

# Chemical Profile and Chemometric Analysis of Genetically Modified Soybeans Produced in the Triângulo Mineiro Region (MG), Brazil

Marco Aurélio Borba Moreira, Amilton Diniz e Souza, Fernanda Barbosa Borges Jardim,  
Luís Carlos Scalon Cunha (Corresponding author)

Federal Institute of Triângulo Mineiro, Av. João Batista Ribeiro, 4000, 38064-790,  
Uberaba, Minas Gerais, Brazil. E-mail: luiscunha@iftm.edu.br

Mário Machado Martins, Luiz Ricardo Goulart

Federal University of Uberlândia, Institute of Biotechnology, Av. Pará, 1720, 38400-902,  
Uberlândia, Minas Gerais, Brazil.

Sérgio Antônio Lemos de Moraes, Francisco José Torres de Aquino, Wellington de  
Oliveira Cruz, Lucas Gustavo da Costa, Waldomiro Borges Neto

Federal University of Uberlândia, Institute of Chemistry, Av. João Naves de Ávila, 2121,  
38400-902, Campus Santa Mônica, 38408-100, Uberlândia, Minas Gerais, Brazil.

Received: January 4, 2021    Accepted: February 12, 2021    Published: February 22, 2021

doi:10.5296/jas.v9i2.18139

URL: <https://doi.org/10.5296/jas.v9i2.18139>

## Abstract

Soy production in Brazil is an important factor for the agro-industrial, economic, and social development of the country. The expansion of soy in the Brazilian territory is mainly due to the incorporation of new genetic characteristics into cultivars that granted resistance to the Cerrado conditions and to herbicides. Currently, Brazilian soy production is the result of genetically modified cultivars. Studies regarding the chemical composition of soybeans show that qualitative and quantitative variations can occur, depending on the region of production. This work aimed to investigate the chemical composition of soybeans produced in different cities of the Triângulo Mineiro region/MG, Brazil (Harvest 2017/2018) and stored in three warehouses located in the city of Uberaba/MG. The grain analysis was made by liquid chromatography coupled to electrospray ionization mass spectrometry (LC-MS-ESI). The classes of metabolites identified from methanolic extraction were organic acids, phenolic

compounds, flavonoids, sugars, amino acids, dipeptides, nitrogenous bases, nucleosides, sphingolipids, and fatty acids. The isoflavones genistein, daidzein, glycitein, genistin, acetylaidzin, and acetylgenistin were identified in soybeans from the three warehouses. The flavonoid eriodictyol-*O*-hexoside was also found. The Principal Component Analysis (PCA) from the mass spectrum data obtained by direct injection in the negative and positive modes evidenced the well-defined separation of three groups, indicating that there was variance among the soy samples from each warehouse. The samples from warehouses 1 and 3 showed greater similarity in the Hierarchical Cluster Analysis (HCA) in negative mode, while in positive mode, the samples from warehouses 2 and 3 presented greater similarity.

**Keywords:** *Glycine max*, chemical composition, multivariate analysis

## 1. Introduction

The global production of soybean (*Glycine max*) for the 2017/2018 harvest was around 342 million tons, with a planted area of 124.5 million hectares. The United States ranked first with a production of 120 million tons (USDA, 2020), followed by Brazil with about 119 million grains, representing 34.7% of the world production (CONAB, 2020). Brazilian production in the 2018/2019 harvest was approximately 115 million tons, remaining as the second largest world producer (CONAB, 2020). The projection for the 2019/2020 harvest is that these two countries will remain world leaders in soy production (USDA, 2020).

Soy is a key component of the process of agro-industrial development in Brazil and its remarkable current production was made possible by the development of genetically modified varieties adapted to the Cerrado conditions, and resistant to herbicides. Genetically modified cultivars are used almost completely in the soybean area planted in Brazil (Dall'agnol, 2016). Among the many producing regions, the State of Minas Gerais, located in Southeast Brazil, contributes with 9% of the total production, being the largest producer in that region (USDA, 2020; CONAB, 2019). It is estimated that Brazil will export about 75 million tons in 2020, and 44.5 million tons will be used in grinding (CONAB, 2020) for the production of soybean meal and oil. Soybean oil is used mainly for human consumption and for the manufacture of biodiesel, while the bran is used for the production of animal feed and industrial human food (CONAB, 2019; CONAB, 2018). Soy has several nutritional properties, being considered a very versatile food rich in essential nutrients such as lipids, proteins, fibers, vitamins, and minerals (Jooyandeh, 2011). Various products and by-products for human consumption can originate from soy, such as flour, butter, cheese, soy flakes, dietary foods, textured protein, soy-burgers, soy-based infant formulas, supplements, soy milk and soy-based drinks, and veggie food, among others. In order to meet a growing demand in human food, soy-based products have been developed in order to provide new healthier foods with a more attractive taste (He & Chen, 2013; Rizzo & Baroni, 2018).

In this perspective, the consumption of soy-based foods has been increasing in recent years due to their benefits for human health. These benefits are often related to bioactive substances present in soybeans, mainly phenolic compounds (Verardo et al., 2015). This class of metabolites is represented in the soy by phenolic acids and flavonoids. Among flavonoids, isoflavones are the most studied components. These compounds are related to the antioxidant

properties of soy (Nam et al., 2014) and with other biological properties like anticancer potential (Ko et al., 2014; Mahmoud et al., 2014) and estrogenic effect (Zaheer & Akhtar, 2015). Soy is one of the richest food sources of isoflavones (Mahmoud et al., 2014) and genistein, daidzein, genistin, daidzin and glycitein, are the most commonly found compounds (Cavaliere et al., 2007; Lee et al., 2008; Dueñas et al., 2012; Mahmoud et al., 2014; Verardo et al., 2015).

Most studies with soybeans are focused primarily on phenolic compounds. However, other classes of metabolites have been found. The list includes vitamins and fatty acids (Chung et al., 2017), amino acids and dipeptides (Forbes et al., 2007; Wang et al., 1978), alkaloids (Wang et al., 2016; Ghahari et al., 2017), proteins (Freiria et al., 2016), polysaccharides (Choct et al., 2010), terpenes and essential oils (Ghahari et al., 2017), saponins, steroids (Kang et al., 2010), sphingolipids (Sullards et al., 2000), and tannins (Malenčić et al., 2007). As with other plants, the chemical composition of soybeans can vary qualitatively and quantitatively depending on the species, type of soil, and climatic conditions of each producing region (Lozovaya et al., 2005; Nam et al., 2014).

Although Brazil is one of the major world producers and considering the economic importance of the various products that are obtained from soybean processing, few studies are found in the literature analyzing the chemical constitution of the organic metabolites of the Brazilian soy, particularly soy produced in the Triângulo Mineiro region - Minas Gerais, Brazil. Thus, the objective of this work was to investigate the chemical constituents of soybeans from the 2017/2018 harvest, produced and stored in the Triângulo Mineiro region, an important soybean production center in the State of Minas Gerais, Brazil. The data presented here represents the first information generated for this region, which highlights the importance of this study.

## 2. Materials and Methods

### 2.1 Study Area

The genetically modified soybeans used in this study were produced in the Triângulo Mineiro region/MG - Brazil (2017/2018 harvest), and the area is outlined in Figure 1. The collections of the soy samples were carried out in three different warehouses located in the city of Uberaba/MG, which receives the main volume of soy produced in this region.



Figure 1. Map of the soybean producing region - 2017/2018 harvest

## 2.2 Warehouses

Three warehouses were selected in the Uberaba region from which to collect the soybean samples. These warehouses were chosen because they receive a considerable volume of genetically modified soybeans from different producers in Uberaba and the nearby region. The selected warehouses, according to latitude and longitude coordinates are called herein warehouse 1 (Lat:-19.707872, Long: -47.979147), warehouse 2 (Lat:-19.700648, Long:-47.977304), and warehouse 3 (Lat:-19.702434, Long:-47.972029).

## 2.3 Sample Collection and Preparation

The sampling was conducted in accordance with the Normative Instruction N°. 60, article 17 of MAPA, from 12/22/2011, with modifications (MAPA, 2011).

Samples of soybeans were collected from the warehouses 1, 2, and 3, at the beginning of the harvest (03/20/18 - 03/29/18). After receiving the material, the impurities (0 - 1% impurities) were removed, and then, the grains were transferred to a grain dryer, where they remained at 110 °C, until reaching the moisture content required for storage (13 - 14%). Then, the collection was carried out at the outlet of the grain dryer, through the hole used to measure the moisture content. A mass of approximately 840 g was collected every 5 minutes, totaling 10 kg per day. At the end, the three samples collected from each warehouse were grouped and homogenized. Then, they were stored in an ultra-freezer until the moment of analysis. During the three days of collection, the receipt of soybeans from all the cities outlined in Figure 1 was also recorded.

## 2.4 Soybean Methanol Extraction

Ground soybeans (5.0 g) from each of the warehouses were weighed in 50 mL beakers, followed by the addition of 15 mL of spectroscopic methanol. Extractions were performed in an ultrasound bath for a period of 30 minutes. The supernatant was first filtered in C18 cartridges for glass syringes containing Millipore filters (0.45 µm), and again filtered into vials (2 mL). This procedure was performed in triplicate for each warehouse. Subsequently, the samples of the vials were analyzed by liquid chromatography coupled to mass spectrometry with electrospray ionization (LC-MS-ESI).

The extractions for the mass spectrometry experiments, by direct injection, followed the same procedure for the LC-MS-ESI experiments. However, ten extractions for each warehouse were carried out. These extractions were done to obtain the grain mass spectra, targeting the chemometric analysis.

## 2.5 Liquid Chromatography Coupled to Electrospray Ionization Mass Spectrometry (LC-MS-ESI)

The soybean analyzes were performed on an Agilent® Infinity 1260 UHPLC coupled to a high resolution mass spectrometer (Agilent® 6520 B) Quadrupole Time of Flight (Q-TOF) with an electrospray ionization source (ESI).

The samples were analyzed at a concentration of 2.0 mg mL<sup>-1</sup>. A volume of 1.0 µL was

injected into the liquid chromatograph on a Poroshell 120EC C18 column (100 mm x 3.0 mm, 2.7  $\mu\text{m}$ ), maintained at 35 °C. The chromatographic conditions were ultrapure water acidified with formic acid (0.1% v/v) (mobile phase A), and methanol (mobile phase B). The solvent gradient system used started with 2% B (0-1 min), 98% B (1-5 min), and 98% B (5-10 min), at a flow rate of 0.7 mL min<sup>-1</sup>.

The mass spectrometer operated with a nebulizing pressure of 20 PSI, drying gas flow of 8.0 L min<sup>-1</sup>, and a temperature of 220 °C, with 4.5 kV energy in the capillary. The measures were taken in both positive and negative modes, obtaining [M+H]<sup>+</sup> and [M-H]<sup>-</sup> ions, respectively. The obtained masses were in high resolution (MS) and the proposed molecular formula followed the lowest difference between the experimental mass and the theoretical mass, double bond equivalence, and nitrogen rule. The error in ppm was calculated according to the equation:  $\text{Eppm} = [\text{experimental mass} - \text{exact mass}/\text{exact mass}] \cdot 10^6$ .

Sequential mass spectrometry (MS<sup>2</sup>) was performed at different collision energies for the positive and negative modes. The structure suggestions were proposed comparing the fragment mass spectra and the high-resolution masses with other works in the literature and Metlin database.

### *2.6 Mass Spectrometry by Direct Injection*

Mass spectrometry by direct injection was used to compare the chemical composition of soybeans between the warehouses. The analyzes were performed on a mass spectrometer Agilent® 6520 Quadrupole Time of Flight (Q-TOF) with an electrospray ionization source (ESI) operating in positive and negative modes. The methanolic solutions were injected directly into the mass spectrometer at a flow of 0.2 mL min<sup>-1</sup>. The general conditions of the equipment were as the follows: drying gas temperature of 220 °C, capillary voltage of 4.5 KV, and 65 V of cone voltage. The mass spectra were measured in scan mode with a ratio/charge ( $m/z$ ) of 100 to 1,000u. The ion masses in high resolution were obtained in positive and negative modes, and presented as [M+H]<sup>+</sup> or [M-H]<sup>-</sup>. A total of 10 samples of soybeans from each warehouse were extracted with methanol and analyzed by mass spectrometry through direct injection.

### *2.7 Statistical Analysis*

The average of ten analyzes for the soybeans in each warehouse were obtained using the MassHunter Workstation Qualitative Analysis software (Agilent®). The ions in each sample were obtained for the positive and negative modes. Although the data were obtained in a high resolution mass analyzer, only the integer values of the mass/charge ratio ( $m/z$ ) were considered. The data were exported to spreadsheets as a table of mass/charge ratio ( $m/z$ ) and absolute abundance. The files were grouped into three different folders: soybeans from Warehouse 1, soybeans from Warehouse 2, and soybeans from Warehouse 3.

The data, in both modes, were transported to the MATLAB environment, version R2015a, standardized, and the ion signals with an intensity lower than or equal to 5% of the maximum abundance were removed, resulting in the matrix for analysis. The data centered on the mean were submitted to unsupervised exploratory analysis of PCA (Principal Component Analysis)

(Hotelling, 1933) and HCA (Hierarchical Cluster Analysis) (Bridges Junior, 1966). The models were built using PLS\_Toolbox, version 8.62.

### 3. Results and Discussion

#### 3.1 Identification of Compounds by LC-MS-ESI

The chemical profile of soybeans from warehouses 1, 2, and 3, in the negative and positive modes, are represented in the base-peak chromatograms (BCP), shown in Figure 2. From the chromatogram analyses obtained by LC-MS-ESI, in the negative and positive modes, it was possible to infer that the soybeans stored in the three warehouses (Harvest 2017/2018) presented similarities regarding their chemical composition. This was confirmed when the chemical composition of the three warehouses was investigated.

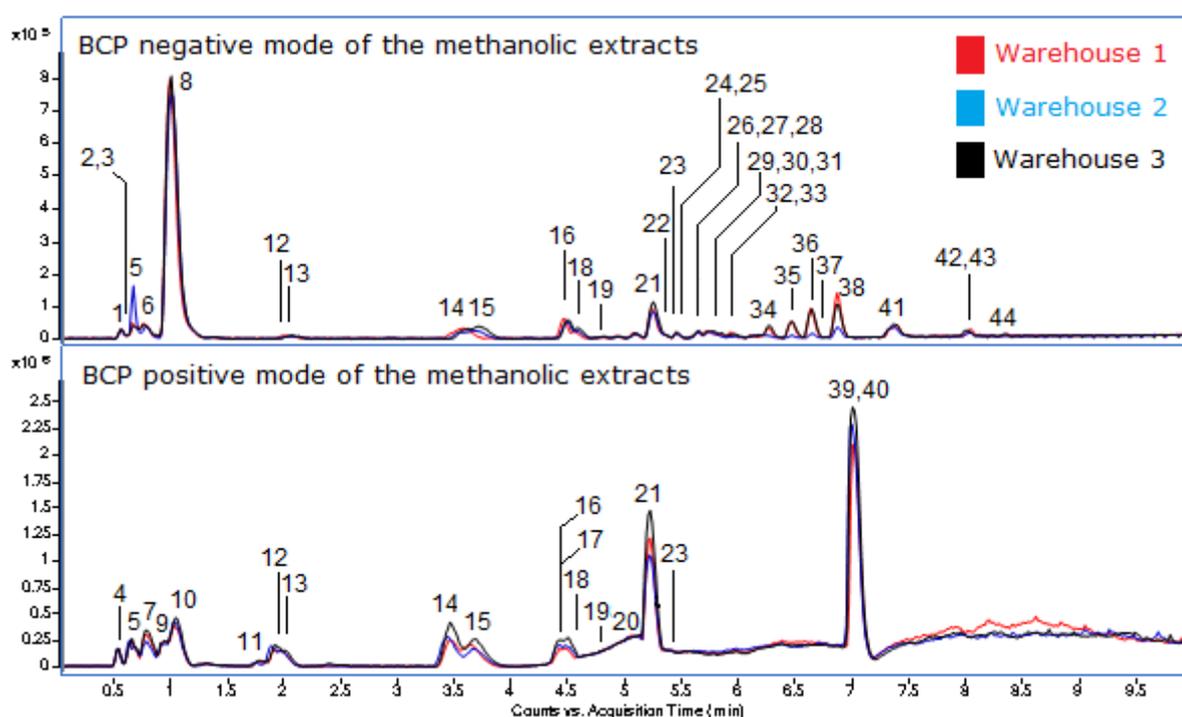


Figure 2. Chromatograms of methanol extracts of soybeans

Table 1 shows the ions in negative and positive modes, error in ppm, fragmentation ( $MS^2$ ), molecular formula, and the attempt to identify the compounds present in the soybean methanol extracts produced in the Triângulo Mineiro region and stored at three different warehouses in the city of Uberaba/MG. According to LC-MS-ESI, no qualitative variations were observed between the grains from the warehouses 1, 2, and 3. However, by inspecting the area of each peak in the chromatogram, it is possible to suggest that there are quantitative variations in the different soy samples. The proposed chemical structures are shown in Figure 3.

A total of 44 compounds were observed in the methanol extracts of soybeans, with 34 being identified. The chemical composition of the soybeans analyzed by LC-MS-ESI from methanolic extracts, consisted of organic acids, phenolic compounds, saccharides, amino acids, dipeptides, nitrogenous bases, nucleosides, sphingolipids, fatty acids, and flavonoids,

such as flavonol, flavanone, flavone, and isoflavones. Most of the identified components found in this study have been previously reported in the composition of soybeans (Wang et al., 1978; Cavaliere et al., 2007; Dueñas et al., 2012; Teixeira et al., 2012; Pereira et al., 2013; Nam et al., 2014; Verardo et al., 2015; Huang et al., 2016).

Some organic acids were observed in the first minute of the chromatogram. The molecular ions  $[M-H]^-$  at  $m/z$  225 and  $m/z$  195, according to Guerreiro et al. (2014), Li et al. (2016), and the Metlin library, were identified as glucoheptonic acid (2) and gluconic acid (3), respectively. Gluconic acid has already been determined in the root and shoot extracts of *Glycine max* L. (cv. Suzuyutaka) by capillary electrophoresis/mass spectrometry (Tawarayama et al., 2014).

The isocitric acid (isocitrate) (6) and citric acid (8) isomers showed the molecular ion  $[M-H]^-$  at  $m/z$  191 and were identified in soybeans from the three warehouses by comparison with Metlin database, Khadi et al. (2017) and Masike et al. (2017). Isocitrate (6) is part of the structure of the isocitratolase enzyme. This enzyme participates in the regulation of the glyoxylate cycle and plays a role in the seed germination process (Martins et al., 2000). The content of isocitrate in soybean seeds can vary with the type of storage and the different cultivars (Carvalho et al., 2014). Both isomers showed a main fragment at  $m/z$  111 for the loss of 80 Da. The fragmentation spectra ( $MS^2$ ) of the two compounds showed the fragments at  $m/z$  173  $[M-H-H_2O]^-$ , 155  $[M-H-2H_2O]^-$ , 147  $[M-H-CO_2]^-$ , and 129  $[M-H-CH_2O_3]^-$ . The fragment ion at  $m/z$  117, only present in the isocitrate (6), was considered a differentiation ion between the two isomers. The position of the hydroxyl group in the isocitrate favors the presence of  $\alpha$  cleavage between carbons 2 and 3, generating the fragment ion at  $m/z$  117, which it is difficult to occur in the citric acid.

Two derivatives of the jasmonic acid have been suggested for the investigated soybeans. The hydroxyjasmonic acid-*O*-hexoside (25) was identified from the molecular ion  $[M-H]^-$  at  $m/z$  387. The loss of 180 Da gave rise to the major fragment at  $m/z$  207  $[M-H-Hexose]^-$ , and its fragmentation spectrum  $MS^2$  is in accordance with Pereira et al. (2013). The compound (30) present at  $m/z$  305 and with a base peak at  $m/z$  225  $[M-H-SO_3]^-$ , has been identified as the hydroxyljasmonic acid-*O*-sulfate (Farg et al., 2017). Among the various studies consulted in the literature on the chemical constituents of soybeans, no reports were found on the presence of these compounds. Jasmonic acid and its derivatives are produced by different species of plants and are related to defense mechanisms. In *Glycine max* (L.) cv. G7R-315, the content of jasmonic acid has been studied in the pericarp of soybean seeds at various stages of development (Creelman et al., 1992).

Amino acids and dipeptides were represented by arginine (4), leucine (9), tyrosine (10), phenylalanine (12), tryptophan (18), glutamyl-tyrosine (14), and glutamyl-phenylalanine (21), while the nitrogenous bases identified were adenine (7) and guanine (11). These compounds are well known in the composition of soybeans (Taira, 1966; Wang et al., 1978; Mu et al., 2012; Teixeira et al., 2012).

Table 1. Proposal for identification of compounds present in the soybeans from warehouses 1, 2, and 3 by LC-MS-ESI

N.	Rt (min)	[M - H] <sup>-</sup> MS	[M + H] <sup>+</sup> MS	Error (ppm) - / +	Fragments MS <sup>2</sup> [M-H] <sup>-</sup> / [M+H] <sup>+</sup>	Molecular Formula	Tentative identity	Reference
1	0.57	272.9590	--	--	--	--	N.I	--
2		225.0614	--	-0.88	10 eV: <b>179</b> , 161, 149, 143	C <sub>7</sub> H <sub>14</sub> O <sub>8</sub>	Glucoheptonic acid	Metlin; Guerreiro et al., 2014
3	0.64	195.0511	--	0.0	15 eV: 177, 159, 129, 99, 87, <b>75</b> , 59	C <sub>6</sub> H <sub>12</sub> O <sub>7</sub>	Gluconic acid	Metlin; Li et al., 2016; Guerreiro et al., 2014
4	--		175.1189	0.0	20 eV: 158, 141, 130, <b>116</b>	C <sub>6</sub> H <sub>14</sub> N <sub>4</sub> O <sub>2</sub>	Arginine	Taira, 1966
5	0.72	341.1088	343.1236	-0.29/0.29	--	C <sub>12</sub> H <sub>22</sub> O <sub>11</sub>	Sucrose	Teixeira et al., 2012
6	0.78	191.0197	--	0.0	10 eV: 183, 173, 155, 147, 129, 117, <b>111</b>	C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	Isocitric acid (Isocitrate)	Metlin; Kadhi et al., 2017; Masike et al., 2017
7	--		136.0618	0.0	20 eV: <b>119</b>	C <sub>5</sub> H <sub>5</sub> N <sub>5</sub>	Adenine	Mu et al., 2012
8	0.98	191.0196	193.0344	-0.52/0.51	10 eV: 173, 155, 147, 130, 129, <b>111</b> / 10 eV: 175, 157, 147, 139, <b>129</b> , 115, 111	C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	Citric acid	Metlin; Kadhi et al., 2017
9	--		132.1019	0.0	20 eV: 121, <b>119</b> , 118	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	Leucine	Wang et al., 1978
10	1.04	--	182.0813	0.55	10 eV: 165, 147, <b>136</b> , 123, 119, 107	C <sub>9</sub> H <sub>11</sub> NO <sub>3</sub>	Tyrosine	Metlin
11	1.76	--	152.0567	0.0	20 eV: 136, <b>135</b> , 134	C <sub>5</sub> H <sub>5</sub> N <sub>5</sub> O	Guanine	Mu et al., 2012
12	1.98	164.0717	166.0862	0.0/-0.6	20 eV: 147, 121, 111, <b>103</b> / 20 eV: 131, 124, 121, <b>120</b> , 107, 103	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	Phenylalanine	Metlin
13	2.00	282.0845	284.0992	0.35/1.0	20 eV: 164, <b>150</b> , 133, 108 / 10 eV: 243, 243, 211, 193, 172, <b>152</b> , 135, 104	C <sub>10</sub> H <sub>13</sub> N <sub>5</sub> O <sub>5</sub>	Guanosine	Abu Reidah et al., 2015a; Metlin
14	3.40	309.1092	311.1238	0.0/0.0	10 eV: 291, 247, 180, 167, <b>128</b> / 10 eV: 293, <b>182</b> , 165, 147, 145, 136	C <sub>14</sub> H <sub>18</sub> N <sub>2</sub> O <sub>6</sub>	Glutamyl-tyrosine I	Harrison, 2004
15	3.70	309.1093	311.1239	0.32/0.32	10 eV: 291, 247, 180, 167, <b>128</b> / 10 eV: 293, <b>182</b> , 165, 147, 145, 136	C <sub>14</sub> H <sub>18</sub> N <sub>2</sub> O <sub>6</sub>	Glutamyl-tyrosine II	Harrison, 2004
16	4.50	483.1256	485.1402	0.0/0.0	20 eV: 367, <b>303</b> , 259, 215, 187, 172, 146, 115 / 10 eV: 323, 305, 263, <b>146</b>	C <sub>20</sub> H <sub>24</sub> N <sub>2</sub> O <sub>12</sub>	N.I	--
17	--		188.0706	0.0	10 eV: 170, <b>146</b> , 142, 129, 118	C <sub>11</sub> H <sub>9</sub> NO <sub>2</sub>	N.I	--
18	4.55	203.0826	205.0973	0.0/0.48	10 eV: 185, 159, 142, 129, <b>116</b> / 10 eV: <b>188</b> , 170, 159, 146, 132, 10 eV: 371, 209, <b>191</b> , 173, 169, 147 / --	C <sub>11</sub> H <sub>12</sub> N <sub>2</sub> O <sub>2</sub>	Tryptophan	Metlin
19	4.83	389.0726	391.0874	0.25/0.76	10 eV: 337, 289, 245, 209, <b>191</b> , 173, 163, 147, 129	C <sub>15</sub> H <sub>18</sub> O <sub>12</sub>	N.I	--
20	5.06	355.0671	--	0.0	10 eV: 275, 257, 164, 147, <b>128</b> / 10 eV: 278, 232, <b>166</b> , 143, 130	C <sub>15</sub> H <sub>16</sub> O <sub>10</sub>	<i>p</i> -Coumaroylglucarate or <i>p</i> -Coumaroylgalactarate	Lin & Risch et al., 1988; Harnley, 2007; Coutinho et al., 2016
21	5.23	293.1143	295.1291	0.0/1.0		C <sub>14</sub> H <sub>18</sub> N <sub>2</sub> O <sub>5</sub>	Glutamyl-phenylalanine	Wang et al., 1978

N.	Rt (min)	[M - H] <sup>-</sup> MS	[M + H] <sup>+</sup> MS	Error (ppm) - / +	Fragments MS <sup>2</sup> [M-H] <sup>-</sup> / [M+H] <sup>+</sup>	Molecular Formula	Tentative identity	Reference
22	5.33	443.1921	--	-0.5	20 eV: <b>413</b> , 377, 299, 253, 195, 119, 101	C <sub>21</sub> H <sub>32</sub> O <sub>10</sub>	Penstemide	Abu Reidah et al., 2015b
23	5.45	563.1618	565.1764	0.0/0.17	20 eV: 501, 461, 431, <b>419</b> , 401, 269, 247, 191, 161, 125 / 5 eV: <b>433</b> , 271, 127	C <sub>23</sub> H <sub>32</sub> O <sub>16</sub>	N.I	--
24	5.55	449.1092	--	0.66	10 eV: <b>449</b> , 287, 269, 227, 170, 151, 125	C <sub>21</sub> H <sub>22</sub> O <sub>11</sub>	Eriodictyol- <i>O</i> -hexoside	Dueñas et al., 2012; Perestelo et al., 2012; Pereira et al., 2013
25		387.1660	--	-0.25	20 eV: <b>387</b> , 369, 231, 225, 207, 163, 119	C <sub>18</sub> H <sub>28</sub> O <sub>9</sub>	Hydroxyjasmonic acid - <i>O</i> -hexoside	Pereira et al., 2013
26		535.1090	--	-1.0	10 eV: <b>535</b> , 367, 341, 297, 193, 173, 129	C <sub>24</sub> H <sub>24</sub> O <sub>14</sub>	N.I	--
27	5.73	565.1199	--	0.0	10 eV: 565, <b>519</b> , 367, 327, 265, 173, 129	C <sub>25</sub> H <sub>26</sub> O <sub>15</sub>	N.I	--
28		595.1302	--	-0.5	10 eV: <b>595</b> , 397, 371, 327, 191, 173, 129	C <sub>26</sub> H <sub>28</sub> O <sub>16</sub>	N.I	--
29		163.0403	--	1.22	20 eV: 153, 138, <b>119</b> , 103	C <sub>9</sub> H <sub>8</sub> O <sub>3</sub>	<i>p</i> -Coumaric acid	Perestelo et al., 2012; Verardo et al., 2015
30	5.85	305.0700	--	0.0	20 eV: 287, <b>225</b> , 214, 163, 143, 131, 109	C <sub>12</sub> H <sub>18</sub> O <sub>7</sub> S	Hydroxyjasmonic acid - <i>O</i> -sulfate	Farag et al., 2017
31		461.1088	--	-0.21	10 eV: 415, <b>253</b>	C <sub>22</sub> H <sub>22</sub> O <sub>11</sub>	Daidzein- <i>O</i> -hexoside derivative	Chen et al., 2005; Cavaliere et al., 2007
32	6.11	431.0984	--	0.0	25 eV: 341, <b>269</b> , 283, 224, 179, 125	C <sub>21</sub> H <sub>20</sub> O <sub>10</sub>	Genistein 7- <i>O</i> -glucoside (Genistin)	March et al., 2004; Dueñas et al., 2012; Verardo et al., 2015
33		477.1040	--	0.0	10 eV: <b>431</b> , 311, 269	C <sub>22</sub> H <sub>22</sub> O <sub>12</sub>	Genistin derivative	Chen et al., 2005; Cavaliere et al., 2007
34	6.26	457.1143	--	0.65	--	C <sub>23</sub> H <sub>22</sub> O <sub>10</sub>	Daidzein acetylglycoside (Acetyl daidzin)	Chen et al., 2005; Cavaliere et al., 2007; Dueñas et al., 2012; Verardo et al., 2015; Chen et al., 2005; Cavaliere et al., 2007; Dueñas et al., 2012; Verardo et al., 2015
35	6.50	473.1089	--	0.0	--	C <sub>23</sub> H <sub>22</sub> O <sub>11</sub>	Genistein acetylglycoside (Acetylgenistin)	Chen et al., 2005; Cavaliere et al., 2007; Dueñas et al., 2012; Verardo et al., 2015
36	6.63	253.0509	255.0652	1.2/0.0	20 eV: <b>253</b> , 225, 224, 223, 209, 197, 196, 180, 167, 152, 133, 132, 117 / 20 eV: <b>255</b> , 227, 199, 181, 171, 157, 153, 137, 119	C <sub>15</sub> H <sub>10</sub> O <sub>4</sub>	Daidzein	Chen et al., 2005; Cavaliere et al., 2007; Kang et al., 2007; Dueñas et al., 2012; Verardo et al., 2015
37	6.70	283.0613	--	0.35	5 eV: <b>283</b> , 268, 240, 208, 186, 164, 151	C <sub>16</sub> H <sub>12</sub> O <sub>5</sub>	Glycitein	Chen et al., 2005; Cavaliere et al., 2007; Dueñas et al., 2012; Verardo et al., 2015
38	6.86	269.0455	271.0603	0.0/0.73	20 eV: <b>269</b> , 252, 239, 241, 224, 225, 201, 197, 181, 180, 170, 171, 159, 133, 117 / 20 eV: <b>271</b> , 253, 243, 229, 215, 197, 187, 165, 153, 135	C <sub>15</sub> H <sub>10</sub> O <sub>5</sub>	Genistein	Chen et al., 2005; Cavaliere et al., 2007; Kang et al., 2007; Dueñas et al., 2012; Verardo et al., 2015; Metlin

N.	Rt (min)	[M - H] <sup>-</sup> MS	[M + H] <sup>+</sup> MS	Error (ppm) - / +	Fragments MS <sup>2</sup> [M-H] <sup>-</sup> / [M+H] <sup>+</sup>	Molecular Formula	Tentative identity	Reference
39	7.02	--	230.2479	0.43	20 eV: <b>230</b> , 212, 175, 159, 132, 117	C <sub>14</sub> H <sub>31</sub> NO	Xestoaminol C	Huang et al., 2016
40	--	--	274.2742	0.36	20 eV: <b>274</b> , 256, 230, 212, 161, 106	C <sub>16</sub> H <sub>35</sub> NO <sub>2</sub>	Hexadecaspheganine	Huang et al., 2016
41	7.33	329.2334	--	0.3	25 eV: 311, 293, 229, <b>211</b> , 199, 171, 129	C <sub>18</sub> H <sub>34</sub> O <sub>5</sub>	Hydroxyoctadecanedioic acid	Gómez-Romero et al., 2010
42	8.04	987.5167	--	--	--	--	N.I	--
43	8.04	941.5123	--	--	--	--	N.I	--
44	8.30	295.2278	--	-0.33	20 eV: 277, 251, 222, 195, 171, 141, 113	C <sub>18</sub> H <sub>32</sub> O <sub>3</sub>	Hydroxyoctadecadienoic acid	Gómez-Romero et al., 2010

-- :Not obtained; N.I: Not identified.

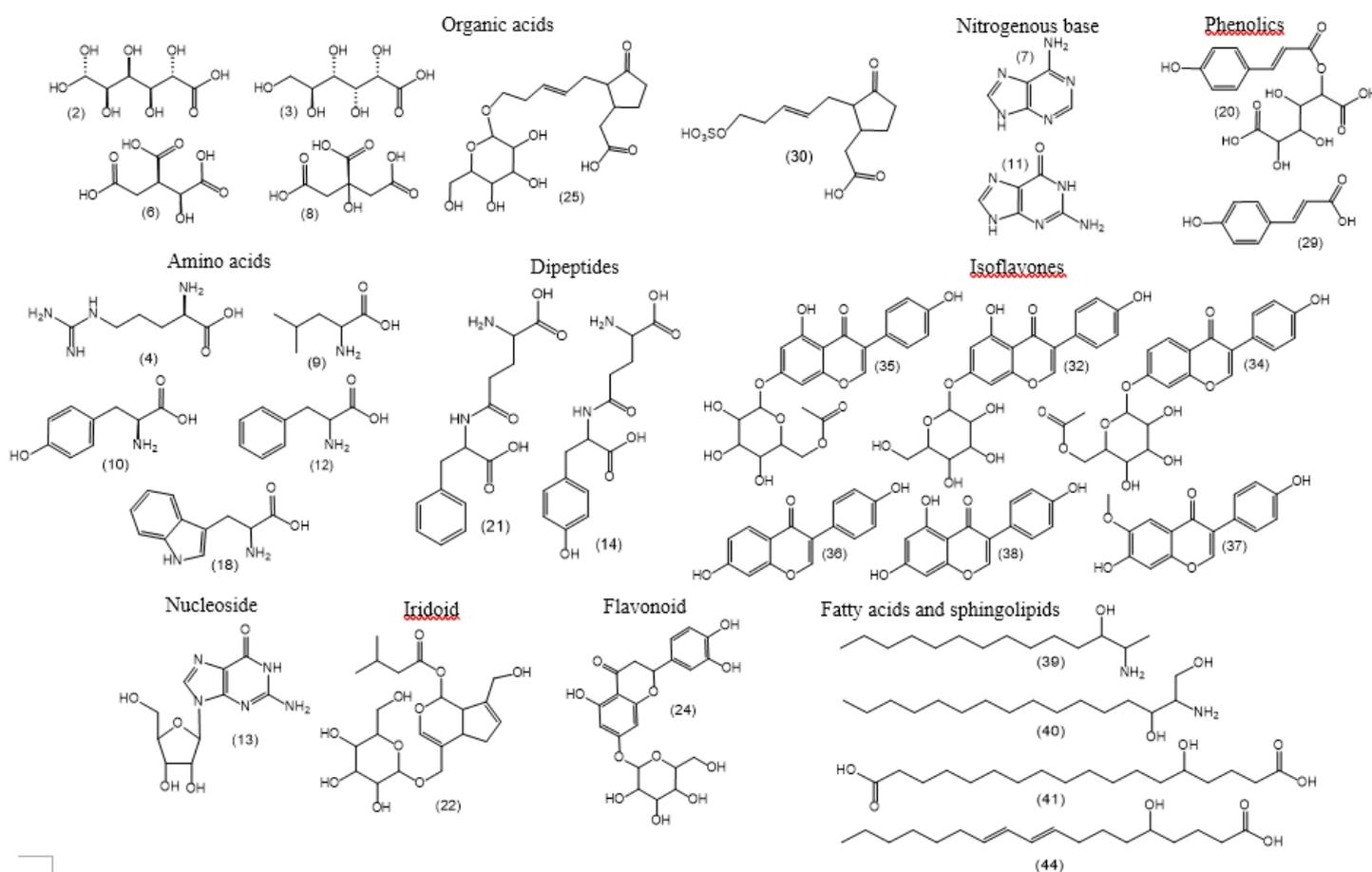


Figure 3. Compounds identified in soybeans from warehouses 1, 2, and 3 by LC-MS-ESI

One flavonoid monohexoside, was observed in the soybeans investigated. The molecular ion [M-H]<sup>-</sup> at *m/z* 449 was identified as eriodictiol-*O*-hexoside (24). This flavonoid showed a loss of 162 Da from the cleavage between hexose and aglycone resulting in the fragment [M-H-Hexose]<sup>-</sup> at *m/z* 287. Dueñas et al. (2012) found the eriodictyol glycoside in soybean seeds that underwent different microbiological fermentations. In studies with soybeans by Cavaliere et al. (2007), Lee et al. (2008) and Guzmán-Ortiz et al. (2017), this metabolite was not observed. The flavonoids kaempferol, naringenin, hesperetin, and their heterosides,

identified in some studies with soybeans (Cavaliere et al., 2007; Lee et al., 2008; Dueñas et al., 2012), were not found in this work, and Verardo et al. (2015) did not identify any of the aforementioned flavonoids. Catechin and epicatechin found in the works of Lee et al. (2008), Nam et al. (2014) and Guzmán-Ortiz et al. (2017), were not observed in this study, nor in those of Cavaliere et al. (2007) and Dueñas et al. (2012).

Compounds 32, 36, 37, and 38 showed the molecular ions  $[M-H]^-$  at  $m/z$  431,  $m/z$  253,  $m/z$  283, and  $m/z$  269, respectively. The fragmentation spectra corresponded to the isoflavones genistin (genistein-7-*O*-glucoside) (32), daidzein (36), glycitein (37), and genistein (37), respectively. The presence of these compounds has already been described in different soy cultivars (Lozovaya et al., 2005; Lin & Harnly, 2007; Dueñas et al., 2012; Verardo et al., 2015). In the isoflavone region in the chromatogram (6.0 min - 7.0 min), the molecular ions  $[M-H]^-$  at  $m/z$  457 and  $m/z$  473 were related to the formulas  $C_{23}H_{22}O_{10}$  (Error 0.65 ppm) and  $C_{23}H_{22}O_{11}$  (Error 0.0 ppm), and to the compounds daidzein acetylglycoside (34) and genistein acetylglycoside (35), respectively. Although these isoflavones did not show their fragment spectra at any collision energy, it was possible to infer their structures, since the exact mass enables the suggestion of the molecular formula and these compounds are well known components of soybeans. In addition, the elution order observed for isoflavones is consistent with other studies in the literature (Chen et al., 2005; Cavaliere et al., 2007; Dueñas et al., 2012; Verardo et al., 2015).

Studies have indicated that the presence and concentration of isoflavones in soybeans depend on environmental and genetic factors. Significant differences in the content of genistein, daidzein, and glycitein were found when variations in soybean type, soil type, and temperature were evaluated. Under controlled conditions, temperature was the factor that influenced the content of isoflavones the most. In general, low temperatures promoted higher levels of isoflavone, and the amplitude of the response depended on the type of cultivar analyzed (Lozovaya et al., 2005). In another study with ten soybean cultivars in two regions of Korea, relevant differences in the content of isoflavones, phenols, and antioxidant activity were observed, depending on the location and type of cultivar. Tryptophan, epicatechin, daidzin, and genistin were the major compounds found in soybean methanol extracts (Nam et al., 2014).

The same qualitative variation observed in the chemical composition of soybeans regarding flavonoids was also found for phenolic compounds when other works in the literature were compared (Cavaliere et al., 2007; Dueñas et al., 2012; Lee et al., 2014; Nam et al., 2014; Verardo et al., 2015; Guzmán-Ortiz et al., 2017). The *p*-Cumaryl (28) acid and *p*-cumaroylglucarate or *p*-cumaroylgalactarate (19) were the phenolic compounds identified in this study. The *p*-Cumaryl acid is a phenolic compound commonly found in soybeans (Lee et al., 2008; Dueñas et al., 2012; Verardo et al., 2015; Guzmán-Ortiz et al., 2017). Regarding *p*-cumaroylglucarate or *p*-cumaroylgalactarate (19), no reports of this compound in the chemical composition of soybeans were found. This compound presented the molecular ion  $[M-H]^-$  at  $m/z$  355 and according to Lin & Harnley (2007) and Coutinho et al. (2016) it was identified as a cumaroyl ester of the glucaric or galactaric acid, the loss of 164 Da resulting from a McLafferty rearrangement in the ester group of the main fragment at  $m/z$  191

[M-H-Cumaroyl]<sup>-</sup>. The *p*-Cumaroylglucarate has already been identified in the hydromethanolic extracts from sugar cane leaves (Coutinho et al., 2016) and orange peels (Lin & Harnley, 2007).

When the various chemical characterization studies with soybeans were compared, qualitative and quantitative variations in the grain composition were observed. Therefore, it is possible to suggest that there is no single composition for soybeans. Environmental and genetic factors strongly affect these variations (Lozovaya et al., 2005; Lee et al., 2008; Nam et al., 2014; Freiria et al., 2016). Thus, climatic conditions, seasonality, soil quality, and growth stage can promote differences in the chemical composition of plants (Dhifi et al., 2016). In addition, the extraction process and all experimental conditions of analysis may have an effect on the presence or absence of a compound or classes of compounds in the chemical composition of a plant. This is the first time that the chemical composition of soy produced in the Triângulo Mineiro region has been reported.

### 3.2 Mass Spectrometry Analysis by Direct Injection and Chemometrics of Soybean Samples From Warehouses 1, 2, and 3

The experiments by direct injection were carried out by directly injecting the methanolic solutions into the mass spectrometer with a continuous flow pumped by the HPLC. In this way, the acquisition of all ions present in the sample in the positive and negative modes was obtained without the previous chromatographic procedure. Figures 4 and 5, below, show the mass spectra containing the soy chemical profile, in positive and negative modes, for each warehouse analyzed.

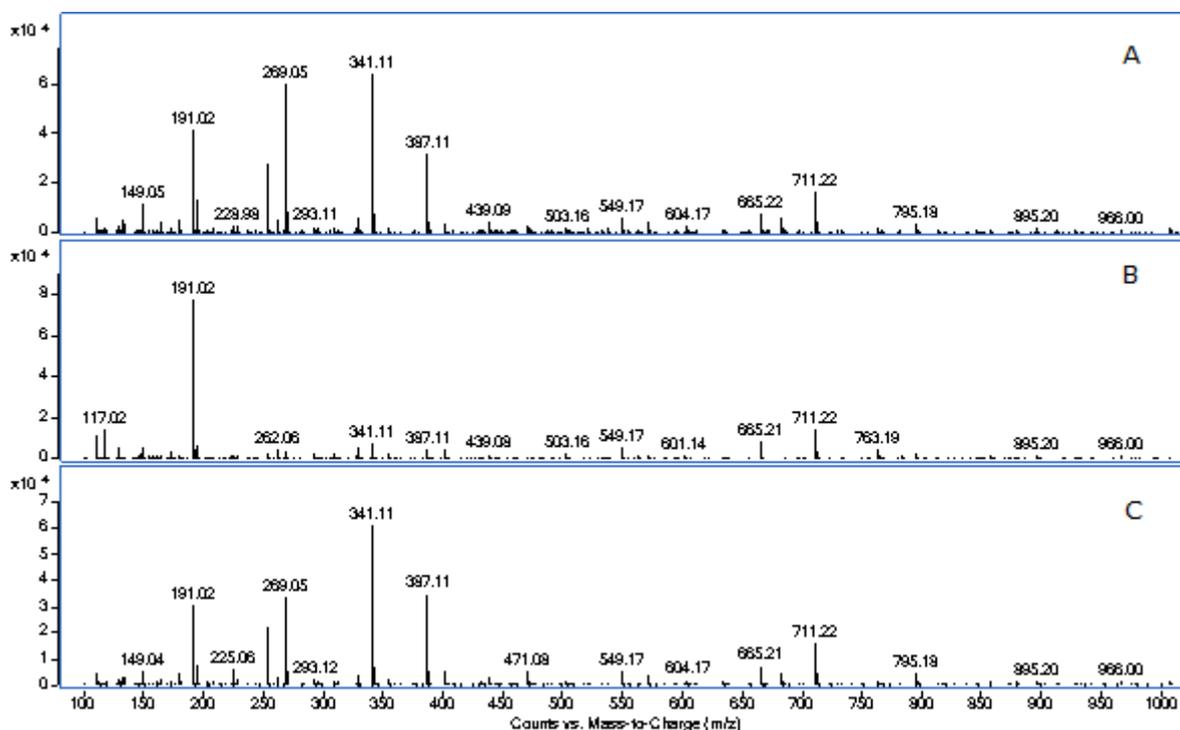


Figure 4. Mass spectra obtained for soybeans in negative mode for (A) warehouse 1, (B) warehouse 2, and (C) warehouse 3

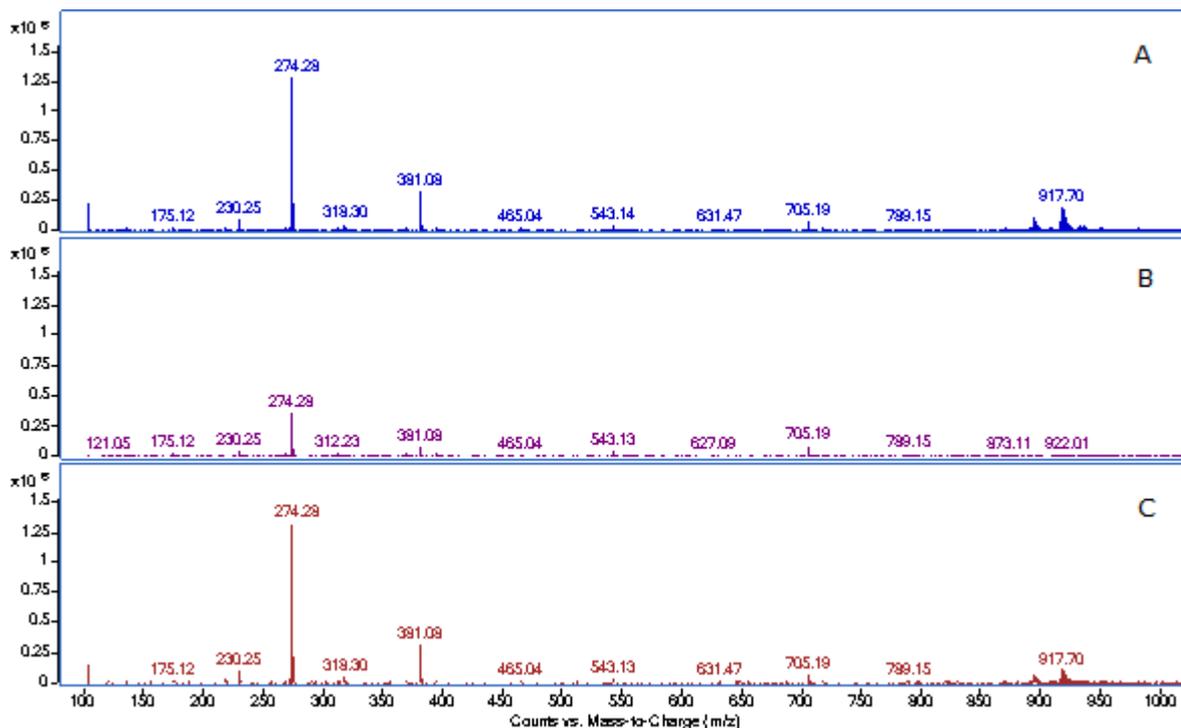


Figure 5. Mass spectra obtained for soybeans in positive mode for (A) warehouse 1, (B) warehouse 2, and (C) warehouse 3

The results generated by LC-EM-ESI, in the negative and positive modes, initially suggested that the major chemical constituents or the components most sensitive to the mass spectrometry technique were the same in all of the three studied warehouses. However, through chemometrics, it was possible to verify that there is a variance between the soybeans collected in the different warehouses.

The results below refer to the PCA analysis with the spectral matrix recorded in the negative mode. Figure 6 presents the plot graph of the PCA scores of the three major Principal Components (PC), showing that they explained a variance of 98.79%, capturing almost all the information from the original data in the dimensional reduction. There is a well-defined separation between the three groups of samples, except for sample 20, which distanced itself from its original group but without interfering in the other ones. This result indicates a great difference in the chemical composition between the grains of warehouses 1, 2, and 3. Figure 7 presents the Loading plot of PC1, a graph that informs the  $m/z$  values used in the separation of the groups observed in Figure 6. Figure 8 presents the HCA graph which groups the samples according to their similarities. Thus, the closer to 100% the linking between two samples, the more chemically similar are their compositions.

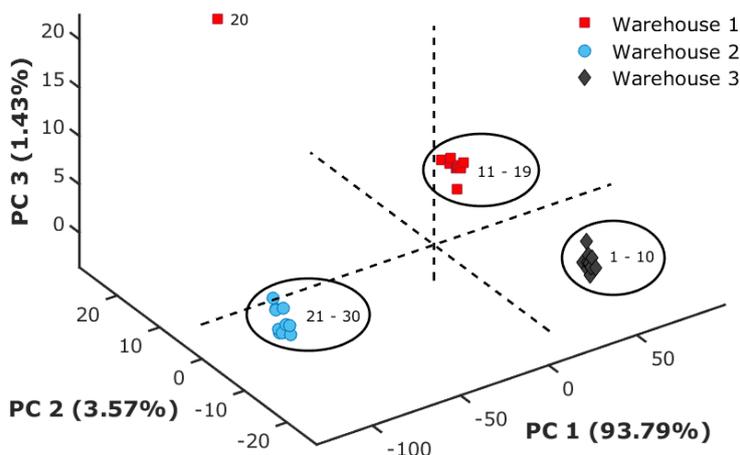


Figure 6. 3D PCA of soybean samples for ions in negative mode

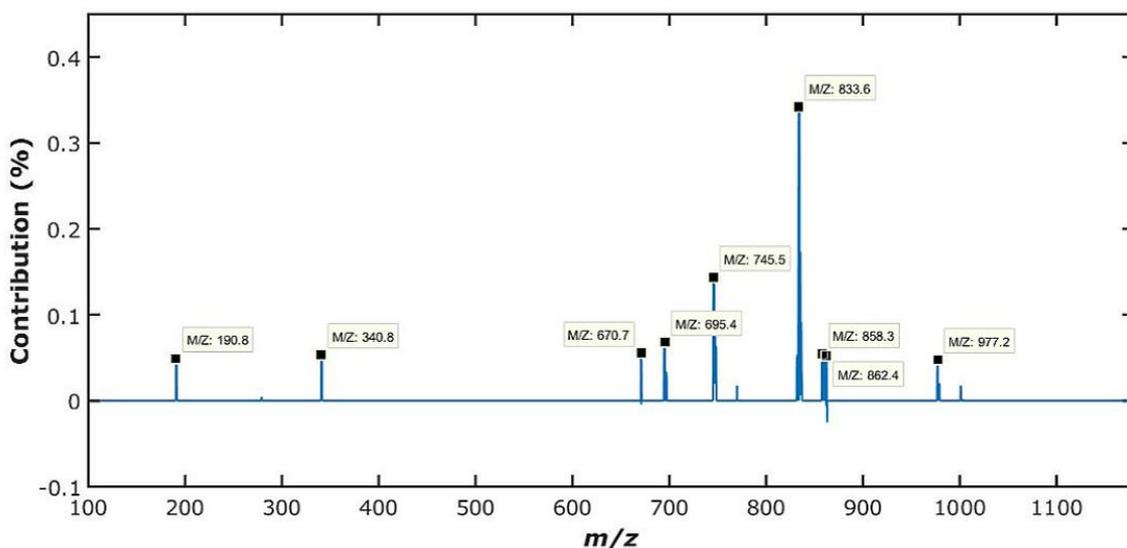


Figure 7. Loading plot of PC1, showing group separation (negative mode)

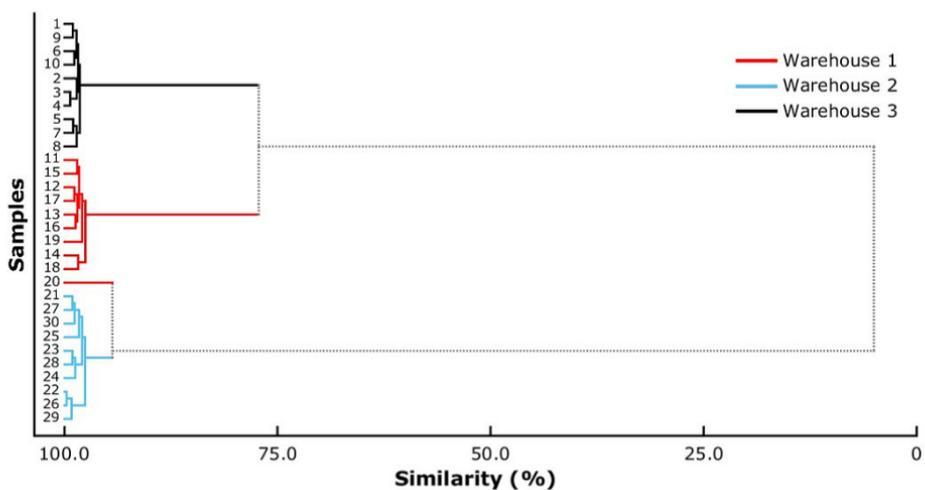


Figure 8. HCA graph for soy samples from warehouses 1, 2, and 3 (negative mode)

The HCA analysis shows that the soy samples collected in warehouses 1, 2, and 3 follow the pattern shown in the PCA, again exhibiting different chemical characteristics for the soybeans of each warehouse. It is also noticed that the grains from warehouses 1 and 3 have a higher similarity in their composition, differing from those of warehouse 2.

The molecular ions  $[M-H]^-$  of  $m/z$  191, 341, 671, 695, 746, 834, 858, 862, and 977 were related to the variance of soybean samples in the negative mode. The molecular ions at  $m/z$  191 and  $m/z$  341 have been identified by LC-MS-ESI and refer to the compounds isocitrate or citric acid, and sucrose, respectively.

The results below refer to the PCA analysis of the spectra in the positive mode (related to ions  $[M+H]^+$ ). Figure 9 shows the 3D PCA graph constructed with the three main components which explained variance of 94.04%. As in the PCA analysis of the data obtained in the negative mode, a distinction was made between the samples of each warehouse and a small difference between the samples that form each group can be observed. This suggests that the better compounds ionized in a positive way (generally bases of Brönsted) have a variation regarding their presence or concentration in each seed, regardless of the warehouse. Figure 10 shows the main ions with  $m/z$   $[M+H]^+$  responsible for the variation between the soy samples from the three warehouses in PC2. Figure 11 represents the HCA graph, which shows the degree of similarity between the samples, and presents the similarity between the seeds from warehouses 2 and 3 regarding their contents.

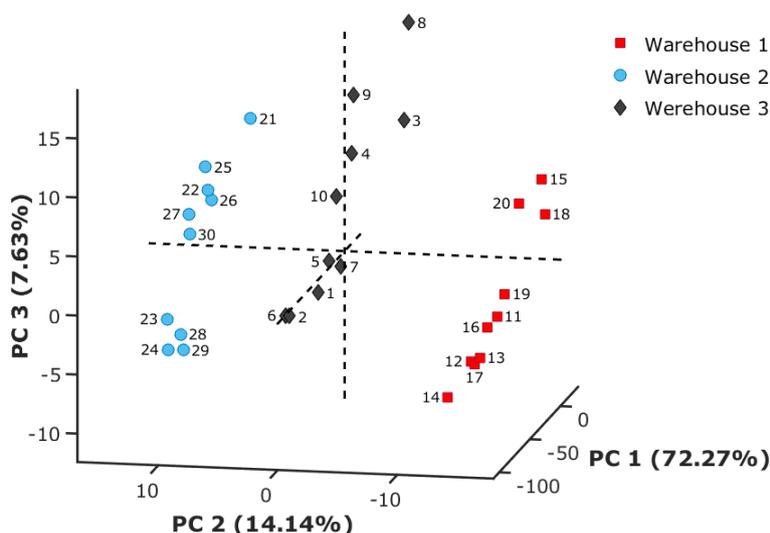


Figure 9. 3D PCA of soybean samples for ions in positive mode

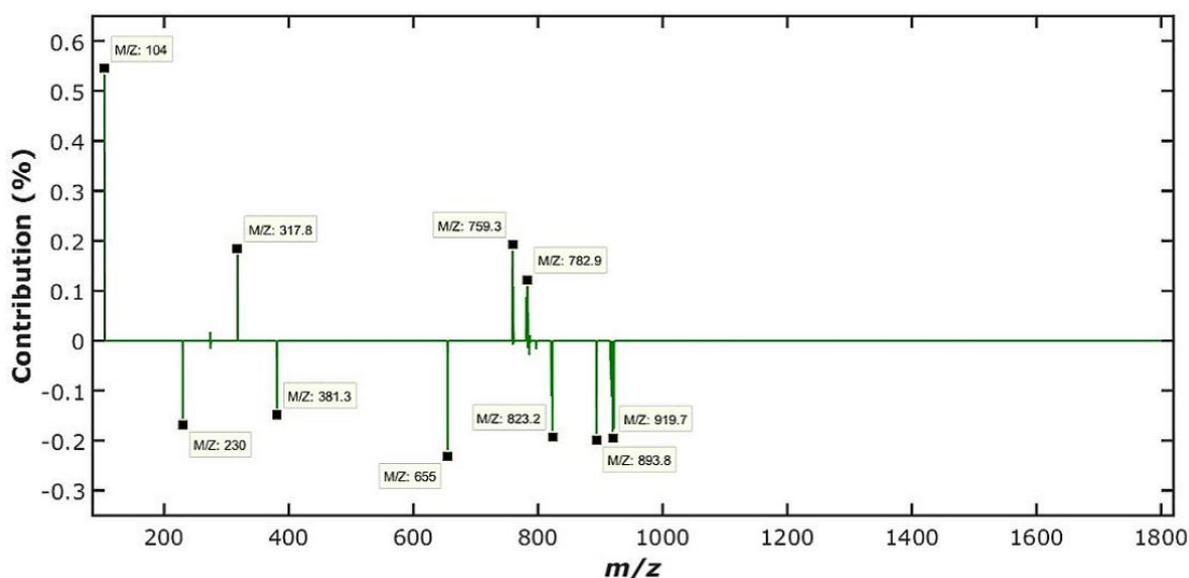


Figure 10. Loading plot of PC2, showing group separation (positive mode)

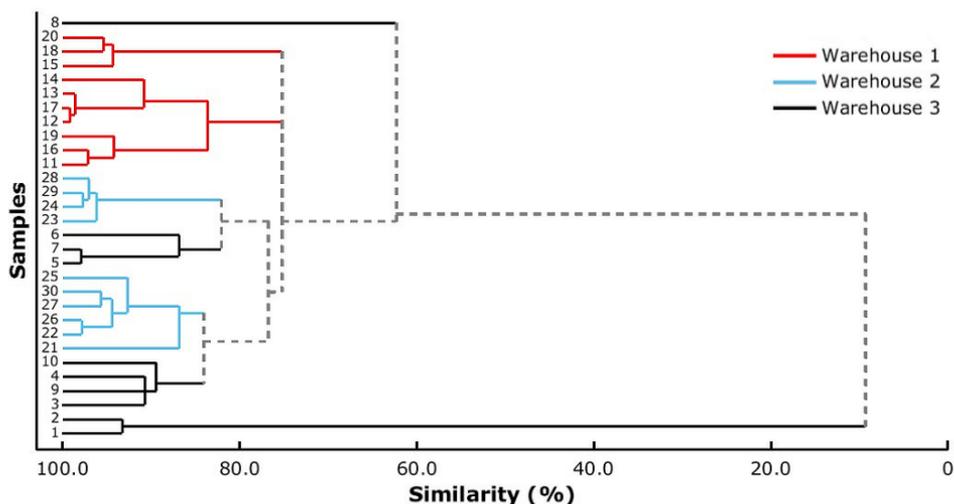


Figure 11. HCA for soy samples from warehouses 1, 2, and 3 (positive mode)

The mass spectrometry experiments by direct injection and by the PCA and HCA of chemometric analyzes showed that the three studied warehouses have different chemical constitutions.

The molecular ions  $[M+H]^+$  that promoted the variance between the soybeans from warehouses 1, 2, and 3 were those at  $m/z$  104, 230, 318, 655, 759, 783, 823, 893, and 920. The molecular ion  $[M+H]^+$  at  $m/z$  230 has been previously identified by LC-MS-ESI as xestoaminol C (Table 1), a sphingolipid (Huang et al., 2016). The molecular ions of  $m/z$  759, 783, and 823 are related to the molecular formulas  $C_{43}H_{87}N_2O_6P$ ,  $C_{45}H_{87}N_2O_6P$ , and  $C_{48}H_{91}N_2O_6P$ , respectively. According to the works of Fred & Tinoco (2015); Huang et al. (2016), and Qu et al. (2018), it is possible to suggest that these compounds are also sphingolipids classified as sphingomyelins. It is noteworthy here that these non-polar high

molecular mass compounds were observed only when the experiment was carried out by direct injection of soybean extracts in the mass spectrometer. The chromatographic conditions used in the LC-MS-ESI experiments hindered the observation of these compounds.

#### 4. Conclusion

The soy produced in the Triângulo Mineiro/MG region is a source of different classes of metabolites, with emphasis on phenolic compounds, flavonoids, and isoflavones. The methodology employed does not allow us to identify all the chemical components present, since the extraction and analytical methods, in addition to environmental conditions, soil composition, type of cultivar, and growth stage, may affect the chemical composition of soybeans. Nevertheless, the chemical profile of the grains produced in this region, which was previously unknown, was determined. This profile was proposed using the extraction with methanol, a polar solvent. The soybeans from the three warehouses investigated showed similar qualitative chemical profiles when analyzed by liquid chromatography coupled with mass spectrometry. However, the PCA and HCA analyses revealed that positive and negative ions contributed to the variance in the chemical composition between the soybeans of each warehouse.

#### Acknowledgments

The authors thank the Foundation for Research Support of the Minas Gerais State (FAPEMIG) and Coordination for the Improvement of Higher Education Personnel (CAPES). We also thank the National Supply Company (CONAB) and the warehouses (CASENG, CERTRIM and GLOBAL) for providing the soy samples.

#### References

- Abu-Reidah, I. M., Ali-Shtayeh, M. S., Jamous, R. M., Arráez-Román, D., & Segura-Carretero, A. (2015a). Comprehensive metabolite profiling of *Arum palaestinum* (Araceae) leaves by using liquid chromatography–tandem mass spectrometry. *Food Research International*, 70, 74-86. <https://doi.org/10.1016/j.foodres.2015.01.023>
- Abu-Reidah, I. M., Ali-Shtayeh, M. S., Jamous, R. M., Arráez-Román, D., & Segura-Carretero, A. (2015b). HPLC–DAD–ESI-MS/MS screening of bioactive components from *Rhus coriaria* L. (Sumac) fruits. *Food Chemistry*, 166, 179-191. <https://doi.org/10.1016/j.foodchem.2014.06.011>
- Bridges Junior, C. C. (1966). Hierarchical cluster analysis. *Psychological Reports*, 18(3), 851-854. <https://doi.org/10.2466/pr0.1966.18.3.851>
- Carvalho, E. R., Mavaieie, D. P. R., Oliveira, J. A., Carvalho, M. V., & Vieira, A. R. (2014). Alterações isoenzimáticas em sementes de cultivares de soja em diferentes condições de armazenamento. *Pesquisa Agropecuária Brasileira*, 49(12), 967-976. <https://doi.org/10.1590/S0100-204X2014001200007>
- Cavaliere, C., Cucci, F., Foglia, P., Guarino, C., Samperi, R., & Laganá, A. (2007). Flavonoid profile in soybeans by high-performance liquid chromatography/tandem mass spectrometry.

*Rapid Communications in Mass Spectrometry*, 21(14), 2177-2187.

<https://doi.org/10.1002/rcm.3049>

Chen, L. J., Zhao, X., Plummer, S., Tang, J., & Games, D. E. (2005). Quantitative determination and structural characterization of isoflavones in nutrition supplements by liquid chromatography-mass spectrometry. *Journal Chromatography A*, 1082, 60-70.

<https://doi.org/10.1016/j.chroma.2005.03.066>

Choct, M., Dersjant-Li, Y., McLeish, J., & Peisker, M. (2010). Soy oligosaccharides and soluble non-starch polysaccharides: a review of digestion, nutritive and anti-nutritive effects in pigs and poultry. *Asian-Australasian Journal of Animal Sciences*, 23(10), 1386-1398.

<https://doi.org/10.5713/ajas.2010.90222>

Chung, I. M., Oh, J. Y., & Kim, S. H. (2017). Comparative study of phenolic compounds, vitamin E, and fatty acids compositional profiles in black seed-coated soybeans (*Glycine Max* L. Merrill) depending on pickling period in brewed vinegar. *Chemistry Central Journal*, 11(64), 1-11. <https://doi.org/10.1186/s13065-017-0298-9>

CONAB. (2018). Companhia Nacional de Abastecimento. Safra 2018/2019. Perspectivas para a agropecuária: Prévia. [Online] Available:

<https://www.conab.gov.br/perspectivas-para-a-agropecuaria>. (Mai 5, 2020).

CONAB. (2019). Companhia Nacional de Abastecimento. Safra 2018/2019: Quarto levantamento. Acompanhamento da Safra Brasileira: grãos. [Online] Available:

<https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos?start=20>. (Mai 5, 2020).

CONAB. (2020). Companhia Nacional de Abastecimento. Safra 2019/2020: Sétimo levantamento. Acompanhamento da Safra Brasileira: grãos. [Online] Available:

<https://www.conab.gov.br/info-agro/safras/graos/boletim-da-safra-de-graos?start=10>. (Mai 5, 2020).

Coutinho, I. D., Baker, J. M., Ward, J. L., Beale, M. H., Creste, S., & Cavalheiro, A. J. (2016). Metabolite profiling of sugarcane genotypes and identification of flavonoid glycosides and phenolic acids. *Journal of Agricultural and Food Chemistry*, 64(21), 4198-4206.

<https://doi.org/10.1021/acs.jafc.6b01210>

Creelman, R. A., Tierney, M. L., & Mullet, J. E. (1992). Jasmonic acid/methyl jasmonate accumulate in wounded soybean hypocotyls and modulate wound gene expression. *Proceedings of the National Academy of Sciences*, 89(11), 4938-4941.

<https://doi.org/10.1073/pnas.89.11.4938>

Dall'agnol, A. (2016). Embrapa soja no contexto do desenvolvimento da soja no Brasil: histórico e contribuições (1st ed.). Brasília, DF: Embrapa.

Dhifi, W., Bellili, S., Jazi, S., Bahloul, N., & Mnif, W. (2016). Essential oils' chemical characterization and investigation of some biological activities. *A Critical Review Medicines*, 3(4), 1-16. <https://doi.org/10.3390/medicines3040025>

- Dueñas, M., Hernandez, T., Robredo, S., Lamparski, G., Estrella, I., & Munoz, R. (2012). Bioactive phenolic compounds of soybean (*Glycine max* cv. Merit): modifications by different microbiological fermentations. *Polish Journal of Food and Nutrition Sciences*, 62(4), 241-250. <https://doi.org/10.2478/v10222-012-0060-x>
- Farag, M. A., Ali, S. E., Rodaya, R. H., El-Seedi, H. R., Sultani, H. N., Laub, A., ... Wessjohann, L. A. (2017). Phytochemical profiles and antimicrobial activities of *Allium cepa* Red cv. and *A. sativum* subjected to different drying methods: a comparative MS-based metabolomics. *Molecules*, 22(5), 1-18. <https://doi.org/10.3390/molecules22050761>
- Forbes, M. W., Jockusch, R. A., Young, A. B., & Harrison, A. G. (2007). Fragmentation of protonated dipeptides containing arginine. Effect of activation method. *Journal of the American Society for Mass Spectrometry*, 18(11), 1959-1966. <https://doi.org/10.1016/j.jasms.2007.08.003>
- Fredi, A. R. O., & Tinoco, L. W. (2015). Esfingolipídios: metabólitos bioativos e modelos para o planejamento de fármacos. *Revista Virtual de Química*, 7(4), 1384-1401. <https://doi.org/10.5935/1984-6835.20150076>
- Freiria, G. H., Lima, W. F., Leite, R. S., Mandarino, J. M. G., Silva, J. B., & Prete, C. E. C. (2016). Productivity and chemical composition of food-type soybeans sown on different dates. *Acta Scientiarum. Agronomy*, 38(3), 371-377. <https://doi.org/10.4025/actasciagron.v38i3.28632>
- Ghahari, S., Alinezhad, H., Nematzadeh, G. A., Tajbakhsh, M., & Baharfar, R. (2017). Chemical composition, antioxidant and biological activities of the essential oil and extract of the seeds of *Glycine max* (Soybean) from North Iran. *Current Microbiology*, 74(4), 522-531. <https://doi.org/10.1007/s00284-016-1188-4>
- Gómez-Romero, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2010). Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry*, 71(16), 1848-1864. <https://doi.org/10.1016/j.phytochem.2010.08.002>
- Guerreiro, T. M., Oliveira, D. N., Ferreira, M. S., & Catharino, R. R. (2014). High-throughput analysis by SP-LDIMS for fast identification of adulterations in commercial balsamic vinegars. *Analytica Chimica Acta*, 838, 86-92. <https://doi.org/10.1016/j.aca.2014.06.009>
- Guzmán-Ortiz, F. A., Martín-Martínez, E. S., Valverde, M. E., Rodríguez-Aza, Y., Berrios, J., & Mora-Escobedo, R. (2017). Profile analysis and correlation across phenolic compounds, isoflavones and antioxidant capacity during germination of soybeans (*Glycine max* L.). *CyTA-Journal of Food*, 15(4), 516-524. <https://doi.org/10.1080/19476337.2017.1302995>
- Harrison, A. G. (2004). Characterization of  $\alpha$ - and  $\gamma$ -glutamyl dipeptides by negative ion collision-induced dissociation. *Journal of Mass Spectrometry*, 39(2), 136-144. <https://doi.org/10.1002/jms.515>
- He, F. J., & Chen, J. Q. (2013). Consumption of soybean, soy foods, soy isoflavones and breast cancer incidence: Differences between Chinese women and women in Western

countries and possible mechanisms. *Food Science and Human Wellness*, 2(3-4), 146-161. <https://doi.org/10.1016/j.fshw.2013.08.002>

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417, 1933. <https://doi.org/10.1037/h0071325>

Huang, H., Tong, T. T., Yau, L. F., Chen, C. Y., Mi, J. N., Wang, J. R., & Jiang, Z. H. (2016). LC-MS based sphingolipidomic study on A2780 human ovarian cancer cell line and its taxol-resistant strain. *Scientific Reports*, 6, 34684. <https://doi.org/10.1038/srep34684>

Jooyandeh, H. (2011). Soy products as healthy and functional foods. *Middle-East Journal of Scientific Research*, 7(1), 71-80.

[https://www.researchgate.net/publication/284893010\\_Soy\\_products\\_as\\_healthy\\_and\\_functional\\_foods](https://www.researchgate.net/publication/284893010_Soy_products_as_healthy_and_functional_foods)

Kadhi, O. M., Melchini, A., Mithen, R., & Saha, S. (2017). Development of a LC-MS/MS Method for the Simultaneous Detection of Tricarboxylic Acid Cycle Intermediates in a Range of Biological Matrices. *Journal of Analytical in Chemistry*, 2017, 1-12.

<https://doi.org/10.1155/2017/5391832>

Kang, J., Badger, T. M., Ronis, M. J. J., & Wu, X. (2010). No-isoflavone phytochemicals in soy and their health effects. *Journal of Agricultural and Food Chemistry*, 58(14), 8119-8133. <https://doi.org/10.1021/jf100901b>

Kang, J., Hick, L. A., & Price, W. E. (2007). A fragmentation study of isoflavones in negative electrospray ionization by MS<sup>n</sup> ion trap mass spectrometry and triple quadrupole mass spectrometry. *Rapid Communications in Mass Spectrometry*, 21(6), 857-868.

<https://doi.org/10.1002/rcm.2897>

Ko, K. P. (2013). Isoflavones: Chemistry, Analysis, Functions and Effects on Health and Cancer. *Asian Pacific Association for Cancer Education*, 15(17), 7001-7010. <https://doi.org/10.7314/APJCP.2014.15.17.7001>

Lee, S. J., Kim, J. J., Moon, H. I., Ahn, J. K., Chun, S. C., Jung, W. S., ... Chung, I. M. (2008). Analysis of isoflavones and phenolic compounds in Korean soybean [*Glycine max* (L.) Merrill] seeds of different seed weights. *Journal of Agricultural and Food Chemistry*, 56(8), 2751-2758. <https://doi.org/10.1021/jf073153f>

Li, S., Lin, Z., Jiang, H., Tong, L., Wang, H., & Chen, S. (2016). Rapid identification and assignation of the active ingredients in fufang banbianlian injection using HPLC-DAD-ESI-IT-TOF-MS. *Journal of Chromatographic Science*, 54(7), 1225-1237. <https://doi.org/10.1093/chromsci/bmw055>

Lin, L. Z., & Harnly, J. M. (2007). A screening method for the identification of glycosylated flavonoids and other phenolic compounds using a standard analytical approach for all plant materials. *Journal of Agricultural and Food Chemistry*, 55(4), 1084-1096. <https://doi.org/10.1021/jf062431s>

Lozovaya, V. V., Lygin, A. V., Ulanov, A. V., Nelson, R. L., Daydé, J., & Widholm, J. M.

- (2005). Effect of temperature and soil moisture status during seed development on soybean seed isoflavone concentration and composition. *Crop Science*, 45(5), 1934-1940. <https://doi.org/10.2135/cropsci2004.0567>
- Mahmoud, A. M., Yang, W., & Bosland, M. C. (2014). Soy isoflavones and prostate cancer: a review of molecular mechanisms. *The Journal of Steroid Biochemistry and Molecular Biology*, 140, 116-132. <https://doi.org/10.1016/j.jsbmb.2013.12.010>.
- Malenčić, D., Popović, M., & Miladinović, J. (2007). Phenolic content and antioxidant properties of soybean (*Glycine max* (L.) Merr.) seeds. *Molecules*, 12(3), 576-581. <https://doi.org/10.3390/12030576>
- MAPA. (2011). Ministério da Agricultura, Pecuária e Abastecimento. [Online] Available: <https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-vegetal/legislacao-1/normativos-cgqv/pocs/instrucao-normativa-no-60-de-22-de-dezembro-de-2011-milho/view>. (Mai 5, 2020).
- March, R. E., Miao, X. S., Metcalfe, C. D., Stobiecki, M., & Marczak, L. (2004). A fragmentation study of an isoflavone glycoside, genistein-7-O-glucoside, using electrospray quadrupole time-of-flight mass spectrometry at high mass resolution. *International Journal of Mass Spectrometry*, 232(2), 171-183. <https://doi.org/10.1016/j.ijms.2004.01.001>
- Martins, C. A. O., Sedyama, C. S., Oliveira, M. G. A., José, I. C., Moreira, M. A., Reis, M. S., & Rocha, V. S. (2000). Atividade da isocitrato-liase durante a germinação de sementes de soja. *Revista Brasileira de Sementes*, 22(1), 42-46. <http://dx.doi.org/10.17801/0101-3122/rbs.v22n1p42-46>
- Masike, K., Mhlongo, M. I., Mudau, S. P., Nobela, O., Ncube, E. N., Tugizimana, F., ... Madala, N. E. (2017). Highlighting mass spectrometric fragmentation differences and similarities between hydroxycinnamoyl-quinic acids and hydroxycinnamoyl-isocitric acids. *Chemistry Central Journal*, 11(29), 1-7. <https://doi.org/10.1186/s13065-017-0262-8>
- Metlin Database. [Online] Available: [https://metlin.scripps.edu/landing\\_page.php?pgcontent=mainPag](https://metlin.scripps.edu/landing_page.php?pgcontent=mainPag). (Oct 18, 2020).
- Mu, G., Luan, F., Xu, L., Hu, F., Liu, H., & Gao, Y. (2012). Determination of purines in soybean milk by capillary electrophoresis in comparison with high performance liquid chromatography. *Analytical Methods*, 4(10), 3386-3391. <https://doi.org/10.1039/C2AY25488C>
- Nam, J. H., Kang, S. W., Hong, S. Y., Kim, S. J., Jin, Y. I., Kim, H. S., ... Ok, H. C. (2014). Analysis of the phenolic content and antioxidant activities of soybean extracts from different regions and cultivars. *Korean Journal of Plant Research*, 27(6), 610-621. <http://dx.doi.org/10.7732/kjpr.2014.27.6.610>
- Pereira, O. R., Peres, A. M., Silva, A. M. S., Domingues, M. R. M., & Cardoso, S. M. (2013). Simultaneous characterization and quantification of phenolic compounds in *Thymus x citriodorus* using a validated HPLC–UV and ESI–MS combined method. *Food Research*

*International*, 54, 1773-1780. <https://doi.org/10.1016/j.foodres.2013.09.016>

Perestelo, R., Lu, Y., Santos, S. A. O., Silvestre, A. J. D., Neto, C. P., Câmara, J. S., & Rocha, S. M. (2012). Phenolic profile of sercial and tinta negra *Vitis vinifera* L. grape skins by HPLC–DAD–ESI–MS<sup>n</sup> - Novel phenolic compounds in *Vitis vinifera* L. grape. *Food Chemistry*, 135(1), 94-104. <https://doi.org/10.1016/j.foodchem.2012.04.102>

Qu, F., Zhang, H., Zhang, M., & Hu, P. (2018). Sphingolipidomic profiling of rat serum by UPLC-Q-TOF-MS: application to rheumatoid arthritis study. *Molecules*, 23(6), 1-14. <https://doi.org/10.3390/molecules23061324>

Risch, B., Herrmann, K., & Wray, V. (1988). (*E*)-*O*-p-coumaroyl-, (*E*)-*O*-feruoyl-derivatives of glucaric acid in citrus. *Phytochemistry*, 27(10), 3327-3329. [https://doi.org/10.1016/0031-9422\(88\)80059-1](https://doi.org/10.1016/0031-9422(88)80059-1)

Rizzo, G., & Baroni, L. (2018). Soy, soy foods and their role in vegetarian diets. *Nutrients*, 10(43), 1-51. <https://doi.org/10.3390/nu10010043>

Sullards, M. C., Lynch, D. V., Merrill Jr, A. H., & Adams, J. (2000). Structure determination of soybean and wheat glucosylceramides by tandem mass spectrometry. *Journal of Mass Spectrometry*, 35(3), 347-353. [https://doi.org/10.1002/\(SICI\)1096-9888\(200003\)35:3<347::AID-JMS941>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1096-9888(200003)35:3<347::AID-JMS941>3.0.CO;2-3)

Taira, H. (1966). Studies on amino acid contents of processed soybean. *Agricultural and Biological Chemistry*, 30(9), 847-855. <https://doi.org/10.1080/00021369.1966.10858695>

Tawarayaya, K., Horie, R., Shinano, T., Wagatsuma, T., Saito, K., & Oikawa, A. (2014). Metabolite profiling of soybean root exudates under phosphorus deficiency. *Soil Science and Plant Nutrition*, 60, 679-694. <https://doi.org/10.1080/00380768.2014.945390>

Teixeira, A. I., Ribeiro, L. F., Rezende, S. T., Barros, E. G., & Moreira, M. A. (2012). Development of a method to quantify sucrose in soybean grains. *Food Chemistry*, 130(4), 1134-1136. <https://doi.org/10.1016/j.foodchem.2011.07.128>

USDA – United States Department of Agriculture. Washington, DC, 2020. <https://apps.fas.usda.gov/psdonline/circulars/production.pdf>

Verardo, V., Riciputi, Y., Garrido-Frenich, A., & Caboni, M. F. (2015). Determination of free and bound phenolic compounds in soy isoflavone concentrate using a PFP fused core column. *Food Chemistry*, 185, 239-244. <https://doi.org/10.1016/j.foodchem.2015.03.090>

Wang, L. C., Cavins, J. F., & Wolf, W. J. (1978).  $\gamma$ -Glutamyl dipeptide content in soybean protein products and other commodities. *Journal of Food Science*, 43, 740-745. <https://doi.org/10.1111/j.1365-2621.1978.tb02406.x>

Wang, T., Zhao, J., Li, X., Xu, Q., Liu, Y., Khan, I. A., & Yang, S. (2016). New alkaloids from green vegetable Soybeans and their inhibitory activities on the proliferation of concanavalin A-activated lymphocytes. *Journal of Agricultural and Food Chemistry*, 64(8), 1649-1656. <https://doi.org/10.1021/acs.jafc.5b06107>

---

Zaheer, K., & Akhtar, M. H. (2017). An updated review of dietary isoflavones: nutrition, processing, bioavailability and impacts on human health. *Critical Reviews in Food Science and Nutrition*, 57(6), 1280-1293. <https://doi.org/10.1080/10408398.2014.989958>

### **Copyright Disclaimer**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).