Food Research International 121 (2019) 12–19

Contents lists available at ScienceDirect

Food Research International

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

^{1}H NMR-based fingerprinting of eleven Mexican *Capsicum annuum* cultivars

Elideth Florentino-Ramos[a], Nemesio Villa-Ruano[b], Diego Hidalgo-Martínez[c], Moisés Ramírez-Meraz[d], Reinaldo Méndez-Aguilar[d], Rodolfo Velásquez-Valle[e], L. Gerardo Zepeda-Vallejo[f], Nury Pérez-Hernández[g], Elvia Becerra-Martínez[a,*]

[a] *Centro de Nanociencias y Micro y Nanotecnologías, Instituto Politécnico Nacional, Av. Luis Enrique Erro S/N, Unidad Profesional Adolfo López Mateos, Zacateco, Delegación Gustavo A. Madero, Ciudad de México 07738, México*

[b] *CONACyT-Centro Universitario de Vinculación y Transferencia de Tecnología, Benemérita Universidad Autónoma de Puebla, CP 72570 Puebla, México*

[c] *Department of Plant and Microbial Biology, University of California, 111 Koshland Hall, MC-3102, Berkeley, CA 94720-3102, USA*

[d] *INIFAP-Campo Experimental Las Huastecas, km 55 Carretera Tampico-Mante, Cuauhtémoc, Tamaulipas, México, CP 89610, México*

[e] *INIFAP-Campo Experimental Zacatecas, Km. 24.5 Carretera Zacatecas-Fresnillo, Apdo. Postal # 18, Calera de V. R., Zacatecas, México, CP 98500, México*

[f] *Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Prol. de Carpio y Plan de Ayala S/N, Col. Santo Tomás, Delegación Miguel Hidalgo, Ciudad de México 11340, México*

[g] *Escuela Nacional de Medicina y Homeopatía, Instituto Politécnico Nacional, Guillermo Massieu Helguera, No. 239, Fracc, "La Escalera", Ticomàn, Ciudad de México 07320, México*

ARTICLE INFO

Keywords:
Capsicum spp.
Cultivars
Metabolomic fingerprinting
Nuclear magnetic resonance spectroscopy
Principal component analysis (PCA)
Linear discriminant analysis (LDA)
Biomarkers

ABSTRACT

Approximately 90% of the chili peppers consumed in the world are harvested in Mexico. The present article describes the untargeted ^{1}H NMR-based metabolomic profiling of 11 cultivars of *Capsicum annuum* species which are routinely consumed worldwide. The metabolomic fingerprinting detected via ^{1}H NMR contained 44 metabolites including sugars, amino acids, organic acids, polyphenolic acids and alcohols which were identified by comparison with the literature data, with Chenomx database and by 2D NMR. Statistical approaches based on principal component analysis (PCA) and linear discriminant analysis (LDA) were used to classify the *Capsicum annuum* cultivars according to their metabolite profile. LDA revealed metabolomic differences and similarities among *Capsicum annuum* cultivars, whereas hierarchical cluster analysis (HCA) significantly separated the cultivars according to the phylogenetic trees obtained. Substantial endogenous levels of free amino acids and carbohydrates were detected in all the studied cultivars but interestingly, *Capsicum annuum* cv. mirasol and *C. annuum* cv. chilaca contained almost three-fold more endogenous levels of vitamin C than the other cultivars. Considering that this antioxidant was found in crude aqueous extracts, its abundance could be directly proportional to its bioavailability for human nutrition. The results suggest that ^{1}H NMR is an effective method to determine differences among cultivars of the *Capsicum annuum* species.

1. Introduction

The genus *Capsicum* is constituted by 25 wild species and 5 domesticated species including *C. annuum*, *C. frutescens*, *C. chinense*, *C. baccatum* and *C. pubescens*. These species contain > 200 cultivars (Troconis-Torres et al., 2012). Three different species of *Capsicum* are extensively cultivated in Mexico (143, 975 ha; 16.22 tons per hectare per year) (Montes-Hernández, 2010). The most common cultivars of chili peppers in Mexico are named agua, anaheim, árbol, caribe, chilaca, chorro, cuaresmeño, húngaro, mirasol, poblano or ancho and serrano. These cultivars belong to the *C. annuum* species, while habanero and manzano cultivars belong to the *C. chinense* and *C. pubescens* species, respectively (González-Zamora et al., 2013). Chili pepper is the most popular and emblematic Mexican condiment which represents an alternative source of vitamins and nutraceuticals such as ascorbate, carotenoids, tocopherols, flavonoids, and capsaicinoids (Hervert-Hernandez, Sayago-Ayerdi, & Goñi, 2010; Maji & Banerji, 2016; Pino et al., 2007). There are several studies regarding the antioxidant activity (Tan, Ali, & Zainal, 2012; Vega-Gálvez et al., 2009) and nutritional content at different stages of chili peppers development (Serrano et al., 2010). Metabolomic studies of chili peppers have mainly focused in the identification and quantification of carotenoids, capsaicinoids, ascorbic acid, and flavonoids (Lee, Howard, & Villalon, 1995). To our knowledge, there is limited information on non-targeted ^{1}H NMR-based metabolomics approaches in chili peppers (Becerra-Martínez et al., 2017; Jang, Jung, Lee, Choi, & Lee, 2015; Villa-Ruano et al., 2018). This

* Corresponding author.
E-mail address: elmartinezb@ipn.mx (E. Becerra-Martínez).

https://doi.org/10.1016/j.foodres.2019.03.025
Received 22 December 2018; Received in revised form 8 March 2019; Accepted 10 March 2019
Available online 12 March 2019

analytical technique is often used for the simultaneous detection of primary and secondary metabolites in biological samples (fluids, tissues, cell cultures and foods) (Picone et al., 2013; Picone et al., 2016; Trimigno et al., 2018). The ^{1}H NMR-based metabolomics profiling and multivariate analyses represent a powerful tool for establishing biochemical associations among metabolites (Wahyuni et al., 2014). It is also a recurrent strategy to identify significant differences based on geographical origin (Jang-Eun et al., 2010; Verpoorte, Choi, Mustafa, & Kim, 2008), taxonomical markers, and physiological conditions (Georgiev, Ali, Alipieva, Verpoorte, & Choi, 2011). This information should be very valuable to determine or even to classify foods produced under different conditions. ^{1}H NMR spectroscopy combined with principal component analysis (PCA) have been extensively used to obtain the metabolomic profiling of honey (Consonni, Cagliani, & Cogliani, 2012), meat (Jung et al., 2010), mango juice (Koda, Furihata, Wei, Miyakawa, & Tanokura, 2012), tea (Van Dorsten, Daykin, Mulder, & Van Dyunhoven, 2006), wine (Anastasiadi et al., 2009) and cheese (Rodrigues et al., 2011). Unlike other analytical techniques, sample preparation for ^{1}H NMR is relatively easy and various organic chemical species (sugars, lipids, amino acids, and organic acids) are accurately and simultaneously detected (Savorani, Rasmussen, Mikkelsen, & Engelsen, 2013). Thus, the aim of this study was to determine the metabolomic fingerprinting of eleven chili pepper cultivars belonging to the *Capsicum annuum* species. We emphasize specific differences in the ^{1}H NMR fingerprinting of these cultivars based on the principal component analysis (PCA).

2. Materials and methods

2.1. Plant source

The chili pepper fruits samples were harvested from an experimental field of INIFAP-Zacatecas (Longitude: 102° 39′ 34.0“, Latitude: 22° 54’ 31.3”, Altitude: 2197 masl) from July to August in 2016. For this study, the following cultivars were considered. (a) *C. annuum* cv. agua, (b) *C. annuum* cv. anaheim, (c) *C. annuum* cv. árbol, (d) *C. annuum* cv. caribe, (e) *C. annuum* cv. chilaca, (f) *C. annuum* cv. chorro, (g) *C. annuum* cv. cuaresmeño, (h) *C. annuum* cv. húngaro, (i) *C. annuum* cv. mirasol, (j) *C. annuum* cv. poblano, (k) *C. annuum* cv. serrano. Ten pepper fruits from each cultivar (110 samples) were processed and analyzed (Table S1 and Fig. S1). All samples were properly washed with deionized water to remove soil particles. The samples were stored at −20 °C prior to sample processing. The fruits were selected based on a similar color, size, texture, and weight. Similarities in color and texture among the samples were estimated by the perception of the harvester. Length, width and weight of the peppers were determined in accordance with the Mexican Official Norms NMX-FF-025-SCFI-2014 (Normas Oficiales Mexicanas, 2015). The length was measured from the base of the fruit (excluding the peduncle) to the apex and the width was measured from the widest point. The weight was estimated using an analytical balance (Table S1).

2.2. Chemicals

Deuterium oxide (D_2O, D 99.9 atom %) was purchased (Cambridge Isotope Laboratories, Inc.) and used as solvent. For ^{1}H NMR analysis, 3-(trimethylsilyl)-1-propanesulfonic acid sodium salt (Sigma-Aldrich Co.; TSP, 97%) was the internal standard; EDTA (ethylenediaminetetraacetic acid) and sodium azide (NaN_3) (Merck®) were also added to prevent bacterial growth. NaOH and HCl (Sigma-Aldrich Co.) were used to adjust the pH.

2.3. Sample preparation

The preparation of the sample was performed in accordance with protocols previously reported and validated for *Capsicum* sp. (Becerra-Martínez et al., 2017; Hohmann, Christoph, Wachter, & Holzgrabe, 2014; Villa-Ruano et al., 2018). Each pepper was squeezed into a mortar, and the juice was centrifuged for 20 min at 15900*g*. For subsequent ^{1}H NMR analysis, 900 µL of the aqueous upper phase was combined with 100 µL of a solution containing 7 mM TSP, 10 mM EDTA and 2 mM NaN_3 in D_2O at pH 5.42 ± 0.05. Finally, 600 µL of this solution was placed into a 5 mm NMR tube.

2.4. Nuclear magnetic resonance (NMR) experiments

^{1}H NMR experiments were performed in a Bruker 750 MHz spectrometer (Bruker Biospin, Rheinstetten, Germany) equipped with a 5 mm TXI cryoprobe. Aqueous extracts of peppers were measured at 298.1 ± 0.1 K without rotation and with 4 dummy scans prior to 64 scans. Acquisition parameters were set as follows: FID size = 64 K, spectral width = 10.00 ppm, receiver gain = 1, acquisition time = 2.18 s, relaxation delay = 10 s, and line width resolution = 0.45 Hz. Data acquisition was achieved with a NOESY presaturation pulse sequence (Bruker 1D noesypr1d) with water suppression via selective irradiation of the water frequency during recycling and mixing time delays (Becerra-Martínez et al., 2017; Villa-Ruano et al., 2018).

Four NMR experiments were performed to corroborate signal assignments: ^{13}C NMR spectroscopy, homonuclear correlation spectroscopy (2D ^{1}H−^{1}H COSY), heteronuclear single quantum correlations (2D ^{1}H−^{13}C HSQC) and heteronuclear multiple bond correlation (2D ^{1}H−^{13}C HMBC). The parameters for the ^{13}C NMR spectrum (188.6 MHz) the following acquisition parameters were used: 2048 scans, acquisition time = 0.72 s, relaxation delay = 2 s, spectral width = 45,459.21 Hz, and FID size = 64 k data (Fig. S3). COSY measurements were made with a spectral width of 7500.00 Hz in either dimension; 2 k × 128 were acquired with 256 scans per increment and a 2 s relaxation delay (Fig. S4). For gHSQC, there were 1024 scans and 128 increments with an acquisition time of 0.065 s and a relaxation delay of 1.5 s. The spectral width was 7500.00 and 45,459.21 Hz for the ^{1}H and ^{13}C dimensions, respectively, with $^{1}J_{CH}$ = 145 Hz (Fig. S5). For gHMBC, there were 128 scans and 128 increments with an acquisition time of 0.26 s and a relaxation delay of 2 s. The spectral widths were the same as for gHSQC (Fig. S6).

2.5. Metabolite profiling

Metabolite identity was assigned in accordance with the literature and it was additionally corroborated by 1D and 2D NMR experiments (Table S3 and Figs. S2–S6). The metabolites identity was performed as previously reported (Becerra-Martínez et al., 2017; Ritota, Marini, Sequi, & Valentini, 2010). The relative abundance of metabolites was determined using the Chenomx NMR suite. This software compares the integral of a known reference signal (TSP) with that of the signals from a library containing known chemical shifts and peak multiplicities for all the studied metabolites (Beltran et al., 2012). The data sets were converted into Microsoft Excel format prior to the chemometric analysis.

2.6. Post-processing of NMR data

Firstly, the ^{1}H NMR spectra were automatically phased. Posteriorly, the baseline was corrected and each spectrum calibrated to the TSP signal at 0.0 ppm with the MestReNova program (version 6.0.2; MestReC, Santiago de Compostela, Spain). The resulting ^{1}H NMR spectra were imported into the processor module of Chenomx NMR Suite version 8.2 (*Chenomx*, Edmonton, *Canada*). In the processor module, the ^{1}H NMR spectra were subjected to baseline correction, line broadening, phase correction and shim. In this module, the spectra were calibrated to the signal of the internal standard (TSP) and the pH was set within a range of 4–9. The relative metabolite abundance was

determined in the Profiler module. From each spectrum, a list of compounds and their relative abundance was produced and then, it was subjected to statistical analysis.

2.7. Multivariate statistical analysis

Chemometric analysis was done using SIMCA version 13.0.3 (Umetrics, Kinnelon, NJ, USA) and Past3.20 software (Øyvind Hammer, April 2018). The concentration data were imported into SIMCA and principal component analysis (PCA), an unsupervised pattern recognition method, was initially performed to examine intrinsic variation in the data set and to obtain an overview of the variation among groups. All variables were UV for multivariate analysis (Eriksson, Johansson, Trygg, & Vikström, 2013). Linear discriminant analysis (LDA) was performed in Past3.20 software. This analysis is commonly used for studying the association between a set of predictors versus a categorical response. Thus, this strategy was applied in order to find the best linear separation between groups by determining the minimal dimensions at which groups can be separated (Gromski et al., 2014). A one-way ANOVA was performed using the GraphPad Prism 7.0 software to determine significant differences in metabolite levels. Tukey-Kramer multiple-comparison tests were performed to reveal pair-wise differences between means ($p < 0.05$).

3. Results and discussion

3.1. Identification of metabolites in Capsicum aqueous extracts by 1D and 2D NMR

The ^{1}H NMR spectrum obtained at 750 MHz revealed a complex pattern of signals in the metabolic fingerprint from *Capsicum annuum* cv. chilaca (Fig. 1). After water suppression, the spectrum revealed the presence of sugars (fructose, glucose, and sucrose) and organic acids (citric acid). The 750 MHz ^{1}H NMR spectrum was divided into three main spectral regions. The first one (0.5–3.0 ppm) contained signals of aliphatic amino acids and organic acids. Signals from 0.90 to 1.10 ppm were related to the presence of valine, leucine, and isoleucine. Signals for alanine and threonine were observed as a doublet at 1.47 ppm and 1.32 ppm, respectively. Asparagine, aspartic acid, γ-aminobutyrate, methionine, and proline were also detected in this region. Proline was the major aliphatic amino acid.

Organic acids such as acetic, citric, lactic, pyruvic and succinic acids were also observed in this region. Citric acid appeared as an AB system at 2.60 and 2.72 ppm and was the most abundant organic acid. It is well known that citric acid has the role of calcium, magnesium and sodium chelator, thus enhancing line widths in a NMR spectrum. To achieve high resolution and optimize line widths, a small amount of EDTA was added to each sample (Dona et al., 2016). Ethanol was observed as a weak triplet signal at 1.17 ppm in the high field. This type of alcohol can be associated with microbial activity in the bacterial symbionts. The second region (3.0–5.5 ppm) exhibited intense and overlapping signals corresponding to the most abundant sugars. These compounds were easily recognizable in the anomeric region at 4.01, 4.63, 5.22, and 5.40 ppm for fructose, β-glucose, α-glucose, and sucrose, respectively. Several weak signals were detected at 4.5–5.5 ppm indicating the presence of galactose and maltose. *Capsicum* cultivars are a good source of ascorbic acid, which was observed as a doublet at 4.57 ppm. The third region was located at 6.0 and 9.0 ppm, these signals were relatively weak and represent aromatic groups from amino acids and phenolic compounds. The most abundant signal in this region was for phenylalanine; however, histidine, tyrosine, tryptophan, adenosine, cytidine, guanosine and uridine were also identified. The characteristic singlets of formic and fumaric acids were observed at 8.44 and 6.52 ppm, respectively, and were additionally detected in the same region.

A similar metabolite profiling was verified in the 1D NMR spectra from the different cultivars of *Capsicum annuum* (Fig. 2). After careful inspection of the ^{1}H NMR spectra of eleven cultivar of *Capsicum annuum*, 44 metabolites were accurately identified. Table S2 shows the differential metabolites stablished in the studied cultivars. The results show that the sugar concentration was very different across the samples (region 3.0–5.5 ppm). Six cultivars, *C. annuum* cv. árbol, *C. annuum* cv. caribe, *C. annuum* cv. chilaca, *C. annuum* cv. cuaresmeño, *C. annuum* cv. mirasol, and *C. annuum* cv. serrano contained the highest concentration of α-glucose and β-glucose. The metabolite profiling of *C. annuum* cv. mirasol had a different content and abundance of aliphatic and aromatic compounds versus the other studied samples; the amino acid content and its abundance was higher than in the other cultivars. The signal at 2.60 and 2.72 ppm assigned to citric acid was remarkably high in *C. annuum* cv. mirasol. Thus, citric acid could be a possible biomarker of this cultivar (Fig. 2 and Table S2).

3.2. Multivariate data analyses and metabolic classification of Capsicum samples

For each cultivar of *C. annuum*, 10 replicates were used and a total of 110 samples were processed in the multivariate analysis. In order to group these data, a principal component analysis was used. The respective clustering of PCA is shown in Fig. 3A and some outliers corresponding to *C. annuum* cv. agua and *C. annuum* cv. mirasol was observed. These data didn't fit well in 95% confident ellipse in which 65.9% of variance is described by this model in accordance with PC1 = 54.9% and PC2 = 11.0%. Further PCA was performed by removing the outliers detected. As expected, a better pattern of clustering was obtained and the main differences according to PC1 = 32.5% were among the cultivars *C. annuum* cv. arbol, *C. annuum* cv. serrano, *C. annuum* cv. chilaca, *C. annuum* cv. caribe (located in the right side of PC1) and the cultivars *C. annuum* cv. anaheim, *C. annuum* cv. chorro, *C. annuum* cv. hungaro and *C. annuum* cv. poblano (located at the left side). For the particular case of the cultivar cuaresmeño, it was difficult to determine a clear separation with this score plot, which had a total of 49.1% of total variance according to PC1 and PC2 (Fig. 3B).

Fig. 3C shows a plot describing which variables were associated to the similarity/dissimilarity between the observations of this study. According to these approaches, the cultivars *C. annuum* cv. arbol, *C. annuum* cv. serrano, *C. annuum* cv. cuaresmeño and *C. annuum* cv. chorro contained similar type an relative abundance of metabolites (16.6% in the loading score plot), while other similarities were observed among the cultivars *C. annuum* cv. caribe, *C. annuum* cv. hungaro, *C. annuum* cv. poblano, *C. annuum* cv. anaheim and *C. annuum* cv. chilaca. In contrast to these results, a significant dissimilarity was determined for *C. annuum* cv. anaheim, *C. annuum* cv. chorro, *C. annuum* cv. hungaro and *C. annuum* cv. poblano and the group composed by *C. annuum* cv. arbol, *C. annuum* cv. serrano, *C. annuum* cv. chilaca and *C. annuum* cv. caribe (in accordance with PC1 = 32.5%; Fig. 3C).

To optimize the separation among cultivars, a Linear Discriminant Analysis (LDA) was carried out. The cultivars *C. annuum* cv. agua and *C. annuum* cv. mirasol were excluded considering its designation as outliers in the PCA previously performed (Fig. 3A). After LDA, four principal groups were delimited according to the two components LD1 = 38.64% and LD2 = 27.09%. In the first quadrant, the cultivars *C. annuum* cv. poblano were found whereas in the second quadrant *C. annuum* cv. arbol, *C. annuum* cv. cuaresmeño, *C. annuum* cv. serrano were located. In the third quadrant only *C. annuum* cv. chilaca was observed and finally, in the fourth quadrant the cultivars *C. annuum* cv. caribe, *C. annuum* cv. chorro were observed. Interestingly, the cultivar *C. annuum* cv. anaheim was located between the first and the fourth quadrant while the *C. annuum* cv. hungaro was placed between the third and fourth quadrants (Fig. 4).

In order to understand the similarities of the studied cultivars, a clustered heatmap hierarchical clustering analysis was done (Fig. 5). The analysis showed the average profile of each compound in the studied cultivars. Red lines in this graphic indicates a high amount of

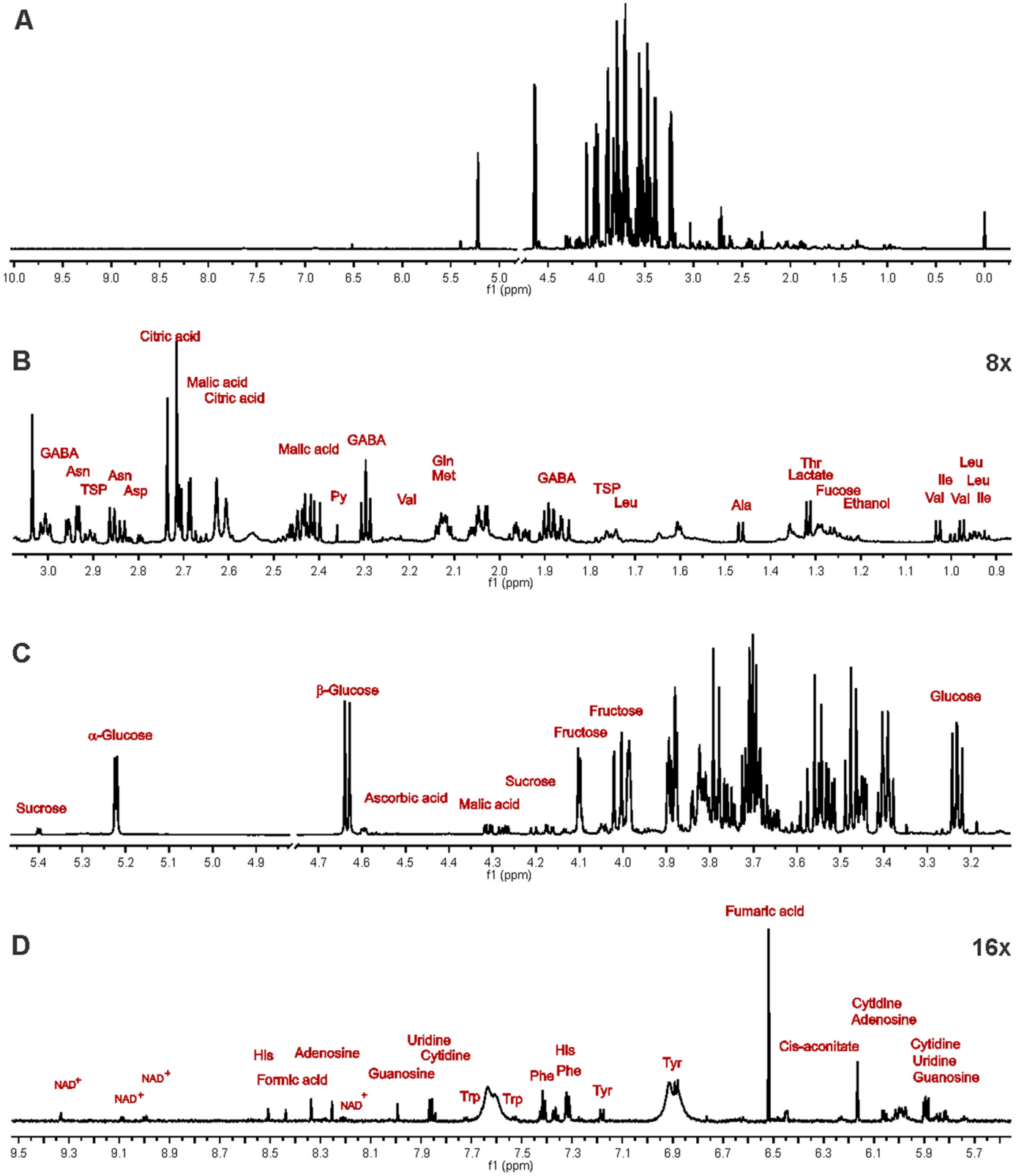

Fig. 1. Characteristic ^{1}H NMR spectrum obtained at 750 MHz from aqueous extracts of *Capsicum annuum* cv. chilaca. Signal assignments were based on 2D NMR experiments and the literature (Becerra-Martínez et al., 2017; Villa-Ruano et al., 2018).

metabolites whereas blue lines indicates the opposite case. The corresponding dendrogram in this figure showed a similar grouping pattern than those of LDA (Fig. 4). Two groups were obtained from this dendrogram, which were denominated as A, and B. In the group A, three subgroups were separated; the first subgroup contained the cultivars *C. annuum* cv. anaheim, *C. annuum* cv. chorro and *C. annuum* cv. poblano. The second subgroup comprised the cultivars *C. annuum* cv. caribe and *C. annuum* cv. hungaro. The third one was conformed for *C. annuum* cv.

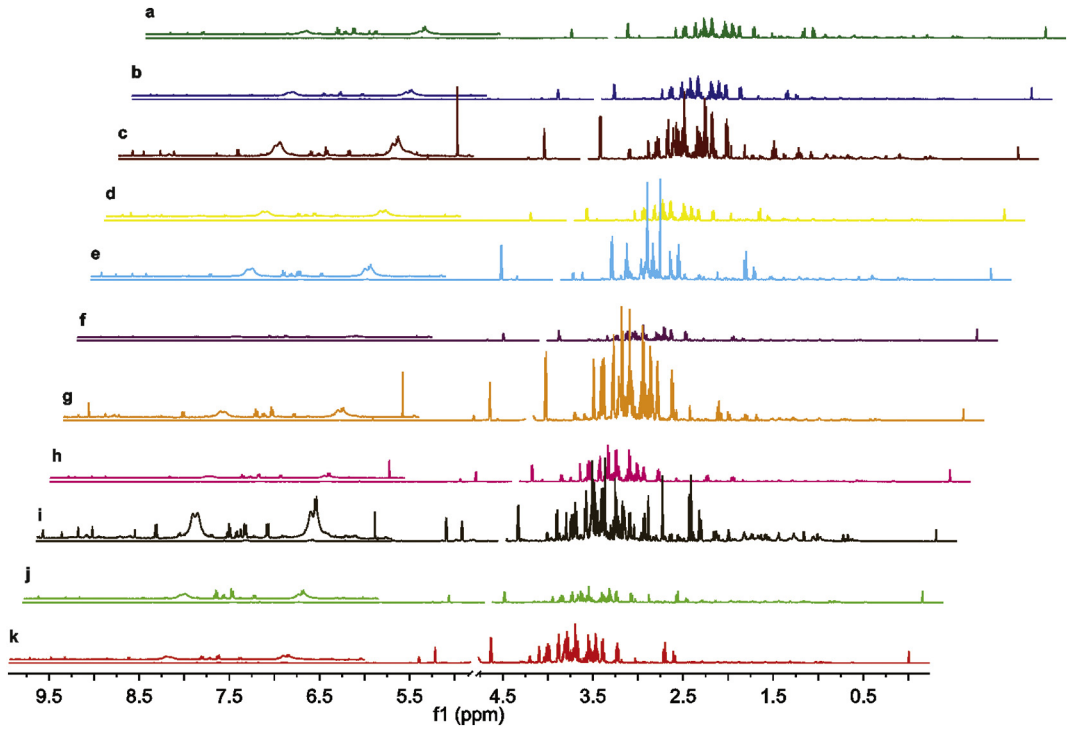

Fig. 2. Representative 750 MHz ^{1}H NMR spectra of eleven chili pepper cultivars belonging to *Capsicum annuum*. (a) *C. annuum* cv. agua, (b) *C. annuum* cv. anaheim, (c) *C. annuum* cv. árbol, (d) *C. annuum* cv. caribe, (e) *C. annuum* cv. chilaca, (f) *C. annuum* cv. chorro, (g) *C. annuum* cv. cuaresmeño, (h) *C. annuum* cv. húngaro, (i) *C. annuum* cv. mirasol, (j) *C. annuum* cv. poblano, (k) *C. annuum* cv. serrano.

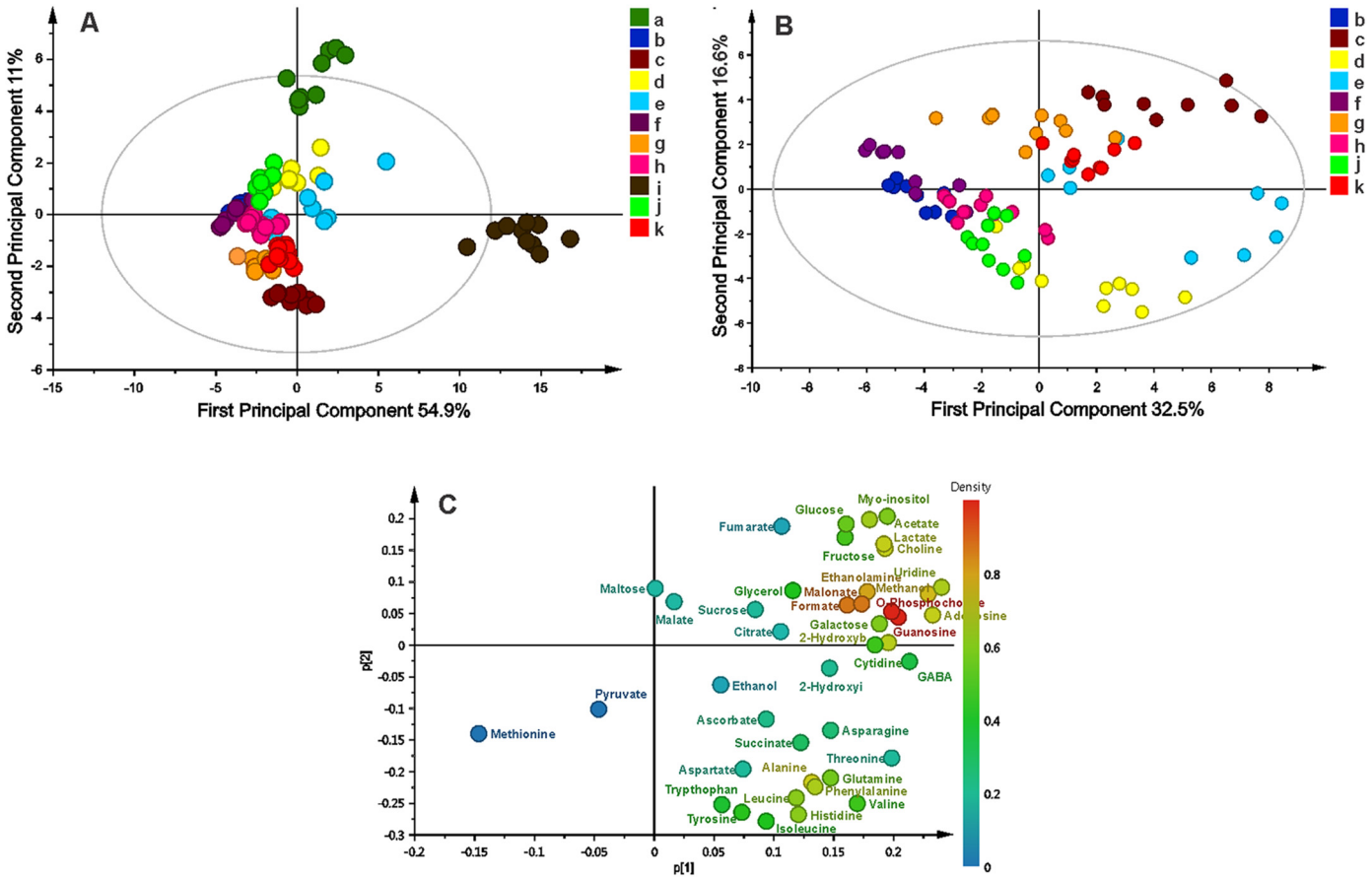

Fig. 3. A) General PCA. B) PCA score plot without outliers and, C) the corresponding loading of scatter plots generated from the 1H NMR spectra (750 MHz) of *Capsicum annuum* cultivars.). (a) *C. annuum* cv. agua, (b) *C. annuum* cv. anaheim, (c) *C. annuum* cv. árbol, (d) *C. annuum* cv. caribe, (e) *C. annuum* cv. chilaca, (f) *C. annuum* cv. chorro, (g) *C. annuum* cv. cuaresmeño, (h) *C. annuum* cv. húngaro, (i) *C. annuum* cv. mirasol, (j) *C. annuum* cv. poblano, (k) *C. annuum* cv. serrano.

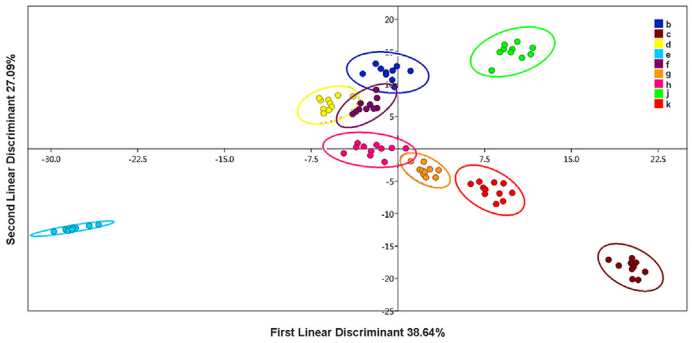

Fig. 4. Linear Discriminant Analysis (LDA) generated from the ^{1}H NMR spectra (750 MHz) of *Capsicum annuum cultivars.* (b) *C. annuum* cv. anaheim, (c) *C. annuum* cv. árbol, (d) *C. annuum* cv. caribe, (e) *C. annuum* cv. chilaca, (f) *C. annuum* cv. chorro, (g) *C. annuum* cv. cuaresmeño, (h) *C. annuum* cv. húngaro, (j) *C. annuum* cv. poblano, (k) *C. annuum* cv. serrano.

cuaresmeño and *C. annuum* cv. serrano. In the group B, only the cultivars *C. annuum* cv. árbol and *C. annuum* cv. chilaca were included.

3.3. Relative abundance of metabolites in Capsicum extracts and NMR analyses

Amino acids are the most abundant group of compounds observed in *Capsicum* species and its cultivars (Table S2). The presence and high abundance of these metabolites has been reported in other works (Jarret, Berke, Baldwin, & Antonious, 2009; Khan et al., 2013; Okunlola, Akinwale, & Adelusi, 2016). Alanine, asparagine, aspartic acid, glutamine, histidine, isoleucine, leucine, phenylalanine, threonine, tryptophan, tyrosine, valine, and GABA were detected in all samples but interestingly methionine was limited to some cultivars of *Capsicum annuum*, such as *C. annuum* cv. anaheim, *C. annuum* cv. caribe, *C. annuum* cv. chorro, *C. annuum* cv. húngaro, *C. annuum* cv. *mirasol, and C. annuum* cv. Poblano (Table S2). Remarkably, the amino acid content was more abundant in both *C. annuum* cv. mirasol and *C. annuum* cv. agua than in the rest of the chili peppers studied (Fig. S7). Thus, *Capsicum annuum* cv. mirasol and *C. annuum* cv. agua were observed as outliers in the PCA analysis (Fig. 3A). From the studied cultivars, *Capsicum annuum* cv. mirasol was the best source of essential and non-essential amino acids for human consumption.

Sugars and organic acids are key factors involved in the organoleptic properties of chili pepper. The content of these compounds usually varies in accordance with the specific genotype and ripening stage of the chili peppers. It is also known that growing conditions and post-harvesting stress can affect the endogenous levels of these metabolites (Ritota et al., 2010). Organic acids also contribute to the characteristic fruity flavor and indicate fruit maturity degree (Aizat et al., 2014; Luning et al., 1994; Soyer, Koca, & Karadeniz, 2003). This study identified some organic acids (Table S2) and their relative ratios among the studied cultivars based on the corresponding peak areas observed in their 750 MHz ^{1}H NMR spectra. The levels of malic acid were particularly high in *C. annuum* cv. húngaro, and fumaric acid was high in *C. annuum* cv. árbol (Fig. S8). These organic acids might be considered as putative biomarkers of these cultivars.

According to López-Hernández, Oruña-Concha, Simal-Lozano, Vázquez-Blanco, and González-Castro (1996), samples of *Capsicum annuum* grown in Galicia contain 45% free sugars including fructose and other reducing sugars such as glucose. Contrary to the present study, sucrose was not reported by these authors. In this and other studies, it was suggested that the presence of these monosaccharides is correlated with sweetness and ripeness (Birch & Pepper, 1983; Luning, Rijk, Withers, & Roozen, 1994; López-Hernández et al., 1996; Osuna-García, Wall, & Waddell, 1998). Glucose and fructose were detected in high concentrations, while galactose and sucrose were only found in a small concentration (Fig. S9). Some cultivars such as *Capsicum annuum* cv. agua and *Capsicum annuum* cv. anaheim revealed a little concentration of sucrose. Fructose and glucose were detected in all samples; however, their lowest abundance was seen in *C. annuum* cv. chorro (Fig. S9). On the contrary, the highest relative abundance of galactose was observed in *C. annuum* cv. mirasol. Interestingly, the sucrose levels were remarkably higher in *C. annuum* cv. mirasol; this strongly suggests than invertase activity was markedly lower in this cultivar at least under the current study conditions. This observation suggests that *C. annuum* cv. mirasol could be the sweetest cultivar studied here; nevertheless, the high endogenous levels of citric and formic acids should affect the

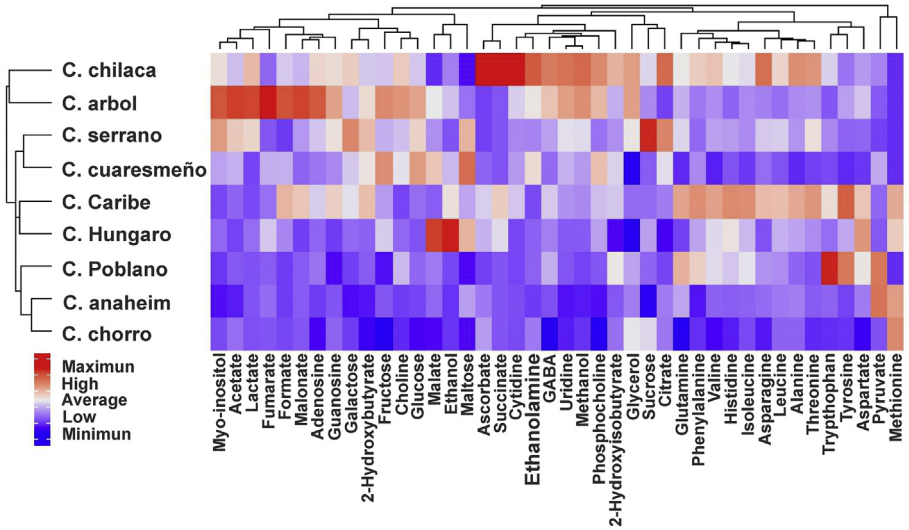

Fig. 5. Heatmap, a clustered heat map for nine cultivars of *Capsicum annuum*. The vertical dendrogram shows the cultivars and the horizontal labels represent the compounds clustered by similarity.

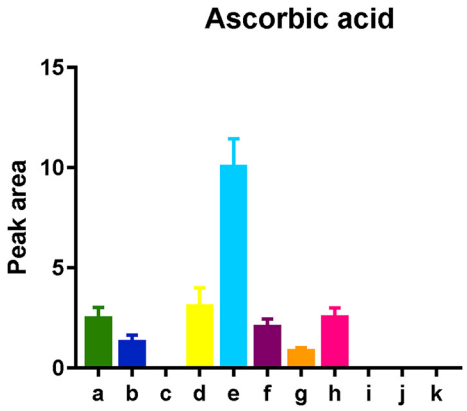

Fig. 6. Relative abundance of ascorbic acid in eleven cultivars of *Capsicum annuum*, which was calculated from the average area of their corresponding signals observed in the ^{1}H NMR spectrum (750 MHz). (a) *C. annuum* cv. agua, (b) *C. annuum* cv. anaheim, (c) *C. annuum* cv. árbol, (d) *C. annuum* cv. caribe, (e) *C. annuum* cv. chilaca, (f) *C. annuum* cv. chorro, (g) *C. annuum* cv. cuaresmeño, (h) *C. annuum* cv. húngaro, (i) *C. annuum* cv. mirasol, (j) *C. annuum* cv. poblano, (k) *C. annuum* cv. serrano.

organoleptic properties of this cultivar (Fig. S8). Despite this effect could be occurring, a sensory evaluation should be performed in order to endorse this hypothesis (Aprea et al., 2017). Considering the high levels of sucrose, glucose, and fructose in *C. annuum* cv. mirasol, osmotic dehydration might be occurring to consequently reduce the shelf life of this chili pepper cultivar. Thus, sugar content could be the key factor to differentiate chili pepper samples. For instance, sucrose and galactose can be considered biomarkers in *C. annuum* cv. mirasol because of their high concentrations (Fig. S9).

Ascorbic acid is clearly the predominant metabolite in some of the studied cultivars (Fig. 6). Surprisingly, *C. annuum* cv. chilaca contained almost three-fold more endogenous levels of vitamin C than the other samples. This result suggested that glucose is rapidly converted into ascorbic acid in *C. annuum* cv. chilaca (Barata-Soares, Gomez, de Mesquita, & Lajolo, 2004). However, the possible influence of biotic and abiotic factors in the biosynthesis of vitamin C in this cultivar must be further studied. Controversially, most popular chili peppers consumed in México such as *C. annuum* cv. árbol, *C. annuum* cv. mirasol, *C. annuum* cv. poblano, and *C. annuum* cv. serrano show no vitamin C levels. According to other studies (Teodoro et al., 2013), ascorbic acid content may vary depending on fruit ripening and genotype with endogenous levels of 43–247 mg/100 g FW. The abundance of this antioxidant in aqueous extracts could be a practical marker to discriminate between species and to separate *C. annuum* cv. chilaca from other chili pepper cultivars harvested under open field conditions.

4. Conclusions

We report the ^{1}H NMR-based metabolomic profiling of eleven cultivars of *Capsicum annum* produced in Mexico and consumed worldwide. The results demonstrate that the combination of ^{1}H NMR spectral data and multivariate analyses are a powerful tool to obtain metabolite fingerprints in chili peppers. The analytical strategy revealed a simultaneous determination of carbohydrates, amino acids, organic acids, and vitamin C. The nutritional content and nutraceutical potential of chili peppers can be easily obtained considering the simplicity, reproducibility, and robustness of this method. This work provides valuable metabolomics data from chili pepper samples cultivated under open field conditions; however, metabolomics fingerprinting under controlled conditions should be studied further.

Acknowledgments

This research received financial support from, SIP-IPN 20170405, SIP-IPN 20180955.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodres.2019.03.025.

References

Aizat, W. M., Dias, D. A., Stangoulis, J. C. R., Able, J. A., Roessner, U., & Able, A. J. (2014). Metabolomics of *Capsicum* ripening reveals modification of the ethylene related-pathway and carbon metabolism. *Postharvest Biology and Technology, 89*, 19–31.
Anastasiadi, M., Zira, A., Magiatis, P., Haroutounian, S. A., Skaltsounis, A. L., & Mikros, E. (2009). ^{1}H NMR-based metabonomics for the classification of Greek wines according to variety, region, and vintage. comparison with HPLC data. *Journal of Agricultural and Food Chemistry, 57*, 11067–11074.
Aprea, E., Charles, M., Endrizzi, I., Corollaro, M. L., Betta, M., Biasioli, F., & Gasperi, F. (2017). Sweet taste in apple: The role of sorbitol, individual sugars, organic acids and volatile compounds. *Scientific Reports, 7*(44950), 1–10.
Barata-Soares, A. D., Gomez, M. L. P. A., de Mesquita, H. C., & Lajolo, F. M. (2004). Ascorbic acid biosynthesis: A precursor study on plants. *Brazilian Journal of Plant Physiology, 16*, 147–154.
Becerra-Martínez, E., Florentino-Ramos, E., Pérez-Hernández, N., Zepeda-Vallejo, L. G., Villa-Ruano, N., Velázquez-Ponce, M., ... Bañuelos-Hernández, A. E. (2017). ^{1}H NMR-based metabolomic fingerprinting to determine metabolite levels in serrano peppers (*Capsicum annuum* L) grown in two different regions. *Food Research International, 102*, 163–170.
Beltran, A., Suarez, M., Rodriguez, M. A., Vinaixa, M., Samino, S., Arola, L., ... Yanes, O. (2012). Assessment of compatibility between extraction methods for NMR-and LC/MS-based metabolomics. *Analytical Chemistry, 84*, 5838–5844.
Birch, G. G., & Pepper, T. (1983). Protection of vitamin C by sugars and their hydrogenated derivatives. *Journal of Agricultural and Food Chemistry, 31*, 980–985.
Consonni, R., Cagliani, L. R., & Cogliani, C. (2012). NMR characterization of saccharides in Italian honeys of different floral sources. *Journal of Agricultural and Food Chemistry, 60*, 4526–4534.
Dona, A. C., Kyriakides, M., Scott, F., Shephard, E. A., Varshavi, D., Veselkov, K., & Everett, J. R. (2016). A guide to the identification of metabolites in NMR-based metabonomics/metabolomics experiments. *Computational and Structural Biotechnology Journal, 14*, 135–153.
Eriksson, L., Johansson, E., Trygg, J., & Vikström, C. (2013). *Multi and megavariate data analysis: Basic principles and application* (4th ed.). Sweden: Umetrics Academy (Chapter 1, 8).
Georgiev, M. I., Ali, K., Alipieva, K., Verpoorte, R., & Choi, Y. H. (2011). Metabolic differentiations and classification of *Verbascum* species by NMR-based metabolomics. *Phytochemistry, 72*, 2045–2051.
González-Zamora, A., Sierra-Campos, E., Luna-Ortega, J. G., Pérez-Morales, R., Rodríguez-Ortiz, J. C., & García-Hernández, J. L. (2013). Characterization of different *Capsicum* varieties by evaluation of their capsaicinoids content by high performance liquid chromatography, determination of pungency and effect of high temperature. *Molecules, 18*, 13471–13486.
Gromski, P. S., Xu, Y., Kotze, H. L., Correa, E., Ellis, D. I., Armitage, E. G., ... Goodacre, R. (2014). Influence of missing values substitutes on multivariate analysis of metabolomics data. *Metabolites, 4*, 433–452.
Hervert-Hernandez, D., Sayago-Ayerdi, S. G., & Goñi, I. (2010). Bioactive compounds of four hot pepper varieties (*Capsicum annuum* L.), antioxidant capacity, and intestinal bioaccessibility. *Journal of Agricultural and Food Chemistry, 58*, 3399–3406.
Hohmann, M., Christoph, N., Wachter, H., & Holzgrabe, U. (2014). ^{1}H NMR profiling as an approach to differentiate conventionally and organically grown tomatoes. *Journal of Agricultural and Food Chemistry, 62*, 8530–8540.
Jang, Y. K., Jung, E. S., Lee, H., Choi, D., & Lee, C. H. (2015). Metabolomic characterization of hot pepper (*Capsicum annuum* "CM334") during fruit development. *Journal of Agricultural and Food Chemistry, 63*, 9452–9460.
Jang-Eun, L., Bum-Jin, L., Jin-Oh, C., Jeong-Ah, H., Sang-Jun, L., Cherl-Ho, L., & Young-Shick, H. (2010). Geographical and climatic dependencies of green tea (*Camellia sinensis*) metabolites: A ^{1}H NMR-based metabolomics study. *Journal of Agricultural and Food Chemistry, 58*, 10582–10589.
Jarret, R. L., Berke, T., Baldwin, E. A., & Antonious, G. F. (2009). Variability for free sugars and organic acids in *Capsicum chinense*. *Chemistry & Biodiversity, 6*, 138–145.
Jung, Y., Lee, J., Kwon, J., Lee, K. S., Ryu, D. H., & Hwang, G. S. (2010). Discrimination of the geographical origin of beef by ^{1}H NMR-based metabolomics. *Journal of Agricultural and Food Chemistry, 58*, 10458–10466.
Khan, A. L., Sang-Mo, K., Dhakal, K. H., Hussain, J., Adnan, M., Kim, M., & Lee, I. (2013). Flavonoids and amino acid regulation in *Capsicum annuum* L. by endophytic fungi under different heat stress regimes. *Scientia Horticulturae, 155*(1–7).
Koda, M., Furihata, K., Wei, F., Miyakawa, T., & Tanokura, M. (2012). Metabolic discrimination of mango juice from various cultivars by band-selective NMR spectroscopy. *Journal of Agricultural and Food Chemistry, 60*, 1158–1166.
Lee, Y., Howard, L. R., & Villalon, B. (1995). Flavonoids and antioxidant of fresh pepper (*Capsicum annuum*) cultivars. *Journal of Food Science, 60*, 473–477.

López-Hernández, J., Oruña-Concha, M. J., Simal-Lozano, J., Vázquez-Blanco, M. E., & González-Castro, M. J. (1996). Chemical composition of padron peppers (*Capsicum annuum* L.) grown in Galicia (N.W. Spain). *Food Chemistry, 51*, 557–559.

Luning, P. A., Vuurst de Vries, R., Yuksel, D., Ebbenhorst-Seller, T., Withers, H. J., & Roozen, J. P. (1994). Combined instrumental and sensory evaluation of flavor of fresh bell peppers *(Capsicum annuum)* harvested at three maturation stages. *Journal of Agricultural and Food Chemistry, 42*, 2855–2861.

Luning, P. A., Rijk, T., Withers, H. J., & Roozen, J. P. (1994). Gas chromatography, mass spectrometry, and sniffing port analyses of volatile compounds of fresh bell peppers *(Capsicum annuum)* at different ripening stages. *Journal of Agricultural and Food Chemistry, 42*, 977–983.

Maji, A. K., & Banerji, P. (2016). Phytochemistry and gastrointestinal benefits of the medicinal spice, *Capsicum annuum* L. (Chilli): A review. *Journal of Complementary & Integrative Medicine, 13*, 97–122.

Montes-Hernández, S. (2010). *Recopilación y análisis de la información existente de las especies del género Capsicum que crecen y se cultivan en México. Informe final de Proyecto.* Campo Experimental Bajío, INIFAP1–84.

Normas Oficiales Mexicanas (2015). Productos alimenticios no industrializados para consumo humano-chile fresco (*Capsicum* spp.). *Technical Report* **NMX-FF-025-SCFI-2014**, Normas Oficiales Mexicanas, México.

Okunlola, G. O., Akinwale, R. O., & Adelusi, A. A. (2016). Proline and soluble sugars accumulation in three pepper species (*Capsicum* spp.) in response to water stress imposed at different stages of growth. *Sciences in Cold and Arid Regions, 8*, 205–211.

Osuna-García, J. A., Wall, M. M., & Waddell, C. A. (1998). Endogenous levels of tocopherols and ascorbic acid during fruit ripening of new mexican-type Chile (*Capsicum annuum* L.) cultivars. *Journal of Agricultural and Food Chemistry, 46*, 5093–5096.

Picone, G., Laghi, L., Gardini, F., Lanciotti, R., Siroli, L., & Capozzi, F. (2013). Evaluation of the effect of carvacrol on the scherichia coli 555 metabolome by using ^{1}H-NMR spectroscopy. *Food Chemistry, 141*, 4367–4374.

Picone, G., Trimigno, A., Tessarin, P., Donnini, S., Rombolà, A. D., & Capozzi, F. (2016). 1H NMR foodomics reveals that the biodynamic and the organic cultivation managements produce different grape berries (Vitis vinifera L. cv. Sangiovese). *Food Chemistry, 213*, 187–195.

Pino, J., González, M., Ceballos, L., Centurion-Yah, A. R., Trujillo-Aguirre, J., Latournerie-Moreno, L., & Sauri-Duch, E. (2007). Characterization of total capsaicinoids, colour and volatile compounds of habanero chilli pepper (*Capsicum chinense*) cultivars grown in Yucatan. *Food Chemistry, 104*, 1682–1686.

Ritota, M., Marini, F., Sequi, P., & Valentini, M. (2010). Metabolomic characterization of Italian sweet pepper (*Capsicum annuum* L.) by means of HRMAS-NMR spectroscopy and multivariate analysis. *Journal of Agricultural and Food Chemistry, 58*, 9675–9684.

Rodrigues, D., Santos, C. H., Rocha-Santos, T. A. P., Gomes, A. M., Goodfellow, B. J., & Freitas, A. C. (2011). Metabolic profiling of potential probiotic or synbiotic cheeses by nuclear magnetic resonance (NMR) spectroscopy. *Journal of Agricultural and Food Chemistry, 59*, 4955–4961.

Savorani, F., Rasmussen, M. A., Mikkelsen, M. S., & Engelsen, S. B. (2013). A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. *Food Research International, 54*, 131–145.

Serrano, M., Zapata, P. J., Castillo, S., Guillén, F., Martínez-Romero, D., & Valero, D. (2010). Antioxidant and nutritive constituents during sweet pepper development and ripening are enhanced by nitrophenolate treatments. *Food Chemistry, 118*, 497–503.

Soyer, Y., Koca, N., & Karadeniz, F. (2003). Organic acid profile of Turkish white grapes and grape juices. *Journal of Food Composition and Analysis, 16*, 629–636.

Tan, C. K., Ali, Z. M., & Zainal, Z. (2012). Changes in ethylene production, carbohydrase activity and antioxidant status in pepper fruits during ripening. *Scientia Horticulturae, 142*, 23–31.

Teodoro, A. F. P., Alves, R., Ribeiro, L. B., Reis, K., Reifschneider, F. J., Fonseca, M. E., ... Agostini-Costa, T. S. (2013). Vitamin C content in habanero pepper accessions (*Capsicum chinense*). *Horticultura Brasileira, 31*, 59–62.

Trimigno, A., Münger, L., Picone, G., Freiburghaus, C., Pimentel, G., Vionnet, N., ... Vergères, G. (2018). GC-MS based metabolomics and NMR spectroscopy investigation of food intake biomarkers for, milk and cheese in serum of healthy humans. *Metabolites, 8*, 1–27.

Troconis-Torres, I. G., Rojas-López, M., Hernández-Rodríguez, C., Villa-Tanaca, L., Maldonado-Mendoza, I. E., Dorantes-Álvarez, L., ... Jaramillo-Flores, M. E. (2012). Biochemical and molecular analysis of some commercial samples of chilli peppers from Mexico. *Journal of Biomedicine and Biotechnology,* 1–11.

Van Dorsten, F. A., Daykin, C. A., Mulder, T. P. J., & Van Dyunhoven, J. M. (2006). Metabonomics approach to determine metabolic differences between green tea and black tea consumption. *Journal of Agricultural and Food Chemistry, 54*, 6929–6938.

Vega-Gálvez, A., Di Scala, K., Rodríguez, K., Lemus-Mondaca, R., Miranda, M., López, J., & Pérez-Won, M. (2009). Effect of air-drying temperature on physico-chemical properties, antioxidant capacity, colour and total phenolic content of red pepper (*Capsicum annuum*, L. var. Hungarian). *Food Chemistry, 117*, 647–653.

Verpoorte, R., Choi, Y. H., Mustafa, N. R., & Kim, H. K. (2008). Metabolomics: Back to basics. *Phytochemistry Reviews, 7*, 525–537.

Villa-Ruano, N., Velásquez-Valle, R., Zepeda-Vallejo, L. G., Pérez-Hernández, N., Velázquez-Ponce, M., Arcos-Adame, V. M., & Becerra-Martinez, E. (2018). ^{1}H NMR-based metabolomic profiling for identification of metabolites in *Capsicum annuum* cv. mirasol infected by beet mild curly top virus (BMCTV). *Food Research International, 106*, 870–877.

Wahyuni, Y., Stahl-Hermes, V., Ballester, A., De Vos, R. C. H., Voorrips, R. E., Maharijaya, A., ... Bovy, A. G. (2014). Genetic mapping of semi-polar metabolites in pepper fruits (*Capsicum* spp.): Towards unravelling the molecular regulation of flavonoid quantitative trait loci. *Molecular Breeding, 33*, 503–518.