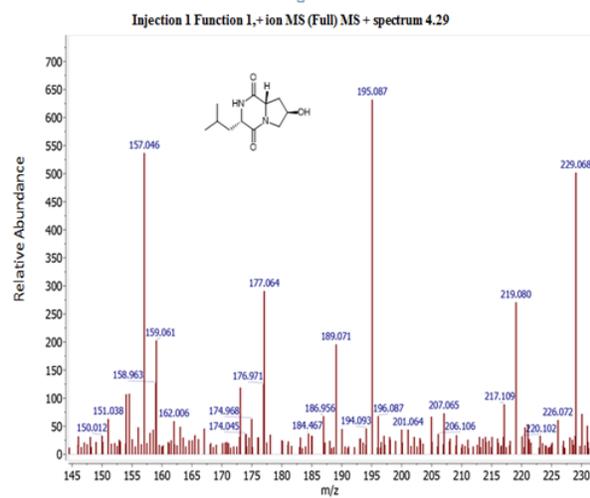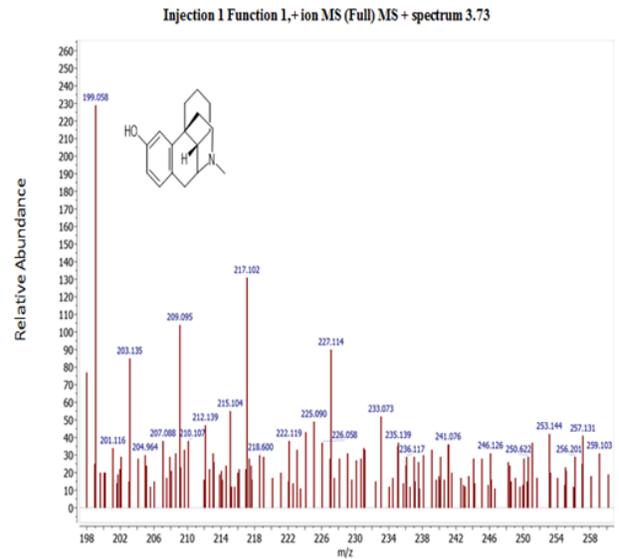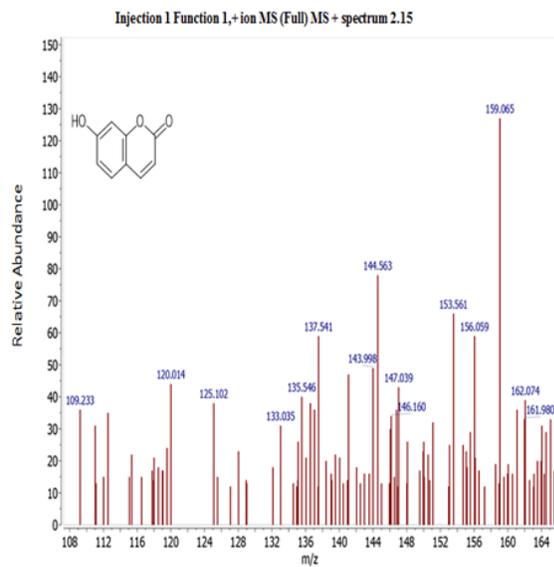Conceptualization of the study was done by Istifanus Y. C. and Famurewa O. J., project was supervised by Istifanus Y. C. and Auwal A. M., validation and visualization were done by Famurewa O. J., manuscript was written, edited and revised by Famurewa O. J., and data curation and formal analysis were done by Famurewa O. J.
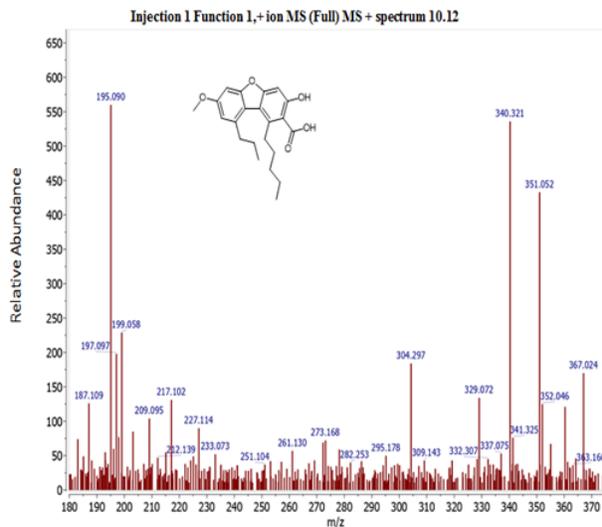
REFERENCES

Beaulieu ED, Ionescu M, Chatterjee S, Yokota K, Trauner D. Lindow S (2013). Characterization of a Diffusible Signaling Factor from *Xylella fastidiosa* mBio 4(1):e00539–e00512. doi: 10.1128/mBio.00539-12. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

Casciaro B, Mangiardi L, Cappiello F, Romeo I, Loffredo MR, Iazzetti A., Quaglio D (2020). Naturally-occurring alkaloids of plant origin as potential antimicrobials against antibiotic-resistant infections. Molecules 25(16):3619.

Chen XL, Geng CA, Chen JJ (2015). A Fragmentation Study of Six C21 Steroidal Aglycones by Electrospray Ionization Ion-Trap Time-of-Flight Tandem Mass Spectrometry. Natural Product Communications (12):2053-8. PMID: 26882663.

Chew KW, Chia SR, Show PL, Ling TC, Chang JS (2018). Biofuels from Microbial Lipids. In: Liao, Q., Chang, JS., Herrmann, C., Xia, A. (eds) Bioreactors for Microbial Biomass and Energy Conversion. Green Energy and Technology. Springer, Singapore. Available at: https://doi.org/10.1007/978-981-10-7677-0-9

Christian KO, Dzuvor SP, Charles AP, Amuzu CA, Francis (2022). Bioactive components from *Moringa oleifera* seeds: production, functionalities and applications – a critical review. Critical Reviews in Biotechnology 42(2):271-293. DOI:10.1080/07388551.2021.1931804

Colnaghi SAV, da Silva DS, Lambais MR, Carrilho E (2007). Characterization of a putative *Xylella fastidiosa* diffusible signal factor by HRGC-EI-MS. Journal of Mass Spectrometry 42(4):490-496. doi: 10.1002/jms.1181.

Cushnie TPT, Taylor PW, Nagaoka Y, Uesato S, Hara Y, Lamb AJ (2008). Investigation of the Antibacterial Activity of 3-O-Octanoyl-(-)-Epicatechin. Journal of Applied Microbiology 105(5):1461-1469. doi:10.1111/j.1365-2672.2008.03881.x CrossRef Full Text | Google Scholar

De Oliveira GG, Carnevale NF, Demarque DP, De Sousa P-JJA, Filho SPRC, De Melo SJ, Guedes da Silva Almeida JR, Callegari Lopes JL, Lopes NP (2017). Dereplication of Flavonoid Glycoconjugates from *Adenocalymma imperatoris-maximilianii* by untargeted tandem mass spectrometry-based molecular networking. Planta Medica 83(07):636-646.

Dieu A, Millot M, Champavier Y, Chulia JA, Vergnaud J, Chaleix V, Bressollier P, Sol V, Gloaguen V (2012). Antimicrobial activity of *Cladonia incrassata* acetone extract. Planta Medica 78(11):PI461.

Di Masi S, De Benedetto GE, Malitesta C, Saponari M, Citti C, Cannazza G, Ciccarella G (2022). HPLC-MS/MS method applied to an untargeted metabolomics approach for the diagnosis of "olive quick decline syndrome. Analytical and Bioanalytical Chemistry 414(1):465-473. doi: 10.1007/s00216-021-03279-7. Epub 2021 Mar 25. PMID: 33765220; PMCID: PMC8748322.

Dowhan W (2016). Synthesis and Structure of Glycerolipids, Encyclopedia of Cell Biology, Academic Press, Pages 160-172, ISBN 9780123947963. Available at: https://doi.org/10.1016/B978-0-12-394447-4.10020-3.

Duan JY, Li ZY, Li J, Song HB, Fu JS, Xie BG (2020). Comparison of nutritional and flavor characteristics between four edible fungi and four fruits and vegetables based on components and characteristics of free amino acids. Mycosystema 39(6):1077-1089. doi: 10.13346/j.mycosystema.20062

El-Haddad AE, El-Deeb EM, Koheil MA, El-Khalik SMA, El-Hefnawy HM (2019). Nitrogenous phytoconstituents of genus Moringa: spectrophotometrical and pharmacological characteristics. Medicinal Chemistry Research 28:1591-1600. https://doi.org/10.1007/s00044-019-02403-8.

Elsayed EA, Sharaf-Eldin MA, Wadaan M (2015). In vitro evaluation of cytotoxic activities of essential oil from Moringa oleifera seeds on HeLa, HepG2, MCF-7, CACO-2 and L929 cell lines. Asian Pacific Journal of Cancer Prevention 16(11):4671-4675.

Esposito M, Nothias LF, Retailleau P,Costa J, Roussi F, Neyts J, Leyssen P, Touboul D, Litaudon M, Paolini J (2017). Isolation of premyrsinane, myrsinane, and tigliane diterpenoids from *euphorbia pithyusa* using a chikungunya virus cell based assay and analogue annotation by molecular networking. Journal of Natural Products 80(7):2051-2059.

Farooqui A, Horrocks L, Farooqui T (2000). Glycerophospholipids in brain: Their metabolism, incorporation into membranes, functions, and involvement in neurological disorders. Chemistry and Physics of Lipids 106(1):1-29.

Flávia MSW, Alessandro FS, Lavínia KBS, Lays CA, Lisiane SF, Alvaro SL, Alini TF, Cláudio D, Cleide MFS (2022). Influence of seasonality on the physicochemical properties of *Moringa oleifera Lam.* Seed oil and their oleochemical potential, Food Chemistry: Molecular Sciences 4:100068. Available at: https://doi.org/10.1016/j.fochms.2021.100068.

Galuppo M, Nicola GR, Iori R, Dell'utri P, Bramanti P, Mazzon E (2013). Antibacterial activity of glucomoringin bioactivated with myrosinase against two important pathogens affecting the health of long-term patients in hospitals. Molecules 18(11):14340-14348.

Goulitquer S, Potin P, Tonon T (2012). Mass spectrometry-based metabolomics to elucidate functions in marine organisms and ecosystems. Marine Drugs10(4):849-880.

Jeyaseelan EC, Jenothiny S, Pathmanathan MK, Jeyadevan JP (2012). Antibacterial activity of sequentially extracted organic solvent extracts of fruits, flowers and leaves of *Lawsonia inermis L.* from Jaffna Asian Pacific Journal of Tropical Biomedicine 2(10):798-802. doi: 10.1016/S2221-1691(12)60232-9. PMID: 23569850; PMCID: PMC3609233.

Liang YX, Zhang L, Gao FH, Zhang XM, Li X, Zhang HH (2019). Changes in NaCl, reducing sugar and amino acids and formation of tastes in Chongqing Shuidouchi during fermentation. Food Fermentation Food and Fermentation Industries 45(18):27-34.

Lindow S, Newman K, Chatterjee S, Baccari C, Lavarone AT, Ionescu M (2014). Production of *Xylella fastidiosa* diffusible signal factor in transgenic grape causes pathogen confusion and reduction in severity of Pierce's disease. Molecular Plant-Microbe Interactions 27(3):244-254. doi: 10.1094/mpmi-07-13-0197-fi.

Lu JJ, Bao JL, Chen XP, Huang M, Wang YT (2012). Alkaloids isolated from natural herbs as the anticancer agents. Evidence-based complementary and alternative medicine.

Lu W, Su X, Klein MS, Lewis IA, Fiehn O, Rabinowitz JD (2017). Metabolite measurement: pitfalls to avoid and practices to follow. Annual Review of Biochemistry 86:277-304.

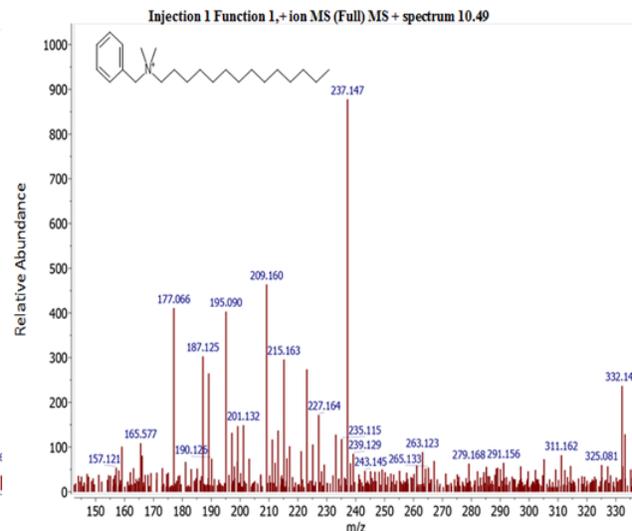Matt B (2022). Integrating LC-MS/MS based molecular networking and

analysis of medicinal plants. Rhode Island IDeA Network of Biomedical Research Excellence.

Messaili S, Colas C, Fougère L, Destandau E (2020). Combination of molecular network and centrifugal partition chromatogrpahy fractionation for targeting and identifying *Artemisia annua L.*antioxidant compounds. Journal of Chromatography A 1615:460785.

Nothias LF, Petras D, Schmid R, Dührkop K, Rainer J, Sarvepalli A (2020). Feature-based Molecular Networking in the GNPS Analysis Environment. Feature-based molecular networking in the GNPS analysis environment. Nature Methods 17(9):905-908.

Patel S, Savjani J (2015). Systematic review of plant steroids as potential anti-inflammatory agents: Current status and future perspectives. The Journal of Phytopharmacology 4(2):121-125. 10.31254/phyto.2015.4212.

Piasecka A, Kachlicki P, Stobiecki M (2019). Analytical Methods for Detection of Plant Metabolomes Changes in Response to Biotic and Abiotic Stresses. International Journal of Molecular Sciences 20(2):379. doi: 10.3390/ijms20020379. PMID: 30658398; PMCID: PMC6358739.

Quinn RA, Nothias LF, Vining O, Meehan M, Esquenazi E, Dorrestein PC (2017). Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. Trends in Pharmacological Sciences 38(2):143-154. doi: 10.1016/j.tips.2016.10.01

Rashid U, Anwar F, Moser BR (2008). *Moringa oleifera* oil: A possible source of Biodiesel. Bioresource Technology 99(17):8175-8179.

Ruslan HJ, Sosidi H, Ridhay A, Agus AY (2021). Production of biodiesel based on moringa seed oil at varied reaction times by using CaO catalyst from chicken eggshells. Journal of Physics: Conference Series 1763(1):012054.

Saa RW, Fombang EN, Ndjantou EB, Njintang, NY (2019). Treatments and uses of *Moringa oleifera* seeds in human nutrition: A review. Food Science and Nutrition 7(6):1911-1919. https://doi.org/10.1002/fsn3.1057

Sawada Y, Hirai M (2013). Integrated LC-MS/MS system for plant metabolomics. Computational and Structural Biotechnology Journal e201301011. 10.5936/csbj.201301011.

Shimizu T, Watanabe M, Fernie AR, Tohge T (2018). Targeted LC-MS Analysis for Plant Secondary Metabolites. Plant Metabolomics: Methods and Protocols, pp. 171-181. Available at: https://doi.org/10.1007/978-1-4939-7819-9_12

Taheri M, Moazeni-Pourasil RS, Sheikh-Olia-Lavasani M, Karami A, Ghassempour A (2016). Central composite design with the help of multivariate curve resolution in loadability optimization of RP-HPLC to scale-up a binary mixture. Journal of Separation Science 39(6):1031-1040. doi: 10.1002/jssc.201500526. Epub. PMID: 26768981.

Takase M, Essandoh PK, Asare RK, Nazir KH (2022). The Value Chain of *Moringa oleifera* Plant and the Process of Producing Its Biodiesel in Ghana. Scientific World Journal 1827514. doi: 10.1155/2022/1827514. PMID: 35898285; PMCID: PMC9314179.

Ullah A, Munir S, Badshah SL, Khan N, Ghani L, Poulson BG, Emwas AH, Jaremko M (2020). Important Flavonoids and Their Role as a Therapeutic Agent. Molecules 25(22):5243. doi: 10.3390/molecules25225243. PMID: 33187049; PMCID: PMC7697716.

Venugopala KN, Rashmi V, Odhav B (2013). "Review on Natural Coumarin Lead Compounds for Their Pharmacological Activity", BioMed Research International. Article ID 963248, 14 pages. Available at: https://doi.org/10.1155/2013/963248

Vincenti F, Montesano C, Di Ottavio F, Gregori A, Compagnone D, Sergi M, Dorrestein P (2020). Molecular Networking: A Useful Tool for the Identification of New Psychoactive Substances in Seizures by LC–HRMS. Frontiers in Chemistry 8:572952, https://www.frontiersin.org/articles/10.3389/fchem.2020.572952, doi:10.3389/fchem.2020.572952.

Wang M, Carver JJ, Phelan VV, Sanchez LM, Garg N, Peng Y, Nguyen DD, Watrous J, Kapono CA, Luzzatto-Knaan T (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. Nature Biotechnology 34(8):828-837. doi: 10.1038/nbt.3597.

Wolfender JL, Nuzillard JM, Vander Hooft JJJ, Renault JH, Bertrand S (2018). Accelerating Metabolite Identification in Natural Product Research: Toward an Ideal Combination of LiquidChromatography–High-Resolution Tandem Mass Spectrometry and NMR Profiling, in SilicoDatabases, and Chemometrics. Analytical Chemistry 91(1):704-742. [CrossRef]

Wu S, Zhou J, Xin Y, Xue S (2015). Nutritional stress effects under different nitrogen sources on the genes in microalga *Isochrysis zhangjiangensis* and the assistance of *Alteromonas macleodii* in releasing the stress of amino acid deficiency. Journal of Phycology 51(5):885-895.

Yang JY, Sanchez LM, Rath C M, Liu X, Boudreau PD, Bruns N (2013). Molecular networking as a dereplication strategy. Journal of Natural Products 76(9):1686-1699. doi: 10.1021/np400413

Yang K, Han X (2016). Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences. Trends in Biochemical Sciences 41(11):954-969. doi: 10.1016/j.tibs.2016.08.010.

Zeki OC, Eylem CC, Reçber T, Kır S, Nemutlu E (2020). Integration of GC-MS and LC-MS for untargeted metabolomics profiling. Journal of Pharmaceutical and Biomedical Analysis 190:113509. doi: 10.1016/j.jpba.2020.113509.

Zhou L, Li J, Yan C (2018). Simultaneous determination of three flavonoids and one coumarin by LC-MS/MS: Application to a comparative pharmacokinetic study in normal and arthritic rats after oral administration of *Daphne genkwa* extract. Biomedical Chromatography 32(7):e4233. doi: 10.1002/bmc.4233. Epub Apr 26. PMID: 29500935.

SUPPLEMENTARY MATERIALS

1. m/z 159.066

2. m/z 259.078

3. m/z 229.068

4. m/z 209.154

5. m/z 205.086

6. m/z 304.298

7. m/z 368.352

8. m/z 332.33
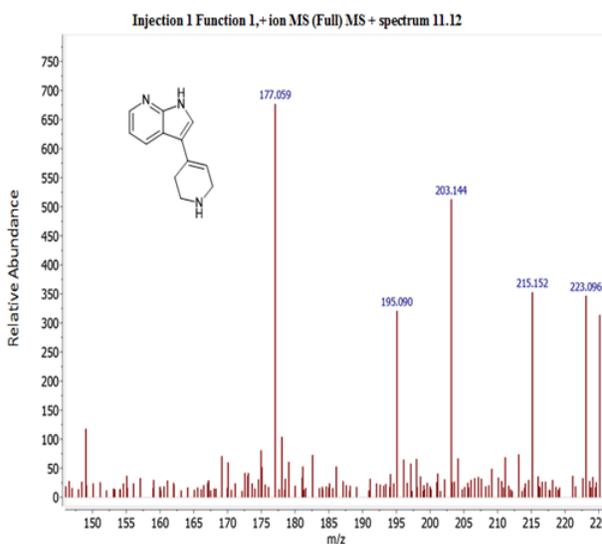
9. m/z 223.06
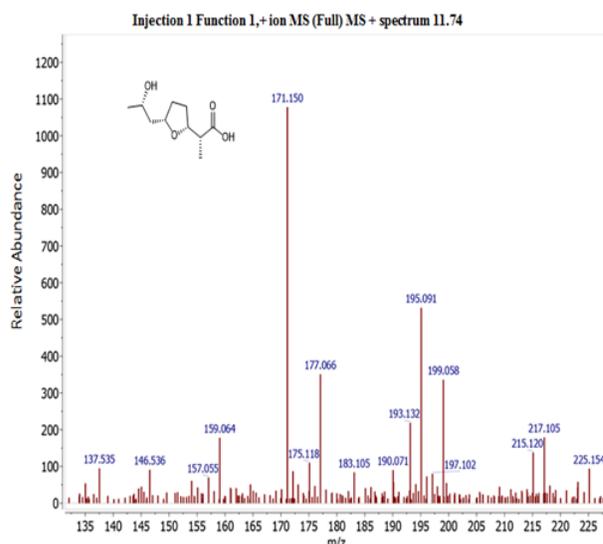
10. m/z 225.194

11. m/z 467.193

12. m/z 445.21

Injection 1 Function 1,+ ion MS (Full) MS + spectrum12.44

13. m/z 425.229

Injection 1 Function 1,+ ion MS (Full) MS + spectrum12.51

14. m/z 1035.42

Injection 1 Function 1,+ ion MS (Full) MS + spectrum12.63

15. m/z 529.206

Injection 1 Function 1,+ ion MS (Full) MS + spectrum12.78

16. m/z 455.244

Injection 1 Function 1,+ ion MS (Full) MS + spectrum12.95

17. m/z 357.287

Injection 1 Function 1,+ ion MS (Full) MS + spectrum13.01

18. m/z 379.266

19. m/z 339.25

20. m/z 313.233

21. m/z 462.307

22. m/z 381.26

23. m/z 342.334

24. m/z 341.264

Injection 1 Function 1,+ ion MS (Full) MS + spectrum13.93

25. m/z 371.274

Injection 1 Function 1,+ ion MS (Full) MS + spectrum14.13

26. m/z 395.274

Injection 1 Function 1,+ ion MS (Full) MS + spectrum14.35

27. m/z 283.259

Injection 1 Function 1,+ ion MS (Full) MS + spectrum14.45

28. m/z 409.288

Injection 1 Function 1,+ ion MS (Full) MS + spectrum14.52

29. m/z 321.238

Injection 1 Function 1,+ ion MS (Full) MS + spectrum14.68

30. m/z 319.222

31. m/z 429.298

32. m/z 437.283

33. m/z 383.198

34. m/z 367.278

35. m/z 469.309

36. m/z 511.319

Figure S1. MS/MS spectra and structures of various metabolites identified in ethyl acetate extract of *M. oleifera* seeds in positive ion mode.

1. m/z 524.334

2. m/z 398.239

3. m/z 476.289

4. m/z 507.341

5. m/z 506.325

6. m/z 747.503

7. m/z 480.304

8. m/z 471.326

9. m/z 597.503

10. m/z 534.374

11. m/z 441.317

12. m/z 634.189

Figure S2. MS/MS spectra and structures of identified lipids in ethyl acetate extract of *M. oleifera* seeds in positive ion mode.