LWT - Food Science and Technology 117 (2020) 108598

Contents lists available at ScienceDirect

LWT - Food Science and Technology

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

Changes in methylxanthines and flavanols during cocoa powder processing and their quantification by near-infrared spectroscopy

Maribel Alexandra Quelal-Vásconez[a], María Jesús Lerma-García[a,*], Édgar Pérez-Esteve[a], Alberto Arnau-Bonachera[b,c], José Manuel Barat[a], Pau Talens[a]

[a] Departamento de Tecnología de Alimentos. Universitat Politècnica de València, Camino de Vera, s/n, 46022, Valencia, Spain
[b] Institute for Animal Science and Technology. Universitat Politècnica de València, Camino de Vera, s/n, 46022, Valencia, Spain
[c] Biomedical Research Institute (PASAPTA-Pathology Group), Veterinary School, Universidad Cardenal Herrera-CEU, CEU Universities, Av. Seminario s/n, 46113, Moncada, Valencia, Spain

ARTICLE INFO

Keywords:
Cocoa powder
Methylxanthines
Flavanols
HPLC
Near-infrared spectroscopy

ABSTRACT

Variation in methylxanthines (theobromine and caffeine) and flavanols (catechin and epicatechin) was studied in a large set of cocoa powders (covering different origins, processing parameters and alkalisation levels). The content of these compounds was established by high-performance liquid chromatography (HPLC), whose results showed that the alkalisation process lowered the content of all analytes, whose loss was more evident in flavanols. Therefore, the determination of these analytes in a huge set of samples allowed not only better knowledge of the concentration variability in natural commercial cocoas from different origins, but also the understanding of the effect that industrial alkalisation has on these contents. The feasibility of reflectance near-infrared spectroscopy (NIRS) combined with partial least square (PLS) to non-destructively predict these contents, was also evaluated. All the analytes were generally well predicted, with predictions for methylxanthines (R^2_P 0.819–0.813 and RMSEP 0.068–0.022%, and bias 0.005 and 0.007 for theobromine and caffeine, respectively) and for flavanols (R^2_P 0.830–0.824; RMSEP 8.160–7.430% and bias −1.440 and −1.034 for catechin and epicatechin, respectively). Thus NIRS could be an alternative fast reliable method for the routine assessment of these analytes in the cocoa industry.

1. Introduction

Cocoa powder is the most important raw material of confectionery products, chocolate-flavoured bakeries, ice-creams and drinks (Miller et al., 2008). Apart from technologic properties, cocoa (*Theobroma cacao* L.) and its derived products are well considered for being a rich source of methylxanthines and polyphenols (Langer, Marshall, Day, & Morgan, 2011).

Methylxanthines (i.e. theobromine and caffeine) are pharmacologically active alkaloids responsible for a bitter cocoa taste and desirable physiological effects; e.g. stimulation of the central nervous system and gastric secretion, diuresis, bronchodilation, and stimulation of skeletal muscles in high doses (Franco, Oñatibia-Astibia, & Martínez-Pinilla, 2013). They also display health benefits in diseases involving cell death in the nervous system (Oñatibia-Astibia, Franco, & Martínez-Pinilla, 2017).

The main type of polyphenols (known for their demonstrated antioxidant and anti-inflammatory properties) in cocoas are flavanols. This family of compounds includes catechin and epicatechin (monomeric species) and procyanidins (oligomeric and polymeric fractions). Among them, epicatechin is the most abundant flavanol in cocoa and accounts for 35% of the total polyphenolic fraction (Lacueva et al., 2008).

While producing cocoa powder from cocoa beans, seeds are primarily fermented, dried and roasted. Then broken beans (nibs) are ground, heated and liquefied. The product of these operations, cocoa liquour, is pressed to obtain two different fractions: cocoa powder and butter. Optionally, nibs or cake can be treated with an alkali dissolved in water. This alkalisation reduces acidity, bitterness and astringency by improving and enhancing the aromatic features of cocoa powder (Kongor et al., 2016). Alkalisation also allows colour development by transforming the typical light brown hues of natural cocoa powder into reddish or very dark ones. These colour changes derive from a sequence of chemical reactions between alkalizing agents and pigments in the presence of water and oxygen at high temperatures. Finally, the dispersibility of cocoa powder increases with alkalisation. This property facilitates the use of cocoa powders in different industries like dairy

* Corresponding author.
E-mail address: malerga1@tal.upv.es (M.J. Lerma-García).

https://doi.org/10.1016/j.lwt.2019.108598
Received 17 December 2018; Received in revised form 4 September 2019; Accepted 5 September 2019
Available online 06 September 2019
0023-6438/ © 2019 Elsevier Ltd. All rights reserved.

Abbreviations

HPLC	high performance liquid chromatography
GC	gas chromatography
IR	infrared spectroscopy
NIRS	near infrared spectroscopy
PLS	partial least square
PCA	principal component analysis
LV	latent variable
RMSEC	root mean square error of calibration
RMSECV	root mean square error of cross-validation
RMSEP	root mean square error of prediction
R^2_C	coefficient of determination for calibration
R^2_{CV}	coefficient of determination for cross-validation
R^2_P	coefficient of determination for prediction
RPD	ratio of prediction deviation
LOD	limit of detection
LOQ	limit of quantification
S-G	Savitzky-Golay
OSC	orthogonal signal correction

products (Afoakwa, Paterson, Fowler, & Ryan, 2008).

Cocoa powder processing not only strongly affects the sensory properties of cocoa and derived products, but also alters flavonoid and methylxanthine fractions (Payne, Hurst, Miller, Rank, & Stuart, 2010). However, studies to date have used small sample numbers and have, thus provided conclusions that can be read only as behaviour tendencies.

The commonest techniques employed to analyse methylxanthines and flavanols from cocoa extracts or isolated fractions are high-performance liquid chromatography (HPLC) or gas chromatography (GC) (Cádiz-Gurrea et al., 2014; Elwers, Zambrano, Rohsius, & Lieberei, 2009; Fayeulle et al., 2018; Humston, Knowles, McShea, & Synovec, 2010; Machonis, Jones, Schaneberg, Kwik-Uribe, & Dowell, 2014; Van Durme, Ingels, & De Winne, 2016), which provide reliable and accurate descriptions of individual cocoa qualities. Recently, the use of novel methods based on on-line comprehensive two-dimensional liquid chromatography coupled to tandem mass spectrometry (LC × LC-MS/MS) have allowed the characterization of new secondary metabolites of cocoa beans (Toro-Uribe, Montero, López-Giraldo, Ibáñez, & Herrero, 2018). However, these methods are not recommended for routine raw material quality controls as they are destructive, require specialised personnel, sample preparation and expensive equipments. Thus simpler, faster and non-destructive techniques are required.

An alternative is infrared spectroscopy (IR is a fast non-destructive analytical tool that needs little samples preparation), which is useful for both qualitative and quantitative analyses of molecules. Finally, the application of chemometric techniques to IR data provides a powerful tool to develop methods capable of classifying or characterising samples (Vergara-Barberán, Lerma-García, Herrero-Martínez, & Simó-Alfonso, 2015).

Recent studies into commercially cocoa derivates have shown the suitability of near IR (NIR) for the quantification of main constituents like fat, protein, moisture and carbohydrates (Veselá et al., 2007). NIRS has also been shown appropriate to determine parameters like geographical origin (Teye, Huang, Dai, & Chen, 2013), fermentation quality (Hue et al., 2014), fermentation index or pH (Sunoj, Igathinathane, & Visvanathan, 2016), and minor valuable components like organic acids (Krähmer et al., 2015), caffeine, theobromine and epicatechin in unfermented and sun-dried beans (Álvarez, Pérez, Cros, Lares, & Assemat, 2012), procyanidins (Whitacre et al., 2003) and adulterations with carob flour (Quelal-Vásconez, Pérez-Esteve, Arnau-Bonachera, Barat, & Talens, 2018) and cocoa shell (Quelal-Vásconez et al., 2019). Despite these advances in cocoa characterisation by NIRS, as far as we know, the effect of cocoa powder processing on methylxanthine and flavanol content by this technique has not yet been studied.

In this context, the objective of this work is twofold. Firstly, to study the effect of cocoa powder processing on methylxanthines (theobromine and caffeine) and flavanols (catechin and epicatechin) contents in a large batch of samples (with different origins, processing parameters and alkalisation levels). Secondly, to evaluate the feasibility of reflectance NIRS combined with partial least square (PLS) to non-destructively predict the content of these compounds in cocoa powders.

2. Materials and methods

2.1. Reagents and samples

The employed reagents were: caffeine, theobromine, catechin and epicatequin (Sigma-Aldrich, St. Louis, Missouri, USA), acetonitrile (J.T. Baker, The Netherlands), methanol (Labkem, Barcelona, Spain) and acetic acid glacial (Sharlau, Barcelona, Spain). Deionised water was obtained using an Aquinity deioniser (Membrapure GmbH, Berlin, Germany).

Samples comprised 86 cocoa powders provided by Olam Food Ingredients Company (Cheste, Spain) or purchased in different nationals and international markets, to assure variability in cocoa origin (Ivory Coast (n = 22), Ghana (n = 8), Indonesia (n = 12), Ecuador (n = 7), Peru (n = 4), or undeclared (n = 33), year of production (2017 (n = 67) or 2018 (n = 19)) and processing practices (natural or alkalised cocoa powders).

Before analysing, samples were characterised according to their extractable pH and colour. Extractable pH determination was performed by the procedure described in the ADM Cocoa Manual (2006). CIE L*a*b colour coordinates were obtained by measuring the reflection spectrum using a 10° observer and D65 illuminant (Minolta, CM 3600D, Tokyo, Japan). Extractable and colour data were used to classify cocoa powders into five categories: natural (NC) (pH 5.0–6.0), light-alkalised (LAC) (pH 6.0–7.2), medium-alkalised (MAC) (pH 7.2–7.6), strong-alkalised (SAC) (pH > 7.6) and black powders (BC) (pH > 7.6 and very low L values) (Miller et al., 2008).

2.2. Instrumentation and experimental conditions

2.2.1. Methylxanthines and flavanols extraction

In order to extract methylxanthines and flavanols from cocoa powders, the protocol of Lacueva et al. (2008) was adapted: 0.5 g of cocoa powder was weighed, suspended in 5 mL H_2O at 100 °C + 20 mL methanol, mixed for 20 min at 36 °C by constant agitation, and centrifuged for 10 min at 10,000 rpm at room temperature. Finally, the obtained supernatants were filtered using 0.22 μm pore size PTFE filters (Scharlab, Barcelona, Spain). The obtained samples were immediately injected into the HPLC system or stored at −20 °C.

2.2.2. HPLC determination of methylxanthines and flavanols

Analytes determination was performed in a liquid chromatograph model LaChrom Elite (Hitachi Ltd., Tokyo, Japan), equipped with an auto-sampler and a UV detector (models L-220 and L-2400, respectively). A 5-μm analytical column was used, model Liquid Purple C18 (250 × 4.6 mm i.d.) from Análisis Vínicos (Tomelloso, Spain). The mobile phase was prepared by mixing 2% aqueous acetic acid (phase A) and a ternary mixture composed of acetonitrile–H_2O–acetic acid in the 40:9:1 v/v/v proportion (phase B). The following gradient (adapted from Elwers et al., 2009) was used: 0–8 min, 10% B; 8–20 min, 10–15% B; 20–35 min, 15–90% B; 35–50 min, 90% B, then back to the initial conditions. Other chromatographic conditions were: UV detection, 280 nm; column temperature, 40 °C; injection volume, 20 μL; flow rate,

1.2 mL min^{-1}.

The differences in analyte content with the various different alkalisation levels were established by a multifactorial ANOVA (95% confidence level of LSD; $p < 0.05$), constructed using Statgraphics Centurion XV from Manugistics Inc. (Rockville, MD, USA).

2.2.3. NIR spectra collection

The 86 cocoa powders were scanned in a FOSS NIR 5000 System spectrometer, equipped with a transport module (Silver Spring, MD, USA), using the *ISIscan*™ Software version 4.10.0 (Infrasoft International, State College, PA, USA). Samples (ca. 5.0 g) were measured by filling a round sample cup (3.8 cm diameter x 1 cm-thick quartz windows) to preserve unvarying surface and thickness during spectral collection. For each sample, 32 sequential scans with 700 wavelengths were measured from 1100 to 2500 nm at 2-nm intervals. Samples were measured twice, and the mean spectra were employed for the statistical analysis.

2.3. Spectral data analysis

Calibration models to predict the methylxanthines (theobromine and caffeine) and flavanols (catechin and epicatechin) contents (previously established by HPLC) were constructed by PLS analysis of the spectral data using Unscrambler v10.4 software from the CAMO Software AS (OSLO, Norway). PLS is a multivariate calibration method that could be used to correlate NIR spectra data with chemical component contents. Spectral data were organised in a matrix containing the number of cocoa samples ($N = 86$) in rows and the X- and Y-variables in columns. The X-variables corresponded to the 700 wavelengths, while the Y-variables were the HPLC-determined analyte contents. For PLS model construction, the 86 samples were separated into two datasets: the 67 samples from, 2017 were employed to create and evaluate the PLS models by leave-one-out cross-validation, while the 19 samples from 2018 were used as an external validation set.

Before any spectral treatment, all spectra were used to construct a principal component analysis (PCA) model, which was employed to identify and remove defective spectral outliers using the Q residual values and the Hotelling T^2 with a 95% confidence limit (Bro & Smilde, 2014). Moreover, with the aim of detecting similarities among the samples, NIR spectra data were employed to build a clustering model by using the hierarchical complete-linkage method (HCL). This method aims to find similarities and patterns between samples, being this information represented in a dendrogram plot (Wajrock, Antille, Rytz, Pineau, & Hager, 2008).

Finally, PLS models were constructed with no spectral pretreatment (raw data) and with three different pre-treatments: 2nd derivative performed with the Savitzky-Golay smoothing algorithm (2nd derivative S-G) (Savitzky & Golay, 1951), orthogonal signal correction (OSC) (Wold, Antti, Lindgren, & Öhman, 1998), and their combination. The PLS models' accuracy was evaluated by: the required number of latent variables (LVs), the root mean square error of calibration (RMSEC) and cross-validation (RMSECV), and the coefficient of determination for calibration (R^2_C) and cross-validation (R^2_{CV}). The PLS models' predictive capability was judged by considering: the coefficient of determination for prediction (R^2_P), the root mean square error of prediction (RMSEP), the bias, the standard error of prediction (SEP) and the ratio of prediction deviation (RPD) (calculated as the ratio between the standard deviation of the reference values -training set- and RMSEP). Performance of the different models was considered good when the number of LVs, and RMSE and SE values were low, when R^2 value tended to unit, and when bias values were near to zero, for both the calibration and external prediction parameters (Quelal-Vásconez et al., 2019). Regarding RPD, a value between 2.0 and 2.5 makes approximate quantitative predictions possible, while values between 2.5 and 3.0, and above 3.0, indicates a prediction that could be considered good and excellent, respectively (Saeys, Mouazen, & Ramon, 2005).

3. Results and discussion

3.1. Alkalisation level evaluation

The 86 cocoa powders were physico-chemically characterised and divided into five categories according to their extractable pH values and colour (Miller et al., 2008). By contemplating these values, 23 cocoa samples were classified as natural, 19 as light-alkalised, 21 as medium-alkalised, 19 as strong-alkalised and 4 as black powders. Considering the linear correlation between alkalisation and pH (Pérez-Esteve, Lerma-García, Fuentes, Palomares, & Barat, 2016), this classification was taken into account to study the relationship between alkalisation intensity and the changes found in caffeine, theobromine, catechin and epicatechin contents.

3.2. HPLC determination of methylxanthines and flavanol contents in cocoa powders

By applying the experimental conditions included in Section 2.2.1, analyte peak identification was achieved by comparing the retention times of the sample peaks with the retention times of the standards. Four peaks were observed at retention times of 5.1, 9.1, 12.0 and 15.8 min, which respectively corresponded to theobromine, catechin, caffeine and epicatechin. These compounds were quantified using external calibration curves. To construct them, six standard solutions at different concentrations within the ranges showed in Table 1 were prepared and injected. In all cases, R^2 above 0.9996 were observed. The other parameters in Table 1 were limit of detection (LOD) and limit of quantification (LOQ), which were estimated following the ICH guidelines (1996). The obtained values ranged between 0.005-0.1 and 0.017–0.33 mg L^{-1} for LODs and LOQs, respectively. These values were lower than those previously reported by others (Gottumukkala, Nadimpalli, Sukala, & Subbaraju, 2014; Risner, 2008; Srdjenovic, Djordjevic-Milic, Grujic, Injac, & Lepojevic, 2008). In order to assure that no matrix effect was observed in the quantification of analytes, standard addition calibration curves (considering the linearity ranges in Table 1) were constructed. The four curves provided R^2 above 0.9995 and similar slopes to the external calibration curves. Therefore, it was concluded that the external calibration curves were correctly used to quantify these analytes in cocoa powders.

Finally, the 86 cocoa powders were subjected to HPLC (see Table 2 for the results). The most abundant alkaloid was theobromine, whose content ranged from 1.53 ± 0.02 to 2.4 ± 0.1 g/100 g when considering all the samples. Statistical differences in theobromine content were found among alkalisation ($p < 0.05$) as BC possessed lower contents than the other categories. The caffeine content ranged from 0.1503 ± 0.0003 to 0.412 ± 0.003 g/100 g. As with theobromine, significant differences were observed among samples with different alkalisation degrees ($p < 0.05$), and this behaviour has been reported by other authors (http://www.sciencedirect.com/science/article/pii/S0889157511001165, Brunetto et al., 2007; Li et al., 2012).

Among flavanols, epicatechin was the most abundant analyte. The highest content (536.59 ± 0.08 mg/100 g) went to a NC sample

Table 1
Linear ranges, determination coefficient, LOD and LOQ of the determination of methylxanthines and monomeric flavanols by HPLC.

Analyte	t_R (min)	Linear range (mg L^{-1})	R^2	LOD (mg L^{-1})	LOQ (mg L^{-1})
Theobromine	5.1	50–500	0.9998	0.005	0.017
Catechin	9.1	1–50	0.9996	0.1	0.33
Caffeine	12.0	5–250	0.9998	0.05	0.17
Epicatechin	15.8	0.5–100	0.9999	0.1	0.33

t_R: retention time; R^2: determination coefficient; LOD: limit of detection; LOQ: limit of quantification.

Table 2
Theobromine, catechin, caffeine and epicatechin content (mean, minimum, maximum values and their standard error) found for the different types of cocoa powders employed in this study.

Analyte	Natural (pH 5–6)	Alkalisation degree*			
		Light (pH 6–7.2)	Medium (pH 7.2–7.6)	Strong (pH > 7.6)	Black (pH > 7.6, low L*)
Theobromine (g/100g)	2.15^{a} (1.53 ± 0.02–2.36 ± 0.02)	2.13^{ab} (1.77 ± 0.01–2.393 ± 0.002)	2.04^{ab} (1.68 ± 0.02–2.4 ± 0.1)	2.06^{ab} (1.76 ± 0.01–2.33 ± 0.09)	1.95^{b} (1.768 ± 0.001–2.07 ± 0.01)
Catechin (mg/100g)	78.21^{a} (15.2 ± 0.5–167.5 ± 1.2)	60.39^{a} (13.3 ± 0.2–132 ± 1)	27.86^{b} (4.9 ± 0.2–65.6 ± 0.5)	23.36^{b} (4.53 ± 0.07–46.3 ± 0.5)	< LOD
Caffeine (g/100g)	0.29_{a} (0.201 ± 0.002–0.412 ± 0.003)	0.24^{b} (0.175 ± 0.001–0.393 ± 0.002)	0.21^{bc} (0.157 ± 0.002–0.334 ± 0.001)	0.20^{c} (0.172 ± 0.002–0.255 ± 0.001)	0.20^{bc} (0.1503 ± 0.0003–0.302 ± 0.002)
Epicatechin (mg/100g)	163.17^{a} (11.5 ± 0.2–536.59 ± 0.08)	80.24^{b} (11.9 ± 0.2–219 ± 1)	33.27^{bc} (8.0 ± 0.3–111.00 ± 0.05)	28.13^{c} (10.5 ± 0.2–54.9 ± 0.2)	< LOD

Within rows, values bearing different letters are significantly different ($p < 0.05$).
LOD = limit of detection.

labelled as Equator origin. Other natural samples exhibited an average content of ca. 160 mg/100 g. Contents statistically lowered as the alkalisation process became more intense, and reached average values of ca. 80, 33, 28 mg/100 g and with LODs in the LAC, MAC, SAC and BC, respectively. The same behaviour was found for catechin contents. In the NC samples, catechin content ranged from 15.2 ± 0.5 to 167.5 ± 1.2 mg/100 g. The highest value went to another sample from Equator. In the other natural samples, the average value was ca. 80 mg/100 g. According to Table 2, no significant differences in catechin content were found in the LAC cocoa powders, but contents statistically decreased in the medium-alkalised powders, with undetected quantities in the black powders. Similar results have been found by other authors. In a study done with 11 cocoa powder samples, Lacueva et al. (2008) reported values of 200 and 70 mg/100 g for epicatechin and catechin in natural powders, and 30 and 25 mg/100 g in alkalised powders, respectively. Similar contents were observed by Payne et al. (2010), who reported 223 and 88 mg/100 g (natural), 69 and 70 mg/100 g (light-alkalised), 26 and 36 mg/100 g (medium-alkalised) and 4 and 9 mg/100 g (strong-alkalised) for epicatechin and catechin, respectively.

3.3. Clustering analysis of the samples

The clustering analysis of the NIR spectra shows four main groups (see Fig. 1). First group comprised only one sample of BC, while the second one included the other BCs and ca. 16% of the SAC samples. The third group included most NCs, ca. 74% of LACs, ca. 20% of MACs and ca. 6% of SACs. Finally, the fourth group included the 81% of MACs, ca. 79% of SACs, 26% of LACs and the rest of NCs. In general, it can be observed that BC samples (minimum content of catechin and epicatechin detected) belonged to the first and second group and that the rest groups are mostly linked to strong (group 4) or mild alkalisation conditions (group 3). However, this grouping cannot be completely linked neither to the content or flavanols nor methylxanines. For instance, samples with the highest flavanols content (Ecuadorian samples –i.e. NC 67 or NC 52-) are not clearly separated from other natural or alkalised samples. Therefore, it could be concluded that in sample clustering, besides quantified analytes or alkalisation degrees, there are other chemical signals (i.e. proteins, sugars, volatiles, etc) that are affecting sample clustering.

3.4. Prediction of theobromine, catechin, caffeine and epicatechin content by PLS

The raw spectra of the samples (between 1100 and 2500 nm) could contain both useful and irrelevant information; thus, it was pre-treated to enhance the final outcome. For all spectra, the first applied pre-treatment was the 2nd derivative S-G, followed by the OSC pre-treatment and their combination. Then, the optimal pre-treatment method to enhance the PLS models' performance was selected by taking into account the values included in Table 3. As can be observed in this table, the best results were observed after applying the 2nd derivative S-G + OSC for all models. Thus, the spectra obtained after applying this pre-treatment is shown in Fig. 2. As seen in this figure, signal peaks at wavelengths of 1728, 1764, 1884, 2312 and 2348 nm were evidenced. The region comprised between 1600 and 1800 nm predominantly corresponds to the first overtone region of carbohydrates (C–H bands) (Bázár et al., 2016), and the region comprised between 1700 and 2300 nm normally attributes to the first overtones of C–H stretching associated with sugars (Osborne, Fearn, & Hindle, 1993, pp. 123–132).

Table 3 contains the results obtained for the PLS models constructed to predict the different analytes. As seen for the PLS models obtained with the optimal pre-treatment (marked in italics), all the PLS models provided satisfactory R^2_C, R^2_{CV} and R^2_P values within ranges 0.940–0.969, 0.802–0.889 and 0.813–0.819, respectively. For methylxanthine content predictions, the RMSEP values were 0.068 and 0.022 for theobromine and caffeine, respectively, and the bias came very close to 0

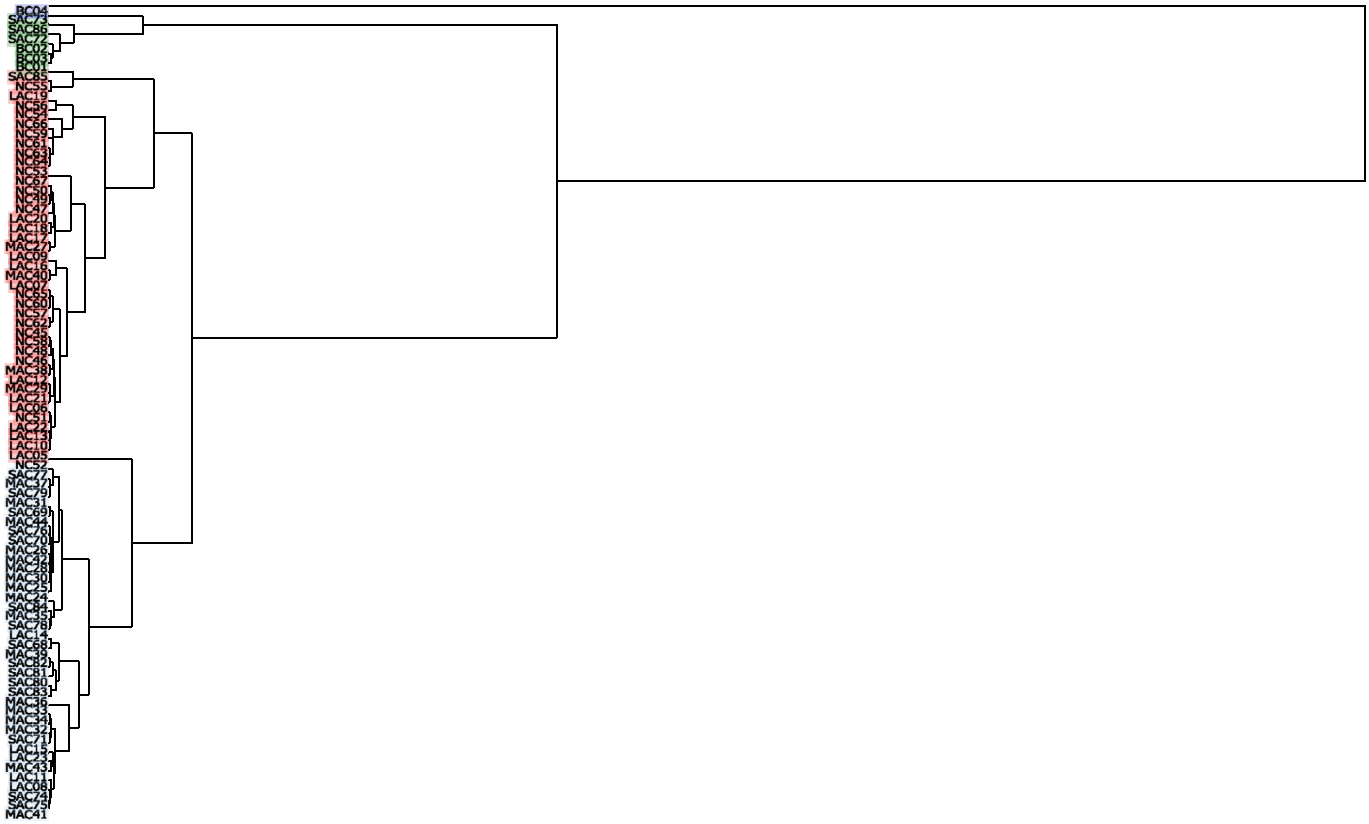

Fig. 1. Dendogram of the clustering analysis performed with the NIR spectra of the entire sample set (n = 86).

(0.005 and 0.007) for theobromine and caffeine, respectively. With flavanols, the models provided RMSEP values of 8.160 and 7.430, and bias values of −1.440 and −1.034 for catechin and epicatechin, respectively. Notwithstanding, the RPD values for all the models were above 2.0, which are acceptable for quantitative predictions according to the literature (Saeys et al., 2005). The good fit between the different analytes content measured by HPLC and the contents predicted by the PLS models for the evaluation set samples is shown in Fig. 3.

Finally, the wavelengths with a better prediction capability according to the b vector profiles are shown in Fig. 4. The most important wavelengths related to variation in theobromine content (see Fig. 4a) are found at 1384, 1414, 1710, 1730, 1742, 1764, 1884, 1906, 1934, 2308, 2322 and 2350 nm, among others, and agree with those described by Álvarez et al. (2012). Among these wavelengths, 1414, 1730, 1764, 1906, 1934, 2308 and 2350 nm increased as theobromine content rose, while the rest lowered. According to literature (Veselá et al.,

Table 3
Results of the PLS models constructed for the prediction of theobromine, catechin, caffeine and epicatechin contents in cocoa powders.

Compound	Pre-treatment	#LV	Calibration		Cross-validation		Prediction			
			R^2_C	RMSEC	R^2_{CV}	RMSECV	R^2_P	RMSEP	bias	RPD
Theobromine (g/100g)	Non-pretreated	4	0.601	0.111	0.520	0.123	0.160	0.143	−0.027	1.22
	2nd Der. S-G	6	0.976	0.027	0.774	0.084	0.780	0.073	0.008	2.07
	OSC	4	0.923	0.048	0.870	0.064	0.793	0.209	0.027	1.38
	2nd Der. S-G + OSC	*3*	*0.969*	*0.030*	*0.889*	*0.058*	*0.819*	*0.068*	*0.005*	*2.13*
Catechin (mg/100g)	Non-pretreated	6	0.510	11.430	0.142	15.460	0.552	13.110	−1.463	1.25
	2nd Der. S-G	2	0.290	13.722	0.170	15.212	0.311	16.272	−1.962	1.01
	OSC	7	0.950	3.680	0.820	7.125	0.158	17.986	1.829	0.91
	2nd Der. S-G + OSC	*1*	*0.997*	*0.787*	*0.932*	*4.349*	*0.830*	*8.160*	*−1.440*	*2.01*
Caffeine (g/100g)	Non-pretreated	7	0.832	0.026	0.763	0.031	N.A	0.058	0.026	1.09
	2nd Der. S-G	5	0.941	0.015	0.763	0.031	0.775	0.025	0.008	2.55
	OSC	3	0.911	0.019	0.869	0.023	0.031	0.051	−0.028	1.22
	2nd Der. S-G + OSC	*1*	*0.940*	*0.015*	*0.802*	*0.028*	*0.813*	*0.022*	*0.007*	*2.81*
Epicatechin (mg/100g)	Non-pretreated	7	0.900	5.210	0.433	12.924	0.690	9.899	−4.690	1.58
	2nd Der. S-G	5	0.972	2.740	0.295	14.412	0.644	10.592	−3.024	1.57
	OSC	1	0.999	0.050	0.999	12.968	0.470	12.970	−0.480	1.29
	2nd Der. S-G + OSC	*6*	*0.999*	*0.262*	*0.988*	*1.810*	*0.824*	*7.430*	*−1.034*	*2.24*

2nd Der. S-G = Second derivative-Savitzky Golay; OSC = Orthogonal signal correction; #LV = latent variables; R^2_C = coefficient of determination for calibration; RMSEC = root mean square error of calibration; R^2_{CV} = coefficient of determination for cross-validation; RMSECV = root mean square error of cross-validation; R^2_P = coefficient of determination for prediction; RMSEP = root mean square error of prediction; RPD = ratio of prediction deviation; N.A. = non available.

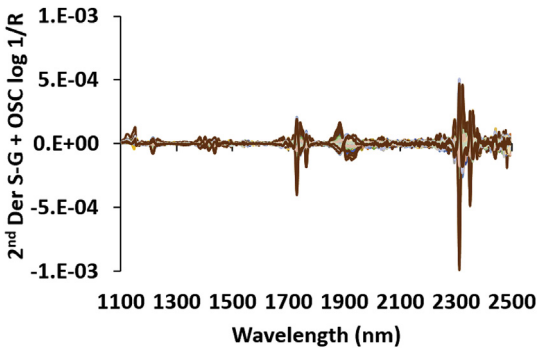

Fig. 2. Spectra of the cocoa powders obtained from after applying 2nd derivative S-G + OSC pre-treatment.

2007), these variations are characterised mainly by the stretching of H_2O of weakly bounded water, proteins and aromatics, the first overtone of stretching of CH of aromatics, the first overtone of the symmetric and asymmetric stretching vibration of CH_2, and the stretching and rocking vibrations of CH_2 of polysaccharides and fats.

For catechin content predictions, the most important wavelengths were 1414, 1440, 1730, 1744, 1880, 1908, 2312, 2322, 2350 and 2360 nm (see Fig. 4b). Among them, 1414, 1744, 1908, 2322 and 2360 nm increased as catechin content rose, while the rest lowered. These variations may be associated with the stretching of H_2O of weakly bounded water, proteins and aromatics and also of non-bounded water, the asymmetric stretching and rocking of H_2O of very weakly bounded water fat, asymmetric stretching and rocking of CH_2 of polysaccharides and the stretching and rocking of CH and CC (Veselá et al., 2007).

In the case of caffeine (see Fig. 4c), the most important wavelengths are located at 1414, 1646, 1940, 2272, 2312 and 2350 nm (which increase as the caffeine content raised) and the wavelengths at 1214, 1440, 1730, 1882, 2250, 2322, and 2360 nm (which decrease as the caffeine content raised). In this case, the wavelengths corresponded to the stretching of H_2O (non-bounded water), the first overtone of asymmetric stretching vibration of CH_2, proteins, combination of amides, the stretching of CH and CC and the stretching and rocking of CH_2 of fats (Krähmer et al., 2015).

Finally, and as shown in Fig. 4d for epicatechin prediction, the most important wavelengths related to the variation of epicatechin are the 1440, 1900, 2014, 2246, 2320 and 2358 nm bands (which increase as the epicatechin content increases), and the 1922, 2040, 2176, 2266, 2312 and 2350 nm bands (which decrease as the epicatechin content raises). These bands could be assigned to the stretching of H_2O (non-bounded water), the asymmetric stretching and rocking of H_2O of very weakly bounded water fat, proteins and the stretching of CH and C═C (Veselá et al., 2007). Similar NIR chemical vibrations has been found in the quantification of total phenols and carotenoids in blackberries (Toledo-Martín et al., 2018), and in the determination of the fat, caffeine, theobromine of sun dried cocoa beans (Álvarez et al., 2012).

4. Conclusions

The HPLC determination of methylxanthines and flavanols of a large collection of cocoa powders provided not only better knowledge of the concentration variability in natural cocoas from different origins, but also the understanding of the effect that industrial alkalisation has on those contents. A wide fluctuation in flavanol content was observed for natural powders, which highlights the natural variability of unprocessed samples given their different origins. Moreover, the effect of

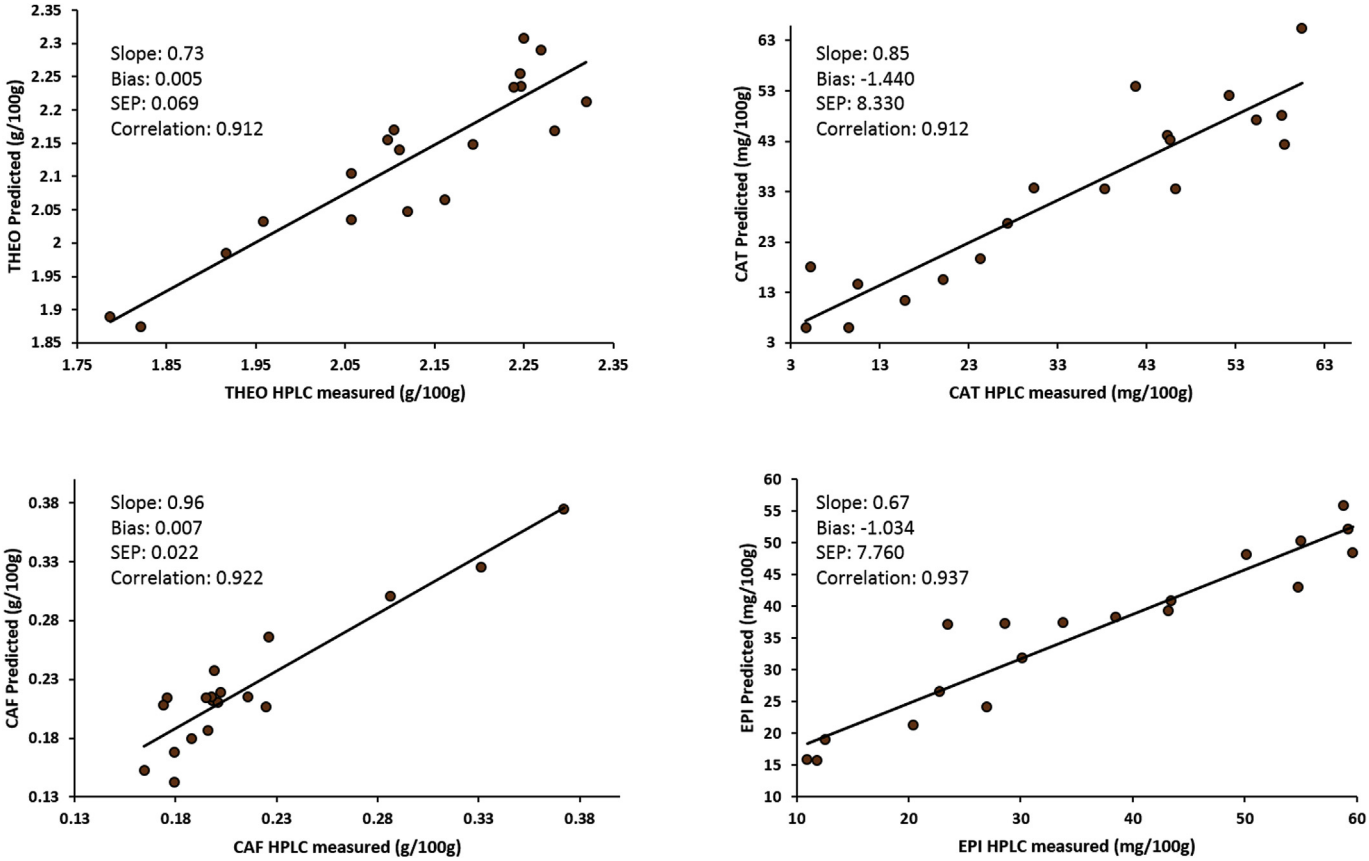

Fig. 3. HPLC measured *versus* NIR predicted (a) theobromine (THEO), (b) catechin (CAT), (c) caffeine (CAF) and (d) epicatechin (EPI) contents by PLS in the prediction set.

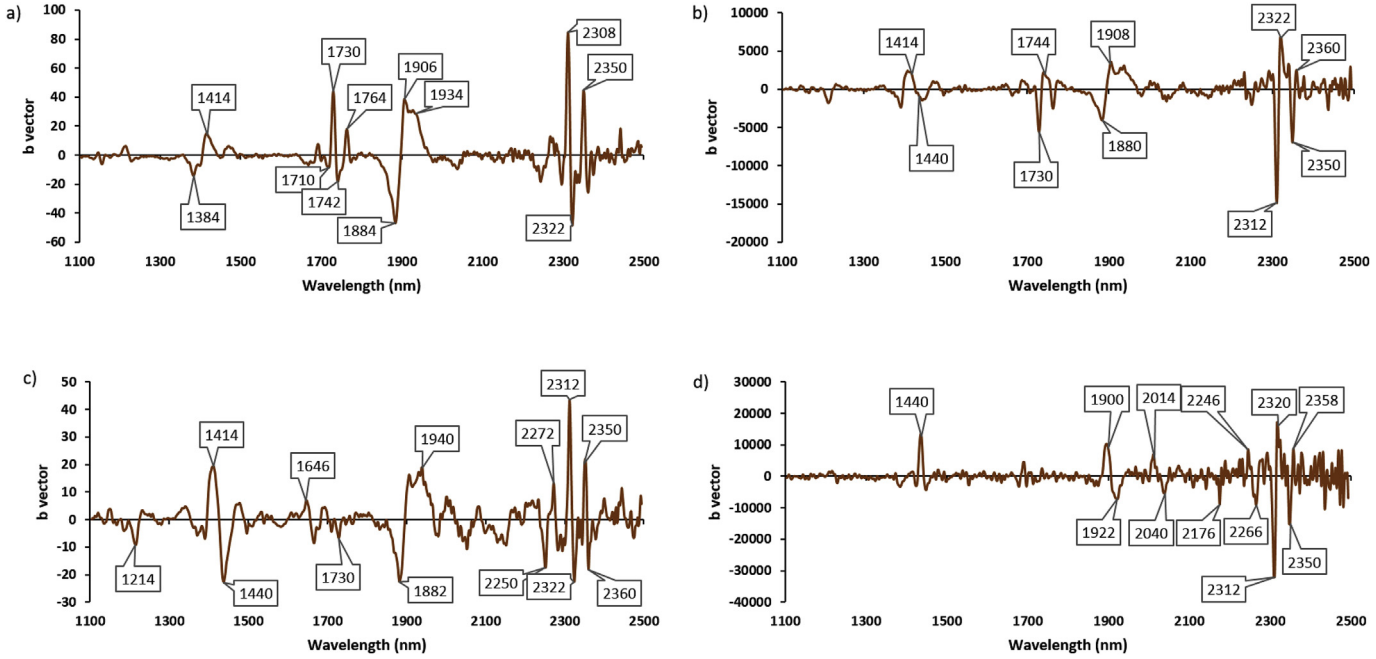

Fig. 4. B vector profiles of the PLS models constructed to predict (a) theobromine, (b) catechin, (c) caffeine and (d) epicatechin contents.

cocoa alkalisation on the content of all analytes was evidenced. Despite such evidence, the many analysed samples indicated strong-alkalised powders with higher analyte contents than some natural cocoa powders. This reinforces the importance of measuring the content of these analytes during raw material selection and in all the industrial processing steps when functional products want to be launched on the market. In line with this, the possibility of predicting the content of these functional analytes by a fast, non-destructive and reliable methodology, such as NIRS, was confirmed. Despite the fact that the clustering analysis did not allowed a sample grouping according to the alkalisation degree or the flavanols and/or methylxantines content by applying PLS models, all analytes were satisfactorily predicted. Therefore, the present results bridge the information gap in the cocoa sector about the variability found in these functional compounds in commercial samples, and also propose a fast reliable methodology to establish the content of these important functional compounds for the cocoa industry.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to acknowledge the financial support of the Spanish Government and European Regional Development Fund (Project RTC-2016-5241-2). M. A. Quelal thanks the Ministry of Higher Education, Science, Technology and Innovation (SENESCYT) of the Republic of Ecuador for her PhD grant. Olam Food Ingredients Company is aknowledged for proving part of the cocoa samples used in the study.

References

ADM Cocoa Manual (2006). *The De Zaan® cocoa manual.* The Netherlands: Archer Daniels Midland Company BV.

Afoakwa, E. O., Paterson, A., Fowler, M., & Ryan, A. (2008). Flavor formation and character in cocoa and chocolate: A critical review. *Critical Reviews in Food Science and Nutrition, 48*(9), 840–857.

Álvarez, C., Pérez, E., Cros, E., Lares, M., Assemat, S., Boulanger, R., et al. (2012). The use of near infrared spectroscopy to determine the fat, caffeine, theobromine and (–)-epicatechin contents in unfermented and sun-dried beans of Criollo cocoa. *Journal of Near Infrared Spectroscopy, 20*(2), 307–315.

Bázár, G., Romvári, R., Szabó, A., Somogyi, T., Éles, V., & Tsenkova, R. (2016). NIR detection of honey adulteration reveals differences in water spectral pattern. *Food Chemistry, 194*, 873–880.

Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods, 6*(9), 2812–2831.

Brunetto, M. del R., Gutiérrez, L., Delgado, Y., Gallignani, M., Zambrano, A., Gomez, A., et al. (2007). Determination of theobromine, theophylline and caffeine in cocoa samples by a high-performance liquid chromatographic method with on-line sample cleanup in a switching-column system. *Food Chemistry, 100*(2), 459–467.

Cádiz-Gurrea, M. L., Lozano-Sanchez, J., Contreras-Gámez, M., Legeai-Mallet, L., Fernández-Arroyo, S., & Segura-Carretero, A. (2014). Isolation, comprehensive characterization and antioxidant activities of Theobroma cacao extract. *Journal of Functional Foods, 10*, 485–498.

Elwers, S., Zambrano, A., Rohsius, C., & Lieberei, R. (2009). Differences between the content of phenolic compounds in Criollo, Forastero and Trinitario cocoa seed (Theobroma cacao L.). *European Food Research and Technology, 229*(6), 937–948.

Fayeulle, N., Vallverdu-Queralt, A., Meudec, E., Hue, C., Boulanger, R., Cheynier, V., et al. (2018). Characterization of new flavan-3-ol derivatives in fermented cocoa beans. *Food Chemistry, 259*, 207–212.

Franco, R., Oñatibia-Astibia, A., & Martínez-Pinilla, E. (2013). Health benefits of methylxanthines in cacao and chocolate. *Nutrients, 5*(10), 4159–4173.

Gottumukkala, R. V. S. S., Nadimpalli, N., Sukala, K., & Subbaraju, G. V. (2014). Determination of catechin and epicatechin content in chocolates by high-performance liquid chromatography. *International Scholarly Research Notices,* 1–15 2014.

Hue, C., Gunata, Z., Bergounhou, A., Assemat, S., Boulanger, R., Sauvage, F. X., et al. (2014). Near infrared spectroscopy as a new tool to determine cocoa fermentation levels through ammonia nitrogen quantification. *Food Chemistry, 148*, 240–245.

Humston, E. M., Knowles, J. D., McShea, A., & Synovec, R. E. (2010). Quantitative assessment of moisture damage for cacao bean quality using two-dimensional gas chromatography combined with time-of-flight mass spectrometry and chemometrics. *Journal of Chromatography A, 1217*(12), 1963–1970.

International Conference on Harmonization (ICH guidelines) (1996). *Validation of analytical procedures: Text and methodology. Geneva: ICH-Q2.*

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

Krähmer, A., Engel, A., Kadow, D., Ali, N., Umaharan, P., Kroh, L. W., et al. (2015). Fast and neat – determination of biochemical quality parameters in cocoa using near infrared spectroscopy. *Food Chemistry, 181*, 152–159.

Lacueva, C. A., Monagas, M., Khan, N., Izquterdo-Pulido, M., Urpi-Sarda, M., Permanyer, J., et al. (2008). Flavanol and flavonol contents of cocoa powder products: Influence of the manufacturing process. *Journal of Agricultural and Food Chemistry, 56*(9), 3111–3117.

Langer, S., Marshall, L. J., Day, A. J., & Morgan, M. R. (2011). Flavanols and methylxanthines in commercially available dark chocolate: A study of the correlation with nonfat cocoa solids. *Journal of Agricultural and Food Chemistry, 59*(15), 8435–8441.

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*(1), 17–23.

Machonis, P., Jones, M., Schaneberg, B., Kwik-Uribe, C., & Dowell, D. (2014). Method for the determination of catechin and epicatechin enantiomers in cocoa-based ingredients and products by high-performance liquid chromatography: First action 2013.04. *Journal of AOAC International, 97*(2), 506–509.

Miller, K. B., Hurst, W. J., Payne, M. J., Stuart, D. A., Apgar, J., Sweigart, D. S., et al. (2008). Impact of alkalization on the antioxidant and flavanol content of commercial cocoa powders. *Journal of Agricultural and Food Chemistry, 56*(18), 8527–8533.

Onatibia-Astibia, A., Franco, R., & Martínez-Pinilla, E. (2017). Health benefits of methylxanthines in neurodegenerative diseases. *Molecular Nutrition & Food Research, 61*(6), 1600670.

Osborne, B. G., Fearn, T., & Hindle, P. H. (1993). *Practical NIR spectroscopy with applications in food and beverage analysis. Burnt Mill* (2nd ed.). Harlow, Essex, England: Longman Group.

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*(19), 10518–10527.

Pérez-Esteve, É., Lerma-García, M. J., Fuentes, A., Palomares, C., & Barat, J. M. (2016). Control of undeclared flavoring of cocoa powders by the determination of vanillin and ethyl vanillin by HPLC. *Food Control, 67*, 171–176.

Quelal-Vásconez, M. A., Lerma-García, M. J., Pérez-Esteve, É., Arnau-Bonachera, A., Barat, J. M., & Talens, P. (2019). Fast detection of cocoa shell in cocoa powders by near infrared spectroscopy and multivariate analysis. *Food Control, 99*, 68–72.

Quelal-Vásconez, M. A., Pérez-Esteve, É., Arnau-Bonachera, A., Barat, J. M., & Talens, P. (2018). Rapid fraud detection of cocoa powder with carob flour using near infrared spectroscopy. *Food Control, 92*, 183–189.

Risner, C. H. (2008). Simultaneous determination of theobromine, (+)-Catechin, caffeine, and (-)-Epicatechin in standard reference material baking chocolate 2384, cocoa, cocoa beans, and cocoa butter. *Journal of Chromatographic Science, 46*(10), 892–899.

Saeys, W., Mouazen, A. M., & Ramon, H. (2005). Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy. *Biosystems Engineering, 91*(4), 393–402.

Savitzky, A., & Golay, M. (1951). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry, 40*(42).

Srdjenovic, B., Djordjevic-Milic, V., Grujic, N., Injac, R., & Lepojevic, Z. (2008). Simultaneous HPLC determination of caffeine, theobromine, and theophylline in Food, drinks, and herbal products. *Journal of Chromatographic Science, 46*(2), 144–149.

Sunoj, S., Igathinathane, C., & Visvanathan, R. (2016). Nondestructive determination of cocoa bean quality using FT-NIR spectroscopy. *Computers and Electronics in Agriculture, 124*, 234–242.

Teye, E., Huang, X., Dai, H., & Chen, Q. (2013). Rapid differentiation of Ghana cocoa beans by FT-NIR spectroscopy coupled with multivariate classification. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 114*, 183–189.

Toledo-Martín, E. M., García-García, M. C., Font, R., Moreno-Rojas, J. M., Salinas-Navarro, M., Gómez, P., et al. (2018). Quantification of total phenolic and carotenoid content in blackberries *(Rubus fructicosus L.)* using near infrared spectroscopy (NIRS) and multivariate analysis. *Molecules, 23*(12), E3191.

Toro-Uribe, S., Montero, L., López-Giraldo, L., Ibáñez, E., & Herrero, M. (2018). Characterization of secondary metabolites from green cocoa beans using focusing-modulated comprehensive two-dimensional liquid chromatography coupled to tandem mass spectrometry. *Analytica Chimica Acta, 1036*, 204–213.

Van Durme, J., Ingels, I., & De Winne, A. (2016). Inline roasting hyphenated with gas chromatography-mass spectrometry as an innovative approach for assessment of cocoa fermentation quality and aroma formation potential. *Food Chemistry, 205*, 66–72.

Vergara-Barberan, M., Lerma-García, M. J., Herrero-Martínez, J. M., & Simó-Alfonso, E. F. (2015). Cultivar discrimination of Spanish olives by using direct FTIR data combined with linear discriminant analysis. *European Journal of Lipid Science and Technology, 117*(9), 1473–1479.

Veselá, A., Barros, A. S., Synytsya, A., Delgadillo, I., Čopíková, J., & Coimbra, M. A. (2007). Infrared spectroscopy and outer product analysis for quantification of fat, nitrogen, and moisture of cocoa powder. *Analytica Chimica Acta, 601*(1), 77–86.

Wajrock, S., Antille, N., Rytz, A., Pineau, N., & Hager, C. (2008). Partitioning methods outperform hierarchical methods for clustering consumers in preference mapping. *Food Quality and Preference, 19*(7), 662–669.

Whitacre, E., Oliver, J., Van Den Broek, R., Engelsen, P., Kremers, B., Van Der Horst, B., et al. (2003). Predictive analysis of cocoa procyanidins using near-infrared spectroscopy techniques. *Food Chemistry and Toxicology Predictive, 68*(9), 2618–2622.

Wold, S., Antti, H., Lindgren, F., & Öhman, J. (1998). Orthogonal signal correction of near-infrared spectra. *Chemometrics and Intelligent Systems, 44*, 175–185.