Your leader in SCIENCE
Your partner in TRUST

June 4 2024
Application Note
Version 1

Authentication and Blend Analysis of Coffee Products

Using ^{1}H Nuclear Magnetic Resonance Spectroscopy and Non-Targeted Chemometrics

Arun Krishnamurthy, Ph.D.
NMR Spectroscopist

Thiru Arunachalam, Ph.D.
NMR Spectroscopist and Phytochemist

PURITY-IQ INC.

150 Research Lane, Suite 102
Guelph ON, N1G4T2, Canada

info@purity-iq.com

www.purity-iq.com

The Most Active Ingredient in Any Product is TRUST

Abstract

This document outlines a methodology that has been specifically developed to meet the needs of the coffee industry. The method employs quantitative ^{1}H NMR and full chemical profiling to accurately authenticate pure coffee and verify the percentage of arabica coffee in arabica and robusta blends of different roast types. This study involves *Coffea arabica* and *Coffea canephora* var. *robusta* beans of different roast types (dark, medium, light, and green). An appropriate extraction solvent was used to quantify six biomolecules in coffee: caffeine, trigonelline, 3- and 5-caffeoylquinic acid, cafestol, kahweol, and 16-*O*-methylcafestol. Among these, kahweol and 16-*O*-methylcafestol concentrations and statistical models built using non-targeted metabolomics were used to determine *C. arabica* and *C. robusta* ratios. The proposed methodology offers a systematic approach and a highly accurate method to determine coffee purity.

Introduction

Coffee, a beloved beverage worldwide [1], is primarily produced in tropical regions of South and Central America, East Africa, and Middle-Southeast Asia [1,2]. There are two types of commercially relevant coffee: Arabica (*Coffea arabica*) and Robusta (*Coffea canephora* var. *robusta*), with significant differences between these in terms of their quality and flavour [3]. Coffee beans from different regions and environments can be blended together to create unique blends [4,5]. However, the lower price of Robusta (for example, Arabica – 5.27 USD/kg, Robusta - 4.23 USD/kg [6]; Arabica – 4.4 USD/kg, Robusta – 2.4 USD/kg [7]) leads to a widespread practice of diluting high-quality Arabica with cheaper Robusta beans, making it imperative for brand owners to identify such adulterations, safeguarding consumer interests and guaranteeing the delivery of high-quality products [5].

With the adulteration of commercially popular products on the rise, consumers expect traceability systems to warrant the origin and compositional value of food and beverage commodities, making it necessary to develop innovative analytical solutions to ensure their authenticity. The same applies for coffee products and several analytical methods such as high-performance liquid chromatography (HPLC), gas chromatography (GC), NMR, near-infrared (NIR), and infrared (IR) spectroscopy have been implemented to study the chemical composition of coffee either by monitoring targeted biomarkers or capturing chemical fingerprints to discriminate coffee

samples based on multivariate statistical methodologies [8]. In a targeted approach, validated HPLC-based methods are available for determining the concentration of biomarkers in coffee to assess adulteration but these suffer from high cost and labour, rendering them less preferable [9]. On the contrary, Nuclear Magnetic Resonance (NMR)-based methods [9–14] have proven to be far more efficient, accurate, and robust in providing an innovative solution for the authentication of coffee products. The technique can be used in both targeted and untargeted approaches, wherein individual biomarkers can be identified and quantified, and the total chemical fingerprint can be captured and analyzed through multivariate statistics, respectively. Several analytical concerns such as product purity (robusta vs. arabica %), roast type, consistency in chemical composition of products from different batches, and presence of biomolecules and their associated concentrations can all be tied up into one analysis, providing a complete picture of the product chemistry while addressing authenticity, purity, and quality.

A brief description of our targeted and untargeted methods and approaches in addressing coffee authenticity is presented in this note.

- **Targeted:** Kahweol and 16-*O*-methylcafestol (16-*O*MC), the latter of which is a methylated congener of cafestol, are the biomolecules unique to arabica and robusta, respectively, and act as chemical markers for determining the adulteration of arabica beans with robusta [15].
- **Non-Targeted:** The chemical fingerprints of pure and adulterated coffee (arabica blended with robusta in different fractions) are captured and a sophisticated spectral library is built, which is used for predicting blend ratios using multivariate statistical analysis.
- **Types of Samples Analyzed:** green beans and dark, medium, and light roasts.

Experimental

Arabica and Robusta beans authenticated through genomics and morphological evaluation were obtained from a research collaborator. These beans were ground and extracted using an appropriate solvent. One-dimensional proton (^{1}H) NMR spectra were acquired using a 400 MHz Bruker Avance III spectrometer.

Six samples of each type were analyzed to build the statistical models:

- 100% *C. arabica* (light, medium, dark, green)
- 100% *C. robusta* (light, medium, dark, green)
- Blends of *C. arabica/C*. robusta dark roast in increments of 10%
- Blends of *C. arabica/C*. robusta medium roast in increments of 10%
- Blends of *C. arabica/C*. robusta light roast in increments of 10%
- Blends of *C. arabica/C*. robusta green beans in increments of 10%

These spectra were processed, baseline and phase-corrected, and binned for multivariate statistical analysis on the ***R*** platform. The characteristic peaks of six coffee constituents (caffeine, trigonelline, 3- and 5-caffeoylquinic acid, cafestol, kahweol, and 16-*O*MC) were assigned based on literature, spiking experiments of reference molecules, and two-dimensional NMR experiments. These molecules were quantified using ERETIC2 [16], a PULCON methodology (Pulse Length–based Concentration determination) [17] built into the data acquisition and processing software supplied by Bruker®.

Data Analysis

Non-targeted: Statistical models employing different methods such as the Principal Component Analysis (PCA), Hierarchical Clustering, and Linear Discriminant Analysis (LDA) were constructed using authentic 100% pure *C. robusta* and *C. arabica* dark, medium, light roasts and green bean samples and their blends in different proportions. These models were used in predicting the actual composition of new in-house blends of *C. arabica* and *C. robusta* of different roast types.

Targeted: The six biomolecules, caffeine, trigonelline, 3- and 5-caffeoylquinic acid, cafestol, kahweol, and 16-*O*MC were quantified in pure and blended samples and statistical models were constructed composing the molecular concentrations. These models were employed in analyzing new in-house blends of *C. arabica* and *C. robusta* of different roast types.

Results and Discussion

Non-targeted approach: A principal component analysis (PCA) plot generated by non-targeted total chemical profiling of pure and blended coffee samples is shown in **Figure 1** wherein, the blend samples of different roasts cluster into different groups, a result of differences in the spectral profile of the coffee types. These differences include the position and intensity of peaks in the spectrum, which reflect the variance in the metabolic makeup of the coffee. Moreover, the samples within each cluster segregate based on their blend ratios, thus demonstrating the efficacy and capability of non-targeted chemometrics to differentiate samples based just on their chemical profiles.

The spectral data of ten new samples (*T1-T10*) of different roast type and blend composition unknown to the analyst (in-house blind study) treated identically to the samples were applied against this model for blend-ratio predictions. These samples align with respective clusters of roast types as well as blend samples within each cluster based on their chemical composition as shown in **Figure 2**, providing key insights into their *C. arabica* and *C. robusta* composition. LDA method was employed in conjunction with PCA to deduce the roast type and the predictions are listed in **Table 1** along with the true chemical nature of the blind samples. The blend composition of blind samples was predicted with high accuracy using non-targeted chemical profiling and multivariate statistical analysis.

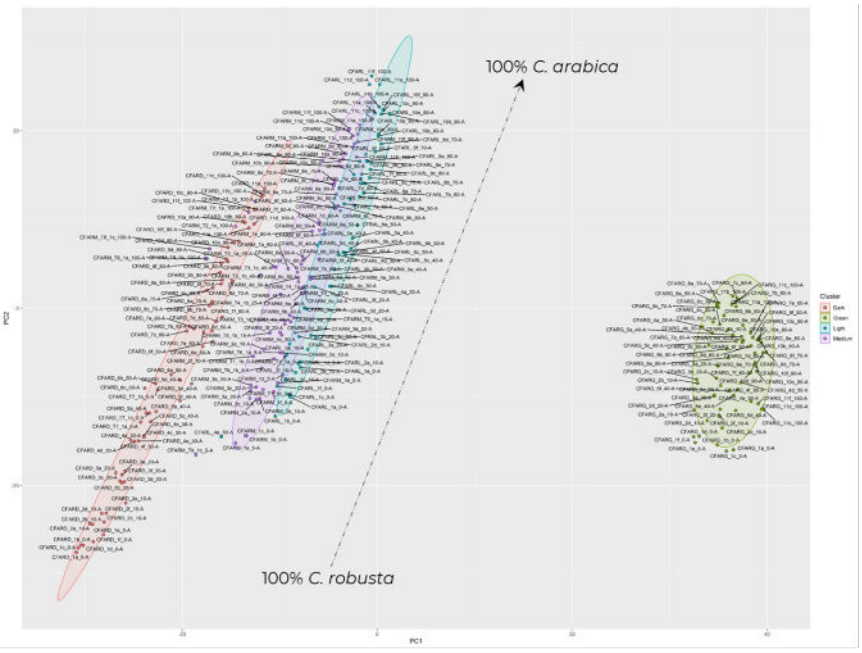

Figure 1. Principal Component Analysis (PCA) plot showing the clustering of datasets into different groups based on the roast type and further differentiation of samples within each group based on their blend ratio (Sample labels – CFARD_1a-1f_0-A: Dark roast 0% *C. arabica*, 1a-1f (6 samples); CFARD_1a-1f_100-A: Dark roast 100% *C. arabica; CFARM – Medium roast; CFARL – light roast; CFARG – Green beans*).

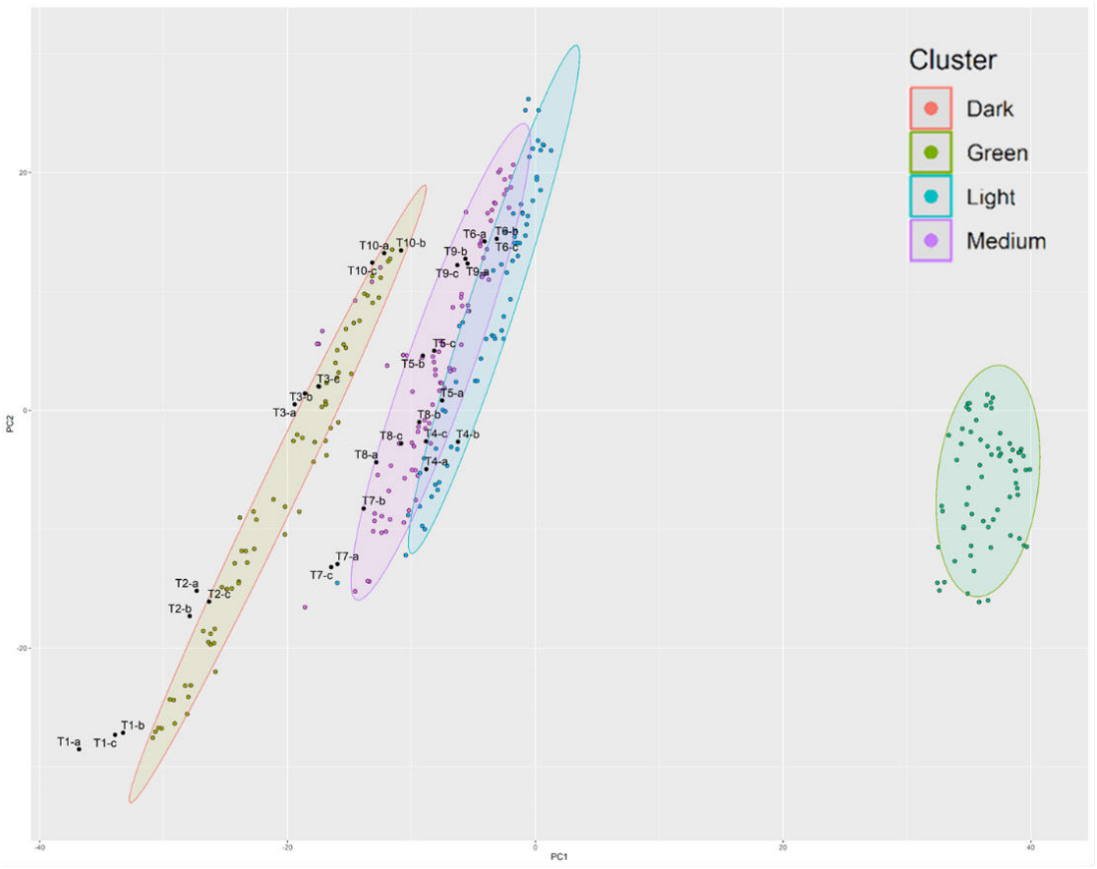

Figure 2. PCA plot showing the alignment of ten blind samples (*T1-T10*) in triplicates (a-c) with different clusters for roast and blend determination.

Table 1. Roast type and blend composition predictions from PCA and LDA methods.

Blind Samples	Dark Probabilities (%)	Green	Light	Medium	Blend Prediction (%) Arabica	Identity (Roast)
T1	1	0	0	0	10	10% Arabica (Dark)
T2	1	0	0	0	30	30% Arabica (Dark)
T3	1	0	0	0	70	70% Arabica (Dark)
T4	0	0	1	0	20	10% Arabica (Light)
T5	0	0	1	0	40	30% Arabica (Light)
T6	0	0	1	0	70	70% Arabica (Light)
T7	0	0	0	1	10	10% Arabica (Medium)
T8	0	0	0	1	30	30% Arabica (Medium)
T9	0	0	0	1	70	70% Arabica (Medium)
T10	0	0	0	1	100	100% Arabica (Medium)

0 – False; 1 – True

Targeted approach: Coffee is a rich source of phytochemicals that offer a multitude of health benefits. These phytochemicals include caffeine, chlorogenic acids (CQAs), diterpenes, trigonelline, and melanoidins, essential contributors to the flavor, aroma, and health advantages of coffee [18–20]. Caffeine acts as a bio-stimulant and contributes to coffee bitterness, while CQAs and their derivatives have antioxidant properties and the diterpenes cafestol, kahweol, and 116-*O*MC are related to increased levels of serum cholesterol [21–24]. Trigonelline, a pyridine alkaloid, contributes to the flavors and aroma of coffee [18–20]. These important biomolecules can be accurately quantified using NMR spectroscopy, and their quantities can further be used in authenticating coffee products for purity and roast type. The linear discriminant analysis plot in **Figure 3** clearly demonstrates this capability wherein samples cluster into different roast types based on the molecular concentrations. This model, which represents a targeted approach can also be employed for coffee authentication.

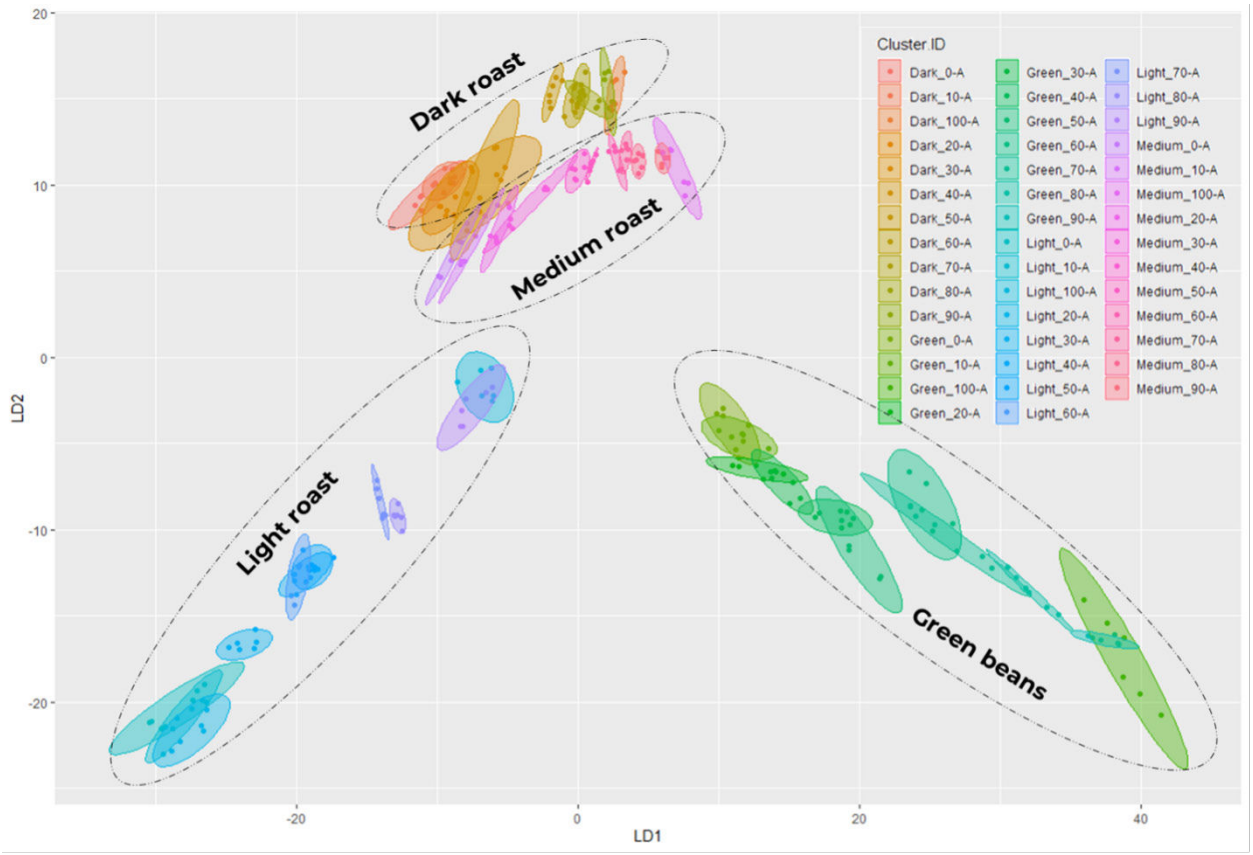

Figure 3. LDA model of coffee blends and their roast built using the concentrations of six biomolecules that were quantified (sample label: Dark_0-A – Dark roast 0% Arabica).

Conclusions

NMR chemometrics-based analytics offer information on:

- Product purity (Robusta vs. Arabica %)
- Consistency in chemical composition of products from different batches
- Presence of biomolecules and their associated concentrations
- A comprehensive understanding of product chemistry while addressing authenticity, purity, and quality.

References

1. A. Samoggia, B. Riedel, Consumers' perceptions of coffee health benefits and motives for coffee consumption and purchasing, Nutrients. 11 (2019) 653.
2. A.A. Nuhu, Bioactive Micronutrients in Coffee: Recent analytical approaches for characterization and quantification, ISRN Nutr. 2014 (2014) 1–13.
3. M. Jeszka-Skowron, A. Zgoła-Grześkowiak, T. Grześkowiak, Analytical methods applied for the characterization and the determination of bioactive compounds in coffee, Eur. Food Res. Technol. 240 (2015) 19–31.
4. N. Núñez, X. Collado, C. Martínez, J. Saurina, O. Núñez, Authentication of the origin, variety and roasting degree of coffee samples by non-targeted HPLC-UV fingerprinting and chemometrics. Application to the detection and quantitation of adulterated coffee samples, Foods. 9 (2020).
5. G. Hu, X. Peng, X. Wang, X. Li, X. Li, M. Qiu, Excavation of coffee maturity markers and further research on their changes in coffee cherries of different maturity, Food Res. Int. 132 (2020).
6. Ycharts. https://ycharts.com/indicators/world_coffee_robusta_price (accessed May 9, 2024).
7. Statista. https://www.statista.com/statistics/675807/average-prices-arabica-and-robusta-coffee-worldwide/ (accessed May 9, 2024).
8. V.A. Arana, J. Medina, P. Esseiva, D. Pazos, J. Wist, Classification of coffee beans by GC-C-IRMS, GC-MS, and 1H-NMR, J. Anal. Methods Chem. 2016 (2016) 8564584.
9. E. Schievano, C. Finotello, E. De Angelis, S. Mammi, L. Navarini, Rapid authentication of coffee blends and quantification of 16- o -methylcafestol in roasted coffee beans by nuclear magnetic resonance, J. Agric. Food Chem. 62 (2014) 12309–12314.
10. Y.B. Monakhova, W. Ruge, T. Kuballa, M. Ilse, O. Winkelmann, B. Diehl, F. Thomas, D.W. Lachenmeier, Rapid approach to identify the presence of Arabica and Robusta species in coffee using 1H NMR spectroscopy, Food Chem. 182 (2015) 178–184.
11. M. Defernez, E. Wren, A.D. Watson, Y. Gunning, I.J. Colquhoun, G. Le Gall, D. Williamson, E.K. Kemsley, Low-field 1H NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees, Food Chem. 216 (2017) 106–113.

References (Continued)

12. C. Finotello, C. Forzato, A. Gasparini, S. Mammi, L. Navarini, E. Schievano, NMR quantification of 16-O-methylcafestol and kahweol in Coffea canephora var. robusta beans from different geographical origins, Food Control. 75 (2017) 62–69.
13. V.A. Arana, J. Medina, R. Alarcon, E. Moreno, L. Heintz, H. Schäfer, J. Wist, Coffee's country of origin determined by NMR: The Colombian case, Food Chem. 175 (2015) 500–506.
14. I.W. Burton, C.F. Martinez Farina, S. Ragupathy, T. Arunachalam, S. Newmaster, F. Berrué, Quantitative NMR methodology for the authentication of roasted coffee and prediction of blends, J. Agric. Food Chem. 68 (2020) 14643–14651.
15. Y. Gunning, M. Defernez, A.D. Watson, N. Beadman, I.J. Colquhoun, G. Le Gall, M. Philo, H. Garwood, D Williamson, A.P. Davis, E.K. Kemsley, 16-O-methylcafestol is present in ground roast Arabica coffees: Implications for authenticity testing, Food Chem. 248 (2018) 52–60.
16. B. Manual, ERETIC2 user's guide-preliminary, in: Billerica (MA), 2012.
17. G. Wider, L. Dreier, Measuring protein concentrations by NMR spectroscopy, J. Am. Chem. Soc. 128 (2006) 2571–2576.
18. D. Perrone, C.M. Donangelo, A. Farah, Fast simultaneous analysis of caffeine, trigonelline, nicotinic acid and sucrose in coffee by liquid chromatography-mass spectrometry, Food Chem. 110 (2008) 1030–1035.
19. N.C. Bicho, A.E. Leitão, J.C. Ramalho, F.C. Lidon, Identification of chemical clusters discriminators of the roast degree in Arabica and Robusta coffee beans, Eur. Food Res. Technol. 233 (2011) 303–311.
20. C.A.. De Maria, L.. Trugo, F.R.A. Neto, R.F.. Moreira, C.. Alviano, Composition of green coffee water-soluble fractions and identification of volatiles formed during roasting, Food Chem. 55 (1996) 203–207.
21. S. Badmos, S.H. Lee, N. Kuhnert, Comparison and quantification of chlorogenic acids for differentiation of green Robusta and Arabica coffee beans, Food Res. Int. 126 (2019).
22. M.N. Clifford, I.B. Jaganath, I.A. Ludwig, A. Crozier, Chlorogenic acids and the acyl-quinic acids: Discovery, biosynthesis, bioavailability and bioactivity, Nat. Prod. Rep. 34 (2017) 1391–1421.
23. P. Esquivel, V.M. Jiménez, Functional properties of coffee and coffee by-products, Food Res. Int. 46 (2012) 488–495.
24. C.-L. Ky, J. Louarn, S. Dussert, B. Guyot, S. Hamon, M. Noirot, Caffeine, trigonelline, chlorogenic acids and sucrose diversity in wild Coffea arabica L. and C. canephora P. accessions, Food Chem. 75 (2001) 223–230.