UPLC-QTOF-MSE-based metabolic profile to screening the chemodiversity of guaraná seeds through the untarget metabolimic approach

Tamyris de Aquino Gondim[1], Jhonyson Arruda Carvalho Guedes[2], Firmino José do Nascimento Filho[3], Gilvan Ferreira da Silva[3], André Luiz Atroch[3], Natasha Veruska dos Santos Nina[4], Rivelino Martins Cavalcante[1], Guilherme Julião Zocolo[5]

[1.] UFC , Universidade Federal de Ceará, Av. Abolição, 3207-Meireles, CEP: 60165-081, Fortaleza, CE, Brazil;
[2.] UFC, Universidade Federal de Ceará, Campus do Pici, Fortaleza - CE, CEP 60440-900, Brazil.;
[3.] Embrapa, Embrapa Amazônia Ocidental, Rodovia AM-010, Km 29, s/n - Zona Rural, CEP 69010-970, Manaus, AM, Brazil.;
[4.] UFAM, Universidade Federal do Amazonas, Av. General Rodrigo Octavio Jordão Ramos, 1200 - Coroado I, CEP 69067-005, Manaus, AM, Brazil;
[5.] Embrapa, Embrapa Soja, Londrina - PR, CEP 86001-970, Brazil;

The increasing focus on health and wellness has driven the development of functional foods, which offer benefits beyond basic nutrition. In this context, guaraná seeds (*Paullinia cupana*) have garnered attention for their high levels of caffeine, phenolic compounds, amino acids, and other nutrients. Therefore, guaraná is a prominent dietary supplement, with Brazil leading the world in its production. Thus, this study evaluated and determined the chemical profile of guaraná seeds from fifty-six different clones from the Active Germplasm Bank, aiming to provide different purposes in the food industry. The guaraná samples were extracted through a liquid-liquid microextraction. Initially, 50 mg of dry, ground, and homogenized plant material samples were weighed, 4 mL of hexane were added and vortexed for 1 min. Samples were placed in an ultrasonic bath for 20 min. Following this, 4 mL of a ethanol:water (7:3) solution was added, vortexed again for 1 min, and placed in the ultrasound bath for 20 min. To complete the separation of the hexane and hydroethanolic phases, the test tube containing the mixture was centrifuged for 10 min. Afterward, 2 mL aliquot was removed from the hydroethanolic phase and filtered before being added to vials. Seed samples of guaraná clones were extracted in biological triplicate, where each extract was analyzed only once by ultra-performance liquid chromatography coupled to high-resolution mass spectrometry (UPLC-QTOF-MS^E) with electrospray ionization (ESI^+ and ESI^-). The data obtained from the UPLC-QTOF-MS^E was processed using the software MS-DIAL to set up the parameters for untargeted metabolomics: deconvoluted spectra, peak alignment, and filtering. The annotation of metabolites was conceived through the joint analysis of MS and MS/MS mass spectra, aided by databases and spectral libraries available in the MS-DIAL and MS-FINDER software, in addition to the PubChem and SciFInder databases The multivariate analysis of guaraná seed samples was performed by MarkerLynx XS software. For data processing: mass range from 110 to 1500 Da; mass tolerance 0.02 Da; and data matrix scaled with the Pareto method. Subsequently, the principal components analysis (PCA) was conceived, being described through graphs of scores and loadings. The chemical profile of the samples was established through an untargeted metabolomics approach, aiming to perform an exploratory screening of the specialized metabolites present in the samples**.** This approach allowed the observation of nineteen specialized metabolites, mainly proanthocyanidins, and methylxanthines. Additionally, the use of unsupervised multivariate analysis tools such as PCA allowed the evaluation of similarities and differences between clones The PCA was redesigned and the scores plot showed a separation of clones. The two principal components present an accumulated explained variance of 31%. The analysis of loadings allowed identifying the main specialized metabolites responsible for the discrimination of the four groups. Therefore, it was possible to differentiate the clones based mainly on the differences in the levels of metabolite concentration. Thus, contributing to the selection of the best clones for different purposes, for example with higher or lower caffeine levels, enabling appropriate use in different types of industry, such as the food industry and pharmaceutical industry, in addition to the

possibility of being a potential source of bioactive compounds.

Agradecimentos: The authors gratefully acknowledge the financial support from the CNPq, National Council for Scientific and Technological Development (303791/2016-0), INCT BioNat, National Institute of Science and Technology (grant # 465637/2014-0). We would also like to thank Embrapa (SEG 03.14.01.012.00.00 and SEG 10.19.00.035.00.00). Part of the work was financed by FAPEAM – Fundação de Amparo à Pesquisa do Estado do Amazonas, through EDITAL Nº. 004/2018 - AMAZONAS STRATEGIC, Project title: Conservation and use of the collection of guarana tree genotypes in Amazonas. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (PROEX 23038.000509/2020-82).