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.

Received 18 July, 2023; Accepted 14 August, 2023

*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 (Wollender et al., 2018), the application of LC–MS/MS-based metabolic profiling of biological systems...
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; Piasaecka 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, anti-diabetic, antioxidant, anti-hypertensive, 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
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 H2O + 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 MS² 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 (tR) 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.

identification of coumarins

Coumarins containing the unique 2H-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, 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$_2$O molecule. The ion at m/z 344 was further fragmented by losses of CH$_3$O and CH$_2$COOH 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$_2$O, CO and C$_2$H$_2$O 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$_2$O molecule, CO and C$_4$H$_8$ respectively. Peak 3 was tentatively identified as Cyclo(L-Leu-trans-4-hydroxy-L-Pro). Peak 4 with m/z 211 was identified as Cyclo(L-Leu-L-Pro) by fragment ions m/z 183 and 155 at $t_{R}$ 5.41 min by losses of CO and C$_4$H$_8$.

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.

<table>
<thead>
<tr>
<th>Peak no.</th>
<th>RT (min)</th>
<th>Putative metabolite</th>
<th>Molecular formula</th>
<th>Adduct</th>
<th>SpecMZ</th>
<th>LibMZ</th>
<th>Mass error (ppm)</th>
<th>MS/MS fragments</th>
<th>Chemical class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>2.15</td>
<td>7-Hydroxy-2H-chromen-2-one</td>
<td>C_{11}H_{10}O_{3}</td>
<td>[M+H]^+</td>
<td>159.066</td>
<td>161.024</td>
<td>-1.958</td>
<td>133, 117, 105</td>
<td>Coumarin</td>
</tr>
<tr>
<td>2.</td>
<td>3.73</td>
<td>Dextrorphan</td>
<td>C_{11}H_{22}NO</td>
<td>[M+H]^+</td>
<td>259.078</td>
<td>258.190</td>
<td>-0.892</td>
<td>240, 227, 201</td>
<td>Alkaloid</td>
</tr>
<tr>
<td>3.</td>
<td>4.29</td>
<td>Cyclo(L-Leu-trans-4-hydroxy-L-Pro)</td>
<td>C_{11}H_{14}N_{3}O_{3}</td>
<td>[M+H]^+</td>
<td>229.068</td>
<td>227.140</td>
<td>1.929</td>
<td>209, 199, 171</td>
<td>Amino acid</td>
</tr>
<tr>
<td>4.</td>
<td>6.41</td>
<td>Cyclo(L-Leu-L-Pro)</td>
<td>C_{11}H_{14}N_{3}O_{2}</td>
<td>[M+H]^+</td>
<td>209.154</td>
<td>211.145</td>
<td>-1.991</td>
<td>183, 155</td>
<td>Amino acid</td>
</tr>
<tr>
<td>5.</td>
<td>8.47</td>
<td>1-[4-hydroxy-3-(3-methylbut-2-enyl)phenyl]ethanone</td>
<td>C_{19}H_{18}O_{2}</td>
<td>[M+H]^+</td>
<td>205.086</td>
<td>205.122</td>
<td>-0.036</td>
<td>187, 163, 149</td>
<td>Organoxygen compound</td>
</tr>
<tr>
<td>6.</td>
<td>9.84</td>
<td>Tricetin</td>
<td>C_{11}H_{14}O_{7}</td>
<td>[M+H]^+</td>
<td>304.298</td>
<td>303.050</td>
<td>1.248</td>
<td>301, 285, 126</td>
<td>Flavonoid</td>
</tr>
<tr>
<td>7.</td>
<td>10.12</td>
<td>Didymic acid</td>
<td>C_{11}H_{22}O_{5}</td>
<td>[M+H]^+</td>
<td>368.352</td>
<td>369.170</td>
<td>-0.818</td>
<td>351, 325, 337</td>
<td>Dibenzoferan</td>
</tr>
<tr>
<td>8.</td>
<td>10.49</td>
<td>Benzyldimethyltetradecylammonium</td>
<td>C_{23}H_{42}N_{3}</td>
<td>[M+H]^+</td>
<td>332.330</td>
<td>333.339</td>
<td>-1.009</td>
<td>304, 290, 278</td>
<td>Benzenoid</td>
</tr>
<tr>
<td>9.</td>
<td>11.12</td>
<td>3-(1,2,3,6-tetrahydropyridin-4-yl)-1H-pyrrolo[2,3-b]pyridine</td>
<td>C_{12}H_{12}N_{3}</td>
<td>[M+Na]^+</td>
<td>223.060</td>
<td>222.100</td>
<td>0.959</td>
<td>183, 171, 157</td>
<td>Alkaloid</td>
</tr>
<tr>
<td>10.</td>
<td>11.74</td>
<td>Nonactic acid</td>
<td>C_{12}H_{24}O_{4}</td>
<td>[M+Na]^+</td>
<td>225.194</td>
<td>225.278</td>
<td>-0.084</td>
<td>185, 157, 129</td>
<td>Polyketide</td>
</tr>
<tr>
<td>11.</td>
<td>11.79</td>
<td>Benzenepropanamide, N-[2-(acetylxylo)-1-(phenoxymethyl)ethyl]-alpha-(benzoylamo)-</td>
<td>C_{23}H_{22}N_{2}O_{4}</td>
<td>[M+Na]^+</td>
<td>467.193</td>
<td>467.194</td>
<td>-0.001</td>
<td>427, 403, 385, 353</td>
<td>Amino acid</td>
</tr>
<tr>
<td>12.</td>
<td>11.87</td>
<td>Aurantiamide acetate</td>
<td>C_{12}H_{18}N_{4}O_{4}</td>
<td>[M+H]^+</td>
<td>445.210</td>
<td>445.212</td>
<td>-0.002</td>
<td>427, 403, 385, 353</td>
<td>Amino acid</td>
</tr>
<tr>
<td>14.</td>
<td>12.51</td>
<td>Esteratin</td>
<td>C_{12}H_{24}O_{6}</td>
<td>[M+Na]^+</td>
<td>1035.420</td>
<td>1035.658</td>
<td>-0.238</td>
<td>489, 434, 377</td>
<td>Steroid</td>
</tr>
<tr>
<td>15.</td>
<td>12.63</td>
<td>(S,S)-asaspernamate</td>
<td>C_{12}H_{24}N_{4}O_{4}</td>
<td>[M+Na]^+</td>
<td>529.206</td>
<td>529.210</td>
<td>-0.004</td>
<td>489, 429, 403</td>
<td>Amino acid</td>
</tr>
<tr>
<td>16.</td>
<td>12.78</td>
<td>3-Deoxycarcynotol</td>
<td>C_{12}H_{24}O_{7}</td>
<td>[M+Na]^+</td>
<td>455.244</td>
<td>457.220</td>
<td>-1.976</td>
<td>417, 375, 315</td>
<td>Terpenoid</td>
</tr>
<tr>
<td>17.</td>
<td>12.95</td>
<td>Monoelaidin</td>
<td>C_{12}H_{24}O_{4}</td>
<td>[M+H]^+</td>
<td>357.287</td>
<td>357.299</td>
<td>-0.012</td>
<td>339, 237</td>
<td>Glycerolipid</td>
</tr>
<tr>
<td>18.</td>
<td>13.01</td>
<td>3-(2-methoxy-3-((2S,3R,4S,5R)-3,4,5,6,7,8,8a-tetrahydro-2H-pyran-2-yl)oxy)phenyl)propanoic acid</td>
<td>C_{13}H_{22}O_{9}</td>
<td>[M+Na]^+</td>
<td>379.266</td>
<td>381.116</td>
<td>-1.850</td>
<td>341, 315, 299</td>
<td>Terpenoid</td>
</tr>
<tr>
<td>19.</td>
<td>13.10</td>
<td>(E)-5-(1,2,4a,5-tetramethyl-7-oxo-3,4,8a-tetrahydro-2H-naphthalen-1-yl)-3-methylpent-2-enolic acid</td>
<td>C_{23}H_{30}O_{3}</td>
<td>[M+Na]^+</td>
<td>339.250</td>
<td>341.209</td>
<td>-1.959</td>
<td>301, 291, 259</td>
<td>Terpenoid</td>
</tr>
<tr>
<td>20.</td>
<td>13.27</td>
<td>(1S,4aS,5R)-5-[2-(furan-3-y1)ethyl]-1,4a-dimethyl-6-methylidene-3,4,5,7,8,8a-hexahydro-2H-naphthalene-1-carboxylic acid</td>
<td>C_{12}H_{20}O_{3}</td>
<td>[M+H]^+</td>
<td>313.233</td>
<td>315.197</td>
<td>-1.964</td>
<td>297, 287, 271</td>
<td>Terpenoid</td>
</tr>
<tr>
<td>21.</td>
<td>13.56</td>
<td>Echinulin</td>
<td>C_{12}H_{20}N_{2}O_{2}</td>
<td>[M+H]^+</td>
<td>462.307</td>
<td>462.313</td>
<td>-0.006</td>
<td>420, 406, 392</td>
<td>Alkaloid</td>
</tr>
<tr>
<td>22.</td>
<td>13.63</td>
<td>Cellexob</td>
<td>C_{11}H_{18}F_{3}N_{2}O_{2}</td>
<td>[M+H]^+</td>
<td>381.260</td>
<td>382.083</td>
<td>-0.823</td>
<td>380, 316, 247</td>
<td>Alkaloid</td>
</tr>
<tr>
<td>23.</td>
<td>13.75</td>
<td>Lauryle L-carnitine</td>
<td>C_{12}H_{20}N_{2}O_{3}</td>
<td>[M+Na]^+</td>
<td>342.334</td>
<td>344.278</td>
<td>-1.944</td>
<td>265</td>
<td>Fatty acid ester</td>
</tr>
<tr>
<td>24.</td>
<td>13.90</td>
<td>Ethyl 3-(2,4-dihydroxy-6-methylpyridin-3-yl)-3-(3-hydroxyphenyl)propanoate</td>
<td>C_{11}H_{18}N_{2}O_{6}</td>
<td>[M+Na]^+</td>
<td>341.264</td>
<td>340.120</td>
<td>1.144</td>
<td>300, 290, 274</td>
<td>Pyridine</td>
</tr>
<tr>
<td>25.</td>
<td>13.93</td>
<td>5-(4-carboxy-3-methylbutyl)-5,6,8a-trimethyl-3-oxo-4a,6,7,8-tetrahydro-4H-naphthalene-1-carboxylic acid</td>
<td>C_{23}H_{30}O_{5}</td>
<td>[M+Na]^+</td>
<td>371.274</td>
<td>373.199</td>
<td>-1.925</td>
<td>333, 323, 305, 291</td>
<td>Terpenoid</td>
</tr>
<tr>
<td>26.</td>
<td>14.13</td>
<td>4-(4-Hydroxy-2,6,6-trimethyl-1-cyclohexen-1-yl)-2-butanyl beta-D-glucopyranoside</td>
<td>C_{16}H_{34}O_{7}</td>
<td>[M+Na]^+</td>
<td>395.274</td>
<td>397.220</td>
<td>-1.946</td>
<td>357, 345, 333</td>
<td>Terpenoid</td>
</tr>
</tbody>
</table>
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 C₂H₂O which resulted in the ion at m/z 403. The next fragmentation was due to the losses of CH₃COOH and C₇H₆ 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 tᵣ 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 H₂O molecule, C₆H₆ and C₇H₄O respectively.

### Identification of terpenoids

Terpenoids are known to display a wide range of biological activities which include cancer chemopreventive 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 tᵣ 12.44 min with sodiumated adduct at m/z 427 displaying protonated adduct at m/z 405. The mass 405 showed MS² fragment ions at m/z 387, 363, 345 and 277 corresponding to the losses of H₂O, C₂H₂O, CH₃COOH and C₆H₆O₃ accordingly. Peak 16 with tᵣ 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, MS² analysis revealed fragment ions at m/z 417, 375 and 315 suggesting losses of H₂O, CH₃COOH and C₂H₄O₃ respectively. Peak 18 at tᵣ 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.

<table>
<thead>
<tr>
<th>No.</th>
<th>tR (min)</th>
<th>Lipids</th>
<th>Molecular formula</th>
<th>Adduct</th>
<th>Parent mass (m/z)</th>
<th>Category</th>
<th>Main class</th>
<th>Sub class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>11.20</td>
<td>LysoPC(0:0/18:0)</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;NO:P</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>524.334</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphocholine</td>
<td>Monoacylglycerophospho-choline</td>
</tr>
<tr>
<td>2.</td>
<td>11.69</td>
<td>LysoPE(12:0/0:0)</td>
<td>C&lt;sub&gt;1&lt;/sub&gt;H&lt;sub&gt;2&lt;/sub&gt;NO:P</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>398.239</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoethanolamine</td>
<td>Monoacylglycerophospho-ethanolamine</td>
</tr>
<tr>
<td>3.</td>
<td>13.18</td>
<td>LysoPE(0:0/18:3 (6Z,9Z,12Z))</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;4&lt;/sub&gt;NO:P</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>476.289</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoethanolamine</td>
<td>Monoacylglycerophospho-ethanolamine</td>
</tr>
<tr>
<td>5.</td>
<td>14.31</td>
<td>LysoPE(0:0/20:1(11Z))</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;NO:P</td>
<td>[M-H]</td>
<td>506.325</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoethanolamine</td>
<td>Monoacylglycerophospho-ethanolamine</td>
</tr>
<tr>
<td>7.</td>
<td>15.35</td>
<td>LysoPE(0:0/18:1(9Z))</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;4&lt;/sub&gt;NO:p</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>480.304</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoethanolamine</td>
<td>Monoacylglycerophospho-ethanolamine</td>
</tr>
<tr>
<td>8.</td>
<td>15.92</td>
<td>LysoPG(15:0/0:0)</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;4&lt;/sub&gt;OP</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>471.326</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoglycerol</td>
<td>Monoacylglycerophospho-glycerol</td>
</tr>
<tr>
<td>9.</td>
<td>16.07</td>
<td>LysoPl(18:1(9Z)/0:0)</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;6&lt;/sub&gt;OP</td>
<td>[M-H]</td>
<td>597.503</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoinositol</td>
<td>Monoacylglycerophospho-inositol</td>
</tr>
<tr>
<td>10.</td>
<td>16.45</td>
<td>LysoPE(0:0/22:2 (13Z,16Z))</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;NO:p</td>
<td>[M+H]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>534.374</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoethanolamine</td>
<td>Monoacylglycerophospho-ethanolamine</td>
</tr>
<tr>
<td>11.</td>
<td>16.77</td>
<td>LysoPG(13:0/0:0)</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;OP</td>
<td>[M-H]</td>
<td>441.317</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoglycerol</td>
<td>Monoacylglycerophospho-glycerol</td>
</tr>
<tr>
<td>12.</td>
<td>16.98</td>
<td>DG(18:3(6Z,9Z,12Z)/18:1(9Z)/0:0)[iso2]</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;5&lt;/sub&gt;Os</td>
<td>[M+NH4]&lt;sup&gt;*&lt;/sup&gt;</td>
<td>634.139</td>
<td>Glycerolipid</td>
<td>Diradylglycerol</td>
<td>Diacylglycerol</td>
</tr>
<tr>
<td>13.</td>
<td>17.04</td>
<td>LysoPl(18:0/0:0)</td>
<td>C&lt;sub&gt;2&lt;/sub&gt;H&lt;sub&gt;6&lt;/sub&gt;OP</td>
<td>[M-H]</td>
<td>599.428</td>
<td>Glycerophospholipid</td>
<td>Glycerophosphoinositol</td>
<td>Monoacylglycerophospho-inositol</td>
</tr>
</tbody>
</table>

[M+Na]<sup>*</sup> ion m/z 381 was identified based on their fragment ions m/z 341, 315 and 299 due to the losses of H<sub>2</sub>O molecule, C<sub>2</sub>H<sub>2</sub>O and CH<sub>3</sub>COOH subsequently. At tR 13.10 min, peak 19 with [M+Na]<sup>*</sup> adduct ion m/z 341 and [M+H]<sup>*</sup> adduct ion m/z 319 was observed to lose H<sub>2</sub>O, C<sub>2</sub>H<sub>2</sub>O and CH<sub>3</sub>COOH to form fragments m/z 301, 291 and 259 respectively. Peak 20 at tR 13.27 min shows deprotonated adduct ion at m/z 315 with subsequent losses of H<sub>2</sub>O, CO and C<sub>2</sub>H<sub>2</sub>O to form fragment ions at m/z 297, 287 and 271 correspondingly. Peak 25 at tR 13.93 min with parent ion m/z [M+Na]<sup>*</sup> 373 displayed m/z [M+H]<sup>*</sup> 351, showing characteristic losses of H<sub>2</sub>O, CO, CH<sub>3</sub>OH and CH<sub>3</sub>COOH to produce product ions at m/z 333, 323, 305 and 291 accordingly.

The peak at tR 14.13 min (peak 26) showed m/z [M+H]<sup>*</sup> 375 and m/z [M+Na]<sup>*</sup> 397. The m/z at 375 dissociated to give fragment ions at m/z 357, 345 and 333 corresponding to losses of H<sub>2</sub>O, CH<sub>3</sub>OH and C<sub>2</sub>H<sub>2</sub>O respectively. Peak 28 at tR 14.45 min having m/z [M+Na]<sup>*</sup> 411 and m/z [M+H]<sup>*</sup> 389 reveals MS/MS fragment ions at m/z 371 and 319 due to the losses of H<sub>2</sub>O molecule and C<sub>3</sub>H<sub>10</sub>O. Peak 29 and 30 with tR 14.52 min and 14.68 min were attributed to pseudo-anisatin. Both peaks showed similar parent ions at m/z 321 [M+Na]<sup>*</sup> and m/z 299 [M+H]<sup>*</sup> displaying fragment ions corresponding to successive losses of H<sub>2</sub>O, CO, CH<sub>3</sub>OH and CH<sub>2</sub>COO 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. Dextorphan with protonated parent ion at m/z 258 was assigned to peak 2 at tR 3.73 min and displaying fragments ions corresponding to successive losses of H<sub>2</sub>O, NH<sub>2</sub>CH<sub>3</sub> and C<sub>4</sub>H<sub>8</sub> at m/z 240, 227 and 201 respectively. Echinulin having an MS parent ion m/z 462 [M+H]<sup>*</sup> at tR 13.56 min (peak 21) produces ions at m/z 420, 406 and 392 suggesting the losses of C<sub>3</sub>H<sub>6</sub>, C<sub>4</sub>H<sub>8</sub> and C<sub>3</sub>H<sub>10</sub> in a relative manner.

The peak with tR 13.63 min (peak 22) gives a mass peak of m/z 382 [M+H]<sup>*</sup> 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...
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$_2$H$_4$, CH$_3$OH and C$_3$H$_7$ 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$_3$CN$^+$ and C$_6$H$_5$N 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

![MS/MS spectrum of 7-Hydroxy-2H-chromen-2-one and proposed fragmentation patterns.](image-url)
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$_2$O molecule, C$_2$H$_5$O and C$_3$H$_6$OH 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$_2$O molecule, C$_3$H$_8$ and C$_5$H$_{10}$ 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$_3$H$_7$N. 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$_2$O molecule and CH$_3$COOH 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$_2$O molecule and C$_4$H$_8$O$_4$ in a relative manner.

**Identification of Organoxygen compounds**

Only two organooxygen compounds were identified in the
ethyl acetate extract of *M. oleifera* seeds. Peak 5 was detected at *t*<sub>r</sub> 8.47 min with protonated adduct [M+H]<sup>+</sup> at m/z 205. The MS<sup>2</sup> spectrum was characterized by losses of H<sub>2</sub>O, C<sub>2</sub>H<sub>8</sub> and C<sub>4</sub>H<sub>8</sub> moieties to exhibit fragments at m/z 187, 163 and 149 correspondingly. Peak 32 at *t*<sub>r</sub> 14.95 min gave an [M+Na]<sup>+</sup> ion at m/z 439 and exhibited an [M+H]<sup>+</sup> ion at m/z 417. The m/z 417 ion yielded product ions at m/z 399, 387 and 373 suggesting losses of H<sub>2</sub>O, CH<sub>2</sub>O and C<sub>2</sub>H<sub>4</sub>O 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 UHPLC-MS/MS analysis of *M. oleifera* seeds extract. Peak 6 with *t*<sub>r</sub> 9.84 min showed an [M+H]<sup>+</sup> 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<sub>2</sub>H<sub>4</sub>O<sub>3</sub> respectively as shown in Figure 4. By referring to mass spectral libraries, peak 6 was tentatively identified as tricetin. Lysionotin at *t*<sub>r</sub> 15.22 min (peak 34) gave an [M+H]<sup>+</sup> ion at m/z 345 and an [M+Na]<sup>+</sup> ion at m/z 367. The [M+H]<sup>+</sup> ion at m/z 345 revealed fragment ion peaks at m/z 327, 313 and 3011 due to the elimination of H<sub>2</sub>O, CH<sub>2</sub>O and C<sub>2</sub>H<sub>4</sub>O accordingly. Peak 39 with *t*<sub>r</sub> 16.68 min produced an [M+H]<sup>+</sup> ion m/z 611 and an [M+NH<sub>4</sub>]<sup>+</sup> ion m/z 628. The [M+H]<sup>+</sup> ion at m/z 611 displayed fragment ions at m/z 593 and 449 due to loss of H<sub>2</sub>O molecule and consequent loss of C<sub>9</sub>H<sub>10</sub>O<sub>5</sub>.

**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*<sub>r</sub> 10.49 min with characteristic fragmentation pattern m/z 304, 290 and 278 is attributed to benzyldimethyltetradecylammonium affirming the losses of C<sub>2</sub>H<sub>5</sub>, C<sub>9</sub>H<sub>17</sub> and C<sub>4</sub>H<sub>8</sub> accordingly from the protonated precursor ion m/z 333.

**Dibenzofuran:** Peak 7 at *t*<sub>r</sub> 10.12 min with m/z [M-H]<sup>-</sup> 369 was putatively identified as didymic acid. The fragment ions at m/z 351, 337 and 325 corresponding to the losses of H<sub>2</sub>O, CH<sub>4</sub>O and CO<sub>2</sub> respectively. Didymic acid is a member of dibenzofuran endowed with antimicrobial activity (Dieu et al., 2012).

**Polyketide:** Nonatic acid (peak 10) at *t*<sub>r</sub> 10.49 min produced [M+Na]<sup>+</sup> ion at m/z 225 displaying [M+H]<sup>+</sup> ion at m/z 203. The ion m/z 203 was subjected to MS<sup>2</sup> analysis yielding product ions m/z 185, 157 and 129 due to the losses of H<sub>2</sub>O molecule, CH<sub>2</sub>O and C<sub>2</sub>H<sub>4</sub>O correspondingly.

**Pyridine:** Peak 24 gave m/z [M+Na]<sup>+</sup> 340 and m/z [M+H]<sup>+</sup> 318 at *t*<sub>r</sub> 13.90 min with MS<sup>2</sup> spectrum characterized by losses of H<sub>2</sub>O molecule, CO and C<sub>2</sub>H<sub>4</sub>O 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]<sup>+</sup> 1035 and m/z [M+H]<sup>+</sup> 507 at *t*<sub>r</sub> 12.51 min was annotated as esterastin. The MS/MS spectra of esterastin displayed product ion signals at m/z 489 (loss of H<sub>2</sub>O), 434 (loss of C<sub>2</sub>H<sub>2</sub>O<sub>3</sub>) and 377 (loss of C<sub>4</sub>H<sub>2</sub>O<sub>3</sub>).

Caudatin (peak 36) at *t*<sub>r</sub> 15.77 min produced [M+Na]<sup>+</sup> ion at m/z 513 displaying [M+H]<sup>+</sup> 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<sup>2</sup> analysis from which the m/z 361 (B) was acquired due to the loss of C<sub>7</sub>H<sub>12</sub>O<sub>2</sub>. 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<sub>2</sub>O molecules from (B). To further the investigations, fragmentation patterns of ions B, C and D were chosen as precursor ions in MS<sup>2</sup> 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<sub>2</sub>H<sub>4</sub>O<sub>2</sub>. Ion G m/z 243 was generated from ion C suggesting the loss of C<sub>2</sub>H<sub>4</sub>O and ion H m/z 283 was produced through elimination of C<sub>2</sub>H<sub>2</sub>O<sub>2</sub> 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*<sub>r</sub> 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...
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, 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 4.** MS/MS spectrum of tricetin 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 phophoethanolamine. 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 of fatty 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
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
A 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 *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 process, the production of biodiesel could be further improved.
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

This research was self-funded.

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


Matt B (2022). Integrating LC-MS/MS based molecular networking and
analysis of medicinal plants. Rhode Island IDeA Network of Biomedical Research Excellence.


Figure S1. MS/MS spectra and structures of various metabolites identified in ethyl acetate extract of *M. oleifera* seeds in positive ion mode.
Figure S2. MS/MS spectra and structures of identified lipids in ethyl acetate extract of *M. oleifera* seeds in positive ion mode.