Chemometric approaches for postharvest quality tracing of cocoa: An efficient method to distinguish plant material origin

Maritza Gil [a,*], Yamile Jaramillo [b], Carolina Bedoya [b], Sandra M. Llano [b], Vanessa Gallego [c], Jairo Quijano [d], Julian Londono-Londono [e]

[a] Universidad Nacional de Colombia. Facultad de Ciencias Agrarias. Medellín- Antioquia Colombia. Researcher Grupo de investigación de Ingeniería de Alimentos, GRIAL. Corporación Universitaria Lasallista, Caldas, Antioquia, Colombia
[b] Corporación Universitaria Lasallista. Researcher Grupo de investigación de Ingeniería de Alimentos, GRIAL. Corporación Universitaria Lasallista, Caldas, Antioquia, Colombia
[c] Universidad de Antioquia. Researcher Grupo de investigación de Ingeniería de Alimentos, GRIAL. Corporación Universitaria Lasallista, Caldas, Antioquia, Colombia
[d] Universidad Nacional de Colombia, Colombia
[e] Regional Director of the Colombian Corporation for Agricultural Research AGROSAVIA, Meta, Colombia

ARTICLE INFO

ABSTRACT

The aim of this study was to compare the quality of a mixture of cocoa harvested and fermented in three sub-regions of Antioquia (Colombia), from the chemometric profile based on multivariate statistical analysis. A mixture of clones CCN-52, ICS-1, FLE-2, and FEC-2 harvested in Bajo Cauca, Uraba and Magdalena Medio were subjected to a spontaneous fermentation. The characterization of raw and well-fermented cocoa was performed through 38 parameters, and results were compared by a Principal Component Analysis (PCA) and a Cluster Analysis (CA), followed by a Principal Factors Analysis (PFA- CA). The CA showed that there are differences among subregions only in raw cocoa from Bajo Cauca. PCA allowed identifying the variability between raw and fermented cocoa in a representative way and these results were consistent with the chemical profile. Besides, the number of parameters to differentiate raw cocoa from different subregions was reduced (11–13 parameters) and it was possible to characterize well fermented cocoa with only 10 parameters of 38. PFA-CA consolidated in three factors a grouping to identify the cocoa quality according to the process or interest of the sensory or functional properties. Factor 1 (cocoa quality indicators with functional properties), Factor 2 (indicators of quality of the beginning of fermentation) and Factor 3 (indicators of quality of well-fermented cocoa) each one with a weight of 39, 35 and 26 respectively.

1. Introduction

There are numerous studies characterizing both raw and fermented cocoa (*Theobroma cacao* L.), in which there is not a consensus in the procedural form, which hinders the standardization or normalization of the significant methods allowing the implementation of a routine for quality control and decision making for the distinction of the wide range of plant materials and traceability of aroma precursor compounds formed during fermentation (Hii et al., 2017; Pereira et al., 2016).

The stages of post-harvest, mainly fermentation and drying, are crucial as initiating phases of precursors and compounds responsible for the sensory characteristics of cocoa as a raw material to be used as an input in the subsequent industrialized stages. These two stages are the topics which more research studies have focused on and which represent a starting point for obtaining current cocoas with high quality.

Previous studies have reported a large number of attributes and defects in the sensory profile-related compounds that are indicators of the correct procedure, either during fermentation or drying (Rodriguez-Campos et al., 2012).

The formation of compounds identified as precursors during fermentation have been explained by means of a series of biochemical, enzymatic and microbiological reactions that lead to the formation of certain key metabolites, not only for aroma but also for taste, some of them with functional properties (Illeghems et al., 2015; Ohene et al., 2011). Most of these metabolites contribute to the overall sensory profile for differentiating quality, and can also be characteristic of special cocoas

---

* Corresponding author.
   E-mail address: magil@unal.edu.co (M. Gil).

(Bertoldi et al., 2016; Krähmer et al., 2015).

In addition to amino acids and sugars, mainly fructose and glucose, organic acids, alkaloids, monomers of procyanidins, anthocyanins, peptides and other macromolecules have been intently studied, in order to deepen understanding of the behavior within the cotyledon of cocoa and keep track of how it will impact on quality at the end of fermentation (Afoakwa et al., 2012; Ortega et al., 2010).

The determination of the chemical composition for traceability poses some challenges in terms of methodologies to follow, which must be supported by robust technologies for the determination of low concentrations, differentiated methods according to the matrix of the cocoa under study (pulp, cotyledon), and according to the moisture level due to the fermentation progress (Bertoldi et al., 2016; Ortega et al., 2010).

The results of this complex characterization can support the unequivocal decision making when performing the analysis of such results for each of the methods applied, supported by the statistical analysis of chemometric substantiation. This will help define the significant compounds for decision making on the profile of aroma and flavor precursors formed after the fermentation of cocoa. Among the most supportive descriptive analysis techniques, the Principal Component Analysis (PCA) reduces the dimension of the data according to its similarities and differences without losing information based on the linear combinations of variables, which are defined as main components (Kaiser, 1961).

Likewise, the Cluster Analysis (CA) can be useful to the grouping into homogeneous classes that allow an identification between the stages of fermentation (beginning and end) or different origins, etc. In addition, the Principal Factor Analysis (PFA) from PCA is part of the methods for sorting variables with a high correlation structure that leads to the determination of few underlying factors. Each one of the factors supported on the knowledge of the information users will enable a classification with different scopes of interest, which for the data collected during fermentation could explain productive aspects of quality and functionality, among others (Wold, 1995).

Therefore, this study proposes as hypothesis the chemometric approximation of a mixture of cocoas grown in three subregions of Antioquia, based on the characterization of the physical, chemical and bromatologic profiles, which involved the standardization of methods. That would allow obtaining reliable information about the characteristics of the materials assessed according to the area where the harvest was obtained, as well as determining quality at the end of fermentation but with the benefit of identifying the constrained and conclusive parameters

in order to reduce the number of tests and thus obtain a descriptive and inferential characterization using statistical tools.

In order to prove the hypothesis, the evaluation of three subregions of the same producer department became important because having agro-ecological differences it was expected that their quality would be different. Additionally, with the collection of a wide range of information from techniques currently used to evaluate the quality of cocoa beans, it was expected to reduce the information reliably with multivariate statistical studies, and thus be able to traceability mainly among raw and well-fermented cocoas during this postharvest stage.

2. Material and methods

Fig. 1 describes the most relevant aspects of the stages that comprise the methodological scheme that is detailed below.

2.1. Chemicals

The following reagents classified as acids were used: acetic US Pharmacopeia (Convention-Rockville, USA), citric, lactic, and malic which were purchased from Supelco (North Harrison, USA); oxalic and succinic acids obtained from Dr. Ehrenstorfer (Augsburg, Germany). Saccharides: fructose from Chromadex (Irvine, CA, USA); glucose and sucrose purchased from Supelco (North Harrison, USA); maltose and polyol mannitol European Pharmacopoeia Reference Standard (Augsburg, Germany). Alkaloids: caffeine and theobromine from Sigma Aldrich (St. Louis, USA). Procyanidins monomers: epicatechin and catechin were obtained from Chromadex (California, USA). Amino acids: asparagine, cysteine, phenylalanine, leucine, lysine, methionine, and valine were commercially available from Sigma Aldrich (St. Louis, USA).

The reagents and analytical HPLC or mass grade solvents used for the extraction and quantification of the studied compounds were: acetone, ortho-phosphoric acid, sulfuric and hydrochloric acid, hexane, chloroform, diethyl ether, petroleum ether, Tashiro's indicator and phenolphthalein, sodium dibasic phosphate dodecahydrate, hexane, sodium hydroxide, methanol, n-octanol, copper sulfate, potassium-acid phthalate purchased from Merck (Darmstadt, Germany). 6-hydroxy-2,5,7-tetramethylchroman-2- carboxylic acid (Trolox), fluorescein, 2,2′-azobis(2-amidinopropane) dihydrochloride (AAPH), formic acid, and sodium formate were obtained from Sigma Chemical Co (Missouri, USA).

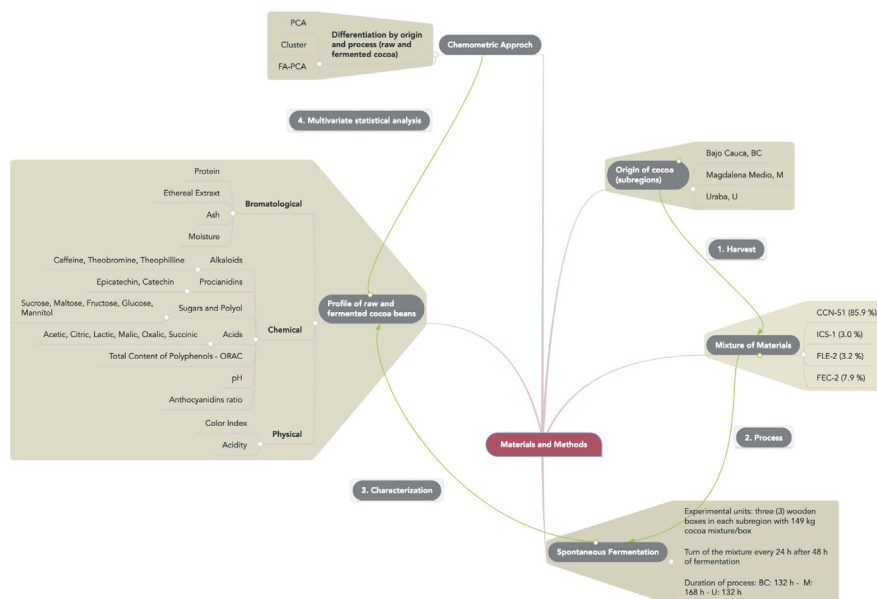The ultrapure water used in most of the tests showed a resistivity of

Fig. 1. Schematic representation of the materials and methodology used to carry out the chemometric approach.

18.2 MΩ cm, Purelab Ultra ELGA (Celle, Germany).

2.2. Plant material

The plant material used in the present study consisted of a mixture of two international clones: CCN-51 (Collection Castro Naranjal), ICS-1 (Imperial College Selection), and two national clones: FEC-2 (Fedecacao El Carmen, Santander) and FLE -2 (Fedecacao Lebrija, Santander), provided by the National Cacao Federation of Colombia - Fedecacao. The four clones were obtained from the production of the secondary harvest of 2017 (March–May) in three subregions of the department of Antioquia, Colombia. Each clone was harvested from different farms classified by Fedecacao as productive material that met the standards of good agricultural practices. The farms were located in the principal cocoa producing municipalities: Bajo Cauca (2 municipalities), Uraba (4 municipalities) and Magdalena Medio (7 municipalities). The sampling was random and representative, so that the amount required for fermentation and analysis was obtained.

2.3. Study location

The fermentation study was carried out in three of the main cocoa producing subregions in the department of Antioquia, Colombia. The experimental units were located as follows: Bajo Cauca (BC): La Candelaria farm located on the road that leads from Caucasia to Nechí N 08° 40″6.57'; W 75° 10″9.74',76.3 above sea level. Urabá (U): Centro de Desarrollo Agrario y Sede de Estudio of the Universidad de

Antioquia para la Región de Urabá - Tulenapa farm of the municipality of Carepa at the geographic coordinates N 07° 46″4.51'; W 76° 39″8.90'; 14.7 m above sea level. Magdalena Medio (M): it was carried out in the Cannes farm located in the Betulia village of Maceo municipality N 06° 31″6.52 '; W 76° 49″4.89'; 1048 m above sea level.

2.4. Fermentation

The opening of the cobs was carried out 36 h after harvest. The cocoa in slime had a draining time of 4–6 h before weighing in order to obtain the mixture of the clones in the following ratios: CCN-51 (85.9%), ICS-1 (3.0%), FLE-2 (3.2%) and FEC-2 (7.9%). The ratios were based on the productive capacity of the clones in the subregions studied. The mixture of cocoa in slime was added in three wooden drawers, considered as an experimental unit arranged in each study location. The shedding of each clone to the drawer was carried out in the same order in which it was shelled and weighed, as mentioned before. The fermenting drawers were built in the three subregions keeping the same design, which consisted of nine oak wood drawers located in the form of a staircase comprising three rows with three drawers in each one. The mixture was arranged only in the central row of three drawers, separated by removable divisions of 2.5 cm. After homogenizing the mass of cocoa in each experimental unit by means of a manual mix with an oak wood stirrer assigned to each drawer, it was covered with Bijao leaves (*Calathea lutea*) and finally with a gunnysack. The dimensions of each drawer were: width 0.65 m, length 0.70 m, and height 0.65 m with a maximum capacity of 153 kg and 149 kg of filling. In order to facilitate agitation, perforations of 0.8 cm were made every 15 cm at the bottom so as to favor the draining of the mucilage during fermentation.

The turning was done using the same oak wood stirrers used from the beginning. The cocoa in slime was left to rest for 48 h to propitiate the micro-anaerobic conditions at the beginning. After this time, the turning began and it was repeated every 24 h, until the producer in each study zone indicated that it had reached fermentation conditions, those being 132 h in Bajo Cauca and Uraba and 168 h in Magdalena Medio. Every 24 h, the sampling (approximately 100 grains) was selected randomly in three heights of the drawer (low, medium, and high). The mucilage was removed from each grain and packed separately from the cotyledon in a conical tube, and was then frozen at -20 °C until analysis.

2.5. Sample treatment

The non-mucilage and frozen cocoa beans were subjected to size reduction in a mill (Model No. 80370- Hamilton Beach, USA) and the particle size was selected between 594 and 420 μm after passing the ground cocoa through sieve N° 40. The grounded cocoa was stored as follows: one gram was placed in a 20 mL gas chromatography vial and hermetically sealed with Teflon-coated white septa lid (Thermo Fisher, Massachusetts, USA) and stored at refrigeration temperature (between 8 and 12 °C) until chromatographic analysis. The remaining sample was placed in a high density polypropylene container with a screw cap and in turn, this package was kept inside a plastic box containing silica gel to avoid moisture absorption of the sample during its resting at -20 °C.

The analysis carried out for the characterization of the raw (R) and well-fermented (WF) grains are based on international standards (AOAC) and on standardized and validated methods, which in total comprise 38 analyses. This information is describe in Table 1.

2.6. Chemometric analysis

The statistical analysis supported by chemometrics was based on the analysis of the profile of taste and aroma precursors that are formed after fermentation from the formation of sugars and amino acids and some parameters that can affect the production of these compounds (acidity, pH, bromatological profile, anthocyanins, alkaloids, procyanidins, and color). The multivariate analyses are described below.

2.6.1. Principal Component Analysis
The Principal Component Analysis was used to reduce the size of the data set based on their similarities and differences by transforming the original variables into a new set of uncorrelated variables without losing information based on linear combinations that maximize variability (Kaiser, 1961).

The PCA started with the analysis of a matrix (18 × 38) consisting of 18 samples (rows) composed of the mixtures of clones: a) raw, evaluated in each zone on day zero of fermentation, coded as 0CF-BC, 0CF-U and 0CF-M; and b) well-fermented, identified by the number of hours fermentation took in each subregion, as follows: 132CF-BC, 132CF-U, and 168CF-M. The analyses were performed in triplicate. The columns (38) formed by the values of the evaluated variables gather the tests described in section 2.5.

2.6.2. Cluster analysis
The clusters were used in the data to contrast similarities or differences between individuals, which in this case were the samples before and after fermentation, thus eliminating redundant information from the set of observations and favoring simpler structures to analyze patterns of behavior, if any (Gentleman et al., 2011).

2.6.3. Principal Factor Analysis with cluster analysis
From the method of PCA previously determined and by the Kaiser - Varimax Method transformation criterion, the analysis of common factors for raw and fermented cocoa was carried out. The factors allowed proposing axes with few large loads and as close to zero as possible for the others, which is achieved by an iterative maximization of the quadratic function of the loads.

For the study of the multivariate analysis, statistical software R was used (Version 0.98.1103 GNU Affero General Public License; https://cran.r-project.org).

3. Results and discussion

The results are displayed in two sections which comprise the description of the profile of the cocoa characteristic components and its respective chemometric analysis.

Table 1
General aspects of the methods applied for the bromatological, chemical and physical characterization.

Analysis	Parameter	Reference /Standard	Method	Optimization	Experimental Design
Fermentation monitoring	Fermentation index (totals anthocyanins)	Gourieva and Tserrevitinov (1979)	Spectrophotometric	Standardization (R^2, residuals <20 %)	N/A
	Color	Vignoni et al. (2006)	Colorimeter	N/A	N/A
Bromatological	Ashes	AOAC 972.15	Gravimetric	N/A	N/A
	Crude fiber	AOAC 930.20	Acid and basic digestion	N/A	N/A
	Moisture	AOAC 931.04	Gravimetric –Constant weight	N/A	N/A
	Protein	AOAC 970.22	Kjeldahl	N/A	N/A
	Ethereal extract by Soxhlet	AOAC 925.07	Soxhlet	Selection of solvent Cycles and time of extractions	Analysis of variance – one way (Factor: solvents) 6 levels
	Ethereal extract assisted by ultrasound	Carrillo et al. (2014)	Ultrasound	N/A	
Physical and chemical analysis	Total acidity and pH	AOAC 970.21 Nazzaduri et al. (2006), Guehi et al. (2010)	Potentiometric	Solids separation by centrifugation	
	Total polyphenols content	Singleton & Rossi, (1965)	Spectrofluorometric	Standardization (residuals <20 %)	
	Oxygen radical scavenging capacity ORAC	Ou et al. (2001)	Spectrofluorometric	Standardization (R^2, residuals <20 %)	
Chromatographic Analysis	**Sugars** Fructose, glucose, sucrose, and maltose **Polyol** Mannitol	Ibañez et al. (2007) Gil (2018)	SPE UHPL-C-CAD	Validation (SANTE, 2015)	Precision intermedia (T test)
	Non-volatiles organics acids Citric, malic, oxalic, and succinic acids	Doores et al. (1993)	SPE UHPLC-C-CAD	Standardization (% recovery, R^2, and residuals <20 %)	
	Acetic and lactic acid	Gil (2018)	SPE UHPLC-DAD	Standardization (% recovery, R^2, and residuals <20 %)	
Analysis	Parameter Acetic acid	Reference/Standard Gil (2018)	Method HS-SPME GC-MS/MS-O	Optimization Standardization of time and temperature of extraction (R^2, residuals <20 %)	Experimental Design
	Aminoacids Phenylalanine, leucine, methionine valine, cysteine, asparagine, and lysine	Jürgen Voigt et al. (1994)	SPE UHPLC-C-CAD	Standardization (% recovery, R^2, and residuals <20 %)	
	Procyanidin monomers Cathequine Epicathequine	Ortega et al. (2010)	UHPLC-MS/MS -QqQ/ESI	Standardization (% recovery, R^2, and residuals <20 %)	Factorial design two-way with 3 levels (extraction temperature, and frequency)
	Alkaloids Caffeine Teophylline Theobromine	Ortega et al. (2010)	UHPLC-MS/MS -QqQ/ESI	Standardization (% recovery, R^2, and residuals <20 %)	Factorial design 2-way with 3 levels (extraction temperature and frequency)

N/A – not available.

3.1. Bromatological, physical and chemical profile of raw and fermented cocoa

Cocoa bean quality parameters are listed in Table 2, in which the results of the analysis are: bromatological, physical and chemical characterstics obtained from the mixture of clones, before fermentation as well as at the end of this stage. Table 3 shows the parameters determined for the standardization of the chromatographic methods with which the described results were obtained. The calculated parameters were: the equation that describes the linear relationship between the response of intensity with the concentration of each external standard prepared in solvent and cocoa matrix obtained from the different stages of post-harvest (fermented or dried) to correct this effect during the chromatographic quantification. Other verified aspects were the recovery, as a measure of assurance of the effectiveness of the method of analyte extraction from samples of cocoa, and finally, quantification ranges that defines the minimum limit of detection for each compound in the matrix.

Due to the number of evaluated parameters, multivariate statistical analysis focused on the main components, followed by an Analysis of Cluster and Principal Factor Analysis, as described below.

3.2. Multivariate statistical analysis

Thirty-eight evaluated parameters to the cocoa comprise the majority of tests defined as minimum requirements of the grain of dry and fermented cocoa by the regulation of the producing countries (Comisión venezolana de Normas Industriales - COVENIN, 1995; Dirección General de Normas, 1980; ICONTEC Internacional, 2003). The scopes of the analysis is part of the compilation of the aspects considered in several previous studies on the requirements identified for traceability of the

Table 2

Bromatological, physical and chemical profile of the mixture of cocoa clones (CCN-51, ICS-1, FLE-2, and FEC-2) before starting fermentation (or time zero) in Bajo Cauca (0-BC), Uraba (0-U), and Magdalena Medio (0-M) and after fermentation at 132 h in Bajo Cauca (132-BC) and Uraba (132-U), and at 168 h in Magdalena Medio (168-M).

Analysis	0-BC	132-BC	0-U	132-U	0-M	168-M
Acidity (% acetic acid w/w)	0.16 ± 0.06	1.56 ± 0.25	0.35 ± 0.04	0.85 ± 0.37	0.30 ± 0.03	0.76 ± 0.02
Acidity (% lactic acid w/w)	3.72 ± 0.91	1.58 ± 0.15	0.52 ± 0.06	1.27 ± 0.55	0.45 ± 0.04	1.13 ± 0.03
Acidity (% citric acid w/w)	2.65 ± 0.65	1.12 ± 0.10	0.37 ± 0.04	0.91 ± 0.39	0.32 ± 0.03	0.81 ± 0.02
Acidity (% malic acid w/w)	2.77 ± 0.68	1.18 ± 0.11	0.39 ± 0.05	0.95 ± 0.41	0.34 ± 0.03	0.85 ± 0.02
Acidity (% oxalic acid w/w)	1.86 ± 0.46	0.79 ± 0.08	0.26 ± 0.03	0.64 ± 0.28	0.23 ± 0.02	0.57 ± 0.01
Acidity (% succinic acid w/w)	4.88 ± 1.20	2.07 ± 0.19	0.68 ± 0.08	1.67 ± 0.73	0.60 ± 0.05	1.49 ± 0.03
pH	6.27 ± 0.15	4.46 ± 0.08	6.22 ± 0.23	5.1 ± 0.11	6.23 ± 0.15	5.23 ± 0.25
Moisture (kg water/kg wet cocoa)	0.66 ± 0.02	0.38 ± 0.01	0.59 ± 0.01	0.41 ± 0.02	0.66 ± 0.02	0.42 ± 0.01
Ashes (%)	1.8 ± 0.4	1.1 ± 0.5	2.3 ± 0.1	1.2 ± 0.5	2.0 ± 0.00	1.4 ± 0.1
Total crude fiber (%)	7.2 ± 0.3	11.8 ± 0.2	8.9 ± 0.2	13.3 ± 1.3	9.9 ± 2.7	11.1 ± 1.4
Protein (%)	14.01 ± 0.27	10.23 ± 0.11	12.7 ± 0.3	10.34 ± 0.06	10.24 ± 0.32	8.19 ± 0.75
Ethereal extract (%)	49.0 ± 1.8	45.6 ± 1.0	46.6 ± 4.2	44.7 ± 4.3	47.3 ± 4.0	52.1 ± 1.4
Anthocyanins (Ratio 430/530 nm)	2.12 ± 0.15	0.73 ± 0.02	1.84 ± 0.30	0.59 ± 0.02	2.63 ± 0.12	0.84 ± 0.04
Color	21.95 ± 1.04	19.62 ± 2.10	16.75 ± 0.89	32.31 ± 1.70	20.90 ± 2.47	55.71 ± 9.38
Totals Polyphenols (Gallic acid equiv./100 g)	16843.67 ± 2087.73	13973.03 ± 4225.89	58185.23 ± 7710.54	16289.87 ± 6645.60	31162.33 ± 2085.16	22663.60 ± 4111.51
ORAC (Trolox equiv./100g)	122460.80 ± 13945.60	95308.53 ± 10785.88	299486.33 ± 19243.35	40616.43 ± 52723.93	160941.87 ± 675.59	129193.47 ± 8866.02
Acetic acid (% w/w) HS-SPME-GC	0.37 ± 0.30	2.11 ± 1.16	0.42 ± 0.05	0.78 ± 0.16	0.04 ± 0.01	1.97 ± 1.12
Lactic acid (% w/w)	0.90 ± 0.02	0.65 ± 0.03	0.82 ± 0.11	0.22 ± 0.01	0.87 ± 0.10	0.78 ± 0.20
Citric acid (% w/w)	0.28 ± 0.03	0.09 ± 0.02	0.63 ± 0.10	0.25 ± 0.07	0.53 ± 0.05	0.18 ± 0.10
Oxalic acid (% w/w)	4.05 ± 0.27	2.20 ± 0.08	3.36 ± 0.29	1.41 ± 0.02	3.68 ± 0.43	2.35 ± 0.17
Succinic acid (% w/w)	12.92 ± 0.88	7.01 ± 0.27	10.92 ± 0.97	4.34 ± 0.06	11.99 ± 1.44	7.49 ± 0.58
Fructose (% w/w)	2.12 ± 0.49	0.48 ± 0.09	0.48 ± 0.36	0.64 ± 0.01	0.64 ± 0.04	0.24 ± 0.03
Glucose (% w/w)	2.09 ± 0.46	0.16 ± 0.03	0.30 ± 0.51	0.79 ± 0.03	0.06 ± 0.02	0.25 ± 0.03
Sucrose (% w/w)	4.74 ± 2.19	0.15 ± 0.12	6.79 ± 0.65	0.27 ± 0.07	3.22 ± 0.38	0.02 ± 0.00
Maltose (% w/w)	n.d	0.11 ± 0.02	n.d.	0.09 ± 0.03	n.d.	0.02 ± 0.01
Mannitol (% w/w)	n.d	0.02 ± 0.02	n.d.	0.24 ± 0.22	n.d.	0.00 ± 0.00
Phenylalanine (% w/w)	n.d	0.08 ± 0.01	n.d.	0.00 ± 0.00	n.d.	0.04 ± 0.00
Leucine (% w/w)	0.03 ± 0.03	0.05 ± 0.01	n.d.	0.11 ± 0.02	n.d.	0.05 ± 0.00
Methionine (% w/w)	0.02 ± 0.01	0.02 ± 0.00	0.01 ± 0.00	0.04 ± 0.01	0.02 ± 0.03	0.01 ± 0.00
Valine (% w/w)	0.08 ± 0.01	0.09 ± 0.02	0.10 ± 0.01	0.14 ± 0.03	0.11 ± 0.06	0.06 ± 0.00
Cysteine (% w/w)	n.d.	0.08 ± 0.03	n.d.	0.17 ± 0.01	0.03 ± 0.05	0.06 ± 0.02
Asparagine (% w/w)	n.d.	0.04 ± 0.02	n.d.	0.08 ± 0.03	n.d.	n.d.
Lysine (% w/w)	0.01 ± 0.00	0.10 ± 0.01	0.02 ± 0.00	0.18 ± 0.01	0.00 ± 0.00	0.08 ± 0.00
Theobromine (% w/w)	0.65 ± 0.10	0.12 ± 0.02	0.18 ± 0.004	0.10 ± 0.01	0.17 ± 0.03	0.11 ± 0.01
Caffeine (% w/w)	0.24 ± 0.03	0.05 ± 0.01	0.08 ± 0.003	0.05 ± 0.001	0.07 ± 0.02	0.05 ± 0.01
Ratio theobromine/caffeine (% w/w)	2.69 ± 0.16	2.52 ± 0.21	2.34 ± 0.10	2.30 ± 0.18	2.33 ± 0.13	2.29 ± 0.14
Epicathequine (% w/w)	0.25 ± 0.01	0.06 ± 0.003	0.07 ± 0.003	0.06 ± 0.009	0.08 ± 0.002	0.05 ± 0.000
Cathequine (% w/w)	0.023 ± 0.01	0.00 ± 0.00	n.d.	n.d.	n.d.	n.d.

% w/w: percentage of mg of the ratio between the mg of the analyte evaluated per one mg of sample on a dry basis.

n.d.: not detected, according to the limit of quantification of the methods used for the analysis.

cocoa quality, from its harvest to the final post-harvest stage, framed on requirements of national and international markets of the companies responsible for their industrialization (Araujo et al., 2014; Badrie et al., 2015; Kongor et al., 2016; Wollgast and Anklam, 2000). For this reason, each of the analyses included criteria such as: contribution in flavor (alkaloids and procyanidins monomers), aroma (sugars and free amino acids), indicator of the fermentation progress (anthocyanins, protein, acidity, and pH), functional properties (antioxidant capacity and total polyphenols) and commercial (ether extract content).

The analysis conducted on the multivariate statistical treatment allowed the extraction and description of major independent data gradients as described below. In terms of the distribution of variance, it allowed the identification of redundancy, that is to say, the amount of information shared between the results of the bromatological, physical and chemical analysis for each sample, both at the beginning and at the end of fermentation in the three subregions.

3.3. Principal components analysis

The distribution of the variances of the resulting components of the three first principal components explained 79.6% of the maximum variation in all the analyzed sample observations. Additionally, it is noted that the two first principal components explained approximately 68.2% of the total variation, with which the analysis was carried out descriptively. The contributions of the parameters conducive to variability in the first two principal components can be seen graphically from the biplot and the circle of correlations in the foreground factorial presented in Fig. 2.

The variables that are grouped with the first and the second principal component are the most important to explain variability in the data set. The variables that are not correlated to any of the principal components, or that are associated with the other components, presented a low contribution.

The compounds with the greatest contribution to the first principal component in the positive X-axis (right upper quadrant) correspond to the percentage of moisture, crude fiber, oxalic and succinic acid, lysine, sucrose, theobromine, caffeine, epicatechin, and anthocyanins. In the negative axis (left quadrant), the compounds with greater contribution were the total acidity, expressed as acetic acid, leucine, lysine, cysteine, maltose, and crude fiber.

Table 3
Equation of the calibration curve, limit of quantification and percentage of recovery of the methods used in the characterization of the chemical profile of unfermented and fermented cocoa in Bajo Cauca, Uraba, and Magdalena Medio.

Analyte	Instrument	Solvent for the preparation of the calibration curve	Intercept	Slope	R^2	Curve concentration range (mg/L)	Recovery percentage
Fructose	UHPLC-C-CAD	Solvent: CH_3CN	0.1261	0.0452	0.998	01–30.0	N/A
		Matrix (raw fermented cocoa)	-0.0006	0.0424	0.998	0.2–30.0	Low level: 117.0
							Intermediate level: 112.2
							High level: 108.4
		Matrix (dry cocoa)	0.0138	0.0431	0.996	0.4–30.0	Low level: 113.9
							Intermediate level: 111.8
							High level: 117.7
Glucose	UHPLC-C-CAD	Solvent: CH_3CN	0.0964	0.0389	0.999	0.1–30.0	N/A
		Matrix (raw fermented cocoa)	0.015	0.0424	0.997	0.5–30	Low level: 87.3
							Intermediate level: 77.8
							High level: 103.4
		Matrix (dry cocoa)	0.0101	0.0432	0.997	0.6–30	Low level: 104.0
							Intermediate level: 79.6
							High level: 95.6
Sucrose	UHPLC-C-CAD	Solvent: CH_3CN	0.1516	0.0534	0.998	0.1–30.0	N/A
		Matrix (raw fermented cocoa)	0.0111	0.0457	0.998	0.4–30.0	Low level: 120.2
							Intermediate level: 110.2
							High level: 116.4
		Matrix (dry cocoa)	0.0288	0.0437	0.997	0.4–30.0	Low level: 84.8
							Intermediate level: 111.2
							High level: 111.1
Maltose	UHPLC-C-CAD	Solvent: CH_3CN	0.0444	0.0373	0.999	0.1–50	N/A
		Matrix (raw fermented cocoa)	0.0188	0.0314	0.996	5.0–50.0	Low level: 112.8
							Intermediate level: 102.0
							High level: 118.1
		Matrix (dry cocoa)	-0.0396	0.0347	0.998	5.0–50.0	Low level: 113.2
							Intermediate level: 91.0
							High level: 119.7
Mannitol	UHPLC-C-CAD	Solvent: CH_3CN	0.0888	0.0515	0.999	0.1–30.0	N/A
		Matrix (raw fermented cocoa)	-0.0161	0.0473	1.000	0.4–30	Low level: 106.4
							Intermediate level: 84.7
							High level: 90.6
		Matrix (dry cocoa)	-0.0146	0.0491	0.998	0.3–30	Low level: 110.9
							Intermediate level: 104.6
							High level: 110.0
Phenylalanine	UHPLC-C-CAD	Matrix (well-fermented and dry cocoa)	-1.0253	0.0761	0.999	1.0–60	68
Leucine			0.1257	0.0627	0.998	1.0–60	96.2
Methionine			-0.2386	0.1547	0.996	1.0–60	<70
Valine			-0.7939	0.1069	0.998	1.0–60	75.2
Cysteine			0.1923	0.0631	0.998	1.0–60	116.3
Asparagine			-0.0707	0.077	0.999	1.0–60	114.6
Lysine			-0.4354	0.1158	0.999	1.0–60	79.9
Citric			0.105	0.0382	0.996	0.1–70	117.9
Malic			0.0897	0.0428	0.997	0.1–70	79.2
Oxalic			-0.015	0.0405	0.995	0.1–70	>120
Succinic			0.0462	0.0121	0.999	0.1–70	
Acetic acid	HPLC-DAD	Matrix (well-fermented and dry cocoa)	-0.0034	0.0161	0.992	3.0–20.0	85
	GC-MS	Matrix (well-fermented and dry cocoa)	215916.7125	1341.7457	0.999	0.5–100	72
Lactic acid	HPLC-DAD	Matrix (well-fermented and dry cocoa)	-0.0253	0.0145	0.995	3.0–20.0	86
Theobromine	UHPLC-TQD/ESI	Matrix (well-fermented and dry cocoa)	425.072	1945.29	0.999	2.5–100	103.87
Caffeine			-2541.82	13804.2	0.999	0.02–50	106.27
Theophylline			-970.582	10170.6	0.999	0.02–50	>120
Epicathequine			-468.281	1546.46	1.000	0.02–50	
Cathequine			-209.54	198.18	0.995	0.02–50	
Total polyphenols	Spectrofluorometer	Solvent	-0.0044	0.0037	0.999	10–100	N/A
ORAC	Spectrofluorometer	Solvent	2.8258	0.0838	0.994	10–600	N/A

The compounds with the greatest contribution to the second main component correspond to the acidity expressed in terms of lactic, oxalic, citric, succinic acid concentration expressed in % w/w of citric acid, glucose, protein, total polyphenols content and ORAC. These compounds in particular have the greatest contribution in the positive X-axis of the second component. In the negative axis (right lower quadrant), the compounds of greatest contribution were: acid oxalic, succinic, citric, anthocyanins, ashes, sucrose, total polyphenols content, and ORAC.

The variables in PCA 1 and PCA 2 allowed distinguishing between unfermented and well-fermented cocoa, as evidenced in the variables representing the positive X-axis of PCA 1. In addition, the representation quality of the sample observations could be analyzed using the square cosine as a measure, which, for high values, gives indications of a good

representation of the variable in the main component. Fig. 3 shows the contribution of each sample observation of the cocoas at the beginning and end of fermentation in each subregion It can be seen that before starting the fermentation process (1, 2 and 3) the cotyledons from Bajo Cauca have the greatest contribution to the first two main components, characterized by having high levels of total acidity expressed in lactic acid, malic, oxalic, succinic, and citric; and high levels of epicatechin, theobromine, and caffeine. The cotyledons of the region of Uraba (7, 8 and 9) and Magdalena Medio (13, 14 and 15) are characterized by high values of citric acid, ORAC, total polyphenols content and ashes, characteristic of unfermented cocoas.
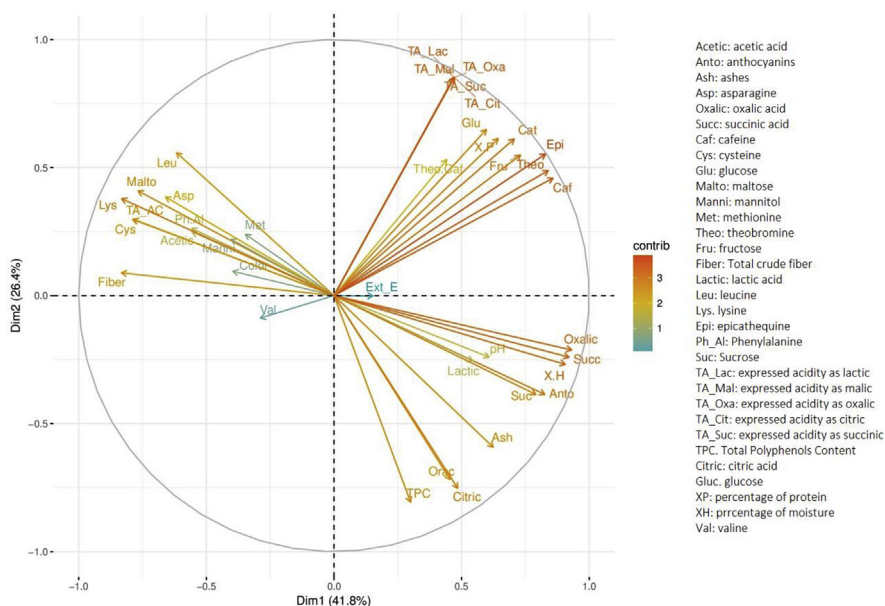
Fig. 2. Biplot of Principal Components Analysis between the bromatological, chemical and physical parameters of raw and well-fermented cocoa.

3.4. Cluster analysis

From the cluster analysis, it was possible to forcefully confirm the distribution by subregions before fermentation, since the profile of the cocoa mixture obtained in Bajo Cauca is clearly separated in the factorial plane, which is grouped in a first cluster identified in blue in the upper left positive axis. The profiles of the mixtures from Uraba and Magdalena Medio match in the second red cluster, distributed in the lower axis in the negative scale. Fig. 4 shows the distribution in the first factorial plane which explains 68.2% of the similarities in two factors or dimensions: Dim1 (41.8%) and Dim2 (26.4%), that are homogeneous within themselves.

After completing fermentation and regardless of the initial conditions of the cocoas grown in different sites, the three subregions were grouped in the same cluster identified in yellow and located in the upper right axis in the middle of the previous clusters. This last group allowed confirming the association of the three subregions at the end of fermentation according to the maximum variability of the identified variables, which

explains the component of these observations despite indicating two vectors in the second principal component near the X axis (valine and total crude fiber).

3.5. Chemometric approximation

The characteristic variables of the unfermented cocoa in Bajo Cauca were the total acidity expressed in citric, lactic, oxalic, and succinic acid, the glucose content, fructose, total protein, theobromine, caffeine, ratio of theobromine/caffeine, catechin, and epicatechin (12 of the 38 parameters analyzed). On the other hand, Uraba and Magdalena Medio were described by 11 variables: citric acid, lactic, oxalic and succinic, pH, sucrose, anthocyanins, ashes, moisture content, total polyphenols content, and antioxidant capacity by ORAC. Most of the characteristics could be exploited or controlled to obtain optimal results in this matrix before going to fermentation: by the knowledge a priori of the bromatological, physical and chemical profile with only these variables selected as significant, which may be susceptible to change through the management of
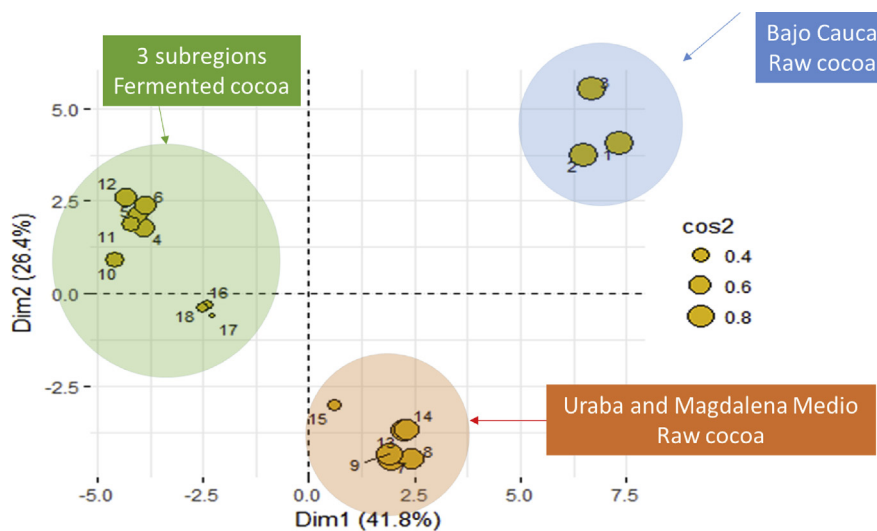

Fig. 3. Contribution of each sample observation in the first factorial plane. Cocoa samples from Bajo Cauca 1, 2 and 3 (unfermented) and 4,5 and 6 (well fermented - 132 h). Uraba 7, 8 and 9 (unfermented) and 10, 11 and 12 (well fermented - 132 h) and Magdalena Medio 13, 14 and 15 (unfermented) and 16, 17 and 18 (well fermented - 168 h).

agronomic conditions or good agricultural practices, which are considered to be some of the causes of continuous improvement (Bedoya, 2016; Hashim et al., 1998; Oliveira et al., 2016).

Similarly, the chemometric study allowed the distinction of the linear combination of 13 variables associated with cocoa after being subjected to the fermentation process that consolidates eight free amino acids, the acidity expressed in acetic acid and its respective determined concentration, mannitol, total crude fiber and color, mainly explained by the main component of the left axis, with valine, methionine, mannitol, and color being the variables of lowest contribution, followed by acetic acid and phenylalanine with a medium contribution, and the others with a high contribution (see Fig. 2).

The differences between the subregions at the beginning of fermentation could be attributed to agroecological conditions since Bajo Cauca was the only area that had two kinds of land, including La Isla, surrounded by the Cauca River, which presented soil characteristics with a high pH (6.7) compared to the other productive areas where the harvest was carried out, these latter with a pH between 4.7 and 5.0; a conductivity index (CI) of 21 and percentage of organic matter above 94%, unlike the other soils (CI: 12 to 18) and (37%–87%), respectively; and also, a soil with a sandy texture in which approximately 50% of the regional clones were obtained for this subregion. Similarly, Bajo Cauca was the area with the lowest altitude (90–340 m above sea level), the highest room temperature (29.1 ± 3.9 °C) and the lowest relative humidity ($73 \pm 18\%$), when compared to Uraba (RH: $90 \pm 9\%$; Troom: 26.8 ± 2.8 °C) and Magdalena Medio (RH: $88 \pm 13\%$; Troom: 22.4 ± 3.1 °C).

One of the parameters that explain variability is the content of alkaloids, specifically theobromine and caffeine. According to previous studies conducted on crude cocoas for the differentiation of geographical areas, it was found that a higher alkaloid content is observed at lower altitude (Carrillo et al., 2014). These approximations coincide with the results obtained in Bajo Cauca, the cocoa productive zone with the lowest altitude and which corresponded to the highest value of theobromine ($0.65 \pm 0.10\%$ w/w), while in Uraba and Magdalena Medio, the concentrations represented approximately 26% (theobromine) and 33% (caffeine) of the concentration in Bajo Cauca.

With respect to the theobromine/caffeine ratio, the decrease in the three locations was low, since Uraba and Magdalena Medio presented a decrease of 1.7% and only Bajo Cauca reached 6.3%. The difference between subregions is related to the change behavior between the alkaloids after fermentation, which is approximately 2:1 in Bajo Cauca in relation to the other two zones. Thus this subregion showed a decrease of 81.5 % in theobromine and 79.2 % in caffeine, while in Uraba there was a decrease of 44.4% and 37.5%, followed by Magdalena Medio with 35.3% and 28.6% for each alkaloid, respectively. A similar decrease was observed in the two methylxanthines and hence the relationship was maintained without major changes. The decrease of more than 20%, reference value of other previous studies for foreign cocoas, could be given by the days of previous storage, as well as the possible permeability of the husk, higher in the Criollo and Trinitario varieties used in the mixture of cocoa (Carrillo et al., 2014). On the other hand, Uraba and Magdalena Medio were characterized by having a total polyphenols content higher than that reported in Bajo Cauca, an expression that may also be associated with agronomic conditions such as shade and irrigation control, which are two factors associated with oxidative stress, which activates the production of polyphenols as a defense and control mechanism (Tomas-Barberán et al., 2007; Zuidema et al., 2005).

In Bajo Cauca, productive units located in areas far from the Cauca river had a shady level of less than 10% (maximum radiation 1032.7 W/m^2), but the Dios con Nosotros farm (maximum radiation 1188.4 W/m^2) is surrounded by the river, has a greater protection with transient and permanent crops (41%) and coincides with the unit where most of the cocoa fruits were harvested. In the comparison of Bajo Cauca with Magdalena Medio (maximum radiation 1118.8 W/m^2) and Uraba (maximum radiation 1181.03 W/m^2), more production units were found with only 10% of shade or even without this protection, despite being crops with more than 4 years, in which a 30% permanent shade ratio is expected. These observations, aligned with the previous theories, can explain why the plant and the fruit are induced to produce greater quantities of polyphenols.

The behavior of the polyphenols and the antioxidant capacity were coordinated in the same ratios, but the antioxidant capacity stands out, since the values are above some materials of different varieties and origins (Carrillo et al., 2014; de Oliveira et al., 2015; Wang et al., 2016). It can be highlighted that an alternative of cocoas with functional properties of this type may have their origin in the subregions of Uraba or Magdalena Medio, or on the contrary and in order to avoid a direct impact on the sensory profile, the option may be the vegetable material
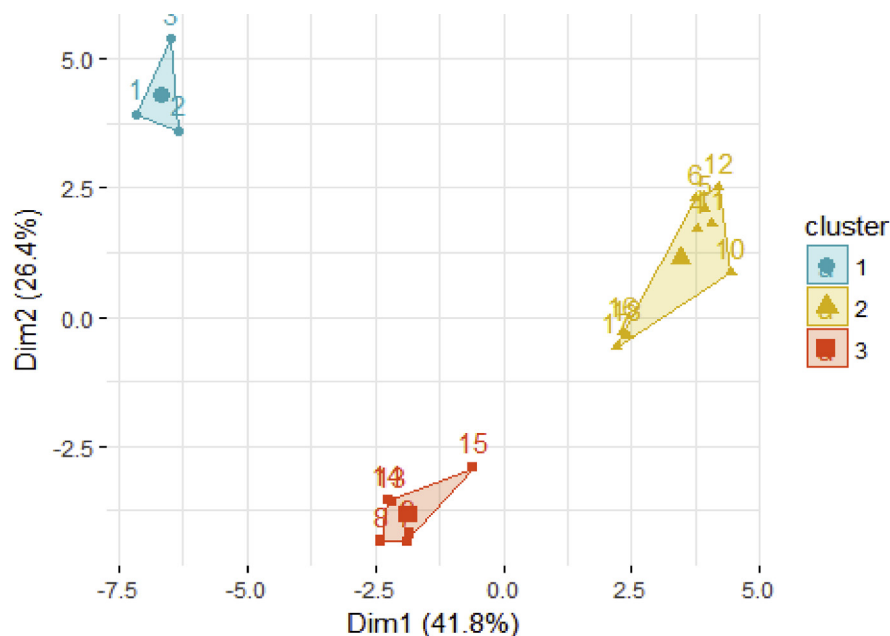

Fig. 4. Cluster of the distribution in the first factorial plane of each sample observation: Bajo Cauca 1, 2 and 3 (unfermented) and 4, 5 and 6 (fermented). Uraba 7, 8 and 9 (unfermented) and 10, 11 and 12 (Fermented) and Magdalena Medio 13, 14 and 15 (unfermented) and 16, 17 and 18 (unfermented).

from Bajo Cauca. However, after fermentation, there was no significant variability requiring a distinction of this type, since the decrease after the process was proportional to the initial amount.

The main procyanidin monomers present in cocoa contribute significantly to the antioxidant capacity and the differentiation between varieties. The behavior of these compounds in Bajo Cauca was similar to that reported by Niemenak et al. (2006), in which the amount of catechin is approximately 10 times lower than epicatechin (0.25–0.023 mg/100 mg). Likewise, alkaloids are within the range published for cocoa from New Guinea (1.5–8.3 mg/g) and Ivory Coast (1.1–7.6 mg/g). Uraba and Magdalena Medio presented lower values than the mentioned ranges.

Another flavonols that are part of the total content of polyphenols are anthocyanins. Anthocyanins area practical indicator for producers because they allow them to follow the evolution of fermentation in a visual way. In raw cocoas, the content of this type of polyphenols can increase as the intensity of solar radiation increases and in the absence of irrigation during the development of the fruit, as part of a protection mechanism (Beer, 1987; Oracz and Nebesny, 2014). Bajo Cauca is the zone with less rainfall frequency in relation to Uraba and Magdalena Medio, which could induce an oxidative stress in the productive units that did not have an irrigation system, thus generating a higher ratio of anthocyanins (2.12 ± 0.15).

Other variables distributed in each component that contribute to the chemometric analysis are acidity and pH. Among the acids expected in raw cocoa, citric acid stands out, which is an oxyacid that greatly contributes to the increase of the acidity of the medium in fresh fruits, since it is bio-synthesized during the ripening of the fruit (Afoakwa et al., 2012; Afoakwa et al., 2009; Afoakwa et al., 2013). Other oxyacids are malic and lactic, although in the mixtures of each subregion under study, only the presence of lactic acid was detected in low concentrations and malic acid did not exceed a concentration of 0.1 mg/L. In parallel with the mentioned acids, it is possible to find some dicarboxylic acids such as oxalic and succinic, the last one being present in the fresh cocoa fruits in greater proportion, as found in the experimental units evaluated. For all the above, the acids mentioned are related to raw cocoa and helped differentiation between subregions. Citric acid was found in a higher concentration in Uraba and Magdalena Medio and, on the contrary, oxalic and succinic acids were lower than in Bajo Cauca.

For its part, acetic acid was present only in concentrations close to the minimum limit of quantification in the raw cotyledon (0.5 mg/L) (see Table 2), which can be attributed to the knowledge of its metabolic pathway, which is not expected in cocoa mainly during the ripening of the fruit but during fermentation (Moreira et al., 2013). This is how the results of the multivariate analysis can be understood since the acetic acid was the only one of the organic acids that did not contribute in the explanation of the variability of the raw cocoas from the three locations, differing from the lactic acid, which was found in low concentrations of 0.86 ± 0.04% w/w, enough to be part of the linear combination of the component that groups the observations of raw cocoa from Uraba and Magdalena Medio with an average contribution (between 2 and 3).

However, the acetic acid at the end of fermentation was the volatile compound included in the variables to explain what happened in the three subregions after the process. Its concentration was below 2.1% w/w, being the one calculated in Uraba the lowest one and Bajo Cauca the highest one, which is reflected in the pH behavior that ends up being higher in Uraba.

The parameters corresponding to the bromatological profile that intervened in the conformation of the main components were the content of ethereal extract, ash, total fiber and moisture. Mainly, the ethereal extract is one of the most important parameters at commercial level and one of the rheological properties required for the control of crystallization in cocoa derivatives (Gu et al., 2013; Araujo et al., 2014). Despite being a variable included in the analysis of components, its vector is in the axis of X with a short length, which induces the identification of the parameter as having a low contribution to the explanation of the difference between subregions. On the other hand, the content of ash in Uraba

and Magdalena Medio was above 2.0% w/w and in Bajo Cauca it remained below this value, results which could be influenced by the days of previous storage of the cobs (Afoakwa et al., 2013).

The content of the aroma precursors marked the difference between the origin of the plant material and the beginning and end of fermentation, especially that of the reducing sugars (fructose and glucose) and the disaccharides from which they come, mainly sucrose and maltose. For the chemometric study, it is evident that the maltose was not significant and therefore it is not part of the first two components. In contrast, in the crude cotyledon, the ratio of sucrose to reducing sugars was 2:1 in Bajo Cauca, followed by Uraba in a lower ratio (sucrose: 6.79 ± 0.65% w/w, fructose: 0.48 ± 0.36% w/w, glucose: 0.30 ± 0.51% w/w) and Magdalena Medio, which showed the lowest values (see Table 2). Finally, as a report on the behavior of the profile obtained at the end of fermentation, it can be observed that fructose is found in a higher concentration than glucose in Bajo Cauca and Uraba, possibly reflecting the higher reactivity of the latter monosaccharide, in line with a similar profile reported by Rodríguez and collaborators on cocoa grown in Mexico (Rodriguez-Campos et al., 2011). Another compound that was part of the profile evaluated was mannitol, which results from the reduction of fructose (Schwan, 2015). Although, it was not significant in the multivariate analysis, it could be a possible explanation for the decrease in fructose.

The other relevant aroma precursor was the presence of free amino acids from protein hydrolysis. The subregion with the highest percentage of proteins was Bajo Cauca (14.01 ± 0.27%), a phenomenon that could be directly marked by the plant nutrition system, according to the type of fertilizer and frequency of this activity. The latter is highlighted, since it agrees with the periodicity of fertilization in Bajo Cauca, which, due to having a lesser shade in areas far from the Cauca River and a high radiation (maximum 1032.7 W/m^2), underwent this practice three times a year. From monitoring the development of total protein content before and after fermentation, it was possible to observe that Bajo Cauca was the subregion where the highest content was found (13.8 ± 0.2%) and which was exposed to greater hydrolysis if associated with the content decrease by 27% after fermentation. The lowest protein content was found in Magdalena Medio and this is evidenced in the final amino acid concentration, which was lower and with a hydrolysis that reached 20%. Finally, the lowest hydrolysis was observed in Uraba with 18%.

In the three subregions, two free amino acids were found in common in raw cocoa: valine and methionine. The latter was reported in a study on the aminolithic profile in raw cocoas in three of the 27 countries evaluated, which can be considered as a differentiating aspect of the Antioquia cocoas (Rohsius et al., 2006). In addition to the above, lysine was the distinctive amino acid in Bajo Cauca and Uraba, while in Magdalena Medio it was cysteine.

After fermentation, an increase in the concentration of lysine was evident in the three subregions, with Uraba being the area with the highest record (see Table 2). Its presence in high concentrations increases the possibility of reaction not only for the formation of the thermal contaminant carboxymethyl lysine but also of other compounds, since this is one of the most reactive amino acids in the glucose-amino acid complex due to the minimization of the steric effect (Jumnongpon et al., 2012; Nguyen et al., 2016). For its part, asparagine was not registered in Magdalena Medio.

The other amino acids present in the three subregions at the end of fermentation, either because they were developing or because they were decreasing throughout the process, consisted of leucine, methionine, valine, and cysteine. Only phenylalanine was not detected in Uraba. The developed profile can be the precursor of properties related to aroma or rheological properties, such as those associated with cysteine for its ability to form S–H bridges and affect texture. On the other hand, the volatile compounds related to aromatic aldehydes are developed from hydrophobic amino acids such as valine, leucine, which undergoes a degradation to form 3-methyl butanal with malt descriptor, and phenylalanine, which produces phenyl acetaldehyde, related to a floral descriptor and honey. In later stages, these compounds will become the

precursors of amino-ketones, main reactants for the formation of pyrazines, compounds of great relevance in the aroma of chocolate (Owusu et al., 2013).

Finally, this chemometric approach allowed to corroborate the hypothesis raised in this study, demonstrating the differentiation in the quality of cocoa beans according to the subregion and the variety between raw and well-fermented cocoas.

3.6. Factor analysis from the Principal Components Analysis

In order to demonstrate the possible intentionality of the information obtained from cocoa according to its origin or treatment for the formation of precursors which could be projected to other studies or used as quality control, the components obtained were used to define a Principal Factors Analysis (PFA-PCA). The analysis led to three groups and could be used as a tool for making decisions about the classification and distinction of cocoa by the actors involved in the early stages of post-harvest.

Table 4 describes the variables that make up the three factors, with the respective weighting and unobservable denomination that was reached.

The three factors together explained 100% of the variation of the whole sample analyzed, with a weight of 39, 35 and 26, which include the following distinctive signs:

Factor 1, called "Cocoa quality indicators with functional properties", was defined in this way to represent the existence of the compounds responsible for the antioxidant activity (catechin and epicatechin) and the content of alkaloids (caffeine and theobromine). Both groups are related to their desirable functional properties, but an effect on taste (astringent and bitter) produced by these compounds makes the industry define what the balance in the content will be. On the other hand, the concentration of acids may alter the pH of the medium, a parameter that may directly influence the inactivation of the enzyme polyphenol oxidase, responsible for the transformation of the procyanidin monomers mentioned. Finally, the presence of reducing sugars in this factor can help to understand how many are available to start the Maillard reaction once the fermentation is finished and the drying begins.

Table 4

Analysis of Main Factors for cocoas before and after fermentation in three subregions of Antioquia.

Factor	Denomination/Weighting	Variables
First Factor	Cocoa quality indicators with functional properties **39**	Total acidity expressed in: lactic, malic, malic, oxalic and succinic acid
		Protein content (%)
		Fructose content (% w/w)
		Glucose content (% w/w)
		Theobromine content (% w/w)
		Caffeine content (% w/w)
		Epicatechin content (% w/w)
		Catechin content (% w/w)
Second Factor	Quality indicators of beginning of fermentation **35**	pH
		Moisture (%)
		Ashes (%)
		Anthocyanins ratio
		Lactic acid content (% w/w)
		Citric acid content (% w/w)
		Oxalic acid content (% w/w)
		Succinic acid content (% w/w)
		Sucrose content (% w/w)
Third Factor	Quality indicators of well-fermented cocoa **26**	Fiber (%)
		Mannitol content (%)
		Leucine content (%)
		Methionine content (% w/w)
		Valine content (% w/w)
		Cysteine content (% w/w)
		Asparagine content (% w/w)
		Lysine content (% w/w)

The denomination associated with the functional properties is due to the fact that health benefits have been attributed to the alkaloids present in cocoa. However their impact at sensory level has been the subject of studies to achieve their reduction, since they are attributed a significant contribution to the astringent taste, which leads to the challenge of achieving a balance (Afoakwa et al., 2012).

Factor 2 was defined as "Indicators of quality of the beginning of fermentation" since this moment of the process (raw cocoa) is characterized by its high humidity and sucrose content, which will be hydrolyzed during fermentation. On the other hand, the pH will be determined by citric acid, characteristic of raw cocoa, and other acids that will vary according to the cocoa origin, as demonstrated by the PCA.

The last factor, "Indicators of quality of well-fermented cocoa", takes into consideration the free amino acids, aroma precursors characteristic of the final phase of fermentation, which are produced by hydrolysis due to the action of the enzymes endoprotease aspartic and carboxypeptidase serine (Schwan, 2015).

4. Conclusion

In general, the chemometric approach led to selection of variables, which define the profile of unfermented cocoa from a mixture of clones according to the origin of three of the main productive zones in the department of Antioquia, relevant findings for the characterization of quality by origen, since as evidenced varies among them, with properties that may lead to different strategies of post-harvest or use in the industry, and a saving in the selection of analysis, going from 38 parameters to 13 or 11 that safely explain the main characteristics of raw cocoa beans.

The profiles of the fermented cocoa were defined spontaneously in the three subregions, which allowed finding a group of associated variables through a linear combination, thus making it possible to reduce the information to the really significant analyses so as to distinguish between the unfermented cocoa and the one that reaches the end of this process.

Markers were identified which allowed a distinction of some descriptors of attributes and defects. These can be established through the monitoring of factors that bring together some of the compounds that provide information on the profiles of specific interest for the cocoa production chain, such as the quality of the cocoa with respect to functional properties, characteristic conditions of the beginning of fermentation or attributes related to well-fermented cocoa (formation of precursors). In this way, it is possible to conduct an evaluation of the quality of this raw material of the chocolate industry obtained before or after fermentation and in different productive zones.

Declarations

Author contribution statement

Maritza Gil: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yamile Jaramillo, Sandra M Llano: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Carolina Bedoya: Performed the experiments; Analyzed and interpreted the data.

Vanessa Gellego: Analyzed and interpreted the data; Wrote the paper.

Jairo Quijano, Julian Londono-Londono: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Colombia (647-2015).

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

Abbreviations

BC   Bajo Cauca
CA   Cluster Analysis
CI   Conductivity Index
CCN-51   Collection Castro Naranja
FEC-2   Fedecacao El Carmen, Santander
FLE -2   Fedecacao Lebrija, Santander
RH   Relative Humidity
ICS-1   Imperial College Selection
M   Magdalena Medio
ORAC   Oxygen Radical Absorbance Capacity
PCA   Principal Component Analysis
PFA   Principal Factors Analysis
Troom   Room Temperature
R   Raw
U   Uraba
WF   Well-Fermented
0CF-BC   raw cocoa on day zero of fermentation in Bajo Cauca
0CF-U: raw cocoa on day zero of fermentation in Uraba
0CF-M: raw cocoa on day zero of fermentation in Magdalena Medio
132CF-BC   well-fermented cocoa at 132 h of fermentation in Bajo Cauca
132CF-U: well-fermented cocoa at 132 h of fermentation in Uraba
168CF-M: well-fermented cocoa at 168 h of fermentation in Magdalena Medio

References

Afoakwa, E.O., Paterson, A., Fowler, M., Ryan, A., 2009. Matrix effects on flavour volatiles release in dark chocolates varying in particle size distribution and fat content using GC-mass spectrometry and GC-olfactometry. Food Chem. 113, 208–215.

Afoakwa, E.O., Quao, J., Takrama, F.S., Budu, a. S., Saalia, F.K., 2012. Changes in total polyphenols, o-diphenols and anthocyanin concentrations during fermentation of pulp pre- conditioned cocoa (Theobroma cacao) beans. Int. Food Res. J. 19 (3), 1071–1077.

Afoakwa, E.O., Quao, J., Takrama, J., Budu, A.S., Saalia, F.K., 2013. Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. J. Food Sci. Technol. 50 (December), 1097–1105.

Araujo, Q.R., Fernandes, C.A.F., Ribeiro, D.O., Efraim, P., Steinmacher, D., Lieberei, R., et al., 2014. Cocoa quality index – a proposal. Food Control 46, 49–54.

Badrie, N., Bekele, F., Sikora, E., Sikora, M., 2015. Cocoa agronomy, quality, nutritional, and health aspects. Crit. Rev. Food Sci. Nutr. 55 (5), 620–659.

Bedoya, C., 2016. Metodologías para el análisis bromatológico, físico y químico del cacao fermentado y seco, dentro del marco normativo internacional. Corporación Universitaria Lasallista.

Beer, J., 1987. Advantages, disadvantages and desirable characteristics of shade trees for coffee, cacao and tea. Agrofor. Syst. 5 (1), 3–13.

Bertoldi, D., Barbero, A., Camin, F., Caligiani, A., Larcher, R., 2016. Multielemental fingerprinting and geographic traceability of Theobroma cacao beans and cocoa products. Food Control 65, 46–53.

Carrillo, L.C., Londoño-Londoño, J., Gil, A., 2014. Comparison of polyphenol, methylxanthines and antioxidant activity in Theobroma cacao beans from different cocoa-growing areas in Colombia. Food Res. Int. 60, 273–280.

Comisión venezolana de Normas Industriales - COVENIN, 1995. Norma venezolana. Granos de Cacao. Covenin 50, 1995.

de Oliveira, T.B., Rogero, M.M., Genovese, M.I., 2015. Poliphenolic-rich extracts from cocoa (Theobroma cacao L.) and cupuassu (Theobroma grandiflorum Willd. Ex Spreng. K. Shum) liquors: a comparison of metabolic effects in high-fat fed rats. PharmaNutrition 3 (2), 20–28.

Dirección General de Normas, 1980. Norma Mexicana NMX-F-352-S. Cacao en grano fermentado.

Doores, S., 1993. In: Organic Acids. YMC Separation Technology, Vol. 18.

Gentleman, Robert, Hornik, Kurt, P, G., 2011. Use R! Springer, New York. Media (First).

Gourieva, K.B., Tserrevitinov, O.B., 1979. 646254. USSR.

Gil, M., 2018. Aproximación quimiométrica del balance entre los compuestos neoformados y los responsables de la calidad desarrollados durante las etapas de poscosecha de cacaos especiales (Theobroma cacao L.) cultivados en Antioquia. Universidad Nacional de Colombia - Medellín.

Guehi, S.T., Dabonne, S., Ban-Koffi, L., Kedjebo, D.K., Zahouli, G.I.B., 2010. Effect of turning beans and fermentation method on the acidity and physical quality of raw cocoa beans. Adv. J. Food Sci. Technol. 2 (3), 163–171.

Gu, F., Tan, L., Wu, H., Fang, Y., Xu, F., Chu, Z., Wang, Q., 2013. Comparison of cocoa beans from China, Indonesia and Papua New Guinea. Foods 2, 183–197.

Hashim, P., Selamat, J., Kharidah, S., Ali, A., 1998. Changes in free amino acid , peptide-N , sugar and P y razine concentration during cocoa fermentation. J. Sci. Food Agric. 78, 535–542.

Hii, C.L., Menon, A.S., Chiang, C.L., Sharif, S., 2017. Kinetics of hot air roasting of cocoa nibs and product quality. J. Food Process. Eng. 40 (3), 1–6.

Ibáñez Martínez, M., Hernández Fernández, F., 2007. Universitat Jaume I Departament de Ciències Experimentals Química Analítica Institut Universitari de Plaguicides i Aigües Potencial de la cromatografía líquida acoplada a espectrometría de para la elucidación, cuantificación y confirmación de plaguicidas. Society. Retrieved from htt p://www.tesisenred.net/TDX-0908110-123653.

ICONTEC Internacional, 2003. Norma Técnica Colombiana NTC 1252. Cacao en grano.

Illeghems, K., Weckx, S., De Vuyst, L., 2015. Applying meta-pathway analyses through metagenomics to identify the functional properties of the major bacterial communities of a single spontaneous cocoa bean fermentation process sample. Food Microbiol. 50, 54–63.

Jumnongpon, R., Chaiseri, S., Hongsprabhas, P., Healy, J.P., Meade, S.J., Gerrard, J.A., 2012. Cocoa protein crosslinking using Maillard chemistry. Food Chem. 134, 375–380.

Kaiser, H.F., 1961. A note on Guttman's lower bound for the number of common factors. Br. J. Stat. Psychol. 14, 1–2.

Kongor, J.E., Hinneh, M., Van de Walle, D., Afoakwa, E.O., Boeckx, P., Dewettinck, K., 2016. Factors influencing quality variation in cocoa (Theobroma cacao) bean flavour profile - a review. Food Res. Int. 82, 44–52.

Krähmer, A., Engel, A., Kadow, D., Ali, N., Umaharan, P., Kroh, L.W., Schulz, H., 2015. Fast and neat – determination of biochemical quality parameters in cocoa using near infrared spectroscopy. Food Chem. 181, 152–159.

Moreira, I.M., Miguel, M.G., Duarte, W., Dias, D., Schwan, R., 2013. Microbial succession and the dynamics of metabolites and sugars during the fermentation of three different cocoa (Theobroma cacao L.) hybrids. Food Res. Int. 54 (1), 9–17.

Nazaruddin, R., Seng, L.K., Hassan, O., Said, M., 2006. Effect of pulp preconditioning on the content of polyphenols in cocoa beans (Theobroma Cacao) during fermentation. Ind. Crops Prod. 24 (1), 87–94.

Nguyen, H.T., van der Fels-Klerx, H.J., van Boekel, M.A.J.S., 2016. Kinetics of N(ε)-(carboxymethyl)lysine formation in aqueous model systems of sugars and casein. Food Chem. 192, 125–133.

Ohene, E., Quao, J., Takrama, J., Simpson, A., Kwesi, F., 2011. Chemical composition and physical quality characteristics of Ghanaian cocoa beans as affected by pulp pre-conditioning and fermentation. J. Food Sci. Technol. 50, 1097–1105.

Oliveira, L.F., Braga, S.C.G.N., Augusto, F., Hashimoto, J.C., Efraim, P., Poppi, R.J., 2016. Differentiation of cocoa nibs from distinct origins using comprehensive two-dimensional gas chromatography and multivariate analysis. Food Res. Int. 90, 133–138.

Oracz, J., Nebesny, E., 2014. Influence of roasting conditions on the biogenic amine content in cocoa beans of different Theobroma cacao cultivars. Food Res. Int. 55, 1–10.

Ortega, N., Romero, M.P., Macià, A., Reguant, J., Anglés, N., Morelló, J.R., Motilva, M.J., 2010. Comparative study of UPLC-MS/MS and HPLC-MS/MS to determine procyanidins and alkaloids in cocoa samples. J. Food Compos. Anal. 23 (3), 298–305.

Ou, B., Hampsch-Woodill, M., Prior, R., 2001. Development and Validation of an Improved Oxygen Radical Ab- sorbance Capacity Assay Using Fluorescein as the Fluorescent Probe. J. Agric. Food Chem. 49, 4619–4626.

Owusu, M., Petersen, M.A., Heimdal, H., 2013. Relationship of sensory and instrumental aroma measurements of dark chocolate as influenced by fermentation method, roasting and conching conditions. J. Food Sci. Technol. 50 (5), 909–917.

Pereira, G. V. de M., Soccol, V.T., Soccol, C.R., 2016. Current state of research on cocoa and coffee fermentations. Curr. Opin. Food Sci. 7, 50–57.

Rodriguez-Campos, J., Escalona-Buendía, H.B., Orozco-Avila, I., Lugo-Cervantes, E., Jaramillo- Flores, M.E., 2011. Dynamics of volatile and non-volatile compounds in cocoa (Theobroma cacao L.) during fermentation and drying processes using principal components analysis. Food Res. Int. 44 (1), 250–258.

Rodriguez-Campos, J., Escalona-Buendía, H.B., Contreras-Ramos, S.M., Orozco-Avila, I., Jaramillo- Flores, E., Lugo-Cervantes, E., 2012. Effect of fermentation time and drying temperature on volatile compounds in cocoa. Food Chem. 132, 277–288.

Rohsius, C., Matissek, R., Lieberei, R., 2006. Free amino acid amounts in raw cocoas from different origins. Eur. Food Res. Technol. 222, 432–438.

Schwan, R.F., 2015. In: Schaw, R., Fleet, G. (Eds.), Cocoa and Coffee Fermentations. CRC Press, Boca Ratón, FL.

Singleton, V.L., Rossi, J.A., 1965. Colorimetry of total phenolics with phosphomolybdicphosphotungstic acid reagents. Am. J. Enol. Vitic. 16, 144–158.

Tomas-Barberán, F.A., Cienfuegos-Jovellanos, E., Marín, A., Muguerza, B., Gil-Izquierdo, A., Cerdá, B., et al., 2007. A new process to develop a cocoa powder with higher flavonoid monomer content and enhanced bioavailability in healthy humans. J. Agric. Food Chem. 3926–3935.

Vignoni, L.A., Césari, R.M., Forte, M., Mirábile, M.L., 2006. Determinación de Indice de Color en Ajo Picado. Información Tecnológica 17 (6), 63–67.

Voigt, J., Heinrichs, H., Voigt, G., Biehl, B., 1994. Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. Food Chem. 50 (2), 177–184.

Wang, L., Chen, C., Su, A., Zhang, Y., Yuan, J., Ju, X., 2016. Structural characterization of phenolic compounds and antioxidant activity of the phenolic-rich fraction from defatted adlay (Coix lachryma-jobi L. var. ma-yuen Stapf) seed meal. Food Chem. 196, 509–517.

Wold, S., 1995. Chemometrics; what do we mean with it, and what do we want from it? Chemometr. Intell. Lab. Syst. 30.

Wollgast, J., Anklam, E., 2000. Review on polyphenols in Theobroma cacao: changes in composition during the manufacture of chocolate and methodology for identification and quantification. Food Res. Int. 33 (6), 423–447.

Zuidema, P.A., Leffelaar, P.A., Gerritsma, W., Mommer, L., Anten, N.P.R., 2005. A physiological production model for cocoa (Theobroma cacao): model presentation, validation and application. Agric. Syst. 84 (2), 195–225.