



# Assessment of cocoa powder changes during the alkalization process using untargeted metabolomics

Maidier Greño<sup>a,b</sup>, Miguel Herrero<sup>b</sup>, Alejandro Cifuentes<sup>b</sup>, María Luisa Marina<sup>a,c</sup>, María Castro-Puyana<sup>a,c,\*</sup>

<sup>a</sup> Universidad de Alcalá. Departamento de Química Analítica, Química Física e Ingeniería Química, Ctra. Madrid-Barcelona Km. 33.600, 28871, Alcalá de Henares (Madrid), Spain

<sup>b</sup> Laboratory of Foodomics, Institute of Food Science Research (CIAL, CSIC-UAM), Nicolás Cabrera 9, 28049, Madrid, Spain

<sup>c</sup> Universidad de Alcalá. Instituto de Investigación Química Andrés M. del Río, Ctra. Madrid-Barcelona Km. 33.600, 28871, Alcalá de Henares (Madrid), Spain

## ARTICLE INFO

### Keywords:

Non-targeted metabolomics  
Liquid chromatography  
High-resolution mass spectrometry  
Cocoa powder  
Alkalization process

## ABSTRACT

This work is focused on the study of the metabolomic changes that occur in cocoa powder throughout the alkalization process by using an untargeted ultra high-performance liquid chromatography coupled to high-resolution mass spectrometry (UHPLC-QTOF-MS) metabolomics approach. With this aim, cocoa powder samples submitted to different alkalization degrees (light, medium, and strong) were analyzed. Metabolite extraction was performed using 75% MeOH in water since it provided the highest number of molecular features. After data processing, non-supervised and supervised methods were applied to carry out the statistical data analysis. Thus, the most significant metabolites that allow establishing differences among the cocoa powder samples submitted to the alkalization processes were pointed out. Thus, 43 and 30 metabolites in positive and negative ionization modes, respectively, demonstrated to be relevant. Among them, 9 compounds were unequivocally identified and 22 tentatively identified. Most of them were amino acids, alkaloids, organic acids or polyphenols, among others.

## 1. Introduction

Cocoa is a highly valued product, consumed all over the world not only for its pleasant flavor and aroma but also for the nutritional and health benefits that it presents. These properties make cocoa a suitable ingredient in the food industry and personal care products. Despite the sweet flavor of chocolate (the most known cocoa-based product) cocoa is characterized by presenting acidity, astringency, and bitterness properties. These non-desirable properties could be reduced using an alkalization (or dutching) process (Sioriki et al., 2021; Valverde et al., 2020) which is based on the combination of an alkali solution (usually NaOH and K<sub>2</sub>CO<sub>3</sub> in concentration from 1 to 6%) with cocoa. A particular temperature, (from 60 to 130 °C), and pressure (from 0.10 to 1.22 MPa) are applied during a period between 5 and 180 min (Valverde-García et al., 2020). In general, the final product is classified according to its pH value, which may range from pH 5.0 to pH 8.0. In this way, the product is labeled as dark natural if the pH is between 5.0 and 6.0, light alkalized cocoa when the pH is from 6.0 to 7.2, medium alkalized cocoa when the pH range is between 7.2 and 7.6, and strongly

alkalized cocoa when the product has a pH value higher than 7.6 (Miller et al., 2008). The use of the different alkalized cocoa powders depends on the final application of the product: natural and light alkalized cocoa powders are employed to produce chocolate, ice creams, and cocoa beverages whereas medium and strong alkalized cocoa powders are mainly used to prepare ice creams and bakery products such as cookies or cakes (Valverde-García et al., 2020).

Depending on the alkali solution, temperature, pressure, and time employed, cocoa solubility, color, and flavor might change. Color changes may be due to Maillard reactions, polyphenol polymerization and oxidation, or the increased activity of polyphenol oxidase enzyme at higher pHs (Valverde-García et al., 2020).

Cocoa is composed of a great variety of compounds such as carbohydrates, proteins, fats, polyphenols, methylxanthines, minerals, vitamins, and amino acids (Martín & Ramos, 2017). Certainly, the alkalization process implies several chemical transformations which can modify the cocoa composition but also its nutritional, sensory, and microbiological characteristics. Several works have investigated the change in the composition of alkalized cocoa (Valverde-García et al.,

\* Corresponding author. Universidad de Alcalá. Departamento de Química Analítica, Química Física e Ingeniería Química, Ctra. Madrid-Barcelona Km. 33.600, 28871, Alcalá de Henares (Madrid), Spain.

E-mail address: [maria.castrop@uah.es](mailto:maria.castrop@uah.es) (M. Castro-Puyana).

<https://doi.org/10.1016/j.lwt.2022.114207>

Received 19 August 2022; Received in revised form 13 October 2022; Accepted 21 November 2022

Available online 22 November 2022

0023-6438/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2020). In general, these works are focused on the determination of: (i) the total content of polyphenols (Andres-Lacueva et al., 2008; Gu et al., 2006; Gültekin-Özgüven et al., 2016; Hurst et al., 2011; Jolić et al., 2011; Kofink et al., 2007; Li et al., 2014; Miller et al., 2008; Payne et al., 2010; Quelal-Vásquez et al., 2020; Rodríguez et al., 2009; Stanley et al., 2015; Stark & Hofmann, 2006; Todorovic et al., 2017), (ii) the mineral content (Adeyeye et al., 2016) (iii) the volatile composition (Alasti et al., 2019, 2020; Li et al., 2012) or (iv) the effect of alkalization on Maillard products (Taş & Gökmen, 2016). However, these studies are based on the determination of a specific class of compounds and just a few articles have reported the characterization of a wide variety of compounds including volatile and nonvolatile compounds (Li et al., 2012; Sioriki et al., 2022) or the identification of new compounds as a result of alkalization treatment (Germann et al., 2019a, 2019b). For this reason, a deeper study focused on a more comprehensive analysis of the set of metabolites could provide relevant information to evaluate changes in the metabolic profile of cocoa powder samples submitted to different alkalization degrees. In this regard, “omics” techniques have gained importance in the last years due to the extensive information that they provide about a biological system. From the perspective of Food Science, metabolomics aims to provide information regarding the quality, safety, and traceability of foods and food products. LC-MS is the most common analytical platform for untargeted metabolomic studies (Castro-Puyana et al., 2017; Zhong et al., 2022). To the best of our knowledge, there are no metabolomics-based strategies reported in the literature focused on the study of the effect of the alkalization process on cocoa powder.

Bearing in mind that the alkalization process to what the cocoa samples are submitted may lead to chemical transformations modifying the cocoa composition, and therefore its nutritional, sensory, and microbiological characteristics, this work was focused on implementing a UHPLC-QTOF-MS untargeted metabolomics approach to study the metabolic profile of cocoa powder samples submitted to different alkalization degrees. In this way, it would be possible to perform a comprehensive metabolomics analysis to identify those metabolites showing significant differences along the alkalization degrees, which can be considered as potential markers of the alkalization process.

## 2. Materials and methods

### 2.1. Reagents and chemicals

Solvents employed were HPLC-grade. Heptane was from Scharlau (Barcelona, Spain). MS grade methanol and acetonitrile were from VWR Chemicals (Barcelona, Spain), whereas formic acid and ethanol were purchased from Thermo Fisher Scientific (Madrid, Spain). Ultrapure water was obtained from a Millipore system (Billerica, MA, USA). Citric acid, adenosine, phenylalanine, theobromine, protocatechuic acid, (+)-catechin hydrate, and caffeine were from Sigma (Madrid, Spain). Pyroglutamic acid and leucine were obtained from Fluka (Buchs, Switzerland).

### 2.2. Samples

A natural cocoa powder sample (CN) and three cocoa powder samples alkalized at different degrees: light (CL), medium (CM), and strong (CS) were kindly provided by Olam International (Valencia, Spain). The alkalized cocoa samples were compared to CN. The pH values for all these samples were as follows:  $5.5 \pm 0.5$  for CN;  $6.5 \leq \text{pH} < 7.3$  for CL;  $6.9 \leq \text{pH} \leq 7.6$  for CM; and  $7.8 \leq \text{pH} \leq 8.7$  for CS.

### 2.3. Sample preparation

Cocoa samples were defatted with heptane (1 g with 10 mL, for three times) and subsequently, they were dried under vacuum (SpeedVac SPD1030, Thermo Fisher Scientific). Then, 100 mg of each cocoa sample were mixed with 1.5 mL of the extraction solvent (75% of

methanol in water). The extraction procedure was performed in a Thermomixer (Thermomixer C Eppendorf AG, Hamburg, Germany) at 750 rpm for 30 min at 25 °C. The extracted samples were centrifuged at 3500 rpm for 10 min at 25 °C (Heraeus Fresco 21 Centrifuge, Thermo Fischer Scientific). Finally, the supernatant was directly injected into the UHLC-MS system.

For the metabolomic sequence, each alkalized cocoa sample was extracted five times, and equal aliquots of each alkalized cocoa extract (CN, CL, CM, and CS) were mixed to prepare a quality control sample (QC). Moreover, a solution containing ten standards was injected during the metabolomic sequence to ensure good analytical performance.

### 2.4. Ultra high-performance liquid chromatography-mass spectrometry conditions

An Agilent 1290 UHPLC system coupled to a 6540 series quadrupole time-of-flight (QTOF) mass spectrometer (Agilent Technologies, Germany) equipped with a Jet Stream thermal orthogonal electrospray ionization (ESI) source was used to carry out all the analysis. Agilent Mass Hunter Qualitative Analysis software (B.10.00) was employed for MS control, data acquisition, and data analysis.

A Zorbax Eclipse Plus C<sub>18</sub> column (100 × 2.1 mm, 1.8 μm particle diameter) was used protected by a Zorbax Eclipse Plus C<sub>18</sub> guard column (5 × 2.1 mm, 1.8 μm particle diameter) both from Agilent Technologies. The chromatographic methodology and mass detection parameters were similar to previous works focused on the use of UHPL-MS for the characterization of natural products (Avanza et al., 2021; Peixoto et al., 2021). Elution was performed employing the following mobile phases: water + 0.1% formic acid (solvent A) and acetonitrile + 0.1% formic acid (solvent B) in positive ionization, and water + formic acid 0.01% (A) and acetonitrile + formic acid 0.01% (B) in negative ionization mode. In both cases, the gradient was as follows: 0–30% B in 0–7 min; 30–80% in 7–9 min, 80–100% in 9–11 min, 100% B in 11–13 min and 0% B in 13–14 min. The flow rate was 0.5 mL/min, the column temperature was 40 °C and the sample injection volume was 2 μL.

The mass spectrometer was used in positive and negative ESI modes with the following parameters: capillary voltage, 3000 V; nozzle voltage, 0 V; nebulizer pressure, 40 psi; sheath gas flow 11 L/min and temperature 350 °C; drying gas flow rate, 8 L/min and gas temperature, 300 °C, skimmer voltage, 45 V; fragmentor voltage, 110 V; octapole voltage, 750 V. MS and auto MS mode was set to acquire *m/z* values ranging from 25 to 1100 at a scan rate of 5 spectra per second. MS/MS analyses were performed by fragmenting the selected precursor ions with voltages of 20 and 40 V in the collision chamber. A solution of two reference ions was used to allow proper mass correction, namely, *m/z* 121.0508 (C<sub>5</sub>H<sub>4</sub>N<sub>4</sub>) and 922.0097 (C<sub>18</sub>H<sub>18</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>24</sub>) in positive mode, and *m/z* 119.0363 (C<sub>5</sub>H<sub>4</sub>N<sub>4</sub>) and 966.0007 (C<sub>18</sub>H<sub>18</sub>O<sub>6</sub>N<sub>3</sub>P<sub>3</sub>F<sub>24</sub> + formate) in negative mode.

### 2.5. Metabolomics sequence

Blanks and the QC sample were injected at the beginning and end of the analytical sequence to assure the good stability of the system. Then, four groups of samples and five replicates of each group were randomly injected in triplicate. Every six samples a QC sample was also injected to assess the stability of the analyses. Moreover, through the metabolomics sequence, the mix solution containing ten standards was injected to evaluate mass accuracy and retention time-shifting.

### 2.6. Data processing and multivariate analysis

Data treatment was carried out using Mass Hunter Qualitative (MHQ) (B.10.00) and Profiler Professional (MPP) (B.07.00) software from Agilent. The molecular Feature Extraction (MFE) tool was used to obtain information related to molecular features. The algorithm used was “small molecules (chromatographic)” and the following parameters

were selected for the molecular features extraction: ions  $\geq 500$  counts; isotope model = common organic molecular; and limited assigned change was established at 2. For positive ionization mode  $H^+$ ,  $Na^+$ , and  $NH_4^+$  and for negative mode  $H^-$ ,  $Cl^-$ ,  $HCOO^-$ , and  $CH_3COO^-$  adducts were considered to identify different ion species coming from the same molecular feature.

After peak alignment (retention time window of 0.15 min, a mass tolerance of 0.02 Da, and a mass window of 15 ppm) of the extracted molecular features using MPP, molecular feature filtering was performed, applying a minimum absolute abundance of 10,000 counts, an ion number of 2, and all charges permitted. Moreover, molecular features present in 100% of the QC samples with a coefficient of variation below 30% were selected for statistical analysis. Multivariate statistical analysis was performed with SIMCA 14.0 software (Umetrics, Umeå, Sweden). Data was centered and divided by the square root of the standard deviation as a scaling factor (Pareto scaling). Both, unsupervised principal component analysis (PCA), to evaluate clustering in the analyzed samples, and partial least squares discriminant analysis (PLS-DA), to discriminate samples according to their alkalization degree, were carried out. Models were evaluated by the good-ness-of-fit parameters  $R^2X$ ,  $R^2Y$ , and  $Q^2$ .

## 2.7. Metabolite identification

Identification was based on the molecular features which presented significant differences in the PLS-DA model. This process was performed using the CEU Mass Mediator (Gil de la Fuente et al., 2018), and FooDB database (<http://foodb.ca>) considering an error of 30 ppm. Metabolites showing significant differences among groups and those commercially available were purchased and analyzed under the same experimental conditions. The comparison of retention times and MS/MS fragmentation was carried out to confirm the metabolite identification. When standards were not commercially available, experimental MS/MS spectra were compared to those described in the HMDB database, literature, and predicted spectra in CFM-ID (<https://cfmid.wishartlab.com/>).

## 3. Results and discussion

### 3.1. Optimization of the extraction procedure

One of the most crucial steps in an untargeted metabolomic analysis is sample preparation since it will directly influence the results obtained. Ideally, the sample preparation should be as non-selective as possible to ensure a broad coverage of the metabolome.

The approach followed for optimizing the extraction procedure was based on the achievement of the greatest number of molecular features from cocoa samples. Previous results obtained by our research group related to the use of an HPLC-MS-based metabolomics approach to reveal cocoa powder adulterations demonstrated that a defatting step (to remove fatty acid content), the use of a thermomixer (as it enabled the control of temperature), and 30 min of extraction time were the most appropriate conditions to carry out the extraction (Greño et al., 2022). Taking into account these conditions, the extraction solvent to perform the metabolite extraction from the samples investigated in this study was optimized.

Thus, different solvents and solvents mixtures, including methanol (25, 50, 75, 80, 90% in water), ethanol (25, 50, 75% in water), and 100% water at pH 2.0 (water with small amounts of phosphoric acid) and 9.0 (water with sodium borate) were evaluated. The largest number of molecular features was obtained using 50% MeOH for positive ionization mode, and 80% MeOH for negative ionization mode (see Table 1). However, as reported in Table 1, the use of 50% MeOH in negative ionization mode revealed 559 less features than with 80% MeOH. However, the use of 80% MeOH as solvent in positive ionization mode gave rise to a high variation in terms of molecular feature extraction

**Table 1**

Average number of molecular features and RSD values of three replicate injections of a QC sample using different extraction solvents.

	Number of molecular features			
	ESI +		ESI-	
	MFE	RSD(%)	MFE	RSD(%)
25% EtOH in water	808	10.2	1833	1.7
50% EtOH in water	787	6.3	2018	1.9
75% EtOH in water	790	4.3	2162	2.1
25% MeOH in water	903	12.0	1880	3.9
50% MeOH in water	1125	1.8	2020	2.8
75% MeOH in water	1124	3.8	2424	4.4
80% MeOH in water	1041	12.0	2579	1.0
90% MeOH in water	833	8.5	2359	1.6
Water pH 2.0	822	5.1	1904	2.6
Water pH 9.0	823	4.3	1826	3.3

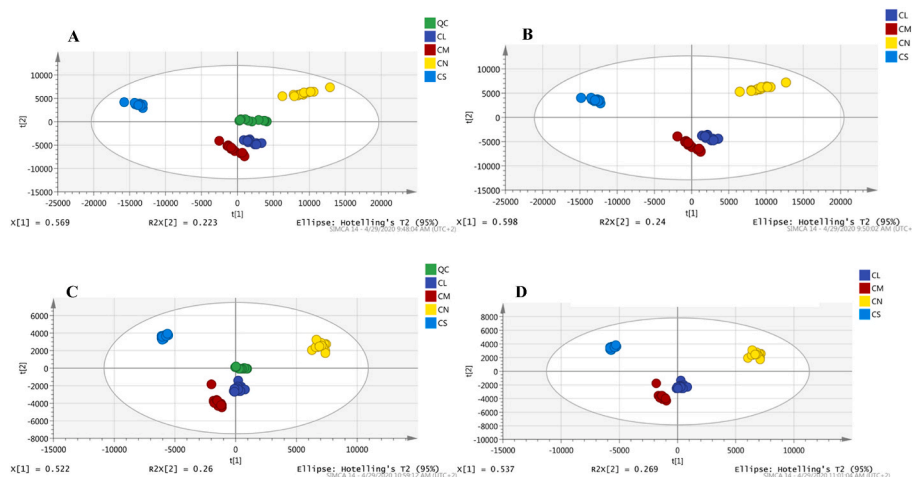
(RSD 12%). Bearing all this in mind, 75% MeOH was selected as the optimum extraction solvent as it enabled a high number of molecular features for both ionization modes with a RSD lower than 4.5. Figs. S1 and S2 show representative total ion chromatograms (TIC) obtained for the samples analyzed (CN, CL, CM, CS) in positive and negative ESI modes.

### 3.2. Data treatment by multivariate statistical analysis

After sample extraction and the UHPLC-MS analysis of all cocoa samples (a total of 60 samples) and QC samples, the data treatment was carried out (see section 2.6). In positive mode, 4026 molecular features were obtained. After filtering by frequency (features present in 100% of QC samples) and variability (features with RSD below 30% in QC samples) the total number of molecular features was 682. Regarding negative mode, 2186 molecular features were achieved which after the filtering by frequency and variability were reduced to a total of 468.

PCA was chosen to carry out the multivariate statistical analysis since it allows reducing huge amounts of data into a simple model to discover the most significant differences among groups of samples. PCA score plots with and without QC samples for positive and negative ionization modes are shown in Fig. 1. As can be seen in this figure, QC samples were clustered and centered in the score plot, demonstrating good analytical consistency of the data during all the metabolomics sequences. Moreover, samples corresponding to the same alkalization degree were also clustered and high differentiation among them was achieved for both ionization modes. The first and second components of the PCA models explained 60 and 24% of the variance, respectively, for positive ionization mode and 54 and 27% of the variance for negative ionization mode.

Once the unsupervised multivariate analysis was performed, PLS-DA was used to discriminate group samples. A PLS-DA model was built with the four groups of cocoa samples, including natural and differently alkalized groups (CN, CL, CM, CS). For both ionization modes, the different groups of samples were well differentiated. Table 2 shows the quality parameters of the PLS-DA models ( $R^2X$ ,  $R^2Y$ , and  $Q^2$ ) and the F and p-values of the ANOVA test. The quality of the model was demonstrated by the high F value and the low p-value obtained together with the results of the permutation test ( $Q^2$  and  $R^2$  were below the original values). Since this model cannot reveal slight differences among groups of samples, several PLS-DA models were built to compare natural cocoa with each alkalization degree in a pairwise way (CN Vs CL, CN Vs CM, CN Vs CS). These models are shown in Fig. 2 whereas their quality parameters are included in Table 2. In the PLS-DA model established to compare CN Vs CL in positive ionization mode (Fig. 2A), a sample can be observed out of the 95% of the confidence limit represented by the hotelling's ellipse; however, taking into account that it was inside the 99% of the hotelling's ellipse it was not considered as an outlier and was kept to carry out the PLS-DA analysis. In all cases, high  $Q^2$  values were



**Fig. 1.** Principal Component Analysis (PCA) score plot obtained in positive ionization mode with (A) and without (B) QC and negative ionization mode with (C) and without (D) QC. Hotelling's ellipses represent the 95% confidence limit. QC, quality control (green); CN, natural cocoa (yellow); CL, light alkalized cocoa (dark blue); CM, medium alkalized cocoa (red); CS, strong alkalized cocoa (light blue). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

**Table 2**

Quality parameters ( $R^2X$ ,  $R^2Y$ , and  $Q^2$ ) of the PLS-DA models and F and p-values of the cross-validated ANOVA obtained in both ESI modes.

	ESI +		ESI-	
	Quality parameters	Cross-Validated ANOVA	Quality parameters	Cross-Validated ANOVA
CN, CL, CM, CS	$R^2X = 0.950$ $R^2Y = 0.991$ $Q^2 = 0.988$	$F = 433.0$ p-value = 0	$R^2X = 0.877$ $R^2Y = 0.984$ $Q^2 = 0.983$	$F = 424.1$ p-value = 0
CN vs. CL	$R^2X = 0.903$ $R^2Y = 0.994$ $Q^2 = 0.992$	$F = 925.8$ p-value = $9.1 \times 10^{-27}$	$R^2X = 0.800$ $R^2Y = 0.996$ $Q^2 = 0.995$	$F = 2627.7$ p-value = $1.2 \times 10^{-31}$
CN vs. CM	$R^2X = 0.774$ $R^2Y = 0.982$ $Q^2 = 0.977$	$F = 563.2$ p-value = $9.7 \times 10^{-23}$	$R^2X = 0.806$ $R^2Y = 0.997$ $Q^2 = 0.997$	$F = 3894.0$ p-value = $5.9 \times 10^{-34}$
CN vs. CS	$R^2X = 0.877$ $R^2Y = 0.990$ $Q^2 = 0.990$	$F = 1362.9$ p-value = $7.7 \times 10^{-28}$	$R^2X = 0.860$ $R^2Y = 0.998$ $Q^2 = 0.998$	$F = 6805.5$ p-value = $1.2 \times 10^{-37}$

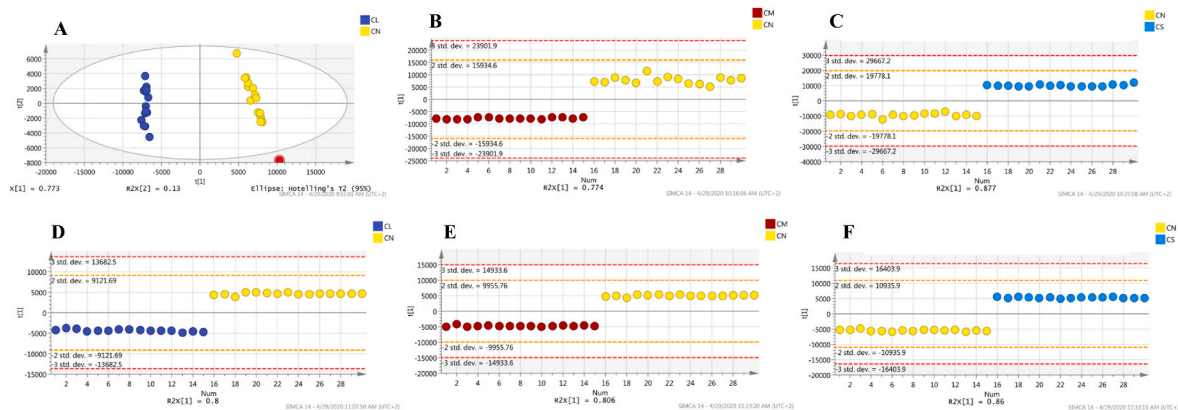
obtained ( $Q^2 > 0.997$ ) along with high F values ( $F > 563.2$ ) and low p-values ( $9.7 \times 10^{-23}$ ) for the cross-validated ANOVA test. Moreover, cross-validation and permutation tests were carried out to reinforce the idea that group separation was owing to differences between metabolic profiles and not to data overfitting. Figs. S3 and S4 showed the PLS-DA models of CV score plot for the pairwise comparison and permutation tests, demonstrating a good separation between groups and a good quality of the built models (in permutation tests  $Q^2$  and  $R^2$  values were

below the original values and a positive slope for both parameters was obtained). All these factors, together with PCA models demonstrated that significant differences among the analyzed groups can be found. The variables responsible for the differentiation among natural cocoa samples and cocoa samples at different alkalization degrees were determined by the variable importance in the projection (VIP) value of PLS-DA models. Thus, those variables with a VIP value higher than 2.0 were selected as the most important variables that differentiate groups of cocoa samples. Following this criteria, a total of 43 variables in positive ionization mode and 30 in negative ionization mode were obtained.

### 3.3. Metabolite identification

The identification of significant metabolites corresponds to the most time-consuming and laborious steps in metabolomics due to several factors; a lack of commercially available standards or the high price of those that are commercially available, and the difficult comparison of the obtained MS/MS spectra with those reported in the literature or libraries. The use of tandem HRMS allows the obtention of accurate mass and MS/MS spectra which facilitates the identification of metabolites.

A summary of retention times, molecular formula, experimental  $m/z$ , mass error, main MS/MS fragments, VIP value of the pairwise PLS-DA models, and the trend of the significant metabolites along the alkalization process is shown in Tables 3 and 4 for positive and negative ionization modes, respectively. Fig. S5 shows the diagrams of the trends observed for the tentatively and unequivocally identified compounds along the alkalization degrees. Sixteen metabolites were identified in the positive mode (seven of them unequivocally identified using commercial



**Fig. 2.** Partial Least Square Discriminant Analysis (PLS-DA) models comparing CN with different alkalized cocoa samples in both positive ionization mode (A) CN vs CL; (B) CN vs CM; (C) CN vs CS and negative ionization mode (D) CN vs CL; (E) CN vs CM; (F) CN vs CS.

**Table 3**

Tentative identification of significant variables in positive ionization mode.

N°	rt	Molecular formula	Tentative Identification	[M-H] <sup>+</sup>	Mass error (ppm)	Main MS/MS fragments	VIP values			
							CN vs CL	CN vs CM	CN vs CS	Alkalization trend
1	0.8	C <sub>5</sub> H <sub>11</sub> N <sub>3</sub> O <sub>2</sub>	4-(1-guanidino)butyric acid	146.0925	1	43.0179, 87.0446	3.48186	3.4172	3.4172	↓
2	0.9		Unknown	182.0516		43.0243, 87.0525, 98.0526	2.548	2.50817	3.96617	↑
3	1	C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub>	Pyroglutamic acid <sup>a</sup>	130.0503	3	41.0384, 56.0495, 84.0445	2.72718	2.72003	1.90019	—
4	1.1	C <sub>6</sub> H <sub>13</sub> NO <sub>2</sub>	Leucine <sup>a</sup>	132.1024	4	44.0501, 69.0705, 86.0974	7.56678	8.71567	11.0701	↓
5	1.2	C <sub>10</sub> H <sub>13</sub> NO <sub>2</sub>	Salsolinol	180.1024	3	91.0551, 136.0757, 145.0649, 163.0747	2.61367	2.78787	2.3043	—
6	1.2		Unknown	345.1845		86.0971, 132.1020		1.91924	2.59051	↓
7	1.3		Unknown	86.0606		44.0132, 56.0500	3.29926	3.84774	3.84774	↑
8	1.4		Unknown	190.1076		69.0703, 98.0967, 144.1017	2.60209	2.56433		↓
9	1.5		Unknown	330.1665		70.0655, 169.1341	2.21483	2.18167	1.75923	—
10	1.5		Unknown	289.1398		72.0809, 143.1185, 189.1028, 235.1041	2.63741			↑
11	1.6	C <sub>10</sub> H <sub>13</sub> N <sub>5</sub> O <sub>4</sub>	Adenosine <sup>a</sup>	268.1043	1	136.0617	1.98507	2.29881	2.2939	↑
12	1.8		Unknown	144.0657		70.0649, 98.0599	2.60686	2.29312	2.15612	—
13	1.8	C <sub>10</sub> H <sub>20</sub> N <sub>2</sub> O <sub>3</sub>	Val-Val	217.1548	1	72.0806			2.49914	↓
14	1.9		Unknown	254.1613		70.0651, 98.0598, 125.0705, 167.1185, 195.1135	4.31301	4.35811	4.35811	—
15	1.9	C <sub>10</sub> H <sub>11</sub> NO <sub>2</sub>	Dehydrosalsolinol	178.0863	0	162.0548			2.093	↑
16	2		Unknown	103.0545		51.0228, 77.0292	2.55347	2.73779	2.868	↓
17	2	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	Phenylalanine <sup>a</sup>	166.0867	3	77.0391, 103.0547, 120.0808	5.73325	6.48246	6.48246	↓
18	2		Unknown	256.1770		169.1334, 197.1285	2.37624	2.28688	1.80134	—
19	2.1		Unknown	199.1082		109.0757, 121.076, 135.0913, 163.0864, 181.0967	2.14039	1.75327		↑
20	2.1	C <sub>7</sub> H <sub>11</sub> NO <sub>2</sub>	L-Hypoglycin A	142.0867	3	41.0385, 69.0333, 74.0601, 96.0812, 124.0761	1.93114	2.27454	2.42329	↑
21	2.4		Unknown	242.1869		84.0805, 141.1027, 169.0972, 197.1643, 225.1591	2.0955	2.11894	1.62489	↓
22	2.5	C <sub>10</sub> H <sub>21</sub> N <sub>5</sub> O <sub>4</sub>	Arg-Thr	276.1710	16	84.0805, 120.0802, 175.08631, 259.1437	3.00609	3.05811	3.05811	↓
23	2.5	C <sub>7</sub> H <sub>8</sub> N <sub>4</sub> O <sub>2</sub>	Theobromine <sup>a</sup>	181.0723	2	42.0340, 110.0710, 138.0655, 163.0606	3.04819	1.97436	1.97436	↑
24	2.6		Unknown	399.2351		158.0926, 197.1282		2.10664	2.48566	↓
25	2.7		Unknown	304.1776		98.0597, 120.0798, 217.1329, 245.1287	3.21448	3.11444	3.11444	—
26	2.8		Unknown	188.0711		118.0646, 146.0587	2.16518	2.29611	2.37581	↓
27	3.4		Unknown	154.0505		94.0288, 122.0231	2.75232	2.86339		↑
28	3.7	C <sub>8</sub> H <sub>10</sub> N <sub>4</sub> O <sub>2</sub>	Caffeine <sup>a</sup>	195.0882	3	110.0716, 138.0661		2.06151	1.54271	↑
29	4.1	C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>	Catechin <sup>a</sup>	291.0868	2	123.0441, 139.0382	1.79712	2.23711	2.20814	↓
30	4.3		Unknown	402.1663		70.0647, 112.0864, 250.1195	2.21883	2.37674	2.12833	↓
31	5.4		Unknown	412.2177		85.02923, 103.0386, 127.0384, 145.0505	3.63488	3.26868	3.26868	↓
32	8.8	C <sub>18</sub> H <sub>37</sub> NO <sub>3</sub>	Dehydrophytosphingosine	316.2852	2	60.04473	2.01273	1.99366	1.66497	—
33	9.1	C <sub>26</sub> H <sub>50</sub> NO <sub>8</sub> P	Lyso/OH-18:2a-PC	536.3349	3	104.1067, 184.0726, 518.3232	2.0289			↓
34	9.2		Unknown	415.2117		119.0857		3.99285	3.26257	↓
35	9.3		Unknown	315.2535		weak signals	2.33307	2.21701	2.46251	↑
36	9.7		Unknown	566.4281		86.0967, 227.1754, 340.2593, 453.3434	2.87346			↓
37	9.7		Unknown	368.3528		309.279	2.85837	2.52488		↓
38	9.9	C <sub>23</sub> H <sub>44</sub> NO <sub>7</sub> P	Lyso PE (18:2)	478.2933	1	44.0494, 306.2792, 337.2737	2.80299	2.79131	2.27551	—
39	10.1		Unknown	541.3351		No fragmentation pattern	2.28089	2.00786		↓
40	10.7		Unknown	637.3061		57.0699, 337.0236, 393.0869, 525.1809, 581.02423		2.55833	2.04382	↓
41	10.7		Unknown	256.2640		weak signals	2.29202	2.8445		↓
42	10.8		Unknown	282.2795		69.0695	2.22081	2.57607		↓
43	11.3		Unknown	284.2953		43.0548	2.05833	2.31216		↓

↑ The level of the compound increases with alkalization. ↓ The level of the compound decreases with alkalization. — The level of the compound does not change significantly with alkalization.

<sup>a</sup> Confirmed with commercial standard.

standards), whereas in the negative mode it was also possible to identify a total of fifteen metabolites (five of them unequivocally identified using commercial standards).

Different protein and non-protein amino acids were unequivocally identified by comparing their retention time and MS/MS fragmentation pattern with those obtained for commercial standards. Thus, it was

possible to identify pyroglutamic acid both in positive (**compound 3**, [M+H]<sup>+</sup> = 130 *m/z*, with an intense fragment at *m/z* 84 (C<sub>3</sub>H<sub>2</sub>NO<sub>2</sub><sup>+</sup>)) and in negative mode (**compound 3**, [M-H]<sup>−</sup> = 128 *m/z*). Also, the protein amino acid phenylalanine was unequivocally identified as **compound 17** ([M+H]<sup>+</sup> = 166 *m/z*) or **compound 5** ([M-H]<sup>−</sup> = 164 *m/z*) in positive and negative ionization mode, respectively, taking into

**Table 4**

Tentative identification of significant variables in negative ionization mode.

N°	rt	Molecular formula	Tentative Identification	[M-H] <sup>-</sup>	Mass error (ppm)	Main MS/MS fragments	VIP values			
							CN vs CL	CN vs CM	CN vs CS	Alkalization trend
1	0.6	C <sub>4</sub> H <sub>6</sub> O <sub>5</sub>	Malic acid	133.0160	13	43.0197, 71.0147, 115.0042	2.48679	2.24862	2.49771	↑
2	0.6	C <sub>6</sub> H <sub>8</sub> O <sub>7</sub>	Citric acid <sup>a</sup>	191.0217	11	57.0352, 58.0299, 87.0097, 111.0094	6.81052	6.81382	5.45201	—
3	0.7	C <sub>5</sub> H <sub>7</sub> NO <sub>3</sub>	Pyroglutamic acid <sup>a</sup>	128.0365	9	52.0202	2.33055	2.68721	1.91661	↓
4	1.2	C <sub>6</sub> O <sub>5</sub> H <sub>10</sub>	3-hydroxy-3-methylglutaric acid	161.0476	13	57.0351	3.07484	2.37583	2.76056	↓
5	1.9	C <sub>9</sub> H <sub>11</sub> NO <sub>2</sub>	Phenylalanine <sup>a</sup>	164.0734	11	72.0097, 103.0565, 147.0426	1.96552	2.11213	2.42047	↓
6	2.3	C <sub>7</sub> H <sub>6</sub> O <sub>4</sub>	Protocatechuic acid <sup>a</sup>	153.0210	11	108.0217, 109.0296			2.47663	↑
7	2.4	C <sub>9</sub> H <sub>14</sub> O <sub>7</sub>	Isopropyl citrate	233.0687	9	43.0189, 59.0141, 101.0613	3.38731	2.49272		↓
8	2.7	C <sub>13</sub> H <sub>13</sub> NO <sub>7</sub>	N-caffeoyl -L-aspartate	294.0635	5	88.0409, 132.0312, 161.0248	1.52599	2.39425	3.28917	↓
9	2.9	C <sub>7</sub> H <sub>12</sub> O <sub>5</sub>	2-isopropylmalic acid	175.0630	10	59.0145, 85.0662, 115.0413, 131.0713	2.63139		2.22225	—
10	3	C <sub>7</sub> H <sub>6</sub> O <sub>3</sub>	Protocatechuic aldehyde	137.0255	8	108.0222	2.72169	2.70208	4.89866	↑
11	3.6		Unknown	304.1058		94.0294, 123.0321, 156.1034, 180.1041	2.44136	2.32191		↑
12	3.7		Unknown	329.0917		96.9601	2.44815	2.88917	2.43585	—
13	3.8		Unknown	407.1574		57.0346, 96.9597, 167.0391	1.59539	1.51575	2.28219	↓
14	3.8		Unknown	329.0917		96.9600	2.41886	2.11289	2.79163	↓
15	3.8		Unknown	131.0726		33.1172, 69.0842, 85.0647	1.82685	2.82985		↓
16	4.1		Unknown	409.1733		59.0135, 101.0242	1.77651	1.6125	2.15905	↓
17	4.1	C <sub>12</sub> H <sub>18</sub> O <sub>7</sub> S	Hydroxy-jasmonic acid sulfate	305.0721	7	59.0139, 96.9600, 181.1226, 225.1134	3.73193	5.31579	5.00874	—
18	4.2	C <sub>15</sub> H <sub>14</sub> O <sub>6</sub>	Catechin <sup>a</sup>	289.0748	10	245.0821, 205.0525, 137.0240, 125.0238, 109.0296	1.69765	1.97921	2.15111	↓
19	4.9		Unknown	286.0948		96.0449, 123.0328, 180.1036	2.38075	2.30051	1.82278	↑
20	5	C <sub>16</sub> H <sub>14</sub> O <sub>8</sub>	Unknown	333.0635	6	163.0033, 315.0517	1.98457	1.51501	2.31205	↑
21	5.3	C <sub>17</sub> H <sub>30</sub> O <sub>10</sub>	Hexenyl primeveroside (isomer 1)	393.1787	5	57.0341, 99.0453, 250.1391	5.80185	5.05596	7.39796	↓
22	5.9	C <sub>17</sub> H <sub>30</sub> O <sub>10</sub>	Hexenyl primeveroside (isomer 2)	393.1787	5	57.0345, 99.0453, 249.1345			2.23818	↓
23	6.6		Unknown	357.1242		96.9595	2.07514	2.19702	2.29994	↓
24	8.5		9,12,13-trihydroxyoctadecanoid acid	329.2361		99.0818, 139.1134, 171.1042, 211.1344, 229.1447	2.08788	1.70717	1.92079	—
25	9.5		Unknown	315.2569		58.0063, 116.9288, 141.1285	3.5022	2.45643		—
26	9.5		Unknown	295.2304		141.12921	1.86755		2.04202	—
27	9.9		Unknown	377.2118		78.9593			2.54471	↑
28	10.8		Unknown	279.2351		44.9981, 96.9597, 211.2690		2.06385	1.87428	↑
29	10.9	C <sub>18</sub> H <sub>36</sub> O <sub>3</sub>	Hydroxystearic acid isomer	299.2613	7	59.0141	2.01305	2.0352		↓
30	11.2		Unknown	281.2503	13	weak signals	2.10682	3.19474	2.88718	↑

↑ The level of the compound increases with alkalization. ↓ The level of the compound decreases with alkalization. — The level of the compound does not change significantly with alkalization.

<sup>a</sup> Confirmed with commercial standard.

account its characteristic fragmentation pattern (fragments at  $m/z$  120 (C<sub>8</sub>H<sub>10</sub>N<sup>+</sup>) and  $m/z$  103 (C<sub>8</sub>H<sub>7</sub><sup>+</sup>) in the positive mode and fragments at  $m/z$  147 (C<sub>9</sub>H<sub>7</sub>O<sub>2</sub><sup>-</sup>), 103 (C<sub>8</sub>H<sub>7</sub><sup>-</sup>) and 72 (C<sub>2</sub>H<sub>2</sub>NO<sub>2</sub><sup>-</sup>) in the negative mode). The MS/MS spectrum of **compound 4** ([M+H]<sup>+</sup> = 132  $m/z$ ) showed the fragments  $m/z$  86 (C<sub>5</sub>H<sub>12</sub>N<sup>+</sup>), 69 (C<sub>5</sub>H<sub>9</sub><sup>+</sup>), and 44 (C<sub>2</sub>H<sub>6</sub>N<sup>+</sup>) (Jiang et al., 2020) which were also found in Leucine commercial standard. The content of pyroglutamic acid in all the alkalization degrees studied was similar and lower than that obtained in CN, but a clear trend with the alkalization degree was not observed. However, the content of leucine and phenylalanine decreased as the alkalization degree increased which was in agreement with findings described by other authors who have related this trend with the fact that these compounds are mainly involved in Maillard and oxidative deamination reactions or they are interacting with peptides and polyphenols (Li et al., 2012; Valverde-García et al., 2020). **Compounds 11, 23, and 28** were also

unequivocally identified in positive mode. The first of them was identified as adenosine ([M+H]<sup>+</sup> = 268  $m/z$ ) due to the loss of C<sub>5</sub>H<sub>8</sub>O<sub>4</sub> moiety which generated a fragment ion at  $m/z$  136 which corresponds to adenine (Van Dycke et al., 2010). Regarding the behavior of this compound during the alkalization, it clearly increased during the process. **Compounds 23** ([M+H]<sup>+</sup> = 181  $m/z$ ) and **28** ([M+H]<sup>+</sup> = 195  $m/z$ ) were identified as the alkaloids theobromine and caffeine, respectively, based on the comparison of their retention times and MS/MS fragmentation patterns to those obtained for the commercial standards. Both compounds are the most abundant methylxanthines in cocoa. As previously reported by Bianco et al. caffeine showed an intense ion fragment at  $m/z$  138 due to the neutral loss of methyl isocyanate (−57 Da) as a consequence of a retro-Diels Alder rearrangement (Bianco et al., 2009). A consecutive loss of CO (−28 Da) from this fragment gave rise to an ion at  $m/z$  110. For theobromine, the neutral loss of water yielded an

ion fragment at  $m/z$  163. Subsequently, the same fragmentation pattern of caffeine was observed (ion fragments at  $m/z$  138 and 110). Although other authors have reported a decrease in the level of these compounds when the alkalization degree increases (Li et al., 2012), the trend observed in this work corresponds to a slight increase of theobromine and caffeine levels at the highest alkalization degree in comparison with that obtained for natural cocoa. Sioriki et al. reported that alkalization did not produce any significant change in the concentration of theobromine in the cocoa powder which is in agreement with our results (Sioriki et al., 2021, 2022). Regarding caffeine, these authors reported a decrease in the caffeine content by 22–35% in the first min of alkalization while no significant changes were observed with higher alkalization temperature and longer times.

**Compound 29** ( $[M+H]^+ = 291$   $m/z$ ), detected in positive ionization mode is the same as **compound 18** ( $[M-H]^- = 289$   $m/z$ ) found under negative ionization mode. The fragmentation pattern obtained experimentally (which was in agreement with that reported by Sánchez-Rabeneda et al. and that obtained by the injection of the pure standard, enabled the identification of this compound as catechin (Sánchez-Rabeneda et al., 2003). The level of this polyphenol decreased along with the increase in the alkalization degree (see Fig. S5). This trend was previously explained by other authors as the increased pH in strong alkalized cocoa powder is responsible for the oxidation and degradation of polyphenols and thus, the content of polyphenols decreased in these samples (Quelal-Vásquez et al., 2020; Valverde-García et al., 2020).

Under MS/MS fragmentation of **compound 1** ( $[M+H]^+ = 146$   $m/z$ ) in positive ionization mode, ion fragments at  $m/z$  87 and  $m/z$  43 were obtained which matched with the predicted ones described in the HMDB database for 4-(1-guanidino)butyric acid (a compound belonging to the family of  $\gamma$ -amino acids). In this case, the alkalization of natural cocoa powder led to an increase in the compound content which is similar at the two first alkalization levels (CL and CM) and decreased for the strong alkalization (CS). The  $[M+H]^+$  at  $m/z$  180 and  $m/z$  178, corresponding to compounds **5** and **15**, respectively, were tentatively identified as salsolinol (a phenolic compound previously identified in cocoa powder) and dehydrosalsolinol, thanks to the comparison of their fragmentation patterns with those previously reported in the literature (Riggin & Kissinger, 1976). On the one hand, Cai & Liu described the same fragment ions for salsolinol (Cai & Liu, 2008) whereas Nikolic et al. reported the intense fragment ion at  $m/z$  162 for dehydrosalsolinol (Nikolić et al., 2012). These compounds showed a different behavior when increasing the alkalization degree. The content of salsolinol was similar in all the alkalized samples and lower than that obtained in CN. On the contrary, the level of dehydrosalsolinol increased from CN to CL and CM (where is practically constant) and then increased with further alkalization.

**Compound 13** ( $[M+H]^+ = 217$   $m/z$ ) and **compound 22** ( $[M+H]^+ = 276$   $m/z$ ) were both tentatively identified as dipeptides by the comparison of their fragmentation patterns with those obtained in FooDB database. Compound **13** showed a fragment ion at  $m/z$  72 and it was assigned to Val-Val, whereas compound **22** was tentatively assigned to Arg-Thr since this compound exhibited fragments at  $m/z$  175 and  $m/z$  120 that correspond to the loss of arginine and threonine, respectively. The trend of both dipeptides was to decrease their amount as the alkalization degree increased, which is in agreement with previous literature where it has been described that the content of proteins, peptides, and amino acids decreases in alkalized cocoa samples since this process changes the media leading to the degradation of this kind of compounds (Adeyeye, 2016; Méndez-Albores et al., 2013; Odus & Longe 1998). The ion at  $m/z$  142 ( $[M+H]^+$ ) was tentatively identified as L-Hypoglycin A (**compound 20**). According to Carlier et al. the breakdown of the C-C bond of the C $\alpha$  amine gave two main fragments at  $m/z$  96 and 74 which are present in the obtained MS/MS fragmentation spectrum (Carlier et al., 2015). Under MS/MS fragmentation of **compound 32** ( $[M+H]^+ = 316$   $m/z$ ) in positive ionization mode, ion  $m/z$  60 ( $C_2H_6NO$  moiety) was obtained. The comparison of its MS/MS fragmentation with that

found in FooDB database allowed tentatively identifying this compound as dehydropyrosphingosine. The level of this compound strongly decreased when the sample was submitted to the alkalization process. However, the variation of the alkalization degree only gave rise to a slight variation in the content of this 1,3-aminoalcohol. The last two compounds tentatively identified in positive ionization mode corresponded to **compounds 33** and compound **38**, that were assigned to lyso/OH-18:2a-phosphatidylcholine and lysophosphatidylethanolamine (18:2), respectively. The MS/MS spectrum of compound **33** showed a fragment at  $m/z$  518 corresponding to the loss of a water molecule  $[M + H-H_2O]^+$  from the  $[M+H]^+$  ion at  $m/z$  536. **Compound 33** also presented an intense fragment at  $m/z$  184 which is characteristic of lysophosphatidylcholines (Fang et al., 2003), enabling the identification of this compound as lyso/OH-18:2a-phosphatidylcholine. **Compound 38** showed a fragment at  $m/z$  337 which was assigned to a neutral loss of 141 Da characteristic of lysophosphatidylethanolamines (Fang et al., 2003). The comparison of the level of these compounds between CN and CL showed an opposite trend since while the level of **compound 33** increased, the level of **compound 38** decreased. However, the level of these two compounds did not change significantly when the alkalization degree was modified.

In negative ionization mode, small organic acids were also highlighted as markers of cocoa alkalization. In particular, **compound 1** ( $[M-H]^- = 133$   $m/z$ ) and **compound 2** ( $[M-H]^- = 191$   $m/z$ ) were identified as malic and citric acids, respectively. Malic acid (**compound 1**) presented an intense fragment at  $m/z$  115 as a consequence of the loss of a water molecule ( $-18$  Da), whereas citric acid (**compound 2**), which had an intense fragment ion at  $m/z$  111 corresponding to  $[M-H-CO_2-2H_2O]^-$ , was unequivocally identified by comparing its fragmentation pattern and retention time with those obtained when the commercial standard was injected into the MS instrument (Fernández-Fernández et al., 2010). These two compounds were related to **compounds 7** ( $[M-H]^- = 233$   $m/z$ ) and **9** ( $[M-H]^- = 175$   $m/z$ ). The MS/MS fragmentation of **compound 7** showed a fragment ion  $m/z$  101 ( $C_4H_5O_3$ ) and  $m/z$  59 characteristic of acetic moiety ( $C_2H_3O_2$ ). According to that and comparing its MS/MS spectrum with that described in the HMDB database, this compound was tentatively identified as isopropyl citrate. **Compound 9**, which was tentatively identified as isopropylmalic acid, showed an intense fragment ion at  $m/z$  115  $[M-H-2CHOH]^-$  ( $-60$  Da, loss of acetic residue) (Gómez-Romero et al., 2010). Even though it is known that citric, lactic, oxalic, and succinic acids contribute to the acidity of cocoa (Valverde-García et al., 2020), to the best of our knowledge, there are no reported works in which the variation of malic and citric acid with the alkalization degrees has been described. The trends observed for these compounds when increasing the alkalization degree were different between them. On the one hand, the content of malic acid slightly increased when the cocoa sample was submitted to the highest alkalization degree whereas the content of isopropyl citrate decreased. On the other hand, the content of citric acid was practically constant in all the alkalization degrees, and the level of isopropylmalic acid was similar in the lowest and the highest alkalization degree and decreased slightly in the medium level.

**Compound 4** which showed a  $[M-H]^-$  ion at  $m/z$  161 and a fragment at  $m/z$  57 was identified as 3-hydroxy-3-methylglutaric acid by the comparison of its fragmentation spectrum with that obtained in the database FooDB. The level of this compound was lower for all alkalization degrees when compared with its level in CN. In addition, its content was similar in CL and CM and lower in CS. **Compounds 6** ( $[M-H]^- = 153$   $m/z$ ) and **10** ( $[M-H]^- = 137$   $m/z$ ) were assigned to protocatechuic acid (which was unequivocally identified) and protocatechuic aldehyde (Barbana et al., 2017), respectively. For both compounds, an increase in their content was observed in strongly alkalized cocoa samples. The increase was more pronounced in the case of protocatechuic aldehyde.

The  $[M-H]^-$  ion at  $m/z$  294 (**compound 8**), was tentatively identified as *N*-caffeoyl-L-aspartate (an *N*-phenylpropenoyl-L-amino acid

hydroxycinnamoyl amino acid conjugate). The MS/MS spectrum of this compound showed an intense fragment ion at  $m/z$  132 corresponding to the double decarboxylation and following  $N$ -CO  $\alpha$ -cleavage which corresponds to aspartate. In addition, the ion at  $m/z$  161 was useful to confirm the hydroxycinnamoyl moiety (Lavorgna et al., 2021). **Compound 17** ( $[M-H]^- = 305$   $m/z$ ) was tentatively assigned to hydroxyjasmonic acid sulfate which had an intense MS/MS fragment at  $m/z$  225 (Patras et al., 2014). This compound has been previously reported in studies related to the effect of the fermentation process in cocoa beans studied (Patras et al., 2014; D'Souza et al., 2017). The level of this compound did not change significantly during the alkalization process.

**Compounds 21 and 22** displayed the same  $[M-H]^-$  ion at  $m/z$  393. The main fragment at  $m/z$  249 observed in their MS/MS spectra allow the identification of these compounds as hexenyl primeveroside isomers, according to the literature (Cádiz-Gurrea et al., 2010, 2020). Compound **24** presenting fragments at  $m/z$  99, 139, 171, 211, and 229 was tentatively identified as 9,12,13-trihydroxyoctadecanoid acid by the comparison of its MS/MS fragmentation with that reported by Vargas-Arana et al. (Vargas-Arana et al., 2022). As in the case of compound 17, the alkalization process did not change significantly the level of 9,12,13-trihydroxyoctadecanoid acid. The last metabolite tentatively identified in negative ionization mode corresponds to **compound 29** which showed the  $[M-H]^-$  ion at  $m/z$  299 and a fragment ion at  $m/z$  59. It was tentatively identified as an isomer of hydroxystearic acid according to Kokotou et al. (Kokotou et al., 2020). Its level was increased from CN to CL and CM and decreased with strong alkalization (although its level continues to be higher than in CN).

From the results obtained, it is demonstrated that the metabolomic assessment of cocoa power during the alkalization process by using the developed analytical methodology allows proposing a set of compounds of different nature as potential markers of the alkalization process. However, to achieve a more detailed information about the potential of these markers, it would be necessary to obtain a more comprehensive analysis so that the metabolomics strategy should be employed to analyze a larger number of samples.

#### 4. Conclusions

In this study, an untargeted metabolomic strategy based on the use of UHPLC-MS was developed to investigate the changes taking place in cocoa powder submitted to different alkalization degrees. Once optimized the extraction protocol to obtain the highest number of molecular features, different samples were analyzed and the data obtained were treated by using chemometrics tools. This approach allowed to suggest 43 metabolites in positive ionization mode and 30 in negative ionization mode as potential markers capable of establishing differences among cocoa powder samples submitted to different alkalization degrees. Just three of these metabolites were observed both in positive and negative modes which demonstrates the relevance of performing the metabolomics in both ionization modes to achieve a broad metabolite coverage. Among all the metabolites, a total of 31 could be identified, being 9 of them unequivocally identified. Mainly, they were identified as amino acids, alkaloids, organic acids, or polyphenols, among others. Different trends related to the effect of the alkalization degree on the level of the identified compounds were observed. The present data provide comprehensive information on the effects of alkalization process on the metabolites of cocoa powder samples demonstrating that this process has a great impact on their chemical composition. From a practical point of view, the evaluation of changes in the chemical composition of cocoa powder samples submitted to different alkalization degrees by using the metabolomics strategy provides valuable information for future development and improvement of the nutritional and sensory characteristics of cocoa powder samples.

#### CRedit authorship contribution statement

**Maidier Greño:** Investigation, Formal analysis, Validation, Data curation, Visualization, Writing – original draft. **Miguel Herrero:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing, Project administration, Funding acquisition. **Alejandro Cifuentes:** Writing – review & editing, Funding acquisition. **María Luisa Marina:** Conceptualization, Methodology, Resources, Supervision, Writing – review & editing, Project administration, Funding acquisition. **María Castro-Puyana:** Conceptualization, Methodology, Visualization, Data curation, Resources, Supervision, Writing – original draft, Writing – review & editing, Project administration, Funding acquisition.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgements

Authors thank the Comunidad de Madrid and European funding from FSE and FEDER programs for project S2018/BAA-4393 (AVANSECAL-II-CM) and the University of Alcalá for project CCGP2017-EXP/057. M.C. P. thanks MINECO for her “Ramón y Cajal” research contract (RYC-2013-12688). M.G. thanks the Spanish Ministry of Science, Innovation, and Universities for her pre-doctoral contract (FPU17/01635). Authors also thank OLAM for the kind donation of cocoa powder samples.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.lwt.2022.114207>.

#### References

- Adeyeye, E. I. (2016). Proximate, mineral and antinutrient compositions of natural cocoa cake, cocoa liquor and alkalized cocoa powders sourced in Nigeria. *Journal of Advance Pharmaceutical Science and Technology*, 1, 12–28. <https://doi.org/10.14302/issn.2328-0182.japst-15-855>
- Alasti, F. M., Asefi, N., Maleki, R., & SeiedlouHeris, S. S. (2019). Investigating the flavor compounds in the cocoa powder production process. *Food Sciences and Nutrition*, 7, 3892–3901. <https://doi.org/10.1002/fsn3.1244>
- Alasti, F. M., Asefi, N., Maleki, R., & SeiedlouHeris, S. S. (2020). The influence of three different types and dosage of alkaline on the inherent properties in cocoa powder. *Journal of Food Science & Technology*, 57, 2561–2571. <https://doi.org/10.1007/s13197-020-04293-w>
- Andres-Lacueva, C., Monagas, M., Khan, N., Izquierdo-Pulido, M., Urpi-Sarda, M., Permanyer, J., & Lamuela-Raventós, R. M. (2008). Flavanol and flavonol contents of cocoa powder products: Influence of the manufacturing process. *Journal of Agricultural and Food Chemistry*, 56, 3111–3117. <https://doi.org/10.1021/jf0728754>
- Avanza, M. V., Álvarez-Rivera, G., Cifuentes, A., Mendiola, J. A., & Ibáñez, E. (2021). Phytochemical and functional characterization of phenolic compounds from Cowpea (*Vigna unguiculata* (L.) Walp.) obtained by green extraction technologies. *Agronomy*, 11, 162. <https://doi.org/10.3390/agronomy11010162>
- Barbana, C., Nardin, T., Pierotti, A., Malacarne, M., & Larcher, R. (2017). Targeted and untargeted characterisation of free and glycosylated simple phenols in cocoa beans using high resolution-tandem mass spectrometry (Q-Orbitrap). *Journal of Chromatography A*, 1480, 41–49. <https://doi.org/10.1016/j.chroma.2016.12.022>
- Bianco, G., Abate, S., Labella, C., & Cataldi, T. R. I. (2009). Identification and fragmentation pathways of caffeine metabolites in urine samples via liquid chromatography with positive electrospray ionization coupled to a hybrid quadrupole linear ion trap (LTQ) and Fourier transform ion cyclotron resonance mass spectrometry and tandem mass spectrometry. *Rapid Communications in Mass Spectrometry*, 23, 1065–1074. <https://doi.org/10.1002/rcm.3969>
- Cádiz-Gurrea, M. L., Fernández-Ochoa, A., Leyva-Jiménez, F. J., Guerrero-Muñoz, N., Villegas-Aguilar, M. C., Pimentel-Moral, S., Ramos-Escudero, F., & Segura-Carretero, A. (2020). LC-MS and spectrophotometric approaches for evaluation of

- bioactive compounds from Peru cocoa by-products for commercial applications. *Molecules*, 25, 3177. <https://doi.org/10.3390/molecules25143177>
- Cádiz-Gurrea, M. L., Lozano-Sánchez, J., Contreras-Gómez, M., Legeai-Mallet, L., Fernández-Arroyo, S., & Segura-Carretero, A. (2010). Isolation, comprehensive characterization and antioxidant activities of Theobroma cacao extract. *Journal of Functional Foods*, 10, 485–498. <https://doi.org/10.1016/j.jff.2014.07.016>
- Cai, M., & Liu, Y.-M. (2008). Quantification of salsolinol enantiomers by stable isotope dilution liquid chromatography with tandem mass spectrometric detection. *Rapid Communications in Mass Spectrometry*, 22, 4171–4177. <https://doi.org/10.1002/rcm.3847>
- Carlier, J., Guitton, J., Moreau, C., Boyer, B., Bevalot, F., Fanton, L., Habyarimana, J., Gault, G., & Gaillard, Y. (2015). A validated method for quantifying hypoglycin A in whole blood by UHPLC-HRMS/MS. *Journal of Chromatography B*, 978–979, 70–77. <https://doi.org/10.1016/j.jchromb.2014.11.029>
- Castro-Puyana, M., Perez-Míguez, R., Montero, L., & Herrero, M. (2017). Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. *TrAC, Trends in Analytical Chemistry*, 93, 102–118. <https://doi.org/10.1016/j.trac.2017.05.004>
- D'Souza, R. N., Grimbs, S., Behrends, B., Bernaert, H., Ullrich, M., & Kuhnert, N. (2017). Origin-based polyphenolic fingerprinting of Theobroma cacao in unfermented and fermented beans. *Food Research International*, 99, 550–559. <https://doi.org/10.1016/j.foodres.2017.06.007>
- Fang, N., Yu, S., & Badger, T. M. (2003). LC-MS/MS Analysis of lysophospholipids associated with soy protein isolate. *Journal of Agricultural and Food Chemistry*, 51, 6676–6682. <https://doi.org/10.1021/jf034793v>
- Fernández-Fernández, R., López-Martínez, J. C., Romero-González, R., Martínez-Vidal, J. L., Alarcón-Flores, M. I., & Garrido-Frenich, A. (2010). Simple LC-MS determination of citric and malic acids in fruits and vegetables. *Chromatographia*, 72, 55–62. <https://doi.org/10.1365/s10337-010-1611-0>
- Germann, D., Stark, T. D., & Hofmann, T. (2019a). Formation and characterization of polyphenol-derived red chromophores. Enhancing the color of processed cocoa powders: Part 1. *Journal of Agricultural and Food Chemistry*, 67, 4632–4642. <https://doi.org/10.1021/acs.jafc.9b01049>
- Germann, D., Stark, T. D., & Hofmann, T. (2019b). Formation and characterization of polyphenol-derived red chromophores. Enhancing the color of processed cocoa powders: Part 2. *Journal of Agricultural and Food Chemistry*, 67, 4643–4651. <https://doi.org/10.1021/acs.jafc.9b01050>
- Gil de la Fuente, A., Godzien, J., Fernández-López, M., Rupérez, F. J., Barbas, C., & Otero, A. (2018). Knowledge-based metabolite annotation tool: CEU mass mediator. *Journal of Pharmaceutical and Biomedical Analysis*, 154, 138–149. <https://doi.org/10.1016/j.jpba.2018.02.046>
- Gómez-Romero, M., Segura-Carretero, A., & Fernández-Gutiérrez, A. (2010). Metabolite profiling and quantification of phenolic compounds in methanol extracts of tomato fruit. *Phytochemistry*, 71, 1848–1864. <https://doi.org/10.1016/j.phytochem.2010.08.002>
- Greño, M., Plaza, M., Marina, M. L., & Castro-Puyana, M. (2022). Untargeted HPLC-MS-based metabolomic approach to reveal cocoa powder adulterations. *Food Chemistry* (under first revision).
- Gu, L., House, S. E., Wu, X., Ou, B., & Prior, R. L. (2006). Procyanidin and catechin contents and antioxidant capacity of cocoa and chocolate products. *Journal of Agricultural and Food Chemistry*, 54, 4057–4061. <https://doi.org/10.1021/jf060360r>
- Gültekin-Özguven, M., Berktaş, I., & Özçelik, B. (2016). Change in stability of procyanidins, antioxidant capacity and in-vitro bioaccessibility during processing of cocoa powder from cocoa beans. *LWT - Food Science and Technology*, 72, 559–565. <https://doi.org/10.1016/j.lwt.2016.04.065>
- Hurst, W. J., Krake, S. H., Bergmeier, S. C., Payne, M. J., Miller, K. B., & Stuart, D. A. (2011). Impact of fermentation, drying, roasting and Dutch processing on flavan-3-ol stereochemistry in cacao beans and cocoa ingredients. *Chemistry Central Journal*, 5, 53. <https://doi.org/10.1186/1752-153X-5-53>
- Jiang, C., Arthur, C. J., & Gates, P. J. (2020). A computational and experimental study of the fragmentation of L-leucine, L-isoleucine and L-allo-isoleucine under collision-induced dissociation tandem mass spectrometry. *The Analyst*, 145, 6632–6638. <https://doi.org/10.1039/D0AN00778A>
- Jolić, M. A., Redovniković, I. R., Marković, K., Šipušić, D. I., & Delonga, K. (2011). Changes of phenolic compounds and antioxidant capacity in cocoa beans processing. *International Journal of Food Science & Technology*, 46, 1793–1800. <https://doi.org/10.1111/j.1365-2621.2011.02670.x>
- Kofink, M., Papagiannopoulos, M., & Galensa, R. (2007). (-)-Catechin in cocoa and chocolate: Occurrence and analysis of an atypical flavan-3-ol enantiomer. *Molecules*, 12, 1274–1288. <https://doi.org/10.3390/12071274>
- Kokotou, M. G., Mantzourani, C., Bourboulia, A., Mountanea, O. G., & Kokotos, G. (2020). A liquid-chromatography-high resolution mass spectrometry (LC-HRMS) method for the determination of free hydroxy fatty acids in cow and goat milk. *Molecules*, 25, 3947. <https://doi.org/10.3390/molecules25173947>
- Lavorgna, M., Pacifico, S., Nugnes, R., Russo, C., Orlo, E., Piccolella, S., & Isidori, M. (2021). Theobroma cacao Criollo var. Beans: Biological properties and chemical profile. *Foods*, 10, 571. <https://doi.org/10.3390/foods10030571>
- Li, Y., Feng, Y., Zhu, S., Luo, C., Ma, J., & Zhong, F. (2012). The effect of alkalization on the bioactive and flavor related components in commercial cocoa powder. *Journal of Food Composition and Analysis*, 25, 17–23. <https://doi.org/10.1016/j.jfca.2011.04.010>
- Li, Y., Zhu, S., Feng, Y., Xu, F., Ma, J., & Zhong, F. (2014). Influence of alkalization treatment on the color quality and the total phenolic and anthocyanin contents in cocoa powder. *Food Science and Biotechnology*, 23, 59–63. <https://doi.org/10.1007/s10068-014-0008-5>
- Martín, M. A., & Ramos, S. (2017). Health beneficial effects of cocoa phenolic compounds: Amini-review. *Current Opinion in Food Science*, 14, 20–25. <https://doi.org/10.1016/j.cofs.2016.12.002>
- Méndez-Albore, A., Campos-Aguilar, A. Z., Moreno-Martínez, E., & Vázquez-Durán, A. (2013). Physical and chemical degradation of B-aflatoxins during the roasting and dutching of cocoa liquor. *Journal of Agricultural Science and Technology A*, 15(3), 557–567.
- Miller, K. B., Hurst, W. J., Payne, M. J., Stuart, D. A., Apgar, J., Sweigart, D. S., & Ou, B. (2008). Impact of alkalization on the antioxidant and flavanol content of commercial cocoa powders. *Journal of Agricultural and Food Chemistry*, 56, 8527–8533. <https://doi.org/10.1021/jf801670p>
- Nikolić, D., Gödecke, T., Chen, S.-N., White, J., Lankin, D. C., Pauli, G. F., & van Breenen, R. B. (2012). Mass spectrometric dereplication of nitrogen-containing constituents of black cohosh (*Cimicifuga racemosa* L.). *Fitoterapia*, 83, 441–460. <https://doi.org/10.1016/j.fitote.2011.12.006>
- Patras, M. A., Milev, B. P., Vrancken, G., & Kuhnert, N. (2014). Identification of novel cocoa flavonoids from raw fermented cocoa beans by HPLC-MS<sup>n</sup>. *Food Research International*, 63, 353–359. <https://doi.org/10.1016/j.foodres.2014.05.031>
- Payne, M. J., Hurst, W. J., Miller, K. B., Rank, C., & Stuart, D. A. (2010). Impact of fermentation, drying, roasting, and Dutch processing on epicatechin and catechin content of cacao beans and cocoa ingredients. *Journal of Agricultural and Food Chemistry*, 58, 10518–10527. <https://doi.org/10.1021/jf102391q>
- Peixoto, J. A. B., Álvarez-Rivera, G., Alves, R. C., Costa, A. S. G., Machado, S., Cifuentes, A., Ibáñez, E., & Oliveira, M. B. P. (2021). Comprehensive phenolic and free amino acid analysis of rosemary infusions: Influence on the antioxidant potential. *Antioxidants*, 10, 500. <https://doi.org/10.3390/antiox10030500>
- Quelal-Vásquez, M. A., Lerma-García, M. J., Pérez-Estève, E., Arnau-Bonachera, A., Barat, J. M., & Talens, P. (2020). Changes in methylxanthines and flavanols during cocoa powder processing and their quantification by near-infrared spectroscopy. *LWT - Food Science and Technology*, 117, Article 108598. <https://doi.org/10.1016/j.lwt.2019.108598>
- Riggin, R. M., & Kissinger, P. T. (1976). Identification of salsolinol as a phenolic component in powdered cocoa and cocoa-based product. *Journal of Agricultural and Food Chemistry*, 24, 900. <https://doi.org/10.1021/jf060206a043>
- Rodríguez, P., Pérez, E., & Guzmán, R. (2009). Effect of types and concentrations of alkali on the color of cocoa liquor. *Journal of the Science of Food and Agriculture*, 89, 1186–1194. <https://doi.org/10.1002/jsfa.3573>
- Sánchez-Rabeneda, F., Jauregui, O., Casals, I., Andres-Lacueva, C., Izquierdo-Pulido, M., & Lamuela-Raventós, R. M. (2003). Liquid chromatographic/electrospray ionization tandem mass spectrometric study of the phenolic composition of cocoa (Theobroma cacao). *Journal of Mass Spectrometry*, 38, 35–42. <https://doi.org/10.1002/jms.395>
- Sioriki, E., Lemarcq, V., Alhakim, F., Triharyogi, H., Tuentner, E., Cazin, C. S. J., Nolan, S. P., Pieters, L., Van de Walle, D., & Dewettinck, K. (2021). Impact of alkalization conditions on the phytochemical content of cocoa powder and the aroma of cocoa drinks. *LWT - Food Science and Technology*, 145, Article 111181. <https://doi.org/10.1016/j.lwt.2021.111181>
- Sioriki, E., Tuentner, E., Van de Walle, D., Lemarcq, V., Cazin, C. S. J., Nolan, S. P., Pieters, L., & Dewettinck, K. (2022). The effect of cocoa alkalization on the non-volatile and volatile mood-enhancing compounds. *Food Chemistry*, 381, Article 132082. <https://doi.org/10.1016/j.foodchem.2022.132082>
- Stanley, T. H., Smithson, A. T., Neilson, A. P., Anantheswaran, R. C., & Lambert, J. D. (2015). Analysis of cocoa proanthocyanidins using reversed phase high-performance liquid chromatography and electrochemical detection: Application to studies on the effect of alkaline processing. *Journal of Agricultural and Food Chemistry*, 63, 5970–5975. <https://doi.org/10.1021/acs.jafc.5b02661>
- Stark, T., & Hofmann, T. (2006). Application of a molecular sensory science approach to alkalized cocoa (Theobroma cacao): Structure determination and sensory activity of nonenzymatically Glycosylated flavan-3-ols. *Journal of Agricultural and Food Chemistry*, 54, 9510–9521. <https://doi.org/10.1021/jf062403>
- Taş, N. G., & Gökmen, V. (2016). Effect of alkalization on the Maillard reaction products formed in cocoa during roasting. *Food Research International*, 89, 930–936. <https://doi.org/10.1016/j.foodres.2015.12.021>
- Todorovic, V., Milenkovic, M., Vidovic, B., Todorovic, Z., & Sobajic, S. (2017). Correlation between antimicrobial, antioxidant activity, and polyphenols of alkalized/nonalkalized cocoa powders. *Journal of Food Science*, 82(4), 1020–1027. <https://doi.org/10.1111/17503841.13672>
- Valverde-García, D., Pérez-Estève, E., & Barat-Baviera, J. M. (2020). Changes in cocoa properties induced by the alkalization process: A review. *Comprehensive Reviews in Food Science and Food Safety*, 19, 2200–2221. <https://doi.org/10.1111/1541-4337.12581>
- Valverde, D., Behrends, B., Pérez-Estève, E., Kuhnert, N., & Barat, J. M. (2020). Functional changes induced by extrusion during cocoa alkalization. *Food Research International*, 136, Article 109469. <https://doi.org/10.1016/j.foodres.2020.109469>
- Van Dycke, A., Verstraete, A., Pil, K., Raedt, R., Vonck, K., Boisson, D., & Boon, P. (2010). Quantitative analysis of adenosine using liquid chromatography/atmospheric pressure chemical ionization-tandem mass spectrometry (LC/APCI-MS/MS). *Journal of Chromatography B*, 878, 1493–1498. <https://doi.org/10.1016/j.jchromb.2010.03.047>
- Vargas-Arana, G., Merino-Zegarra, C., Tang, M., Pertino, M. W., & Simirgiotis, M. J. (2022). Characterization and antioxidant nutritional analysis of cocoa waste flours from the peruvian amazon. *Antioxidants*, 11, 595. <https://doi.org/10.3390/antiox11030595>
- Zhong, P., Wei, X., Li, X., Wei, X., Wu, S., Huang, W., Koides, A., Xu, Z., & Lei, H. (2022). Untargeted metabolomics by liquid chromatography-mass spectrometry for food authentication: A review. *Comprehensive Reviews in Food Science and Food Safety*. <https://doi.org/10.1111/1541-4337.12938>