Research Article

Phytochemical Analysis

Received: 10 January 2017, Revised: 28 March 2017, Accepted: 28 March 2017 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/pca.2692

Untargeted Metabolomic Analysis of *Capsicum* spp. by GC–MS

Bianca Camargo Aranha,[a] Jessica Fernanda Hoffmann,[a] Rosa Lia Barbieri,[b] Cesar Valmor Rombaldi[a] and Fábio Clasen Chaves[a]*

ABSTRACT:

Introduction – In order to conserve the biodiversity of *Capsicum* species and find genotypes with potential to be utilised commercially, Embrapa Clima Temperado maintains an active germplasm collection (AGC) that requires characterisation, enabling genotype selection and support for breeding programmes.

Objective – The objective of this study was to characterise pepper accessions from the Embrapa Clima Temperado AGC and differentiate species based on their metabolic profile using an untargeted metabolomics approach.

Material and Methods – Cold (-20°C) methanol extraction residue of freeze-dried fruit samples was partitioned into water/methanol (A) and chloroform (B) fractions. The polar fraction (A) was derivatised and both fractions (A and B) were analysed by gas chromatography coupled to mass spectrometry (GC–MS). Data from each fraction was analysed using a multivariate principal component analysis (PCA) with XCMS software.

Results – Amino acids, sugars, organic acids, capsaicinoids, and hydrocarbons were identified. Outlying accessions including P116 (*C. chinense*), P46, and P76 (*C. annuum*) were observed in a PCA plot mainly due to their high sucrose and fructose contents. PCA also indicated a separation of P221 (*C. annuum*) and P200 (*C. chinense*), because of their high dihydrocapsaicin content.

Conclusions – Although the metabolic profiling did not allow for grouping by species, it permitted the simultaneous identification and quantification of several compounds complementing and expanding the metabolic database of the studied *Capsicum* spp. in the AGC. Copyright © 2017 John Wiley & Sons, Ltd.

Keywords: Metabolic profiling; organic acids, capsaicinoids; sugars; amino acids

Introduction

The genus *Capsicum* is native to humid tropical zones of Central and South America, belongs to the Solanaceae family, and encompasses more than 30 perennial species, five of which are domesticated: *C. annuum*, *C. frutescens*, *C. chinense*, *C. pubescens*, and *C. baccatum* (Heiser and Pickersgill, 1969). Since the original proposition by Linnaeus of two species (*C. annuum* and *C. frutescens*), many studies have dealt with speciation at different levels (genetic, chemical, molecular) and the current classification is primarily based on morphological characteristics. These species bear fruit varying in shape, colour, size, and degree of pungency (Smith and Heiser, 1951). *Capsicum* fruit are a source of many health beneficial compounds, such as ascorbic acid (vitamin C), carotenoids (provitamin A), tocopherols (vitamin E), flavonoids, and capsaicinoids (Howard and Wildman, 2006). The presence and content of these compounds can vary depending on biotic and abiotic factors including species, climate, and soil conditions. Metabolomic analysis of a complex matrix allows for the characterisation and quantification of the majority of the compounds present in an extract (De Vos *et al*., 2007; Hoffmann *et al*., 2017).

Metabolomic analysis emerged as an important tool for identification and comparison of a large number of compounds in biological systems. However, plants are particularly complex as they are composed of a wide variety of metabolites, which makes it impossible for any analytical method to simultaneously extract and detect all of the compounds present in the matrix (De Vos *et al*., 2007). Targeted or untargeted metabolomics has been used to obtain information on metabolic variability of germplasm, both to compare species within a genus and to establish the origin of different individuals (Hoffmann *et al*., 2017). Evidence of an effect of geographic environment on pungency within species has been demonstrated (Tewksbury *et al*., 2006). In addition, the capsaicinoid profile has been ruled out as a chemotaxonomic indicator for *Capsicum* species (Zewdie and Bosland, 2001). A group of accessions of different *Capsicum* species from various locations were analysed by a targeted approach, and results indicated that the metabolic variation was independent of species and geographical location (Wahyuni *et al*., 2011). This collection consisted of 32 accessions and when later analysed for their volatile profile, the profile also did not provide chemotaxonomic separation by species (Wahyuni *et al*., 2013). However, when the same 32 genotypes were further investigated using an untargeted metabolomics approach by liquid chromatography coupled to mass spectrometry (LC–MS), the investigators proposed they

* Correspondence to: Fábio Clasen Chaves, Graduate Programme in Food Science and Technology, Department of Agroindustrial Science and Technology, Faculty of Agronomy Eliseu Maciel, Federal University of Pelotas, Caixa Postal 354, CEP 96010-900, Pelotas, Brazil.
Email: fabio.chaves@ufpel.edu.br

[a] Graduate Programme in Food Science and Technology, Department of Agroindustrial Science and Technology, Faculty of Agronomy Eliseu Maciel, Federal University of Pelotas, Caixa Postal 354, CEP, 96010-900, Pelotas, Brazil

[b] Embrapa Clima Temperado, Caixa Postal 403, CEP, 96001-970, Pelotas, Brazil

grouped by species according to their semi-polar profile which did not include capsaicinoid but did include phenolics, flavonoids, and capsianosides (Wahyuni *et al.*, 2013). Chromatography is a separation technique that becomes increasingly more powerful when coupled to MS, for it can also provide structural information and metabolite identification. While LC–MS usually employs a more accurate mass analyser with higher mass resolution and milder ionisation steps, gas chromatography coupled to mass spectrometry (GC–MS) is applicable to volatile and thermally stable compounds and derivatisation reactions may make some compounds more amenable to GC conditions. GC–MS commonly utilises electron ionisation (EI) with constant energy to create reproducible fragmentation patterns, facilitating compound identification (Lisec *et al.*, 2006; De Vos *et al.*, 2007). In this context, untargeted metabolomics using GC–MS were used in the current study to determine the metabolic profile of different *Capsicum* species in order to explore the metabolic potential of accessions of the active germplasm collection and to establish whether species can be distinguished based on their metabolic profile.

Experimental

Chemicals

All the reagents and solvents used in this study were HPLC grade purchased from Sigma-Aldrich (St Louis, MO, USA).

Samples

Sixty accessions from four different *Capsicum* species [*C. baccatum* (34), *C. chinense* (12), *C. annuum* (13), and *C. frutescens* (1)] from the Embrapa Clima Temperado active germplasm collection (AGC) with high genetic variability originally collected from different regions in Brazil were used. Ripe fruit were collected at 9 a.m., frozen in liquid nitrogen, chopped, freeze-dried (Liobras, L101; Liobras, Brazil), and macerated in a ball mill (Marconi, MA350; Marconi, Brazil). Tissue was stored at -80°C until analysed.

Extract preparation

Capsicum fruit were extracted following the methodology proposed by Lisec *et al.* (2006). Therefore, 30 mg of lyophilised sample was extracted with 1.4 mL pre-cooled methanol (-20°C). Then 60 μL of ribitol (0.2 mg/mL in ultrapure water) and 10 μL of anthrone (1 mg/mL in chloroform) were added and vortexed for 10 s. The mixture was incubated in a thermomixer (Eppendorf thermomixer 5436, Eppendorf, Germany) for 10 min at 70°C and 900 rpm. Samples were centrifuged for 10 min at 11000×*g* and the supernatant was mixed with 1.5 mL of ultrapure water (cooled to 4°C) and 0.75 mL chloroform (-20°C) and vortexed for 10 s. The sample was centrifuged (Sorvall, RC5C, DuPont, USA) for 15 min at 22000×*g* and 150 μL of the upper phase (fraction A) was saved for a later derivatisation step. Thus, 650 μL of the bottom layer (fraction B) was placed in GC vials for injection. Subsequently, fraction A (150 μL) was concentrated to dryness by applying nitrogen gas and 40 μL of methoximation reagent (20 mg/mL methoxyamine hydrochloride in pyridine) was added and the mixture was stirred for 2 h at 37°C in a thermomixer. Then 70 μL of MSTFA [*N*-methyl-*N*-(trimethylsilyl) trifluoroacetamide] was added and the sample was stirred for 30 min at 37°C. The reaction was transferred to GC vials with 250 μL inserts and immediately injected into the GC.

GC–MS parameters for metabolic profiling

A Shimadzu GCMS QP2010 Ultra (Shimadzu, Japan) equipped with auto injector AOC-20i and NIST 2011 mass spectrum library was used. Chromatography and MS parameters followed the protocol proposed by Lisec *et al.* (2006). Samples were injected (1 μL) with the injector temperature set at 230°C, in either a split mode (1:50) for fraction A or a splitless mode for fraction B, using helium as the carrier gas at 2 mL/min flow and linear velocity as flow control mode. The capillary column used was an Rtx-5MS (30 m × 0.25 mm × 0.25 μm) with temperature programming set at 80°C for 2 min, and a ramp temperature of 15°C per minute until 320°C, then maintained for 6 min. MS parameters were: ion source and interface set at 250°C, mass range 70–600 *m/z* scanned at 0.2 scans per second. Fatty acid methyl esters (FAMEs C8–24) were used to determine a retention time index calculated automatically by the Shimadzu software.

Principal component analysis

Principal component analysis (PCA) was performed using free online software – XCMS Online (access link: https://xcmsonline.scripps.edu/). A multi group job was created selecting databases preloaded in the software in mzXML format. Databases for each species were created and three replicates of each accession belonging to each species were included. No database was selected for the QC Dataset option. Instrumental parameters were GC/single quadrupole, GC-EI-MS of single quadrupole, retention time in minutes, positive polarity, and subsequent submission for parametric analysis of variance (ANOVA).

Statistical analyses

Means and standard deviations were calculated for each parameter by accession. ANOVA and means comparisons by Fisher's least significant difference (LSD) test ($p \leq 0.05$) were performed with Statistical Analysis System (SAS) program.

Results and discussion

Untargeted metabolomic analysis of 60 *Capsicum* spp. accessions was performed on the partitioned fractions, water/methanol (fraction A) and chloroform (fraction B), derived from the resuspended residue of a cold methanol extraction of freeze-dried pepper fruit. PCA revealed the distribution of 60 *Capsicum* accessions based on metabolic profiling of fractions A and B [Figs 1(a) and (b)]. Figure 1(a) shows the PCA score for principal components one (PC1) and two (PC2). The first and second components of the PCA model explained 52% of the total variance in metabolite content among *Capsicum* accessions, where PC1 and PC2 represented 42% and 10%, respectively, of the total variability in the data set. Outlying accessions included P116 (*C. chinense*), P46 (*C. annuum*), and P76 (*C. annuum*). P116 did not group with the majority of accessions because of its elevated sucrose content. Meanwhile, fructose was the metabolite present in high titers in P46 and P76 that influenced their separation from the rest of the accessions.

For fraction B [Fig. 1(b)], the first two principal components represented 46% of the total variation. PC1 had the highest accumulated variation and represented 34% of the total variability in the data set. P221 (*C. annuum*) and P200 (*C. chinense*) diverged from the main cluster, and the compound responsible was dihydrocapsaicin.

Multivariate analysis of compounds including general and specialised metabolites present in each fraction studied did not show separation by species; however, outliers rich in specific compounds were identified. These findings are in agreement with Zewdie and Bosland (2001) and Wahyuni *et al.* (2013), who could not differentiate among species based on their metabolic profile. *Capsicum* taxonomy is primarily based on morphological characteristics and the outcrossing capabilities among species lead to genotype variability. Therefore, the general and specialised

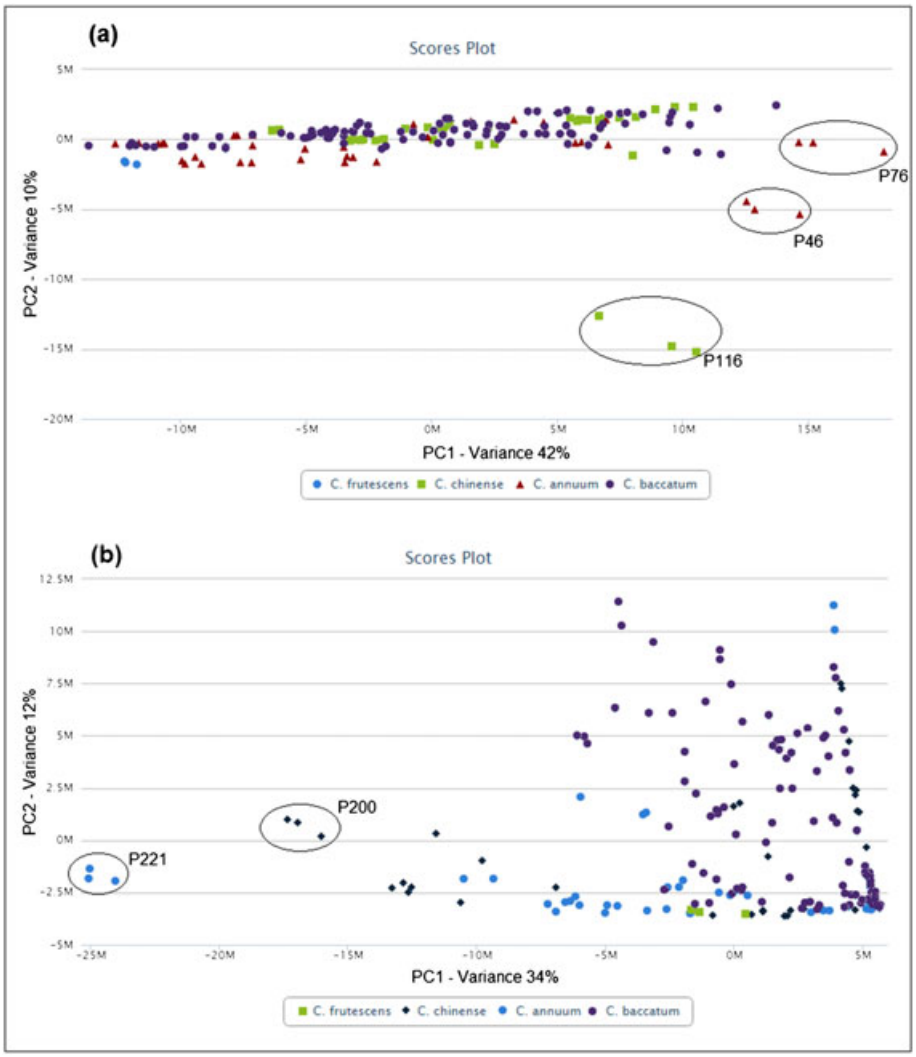

Figure 1. Plotting scores of the principal component analysis (PCA) of polar fraction A – water/methanol (a) and polar fraction B – chloroform fraction from the methanol extraction (b) of four *Capsicum* spp. (*C. annuum*, *C. baccatum*, *C. chinense*, and *C. frutescens*) of the Embrapa Clima Temperado active germplasm collection, Pelotas, RS, Brazil. PC1, principal component one; PC2, principal component two. [Colour figure can be viewed at wileyonlinelibrary.com]

metabolic contents also tend to be variable and not species specific.

Genetic and environmental factors influence the metabolic responses of plants and may be associated with compositional variation among individuals within the same species. Thus, it is ideal to evaluate the maximum possible number of genotypes to better elucidate the characteristics of the species and genus of interest. There are many reports on the capsaicinoid profile of pepper fruit determined by various methodologies (Zewdie and Bosland, 2001). However, few studies both focus on the metabolic profile of pepper extracts of different polarities and include a large sample set.

In this study, several compounds (Tables 1 and 2) were identified in methanolic extracts partitioned into two fractions, methanol/water (fraction A) and chloroform (fraction B). Six classes of compounds were pooled together and quantified: amino acids, sugars, organic acids, hydrocarbons, esters, and capsaicinoids (Figs 2 and 3). Aizat *et al.* (2014) accessed differentially abundant compounds involved in the regulation of non-climacteric fruit ripening via a metabolomics analysis of *Capsicum* by GC–MS and identified sugars, amino acids, organic acids.

Figure 2 shows the quantification of total amino acids, sugars, and organic acids for each genotype as result of the metabolic profile of fraction A. Asparagine was the predominant amino acid in 76% of the studied accessions. It represented on average 27% of the total amino acid content, and varied from 0.7 to 6.17 mg/g. Norvaline was the predominant amino acid in: *C. baccatum* accessions P219 (1.09 mg/g), P274 (1.29 mg/g), P280 (1.03 mg/g), P287 (1.44 mg/g), and P294 (1.78 mg/g); *C. chinense* accessions P236 (0.61 mg/g) and P247 (1.06 mg/g); *C. annuum* accessions P136 (1.16 mg/g) and P189 (1.88 mg/g); and *C. frutescens* accession P82 (1.14 mg/g). In *C. baccatum* accessions P14 and P108 and *C. annuum* accessions P46 and P290, the predominant amino acids were isoleucine (2.25 mg/g), serine (0.76 mg/g and 1.63 mg/g), and proline (1.78 mg/g), respectively. Other amino acids identified included: alanine, valine, tyrosine, glutamine, phenylalanine, threonine, lysine, cysteine, glycine, aspartic acid, and glutamic acid (Table 1). The average content of total amino acids for the studied

Table 1. Derivatised metabolites identified in 60 *Capsicum* spp. accessions by GC–MS analysis

No.	Compound	Class	R_t	Derivatised MW	Reference ions[a]	Retention index[b]
1	Alanine	Amino acid	4.557	233	73 (46); 116 (100); 117 (12); 147 (15); $M^{+\cdot}$ 233 (<1)	780
2	Valine	Amino acid	5.802	261	73 (57); 144 (100); 145 (12); 218 (18); $M^{+\cdot}$ 261 (<1)	901
3	Leucine	Amino acid	6.370	275	73 (38); 102 (10); 158 (100); 159 (14); $M^{+\cdot}$ 275 (<1)	956
4	Isoleucine	Amino acid	6.591	275	73 (53); 158 (100); 159 (14); 218 (19); $M^{+\cdot}$ 275(<1)	977
5	Proline	Amino acid	6.627	259	73 (40); 142 (100); 143 (13); 147 (6); $M^{+\cdot}$ 259 (<1)	981
6	Glycine	Amino acid	6.717	291	73 (48); 86 (21); 147 (26); 174 (100); $M^{+\cdot}$ 291 (<1)	989
7	Serine	Amino acid	7.236	321	73 (94); 100 (25); 204 (100); 218 (61); $M^{+\cdot}$ 321 (<1)	1046
8	Threonine	Amino acid	7.491	335	73 (100); 117 (44); 218 (51); 219 (44); $M^{+\cdot}$ 335 (<1)	1075
9	Norvaline	Amino acid	7.914	273	73 (21); 82 (80); 172 (100); 173 (15); $M^{+\cdot}$ 273 (<1)	1123
10	Butanodioic acid	Organic acid	8.367	350	73 (100); 147 (48); 233 (17); 245 (9); $M^{+\cdot}$ 350 (<1)	1174
11	Aspartic acid	Amino acid	8.634	349	73 (98); 100 (31); 232 (100); 233 (21); $M^{+\cdot}$ 349 (<1)	1202
12	Proline	Amino acid	8.667	273	73 (55); 147 (15); 156 (100); 157 (12); $M^{+\cdot}$ 273 (<1)	1209
13	Butanoic acid	Organic acid	8.704	319	73 (51); 147 (37); 174 (100); 304 (28); $M^{+\cdot}$ 319 (<1)	1214
14	Glutamic acid	Amino acid	9.419	363	73 (69); 128 (39); 147 (22); 246 (100); $M^{+\cdot}$ 363 (3)	1305
15	Phenylalanine	Amino acid	9.506	309	73 (100); 100 (292); 192 (69); 218 (94); $M^{+\cdot}$ 309 (<1)	1317
16	D-(-)-Xylose	Sugar	9.793	463	73 (100); 103 (86); 217 (55); 307 (26); $M^{+\cdot}$ 463 (<1)	1353
17	Asparagine	Amino acid	9.838	348	73 (100); 116 (70); 132 (31); 231 (34); $M^{+\cdot}$ 348 (<1)	1359
18	Xylitol	Sugar	10.306	512	73 (100); 103 (45); 147 (44); 73 (100); $M^{+\cdot}$ 512 (<1)	1421
19	Glutamine	Amino acid	10.581	362	73 (94); 155 (38); 156 (100); 245 (23); $M^{+\cdot}$ 362 (1)	1460
20	1-Cyclohexene-1-carboxylic acid	Organic acid	10.834	462	73 (67); 147 (21); 204 (100); 205 (20); $M^{+\cdot}$ 462 (1)	1496
21	Fructose (isomer 1)	Sugar	10.882	540	73 (100); 147 (23); 217 (69); 437 (18); $M^{+\cdot}$ 540 (<1)	1503
22	Citric acid	Organic acid	10.951	480	73 (100); 147 (56); 273 (94); 347 (20); $M^{+\cdot}$ 480 (<1)	1513
23	Fructose (isomer 2)	Sugar	11.002	540	73 (100); 147 (30); 204 (73); 217 (26); $M^{+\cdot}$ 540 (<1)	1521
24	Fructose (isomer 3)	Sugar	11.380	569	73 (100); 103 (98); 217 (70); 307 (42); $M^{+\cdot}$ 569 (<1)	1574
25	Fructose (isomer 4)	Sugar	11.448	569	73 (100); 103 (87); 217 (70); 307 (42); $M^{+\cdot}$ 569 (<1)	1584
26	D-Allose (isomer 1)	Sugar	11.519	569	73 (100); 147 (40); 205 (49); 319 (51); $M^{+\cdot}$ 569 (<1)	1594
27	Talose (isomer 1)	Sugar	11.565	569	73 (100); 147 (48); 205 (59); 319 (64); $M^{+\cdot}$ 569 (<1)	1600
28	Talose (isomer 2)	Sugar	11.689	569	73 (100); 147 (38); 205 (44); 319 (52); $M^{+\cdot}$ 569 (<1)	1620
29	Tyrosine	Amino acids	11.744	397	73 (42); 218 (100); 219 (17); 280 (13); $M^{+\cdot}$ 397 (<1)	1629
30	3-α-Mannobiose	Sugar	11.923	918	73 (100); 147 (21); 217 (66); 361 (24); $M^{+\cdot}$ 918 (<1)	1657
31	D-Glalactose	Sugar	12.067	540	73 (100); 191 (40); 204 (95); 217 (34); $M^{+\cdot}$ 540 (<1)	1678
32	D-Glucose (isomer)	Sugar	12.393	540	73 (99); 204 (100); 205 (23); 220 (22); $M^{+\cdot}$ 540 (<1)	1730
33	D-Allose (isomer 2)	Sugar	13.092	627	73 (100); 147 (40); 205 (51); 319 (81); $M^{+\cdot}$ 627 (<1)	1844
34	D-(+)-Mannose	Sugar	13.124	627	73 (100); 103 (36); 205 (38); 319 (58); $M^{+\cdot}$ 627 (<1)	1849
35	D-Glucuronic acid	Organic acid	14.415	554	73 (100); 147 (31); 204 (60); 217 (64); $M^{+\cdot}$ 554 (<1)	2075
36	Sucrose	Sugar	15.871	918	73 (60); 217 (31); 661 (100); 662 (33); $M^{+\cdot}$ 918 (<1)	2357
37	Lactose (isomer 1)	Sugar	16.427	918	73 (68); 204 (100); 205 (20); 217 (17); $M^{+\cdot}$ 918 (<1)	2469
38	Galactinol	Sugar	17.553	990	73 (44); 191 (24); 204 (100); 217 (30); $M^{+\cdot}$ 990 (<1)	2695
39	Melibiose	Sugar	16.718	918	73 (74); 204 (100); 205 (22); 217 (46); $M^{+\cdot}$ 918 (<1)	2527
40	β-Gentiobiose	Sugar	16.760	947	73 (69); 204 (100); 205 (21); 361 (39); $M^{+\cdot}$ 947 (<1)	2536
41	Lactose (isomer 2)	Sugar	16.770	918	73 (67); 204 (100); 205 (20); 217 (15); $M^{+\cdot}$ 918 (<1)	2538

Note: R_t, retention time; MW, molecular weight.
[a]Mass spectrometry library NIST 2011. $M^{+\cdot}$ molecular ion.
[b]Fatty acid methyl esters (FAMEs) C8–C24.

species was 6.7 mg/g in *C. baccatum*, 8.2 mg/g in *C. chinense*, 8.1 mg/g in *C. annuum* and 4.1 mg/g in *C. frutescens*. Amino acids are precursors of essential metabolites and are involved in response to stress. Studies evaluating the profile in Italian sweet pepper by HRMAS-NMR spectroscopy showed the presence of the amino acids alanine (Ala), arginine (Arg), asparagine (Asn), γ-amino butyrate acid (GABA), glutamate (Glu), glutamine (Gln), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), and valine (Val) (Ritota *et al.*, 2010). A similar amino acid profile was observed in this study (Table 1).

In an evaluation of the amino acid profile of *C. chinense*, Ananthan *et al.* (2016) reported the predominance of aspartic acid and glutamic acid (185 and 195 mg/100 g of protein), while in this study asparagine and norvaline predominated for the same species. Proline was the predominant amino acid in accession

Table 2. Non-derivatised metabolites identified in 60 *Capsicum* spp. accessions by GC–MS analysis

N°	Compound	Class	R_t	MW	Reference ions[a]	Retention index[b]
1	Pentadecane	Hydrocarbon	6.428	212	70 (16); 71 (100); 85 (65); 99 (14); $M^{+\cdot}$ 212 (34)	956
2	Decanoic acid	Fatty acid	6.933	172	71 (43); 73 (100); 87 (24); 129 (62); $M^{+\cdot}$ 172 (<1)	1009
3	Heneicosane	Hydrocarbon	8.731	296	71 (100); 85 (77); 99 (23); 113 (16); $M^{+\cdot}$ 296 (<1)	1216
4	Methyl tetradecanoate	Ester	10.150	242	71 (32); 74 (100); 85 (18); 87 (61); $M^{+\cdot}$ 242 (3)	1398
5	Tetradecanoic acid	Fatty acid	10.397	228	71 (35); 73 (100); 129 (52); 85 (27); $M^{+\cdot}$ 228 (10)	1433
6	Octadecane	Hydrocarbon	10.675	254	71 (100); 83 (14); 85 (71); 99 (22); $M^{+\cdot}$ 254 (3)	1473
7	2-hydroxy-Cyclopentadecanone	Hydrocarbon	10.785	240	81 (65); 83 (91); 84 (66); 97 (100); $M^{+\cdot}$ 240 (8)	1489
8	Pentadecanoic acid	Fatty acid	10.895	242	71 (53); 73 (100); 83 (35); 129 (47); $M^{+\cdot}$ 242 (11)	1504
9	Nonadecane	Hydrocarbon	11.370	268	71 (100); 85 (69); 99 (24); 113 (17); $M^{+\cdot}$ 268 (1)	1572
10	Methyl palmitoleate	Ester	11.429	268	74 (96); 83 (100); 84 (80); 87 (72); $M^{+\cdot}$ 268 (4)	1581
11	Methyl hexadecanoate	Ester	11.552	270	71 (20); 74 (100); 75 (19); 87 (70); $M^{+\cdot}$ 270 (5)	1598
12	Palmitoleic acid	Fatty acid	11.664	254	70 (56); 83 (100); 84 (70); 97 (74); $M^{+\cdot}$ 254 (4)	1615
13	*n*-Hexadecanoic acid	Fatty acid	11.792	256	71 (46); 73 (100); 85 (32); 129 (48); $M^{+\cdot}$ 256 (17)	1636
14	9,10-Anthracenedione	Anthraquinone	12.102	208	76 (47); 151 (33); 152 (66); 180 (85); $M^{+\cdot}$ 208 (100)	1685
15	Methyl linoleate	Ester	12.660	294	81 (100); 82 (58); 95 (62); 96 (38); $M^{+\cdot}$ 294 (6)	1772
16	Methyl oleate	Ester	12.688	296	74 (100); 83 (98); 84 (89); 97 (81); $M^{+\cdot}$ 296 (4)	1777
17	Oleic acid	Fatty acid	12.989	282	81 (84); 82 (67); 83 (100); 97 (73); $M^{+\cdot}$ 282 (2)	1826
18	Ethyl oleate	Ester	13.077	310	83 (90); 84 (87); 88 (100); 97 (91); $M^{+\cdot}$ 310 (4)	1841
19	Ethyl linolenate	Ester	13.127	306	79 (100); 81 (43); 93 (47); 95 (52); $M^{+\cdot}$ 306 (2)	1849
20	2-Hydroxyethyl hexadecanoate	Ester	13.509	300	84 (45); 98 (91); 104 (100); 117 (65); $M^{+\cdot}$ 300 (1)	1915
21	1-Heneicosanol	Alcohol	13.778	312	82 (43); 83 (100); 97 (90); 111 (47); $M^{+\cdot}$ 312 (<1)	1960
22	bis(2-ethylhexyl) hexanedioate	Ester	14.402	370	70 (34); 71 (29); 112 (25); 129 (100); $M^{+\cdot}$ 370 (<1)	2073
23	2-hydroxyethyl octadecanoate	Ester	14.640	328	86 (41); 98 (93); 104 (100); 117 (64); $M^{+\cdot}$ 328 (1)	2117
24	1,3-dihydroxypropan-2-yl hexadecanoate	Ester	15.016	330	74 (74); 84 (73); 98 (100); 239 (57); $M^{+\cdot}$ 330 (<1)	2187
25	Nonivamide	Capsaicinoid	15.137	293	137 (100); 138 (12); 151 (16); 195 (14); $M^{+\cdot}$ 293 (19)	2210
26	Capsaicin	Capsaicinoid	15.360	305	122 (7); 137 (100); 138 (12); 152 (11); $M^{+\cdot}$ 305 (7)	2255
27	Dihydrocapsaicin	Capsaicinoid	15.473	307	137 (100); 138 (14); 151 (16); 195 (15); $M^{+\cdot}$ 307 (17)	2278
28	Homocapsaicin	Capsaicinoid	15.930	319	71 (15); 137 (100); 138 (13); 152 (10); $M^{+\cdot}$ 319 (5)	2369
29	Homodihydrocapsaicin	Capsaicinoid	16.083	321	137 (100); 138 (14); 151 (19); 195 (15); $M^{+\cdot}$ 321 (14)	2400
30	γ-Tocopherol	Phytosterol	17.657	416	150 (23); 151 (100); 191 (23); 417 (25); $M^{+\cdot}$ 416 (81)	2716
31	DL-α-Tocopherol	Phytosterol	18.013	430	164 (39); 165 (100); 205 (11); 431 (25); $M^{+\cdot}$ 430 (76)	2787
32	Ergost-5-en-3-ol	Phytosterol	18.592	400	81 (94); 95 (92); 105 (88); 107 (97); $M^{+\cdot}$ 400 (100)	2903
33	Stigmasterol	Phytosterol	18.744	412	81 (81); 83 (100); 97 (53); 133 (54); $M^{+\cdot}$ 412 (50)	2934
34	γ-Sitosterol	Phytosterol	19.052	414	81 (97); 95 (90); 105 (82); 107 (100); $M^{+\cdot}$ 414 (97)	2996

Note: R_t, retention time; MW, molecular weight.
[a]Mass spectrometry library NIST 2011. $M^{+\cdot}$ molecular ion.
[b]Fatty acid methyl esters (FAMEs) C8–C24.

P290 (*C. annuum*). Proline accumulation has been observed in plants under stress conditions and attenuates the redox potential of free radicals and serves as an energy reserve and a nitrogen source during salt stress, drought, and temperature stresses (Verbruggen and Hermans, 2008). Another relevant amino acid found in the genus *Capsicum* is phenylalanine, which is a precursor of capsaicinoids. Studies have shown that phenylalanine ammonia lyase activity peaks and phenylalanine accumulation is channelled for capsaicinoid biosynthesis during fruit development (Castro-Concha *et al*., 2016).

Fructose was the most abundant sugar in all species tested, representing on average 49% of the total sugar content in *C. baccatum* (fructose content ranged from 55.8 to 85.4 mg/g), 55% in *C. chinense* (fructose content ranged from 40.0 to 93.3 mg/g), 53% in *C. annuum* (fructose content ranged from 28.6 to 84.9 mg/g), and 44% in *C. frutescens* (27.8 mg/g). Sucrose comprised on average 1 to 7% of the total sugar content, except for accession P116 (*C. chinense*) which contained 18% sucrose. *Capsicum baccatum* accessions showed the highest total sugar content (136.1 mg/g) followed by *C. chinense* (126.7 mg/g), *C. annuum* (100.4 mg/g), and *C. frutescens* (64.2 mg/g) (Fig. 2). Studies report that as *Capsicum* fruit matured, sucrose levels decreased to undetectable levels (Navarro *et al*., 2006). The current study determined the metabolic profiles of 60 genotypes, and examined only fully ripe fruit. Therefore, a low concentration of sucrose was expected. Sucrose represented between 1 and 7% of total sugar content with the exception of *C. chinense* P116, with 18%. A previous study showed that sucrose content in *C. chinense* fruit averaged 0.18 mg/g fresh weight (FW) and ranged from below detection limits to 1.50 mg/g FW, and that glucose or fructose concentrations were always greater than sucrose, as was observed in the present study (Jarret *et al*., 2009). Derivatisation reactions

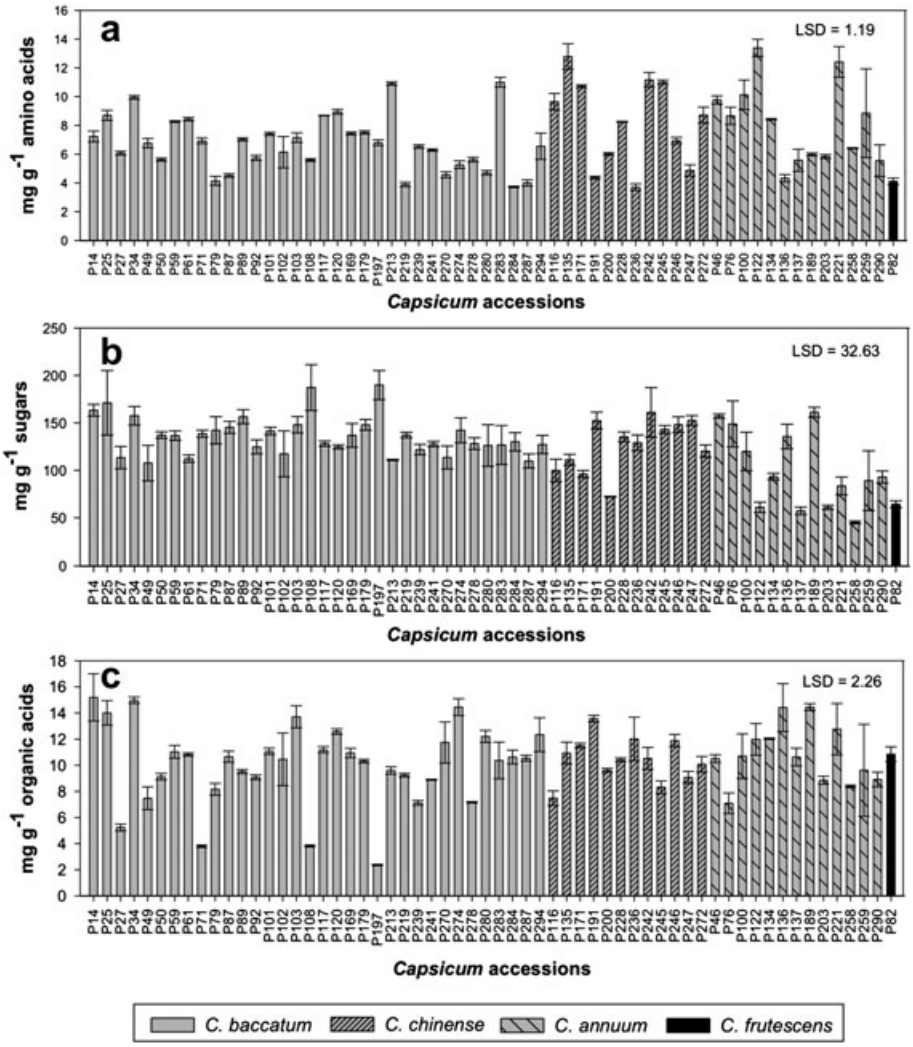

Figure 2. Graphic representation of metabolite content (in mg/g dry weight) of 60 pepper (*Capsicum* spp.) accession Embrapa Clima Temperado active germplasm collection, Pelotas, RS, Brazil. Results are mean ± standard error. (a) Total amino acid, (b) total sugars, and (c) total organic acids. LSD, least significant difference.

such as silylation provide thermal stability to compounds by the addition of a trimethylsilyl radical to hydroxyl, amine, and thiol groups. Derivatisation by oximation is one way to avoid the formation of isomers of sugars and consequent formation of more than one peak for each compound. Typically, up to four sugar isomers are formed in aqueous solution (Halket and Zaikin, 2003). Although an oximation derivatisation step was used in this study, sugar isomers were formed.

The mean total organic acid content for the species studied here was 10.8 mg/g in both *C. frutescens* and *C. annuum*, 10.5 mg/g in *C. chinense*, and 10.0 mg/g in *C. baccatum*. Analysis of peppers from the same Embrapa Clima Temperado AGC showed total acidity to be on average 20.7 mg/g dry weight (DW) in *C. annuum*, 16.4 mg/g DW in *C. baccatum*, 13.6 mg/g DW in *C. chinense*, and 30 mg/g DW in *C. frutescens* determined by titration and expressed as citric acid equivalents (Acunha *et al*., 2017).

Citric acid was the predominant organic acid in *C. chinense* and *C. annuum*. In *C. chinense* its concentration varied from 3.65 to 8.31 mg/g and represented on average 56% of the total organic acid content, while in *C. annuum* citric acid content varied from 4.38 to 11.32 mg/g representing on average 66% of total organic acids. P82, the only *C. frutescens* accession tested, had 8.3 mg/g of succinic acid (butanedioic acid), which represented 77% of its total organic acid content. *Capsicum baccatum* accessions P27 (2.41 mg/g), P71 (2.85 mg/g), P108 (1.93 mg/g), and P197 (1.59 mg/g) also had succinic acid as the predominant organic acid. The remaining *C. baccatum* accessions had citric acid as the primary organic acid, with a maximum content of 10.18 mg/g (P280) [Fig. 2(c)]. Jarret *et al*. (2009) reported that among organic acids, *C. chinense* had higher concentrations of citric acid (2.44 mg/g FW), followed by malic (0.69 mg/g FW), fumaric (0.49 mg/g FW), and succinic acids (0.24 mg/g FW) at lower concentrations. In our study, citric acid was highlighted in accessions of *C. chinense* and *C. annuum*; however, the *C. frutescens* accession and some accessions of *C. baccatum* presented succinic acid as a major organic acid.

Photosynthesis is the most vital activity of a plant and leads to the formation of sugars and carbon backbones. In plants, these general metabolites act as structural components, intermediates, or storage compounds. Specialised metabolites, however, play a

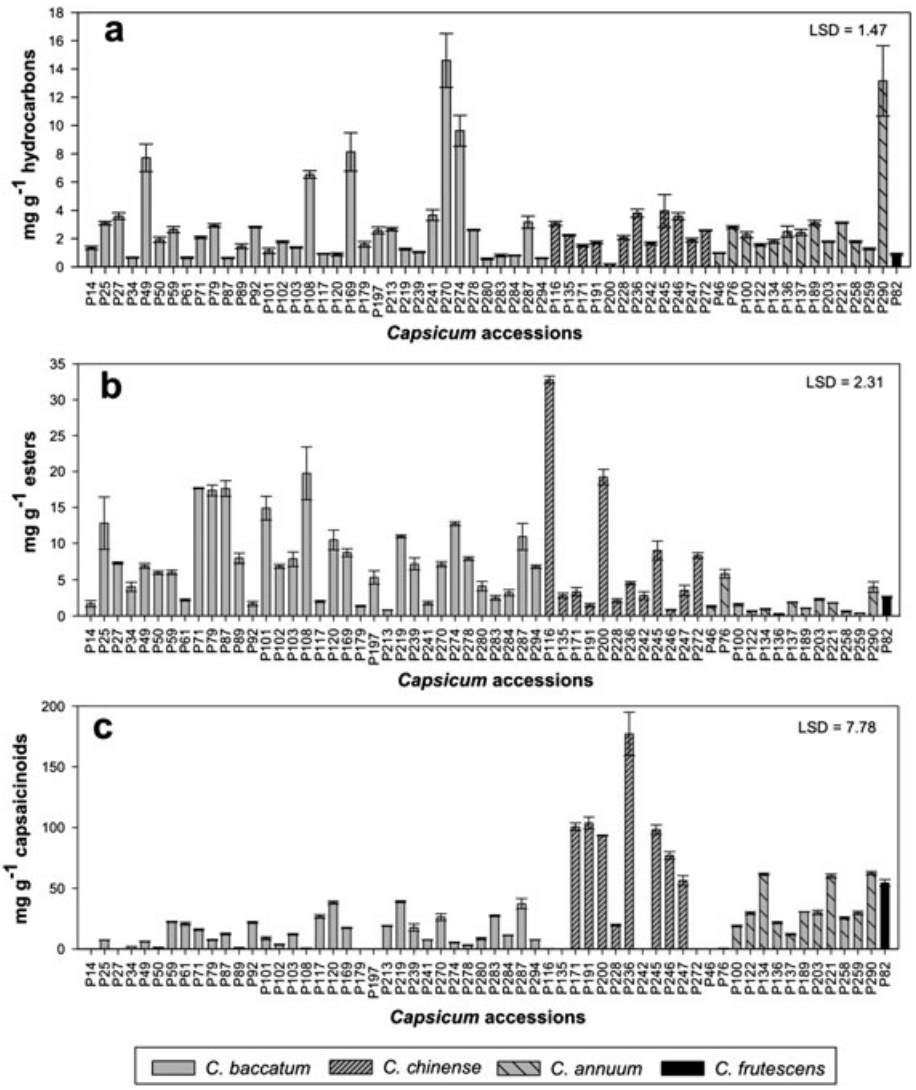

Figure 3. Graphic representation of metabolite content (in mg/g dry weight) of 60 pepper (*Capsicum* spp.) accession Embrapa Clima Temperado active germplasm collection, Pelotas, RS, Brazil. Results are mean ± standard error. (a) Hydrocarbons, (b) esters, and (c) capsaicinoids. LSD, least significant difference.

role in ecological interactions of specific groups of organisms contributing to their fitness under adverse environmental conditions. In the accessions tested, hydrocarbon content varied from 0.2 to 14.6 mg/g. *Capsicum baccatum* accessions P49, P108, P169, P270, and P274 and *C. annuum* accession P290 had more than 6 mg/g [Fig. 3(a)]. Total ester content in fraction B of the 60 accessions tested ranged from 0.4 to 32.8 mg/g, with a mean of 6.3 mg/g. Generally, the most abundant compound was 1,3-dihydroxypropan-2-yl hexadecanoate, which represented on average 30% of the total ester content. Its content varied from 0.32 mg/g (P134) to 6.57 mg/g (P116) (Fig. 3).

Capsaicin was the predominant capsaicinoid in most pepper accessions tested, and represented on average 59% of the total capsaicinoid content. Together capsaicin and dihydrocapsaicin represented 89% of the total capsaicinoid content of the studied accessions. Capsaicin content (when present above quantitation limits) varied from 0.17 mg/g in accession P116 (*C. chinense*) to 140.40 mg/g in accession P236 (*C. chinense*). Dihydrocapsaicin content (when present above quantitation limits) varied from 0.06 mg/g in accession P116 to 31.80 mg/g in accession P236 [Fig. 3(c)]. Among the four species tested, the greatest mean total capsaicinoid content was observed in *C. chinense* (66.1 mg/g), followed by *C. annuum* (31.9 mg/g), and *C. baccatum* (13.6 mg/g). The only *C. frutescens* accession surveyed had 54.4 mg/g. *Capsicum baccatum* accessions P14, P108, and P97, *C. chinense* accessions P116, P242, and P272, and *C. annuum* accession P76 had less than 1 mg/g of total capsaicinoid content. For *C. baccatum* accessions P27 and P179, *C. chinense* accession P135, and *C. annuum* accession P46, capsaicinoids, if present, were below the detection limit.

Capsaicinoids are the compounds responsible for the pungency of *Capsicum* fruit and capsaicin is usually the predominant capsaicinoid. These alkaloids are synthesised from the gene encoding Pun1 enzyme controlling the biosynthetic pathway of capsaicin. This acyltransferase is responsible for transferring an acyl radical to vanillin forming a vanillin amide, which is then joined to a fatty acid (Mazourek *et al*., 2009). Mutations and loss of function of *Pun1* result in the production of capsaicinoids at levels below perception or the complete absence of pungent capsaicinoids, which is the case in most sweet peppers (Wahyuni *et al*., 2013;

Mazourek *et al.*, 2009). Extraction of capsaicinoids by Soxhlet followed by GC–MS analysis found 0.15 to 5.93 mg capsaicin/g and 0.10 to 1.33 mg dihydrocapsaicin/g in cultivars Cascabel (*C. annuum*) and Chilpaya (*Capsicum* sp.), respectively (Peña-Alvarez *et al.*, 2009). Capsaicin has been shown to represent 50% of the total capsaicinoid content in cultivars Jalapeño and Serrano (*C. annuum*), and 83% in Golden and Chocolate Habanero peppers (*C. chinense*) (Giuffrida *et al.*, 2013). The same cultivars possessed approximately 39% and 13% of dihydrocapsaicin, respectively. In this study, two *C. baccatum* genotypes had higher levels of dihydrocapsaicin than capsaicin. Accessions P270 and P284 had 2.6 and 1.13 mg/g DW more dihydrocapsaicin than capsaicin, respectively. Similar results were reported by Peña-Alvarez *et al.* (2009), who observed a higher dihydrocapsaicin content (0.82 mg/g) than capsaicin content (0.71 mg/g) in the variety Canica (*Capsicum* sp.). Zewdie and Bosland (2001) found variation in the predominance of individual capsaicinoids in different *Capsicum* genotypes. In *C. pubescens*, for example, dihydrocapsaicin was the predominant capsaicinoid (Zewdie and Bosland, 2001). More than 10 capsaicinoids have been identified in the *Capsicum* spp. (Mazourek *et al.*, 2009). In this, study eight capsaicinoids were identified (Fig. 3, Table 2).

Scoville (1912) developed the first methodology for quantification of pungency levels in *Capsicum* using sensory analysis. By this method, pungency is expressed in Scoville Heat Units (SHUs) and represents the number of times an extract is diluted until pungency is no longer perceived. The pungency levels expressed on the Scoville scale of the 60 *Capsicum* pepper accessions tested are presented in Table 3. SHUs were estimated based on capsaicin and dihydrocapsaicin contents according to Todd *et al.* (1977). Cultivar Naga King (*C. chinense*) has been reported to have highly pungent fruit, with more than one million SHUs (Ananthan *et al.*, 2016). Studies evaluating the pungency of different cultivars of *C. chinense* observed that the variety Trinidad Moruga Scorpion presented fruit pungency levels higher than two million SHUs (Bosland *et al.*, 2012). These results are in accordance with the findings of the present study that showed accessions of *C. chinense* that stood out for their high levels of pungency. For example, P236 (*C. chinense*) had 2772533.5 SHUs (Table 3).

This untargeted metabolomics analysis performed by GC–MS in polar fractions A and B demonstrated the variability among genotypes of *Capsicum* species. These results further expanded the Embrapa Clima Temperado AGC metabolite database not previously surveyed for amino acid, sugars, organic acids, hydrocarbons, and esters, and provided further evidence of the limited application of metabolic profiling for species differentiation in *Capsicum*.

Table 3. Pungency of 60 *Capsicum* accessions of the active germplasm collection of Embrapa Clima Temperado, Pelotas, Brazil

Accession	Species	SHU[a]	Accession	Species	SHU[a]
P14	*C. baccatum*	3656.4	P283	*C. baccatum*	380419.7
P25	*C. baccatum*	107683.8	P284	*C. baccatum*	140879.1
P27	*C. baccatum*	ND	P287	*C. baccatum*	445121.4
P34	*C. baccatum*	30468.8	P294	*C. baccatum*	94760.1
P49	*C. baccatum*	99651.1	P116	*C. chinense*	4454.5
P50	*C. baccatum*	22247.4	P135	*C. chinense*	ND
P59	*C. baccatum*	312916.9	P171	*C. chinense*	1528963.3
P61	*C. baccatum*	281398.2	P191	*C. chinense*	1569433.0
P71	*C. baccatum*	233642.6	P200	*C. chinense*	1427024.7
P79	*C. baccatum*	112034.1	P228	*C. chinense*	257989.4
P87	*C. baccatum*	184829.0	P236	*C. chinense*	2772533.5
P89	*C. baccatum*	18178.9	P242	*C. chinense*	1497.8
P92	*C. baccatum*	278783.7	P245	*C. chinense*	1512695.7
P101	*C. baccatum*	121248.4	P246	*C. chinense*	1145864.2
P102	*C. baccatum*	53906.0	P247	*C. chinense*	841602.5
P103	*C. baccatum*	186764.1	P272	*C. chinense*	4062.2
P108	*C. baccatum*	11170.5	P46	*C. annuum*	ND
P117	*C. baccatum*	369022.3	P76	*C. annuum*	8818.3
P120	*C. baccatum*	530937.1	P100	*C. annuum*	265563.0
P169	*C. baccatum*	249347.2	P122	*C. annuum*	411088.1
P179	*C. baccatum*	ND	P134	*C. annuum*	675927.7
P197	*C. baccatum*	2409.2	P136	*C. annuum*	299327.4
P213	*C. baccatum*	275067.2	P137	*C. annuum*	163006.8
P219	*C. baccatum*	504567.7	P189	*C. annuum*	458080.9
P239	*C. baccatum*	245245.2	P203	*C. annuum*	382423.1
P241	*C. baccatum*	95094.5	P221	*C. annuum*	591443.6
P270	*C. baccatum*	367186.5	P258	*C. annuum*	366316.4
P274	*C. baccatum*	70916.1	P259	*C. annuum*	412225.3
P278	*C. baccatum*	45985.2	P290	*C. annuum*	870679.6
P280	*C. baccatum*	112505.6	P82	*C. frutescens*	755282.7

[a]Scoville Heat Unit (SHU) (Scoville, 1912); ND, not detected.

Acknowledgements

The authors would like to acknowledge Secretaria de Desenvolvimento Econômico, Ciência e Tecnologia do Governo do Estado do RS (Programa de Apoio aos Polos Tecnológicos), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) for financial support for research.

References

Acunha TS, Crizel RL, Tavares IB, Barbieri RL, Pereira CMP, Rombaldi CV, Chaves FC. 2017. Bioactive compound variability in a Brazilian Capsicum pepper collection. *Crop Sci* **65**: 523–532.

Aizat WM, Dias DA, Stangoulis JCR, Able JA, Roessner U, Able AJ. 2014. Metabolomics of capsicum ripening reveals modification of the ethylene related-pathway and carbon metabolism. *Postharvest Biol Technol* **89**: 19–31.

Ananthan R, Subhash K, Longvah T. 2016. Capsaicinoids, amino acid and fatty acid profiles in different fruit components of the world hottest Naga king chilli (*Capsicum chinense* Jacq). *Food Chem*. https://doi.org/10.1016/j.foodchem.2016.12.073.

Bosland PW, Coon D, Reeves G. 2012. 'Trinidad Moruga Scorpion' pepper is the world's hottest measured chile pepper at more than two million Scoville heat units. *Hort Technology* **22**: 534–538.

Castro-Concha LA, Baas-Espinola FM, Ancona-Escalante WR, Vázquez-Flota FA, Miranda-Ham ML. 2016. Phenylalanine biosynthesis and its relationship to accumulation of capsaicinoids during *Capsicum chinense* fruit development. *Biol Plantarum* **60**: 579–584.

De Vos RC, Moco S, Lommen A, Keurentjes JJ, Bino RJ, Hall RD. 2007. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat Protoc* **2**: 778–791.

Giuffrida D, Dugo P, Torre G, Bignardi C, Cavazza A, Corradini C, Dugo G. 2013. Characterization of 12 *Capsicum* varieties by evaluation of their carotenoid profile and pungency determination. *Food Chem* **140**: 794–802.

Halket JM, Zaikin VG. 2003. Derivatization in mass spectrometry – 1. Silylation. *Eur J Mass Spectrom* **9**: 1–21.

Heiser CB, Jr, Pickersgill B. 1969. Names for the cultivated *Capsicum* species (Solanaceae). *Taxon* **18**: 277–283.

Hoffmann JF, Carvalho IR, Barbieri RL, Rombaldi CV, Chaves FC. 2017. *Butia* spp. (Arecaceae) LC-MS-based metabolomics for species and geographical origin discrimination. *J Agric Food Chem* **65**(2): 523–532. https://doi.org/10.1021/acs.jafc.6b03203.

Howard LR, Wildman REC. 2006. Chapter 13. In *Handbook of Nutraceuticals and Functional Foods*, second edn. CRC Press: New York; 560.

Jarret RL, Berke T, Baldwin EA, Antonious GF. 2009. Variability for free sugars and organic acids in *Capsicum chinense*. *Chem Biodivers* **6**: 138–145.

Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. 2006. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat Protoc* **1**: 387–396.

Mazourek M, Pujar A, Borovsky Y, Paran I, Mueller L, Jahn MM. 2009. A dynamic interface for capsaicinoid systems biology. *Plant Physiol* **150**: 1806–1821.

Navarro JM, Flores P, Garrido C, Martinez V. 2006. Changes in the contents of antioxidant compounds in pepper fruits at different ripening stages, as affected by salinity. *Food Chem* **96**: 66–73.

Peña-Alvarez A, Ramirez-Maya E, Alvarado-Suarez LA. 2009. Analysis of capsaicin and dihydrocapsaicin in peppers and pepper sauces by solid phase microextraction – gas chromatography – mass spectrometry. *J Chromatogr A* **1216**: 2843–2847.

Ritota M, Marini F, Sequi P, Valentini M. 2010. Metabolomic characterization of Italian sweet pepper (*Capsicum annum* L.) by means of HRMAS-NMR spectroscopy and multivariate analysis. *J Agr Food Chem* **58**: 9675–9684.

Scoville WL. 1912. Note on Capsicums. *J Pharm Sci* **1**: 453–454.

Smith PG, Heiser CB, Jr. 1951. Taxonomic and genetic studies on the cultivated peppers, *Capsicum annuum* L. and *C. frutescens* L. *Am J Bot* **38**: 362–368.

Tewksbury JJ, Manchego C, Haak DC, Levey DJ. 2006. Where did the chili get its spice? Biogeography of capsaicinoid production in ancestral wild chili species. *J Chem Ecol* **32**: 547–564.

Todd PH, Bensinger MG, Biftu T. 1977. Determination of pungency due to *Capsicum* by gas-liquid chromatography. *J Food Sci* **42**: 660–665.

Verbruggen N, Hermans C. 2008. Proline accumulation in plants: A review. *Amino Acids* **35**: 753–759.

Wahyuni Y, Ballester AR, Sudarmonowati E, Bino RJ, Bovy AG. 2011. Metabolite biodiversity in pepper (*Capsicum*) fruits of thirty-two diverse accessions: Variation in health-related compounds and implications for breeding. *Phytochemistry* **72**: 1358–1370.

Wahyuni Y, Ballester AR, Tikunov Y, De Vos RC, Pelgrom KT, Maharijaya A, Sudarmonowati E, Bino RJ, Bovy AG. 2013. Metabolomics and molecular marker analysis to explore pepper (*Capsicum* sp.) biodiversity. *Metabolomics* **9**: 130–144.

Zewdie Y, Bosland PW. 2001. Capsaicinoid profiles are not good chemotaxonomic indicators for Capsicum species. *Biochem Syst Ecol* **29**: 161–169.