World Journal of Analytical Chemistry, 2013, Vol. 1, No. 4, 69-72
Available online at http://pubs.sciepub.com/wjac/1/4/5
© Science and Education Publishing
DOI:10.12691/wjac-1-4-5

Untargetted Metabolomics in Assessment of Variations among Kenyan Arabica Coffee Genotypes Using Organic Compounds in the Brew

Cecilia Kathurima*, Fredrick Ng'ang'a

Coffee Research Foundation Ruiru-Kenya
*Corresponding author: cwkathurima@yahoo.com

Received September 09, 2013; Revised November 19, 2013; Accepted November 27, 2013

Abstract Genotypical variations between two commercial coffee cultivars considered as controls (Ruiru 11 and SL28) and five other genotypes under investigation (Cr8, Cr22, Cr23, Cr27 and Cr30) all grown under similar conditions were evaluated. Organic compounds were extracted using C18 Solid Phase Extraction (SPE) cartridges and analyzed by Gas Chromatography- Mass Spectrometry. Untargeted metabolomics data processing of the mass spectra was carried out on XCMS online (a web based platform). The analysis revealed absence of any significant variation in composition of the organic compounds at $p = 0.05$. However, from the Principal Component Analysis plots the genotypes Cr27, Cr22 and Cr23 clustered close together with Ruiru 11 while the other genotypes showed no distinct clustering pattern. Overall, there was no significant variation in the organic compounds composition across the coffee genotypes grown under similar agronomic practices in the same region and processed uniformly.

Keywords: *untargeted metabolomics, XCMS online, organic compounds, genotypes, coffee*

Cite This Article: Cecilia Kathurima, and Fredrick Ng'ang'a, "Untargetted Metabolomics in Assessment of Variations among Kenyan Arabica Coffee Genotypes Using Organic Compounds in the Brew." *World Journal of Analytical Chemistry* 1, no. 4 (2013): 69-72. doi: 10.12691/wjac-1-4-5.

1. Introduction

Metabolomics is based on biology, analytical chemistry, and information science, and it has become an important tool in many research areas [1]. Metabolomics is a rapidly developing field of 'omics' research, dealing with the detection, identification and quantification of low molecular-weight compounds (typically below 1000 Da) in cells, organs or organisms. The analysis and identification of small molecules is important in many areas such as Food Science [2] flavor research, biology and medicine such as biomarker discovery, diagnostics, pharmaceutical chemistry and functional genomics [3,4]. Metabolomic analyses have been generally classified as targeted or untargeted. Targeted analyses focus on a specific group of intended metabolites with most cases requiring identification and quantification of as many metabolites within the group [5]. Targeted analyses are important for assessing the behavior of a specific group of compounds in the sample under determined conditions, this type of analysis requires higher level of purification and a selective extraction of metabolites [1].

Untargeted (comprehensive) metabolomics is the global and simultaneous profiling of as many metabolites as possible in a search to identify altered pathways that provide a phenotypic signature for the biological system of interest [6] Untargeted metabolomics focuses on the detection of as many groups of metabolites as possible to obtain patterns or fingerprints without necessarily identifying nor quantifying a specific compound(s) [7].

Based on the specific objective of the analysis and data manipulation, most metabolomic studies can be classified as discriminative, informative, and/or predictive. Discriminative analyses have been aimed at finding the differences between sample populations without necessarily creating statistical models or evaluating possible pathways that may elucidate such differences. Discrimination is usually achieved by the use of multivariate data analysis (MVDA) techniques intended to maximize classification; principal components analysis (PCA) being the most used tool [8]. Informative metabolomics has been used in the development and continuous update of metabolite databases such as the human metabolome database [4] in predictive metabolomics statistical models based on metabolite profile and abundance are created to predict a variable that is difficult to quantify by other means.

In Food Science, metabolomics has been applied in various aspects such as: classification of grape wine by variety and production area [9], prediction of green tea sensory quality [10], evaluation of origin of extra virgin olive oil [11] verification of Corsican honey [12] and most interesting work by [13] on representation, comparison, and interpretation of metabolome fingerprint data for total composition analysis and quality trait investigation in potato cultivars.

The most commonly used analytical chromatographic platforms in metabolomics studies are based on hybrid

systems such as GC/MS, liquid chromatography (LC)/MS, and capillary electrophoresis (CE)/MS. Among these platforms, GC/MS is a relatively mature method because the reproducible measurement is possible and many peaks (200 to 500) can be reliably obtained from a biological sample [14]. In addition, peak identification is straightforward when retention time and mass spectra data are compared to those of accumulated compound information in a laboratory. For these reasons, GC/MS is generally recognized as one of the most versatile and applicable platform in metabolomics [1].

Over the past decade, several software programs for automated processing of mass spectrometry-based metabolomic data have been introduced. They range from the commercial ones such as: metAlign, MS rResolver, Marker View, MarkerLynx to the freely available ones such us MZmine, MSFACTs, MathDAMP, COMSPARI and XCMS [15].

XCMS has been widely accepted by the metabolomic community more so in the field of untargeted metabolomics, with over 350 citations of the original paper and more than 45,000 downloads [16]. The introduction of the XCMS online, a web based platform for processing untargeted metabolomics data, offers a more simplified route for analysis of metabolomics data for scientist with little or no technical knowledge on programming.

The objective of this study was to assess the difference in the Volatile Organic Compound composition of five breeding lines under trial as compared to two commercial cultivars predominantly grown in Kenya. Solid Phase Extraction using C18 cartridges was used for extraction of the organic compounds present in the coffee brew and later followed by Gas Chromatoghy –Mass Spectrometry analysis.

2. Materials and Methods

2.1. Study Site

The coffee materials used in this study were obtained from Machakos Agricultural Training Centre (ATC) in Eastern Kenya. This site lies at latitude 1°31′S and longitude 37°16′E and has an altitude of 1600 Metres Above Sea Level. The area is semi-arid with mean annual rainfall of 750 mm and mean annual temperature of 20.9 ^{0}C. The soils are luvisols, well drained, moderately deep to deep, dark red to yellowish red, friable to firm, sandy clay often with a topsoil of loamy sand and are strongly leached soils [17].

2.2. Test Materials

Two Kenyan commercial cultivars; SL28 and Ruiru 11 were assessed alongside five advanced breeding lines coded as Cross 8 (Cr8), Cross 22 (Cr22), Cross 23 (Cr23), Cross 27 (Cr27) and Cross 30 (Cr30).

2.3. Experimental Layout

The coffee genotypes evaluated in this study were established in a Randomized Complete Block Design (RCBD) with three replications. The trial was established in 2007.

2.3.1. Processing of the Samples

Coffee cherries were harvested from a sample size of 20 trees during the peak period in 2011. The cherries were bulked and wet processed using standard recommended procedures [18]. The cherry samples were pulped, fermented, washed and dried to a final moisture content of 10.5 to 11%. The parchment was then hulled and graded to seven grades based on size, shape and density. Grade AB was used as a representative grade for the characterization of volatile compounds.

2.3.2. Roasting Green Coffee and Brew Preparation

Roasting of the green coffee was done to attain a medium roast level using laboratory roaster (Probat BRZ 4, Rhein, Germany), within 24 hour of evaluation and allowed to rest for at least eight hours. The coffee brew was prepared as described in the Specialty Coffee Association of America protocol [19]. Samples were weighed out to the predetermined ratio of 8.25g per 150 ml of water. Each batch of the roasted coffee genotypes was ground separately using a sample grinder (Probat vtv-633T, Rhein, Germany) directly into the cup. Boiled deionized water was gently added to the cup taking care not to spill over while filling the cup. The brewed coffee was allowed to cool to room temperature (22-24°C), filtered under vacuum through a whatman filter paper (No. 42) and extracted immediately with C18 (reverse phase) Solid Phase Extraction cartridges.

2.3.3. Solid Phase Extraction Procedures

In this study, 1000mg/6ml strata C18-E SPE (Phenomenex) cartridges were used and they were conditioned before use. In the conditioning, 10 ml of methanol were passed through the cartridge at a flow rate of approximately 1ml/min and after that 10 ml of distilled were passed through the cartridge for maximum wetting. Without letting the cartridge to dry, 40ml of the brewed coffee from the test genotypes was allowed to run through each cartridge at a flow rate of 2ml/min. The pressure inside the manifold was gradually increased to 60 bars and maintained there for ten minutes. The cartridges were finally flashed with high pressure jet of nitrogen gas to dry them.

The organic analytes were later eluted with 10 ml of dichloromethane at a flow rate of 1ml/min followed by further pre-concentration to 1ml under a stream of nitrogen gas at room temperature [20].

2.3.4. Chromatographic Conditions

GC-MS analyses were performed in a Konic HRGC 400B Gas Chromatograph coupled to a Konic MSQ12 (Sant Cugat, Barcelona, Spain) quadrupole mass spectrometer. One micro liter (1µl) of each extracts were injected into the splitless mode in a TechnoKroma TRB5 (Cross-linked 5% Phenyl-95% Methyl Siloxane) capillary column (15m × 0.25mm i.d × 0.1µm film thickness). Helium was used as the carrier gas at a flow rate of 1ml/min. The injection temperature was maintained at 200°C, while the oven temperature was kept at 60°C and programmed to rise at 4°C/min to 150°C and finally to 240°C at a rate of 6°C/min. Mass spectra were recorded in the Electron Ionization mode at 70 eV scanning from 35-450m/z range, the ion source and transfer line temperature were maintained at 200°C and 250°C respectively.

2.3.5. Untargetted Metabolomic Analysis Using xcms Online

Data files in the cdf formats corresponding to chromatograms from the various cultivars were obtained and transferred into a single distinct folder. A user account was first created on https://xcmsonline.scripps.edu so as to utilise XCMS online platform. The data was uploaded as follows: data Set 1 – SL28 and Ruiru 11 (commercial cultivars); data set 2 – Cr8, Cr22, Cr23, Cr27, Cr28, Cr30 (advanced breeding lines under study). The upload process was allowed to continue to completion. The GC/single quad parameter was later selected as it exactly matched our instrument operating conditions. The results were submitted for analysis on the XCMS online servers.

3. Results and Discussion

From the results obtained there were less than 10 dysregulated i.e less than 10 features with p-value ≤ threshold (0.05). On examination of the online results table, only one feature had a p-value less than 0.05 which was 0.03366. This dysregulated feature correspond to the m/z of 121 and it was from Cr8. The presence of one dysregulated feature was an indicator of the lack of diversity in the composition of organic components across all the genotypes. The composition of organic compounds within the five genotypes under investigation was similar to the two existing commercial cultivars (Ruiru 11 and SL28). However after retention time correction and alignment of the total ion chromatograph (Figure 1), a closer look showed there was observable variation in the intensities of the individual peaks. This showed that there were differences in concentration of the various organic compounds extracted. From the PCA plot (Figure 2), the coffee genotypes Cr27, Cr22 and Cr23 clustered closer to the commercial cultivar Ruiru 11 indicating a higher level of similarity in concentration of the various organic compounds.

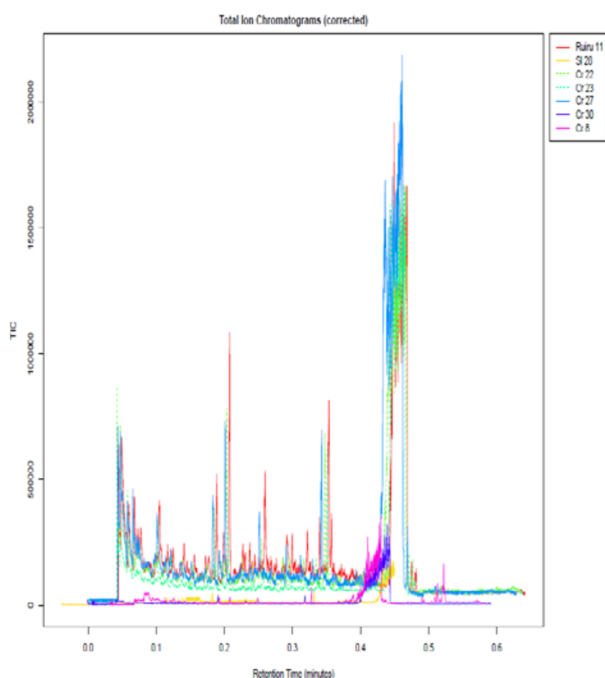

Figure 1. Aligned and retention time corrected total ion chromatogram of the various genotypes under study

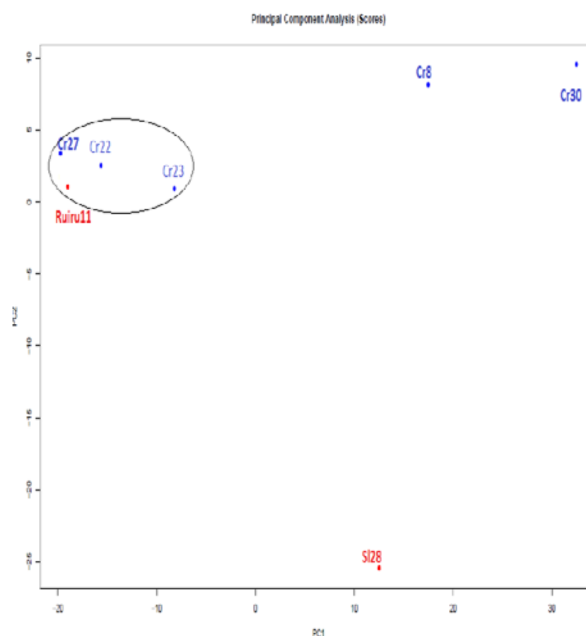

Figure 2. PCA plot showing the clustering of the various genotypes under study

4. Conclusions

Through the application of XCMS online in untargeted metabolomic data processing of GC-MS data, it was possible to establish some diversity in composition of organic compounds present in coffee brews from various genotypes. This has reinforced our believe that genetic cross breeding does not significantly influence the composition of the various organic compounds that influence the cup quality of coffee across various genotypes grown under similar conditions such as soils, rainfall, crop husbandry practices. This is could also be an another indication of the low genetic diversity in Arabica coffee.

Statement of Competing Interest

The authors declare no conflict of interest.

Acknowledgement

The authors extend sincere appreciation to Coffee Research Foundation (CRF) for funding this work. This work is published with the permission of the Director of Research, CRF, Kenya.

References

[1] Tsugawa, H., Tsujimoto, Y., Arita, M, Bamba, T. and Fukusaki, E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC Bioinformatics*, 12 (131). 1-13. 2011.

[2] Cevallos-Cevallos J., Reyes-De-Corcuera J., Etxeberria E, Danyluk M., and Rodrick G. Metabolomic analysis in food science: a review. *Trends in Food Science and Technology*, 20. 557-566. 2009.

[3] Roessner-Tunali, U., Beale M. H., Trethewey R. N., Lange B. M., Wurtele, E. S. and Sumner L. Potential of metabolomics as a

functional genomics tool. *Trends Plant Science.* 9(9):418-425, 2004.

[4] Wishart D. S. Current progress in computational metabolomics. *Briefs in Bioinformatics*, 8(5):279-293, 2007.

[5] Ramautar, R., Demirci, A., and Jong, G. J.d. Capillary electrophoresis in metabolomics. *Trends in Analytical Chemistry,* 25(5), 12. (2006).

[6] Vinayavekhin, N. and Saghatelian, A. Untargeted metabolomics. *Current Protocols in Molecular Biology.* 90 (30) 1-24 .2010.

[7] Monton, M. and Soga, T. Metabolome analysis by capillary electrophoresise mass spectrometry. *Journal of Chromatography* 1168 (1), 237-246. 2007.

[8] Kemsley, E. K., Le Gall, G., Dainty, J. R., Watson, A. D., Harvey, L. J., Tapp, H. S., Multivariate techniques and their application in nutrition: a metabolomics case study. *British Journal of Nutrition,* 98 (1), 1-14. (2007).

[9] Son, H. S., Kim, K. M., Van den Berg, F., Hwang, G. S., Park, W. M., Lee, C. H., H-1 nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. *Journal of Agricultural and Food Chemistry,* 56(17), 8007-8016.

[10] Ikeda, T., Kanaya, S., Yonetani, T., Kobayashi, A. and Fukusaki, E. Prediction of Japanese green tea ranking by fourier transform near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry*, 55. 9908-9912 (2007).

[11] Cavaliere, B., De Nino, A., Hayet, F., Lazez, A., Macchione, B., and Moncef, C., (2007). A metabolomic approach to the evaluation of the origin of extra virgin olive oil: a convenient statistical treatment of mass spectrometric analytical data. *Journal of Agricultural and Food Chemistry,* 55(4), 1454-1462.

[12] Donarski, J. A., Jones, S. A., and Charlton, A. J. (2008). Application of cryoprobe H-1 nuclear magnetic resonance spectroscopy and multivariate analysis for the verification of Corsican honey. *Journal of Agricultural and Food Chemistry,* 56(14), 5451-5456.

[13] Beckmann, M., Enot, D. P., Overy, D. P. and Draper, J. (2007). Representation, comparison, and interpretation of metabolome finger print data for total composition analysis and quality trait investigation in potato cultivars. *Journal of Agricultural and Food Chemistry,* 55 (9), 3444-3451.

[14] Denkert C., Budczies J., Kind T., Weichert W., Tablack P., Sehouli J., Niesporek S., Könsgen D., Dietel M. and Fiehn O. Mass spectrometry-based metabolic profiling reveals different metabolite patterns in invasive ovarian carcinomas and ovarian borderline tumors. *Cancer research,* 66. 10795-10804. 2006.

[15] Katajama, M. and Orešič, M., Data processing for mass spectrometry-based metabolomics. *Journal of chromatography,* 1158, 318-328, (2007).

[16] Patti, G., Yanes, O., Shriver, L., Courade, J., Tautenhahn, R., Manchester, M. and Siuzdak, G. *Nature Chemical Biology*, 8 (3), 232-234. 2012.

[17] Jaetzold, R. and Schmidt H. (1993). Farm Management Handbook of Kenya Vol. IIB Natural conditions and farm management information Central Kenya, Published by Ministry of Agriculture Kenya in cooperation with the German Agricultural Team (GAT) of the German Agency for Technical Cooperation GTZ.

[18] Lingle, T.R., The Cuppers Handbook. Systematic Guide to the Sensory Evaluation of Coffee`s Flavour, Third edition 2001.

[19] Mburu, J.K. The current recommendations for the processing of high quality and safe coffee in Kenya. In the Proceedings of 20th International Scientific Colloquium on Coffee. Bangalore India. 509-512. 2004.

[20] Kathurima C., Kenji G., Muhoho S., Boulanger R. and Ng'ang'a F. Comparative study of genotypical variation of volatile organic composition in brewed Kenyan coffee through solid phase extraction gas chromatography mass spectrometry. *Food Science and Quality Management.* 8, 18-26. (2012).