Food Chemistry 188 (2015) 576–582

Contents lists available at ScienceDirect

Food Chemistry

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

Multi-element, multi-compound isotope profiling as a means to distinguish the geographical and varietal origin of fermented cocoa (*Theobroma cacao* L.) beans

Didier Diomande [a], Ingrid Antheaume [a], Maël Leroux [a], Julie Lalande [a], Stéphane Balayssac [b], Gérald S. Remaud [a,*], Illa Tea [a,*]

[a] Elucidation of Biosynthesis by Isotopic Spectrometry (EBSI) Group, CEISAM, University of Nantes–CNRS UMR6230, 2 rue de la Houssinière, BP 92208, F-44322 Nantes, France
[b] Groupe de RMN Biomédicale, Laboratoire de Synthèse et Physico-Chimie de Molécules d'Intérêt Biologique, UMR CNRS 5068, Université de Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex, France

ARTICLE INFO

Article history:
Received 24 November 2014
Received in revised form 21 April 2015
Accepted 11 May 2015
Available online 12 May 2015

Keywords:
Isotope ratio monitoring by mass spectrometry
Cocoa
Geographical origin
Environment effect

ABSTRACT

Multi-element stable isotope ratios have been assessed as a means to distinguish between fermented cocoa beans from different geographical and varietal origins. Isotope ratios and percentage composition for C and N were measured in different tissues (cotyledons, shells) and extracts (pure theobromine, defatted cocoa solids, protein, lipids) obtained from fermented cocoa bean samples. Sixty-one samples from 24 different geographical origins covering all four continental areas producing cocoa were analyzed. Treatment of the data with unsupervised (Principal Component Analysis) and supervised (Partial Least Squares Discriminant Analysis) multiparametric statistical methods allowed the cocoa beans from different origins to be distinguished. The most discriminant variables identified as responsible for geographical and varietal differences were the $\delta^{15}N$ and $\delta^{13}C$ values of cocoa beans and some extracts and tissues. It can be shown that the isotope ratios are correlated with the altitude and precipitation conditions found in the different cocoa-growing regions.

1. Introduction

Cocoa beans as used in the preparation of chocolate and related products are the seeds from fruit pods of the tree *Theobroma cacao* Linné, which grows in tropical regions throughout the world (in particular Central and South America, Southeast Asia and West Africa). Although *T. cacao* is native to Central and northern South America, nearly 70% of the world crop is currently produced in West Africa. Three major varieties of cocoa plant are cultivated: Forastero, Criollo, and Trinitario. Forastero is the most widely used, comprising 95% of the world production of cocoa, and is principally cultivated in Africa, but with some production also in Central and South America. This tree grows faster and gives a higher yield than other types of cocoa. Overall, however, the highest quality cocoa beans come from the Criollo variety, which is mainly produced in Venezuela, but which gives a relatively low yield. Trinitario, a hybrid between Criollo and Forastero, is considered to be of much higher quality than the latter, but shows a better production and is more resistant to disease than is Criollo.

Before being exported, the cocoa beans undergo fermentation following a drying process. Both these steps are generally conducted as traditional, indigenous processes (Ho, Zhao, & Fleet, 2014), the details of which depend on the country of origin, but which influence the taste and flavor of the cocoa products. Characteristics of geographical origins of the fermented cocoa beans could thus be linked to both the fermentation and drying procedure and to the cocoa varieties, giving a range of economic values to the product, hence a need for adequate methods for traceability.

The processed cocoa bean is composed primarily of triacylglycerides (cocoa butter, 45–54%) and proteins (11.5%) and with sufficient levels of the xanthine alkaloids theobromine (1.2–1.8%) and caffeine (≈0.25%) for these to be readily detected and quantified. The chemical composition depends on the type of bean, the variety, and the quality of the fermentation and drying. Thus, most publications in the area of cocoa traceability have focused primarily on the analysis of the chemical composition of processed cocoa beans (Crews, 2002; Oracz, Zyzelewicz, & Nebesny, 2013) by several methods, including infrared and NMR spectroscopy

* Corresponding authors.
E-mail addresses: gerald.remaud@univ-nantes.fr (G.S. Remaud), illa.tea@univ-nantes.fr (I. Tea).

http://dx.doi.org/10.1016/j.foodchem.2015.05.040
0308-8146/

(Caligiani, Acquotti, Cirlini, & Palla, 2010; Caligiani, Palla, Acquotti, Marseglia, & Palla, 2014), HPLC (Oracz et al., 2013) and mass spectrometry (Crews, 2002; Oracz et al., 2013). These various studies have shown that it is possible to differentiate the 3 cocoa varieties (Forastero, Criollo and Trinitario) and to distinguish 3 zones of cultivation in the Ivory Coast (Aboisso, Daloa and Divo), particularly on the basis of the caffeine and theobromine content. More recently, ^{1}H NMR spectra of hydroalcoholic extracts of traded cocoa beans have been used to characterize the fermented cocoa beans as a function of the variety and geographical origin (Caligiani et al., 2014), although satisfactory discrimination was only possible after the complete assignment of the ^{1}H NMR spectra. Overall, these quantitative compositional approaches have proved inadequate as a general approach for assessing geographical origin of fermented cocoa beans.

An alternative approach, widely used in the area of food traceability, is to examine the variation in the distribution of stable isotopes within the target foodstuff. While such an approach has proved an efficient means to classify the geographic and varietal origins of a wide variety of food products: olive oil (Longobardi et al., 2012), pistachios (Anderson & Smith, 2004), coffee (Rodrigues et al., 2009), tea (Pilgrim, Watling, & Grice, 2010) wine (Camin et al., 2013), orange juice (Doner & Bills, 1981; Rummel, Hoelzl, Horn, Rossmann, & Schlicht, 2010), vanilla (John & Jamin, 2004), to our knowledge there are no studies using stable isotopes to provide information about the geographic or varietal origin of fermented cocoa beans.

This approach is based on the fact that the stable isotopes ratios of the elements (e.g. $^{2}H/^{1}H$, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{18}O/^{16}O$) of which the compounds of the organism are composed reflect characteristics of the plant's environment and physiology. Chemical, physical and biological processes can all cause significant fractionation in isotope ratios. The distributions of isotopes in organic matter are a function of photosynthetic fixation, temperature, plant type (C_3 vs. C_4 plants) (Farquhar, Ehleringer, & Hubick, 1989) and/or environment (latitude) (Farquhar et al., 1989). Plants in humid environments, taking in more atmospheric CO_2 (Farquhar et al., 1989), tend to develop a lower $^{13}C/^{12}C$ ratio than plants in drier environments, wherein internal recycling is prevalent. Many chemical processes, such as denitrification and mineralization, affect nitrogen isotopic composition. Climate and ecosystem variations such as soil types and annual temperatures and precipitation have been reported to affect nitrogen isotope ratios (Amundson et al., 2003). ^{15}N signatures provide also information about regional agriculture practices. Similarly, ^{18}O and ^{2}H isotope signatures reflect water status during synthesis (Schmidt, Robins, & Werner, 2015).

The advantage of multi-element, multi-compound isotope profiling is that a number of parameters are simultaneously used to detect external influences, which can hence reinforce their overall contribution to highlight influences that relate to the growth area and growing conditions of the plant. Furthermore, the approach generates large data sets that are amenable to statistical treatment. Typically, multivariate statistics, which involves the simultaneous observation and analysis of more than two statistical variables is used. In the context of distinguishing varietal and geographical differences, two widely used multivariate methods are used for exploratory data analysis through dimensional reduction and data visualization: the non-supervised Principal Component Analysis (PCA), and the supervised Partial Least Squares-Discriminant Analysis (PLS-DA). These have been exploited to assess the potential for stable isotope ratios measured by isotope ratio monitoring by mass spectrometry (irm-MS) to provide information about the geographical and varietal origin of fermented cocoa beans.

2. Materials and methods

2.1. Samples

Cocoa beans from the 2008 to 2010 harvest periods from 24 different geographical origins, giving in total 61 samples (Table SI1), were provided from CIRAD or were obtained directly from growers in the Ivory Coast.

2.2. Chemicals

Sulfuric acid, chloroform, and cyclohexane were purchased from VWR Prolabo, ammonia solution (25% v/v) from Merck, propan-2-ol from Fluka and ethanol (99.9%) from Docks Des Alcools (France).

2.3. Extraction and separation of the fractions studied

The extraction method for the separation of each sub-matrix from cocoa bean is schematically illustrated in Fig. 1. The first three matrices are the whole beans, the shells and the cotyledons. To obtain a lipid fraction, cotyledons (10 g) were ground and extracted under reflux with cyclohexane (100 ml) for 2 h at 90 °C. After cooling and filtration, the lipids were recovered by evaporation at reduced pressure on a rotary evaporator. Residual cyclohexane was removed by using a vacuum pump. The solid residue insoluble in cyclohexane obtained after filtration consisted of fat-free cocoa powder.

Theobromine was extracted from the cocoa powder in two steps. The fat-free cocoa powder (5 g) was submitted to solid–liquid extraction by sulfuric acid (150 ml, 2 M) at reflux (110 °C) during 1 h. Following cooling, and filtration, a violet colored solution and a solid residue were obtained. Following filtration, the solution was basified with ammonia (28% v/v) until a brown solution was obtained. This was subjected to liquid–liquid extraction with chloroform/isopropanol (3:1 v/v). After evaporation of the solvent and an ethanol wash to remove impurities, a white powder was obtained after drying in an oven 55 °C which was shown by spectroscopic analysis to be principally theobromine. The residue was primarily composed of proteins.

2.4. Elemental analysis (EA)

Elemental analysis was carried out in an elemental analyzer NA2100 (www.thermo.com). For C and N elemental analysis, each fraction of cocoa, obtained according to the analytical procedure (Fig. 1) was weighed with 10^{-6} g precision (cocoa bean, cotyledon, defatted cocoa, shell, protein: 1 mg; lipids sub-matrix: 0.8 mg; theobromine: 0.4 mg) into tin capsules and crimped. The capsules were individually introduced into a combustion furnace (1020 °C) in an excess of oxygen. Chromium oxide and silvered cobalt (II, III) oxide were used as oxidation catalysts and He as carrier gas. Water was removed with a magnesium perchlorate trap. Reduction of NO_x and removal of excess O_2 was achieved with reduced copper. Calibration was with glutamic acid (GA) (%C = 40.78; %N = 9.52). C and N percentage (w/w) were calculated based on the initial weight (mg) of the sample.

2.5. Isotope ratio monitoring by mass spectrometry (irm-MS)

The global isotope compositions, $\delta^{13}C$ and $\delta^{15}N$, were determined using a Delta-V Advantage isotope ratio mass spectrometer (www.thermo.com) coupled to an NA2100 elemental analyzer (irm-EA/MS). Compound (typically 0.8–1.0 mg) was encapsulated

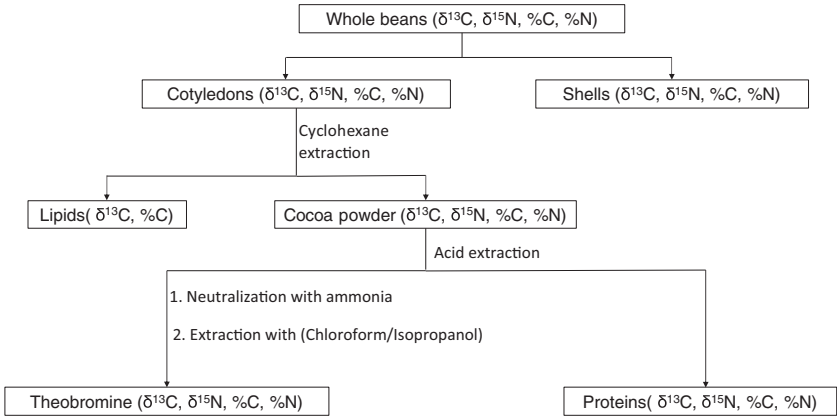

Fig. 1. Scheme showing the separation/extraction of each sub-matrix from cocoa beans, as used in the present study.

in a tin capsule, and combusted in an EA as described. The isotope composition of the resulting gases, CO_2 and N_2, was determined by reference to a working standard of GA standardized against calibrated international reference material (IAEA CH_6 and IAEA CH_7 for carbon isotope ratio and IAEA N1 for nitrogen isotope ratio, International Atomic Energy Agency, Austria). Analytical performance was checked by inserting laboratory standards of GA ($^{13}C = -27.30‰$, (-0.45 as correction factor); $\delta^{15}N = 4.85‰$, (-0.14 as correction factor)) between samples to check for stability and to allow drift correction to be made when necessary.

2.6. Statistical methods

First, a package developed in the R environment was used for a univariate approach based on analysis of variance for each variable, and discriminating variables were uncovered through a supervised univariate approach with *t*-tests and boxplots (Balayssac, Déjean, Lalande, Gilard, & Malat-Martino, 2013). After mean-centering and auto-scaling, the data matrix was subjected to several multivariate statistical analyses using SIMCA-P+ 12.0 software (Umetrics, Umeå, Sweden). PCA was used first to provide an unsupervised analysis that requires no prior knowledge of the data set, and condenses the multivariate data into a reduced number of variables, known as principal components, which describe the greatest amount of variance. PCA is useful for the detection of outliers and for finding patterns and trends. However, it cannot assign class membership to unknown test samples. To narrow the results of PCA, we used a sparse version of PCA (SPCA) which allows statistical deletion of data considered uninteresting in the PCA. The SPCA thus makes it possible to select the most discriminating variables for the first two axes. In these analyses, we set the number of variables selected at 5 and, to the extent that we consider two axes, we have thus 10 variables (Balayssac et al., 2013). The dataset was then handled using a supervised method such as projections to latent structure by PLS-DA. PLS-DA uses prior class information (e.g. geographical origin) to optimize the separation between sample groups, focusing on differences according to sample variations (Szymańska, Saccenti, Smilde, & Westerhuis, 2012).

The predictive ability of the PLS-DA models was validated using the values of Q^2_{cum}, R^2Y and R^2X parameters. The R^2 score represents the variance of X (data) and Y (class of each group) which is variable in the model. The Q^2 score is a cross-validated R^2 which reflects the model's potential for predicting class membership (Roy & Mitra, 2011). The statistical significance of the R^2Y and Q^2 parameters was estimated through the response permutation test, where the Y matrix is randomly permuted 999 times when the X matrix is fixed (Lalande et al., 2014). Finally, a *p*-value was generated from a cross-validated analysis of variance (CV-ANOVA) (Lalande et al., 2014). Loading plot and variable importance in the projection (VIP) from PLS-DA models were used to identify the variables driving the separation between classes, i.e. the metabolites that can be potential biomarkers. In this study, four criteria were used to validate the PLS-DA models (i) $Q^2 > 0.5$, (ii) CV-ANOVA < 0.05, (iii) response permutation test, and (iv) visual analysis of the boxplots and $p < 0.05$ for the data values arising from the coefficient plot and VIP. The PLS-DA models were considered predictive if 3 out of these 4 criteria were met.

Statistical correlations based on Pearson correlation coefficients were performed on isotope ratio. This coefficient captures linear dependency between a variable and two parameters (the altitude and the precipitation) providing a statistical indicator (*p*-value) and relevant clues on the impact on isotope variation of environmental parameters. Correlations rang from -1 (anticorrelated) to 1 (correlated).

3. Results and discussion

3.1. Analytical protocol

In order to test the possibility of differentiating cocoa beans from different origins, we collected fermented cocoa beans from 24 different geographical origins, in total 61 samples (Table SI1). Each sample was submitted in a series to isotopic analyses by irm-EA/MS in triplicate capsule samples. The isotope signatures of whole cocoa bean samples varied from $-32.2‰$ to $-28.1‰$ for $\delta^{13}C$ and from 0.9‰ to 9.1‰ for $\delta^{15}N$. The %C and %N of whole cocoa beans varied from 54.8% to 61.2% for C and from 1.8% to 2.7% for N.

These ranges encouraged a deeper analysis of sub-fractions. Even for the same overall cocoa bean isotope signature, fractions of tissue such as cotyledons and shells, and extracted products, such as theobromine, cocoa powder (defatted cocoa), protein, and lipids could have different isotope signatures. Therefore, several sub-matrices of the cocoa beans were obtained (Fig. 1), each of which was analyzed in duplicate by irm-EA/MS (Tables SI2–SI6).

The advantage of using several parameters such as percentage N and C and $\delta^{13}C$ and $\delta^{15}N$ isotope ratios and from different extracts of fermented cocoa beans is that specific parameters could distinguish the cocoa beans from different origins. In total, 28 parameters were used for this study. The practical application of the method depends on the accuracy of the analysis. The repeatability of the determination of $\delta^{13}C$ and $\delta^{15}N$, based on 3 replicates, was <0.4‰ and of the N and C percentages was <5% for each analyzed cocoa bean and their sub-matrices.

In order to evaluate the influence of the extraction from different fermented cocoa beans on ^{13}C and ^{15}N isotope fractionation, triplicate extractions were made on three cocoa bean samples from different origins, one on each of three different days. Each extract was analyzed in triplicate by irm-EA/MS. As shown in Table 1, the standard deviation for each parameter was inferior to 0.7‰.

3.2. Isotopes ratios in relation to geographical origin

Differences between fermented cocoa beans were assessed using PCA on parameters measured on whole beans and their various sub-fractions. In our study, PCA quantifies the amount of useful information that is contained in the data of the samples collected from different geographical origins (Africa, North and South America, Asia, and Oceania). The representation of all samples and 28 variables on the plot of the two first components indicated that 46% of variance is accounted for. In order to select the variables that contribute the most to the variance in the data set, an analysis by SPCA was performed with the first two components and the five variables of the most variance, obtaining a total of 10 variables: $\delta^{13}C$ values of proteins, cotyledons, defatted cocoa, shells and theobromine, and $\delta^{15}N$ values of proteins, cotyledons, defatted cocoa, shells and cocoa beans. SPCA indicated a variability from 85% on the score plot (Fig. 2A). Moreover, the score plot of PCA showed dominant trends with two clusters: the first corresponds to Africa and the second to the others countries. The influence of the most discriminating five variables can be seen on the loading plot (Fig. 2B). These observations were corroborated by the supervised univariate approach (p-value <0.05) and the boxplots illustrated in Fig. 3. The $\delta^{15}N$ values of proteins, cotyledons, defatted cocoa, shells and cocoa beans could discriminate the geographical origin of cocoa and are higher in Africa compared with others countries (Fig. 3).

The majority of samples were obtained from Africa, and of these, the majority from different towns of the Ivory Coast. When the number of samples from each town is superior to 5, comparison between them is statistically possible. Thus, three

Table 1
Experimental values and standard deviation (SD) of isotope compositions $\delta^{13}C$ and $\delta^{15}N$ (‰) and percentages of C and N (%), obtained over three repetitions of the extraction procedure described in Fig. 1, for each sub-matrix of cocoa bean samples.

Matrix	Whole cocoa bean	Cotyledon	Shell	Lipids	Defatted cocoa	Theobromine	Proteins
$\delta^{15}N$	5.4	5.5	4.4	–	4.9	3.4	6.1
	5.3	5.3	3.8	–	4.3	3.5	5.1
	5.3	5.4	4.0	–	4.6	3.4	5.6
SD (‰)	0.1	0.1	0.3	–	0.3	0.1	0.5
$\delta^{13}C$	−29.1	−29.1	−27.8	−30.5	−27.6	−27.0	−29.2
	−29.0	−29.3	−27.8	−30.6	−27.9	−27.0	−29.1
	−29.0	−29.3	−27.8	−30.6	−27.8	−27.0	−29.1
SD (‰)	0.1	0.1	0.0	0.0	0.2	0.0	0.1
%N	2.6	2.3	1.5	–	4.9	30.4	2.2
	2.5	2.5	1.8	–	4.9	30.4	2.2
	2.6	2.5	1.6	–	4.9	30.4	2.1
SD (%)	0.0	0.2	0.2	–	0.0	0.0	0.1
%C	58.6	61.6	40.6	75.9	46.0	45.3	28.7
	58.6	59.9	39.7	75.8	42.2	45.6	29.0
	58.5	59.4	40.2	75.8	45.6	45.4	28.9
SD (%)	0.1	1.1	0.5	0.1	2.2	0.2	0.2

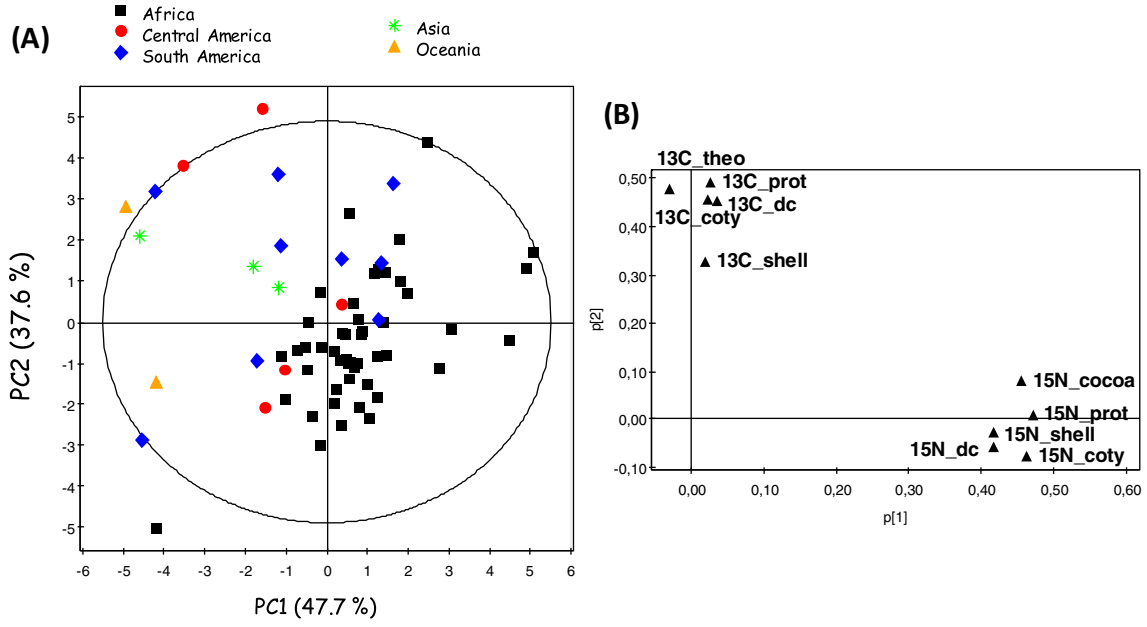

Fig. 2. Score plot (A) and loading plot (B) of the spare principal analysis component SPCA along the two principal axes, using $\delta^{13}C$ values of proteins, cotyledons, defatted cocoa, shells and theobromine and $\delta^{15}N$ values of proteins, cotyledons, defatted cocoa, shells and cocoa beans. Color coding is: Africa, black; Central America, red; South America, blue; Asia, green; Oceania, orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

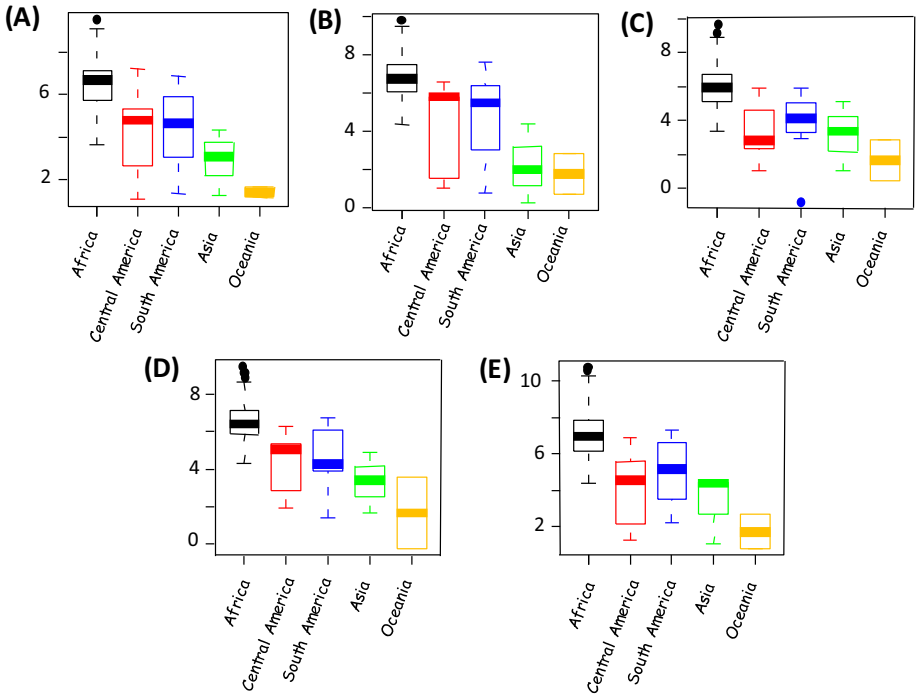

Fig. 3. Boxplots of (A) $\delta^{15}N$ of cocoa beans, (B) $\delta^{15}N$ of cotyledons, (C) $\delta^{15}N$ of shells, (D) $\delta^{15}N$ of defatted cocoa and (E) $\delta^{15}N$ of the protein fraction. Color coding as for Fig. 2. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

towns were selected (Abidjan, Gagnoa and Oumé) in order to investigate the discriminating parameters. The supervised PLS-DA analysis led to a model with 2 principal PLS components with a good predictive value ($Q^2 = 0.63$ and $R^2Y = 0.77$) which indicates its robustness. The first PLS component clearly separated Abidjan and Gagnoa and the second PLS separated Oumé from the others towns (Fig. 4A). The *p*-values were obtained when the samples were compared between the three towns. The most significant parameters were: $\delta^{15}N$ of theobromine ($p < 0.03$), $\delta^{15}N$ of defatted cocoa ($p < 0.01$), $\delta^{13}C$ of proteins ($p < 0.01$), $\delta^{13}C$ of defatted cocoa ($p < 0.01$), $\delta^{13}C$ of lipids ($p < 0.04$) (Fig. 4B).

3.3. Isotopes ratios in relation to cocoa variety

Significant differences between the two varieties (Forastero and Trinitario) were revealed by PLS-DA (Fig. 5) of all samples. The validation parameters ($Q^2 = 0.67$; $R^2X = 0.50$; $R^2Y = 0.80$), the *p*-value of the CV-ANOVA ($3.9.10^{-10}$) and response permutation test (Fig. 5D), clearly indicate the robustness of the model. The most discriminating variables are represented in bold and italics on the loading plot of PLS-DA (Fig. 5B), in the VIP (Fig. 5C) and by univariate statistical tests (data not shown). Compared with Trinitario, the Forastero variety had significantly higher $\delta^{15}N$ values for cocoa beans and their extracts. However, these findings could be due to geographical rather than varietal differences in the cocoa, since the majority of Forastero cocoa is cultivated in Africa, while the variety Trinitario is found in America and Asia. In order to be sure that the observed differences come from either geographical or varietal origin of cocoa beans, the same variety should be cultivated in the same localities in Africa and in America and Asia. Nonetheless, the isotopic analyses clearly distinguish the different producing areas.

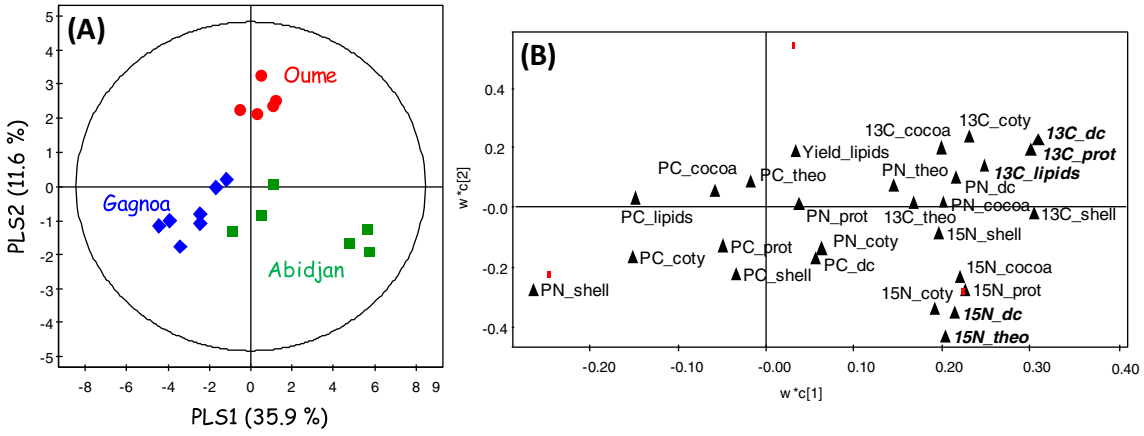

Fig. 4. Statistical analyses of the isotopic data of the three city of Ivory Coast: Abidjan, Gagnoa and Oumé: (A) score plot of the PLS-DA along the two principal PLS axes and (B) loading plot of the PLS-DA. The principal discriminating metabolites are: $\delta^{15}N$ of theobromine and defatted cocoa. Statistical parameters for validation: Axes = 2; $Q^2 = 0.63$; $R^2X = 0.48$; $R^2Y = 0.77$; CV-ANOVA = 6.10^{-4}.

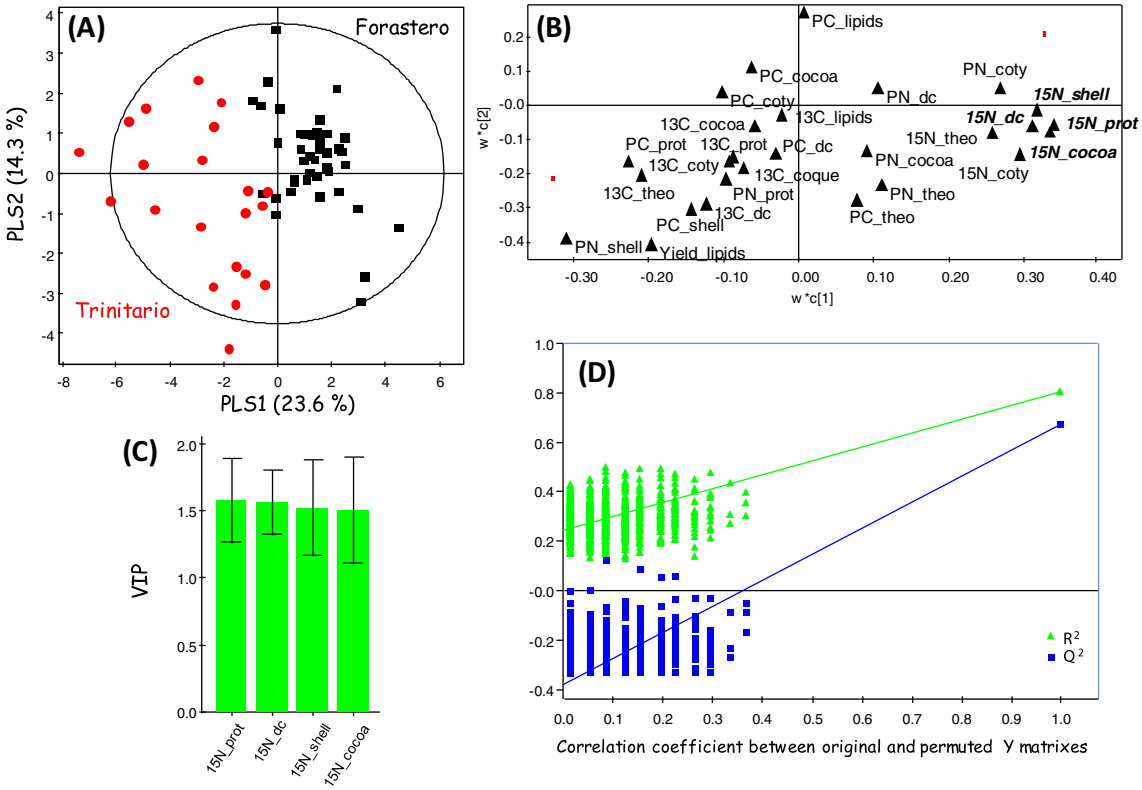

Fig. 5. Statistical analyses of the isotopic data of the variety of cocoa, Forastero and Trinitario: (A) score plot of the PLS-DA along the two principal PLS axes. (B) Loading plot of the PLS-DA. The principal discriminating metabolites are: $\delta^{15}N$ of shells, proteins, cocoa beans and defatted cocoa. (C) VIP of the discriminating parameters and (D) results from the response permutation test (999 permutations). Statistical parameters for validation: Axes = 3; $Q^2 = 0.67$; $R^2X = 0.50$; $R^2Y = 0.80$; CV-ANOVA = $3.9.10^{-10}$ (see the Section 3).

3.4. Isotopes ratios in relation to the environmental

The plant environment can affect the isotope ratios and can be responsible for observed differences. The altitude and precipitation parameters were collected for each of the studied samples. A correlation matrix was analyzed with a 28×2 matrix corresponding to pairwise correlations between the 28 variables and the two environmental parameters, altitude and precipitation (not shown). Correlation levels were around 0.3 (−0.3 when anti-correlated) and statistically significant. Good correlation was found with the altitude parameter. The N isotope signatures of shells ($r = -0.33$; $p = 0.006$) and defatted cocoa ($r = -0.39$; $p = 0.001$) were negatively correlated to altitude, whereas the N of shells ($r = 0.26$; $p = 0.03$) and the N and C percentages of proteins (%N: $r = 0.28$; $p = 0.02$ and %C: $r = 0.25$; $p = 0.04$) were positively correlated. The precipitation parameter showed less correlation, being only linked to %N of protein ($p = 0.02$).

While this aspect of the study merits further investigation, it is apparent that these two parameters could play an important role in explaining the observed N and C isotope and content signatures.

4. Conclusion

In this study it is shown that nitrogen and carbon isotope signatures obtained from cocoa beans and their different sub-matrices can be exploited to differentiate their geographical origins. $\delta^{15}N$ values of proteins, cotyledons, defatted cocoa, shells and cocoa beans discriminate samples from the different continents of Africa, America, Asia and Oceania. Similarly, $\delta^{15}N$ values of theobromine and defatted cocoa, and the $\delta^{13}C$ values of proteins, defatted cocoa and lipids can differentiate between cocoa beans from three towns in the Ivory Coast. Moreover, the isotopic parameters make it possible to distinguish between beans from different varieties, which again command different prices. Notably, high ^{15}N values of cocoa beans and their extracts are found in the Forastero variety compared with the Trinitario variety. These results are particularly pertinent, as they indicate that isotopic analysis on cocoa extracts and cocoa-containing products should also show good discrimination, since key parameters are the $\delta^{15}N$ values of theobromine and defatted cocoa powder, and the $\delta^{13}C$ values of defatted cocoa powder and of the lipids present in cocoa butter.

To some degree, the observed differences of isotope values can be explained on the basis of the altitudinal and precipitation parameters. However, an aspect still to be addressed more closely is the fermentation processes, which could also introduce isotope fractionation, due, for example to different rates of enzymatic reactions during fermentation processes and incomplete fermentation.

A particularly interesting target identified in this study is the alkaloid, theobromine. This is present in all cocoa sources at relatively high levels. Not only does it contain the elements H, C, N and O, but different atoms are derived from a number of distinct metabolic pathways. Position specific natural isotope analysis of theobromine, by ^{13}C or ^{2}H NMR should therefore be a powerful tool to obtain further parameters by which to enhance the differentiation between the origins of cocoa beans. These results will be presented elsewhere.

Acknowledgments

D. Diomande thanks the University of Nantes and the Région des Pays de Loire for funding his Ph.D. The authors thanks also CIRAD for purchasing cocoa samples. We are grateful to our colleague Richard Robins for assistance with the preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2015.05.040.

References

Amundson, R., Austin, A., Schuur, E., Yoo, K., Matzek, V., Kendall, C., et al. (2003). Global patterns of the isotopic composition of soil and plant nitrogen. *Global Biogeochemical Cycles, 17*, 1031–1041.

Anderson, K. A., & Smith, B. W. (2004). Use of chemical profiling to differentiate geographic growing origin of raw pistachios. *Journal of Agricultural and Food Chemistry, 53*, 410–418.

Balayssac, S., Déjean, S., Lalande, J., Gilard, V., & Malat-Martino, M. (2013). A toolbox to explore NMR metabolomic data sets using the R environment. *Chemometrics and Intelligent Laboratory Systems, 126*, 50–59.

Caligiani, A., Acquotti, D., Cirlini, M., & Palla, G. (2010). ^{1}H NMR study of fermented cocoa (*Theobroma cacao* L.) beans. *Journal of Agricultural and Food Chemistry, 58*, 12105–12111.

Caligiani, A., Palla, L., Acquotti, D., Marseglia, A., & Palla, G. (2014). Application of ^{1}H NMR for the characterisation of cocoa beans of different geographical origins and fermentation levels. *Food Chemistry, 157*, 94–99.

Camin, F., Bontempo, L., Perini, M., Tonon, A., Breas, O., Guillou, C., et al. (2013). Control of wine vinegar authenticity through δ^{18}O analysis. *Food Control, 29*, 107–111.

Crews, C. (2002). Authentication of cocoa butter. In M. Jee (Ed.), *Oils and fats authentication*. CRC Press [66 p.].

Doner, L. W., & Bills, D. D. (1981). Stable carbon isotope ratios in orange juice. *Journal of Agricultural and Food Chemistry, 29*, 803–804.

Farquhar, G. D., Ehleringer, J. R., & Hubick, K. T. (1989). Carbon isotope discrimination and photosynthesis. *Annual Review of Plant Biology, 40*, 503–537.

Ho, V. T. T., Zhao, J., & Fleet, G. (2014). Yeasts are essential for cocoa bean fermentation. *International Journal of Food Microbiology, 174*, 72–87.

John, T. V., & Jamin, E. (2004). Chemical investigation and authenticity of Indian vanilla beans. *Journal of Agricultural and Food Chemistry, 52*, 7644–7650.

Lalande, J., Halley, H., Balayssac, S., Gilard, V., Déjean, S., Martino, R., et al. (2014). ^{1}H NMR metabolomic signatures in five brain regions of the AβPPswe Tg2576 mouse model of Alzheimer's disease at four ages. *Journal of Alzheimer's Disease, 39*, 121–143.

Longobardi, F., Ventrella, A., Casiello, G., Sacco, D., Tasioula-Margari, M., Kiritsakis, A. K., et al. (2012). Characterisation of the geographical origin of Western Greek virgin olive oils based on instrumental and multivariate statistical analysis. *Food Chemistry, 133*, 169–175.

Oracz, J., Zyzelewicz, D., & Nebesny, E. (2013). The content of polyphenolic compounds in cocoa beans (*Theobroma cacao* L.), depending on variety, growing region and processing operations: A review. *Critical Reviews in Food Science and Nutrition, 55*, 1176–1192.

Pilgrim, T. S., Watling, R. J., & Grice, K. (2010). Application of trace element and stable isotope signatures to determine the provenance of tea (*Camellia sinensis*) samples. *Food Chemistry, 118*, 921–926.

Rodrigues, C. I., Maia, R., Miranda, M., Ribeirinho, M., Nogueira, J. M. F., & Máguas, C. (2009). Stable isotope analysis for green coffee bean: A possible method for geographic origin discrimination. *Journal of Food Composition and Analysis, 22*, 463–471.

Roy, K., & Mitra, I. (2011). On various metrics used for validation of predictive QSAR models with applications in virtual screening and focused library design. *Combinatorial Chemistry & High Throughput Screening, 14*, 450–474.

Rummel, S., Hoelzl, S., Horn, P., Rossmann, A., & Schlicht, C. (2010). The combination of stable isotope abundance ratios of H, C, N and S with ^{87}Sr/^{86}Sr for geographical origin assignment of orange juices. *Food Chemistry, 118*, 890–900.

Schmidt, H.-L., Robins, R. J., & Werner, R. A. (2015). Multi-factorial in vivo stable isotope fractionation: Causes, correlations, consequences and applications. *Isotopes in Environmental and Health Studies, 51*, 155–199.

Szymańska, E., Saccenti, E., Smilde, A., & Westerhuis, J. (2012). Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics, 8*, 3–16.