Journal of the Korean Magnetic Resonance Society **2017**, *21*, 44-49
DOI 10.6564/JKMRS.2017.21.2.044

Metabolic Features of Coffee Beans Depending on Planted Areas

Won Seok Choi†, Yong Woo In†, Hyun Hwi Kim, Ja-shil Hyun*, and Sung Jean Park*

Gachon Institute of Pharmaceutical Sciences, College of Pharmacy, Gachon University, 534-2 Yeonsu-dong, Yeonsu-gu, Incheon, Republic of Korea

Received Mar 25, 2017; Revised Apr 10, 2017; Accepted Apr 15, 2017

Abstract Coffee is one of the top selling products in the world. There are various coffee bean species around the world. Among them, *Coffea Arabica* is the most popular species. However, there are few studies on the metabolites of coffee beans so far. This study demonstrates effects of the planted regions on the metabolite concentrations of coffee beans. The metabolites of coffee beans can be affected by growing area even although same species are grown. Accordingly, we studied coffee bean metabolites extracted from the same species in different regions (The brand names, Kona from Hawaii, Mocha Matari from Yemen, and Blue Mountain from Jamaica) by using mixed solvent of methanol: water: chloroform. A comparative analysis by NMR spectroscopy was performed and the statistical techniques were used to figure out the differences. As a result, we found that chlorogenic acid, caffeine, citrate, and sucrose mainly contributed to the separation of the three groups. When compared with Kona and Blue Mountain, concentrations of chlorogenic acid, caffeine, and sucrose in Mocha Matari were observed to be relatively down-regulated. In addition, compared with the two other groups, concentration of citrate in Kona was observed to be up-regulated.

Keywords *Coffea Arabica*, metabolites, NMR

Introduction

Coffee is a brewed drink prepared from roasted coffee beans, which are the seeds of berries from the Coffea plant. There are 3 main species in the coffee. They are called *Coffea Arabica*, *Coffea Robusta*, and *Coffea Liberica*. Among them, *Coffea Arabica* is a species of coffee originally indigenous to the forests of the southwestern highlands of Ethiopia. It accounts for 70% of the world's coffee production. The main chemical ingredients in coffee beans are caffeine, chlorogenic acid, and so on. Interestingly, even though the species of coffee bean is exactly same, the flavor and taste are highly different, depending on the planted region. We supposed that these differences may be related with the biological contents of beans and were especially interested in the metabolites of beans.

Metabolomics is the "systematic study of the unique chemical fingerprints that specific cellular processes leave behind". That is, the study of their small-molecule metabolite profiles. Since ^{1}H NMR based metabolomics is a non-biased, non-destructive, high-throughput and easy quantification technique, it is widely applied in toxicology, environmental effects of pesticide residues, and other hazardous substances on organisms.[1-4] It also provides a

† These two authors contributed equally to this work.

* Correspondence to: **Sung Jean Park** or **Ja-shil Hyun,** College of Pharmacy, Gachon University, 534-2 Yeonsu-dong, Yeonsu-gu, Incheon, Republic of Korea, Tel: 82-32-820-4957; E-mail: psjnmr@gachon.ac.kr **for SungJean Park** or Tel: 82-32-820-4958; E-mail: jashilhyun@gachon.ac.kr **for Ja-shil Hyun**

powerful and versatile tool to evaluate the mechanism and diagnosis of a disease.[5-9]
This work focuses on the investigation on metabolites of *Arabica* coffee beans from different regions (Kona from Hawaii, Mocha Matari from Yemen, Blue Mountain from Jamaica), based on the NMR-based metabolomics.

Experimental Methods

Sample preparation- Briefly, the powder of *Arabica* coffee beans from different regions (Kona from Hawaii, Mocha Matari from Yemen, Blue Mountain from Jamaica, n=5 for each group) were divided into 150 mg for each. For extraction of metabolites, 1 mL of a mixed solvent of methanol: water: chloroform = 2: 1: 2 was added in each of 15 tubes and incubated in the ice for 30 min. The tubes were centrifuged at 4℃ with 13,000 rpm for 20 min. The polar supernatant was transferred into new tubes and the extraction solvent was evaporated with vacuum.

NMR measurement and Pre-processing - In NMR experiment, each sample was redissolved with 495 μl sodium phosphate buffer solution (100 mM Na_2HPO_4/100 mM $NaHPO_4$, pH 7.0), 50 μl D_2O and 5 μl TSP (Trimethylsilylpropanoic acid, 2 mM).[10] The cryoprobe-equipped Bruker 600 MHz NMR was used for spectra acquisition. The MestReNova software was used for all these data processing steps such as data binning and baseline correction. ^{1}H NMR spectra were manually phased and baseline corrected using Bruker Topspin 3.1 software (Bruker GmbH, Karlsruhe, Germany) and referenced to TSP at 0.0 ppm. The baseline model was built in each spectrum using the algorithm of multipoint baseline correction. Data binning is data processing technique which splits raw spectral peak in small intervals. To reduce analytical errors, peak alignment was applied before binning and a variable bin size ranging from 0.005 ppm to 0.03 ppm was used so that each single bin contains single metabolic information as much as possible. The chemical shift region of 4.6-5.0 ppm containing residual water was excluded. The lipid or protein contaminated region (1.20–1.34, 1.56–1.66, and 1.98–2.04 ppm) and noise region were also removed from the spectra to clarify the contribution of metabolites. The assignment of bins was achieved using Chenomx NMR suite 7.7 (Chenomx Inc., Edmonton, Canada) and evaluated in ^{1}H-^{13}C HSQC and 2D ^{1}H-^{1}H TOCSY spectra.

Multivariate analysis- The statistical analysis was performed using the MetaboAnalyst (v3.0)[11,12] a web server for metabolic data analysis and interpretation, and SPSS 23 (SPSS, Inc., an IBM Company, Chicago, Illinois, USA). The spectra were classified into three groups: blue-kona-mocha. The spectral integrals (bins) were normalized using sum average algorithm, thus facilitating comparability of samples. The spectral integrals (bins) were normalized using the internal reference, TSP, thus facilitating comparability of samples. The bins were also scaled with pareto-scaling to provide a reasonable balance of contributions from high and low amplitude signals. In the pareto-scaling, each bin is mean-centered and divided by the square root of the standard deviation.
To clarify the separation between groups, data were further processed using a supervised pattern recognition method orthogonal partial least squares discriminant analysis (OPLS-DA).[13,14] The efficiency and reliability of OPLS-DA models were validated by an external validation and a permutation test.[15]

Univariate analysis- The univariate analysis was performed to identify metabolites contributing to the discrimination among groups. The parametric ANOVA test was adopted for this purpose. The multiple comparison between groups was adjusted with Bonferroni's correction.[16] For the multiple testing correction, acquired p-values were adjusted using Benjamini and Hochberg False Discovery Rate (FDR).[17]

Results and Discussion

Group characteristics based on multivariate analysis- To explore the basic group characteristics,

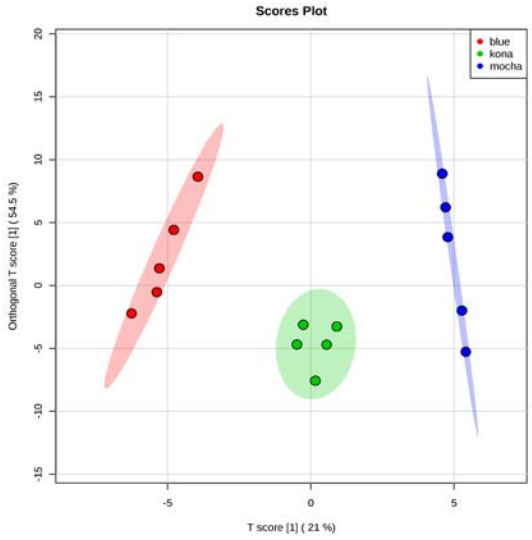

Figure 1. OPLS-DA score plot of three groups. The red represents the 'blue' sample, the green represents the 'kona' and the blue represents the 'mocha'. The ellipse of each group is 95 % confidence region. OPLS-DA for these three groups showed clear separation between three groups.

PCA was performed using the processed bin data. The PCA scatter plot among the two principal components covers 86 % of the quantified bin data. The samples of three groups (blue-kona-mocha) were within the border of the 95 % confidence ellipse and we could not find any reason that these samples can be considered outliers. Thus, these samples were included for further statistical analysis (data not shown).

OPLS-DA which is a supervised statistical method was employed to clarify the separation between groups. OPLS-DA score plots showed the clear clustering between three groups (Figure 1). The components with predictive abilities were R2Y=0.981 and Q2=0877, which means the model has good predictability and goodness of fit. Furthermore, The efficiency and reliability of OPLS-DA models were validated by an external validation and a permutation test. The observed R2Y and Q2 values of the OPLS-DA model were higher and lie in the upper extremes in the permutated distribution, with the empirical *p*-value of R2Y < 0.001 and *p*-value of Q2 < 0.001 (Figure 2).

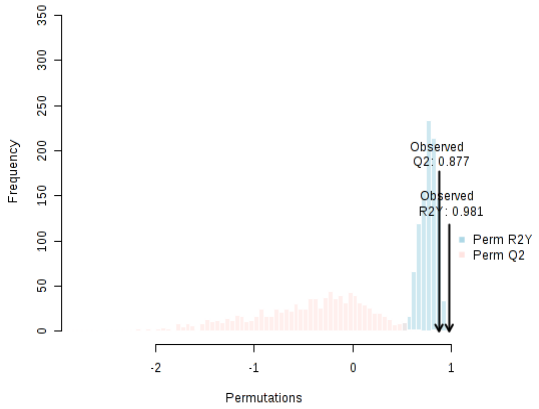

Figure 2. The result of permutation test. The OPLS-DA model was evaluated with the 1,000 permutation test. The red bars represent the permutated Q2 distribution and the blue bars represent the permutated R2Y distribution. The evaluated *p*-vlaues of the observed R2 and Q2 was below 0.05, which means the model could have a potential for discrimination between three groups.

Therefore, the metabolic states of each bean should be quite different and could be used the differentiation of beans.

The univariate analysis using ANOVA- To further assess the statistical capability of the metabolites to differentiate between groups, the mean values of spectral bins were compared between groups using ANOVA test. The result of univariate analysis, that is, ANOVA test was followed by Bonferroni's correction for multiple comparisons among the three groups. The adjusted *p*-values were corrected by the multiple testing corrections, Benjamini and Hochberg FDR. Null hypotheses of no difference were rejected if adjusted p-values were less than 0.05. As a result, as shown in the Table 1, all bins listed in the Table 1 differed significantly between groups. Among the total 104 spectral bins, 55 bins showed significant changes between groups, which revealed the great differences exist in the metabolic states of beans.

Among the perturbed bins, we could clearly identify five metabolites, 2-hydroxybutyrate, citrate, caffeine, sucrose, and chlorogenic acid. The resonance assignment was performed using Chenomx NMR suite 7.7 as well as ^{1}H-^{13}C HSQC and 2D ^{1}H-^{1}H

TOCSY spectra.[18]

The bar charts of identified metabolite are shown in Figure 3. 2-hydroxybutyrate, which is related to oxidative stress in cell metabolism, was highly up-regulated in Mocha beans while citrate, caffeine, and chlorogenic acid were lower than those of other beans. Kona beans showed the relatively high level of citrate and caffeine among beans. Blue beans possess the high level of caffeine, sucrose, and chlorogenic acid.

In conclusion, our study demonstrated that the NMR-based metabolomics could be a useful tool for finding difference among three coffee groups and there are differences in the coffee bean metabolites which are from different regions although they are same species. Multivaraiate analysis, OPLS-DA, was efficient for the classification of coffee beans according to their different regions. When compared with Kona and Blue Mountain, concentration of chlorogenic acid, caffeine, and sucrose in Mocha Matari was observed to be relatively down-regulated. Also, when compared with the two other groups, concentration of citrate in Kona was observed to be up-regulated.

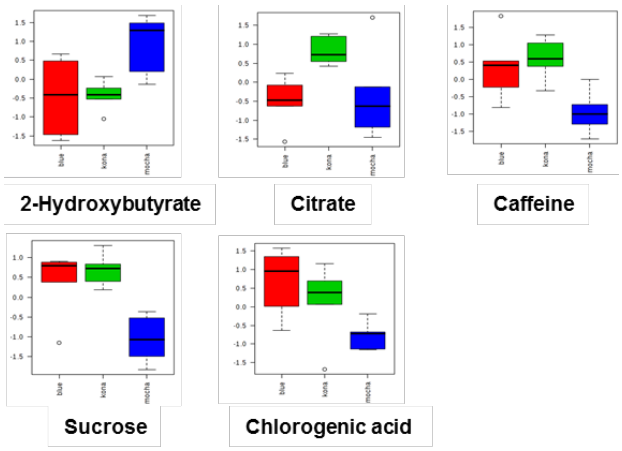

Figure 3. Five metabolites with significant difference between the three groups. Box and whisker plots of five metabolites using quantified concentrations are illustrated. Fight metabolites including 2-hydroxybutyrate, citrate, caffein, sucrose, and Chlorogenic acid showed significant changes. The horizontal line in the middle portion of the box is median value. The bottom and top boundaries of boxes represent lower and upper quartile. The open circles represent outliers.

Table 1. ANOVA result between three groups

Bin (ppm)	*p*-value	FDR	Fisher's LSD (Post hoc)
4.23 .. 4.21	2.50E-07	2.20E-05	kona - blue; blue - mocha; kona - mocha
8.59 .. 8.56	7.52E-06	0.000331	blue - mocha; kona - mocha
4.25 .. 4.23	1.90E-05	0.000507	kona - blue; blue - mocha; kona - mocha
2.59 .. 2.56	2.30E-05	0.000507	kona - blue; mocha - blue; kona - mocha
6.36 .. 6.31	2.90E-05	0.00051	kona - blue; blue - mocha; kona - mocha
2.56 .. 2.53	4.43E-05	0.000649	kona - blue; mocha - blue; kona - mocha
5.43 .. 5.37	5.16E-05	0.000649	kona - blue; blue - mocha; kona - mocha
4.37 .. 4.34	0.000107	0.001132	kona - blue; blue - mocha; kona - mocha
3.33 .. 3.32	0.000116	0.001132	kona - blue; mocha - blue
4.29 .. 4.25	0.000298	0.002531	kona - blue; blue - mocha; kona - mocha

7.09 .. 7.01	0.000331	0.002531	blue - mocha; kona - mocha
7.16 .. 7.09	0.000368	0.002531	blue - mocha; kona - mocha
2.23 .. 2.20	0.000379	0.002531	kona - blue; blue - mocha; kona - mocha
5.37 .. 5.30	0.000403	0.002531	kona - blue; blue - mocha; kona - mocha
6.93 .. 6.85	0.000453	0.002657	blue - mocha; kona - mocha
3.48 .. 3.47	0.000488	0.002682	kona - blue; blue - mocha; kona - mocha
8.79 .. 8.75	0.000542	0.002805	mocha - blue; mocha - kona
8.42 .. 8.40	0.000657	0.003211	blue - mocha; kona - mocha
9.46 .. 9.42	0.000705	0.003224	kona - blue; kona - mocha
3.70 .. 3.69	0.000733	0.003224	blue - mocha; kona - mocha
6.68 .. 6.66	0.000979	0.003767	kona - blue; kona - mocha
4.71 .. 4.69	0.001012	0.003767	kona - blue; kona - mocha
2.20 .. 2.18	0.001026	0.003767	kona - blue; blue - mocha; kona - mocha
4.09 .. 4.06	0.001027	0.003767	blue - mocha; kona - mocha
2.04 .. 2.02	0.001272	0.004478	kona - blue; blue - mocha; kona - mocha
2.26 .. 2.23	0.001355	0.004585	kona - blue; blue - mocha; kona - mocha
4.67 .. 4.66	0.002204	0.007183	kona - blue; kona - mocha
2.15 .. 2.12	0.002326	0.00731	kona - blue; blue - mocha; kona - mocha
8.55 .. 8.49	0.003063	0.00924	mocha - blue; mocha - kona
2.07 .. 2.04	0.00315	0.00924	blue - mocha; kona - mocha
3.26 .. 3.24	0.006479	0.018146	blue - mocha; kona - mocha
3.75 .. 3.71	0.006599	0.018146	kona - mocha
4.06 .. 4.00	0.007635	0.020359	kona - mocha
3.21 .. 3.19	0.008532	0.021682	blue - mocha; kona - mocha
3.71 .. 3.70	0.008776	0.021682	mocha - blue; mocha - kona
4.32 .. 4.30	0.008873	0.021682	kona - blue; kona - mocha
2.79 .. 2.78	0.009116	0.021682	kona - blue; kona - mocha
1.16 .. 1.13	0.009852	0.022814	kona - blue; mocha - blue
8.10 .. 8.05	0.010218	0.023056	kona - blue; kona - mocha
2.02 .. 2.00	0.010589	0.023295	kona - mocha
2.18 .. 2.15	0.011857	0.024414	blue - mocha; kona - mocha
2.89 .. 2.88	0.01191	0.024414	mocha - blue; mocha - kona
8.49 .. 8.44	0.012165	0.024414	blue - mocha; kona - mocha
5.55 .. 5.52	0.012207	0.024414	blue - mocha; kona - mocha
5.88 .. 5.87	0.013022	0.024954	blue - mocha; kona - mocha
3.97 .. 3.95	0.013048	0.024954	blue - mocha; kona - mocha
7.83 .. 7.76	0.013328	0.024954	kona - mocha
3.49 .. 3.48	0.016117	0.029548	kona - blue; kona - mocha
2.32 .. 2.26	0.01721	0.030909	kona - mocha
1.89 .. 1.87	0.019504	0.03367	kona - mocha
4.74 .. 4.72	0.019514	0.03367	kona - mocha
7.95 .. 7.93	0.020022	0.033884	blue - kona; blue - mocha
8.87 .. 8.79	0.0223	0.036848	kona - blue; kona - mocha
3.18 .. 3.16	0.022611	0.036848	blue - mocha; kona - mocha
3.57 .. 3.55	0.023099	0.036958	blue - mocha

Acknowledgements

This research was supported by the Bio & Medical Technology Development Program of the NRF funded by the Korean government, MSIP (NRF-2014M3A9B6069340). This work was also supported by grants from the Basic Science Research Program (2013R1A1A2064418 and 2016R1D1A1B03931783) through the National Research Foundation of Korea, funded by the Ministry of Education, Science, and Technology.

References

1. J. L. Griffin, *Curr. Opin. Chem. Biol.* **7**, 648 (2003)
2. D. S. Wishart, *TrAC-Trend Anal. Chem.* **27**, 228 (2008)
3. M. Li, J. Wang, Z. Lu, D. Wei, M. Yang, and L. Kong, *Aquat. Toxicol.* **146**, 82 (2014)
4. L. M. Samuelsson, L. Förlin, G. Karlsson, M. Adolfsson-Erici, and D. J. Larsson, *Aquat. Toxicol* **78**, 341 (2006)
5. A-H. Emwas, C. Luchinat, P. Turano, L. Tenori, R. Roy, and R. M. Salek, *Metabolomics* **11**, 872 (2015)
6. D. Kumar, A. Gupta, and K. Nath, *Expert Rev. Mol. Diagn.* **16**, 651 (2016)
7. A-H. Emwas, R. Roy, R. T. McKay, D. Ryan, L. Brennan, and L. Tenori, *J. Proteome Res.* **15**, 360 (2016)
8. A. W. Chan, P. Mercier, D. Schiller, R. Bailey, S. Robbins, and D. T. Eurich, *Brit. J. Cancer* **114**, 59 (2016)
9. Y. Tian, X. Nie, S. Xu, Y. Li, T. Huang, H. Tang, and Y. Wang, *Sci. Rep.* **5**, 14869 (2015)
10. J-S. Hyun and S. J. Park, *J. Kor. Magn. Reson. Soc.* **20**, 76 (2016)
11. J. Xia, N. Psychogios, N. Young, and D. S. Wishart, *Nucleic Acids Res.* **37**, W652 (2009)
12. J. Xia, I. V. Sinelnikov, B. Han, and D. S. Wishart, *Nucleic Acids Res.* **43**, W251 (2015)
13. J. Trygg, E. Holmes, and T. Lundstedt, *J. Proteome Res.* **6**, 469 (2007)
14. A. Kullgren, L. M. Samuelsson, D. J. Larsson, B. T. Björnsson, and E. J. Bergman, *Am. J. Physiol-Reg. I.* **299**, R1440 (2010)
15. J. Kang, M-Y. Choi, S. Kang, H. N. Kwon, H. Wen, and C. H. Lee, *J. Agric. Food Chem.* **56**, 11589 (2008)
16. J. D. Spurrier, *J. Nonparametr. Stat.* **15**, 685 (2003)
17. Y. Benjamini and Y. Hochberg, *J. R. Stat. Soc. B.* **57**, 289 (1995)
18. J-S. Hyun and S. J. Park, *J. Kor. Magn. Reson. Soc.* **19**, 149 (2015)