Metabolomics-Based Discovery of Biomarkers with Cytotoxic Potential in Extracts of *Myracrodruon urundeuva*

**Caio B. Castro,**a **Licia R. Luz,**b,c **Jhonyson A. C. Guedes,***c **Diogo D. Porto,**d **Maria Francilene S. Silva,**a **Gisele S. Silva,**a **Paulo R. V. Ribeiro,**c **Kirley M. Canuto,**c **Edy S. Brito,**c **Dávila S. Zampieri,**f **Claúdia Ó. Pessoa**a and **Guilherme J. Zocolo**a,*c

a**Departamento de Química, Universidade Federal de São Carlos, 13565-905 São Carlos-SP, Brazil**

b**Departamento de Química Analítica e Físico-Química, Universidade Federal do Ceará,** Av. Humberto Monte, s/n, Pici, 60455-760 Fortaleza-CE, Brazil

c**Embrapa Agroindústria Tropical, Rua Dra. Sara Mesquita, 2270, Pici,** 60020-181 Fortaleza-CE, Brazil

d**Embrapa Semiárido, Rodovia BR-428, km 152, Zona Rural, 56302-970 Petrolina-PE, Brazil**

e**Núcleo de Pesquisa e Desenvolvimento de Medicamentos (NPDM), Universidade Federal do Ceará,** Rua Coronel Nunes de Mello, 1000, 60420-275 Fortaleza-CE, Brazil

f**Departamento de Química Orgânica e Inorgânica, Universidade Federal do Ceará,** Av. Humberto Monte, s/n, Pici, 60455-760 Fortaleza-CE, Brazil

*Myracrodruon urundeuva* ("aroeira-do-sertão") is a species threatened with extinction due to anthropogenic exploitation. Phytochemical analysis of bark, branch and leaf extracts revealed the presence of several compounds such as flavonoids, phenols, tannins, quercetin derivatives and anacardic acids. Dereplication methodology was performed to tentatively identify 50 compounds analyzed by ultra-performance liquid chromatography coupled with an electrospray ionization quadrupole time-of-flight mass spectrometry operating in MS² mode (UPLC-QTOF-MS²). The extracts exhibited anti-tumor effect in cancer cells HCT-116 (colorectal), SF-295 (glioblastoma), HL-60 (leukemia), and RAJI (leukemia). Also, these results correlate with the principal component analysis (PCA) data that identified three distinct groups indicating, efficiently, metabolic differences between organs of *M. urundeuva*. Through discriminatory analysis of the orthogonal partial least squares (OPLS-DA), the variable of importance in the projection (VIP) and S-Plot, we were able to determine 30 potential biomarkers. The fingerprint of hydroethanolic extracts was correlated with the cytotoxicity assay and demonstrated a significant difference in the composition of plant extract.

**Keywords:** *Myracrodruon urundeuva*, dereplication, UPLC-MS/MS, chemometrics, cytotoxic activity

**Introduction**

*Myracrodruon urundeuva* Fr.All. (Anacardiaceae family), popularly known as “aroeira-do-sertão”, is a medicinal tree found in several regions of Brazil, especially in the caatinga.¹ Currently, it is included in the official list of Brazilian flora species threatened with extinction² in the vulnerable category due to indiscriminate use of the species for several purposes in the wood and pharmaceutical area.³

The plant raises the researchers interest due to anti-inflammatory properties of its extracts, notably associated with the presence of bioactive phenolic compounds such as tannins, polyphenols, ellagitannins and, mainly, dimeric chalcones.⁴ Previous studies have shown that the chemical properties of these substances may be associated with antitumor activity in lung cells and leukemia.⁵ Pharmacological studies revealed a wide variety of pharmacological activities including cytotoxic,⁶,⁷ anti-inflammatory and analgesic.⁸ Besides, *M. urundeuva* may prevent cancer indirectly due to antioxidant and anti-inflammatory activity of its compounds.⁴,⁶-¹⁰

*e-mail: guilherme.zocolo@embrapa.br*
Previous studies have reported antitumor activity of ethanolic extracts from different sections of the plant. Extract dilutions yielded a half maximal inhibitory concentration (IC\textsubscript{50}) between 9.5-16.7 μg mL\textsuperscript{-1} against leukemic HL-60 and, among other types such as SF-295 glioblastoma of IC\textsubscript{50} 17.3-36.3 μg mL\textsuperscript{-1}. The activity of this extract was reported to occur via an apoptotic mechanism, which results in a reduction of cell numbers, cell volume, and viability in addition to internucleosomal DNA fragmentation.

Pessoa et al.\textsuperscript{5} reported the action of ethanolic extracts of the Myracrodruon urundeuva leaf against the HL-60 and SW-1573 lines with IC\textsubscript{50} of 7.4 and 8.5 μg mL\textsuperscript{-1}, respectively. Therefore, the literature shows several shreds of evidence of significant antitumor activity from M. urundeuva extracts, which motivated the study of its chemical composition. No previous research has been conducted with the identification of compounds associated with the biological activity of M. urundeuva extracts. The metabolomic study of plant samples is of great importance when one wants trying to associate certain bioactivity with the chemical composition of the extract. In this regard, the metabolomics focuses on the study of low molecular weight compounds that may be established as biomarkers\textsuperscript{11} by means of metabolic fingerprinting and profiling.\textsuperscript{12} This area of study covers a set of analyses from extraction methods to the statistical analysis of the data in order to identify molecules that can function as biomarkers or from genome alteration.\textsuperscript{13} This is because it can be used in different spheres, such as metabolic fingerprint, metabolic profile, and metabolomics.\textsuperscript{14} The ultra-performance liquid chromatography (UPLC) coupled to high-resolution mass spectrometry (HRMS) has the ability and sensitivity to provide a high-resolution mass spectrum for a complex matrix as a plant extract. Therefore, it is widely used in metabolomics studies involving identification of substances from high complexity extracts.\textsuperscript{14-16}

Ultra-performance liquid chromatography coupled with an electrospray ionization quadrupole time-of-flight mass spectrometry operating in MS\textsuperscript{E} mode (UPLC-QTOF-MS\textsuperscript{E}) allied to chemometric analysis is very useful for the identification of compounds by comparison of different matrices. Further analyses as principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA) identify groups that differ from each other, as well as presenting the responsible components that cause these differences, which are recognized as discriminant.\textsuperscript{17} Thus, these analyses help to provide information about compounds that could be used as diagnostic to each sample type, therefore potential biomarkers.

Present work aimed to explore differences in metabolic fingerprints of M. urundeuva leaves, branches, and bark ethanolic extracts by using UPLC-QTOF-MS\textsuperscript{E} and multivariate modeling (PCA and OPLS-DA) in order to search associations between chemical composition and cytotoxic effect.

**Experimental**

**Plant material**

Samples of leaf, bark and branch from M. urundeuva were collected from naturally occurring young plants from the Embrapa semi-arid experimental field, close to the border between the municipalities of Petrolina and Lagoa Grande (Pernambuco State, Brazil, 09°04’16.4”S, 40°19’5.37”W) on August 24, 2016, between 9 and 10 o’clock in the morning. The voucher specimens have been deposited in the Herbarium with number HTS4978. Samples of leaves, bark, and branches were collected in biological quintuplets (in five different trees) taking into account the four quadrants of the tree, in the north, south, east and west directions. The four quadrants were assembled in single samples for each section of the plant. At the time of collection, the liquid nitrogen cooling process was performed at –80 °C. After that, the material was dried in a forced circulation oven at 40 °C for 168 h (one week). Prior to extraction, the samples were ground in a knife mill and stored in a plastic bag at room temperature.

**Chemicals**

The solvents used were from LiChrosolv\textsuperscript{®} of the Sigma-Aldrich Chemical Company (St. Louis, MO, USA). In all methods, high purity Milli-Q water (Billerica, MA, USA) was used. The standards for chlorogenic acid and corilagin were obtained from Sigma-Aldrich Chemical Company (St. Louis, MO, USA) and urundeuvine A and B were previously isolated by our laboratory.

**Sample preparation**

The method used was adapted for the preparation of extracts by liquid-liquid partition.\textsuperscript{18,19} Leaves, branches, and bark (50 mg) were added in Falcon (15 mL) tube and extracted with 4 mL hexane, at room temperature, for 20 min in ultrasound bath. Afterward, 4 mL of EtOH:H\textsubscript{2}O (7:3) solution was added. The samples were extracted again with hexane, and the hydroethanolic portion was collected to yield the corresponding EtOH extract. Finally, a 1 mL aliquot of the lower (hydroethanolic) phase was filtered
(0.20 μm polytetrafluoroethylene (PTFE)), collected in flasks and stored at ~80 °C until further UPLC analysis.

Chromatographic conditions

The analysis was performed using an Acquity UPLC (Waters) system, coupled with a quadrupole/TOF (Waters) system. A Waters Acquity UPLC BEH column (150 × 2.1 mm, 1.7 μm) was used, with the column temperature set at 40 °C. The binary gradient elution system consisted of 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B). The UPLC elution conditions were optimized as follows: linear gradient from 2 to 95% B (0-15 min), 100% B (15-17 min), 2% B (17.01 min), 2% (17.02-19.01 min), a flow of 0.4 mL min⁻¹, and a sample injection volume of 5 μL.

Mass spectrometry conditions

The chemical profiling of *M. urundeuva* leaves, branches, and bark extracts was performed by coupling the Waters Acquity UPLC system to the QTOF mass spectrometer (Waters, Milford, MA, USA) with the electrospray ionization interface (ESI) in positive and negative ionization modes. The ESI⁺ and ESI⁻ data was acquired in the range of 110-1180 Da, with a fixed source temperature of 120 °C, and a desolvation temperature of 350 °C. A desolvation gas flow of 350 L h⁻¹ was used for the ESI⁺ mode and the 500 L h⁻¹ for the ESI⁻ mode. The capillary voltage was 3 kV. Leucine enkephalin was used as a lock mass. The MS model used was Xevo G2-XS QTOF. The spectrometer operated with MS² centroid programming using a tension ramp from 20 to 40 eV. The instrument was controlled by MassLynx 4.1 software (Waters Corporation).

Chemometric data analysis

The UPLC-MS data of all samples were analyzed using the MarkerLynx XS software to identify potential discriminatory chemical markers in different extracts. For data collection, the method parameters were set as retention time (t_R) range, 0.88-17.0 min, and mass range of 110-1180 Da. For data analysis, a list composed of the identities of the detected peaks was generated using retention time (t_R)-mass data (m/z) pairs as the identifier for each peak. An arbitrary ID was assigned to each of this t_R-m/z pairs based on their order of elution from the UPLC system. The ion intensity for each detected peak was normalized against the sum of the peak intensities within that sample. Ion identification was based on the t_R and m/z values. The resulting three-dimensional data comprising peak number (t_R-m/z pair), sample name, and ion intensity were analyzed by PCA and OPLS-DA using MarkerLynx.²⁰

Cytotoxicity of leaf, bark, and branch samples from *M. urundeuva*

Cell lines and cultures

Cytotoxicity tests were performed against HCT-116 and SW-620 (colorectal), SF-295 (glioblastoma), HL-60 and RAJI (leukemia), PC3 (prostate) and L929 (murine fibroblast) cell lines, which were obtained from the National Cancer Institute (Washington, DC, USA). All cells were cultured in Roswell Park Memorial Institute (RPMI) 1640, except for L929, which was cultivated in Dulbecco’s Modified Eagle Medium (DMEM). Both mediums were supplemented with 10% fetal bovine serum (FBS) and 1% antibiotics (100 U mL⁻¹ penicillin and 100 μg mL⁻¹ streptomycin) at 37 °C with 5% CO₂. The L929 cell line was used to evaluate the selectivity of the extracts and these assays, the anticancer drug doxorubicin was used as positive control.

Determination of cytotoxicity - MTT assay

The determination of cytotoxicity was performed by the MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazoliumbromide) colorimetric method.²¹ The samples were tested at 100 μg mL⁻¹ in six tumor cell lines for initial screening; the IC50 concentration was determined for those samples that showed positive results (growth inhibition > 75%) in at least three cell lines. The cells were plated in 96-well plates at the following concentrations: HCT-116 / SW-620: 0.7 × 10⁵ cells mL⁻¹; SF-295 / PC3 / L929: 0.1 × 10⁶ cells mL⁻¹; HL-60: 3 × 10⁵ cells mL⁻¹; RAJI: 4 × 10⁵ cells mL⁻¹. The cells were treated with the extracts for 72 h. At the end of the treatment, the plates were centrifuged, and the supernatant removed. Then, 150 μL of MTT solution (0.5 μg mL⁻¹) was added and incubated for 3 h. After incubation, the MTT solution was removed, and the precipitated formazan was dissolved with 150 μL of dimethyl sulfoxide (DMSO). The absorbances were read using a plate spectrophotometer (Multimode Detector, DTX 880, Beckman Coulter) at 595 nm.

Statistical analysis of data activity

All experiments were performed in duplicate and repeated three times. For all samples, the selectivity index (SI) was calculated. The calculation of this index corresponds to the division between the IC₅₀ value of each test compound in the L929 non-tumor cell line and the IC₅₀ value of each compound in the tumor cell line.
(SI = IC₅₀ L929 / IC₅₀ neoplastic cells). The experiments were analyzed according to the mean ± standard deviation (SD) of the percentage of cell growth inhibition using the GraphPad Prism software.

Results and Discussion

Chemical profile by UPLC-QTOF-MS

The ethanolic extracts of the three sections of M. urundeuva were obtained from the methodology described in “Sample preparation” sub-section. The extracts were analyzed by UPLC-QTOF-MS following the parameters described in “Chromatographic conditions” and “Mass spectrometry conditions” sub-sections only in the negative mode. In all, about 50 compounds were tentatively identified, covering the three sections of the species studied using MS and MS/MS from the chromatographic analysis (Figure 1). These results were compared to the data reported in the literature (chemotaxonomic) referring to the family (Anacardiaceae) and the genus (Myracrodruon) because there are few reports concerning the species. We used databases such as PubChem, ChemSpider, and Scifinder to support the results.

A wide range of phenolic compounds was identified, mainly derived from flavonoids and tannins. The predominant compounds in leaves were corilagin, firstly reported to the species, as well as geraniinic acid, and compounds well known in the literature as quercetin, gallic acids, and anacardic acids derivatives. The ethanolic extract of branches presented predominantly chlorogenic acid, quinic acid derivatives and the dimeric chalcones, urundeuvines A and B. The bark presented mostly catechin derivatives, in addition to the compounds contained in the branch.

A fragmentation study of the possible biomarkers tentatively identified was performed, presented below. The remaining substances have been tentatively identified and are presented in Table S1 and Figure S1 (Supplementary Information section).

Hydrolysable tannins

Peaks 13, 14 and 19 showed a precursor ion with [M – H]⁻ at m/z 951.0764, 633.0710 and 953.0887, respectively. Peak 13 showed fragments in 933.0729 [M – H – H₂O]⁻ and 300.9977 [M – H – 633 – H₂O]⁻ indicating losses of a water and corilagin unit with water. This compound was identified as geraniin from the literature data with a molecular mass of C₄₁ H₂₈ O₂₇.²⁴ Peak 14 presented MS/MS fragmentation pattern with loss of galloyl unit and water that showed fragment in 463.0564 [M − H − 152 − H₂O]⁻ and fragment in 300.9953 [M − H − 152 − 180]⁻ that corresponds to the loss of a galloyl unit with a hexose. This compound was identified as corilagin according to information previously reported in the literature,²⁴ with a molecular formula of C₂₇ H₂₂ O₁₈. Corilagin identification was confirmed by the comparison with molecular ion and fragmentation pattern of an analytical standard sample presented in Figure S2 (Supplementary Information section). Peak 19 showed similar fragments of the compound 13 at m/z 300.9953 [M − H – 3 × 152 − 180 − H₂O]⁻ and m/z 169.0121 as deprotonated gallic acid. This compound was identified as geraniinic acid with a molecular formula of C₄₁ H₃₀ O₂₇.

Figure 1. Typical base peak intensity (BPI) chromatograms showing the profiles of leaf, branch, and bark of M. urundeuva in the negative ionization mode.
The fragmentation pattern was very similar, and they are derived from the ellagitannin common in the genus *Phyllanthus*. These compounds are being reported for the first time to the Anacardiaceae family.

**Flavonols**

Peaks 24, 26, 27, 33 and 46 were identified as quercetin derivatives. Peak 24 presented precursor ion [M − H]⁻ at m/z 497.1338 (C₂₂H₂₆O₁₃). It is an unidentified compound but may be considered quercetin derivative because it presents a characteristic fragment ion at m/z 301. Peak 26 presented precursor ion [M − H]⁻ at m/z 463.0867 (C₂₁H₂₀O₁₂). This compound presented fragments in 301 [M − H − 162]⁻ indicating loss of hexoside unit. The identification of quercetin-3-O-galactoside was based on the work of Erşan et al. of *Pistacia vera* L., which belongs to Anacardiaceae family. Peak 27 was identified as quercetin-3-O-glucuronide similarly to the above compounds with precursor ion [M − H]⁻ at m/z 477.0648 (C₂₁H₁₈O₁₃). The main fragment at m/z 301 [M − H − 176]⁻ confirms the quercetin derivatives. Peak 33 presented precursor ion [M − H]⁻ at m/z 433.0777 (C₂₀H₁₈O₁₁) and was identified as quercetin-3-O-arabinopyranoside. This determination was performed by comparison to the work of Schieber et al. while was identified these compounds in the respective elution order. The main fragment at m/z 301 [M − H − 132]⁻ confirms the quercetin derivatives by the loss of pentoside unit. Peak 46 presented the deprotonated ion [M − H]⁻ at m/z 599.0995 (C₂₈H₂₄O₁₅) with fragment MS/MS at m/z 301 [M − H − 298]⁻ referring to the loss of one unit of raminoside-gallate, characteristic of quercitrin 2-O-galate, previously identified by Abu-Reidah et al.

Peaks 9 and 15 were identified as gallocatechin derivatives. All gallocatechin derivatives have a 125 Da fragment. This fragmentation is shown in Figure 2 by the formation of free phenol and the non-formation of a fragment of the gallic acid, indicating the gallocatechin and epigallocatechin compounds. This proposal was based on Miketova et al. Peak 9 presented precursor ion [M − H]⁻ at m/z 305.0653 (C₁₅H₁₀O₁) which, according to the previous reference, was identified as gallocatechin/epigallocatechin. Identification of the correct stereochemistry is not possible only by the fragmentation pattern. The comparison with an analytical standard as well as isolation of the compound and analysis by other techniques would be necessary. Peak 15 presented precursor ion [M − H]⁻ at m/z 457.0768. The fragmentation of peak 15 differs from 9 by the presence of the fragments at m/z 305.0664 [M − H − 152]⁻ and 169.0125 [M − H − 288]⁻ indicating the loss of a galloyl unit with the formation of the epigallocatechin unit and gallic acid deprotonated unit, respectively. The formation of this compound fragments is presented in Figure 2.

**Other compounds**

Peak 3 presented precursor ion [M − H]⁻ 133.0129 (C₇H₄O). In the MS/MS spectrum was at m/z 115.0042

![Figure 2. Proposed fragmentation of epigallocatechin 3-O-gallate, collision energy ramp 20–40 eV (adapted from reference 29).](image-url)
[M − H − H2O]. Based on Abu-Reidah et al. work, it was identified as malic acid.

Peak 5 demonstrated in its first-order spectrum the molecular ion m/z [M − H]− at 191.0192 (C6H8O7). This compound presented fragment in 111.0079 Da and it was possible to identify as citric acid as suggested by Lafontaine and co-workers.

Peaks 10, 20, 25, 30, 35, 36, 42, 45 and 50 presented precursor ion [M − H]− at m/z 483.0813, 785.0815, 497.1295, 939.1104, 1091.1266, 341.0647, 629.1295, 603.1172 and 447.0757, respectively. All compounds showed the presence of the same fragment in 169 Da corresponding to deprotonated gallic acid. Peaks 10, 30 and 35 presented with loss of 152 Da each that correspond to galloyl unit. The precursor ions of the peaks 30 and 35 were reported in the literature to the methanolic extracts leaf of M. urundeuva as a series of galloylglucose. Based on other reports, these compounds were identified as pentagalloyl hexoside and hexagalloyl hexoside, respectively.

Peak 12 showed a precursor ion [M − H]− at m/z 353.0861 (C16H18O9), that was identified as chlorogenic acid. To confirm the identification, the mass spectrum (MS/MS) of the analytical standard was compared with that of the extract, where similarity in retention time was observed, and the same fragment at m/z 191.0529 [M − H − 162]−.

Peak 16 exhibited deprotonated ion [M − H]− 337.0910 (C16H18O8). The MS/MS spectrum presented as the base peak the ion fragment at m/z 191.0523 corresponding to the deprotonated quinic acid. The characteristic fragment ion at m/z 163.0423 corresponds, undoubtedly, to coumaric acid. According to Plazonić et al., the compound was tentatively identified as 5-p-coumaroil quinic acid presented in Figure S3 (Supplementary Information section).

Peaks 29 and 51 showed precursor ions [M − H]− at m/z 467.0902 (C31H10O5) and 193.0869 (C11H14O3), respectively. Despite a broad review of the literature, these compounds remained unidentified.

Peaks 55 and 56 were identified as urundeuvine isomers from the comparison with analytical standards. Peaks 55 and 56 presented precursor ions [M − H]− at m/z 525.1268 and 525.1212 (C30H22O9). Both peaks have the same fragmentation pattern corresponding to urundeuvine A analytical standard as can be seen in the MS/MS spectrum (Figure 4).

In addition, the assignment was corroborated with a comparison between the retention times in the extract and in the analytical standard that was 6.24 and 6.63 min; and 6.24 and 6.69 min, respectively. Therefore, these compounds were identified as urundeuvines A isomers II and III. Similarly, MS/MS spectra and retention time based the identification of peak 57 on the comparison of extract with urundeuvine B analytical standard. This peak presented precursor ion at m/z 523.1036 [M − H]− and retention time of 6.78 min. Based on this information, the compounds were identified as urundeuvines B isomer II.

Peaks 60 and 62 were identified as anacardic acid (17:3) and (17:2), respectively presenting at m/z 369.2400 (C24H34O3) and 371.2587 (C24H36O3). All of them presented a loss of one molecule of CO2 equivalent to 44 Da in MS/MS, common in this type of molecule, giving the fragments at m/z 325.2519 and 327.2703.

Chemometric analysis

The main objective of using PCA analysis is to transform large amounts of complex analytical data into easily understood data. The analysis allowed to observe...
more clearly the metabolic differences between the different *M. urundeuva* samples: leaf, branch, and bark. The PCA plot (Figure 5) represented 83.34% of the total variance ($R^2_X [1] = 0.5959$ and $R^2_X [2] = 0.2375$), using the Pareto scale. The groups formed to indicate that the secondary metabolites of the leaf, branch, and bark differ significantly. The first major component (PC1) represents the largest variation in the dataset; the branch and bark samples are on the PC1 positive side, while the leaf samples are on the negative side. The second main component (PC2) corresponds to the maximum amount of variance not explained by PC1, in which case the branch is positive for PC2 and leaf, and bark is negative. Therefore, according to the PCA data, it is evident that the three parts of the plant differ from the respective chemical profiles.

After analysis of principal components for all samples, OPLS-DA was performed among the three groups (leaf-bark, leaf-branch, and bark-branch). It was possible to verify clearly the formation of distinct groups, also observed in the PCA, demonstrating the dissimilarity between leaves, branches, and barks. In addition, in the OPLS-DA, the intra-group variation can be observed, that is, how much the samples from the same tissues may differ from each other, and in this case, a greater homogeneity occurred in the leaf samples compared to bark and branch samples as shown in the OPLS-DA graphs (Figure 6). The good quality of the model is expressed in $R^2_Y$ (explained variance) and $Q^2$ (predicted variance), where the values must be above 0.5 and the closer to 1 the more reliable.34 For the analysis, $R^2_Y$ and $Q^2$ ranged from 0.98 to 0.99, indicating that the results are highly reliable. In order to identify the metabolites that have the greatest contribution to the distinction between the parts of the plant, other statistical tools derived from the OPLS-DA were used: VIP (variable of importance in projection) and S-Plot. By employing VIP, it is possible to predict which are the most significant variables for the selection of biomarkers, in general, VIP $> 1$ is considered statistically significant.35 In the present study, VIP $> 1$ and $p < 0.05$ were used. S-Plot highlights the discriminant variables, that is, those that move away from the common axis between the two groups compared. Figure 7 presents the VIP and S-Plot graphs for the leaf-bark group. The complete data containing all
Figure 6. OPLS-DA analysis of groups: (a) leaf-branch; (b) leaf-bark; (c) bark-branch of *M. urundeuva* (negative mode, *t*ₚ range: 0.88-17.0 min).

Figure 7. VIP of leaf samples (a), bark (b), S-Plot of leaf samples (c) and bark (d) and bar graph of leaf samples (e) and bark (f) of *M. urundeuva*. 
other comparisons made through the S-Plot and bar charts are presented in the Figures S4 and S5 (Supplementary Information section).

After the combination of OPLS-DA, VIP and S-Plot it was possible to tentatively identify the possible biomarkers (Table 1) that may be associated with the highest cytotoxic activity presented in all biomarkers of leaf, bark and branch extracts.

Table 1. Biomarkers present in leaf, branch and bark ethanolic extracts of *M. urundeuva* by UPLC-QTOF-MS

| Peak | tR / min | MSa [M – H]+ | MS/MS | Molecular formula | Tentative identification | M. urundeuva | VIPb | p-valuec |
|------|----------|-------------|-------|------------------|-------------------------|---------------|------|----------|
| 3    | 0.96     | 133.0129    | 115.0042 | C4H8Oa malic acid | +                       | 3.67          | 1.09 x 10^-4 |
| 5    | 0.98     | 191.0192    | 111.0079 | C5H6O3 citric acid | +                       | 5.00          | 2.55 x 10^-5 |
| 9    | 2.24     | 305.0653    | 125.0232 | C6H10O7 galloccatechin / epigallocatechin | +                      | 3.67          | 5.28 x 10^-10 |
| 10   | 2.77     | 483.0813    | 331.0816; 169.0153 | C14H8O7 malic acid + | 3.67          | 4.59          | 1.34 x 10^-9 |
| 12   | 3.00     | 353.0896    | 191.0493; 127.0390 | C6H8O7 citric acid + | 5.00          | 2.55 x 10^-5 |
| 13   | 3.26     | 951.0764    | 933.0729; 300.9977 | C15H14O7 gallocatechin / epigallocatechin | +                      | 3.67          | 5.28 x 10^-10 |
| 14   | 3.30     | 471.0744    | 463.0564; 300.9948 | C6H10O7 galloccatechin | +                       | 6.23          | 6.28 x 10^-6 |
| 15   | 3.48     | 457.0678    | 305.0664; 169.0125; 125.0248 | C7H12O7 gallocatechin-3-O-gallate | +                       | 3.77          | 1.66 x 10^-10 |
| 16   | 3.50     | 337.0910    | 191.0525; 163.0423; 93.0346 | C8H10O7 5-p-coumarol quinic acid | +                       | 5.34          | 6.95 x 10^-12 |
| 19   | 3.80     | 953.0887    | 300.9953; 169.0121 | C14H12O7 geraniinic acid | +                       | 5.96          | 3.56 x 10^-7 |
| 20   | 3.87     | 785.0815    | 300.9965; 169.0135 | C6H8O7 gallic acid derivative I | +                       | 6.34          | 1.69 x 10^-11 |
| 24   | 4.08     | 497.1338    | 301.0029 | C16H18O9 chlorogenic acid | +=                      | 3.18          | 9.02 x 10^-9 |
| 25   | 4.11     | 497.1295    | 313.0569; 217.0141; 169.0125 | C8H12O7 digalloyl hexoside | +                       | 4.81          | 4.15 x 10^-12 |
| 26   | 4.49     | 343.0777    | 301.0091; 300.0276 | C7H12O7 quercetin 3-O-glucoside | +                       | 3.77          | 2.46 x 10^-10 |
| 27   | 4.58     | 1091.1266   | 939.1158; 769.0972 | C41H28O27 geraniinic acid | +                       | 4.00          | 1.06 x 10^-9 |
| 29   | 4.74     | 341.0647    | 217.0136; 189.0191; 169.0123 | C8H12O7 gallic acid derivative | +                       | 7.70          | 4.34 x 10^-9 |
| 30   | 4.74     | 939.1104    | 769.0933; 617.0834; 169.0125 | C18H14O7 pentagalloyl hexoside | +                      | 3.19          | 7.69 x 10^-8 |
| 33   | 4.43     | 343.0777    | 301.0091; 300.0276 | C7H12O7 quercetin 3-O-arabinopyranoside | +                       | 3.32          | 1.28 x 10^-9 |
| 34   | 5.48     | 1091.1266   | 939.1158; 769.0972 | C41H28O27 quercetin 3-O-glucuronide | +                       | 4.11          | 4.1 x 10^-7 |
| 35   | 5.44     | 341.0647    | 217.0136; 189.0191; 169.0123 | C8H12O7 gallic acid derivative | +                       | 6.72          | 3.35 x 10^-12 |
| 42   | 5.18     | 629.1295    | 519.0919; 467.1021; 169.0126 | C10H14O7 gallic acid derivative | +                       | 4.22          | 5.50 x 10^-9 |
| 45   | 5.34     | 603.1172    | 341.0698; 323.0532; 169.0111 | C8H12O7 gallic acid derivative | +                       | 3.35          | 3.21 x 10^-8 |
| 46   | 5.44     | 599.0995    | 301.0322 | C9H12O7 quercetin O-galate isomer I | +                       | 4.79          | 5.02 x 10^-11 |
| 50   | 5.59     | 447.0757    | 295.0283; 169.0137 | C8H12O7 gallic acid derivative | +                       | 3.41          | 3.35 x 10^-12 |
| 51   | 5.79     | 193.0869    | 178.0597; 163.0395 | C8H12O7 gallic acid derivative | +                       | 3.12          | 9.04 x 10^-9 |
| 55   | 6.24     | 525.1195    | 389.1032; 371.0961; 135.0076 | C18H14O7 pentagalloyl hexoside | +                       | 4.74          | 5.59 x 10^-11 |
| 56   | 6.63     | 525.1180    | 389.1053; 371.0939; 135.0079 | C8H12O7 urundeuvine A | +                       | 4.79          | 1.18 x 10^-12 |
| 57   | 6.78     | 523.1036    | 521.0576; 387.0816; 371.0969; 135.0080 | C8H12O7 urundeuvine B | +                       | 3.89          | 1.3 x 10^-8 |
| 60   | 15.38    | 369.2400    | 325.2519; 255.2318; 183.0120; 133.0176 | C8H12O7 anacardic acid (17:3) | +                       | 1.89          | 0.0324 |
| 62   | 16.05    | 371.2587    | 327.2703 | C8H12O7 anacardic acid (17:2) | +                       | 4.10          | 1.94 x 10^-8 |

*aRetention time; bmass spectrometry; cvariable of importance in projection; dpprobability value; ecompounds that were compared with an analytical standard; fVIP and p-value values of the part of the plant that has the highest relative concentration of the compound tentatively identified; gcompounds reported for the first time, in Myracrodruon urundeuva; hnot identified. +: lower relative concentration; ++: highest relative concentration.*
inhibition above 70% in all cell lines exposed to ethanolic bark extracts. The leaf extracts were toxic only to leukemia (HL-60) cells. Table 2 shows the results of the cytotoxicity assays.

After initial screening, IC₅₀ tests were performed with the bark and leaf extracts. The extracts showed higher cytotoxic potential against the leukemic cell line with IC₅₀ ranging from 17.46 (bark) to 18.55 μg mL⁻¹ (leaf) (Table 3).

Other studies show in vitro cytotoxic effects of plant ethanolic extracts. IC₅₀ of 38.1 μg mL⁻¹ was found after treatment of leukemic cells with ethanolic extract of *M. urundeuva* seeds. The authors showed DNA fragmentation and mitochondrial depolarization caused by seed extracts.

Studies on the inhibition of tumor cells growth under the effect of the ethanolic extract of *M. urundeuva* were carried out. IC₅₀ values for this study were 7.4 μg mL⁻¹ against HL-60 and 8.5 μg mL⁻¹ against SW-1573. Bark ethanolic extract, for example, presented 95.6% growth inhibition for the breast, colon, and glioblastoma lines.

Viana et al. demonstrated that hydroethanolic extract of *M. urundeuva* bark exerts anti-inflammatory and analgesic effects related to chalcones. The extracts have antioxidant properties attributed to flavonoids. Souza et al. demonstrated the anti-inflammatory and protective effects against gastric ulcer in mice or rat after treatment with fraction rich in tannins extracted from the “aroeira” using ethyl acetate as the solvent.

The selectivity index (SI) of each sample was evaluated. The SI measures how much a compound is active against tumor cells without causing damage to non-tumor ones, and it is interesting when it presents values greater than 2.0. In the cell lines tested the leaf extract showed selectivity to the HL-60 line with an index higher than 2. This result can be correlated by the possible biomarkers, tentatively identified via chemometric analysis, present in the leaf, such as corilagin, geranin, geraniinic acid, quercetin derivatives, among others.

Corilagin, a compound well described in the literature, presents a variety of pharmacological effects, such as anti-tumor, anti-inflammatory, antioxidant, and hepatoprotective. Also, the literature reports good antitumor activity along with low toxicity to healthy cells and tissues, making corilagin a promising anticancer lead molecule. The geranin, another possible adjuvant compound, is known to exert antitumor, antibacterial, antioxidant, and antiviral activities. It has already been reported in the literature that quercetin and its derivatives are well known for their antioxidant, antihistaminic and anti-inflammatory properties. The quercetin is being considered a promising new chemotherapeutic agent, and several studies are underway to explore molecules derived from quercetin for cancer-directed chemotherapy. In addition, there may be other substances, as well as their synergistic compounds, that play significant roles in the reported bioactivity but were not identified in the present work due to limitations of the technique chosen.

The bark showed very promising results. The SI was

---

**Table 2.** Average cell growth inhibition (GI) of *M. urundeuva* extracts at 100 μg mL⁻¹

| Extract | HL-60 / % | SD / % | HCT-116 / % | SD / % | SF-295 / % | SD / % | RAJI / % | SD / % |
|---------|-----------|--------|-------------|-------|-----------|-------|----------|-------|
| Leaves  | 86.14     | 4.62   | 69.48       | 4.23  | 64.93     | 3.24  | 41.84    | 3.38  |
| Branch  | 41.98     | 4.94   | −           | −     | 10.52     | 5.47  | −        | −     |
| Bark    | 88.56     | 1.99   | 80.47       | 2.46  | 83.41     | 1.87  | 77.56    | 0.70  |

Leukemia; standard deviation; colorectal; glioblastoma. Results are expressed as mean percent cell growth inhibition (IC) and standard deviation for two independent experiments in triplicate.

---

**Table 3.** IC₅₀ values with a 95% confidence interval of *M. urundeuva* extracts in tumor and non-tumor cell lines

| Sample | SF-295 (glioblastoma) | PC3 (prostate) | HL-60 (leukemia) | RAJI (leukemia) | HCT-116 (colorectal) | SW-620 (colorectal) | L929 (murine fibroblast) |
|--------|-----------------------|----------------|------------------|-----------------|----------------------|--------------------|------------------------|
| Bark   | 38.86 (31.03-48.67)   | > 100          | 17.46 (16.29-19.70) | 65.73 (56.77-76.11) | 32.50 (29.44-35.87) | 40.68 (30.64-54.00) | > 100                  |
| Leaf   | 72.61 (55.80-94.49)   | > 100          | 18.55 (14.80-23.26) | 55.42 (44.49-69.04) | 55.60 (47.24-65.45) | > 100 (47.06-65.26) | 55.42                  |
| Doxorubicin | 0.25 (0.22-0.28) | 0.44 (0.34-0.54) | 0.01 (0.005-0.01) | 0.46 (0.45-0.47) | 0.11 (0.08-0.14) | 0.03 (0.02-0.05) | 0.99 (0.92-1.08) |

IC₅₀ (interval) / (μg mL⁻¹)
higher than 2 for all cell lines tested except for prostate (PC3) and leukemia (RAJI) (Table S2, Supplementary Information section). Eventually, in the extract, it is possible to highlight the discriminant biomarkers belonging to bark that may be related to the cytotoxic activity presented, such as urundeuvine A, galloy derivatives, catechins, and phenolic acids. Bandeira et al.52 had successfully isolated dimeric chalcones: urundeuvine A, B, C and matosin from the internal bark of M. urundeuva, and Souza et al.53 had demonstrated antimicrobial and anti-inflammatory activity for M. urundeuva. The literature also has evidence for the antitumor activity being promoted by specific classes of flavonoids such as chalcones, flavonones, and flavones. Many derivatives of these classes showed significant activities against some tumoral cell lines such as human colon, breast, and kidneys.53 Different polyphenols from “aroeira-vermelha” (Schinus terebinthifolius Raddi) induced cell death of human prostate carcinoma and were considered capable of modulating cell proliferation according to the test concentration. The use of catechins has shown inhibition of prostate and colon cancer.54,55 The combination of classical chemotherapy with nutrients and especially with polyphenols may decrease the pressure and the adverse effects of the antineoplastic drug.56 Therefore, three of the polyphenols present in the bark of the “aroeira” tree are promising compounds for isolation or synthesis into the development of phytopharmaceutical products from natural extracts. Chemical investigations of this extract can be a promising strategy for the discovery of phytotherapeutic agents. Also, the chemical profile comparison of bark and branch extracts revealed compounds that may be important for their biological activity.

Conclusions

From a simple and rapid extraction method, it was possible to trace the chemical profile of the three plant organs of M. urundeuva as leaf, branch and bark using the analytical technique UPLC-QTOF-MS®, which allowed the tentative identification of 50 compounds which covered several classes of compounds as flavonoids, flavanoids, hydrolysable tannins and anacardic acid. From the multivariate data analyses presented, it was possible to have information about the metabolic differences between the extracts compared. Such an association has been significant in the discussion of observed activities because the extracts obtained different responses against the tested lines.

The bark and leaf extract showed high toxicity and low IC50 values against the HL-60 (leukemia), HCT-116 (human colon) and RAJI (leukemia) cell lines compared to the branch. The higher relative concentration of compounds derived from quercetin, galloy derivatives, and phenolic acids present in these extracts may contribute to the understanding of the observed high cytotoxic activity. Some of the compounds identified, such as quercetin derivatives, corilagin, and chlorogenic acid, already has activities recognized as anti-tumor, antioxidants, and anti-inflammatory, among others, which may explain the promising activities observed here compared to literature.

Besides, through the statistical analysis, it was possible to observe the separation of the groups concerning each part of the plant and the identification of the 30 possible biomarkers. Therefore, this metabolic study notes the importance and value of the M. urundeuva plant as a possible source of secondary metabolites that are likely to act to inhibit certain types of cancer cells.

Supplementary Information

Table of tentative identification of the secondary metabolites present in the ethanolic extracts of leaf, bark, and branches of M. urundeuva; values of IC50 (non-tumor cells) / IC50 (tumor cell) selectivity index in tumor lines; spectrum in negative ion mode (ESI-MS/MS) of corilagin and 5-p-coumaroquinic acid; S-Plot of leaf samples, branch, bar graph of leaf samples and branch; S-Plot of bark samples, branch and bar graph of bark samples and branch of M. urundeuva are available free of charge at http://jbcs.sbq.org.br as PDF file.

Acknowledgments

The authors gratefully acknowledge financial support from the Embrapa (SEG 03.14.01.012.00.00), National Council for Scientific and Technological Development (CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), DCR (Programa de Desenvolvimento Científico e Tecnológico Regional), National Institute of Science and Technology (INCT BioNat, grant No. 465637/2014-0, Brazil). Authors thank the CNPq for the financial support and scholarships 303791/2016-0 and DCR-0024-01686.01.00/15.

References

1. Machado, A. C.; Souza, L. P.; Saldanha, L. L.; Pieroni, L. G.; Matos, A. A.; de Oliveira, F. A.; Vilegas, W.; Damante, C. A.; Dokkedal, A. L.; de Oliveira, R. C.; Pharm. Biol. 2016, 54, 2737.
2. Centro Nacional de Conservação da Flora (CNCFlora); Myracrodruon urundeuva in Lista Vermelha da Flora.
Metabolomics-Based Discovery of Biomarkers with Cytotoxic Potential in Extracts of *Myracrodruon urundeuva* J. Braz. Chem. Soc.

1. Viana, G. S. B.; Bandeira, M. A. M.; Matos, F. J. A.; Mahmoud, T. S.; Marques, M. R.; Pessoa, C. Ó.; Lotufo, L. V. C.; Souza, S. M.; Aquino, L. C.; Milach Jr., A. C.; Bandeira, M. A.; Calou, I.; Bandeira, M. A.; Aguiar-Galvão, W.; Cerqueira, G.; Takayama, T.; Mochizuki, T.; Todoroki, K.; Min, J. Z.; Mizuno, M.; Madsen, R.; Lundstedt, T.; Trygg, J.; Ernst, M.; Silva, D. B.; Silva, R. R.; Vêncio, R. Z. N.; Lopes, Wan, J. B.; Bai, X.; Cai, X. J.; Rao, Y.; Wang, Y.; Tugizimana, F.; Steenkamp, P. A.; Piater, L. A.; Dubery, I. A.; Citti, C.; Battisti, U. M.; Braghiroli, D.; Ciccarella, G.; Schmid, M.; Vandelli, M. A.; Cannazza, G.; Phytochem. Anal. 2013, 83, 1046.

2. Chagas-Paula, D. A.; Zhang, T.; da Costa, F. B.; Edrada-Ebel, R. A.; Metabolites 2015, 5, 404.

3. Luz, L. R.; Porto, D. D.; Castro, C. B.; Silva, M. F. S.; Alves Filho, E. G.; Canuto, K. M.; Brito, E. S.; Becker, H.; Pessoa, C. Ó.; Zoccolo, G. J.; J. Chromatogr. B 2018, 1099, 97.

4. MarkerLynx XS v4.1 SCN719; Waters Corporation, Milford, USA, 2009.
45. Londhe, J. S.; Devasagayam, T. P. A.; Foo, L. Y.; Ghaskadbi, S. S.; *Redox Rep.* **2008**, *13*, 199.
46. Li, J.; Wang, S.; Yin, J.; Pan, L.; *Can. J. Physiol. Pharmacol.* **2013**, *91*, 1016.
47. Li, J.; Huang, H.; Feng, M.; Zhou, W.; Shi, X.; Zhou, P.; *Antiviral Res.* **2008**, *79*, 114.
48. Lesjak, M.; Beara, I.; Simin, N.; Pintač, D.; Majkić, T.; Bekvalac, K.; Orčić, D.; Mimica-Dukić, N.; *J. Funct. Foods* **2018**, *40*, 68.
49. Mamani-Matsuda, M.; Kauss, T.; Al-Kharrat, A.; Rambert, J.; Fawaz, F.; Thiolat, D.; Moynet, D.; Coves, S.; Malvy, D.; Mossalayi, M. D.; *Biochem. Pharmacol.* **2006**, *72*, 1304.
50. Kahraman, A.; Erkasap, N.; Köken, T.; Serteser, M.; Aktepe, F.; Erkasap, S.; *Toxicology* **2003**, *183*, 133.
51. Hirpara, K. V.; Aggarwal, P.; Mukherjee, A. J.; Joshi, N.; Burman, A. C.; *Anti-Cancer Agents Med. Chem.* **2009**, *9*, 138.
52. Bandeira, M. A. M.; Matos, F. J. D. A.; Braz-Filho, R.; *Magn. Reson. Chem.* **2003**, *41*, 1009.
53. Queires, L. C. S.; Fauvel-Lafève, F.; Terry, S.; De La Taille, A.; Kouyoumdjian, J. C.; Chopin, D. K.; Vacherot, F.; Rodrigues, L. E. A.; Crépin, M.; *Anticancer Res.* **2006**, *26*, 379.
54. Brizuela, L.; Dayon, A.; Doumerc, N.; Ader, I.; Golzio, M.; Izard, J. C.; Hara, Y.; Malavaud, B.; Cuvillier, O.; *FASEB J.* **2010**, *24*, 3882.
55. Takeda, H.; Sonoshita, M.; Oshima, H.; Sugihara, K. I.; Chulada, P. C.; Langenbach, R.; Oshima, M.; Taketo, M. M.; *Cancer Res.* **2003**, *63*, 4872.
56. Dei Cas, M.; Ghidoni, R.; *Nutrients* **2018**, *10*, 940.

Submitted: July 8, 2019
Published online: October 10, 2019

This is an open-access article distributed under the terms of the Creative Commons Attribution License.