CrossMark

Simultaneous Determination of Alkaloids in Green Coffee Beans from Ethiopia: Chemometric Evaluation of Geographical Origin

Bewketu Mehari[1,4] · Mesfin Redi-Abshiro[2] · Bhagwan Singh Chandravanshi[2] ·
Minaleshewa Atlabachew[1] · Sandra Combrinck[3] · Rob McCrindle[1]

Abstract The alkaloid compositions of 99 green coffee (*Coffea arabica* L.) bean samples comprising eight varieties (Harar, Jimma, Kaffa, Wollega, Sidama, Yirgachefe, Benishangul and Finoteselam) from the major production regions of Ethiopia were investigated. High performance liquid chromatography was applied for the simultaneous determination of four coffee alkaloids in the aqueous extracts of the beans. The limits of detection for the method were established as 13 mg kg^{-1} for trigonelline, 7 mg kg^{-1} for theobromine, 8.5 mg kg^{-1} for caffeine and 4 mg kg^{-1} for theophylline in the dry coffee beans. Theophylline was not detected in any of the samples. The determined concentrations (% *w/w* dry coffee beans) ranged from 0.98 to 1.32 % for trigonelline, 0.0048 to 0.0094 % for theobromine and 0.87 to 1.38 % for caffeine. The concentrations of the alkaloids varied significantly, depending on the geographical origin of the beans. Theobromine was not detected in coffee beans from the East (Harar coffees), and its absence in samples can be used to ascertain whether the coffee originates from this region. Coffee beans from the Northwest were characterized by higher concentrations of caffeine. Application of linear discriminant analysis provided 75 % correct classification of samples into the respective production regions, with a 74 % prediction success rate. The moderate classification efficiency obtained when using alkaloid data demonstrates the potential of using this class of compounds in discriminant models for determination of the geographical origin of green coffee beans from Ethiopia.

Keywords High performance liquid chromatography · Alkaloids · Principal component analysis · Linear discriminant analysis · Coffee · Ethiopia

Introduction

The methylxanthine class and trigonelline comprise the main alkaloids in coffee. Methylxanthines are purine alkaloids, which are plant secondary metabolites derived from purine nucleotides (Crozier et al. 2006). Caffeine (1,3,7-trimethylxanthine), theobromine (3,7-dimethylxanthine) and theophylline (1,3-dimethylxanthine) occur mainly in coffee, tea, cacao, cola and mate (Crozier et al. 2006; Ashihara et al. 2008; Alonso-Salces et al. 2009). Among these, caffeine is the most abundant alkaloid in coffee (Farah et al. 2006b). The second most abundant alkaloid in coffee is trigonelline (*N*-methylpyridinium-3-carboxylate), a secondary metabolite of the coffee plant derived from pyridine (Rodrigues and Bragagnolo 2013). Although barley, corn, onion, pea, soybean and tomato (Gichimu et al. 2014) contain trigonelline, coffee is one of the most important sources of trigonelline among plant-derived foods and beverages for human consumption (Duarte et al. 2010).

The methylxanthine and trigonelline contents of green coffee beans are affected by several factors including variety,

✉ Sandra Combrinck
  combrincks@tut.ac.za

1 Department of Chemistry, Faculty of Science, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

2 Department of Chemistry, College of Natural Sciences, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

3 Department of Pharmaceutical Sciences, Faculty of Science, Tshwane University of Technology, Private Bag X680, Pretoria 0001, South Africa

4 Department of Chemistry, Faculty of Natural and Computational Sciences, University of Gondar, P.O.Box 196, Gondar, Ethiopia

genetic properties of the cultivar, maturity of the beans at harvest, harvesting method and postharvest processing conditions (fermentation, washing, drying, storage), agricultural practices (shade, pruning, fertilization), environmental factors (soil, altitude, sun exposure) and climatic parameters (rainfall, temperature) (Casal et al. 2000; Alonso-Salces et al. 2009). The caffeine content of green coffee beans depends mainly on the coffee variety and location of origin (Casal et al. 2000; Alonso-Salces et al. 2009). Concentrations of caffeine in Arabica coffee were reported to range from 0.8 to 1.4 % (w/w), but were found to be substantially higher in Robusta coffee, ranging from 1.7 to 4.0 % (Cruz et al. 2012). Concentrations of trigonelline in the range 0.74–1.54 % (w/w) were reported in green beans of Arabica coffee (Duarte et al. 2010; Gichimu et al. 2014). Farah (2012) indicated that the amount of trigonelline in Robusta coffee is about two-thirds of that present in Arabica coffee.

According to studies, the caffeine content of coffee is unaffected by the roasting process (Cruz et al. 2012). However, trigonelline undergoes thermal degradation during roasting and, on average, 50–80 % of the compound is converted to other products (Farah et al. 2006b; Kalaska et al. 2014). Nicotinic acid (niacin), a water-soluble vitamin B, forms through demethylation, while decarboxylation of trigonelline results in the formation of methyl pyridine (Perrone et al. 2008; Rodrigues and Bragagnolo 2013).

Determination of the concentrations of the methylxanthines and trigonelline in green beans is important for the coffee industry as these compounds determine the flavour and ultimately, the quality of the final coffee beverage. Firstly, the methylxanthines and trigonelline are associated with the bitterness of the coffee (Farah et al. 2006b). Trigonelline also contributes to the flavour of the beverage by acting as the precursor for different classes of aroma compounds, including the furans, pyrazines, pyrroles and pyridines, formed during coffee roasting (Ky et al. 2001; del Campo et al. 2010; Rodrigues and Bragagnolo 2013; Kalaska et al. 2014). The trigonelline content has been highly correlated with good cup quality of coffee (Farah et al. 2006b). Secondly, the consumption of coffee and the constituent methylxanthines and trigonelline have been associated with several physiological effects. Stimulation of the central nervous system (CNS), acute elevation of blood pressure, gastrointestinal noncompliance and diuresis have been ascribed to the consumption of caffeine (Alpdogan et al. 2002; Farah et al. 2006a; Al-Othman et al. 2012). Some studies have also indicated a significant correlation between high caffeine intake from coffee and the risk of spontaneous abortion (Paradkar and Irudayaraj 2002). Caffeine has been reported to have some health benefits, such as the prevention of type 2 diabetes mellitus by impairing glucose tolerance, decreasing insulin sensitivity and increasing energy expenditure and weight loss (Petrie et al. 2004; Higdon and Frei 2006). Some studies have also associated

caffeine intake with the prevention of Parkinson's disease (Ascherio et al. 2004). Theobromine and theophylline, although to a lesser degree than caffeine, have stimulatory effects on the CNS and cause diuresis, peripheral vasoconstriction, smooth muscle relaxation and myocardial stimulation (Srdjenovic et al. 2008; Rodrigues and Bragagnolo 2013). Trigonelline in consumed coffee has been associated with several health benefits. Primarily, it serves as a source of vitamin B, i.e. nicotinic acid, but the compound has also been indicated to inhibit the invasive activity of cancer cells (Hirakawa et al. 2005) and to regenerate dendrites and axons to improve memory (Tohda et al. 2005).

Several methods have been reported for the determination of methylxanthines in coffee and include separation techniques such as high performance liquid chromatography (HPLC) with ultraviolet (UV) detection (Srdjenovic et al. 2008; Al-Othman et al. 2012), electrochemical methods such as voltammetry (Svorc et al. 2012), and spectroscopic techniques including nuclear magnetic resonance spectroscopy (del Campo et al. 2010), near infrared (Huck et al. 2005) and UV-visible spectroscopy (Alpdogan et al. 2002; Lopez-Martinez et al. 2003; Belay et al. 2008). However, the chromatographic methods provide accurate results and can be used to determine individual methylxanthines in complex mixtures (Huck et al. 2005; Srdjenovic et al. 2008).

Ethiopia is widely regarded as the birth place of coffee (Kebede and Bellachew 2008; Amamo 2014). The province of Kaffa in the western part of the country is considered as the origin of Arabica coffee (Murthy and Naidu 2012). This variety represents Ethiopia's original bean and is still the only one grown in the country (Tessema et al. 2011; Amamo 2014). Currently, Ethiopia is recognized as the largest producer and exporter of Arabica coffee in Africa and among the leading ones in the world (Kufa et al. 2011). In addition, a number of Arabica coffee varieties with distinct flavour profiles are produced in different regions of the country. Factors including genetic properties, climate, altitude and soil type contribute to the unique characteristic taste of each region's coffee (Amamo 2014). Owing to this, the production region of the coffee beans plays an important role in determining the price of Ethiopian coffees in the market.

Of the coffee alkaloids, only caffeine has been investigated in Ethiopian green coffee beans (Belay et al. 2008; Alonso-salces et al. 2009). However, these studies included only limited samples, with regard to both total number and sample type, and hence, no assumptions could be made regarding the variation of caffeine among coffee beans cultivated in the different regions of the country. Hence, the aim of this study was to evaluate the diversity of methylxanthines and trigonelline in green coffee beans produced in different parts of Ethiopia and evaluate the feasibility of using the determined alkaloid concentrations as chemical descriptors of the geographical origin of the coffee beans. As to our knowledge, this

is the first report of the simultaneous determination of all four alkaloids in coffee using HPLC.

Materials and Methods

Coffee Samples

Although coffee is cultivated throughout Ethiopia, the major coffee production areas tend to be concentrated within four regions (East, Northwest, South and West) of the country (Fig. 1). Within each of these regions, representative samples were selected, and the number was based on the commercial importance of the region, in terms of both value and volume of production. Accordingly, a total of 99 samples, 26 from East, 6 from Northwest, 49 from South and 18 from West, were collected (Table 1). At least three 250 g samples of green coffee beans were collected from each of the selected areas, either directly from farms or from the Ethiopia Commodity Exchange (ECX), the country's only governmental organization through which Ethiopia trades its coffees. Samples from Northwest included Benishangul and Finoteselam coffees. These coffees are commercially less important and are exclusively produced for consumption by local communities. A number of coffee varieties are produced in each of the other coffee production regions. The western region produces three commercially important varieties, Jimma, Kaffa and Wollega

coffees, the southern region two varieties, Sidama and Yirgachefe coffees, while the eastern region produces a single variety, Harar coffee. All of the green bean samples were obtained from ripe coffee cherries treated either by washing or sun drying, depending on the practice in the areas where samples were collected. All samples were from the same harvest season and were collected throughout January and February 2014. Samples were stored in plastic bags under airtight conditions and transported to the laboratory for chemical analysis. For each sample, 50 g of the green beans were fine powdered, sieved and stored at room temperature.

Chemicals

A standard of caffeine purchased from Merck (Darmstadt, Germany); standards of theobromine and trigonelline hydrochloride obtained from Fluka (St. Gallen, Switzerland); a reference standard of theophylline, HPLC grade acetonitrile and formic acid supplied by Sigma-Aldrich (Johannesburg, South Africa); and lead acetate purchased from PAL (Manchester, UK) were used as received.

Extraction of Alkaloids

Alkaloids were extracted from the green coffee powders by using the procedure described by Srdjenovic et al. (2008). However, the method was modified by extracting twice by

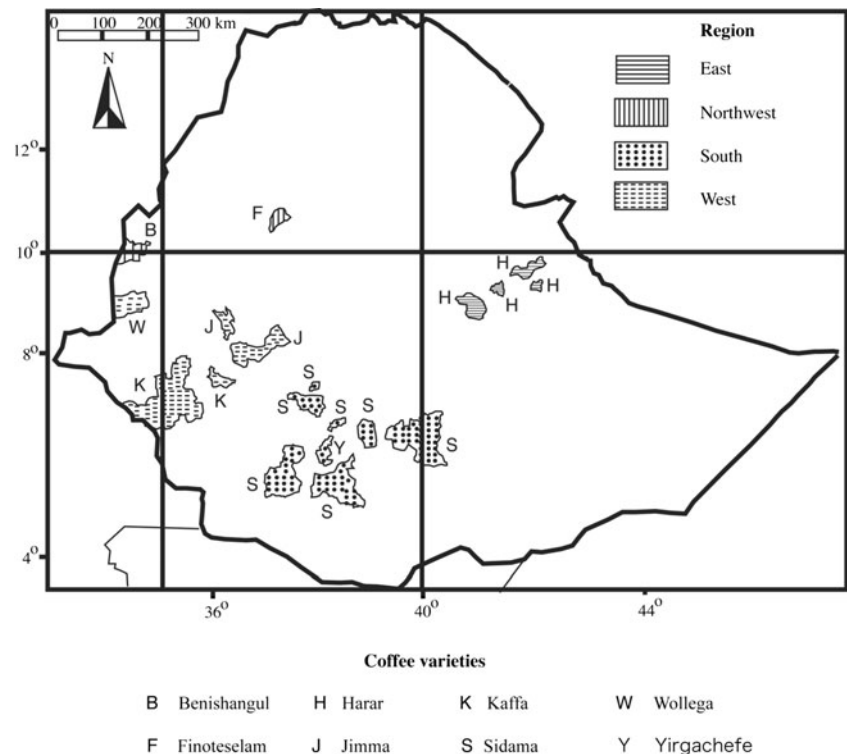**Fig. 1** Map of Ethiopia depicting the coffee sample collection regions

Table 1 Geographical regions of origin, varieties and numbers of the studied green coffee bean samples

Region	Variety	Number of samples
East	Harar	26
Northwest	Benishangul	3
	Finoteselam	3
West	Jimma	6
	Kaffa	10
	Wollega	2
South	Sidama	39
	Yirgachefe	10

shaking for 30 min, rather than once for 5 min by stirring. Briefly, a 0.200-g portion of each coffee powder, weighed directly into a nylon centrifuge tube, was extracted with 5.00 mL of boiling distilled water by shaking for 30 min on a platform shaker at 200 rpm. The mixture was centrifuged for 5 min at 3600 rpm (Dongwang, TD4K-Z, China), where after the supernatant was decanted carefully into a second centrifuge tube. The residue was returned to the tube and extracted a second time with 5.00 mL of boiling water. After combination of the supernatants, the volume was adjusted to 10.0 mL. A 1.50-mL portion of the extract was treated with 30 μL of 20 % aqueous lead acetate solution to precipitate out polysaccharides, proteins and other colloidal material from the extract solution (Perrone et al. 2008). After centrifuging the mixture for 5 min at 12,000 rpm (Hermle Labortechnik GmbH, Z216 MK, Germany), the supernatant was filtered directly into a chromatographic vial through a 0.22-μm syringe filter (Acrodisc; Sigma-Aldrich, Johannesburg, South Africa) for chromatographic analysis.

Determination of Alkaloids

An Agilent Model 1260 Infinity (Germany) HPLC system, coupled to a diode array detector (DAD), was used to determine the alkaloids in the green coffee extracts. After introduction of the sample (5 μL), separation was achieved on a reversed phase C8 column (Supelco, 15 cm×4.6 mm, 5 μm, USA) maintained at 25 °C in a column thermostat. The analysis was carried out under isocratic conditions using 90 % deionized water (acidified as 0.1 % aqueous formic acid) and 10 % acetonitrile at a flow rate of 0.3 mL min^{-1}. The mobile phase was allowed to flow for 2 min between each analysis to wash and recondition the column. Chromatographic data for caffeine, theobromine and theophylline were collected at 272 nm, whereas for trigonelline, it was collected at 264 nm. In addition, the DAD was set to collect UV-vis spectral data in the wavelength range 200–400 nm. Peak identification was performed by comparing

the retention times and the UV-vis spectral profiles of analytes to those of pure standards. Moreover, the identities were confirmed by spiking some of the coffee extracts with standard solutions of the alkaloids. For this, a 1.00-mL aliquot (20 mg mL^{-1}) of the extract was spiked with 12 μL (0.1 mg mL^{-1}) of theobromine, and 20 μL (10 mg mL^{-1}) each of trigonelline and caffeine and was subsequently analysed by HPLC. An increase in peak area for each alkaloid after standard addition supported the identification of the compounds. Quantitative determination of caffeine, theobromine, theophylline and trigonelline was done after constructing calibration curves from the chromatographic peak areas obtained after analysis of standard solutions of the respective alkaloids. All samples were analysed in triplicate from three separate extracts, and each of the triplicate samples was also analysed twice by HPLC.

The identity of theobromine was confirmed by mass spectra obtained from the analyses of the pure standard and selected coffee samples using an ultra performance liquid chromatograph (UPLC) coupled to a photo diode array (PDA) detector and a quadrupole time of flight mass spectrometer (Waters, Acquity UPLC-qTOF-MS, USA). A sample injection volume of 2 μL was used throughout. Separation was performed on a bridged ethylene hybrid C18 column (Waters, Acquity UPLC BEH C18, 1.7 μm, USA) held at 40 °C. A binary solvent system consisting of 0.1 % aqueous formic acid (solvent A) and acetonitrile (solvent B) was used in the gradient mode at a 0.3-mL min^{-1} mobile phase flow rate. The initial concentration of the mobile phase (5 % B) was increased to 10 % B over 2 min, and then to 25 % B over 6 min and finally to 50 % B over 8 min, before returning to 5 % B over 9 min. Mass spectral data were collected in the positive mode after electrospray ionization using a source temperature of 100 °C, desolvation temperature of 400 °C and capillary voltage of 3.5 kV. Argon was used as the collision gas at a collision energy of 6 eV and nitrogen as desolvation gas at 600 L h^{-1}. Theobromine was identified by comparing its retention time, UV-vis spectral profile and mass spectral characteristics to that of the pure standard.

Method Validation

The validity of the method for determining each of the alkaloids was assessed in terms of linearity, limit of detection, limit of quantitation, precision and accuracy. Stock solutions (5000 mg L^{-1} each) of caffeine, theophylline and trigonelline were prepared by dissolving 50 mg of the pure standards in 10 mL of distilled water. A stock solution (300 mg L^{-1}) of theobromine was prepared by dissolving 30 mg of the standard in 100 mL of distilled water. By diluting the stock solutions, 10 standards were prepared with concentrations in the range of 0.25–500 mg L^{-1} for caffeine, theophylline and trigonelline and in the range of 0.25–300 mg L^{-1} for

theobromine. Calibration curves were constructed from the chromatographic peak areas obtained after analysis of the standard solutions of the respective alkaloids. The linearity of the calibration curves was evaluated based on the magnitude of the coefficient of determination (R^2). The limit of detection (LOD) and limit of quantitation (LOQ) of the method for each alkaloid were calculated as three times and ten times, respectively, the standard deviation of the background signal from nine measurements divided by the slope of the calibration equation. The precision of the method was evaluated by analysing samples in triplicates from three separate extracts, each of the extracts were also analysed twice by HPLC, and calculating the relative standard deviation (RSD) for each sample. The accuracy of the analytical procedure was evaluated by recovery studies. For this, 0.200 g of a powdered green coffee sample was spiked with 80, 100 or 120 μL of 0.1 mg mL^{-1} theobromine and 135, 165 or 200 μL each of 10 mg mL^{-1} trigonelline and caffeine and submitted to the extraction and clean-up procedures used for sample preparation. The spiked amounts represent about 65, 85 and 100 % of the concentration of the alkaloids found in the coffee sample. Subsequently, the spiked samples were analysed by HPLC under identical conditions to that of the samples. All of the spiking experiments were performed in triplicates, and the overall average percent recovery values were calculated.

Water Content Determination

In order to express the results on a dry weight basis, the water content of the samples was determined after drying 0.200-g aliquots of powdered green beans in an oven at 105 °C for 12 h (Bertrand et al. 2008).

Chemometric Data Analysis

Data were analysed using one-way analysis of variance (ANOVA) to test for significant differences between the mean concentrations of alkaloids among the regional green coffee beans. Differences were considered significant when $\alpha < 0.05$. Multivariate chemometric analysis was performed using the statistical software package SPSS 20 (IBM Corp, USA). Principal component analysis (PCA) and linear discriminant analysis (LDA) models were constructed after arranging the dataset in the form of a matrix, of which the rows represented the green coffee bean samples and the columns the concentrations of the alkaloids. Before analysis, the data was standardized to a mean of zero and variance of one, by subtracting from each alkaloid the mean concentration of the alkaloid in the data set and dividing by the corresponding standard deviation. Unsupervised PCA models were obtained to visualize sample trends and evaluate the discriminatory characteristics of the determined alkaloids. In contrast, the supervised

LDA models were constructed to study the classification of samples according to geographical origins. The validity of the LDA model was assessed by means of the leave-one-out cross-validation method. In this method, each sample was classified by the discriminant functions computed from the rest of samples in the data set, i.e. each sample is treated as unknown and its class is determined based on the discriminant functions computed from the remainder of the samples. The percentage of cross-validated samples that were correctly classified provided an indication of the number of new samples, belonging to the groups of samples studied, which can be correctly classified by the LDA model. In addition, the reliability of the LDA model was tested in terms of recognition and prediction abilities. For this, the entire sample set was divided as training and validation sets. The validation set consisted of 24 randomly selected samples, while the remaining 75 samples were used as a training set to construct the LDA model. From the average of five iterations, the recognition ability of the model was evaluated from the percentage of the samples in the training set that were correctly classified in to their respective groups by the LDA model constructed based on the samples in the training set. Samples of the validation set were considered completely unknown and were used to evaluate the prediction ability of the LDA model constructed from samples of the training set. The prediction ability of the LDA model was evaluated in terms of the percentage of the samples in the validation set that were correctly classified by using the LDA model constructed in the training step.

Results and Discussion

Characterization of Alkaloids

The extracts of the green coffee samples were analysed to determine the concentrations of caffeine, theobromine, theophylline and trigonelline (Fig. 2). Analysis of a standard mixture (10.0 mg L^{-1} each) of the four alkaloids by HPLC provided well separated peaks within 7 min (Fig. 3). Srdjenovic et al. (2008) reported a HPLC method for the separation of caffeine, theobromine and theophylline within 8 min. In this study, excellent separation was achieved within a shorter time, even after including trigonelline in the mixture.

Theophylline was not detected in any of the green coffee samples. Alonso-Salces et al. (2009) reported that theophylline was not present in green Arabica coffee beans from Africa, Asia and America. Clifford and Kazi (1987) reported that the alkaloid was present only in Robusta beans. However, trigonelline, theobromine and caffeine were detected in the coffee samples (Fig. 4), with the exception of theobromine in Harar coffees. Theobromine was not detected, within the
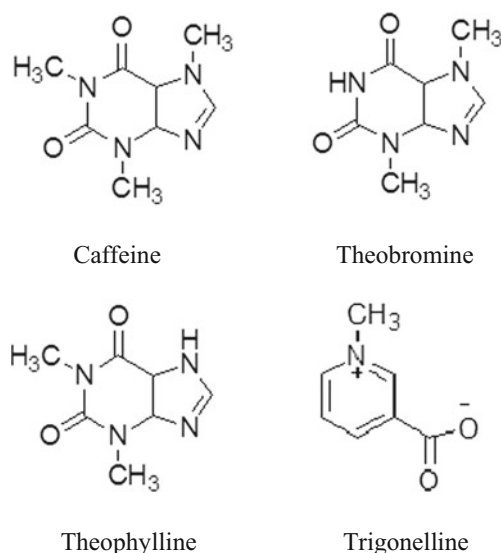
Fig. 2 Structures of the four coffee alkaloids studied

Fig. 4 HPLC chromatogram of a Sidama coffee extract (20 mg mL^{-1})

detection limit of the method, in any of the 26 samples of Harar coffees (Fig. 5). This finding implies that the absence or presence of theobromine can be used as a chemical marker for ascertaining whether the coffee originates from Harar or not.

The identities of the eluting compounds were determined by comparing the retention times and UV-vis spectra with those of pure standards. Moreover, the identities were confirmed by spiking some of the coffee extracts with standard solutions of the alkaloids. An increase in peak area for each alkaloid after standard addition supported the identification of the compounds (Fig. 6). The identity of theobromine was verified by UPLC-MS since the mass spectra obtained for the samples provided an accurate molecular mass, corresponding to the molecular ion at m/z 181.0724106658, and other
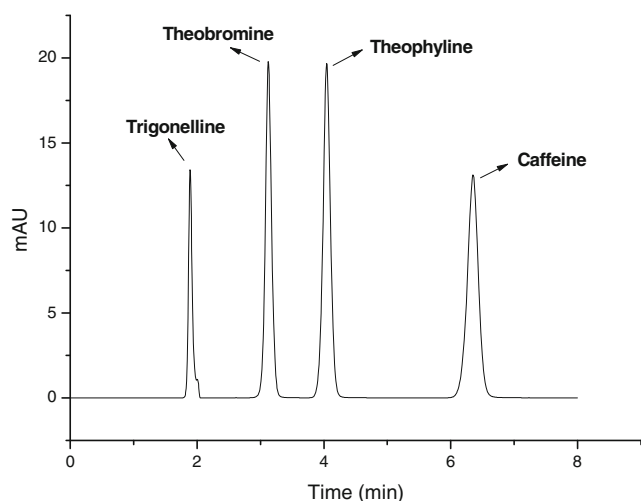
features that exactly matched that of the pure theobromine standard.

Validation of the Method

The method was validated with respect to each of the alkaloids. Calibration curves were found to be linear over the range 0.25–400 mg L^{-1} for theophylline, 0.5–400 mg L^{-1} for caffeine, 0.5–300 mg L^{-1} for theobromine and 1–400 mg L^{-1} for trigonelline. Regression coefficients (R^2) obtained for the calibration curves were better than 0.999, indicating good linearity within the concentration ranges studied. The LODs of the method were 0.26, 0.14, 0.17 and 0.08 mg L^{-1} for trigonelline, theobromine, caffeine and theophylline, respectively. This is equivalent to 13 mg kg^{-1} for trigonelline, 7 mg kg^{-1} for theobromine, 8.5 mg kg^{-1} for caffeine and 4 mg kg^{-1} for theophylline with respect to the dry weight of green coffee beans. These values were lower for theobromine and theophylline and higher for caffeine than

Fig. 3 HPLC chromatogram of a mixture of standard solutions (10 mg L^{-1} each) of the four alkaloids determined in the green coffee beans
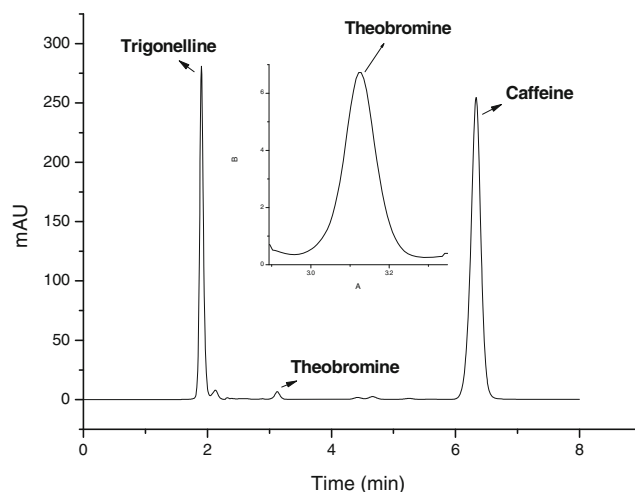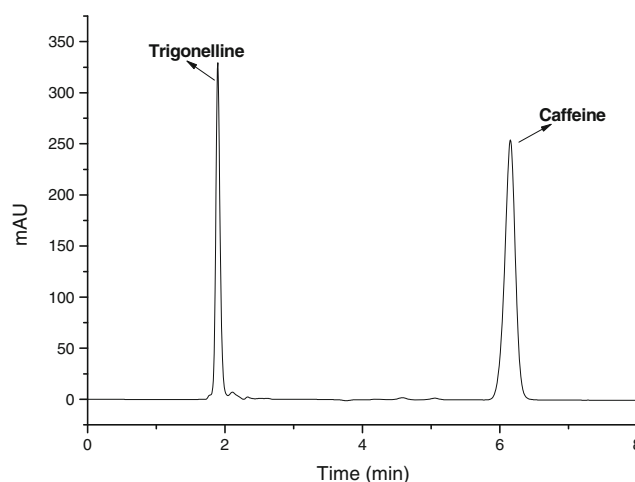

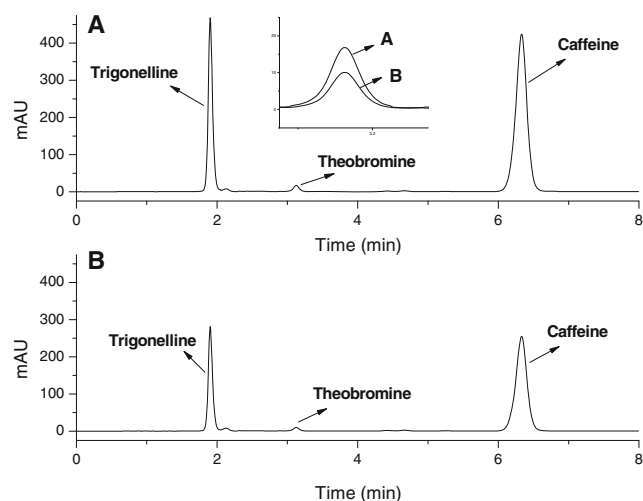Fig. 5 Typical HPLC chromatogram of a Harar coffee extract (20 mg mL^{-1})

Fig. 6 HPLC chromatogram of **a** a 1.00 mL aliquot of a Sidama coffee extract (20 mg mL^{-1}) spiked with 20 μL (10 mg mL^{-1}) trigonelline, 12 μL (100 mg mL^{-1}) theobromine and 20 μL (10 mg mL^{-1}) caffeine, and **b** the unspiked coffee extract. Inset in **a** indicates the overlay and enlarged peaks of theobromine

those obtained by Srdjenovic et al. (2008) after the analysis of coffee beans using HPLC with UV detection. They reported LODs of 0.2 mg L^{-1} for theobromine, 0.1 mg L^{-1} for theophylline and 0.07 mg L^{-1} for caffeine.

The precision of the method expressed in terms of relative standard deviation varied in the range 1.1–5.0 % for trigonelline, 2.8–9.1 % for theobromine and 0.20–6.2 % for caffeine. Relatively higher measurement uncertainties were associated with theobromine, probably due to the low concentrations in the samples. The levels of theobromine present in the coffee beans were, however, well above the LOQ of the method. The LOQ of the method, calculated as ten times the standard deviation of the background signal from nine measurements divided by the slope of the calibration equation, when expressed as % w/w of dry green coffee beans is 0.0026 %. On the other hand, the amounts of theobromine found in the coffee samples (0.0048–0.0094 %) were well above the LOQ of the method, except in Harar coffees where the concentrations of theobromine were below the LOD. The accuracy of the analytical procedure was evaluated by recovery studies performed in triplicates. The average percent recovery values were 98.6(±4.7)% for trigonelline, 99.4(±2.8)% for caffeine and 103.2(±6.3)% for theobromine. The variations of the percent recovery values from 100 % are all within one standard deviation values, indicated in parenthesis, assuring the accuracy of the method for determining each of the alkaloids (Muniz-Valencia et al. 2014).

Variations of Alkaloids in the Coffee Beans

The determined concentrations (% w/w dry coffee) varied in the range of 0.98–1.32 % for trigonelline, below LOD to

0.0094 % for theobromine and 0.87–1.38 % for caffeine across the entire sample set (Table 2). The concentrations of the alkaloids determined in the coffee samples are within the range of values previously reported for green beans of the Arabica variety (Belay et al. 2008; Perrone et al. 2008; Alonso-Salces et al. 2009; de Santos et al. 2012).

Statistical analysis with one-way ANOVA (α=0.05) revealed the presence of significant differences between the mean concentrations of the alkaloids found in samples from the four regions studied (Table 2). The mean concentrations of trigonelline in Northwest and East green coffee samples were significantly higher than that found in West and South samples. Coffee samples from Northwest contained significantly more theobromine than West coffees. On the other hand, as stated earlier, theobromine was not found in East (Harar) coffees. Regarding caffeine, samples from Northwest contained significantly higher amounts than samples from the other regions. Northwest coffees tend to be distinguished by their higher contents of caffeine as compared to those from the other regions. The distribution of the caffeine among samples within each region is displayed using box and whiskers plots (Fig. 7). The plots indicate a summary of the median, the first and third quartiles and the minimum and maximum observed values that are not statistical outliers.

The methylxanthine and trigonelline composition of green beans depends mainly on genetic properties, agricultural practices, harvesting and postharvest processing methods and environmental growing conditions (Casal et al. 2000; Alonso-

Table 2 The mean concentration (% w/w dry weight) of alkaloids determined in green coffee beans from the four regions studied

Region		Trigonelline	Theobromine	Caffeine
East (*n*=26)	Mean	1.19	<LOD[a]	1.05
	SD	0.08	–	0.08
	Min	1.06	–	0.87
	Max	1.32	–	1.31
Northwest (*n*=6)	Mean	1.22	0.0075	1.25
	SD	0.06	0.0007	0.10
	Min	1.15	0.0066	1.12
	Max	1.29	0.0084	1.38
South (*n*=49)	Mean	1.15	0.0067	1.09
	SD	0.06	0.0009	0.06
	Min	1.01	0.0054	0.95
	Max	1.31	0.0094	1.20
West (*n*=18)	Mean	1.10	0.0062	1.11
	SD	0.06	0.0011	0.06
	Min	0.98	0.0048	0.98
	Max	1.20	0.0088	1.24

The minimum (min), maximum (max) and standard deviation (SD) values are also indicated

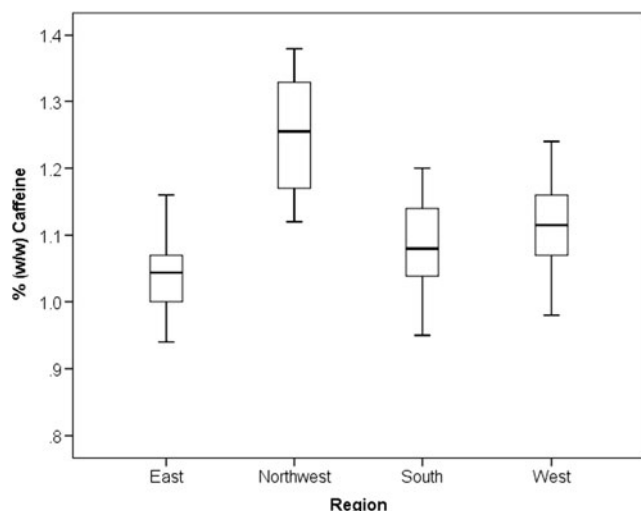[a] <LOD is below limit of detection

Fig. 7 Box plot showing the distribution of caffeine determined in green coffee beans from the four regions studied

Salces et al. 2009). Ethiopian coffee is predominantly produced by small-scale farmers using traditional farming practices, which are normally considered to be organic (Kufa et al. 2011). Furthermore, coffee harvesting in Ethiopia is typically performed by hand picking of ripe coffee cherries. Therefore, differences arising from agricultural processes and harvesting methods probably do not attribute to the observed differences in the methylxanthine and trigonelline contents of the studied coffee samples. Regarding postharvest processing methods, the samples used in the study were of both wet processed (26 samples) and dry processed (73 samples) coffee beans. The samples were collected in accordance with the proportion of commercial coffees produced by the country. Joet et al. (2010) found that wet processing did not affect the caffeine concentration of green Arabica coffee beans from Reunion Island. A study by Duarte et al. (2010) also indicated that caffeine levels are not affected by the processing method used. However, in contrast to this, Duarte et al. (2010) did find that the postharvest processing method had a significant influence on the concentration of trigonelline in green Arabica beans from Brazil. Wet processed beans were found to contain higher amounts of trigonelline than semi-dry processed coffee beans. Rather than the method having a direct effect on trigonelline, the authors ascribed the observed variation in the wet method to the loss of other water-soluble components, resulting in higher concentrations of trigonelline. In relation to geographical origin, Gichimu et al. (2014) indicated the presence of significant variation in the caffeine and trigonelline contents of green Arabica beans cultivated in different parts of Kenya. Alonso-Salces et al. (2009) also reported considerable variation in the concentration of caffeine present in green Arabica coffee beans from different countries and continents. Several investigators have also pointed out the significant influence of genetic traits on the caffeine and trigonelline composition of green coffee beans (Casal et al. 2000; Alonso-

Salces et al. 2009). Concerning Ethiopian coffees, investigators have indicated the existence of high genetic variation between coffees grown in different regions of the country (Kebede and Bellachew 2008). Therefore, the observed variation in the methylxanthine and trigonelline contents of the studied samples can be ascribed mainly to differences in the genetic properties and environmental growing conditions of the coffee plants. The postharvest processing method used may also have contributed to the observed variation of the trigonelline concentrations.

Relative standard deviation values were used to evaluate the extent of variability of the alkaloids among coffee samples

Table 3 The mean concentrations (% *w/w* dry weight) of alkaloids determined in green coffee beans of the eight major varieties studied

Variety		Trigonelline	Theobromine	Caffeine
Benishangul (*n*=3)	Mean	1.20	0.0081	1.16
	SD	0.06	0.0003	0.04
	Min	1.15	0.0079	1.12
	Max	1.26	0.0084	1.19
Finoteselam (*n*=3)	Mean	1.25	0.0070	1.34
	SD	0.07	0.0004	0.03
	Min	1.17	0.0066	1.32
	Max	1.29	0.0074	1.38
Harar (*n*=26)	Mean	1.19	<LOD[a]	1.04
	SD	0.08	–	0.08
	Min	1.06	–	0.87
	Max	1.32	–	1.31
Jimma (*n*=6)	Mean	1.15	0.0064	1.15
	SD	0.04	0.0012	0.07
	Min	1.10	0.0048	1.06
	Max	1.20	0.0084	1.24
Kaffa (*n*=10)	Mean	1.06	0.0057	1.09
	SD	0.04	0.0005	0.06
	Min	0.98	0.0049	0.98
	Max	1.13	0.0063	1.16
Sidama (*n*=39)	Mean	1.14	0.0066	1.09
	SD	0.06	0.0009	0.06
	Min	1.01	0.0054	1.00
	Max	1.26	0.0094	1.20
Wolega (*n*=2)	Mean	1.13	0.0081	1.14
	SD	0.04	0.0011	0.01
	Min	1.10	0.0073	1.13
	Max	1.16	0.0088	1.15
Yirgachefe (*n*=10)	Mean	1.16	0.0070	1.07
	SD	0.08	0.0009	0.05
	Min	1.04	0.0055	0.95
	Max	1.31	0.0087	1.13

The minimum (min), maximum (max) and standard deviation (SD) values are also indicated

[a] <LOD is below limit of detection

Table 4 The concentrations (% dry weight) of caffeine (CAF), theobromine (TBR) and trigonelline (TRG) determined in this study compared with values reported for green coffee beans of the Arabica variety

Sample origin	CAF	TBR	TRG	Method	Ref.
Africa	1.22–1.68	–	–	HPLC-UV	Alonso-Salces et al. (2009)
Central America	1.09–1.74	–	–		
South America	1.05–1.72	–	–		
Asia	1.43–1.69	–	–		
Brazil	0.92	–	1.03	LC-MS	Perrone et al. (2008)
Commercial roasted	1.39–2.21	0.0024–0.011	–	HPLC-UV	Santos and Rangel (2012); Lopez-Martinez et al. (2003)
Brazil	1.05–1.54	–	0.74–0.92	HPLC-UV	Duarte et al. (2010)
Kenya	0.95–1.23	–	1.01–1.54	HPLC-UV	Gichimu et al. (2014)
Ethiopia	0.87–1.38	0.0048–0.0094	0.98–1.32	HPLC-UV	This study

produced within a given region. Trigonelline displayed the highest variability within East samples. Theobromine was highly variable within West, while caffeine was highly variable within the East and Northwest samples. The observed variation is due to the fact that a number of distinct coffee varieties are produced within each region. Furthermore, within an area that produces a given variety of coffee, the presence of several coffee landraces that can be distinguished by their distinct agromorphological and organoleptic characteristics has been indicated (Kebede and Bellachew 2008; Tessema et al. 2011). These unique organoleptic properties are indicative of the presence of genetic variability.

The eight major coffee varieties, i.e. Harar, Jimma, Kaffa, Wollega, Sidama, Yirgachefe, Benishangul and Finoteselam, were also tested for the presence of significant differences in their alkaloid contents. Accordingly, Kaffa coffee was found to have a significantly lower content of trigonelline than the other coffee varieties, with the exception of Jimma and Wellega coffees (Table 3). Furthermore, the Kaffa variety contains smaller amounts of theobromine; however, the difference was significant only compared to Benishangul, Wollega and Yirgachefe coffees. Caffeine was found in significantly higher concentrations in Finoteselam than in the other coffee varieties. Harar was found to contain lower levels of caffeine; however, the difference is significant only compared to Finoteselam, Benishagul, Jimma and Wollega varieties.

The average concentrations of the alkaloids determined in the coffee bean samples from the four regions studied are, generally, comparable with values reported for green Arabica beans from different countries and continents (Table 4). The determined concentrations of caffeine (0.87–1.36 %) are comparable to those (0.95–1.23 %) reported for Kenyan coffees by Gichimu et al. (2014). However, the caffeine contents are considerably lower than those (1.43–1.69 %) reported for Asian coffees by Alonso-Salces et al. (2009). Quantitative data, in the literature, regarding the concentration of theobromine in green coffee beans is scarce.

Some studies have detected, but not quantified, theobromine in green coffee beans (Alonso-Salces et al. 2009). On the other hand, several investigators have reported quantitative data for theobromine in roasted coffee beans (Lopez-Martinez et al. 2003; Santos and Rangel 2012) and brews prepared from roasted ground coffee (Rodrigues and Bragagnolo 2013). Previous studies have indicated that the caffeine content of coffee is unaffected by the roasting process (Cruz et al. 2012). From its structural similarity with caffeine, theobromine may also be stable during coffee roasting. Hence, the amounts (0.0048–0.0094 %) determined in green beans in this study can be compared with values (0.0024–0.011 %) reported for commercial roasted coffees (Lopez-Martinez et al. 2003; Santos and Rangel 2012). The concentrations of trigonelline measured in this study (0.98–1.32 %) agree well
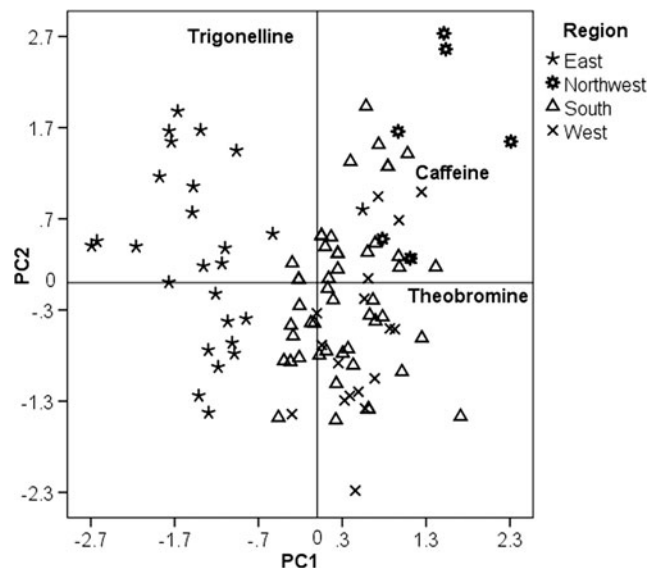
Fig. 8 Scores plot of the constructed the principal component analysis model constructed from the alkaloid concentrations determined in green coffee beans from Ethiopia

with those (1.01–1.54 %) determined by Gichimu et al. (2014) in Kenyan green beans.

Principal component analysis (PCA) was used to visualize sample trends and evaluate the discriminatory characteristics of the determined alkaloids. Two principal components (PCs), with Eigen values >1, were computed. PC1 and PC2 explained 48 and 35 %, respectively, of the overall variance in the dataset. The first component, PC1, was highly influenced to the positive side by both caffeine and theobromine, whereas PC2 was highly influenced to the positive side by trigonelline. Caffeine also contributed to the positive side of PC2. The distribution of samples on the scores of the two PCs is illustrated by Fig. 8. East (Harar) coffees are clearly separated to the negative side of PC1, mainly due to the absence of theobromine in these coffees. This can be a useful analytical tool to authenticate Harar coffees since these coffees can be characterized by the absence of theobromine. Northwest coffees tend to separate far to the positive side of both PC1 and PC2. This is mainly due to the higher content of caffeine in Northwest varieties, which is also clearly seen in the box plots (Fig. 7). On the other hand, South and West coffees tend to exhibit a similar alkaloid composition that is evident from the overlap of samples from these regions (Fig. 8).

Linear discriminant analysis (LDA) was applied to assess the usefulness of the alkaloids as indicators for the geographical of origin of the coffee beans. For this, the alkaloid contents of the coffee beans from the four regions, i.e. East, Northwest, South and West, were used. From the three discriminant functions computed, the first two discriminant functions together accounted for 99 % of the variation in the data set. Theobromine contributed highly to the first function, while caffeine contributed more to the second function. The overall proportion of the sample set correctly classified into the respective regions was 75 %. With respect to each region, 67 % of the South samples were correctly classified, while 100, 67 and 61 % of the samples from the East, Northwest and West samples were correctly classified, respectively.

The LDA model was tested with leave-one-out cross-validation, in which each sample was classified by the discriminant functions computed from the remainder of the samples in the dataset. Accordingly, 74 % of cross-validated samples were correctly classified into their respective regions. The reliability of the LDA model was assessed in terms of its recognition and prediction abilities. The recognition ability of the model was 75 %, whereas its prediction ability was 74 %. This shows that the LDA model has only moderate efficiency. Hence, alkaloids alone are not sufficient to build accurate predictive models for the origin of coffee beans.

However, the results obtained in this study demonstrates the potential of using alkaloids, in conjunction with other chemical constituents such as phenolic compounds, for the construction of classification models and generation of databases useful for the geographical origin discrimination of Ethiopian green coffee beans.

Conclusion

HPLC was successfully applied for the simultaneous determination of coffee alkaloids (caffeine, theobromine, theophylline and trigonelline) in the aqueous extracts of green coffee beans from various regions of Ethiopia. Theophylline was not detected in any of the coffee samples. On the other hand, significant variations in the concentration of the other alkaloids were observed depending on the geographical origin of the coffee beans. Theobromine was not found in coffee beans from the East (Harar coffees), while coffee beans from the Northwest were characterized by higher concentrations of caffeine. Therefore, the absence or presence of theobromine can be used as a chemical marker to ascertain whether beans are of the Harar variety or not. Application of linear discriminant analysis using the alkaloid contents of the beans provided 75 % correct classification of samples into their respective production regions with a 74 % prediction success rate. The moderate classification efficiency obtained with alkaloids demonstrates the potential of using this class of compounds in discriminant models useful for tracing the geographical origin of green coffee beans from Ethiopia.

References

Alonso-Salces RM, Serra F, Reniero F (2009) Botanical and geographical characterization of green coffee (*Coffea arabica* and *Coffea canephora*): chemometric evaluation of phenolic and methylxanthine contents. J Agric Food Chem 57:4224–4235

Al-Othman ZA, Aqel A, Alharbi MKE, Badjah-Hadj-Ahmed AY, Al-Warthan AA (2012) Fast chromatographic determination of caffeine in food using a capillary hexyl methacrylate monolithic column. Food Chem 132:2217–2223

Alpdogan G, Karabina K, Sungur S (2002) Derivative spectrophotometric determination of caffeine in some beverages. Turkish J Chem 26:295–302

Amamo AA (2014) Coffee production and marketing in Ethiopia. Eur J Bus Manag 6:109–121

Ascherio A, Weisskopf MG, O'Reilly EJ (2004) Coffee consumption, gender, and Parkinson's disease mortality in the cancer prevention study II cohort: the modifying effects of estrogen. Am J Epidem 160:977–984

Ashihara H, Sano H, Crozier A (2008) Caffeine and related purine alkaloids: biosynthesis, catabolism, function and genetic engineering. Phys Chem Chem Phys 69:841–856

Belay A, Ture K, Redi M, Asfaw A (2008) Measurement of caffeine in coffee beans with UV/vis spectrometer. Food Chem 108:310–315

Bertrand B, Villarreal D, Laffargue A, Posada H, Lashermes P, Dussert S (2008) Comparison of the effectiveness of fatty acids, chlorogenic acids, and elements for the chemometric discrimination of coffee (*Coffea arabica* L.) varieties and growing origins. J Agric Food Chem 56:2273–2280

Casal S, Oliveira MBPP, Alves MR, Ferreira MA (2000) Discriminate analysis of roasted coffee varieties for trigonelline, nicotinic acid, and caffeine content. J Agric Food Chem 48:3420–3424

Clifford MN, Kazi T (1987) The influence of coffee bean maturity on the content of chlorogenic acids, caffeine and trigonelline. Food Chem 26:59–69

Crozier A, Clifford MN, Ashihara H (2006) Plant secondary metabolites: occurrence, structure and role in the human diet, 1st edn. Blackwell, Oxford, pp 1–22

Cruz R, Cardoso MM, Fernandes L, Oliveira M, Mendes E, Baptista P, Morais S, Casal S (2012) Espresso coffee residues: a valuable source of unextracted compounds. J Agric Food Chem 60:7777–7784

de Santos WJR, Santhiago M, Yoshida IVP, Kubota LT (2012) Electrochemical sensor based on imprinted sol–gel and nanomaterial for determination of caffeine. Sensors Actuators B 167:739–745

del Campo G, Berregi I, Caracena R, Zuriarrain J (2010) Quantitative determination of caffeine, formic acid, trigonelline and 5-(hydroxymethyl)furfural in soluble coffees by ^{1}H-NMR spectrometry. Talanta 81:367–371

Duarte GS, Pereira AA, Farah A (2010) Chlorogenic acids and other relevant compounds in Brazilian coffees processed by semi-dry and wet post-harvesting methods. Food Chem 118:851–855

Farah A (2012) Coffee: emerging health effects and disease prevention, 1st edn. John Wiley & Sons, Oxford, pp 21–58

Farah A, de Paulis T, Moreira DP, Trugo LC, Martin PR (2006a) Chlorogenic acids and lactones in regular and water-decaffeinated Arabica coffees. J Agric Food Chem 54:374–381

Farah A, Monteiro MC, Calado V, Franca AS, Trugo LC (2006b) Correlation between cup quality and chemical attributes of Brazilian coffee. Food Chem 98:373–380

Gichimu BM, Gichuru EK, Mamati GE, Nyende AB (2014) Biochemical composition within *Coffea arabica* cv. Ruiru 11 and its relationship with cup quality. J Food Res 3:31–44

Higdon J, Frei B (2006) Coffee and health: a review of recent human research. Crit Rev Food Sci Nutr 46:101–123

Hirakawa N, Okauchi R, Miura Y, Yagasaki K (2005) Anti-invasive activity of niacin and trigonelline against cancer cells. Biosci Biotechnol Biochem 69:653–658

Huck CW, Guggenbichler W, Bonn GK (2005) Analysis of caffeine, theobromine and theophylline in coffee by near infrared spectroscopy (NIRS) compared to high-performance liquid chromatography (HPLC) coupled to mass spectrometry. Anal Chim Acta 538:195–203

Joet T, Laffargue A, Descroix F, Doulbeau S, Bertrand B, de Kochko A, Dussert S (2010) Influence of environmental factors, wet processing and their interactions on the biochemical composition of green Arabica coffee beans. Food Chem 118:693–701

Kalaska B, Piotrowski L, Leszczynska A, Michalowski B, Kramkowski K, Kaminski T, Adamus J, Marcinek A, Gebicki J, Mogielnicki A, Buczko W (2014) Antithrombotic effects of pyridinium compounds formed from trigonelline upon coffee roasting. J Agric Food Chem 62:2853–2860

Kebede M, Bellachew B (2008) Phenotypic diversity in the Hararge coffee (*Coffea arabica* L.) germplasm for quantitative traits. East Afr J Sci 2:13–18

Kufa T, Ayano A, Yilma A, Kumela T, Tefera W (2011) The contribution of coffee research for coffee seed development in Ethiopia. E3 J Agric Res Develop 1:9–16

Ky C-L, Dussert S, Guyot B, Hamon S, Noirot M (2001) Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild *Coffea arabica* L. and *C. canephora* P. accessions. Food Chem 75:223–230

Lopez-Martinez L, Lopez-de-Alba PL, Garcia-Campos R, De Leon-Rodriguez LM (2003) Simultaneous determination of methylxanthines in coffees and teas by UV–vis spectrophotometry and partial least squares. Anal Chim Acta 493:83–94

Muniz-Valencia R, Jurado JM, Magana SGC, Alcazar A, Diaz JH (2014) Characterization of Mexican coffee according to mineral contents by means of multilayer perceptrons artificial neural networks. J Food Compos Anal 34:7–11

Murthy PS, Naidu MM (2012) Sustainable management of coffee industry by-products and value addition—a review. Resourc Conserv Recycl 66:45–58

Paradkar MM, Irudayaraj J (2002) Rapid determination of caffeine content in soft drinks using FTIR–ATR spectroscopy. Food Chem 78:261–266

Perrone D, Donangelo CM, Farah A (2008) Fast simultaneous analysis of caffeine, trigonelline, nicotinic acid and sucrose in coffee by liquid chromatography-mass spectrometry. Food Chem 110:1030–1035

Petrie HJ, Chown SE, Belfie LM (2004) Caffeine ingestion increases the insulin response to an oral-glucose-tolerance test in obese men before and after weight loss. Am J Clin Nutr 80:22–28

Rodrigues NP, Bragagnolo N (2013) Identification and quantification of bioactive compounds in coffee brews by HPLC–DAD–MS^n. J Food Compos Anal 32:105–115

Santos JR, Rangel AOSS (2012) Development of a chromatographic low pressure flow injection system: application to the analysis of methylxanthines in coffee. Anal Chim Acta 715:57–63

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. J Chromatogr Sci 46:144–149

Svorc L, Tomcik P, Svitkova J, Rievaj M, Bustin D (2012) Voltammetric determination of caffeine in beverage samples on bare boron-doped diamond electrode. Food Chem 135:1198–1204

Tessema A, Alamerew S, Kufa T, Garedew W (2011) Variability and association of quality and biochemical attributes in some promising *Coffea arabica* germplasm collections in southwestern Ethiopia. Int J Plant Breed Genet 1–15

Tohda C, Kuboyama T, Komatsu K (2005) Search for natural products related to regeneration of the neuronal network. Neurosignals 14:34–45