Food Chemistry 240 (2018) 496–504

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Substantial equivalence analysis in fruits from three *Theobroma* species through chemical composition and protein profiling

Walter Pérez-Mora[a,*], Jesús V. Jorrin-Novo[b,*], Luz Marina Melgarejo[a,*]

[a] Laboratorio de Fisiología y Bioquímica Vegetal, Departamento de Biología, Universidad Nacional de Colombia, Bogotá, Colombia
[b] Agroforestry and Plant Biochemistry, Proteomics and Systems Biology Research Group, Dpt. of Biochemistry and Molecular University of Córdoba-CeiA3 Córdoba, Spain

ARTICLE INFO

Keywords:
Theobroma bicolor
Theobroma cacao
Theobroma grandiflorum
Protein profiling
Traceability

ABSTRACT

Substantial equivalence studies were performed in three *Theobroma* spp., *cacao*, *bicolor* and *grandiflorum* through chemical composition analysis and protein profiling of fruit (pulp juice and seeds). Principal component analysis of sugar, organic acid, and phenol content in pulp juice revealed equivalence among the three species, with differences in some of the compounds that may result in different organoleptic properties.

Proteins were extracted from seeds and pulp juice, resolved by two dimensional electrophoresis and major spots subjected to mass spectrometry analysis and identification. The protein profile, as revealed by principal component analysis, was variable among the three species in both seed and pulp, with qualitative and quantitative differences in some of protein species. The functional grouping of the identified proteins correlated with the biological role of each organ. Some of the identified proteins are of interest, being minimally discussed, including vicilin, a protease inhibitor, and a flavonol synthase/flavanone 3-hydroxylase.

Biological significance: Theobroma grandiflorum and *Theobroma bicolor* are endemic Amazonian plants that are poorly traded at the local level. As close relatives of *Theobroma cacao,* they may provide a good alternative for human consumption and industrial purposes. In this regard, we performed equivalence studies by conducting a comparative biochemical and proteomics analysis of the fruit, pulp juice and seeds of these three species. The results indicated equivalent chemical compositions and variable protein profiles with some differences in the content of the specific compounds or protein species that may result in variable organoleptic properties between the species and can be exploited for traceability purposes.

1. Introduction

The *Theobroma* genus is composed of 22 species, nine of which are native to the Amazon region. These species constitute a vital resource for the indigenous communities of Latin America and other tropical countries in Asia and Africa. They are part of the wild animal and human Amazonian diet and are of potential agroindustrial interest (Motamayor et al., 2013). *Theobroma* fruits have a pleasing smell and flavor and are used in making juices, gelatins, ice cream, liquor, marmalades and other products (Rogez et al., 2004; González, Moncada, Idarraga, Rosenberg, & Cardona, 2016). Although only cultivated at the regional level, some species such as *T. bicolor* (macambo or bacao) and *T. grandiflorum* (cupuaçu) are in the process of domestication and technification (Hernández & Hernández, 2012). They may constitute an alternative to *T. cacao*, the most cultivated and commercialized species of this genus. *T. cacao* as well as *T. bicolor* and *T. grandiflorum* seeds are used worldwide in the production of chocolate (toasted seeds) and locally in the production of candy and cosmetics (Motamayor et al., 2013).

In addition, as close relatives of cocoa, other *Theobroma* species may be used in breeding programs as a source of genes related to desired phenotypes, such as disease resistance, drought tolerance, high protein content, and other agronomic and economic traits. The recent sequencing of the *T. cacao* genome (Argout et al., 2011) will favor research in this direction. The efficient utilization of these species requires detailed knowledge of their biology, especially from a genetic and molecular point of view; however, to date, few studies have been published on this topic. A substantial equivalence analysis between commercial cocoa and its wild relatives is necessary as a preliminary step prior to the validation, production and commercialization of Amazonian *Theobromas* as a food alternative. This result will improve consumer confidence and acceptance, open new markets to those interested in exotic fruits, avoid illicit crop problems and increase the incomes of indigenous farmers (Motamayor et al., 2013; Rogez et al., 2004).

In the present study, partial chemical composition (sugars, organic acids, and total phenolics) as well as protein content and profiling

* Corresponding authors.
E-mail addresses: whperezm@unal.edu.co (W. Pérez-Mora), bf1jonoj@uco.es (J.V. Jorrin-Novo), lmmelgarejom@unal.edu.co (L.M. Melgarejo).

http://dx.doi.org/10.1016/j.foodchem.2017.07.128
Received 13 April 2017; Received in revised form 19 July 2017; Accepted 25 July 2017
Available online 29 July 2017
0308-8146/

analyses were performed on the pulp juice and seeds of mature *T. bicolor* and *T. grandiflorum* fruits, and the results were compared to those obtained from *T. cacao*. Proteomics is a useful approach for translational purposes and is used in food and plant material traceability (Jorrin Novo & Valledor, 2013). This proteomics study will provide a basis for future traceability and functional studies on the differences among provenances and the effects of environmental conditions and agronomic practices on the chemical composition of the fruit. To date, there are no published proteomics studies on wild *Theobroma* spp. or on the pulp and seeds of *T. cacao* fruits. A few proteomics studies on *T. cacao* have examined changes that occur in the seeds during chocolate production (Lerceteau, Rogers, Pétiard, & Crouzillat, 1999), changes in the fruit pericarp (Awang, Karim, & Mitsui, 2010), differences between zygotic and somatic embryos (Noah et al., 2013), characteristics of the leaves and roots of soil-flooded plants (Bertolde, Almeida, & Pirovani, 2014) and responses to *Moniliophthora perniciosa* (Pirovani et al., 2008).

2. Materials and methods

2.1. Plant material

Fruits were obtained from the municipalities of Albania, Caquetá – Colombia (*T. grandiflorum*); Belén de los Andaquíes, Caquetá (*T. bicolor*); and Florencia, Caquetá (*T. cacao*). The fruits were mature (approximately 140 days after fruit set), and their condition was optimal for consumption. The fruits were sampled using a completely randomized design. Each fruit represented a single sample unit or biological replicate. Four biological replicates were obtained for each species, and 3 analytical replicates were conducted, yielding a total of 12 replicates per specie.

The pulp was manually separated from the seeds and squeezed using a commercial blender to obtain the juice. The seeds were thoroughly washed with tap water and dried with filter paper.

Both the seeds and the juice were lyophilized using a Labconco Freezone 4.5 device (Labconco Corporation; Kansas City, USA) ground into small particulates and, stored at −80 °C until analysis. The pulp fresh weight/volume of juice/lyophilized powder ratio was 1 g/0.87 mL/165 mg on average.

2.2. Chemical analysis of pulp juice

The chemical composition of the pulp juice was determined as these parameters are related to the quality of the fruit. The pH, total titratable acidity and total soluble solids of the fresh juice were determined directly before lyophilizing. Sugars, organic acids and phenolics were analyzed in the lyophilized juice powder.

The pH was measured using the potentiometric method in accordance with AOAC 981.12 (AOAC, 2005). The total titratable acidity (TTA) was determined according to AOAC 942.15 (AOAC, 2005). The total soluble solids (TSS) expressed in units of °Brix units was determined using a Hanna Instruments HI96802 refractometer (Hanna instruments, Woonsocket, USA.) according to AOAC 932.12 (AOAC, 2005).

The following biomolecules were analyzed in the lyophilized juice powder from the three *Theobroma* spp.: total phenols, reducing sugars, sucrose, glucose, fructose, oxalic acid, malic acid and citric acid. The data are presented as relative to dry weight DW.

The extraction and quantitative analysis of the phenolics and reducing sugars were performed as reported by Zeng et al. (2017). The total phenols were quantified by the Folin-Ciocalteu method using tannic acid as the standard. The reducing sugar content was determined by the 3,5-dinitrosalicilic method using glucose as the standard. The sucrose, glucose and fructose content was measured by HPLC chromatography (column Phenomenex Ca^{++} monosaccharide; mobile phase: water; and refractive index detector) as previously reported (Solarte, Melgarejo, Martínez, Hernández, & Fernández-Trujillo, 2014). The organic acid contents (oxalic, malic and citric acids) were measured by HPLC chromatography (column ROA acid organic H^+; mobile phase: 5 mM H_2SO_4; and a photodiode array detector at 207 nm) (Solarte et al., 2014). The HPLC analysis was performed using a Waters instrument (Waters, Milford, Massachusetts, USA).

2.3. Protein extraction and quantification in seeds and pulp juice

Proteins were extracted from 0.2 and 0.1 g of lyophilized powder of the seeds and the pulp juice, respectively. The preliminary experiments did not yield 2-DE patterns from seeds of sufficient quality or resolution. For this reason, prior to protein extraction, the lyophilized powder was delipidated in *n*-hexane (1:5, p/v with constant shaking) for 5 min. The TCA-acetone protocol reported by Maldonado, Echevarría-Zomeño, Jean-Baptiste, Hernández, and Jorrín-Novo (2008) was employed to extract proteins. The final pellet was solubilized in 2 mL of rehydration solution (7 M urea, 1.2% (p/v) CHAPS, 43 mM DTT, 30 mM Tris HCl pH 8.5). Insoluble material was removed by centrifugation (15500*g*, 10 min, 4 °C), and the protein content was quantified in the supernatant according to the Bradford method (Valero, Valledor, Navarro, Gil Pelegrín, & Jorrín-Novo, 2011) using a Bio-Rad Protein Assay kit and bovine serum albumin pattern (BSA) as the standard.

2.4. Two-dimensional electrophoresis

Protein extracts were subjected to 2D electrophoresis as reported by Maldonado et al. (2008) and Valero et al. (2011). In seeds, 4–7 pH IPG strips 18 cm in length were used, and 3–10 pH strips were used for the pulp juice (Bio-Rad, Hercules, USA). The pH gradient was chosen on the bases of preliminary experiments and after visualizing the pH range distribution of the spots. Strips were actively rehydrated at 50 V for 20 h with 500 μg of extracted protein in 400 μL of IEF buffer (7 M urea, 2% p/v CHAPS, 0.2% (v/v) ampholytes (Bio-Lyte ampholytes 3–10, BioRad, Hercules, USA), 20 mM DTT and 0.01% (p/v) bromophenol blue). IEF was carried out in a Protean IEF Cell system (BioRad, Hercules, USA) at 20 °C using a gradual voltage increase between 250 V–10,000 V until reaching 55,000 V/h. The strips were immediately reduced and alkylated after IEF finished. Second dimension SDS-PAGE was carried out in a Protein Dodeca Cell (BioRad, Hercules, USA) using a 12% gel at 40 V for 3 h and 80 V until the dye reached the bottom of the gel. Seed gels were stained with G-250 colloidal Coomassie blue (Noah et al., 2013), while the pulp gels were silver stained (Blum, Beier, & Gross, 1987) as a more sensitive staining protocol was required to visualize most of the spots. The gels were digitized (UMAX PowerLook 2100XL-USB scanner), and images were analyzed with PDQuest 8.0.1 (BioRad, Hercules, USA) software using tenfold over the background as a minimum criterion for the presence or absence. The analysis was re-evaluated by visual inspection. For each spot, normalized volumes (individual spot intensity/normalization factor calculated for each gel based on the total quantity in valid spots) were used for statistical analysis. The spots were manually cut with an EXQuest spot-cutter system and subjected to MS analysis.

2.5. Mass spectrometry analysis and identification of the proteins

The spots were digested with trypsin, and the resulting peptides were subjected to MS analysis (4700 Proteomics Analyzer MALDI–TOF/TOF Mass Spectrometer, Applied Biosystems) (Maldonado et al., 2008). A range of 800–4000 m/z with an acceleration voltage of 20 kV in the reflectron mode was employed. The spectrometer was calibrated using the trypsin autolysis peaks of m/z = 842.51 and m/z = 2211.10 as the internal standard. The 3 most abundant peptide ions were then subjected to MS/MS analysis, providing information that can be used to determine the peptide sequence. A peptide mass fingerprinting search and a combined search (+ MS/MS) were conducted. The peptide search was carried out by GPS Explorer TM v 3.5 (Applied Biosystems) from