MOLEKUL
eISSN: 2503-0310

Articles

https://doi.org/10.20884/1.jm.2022.17.1.5603

# Discrimination of Metabolite Profiles of Gayo Roasted Arabica and Robusta Coffees

**Nizar Happyana*, Yana Maolana Syah, Euis Holisotan Hakim**

Organic Chemistry Division, Faculty of Mathematics and Natural Sciences, Bandung Institute of Technology, Jl. Ganesha 10, Bandung, West Java, Indonesia

*Corresponding author email: nizar@chem.itb.ac.id

**Received** August 16, 2021; **Accepted** December 29, 2021; **Available online** March 20, 2022

**ABSTRACT.** Gayo (Aceh) coffee is one of the best coffees from Indonesia. In this work, metabolites in the Gayo roasted arabica and robusta coffees were identified with $^1$H NMR spectroscopy analysis. Accumulatively 28 compounds were successfully detected, including the major and minor metabolites of the roasted coffee. Multivariate data analysis was applied to evaluate the dataset extracted from $^1$H NMR spectra of the coffee samples, result in the disclosure of the differences in the chemical profiles between the Gayo roasted coffees of arabica and robusta. Score plots obtained from the models of principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLSDA), classified the roasted coffee samples based on their species. Loading plot and S-plot revealed the discriminant compounds for each coffee. Gayo roasted arabica coffee was characterized with acetic acid and trigonelline, while the robusta coffee was discriminated with fatty acids. This report revealed the chemical differences of both coffees and confirmed the diversity of Gayo coffees.

**Keywords**: $^1$H NMR-based metabolomics, arabica coffee, Gayo roasted coffee, multivariate data analysis, robusta coffee

## INTRODUCTION

Coffee is one of the most popular beverages in the world. Over 3 billion cups of coffee are consumed in the world every day. The high coffee consumption is mainly caused by social, historical and cultural factors, the distinctive flavor, and the health benefits (Ayelign & Sabally, 2013). International Coffee Organization (ICO) reported that the coffee production increased annually from 100 million bags (60-kg) in 1990 to over 165 million bags in 2018 (ICO, 2019). The coffee plant is only cultivated in "the bean belt" region and produced by more than 50 countries, involving around 25 million farmers (ICO, 2019).

Indonesia is the fourth biggest coffee producers and exporters in the world after Brazil, Vietnam and Columbia (ICO, 2019). In 2017, Indonesia produced 717,962 tons of coffee (Crops, 2018) and exported around 467 thousand tons of those worldwide with the total value of 1.18 billion USD (Statistics, 2018). Two major coffee species cultivated in Indonesia are *Coffea canephora* (robusta, around 70%) and *Coffea arabica* (arabica, almost 28%) (Crops, 2018). Aceh is one of the main coffee producers in Indonesia. In 2017, the coffee plantation in Aceh was around 123 thousand acres producing almost 69 thousand tons of the coffee (Crops, 2018). Most coffee cultivated in Aceh (Gayo) is arabica (almost 82%) and the rest is robusta (Crops, 2018). Interestingly, Gayo arabica coffee recently received a geographical indication certification from European Union (EU, 2016). It indicated this coffee is a popular coffee not only in Indonesia but also in international. The taste of this coffee is recognized with a strong body, a low acidity, and savory (Wahyudi & Jati, 2012).

Recently, metabolite profile of Gayo roasted arabica coffee was evaluated and differentiated from the other Indonesian arabica coffees (Happyana, Hermawati, Syah, & Hakim, 2020a). This roasted coffee was characterized with acetic acid, lactic acid, and lipids (Happyana et al., 2020a). Meanwhile, metabolite profile of Gayo roasted robusta coffee was successfully discriminated from the Lampung robusta coffee (Happyana, Hermawati, Syah, & Hakim, 2020b). Interestingly, lactic acid and lipids were also identified as the most discriminant metabolites for Gayo robusta coffee (Happyana et al., 2020b). Both compounds were potential markers for either arabica or robusta coffee from Gayo, Aceh. However, the differences in the metabolite profiles, between Gayo arabica and robusta coffees, were still not well-explained yet.

Various studies had been performed to evaluate chemical contents of both arabica and robusta coffees. Arabica coffee was discriminated from robusta coffee based on their tocopherol profiles (Alves, Casal, Alves, & Oliveira, 2009), volatile organic compounds (Konieczka, Aliaño-González, Ferreiro-González, Barbero, & Palma, 2020 ; Procida, Lagazio, Cateni, Zacchigna, & Cichelli, 2020), contents of chlorogenic acids (Bicchi, Binello,

Pellegrino, & Vanni, 1995), levels of trigonelline, nicotinic acid, and caffeine (Casal, Oliveira, Alves, & Ferreira, 2000), concentrations of amino acids (Casal, Alves, Mendes, Oliveira, & Ferreira, 2003), and profile of fatty acids (Alves, Casal, Oliveira, & Ferreira, 2003). Comparative studies of more widely compound composition between arabica and robusta coffee had been reported as well. Metabolite profiles of arabica and robusta coffees were compared by HPLC analysis coupled with a chemometric approach (Nuñez, Collado, Martínez, Saurina, & Nuñez, 2020). Chemical profiles of arabica and robusta were analyzed with ESI(-)FT-ICR MS and ATR-FTIR (Correia et al., 2016). A quantitative NMR method was applied for the authentication of roasted arabica and robusta coffees, and prediction of the blends (Burton et al., 2020). The performances of ATR-mIR, NIR, and NMR spectroscopies combined with the multivariate data analysis, were compared in discriminating the metabolite profile of arabica coffee from the robusta coffee (Medina et al., 2017).

This study aims to compare the metabolite profiles of the roasted coffees of arabica and robusta originated from Gayo, Aceh, Indonesia. Roasted coffee samples were extracted with deuterated water and then analyzed with $^1$H NMR spectroscopy. The obtained data were further evaluated with multivariate data analysis to reveal the differences in the samples. This study shed more light on the uniqueness of metabolite profiles of each Gayo coffee.

## EXPERIMENTAL SECTION

### Materials

Gayo roasted coffee beans used as the samples in this work were obtained from various coffee suppliers and the detail information was described in **Table 1**. Deuterium oxide ($D_2O$) used for the extraction, sodium-3-(trimethylsilyl)-2,2,3,3-tetradeuteriopropionate (TSP) applied for the calibration of the chemical shift, $KH_2PO_4$ and $K_2HPO_4$ as the buffer solution, were bought from Merck (Darmstadt, Germany).

### Extraction

The roasted coffee beans were ground with an Encore mill (Baratza, United States). The powder of roasted coffee in a 2 mL plastic tube was mixed with 1 mL of $D_2O$ containing TSP (1.00 mM). The tube was closed properly, sonicated with an ultrasonic bath (Krisbow, Jakarta, Indonesia) for 20 minutes at 27 °C, and then incubated in a hot water for 30 minutes at 90 °C. The sample was chilled on the water (room temperature) for 10 minutes. Afterward, it was centrifuged with an MC-12 High Speed Microcentrifuge (Benchmark Scientific, United States) at 12,000 rpm, room temperature for 5 minutes. 400 mL of the supernatant was mixed with 100 µL of phosphate buffer (pH 5) and then placed in a 5 mm NMR tube.

### $^1$H NMR Measurements

$^1$H NMR spectra acquisition were conducted with a 500 MHz Varian Unity INOVA spectrometer (Agilent Technologies, United States). $^1$H NMR measurements were performed with a presaturation method to suppress the $H_2O$ signal. The spectra were recorded into 64 K complex points over an 8012 Hz spectral width and 128 scans. The acquisition time and the recycle delay were 2.72 s, and 2 s, respectively. The free-induction decay (FID) files of obtained $^1$H NMR spectra were further processed using ACD/Labs 12.0 software (Advanced Chemistry Development, Inc., Toronto, Canada). The chemical shift calibration was performed by referencing to the TSP signal.

### Data Extraction of $^1$H NMR Spectra

All $^1$H NMR spectra were aligned and processed with the bucketing technique using ACD/Labs 12.0 software. Bucketing was carried out by cutting the spectra within the region 0.50-10.00 ppm, into integrated bins with the equal width of 0.04 ppm. In this process, intelligent bucketing mode was operated. The buckets containing residual water signal (δ 4.73 - 5.22 ppm) were removed.

**Table 1.** Origins of arabica (A1-A6) and robusta (R1-R6) coffees

| Sample code | Coffee origin | Supplier |
|---|---|---|
| A1 | Atu Lintang, Aceh Tengah | Ottencoffee |
| A2 | Takengon, Aceh Tengah | JPW Coffee |
| A3 | Takengon, Aceh Tengah | Infokopi |
| A4 | Takengon, Aceh Tengah | Fulcaf Coffee |
| A5 | Bintang, Aceh Tengah | Mr. O Coffee |
| A6 | Takengon, Aceh Tengah | Coffindo |
| R1 | Blangkejeren, Gayo Lues | Fry Roast |
| R2 | Linge, Aceh Tengah | Rebbe Coffee |
| R3 | Pintu Rime Gayo, Bener Meriah | Serenade |
| R4 | Takengon, Aceh Tengah | Tampah Kopi Gayo |
| R5 | Takengon, Aceh Tengah | Raja Kopi Aceh |
| R6 | Pintu Rime Gayo, Bener Meriah | Garasco |

The buckets at δ 3.22 - 3.49 ppm and δ 3.82 - 3.88 ppm corresponding to caffeine were also excluded since the changeable signals. The obtained buckets were then normalized with the sum technique to neglect the bias results.

### Multivariate Statistical Analysis

The normalized data were exported into SIMCA-P version 12.0 (Umetrics, Umeå, Sweden) for the multivariate statistical analysis. To reduce the mask effect in the analysis, the normalized data were scaled with Pareto technique. Principal component analysis (PCA) technique was applied for analyzing metabolite profiles of the roasted coffee samples. The total variation explained by the model ($R^2X$) and the cumulative variation driven from the cross validation ($Q^2$) were computed. Hotelling's T2 regions, described as an ellipse in the score plot, explained the 95% confidence interval of the PCA model variation. Orthogonal projection to latent structure-discriminant analysis (OPLSD) model was created to achieve a better coffee discrimination.

## RESULTS AND DISCUSSION

### Detected Metabolites

The obtained $^1$H NMR spectra were evaluated for detecting metabolites in the roasted coffees, either arabica or robusta samples. The metabolites were identified by detecting their fingerprint peaks in the $^1$H NMR spectra. The detected metabolite signals were further verified by comparing with the data from the literature (Burton et al., 2020; Consonni, Cagliani, & Cogliati, 2012; Toci et al., 2017; Wei, Furihata, Hu, Miyakawa, & Tanokura, 2011; Wei, Furihata, Miyakawa, & Tanokura, 2014) and the human metabolome database (HMDB, www.hmdb.ca). As the results, 28 metabolites were successfully identified in the spectra. $^1$H NMR spectra of roasted coffee samples were depicted in **Figure 1**.

The signals correspondence to trigonelline were obviously detected in the spectra, confirming as the major compound in the roasted coffees. The strong singlet peak at δ 4.43 ppm was assigned as the proton signal (H-8) of trigonelline methyl group. Another singlet signal at δ 9.12 ppm was designed as an aromatic proton (H-2) of trigonelline. Meanwhile, other aromatic protons of trigonelline were detected at δ 8.82 (*m*) and 8.84 (*m*) ppm belonging to H-6 and H-4, respectively, and at δ 8.07 ppm with the triplet multiplicity corresponded to H-5. The molecular structure of trigonelline could be found in **Figure 2**. Interestingly, 2 degradation products of trigonelline yielded by the roasting process (Wei et al., 2012), including *N*-methyl-pyridinium and nicotinic acid, were successfully detected in the $^1$H NMR spectra. The proton signals of *N*-methyl-pyridinium in the $^1$H NMR spectra were found at δ 4.37, 8.02, 8.51 and 8.75 ppm, while the peaks of nicotinic acid were discovered at δ 8.27, 8.66 and 8.97 ppm.

Caffeine signals were explicitly detectable in the spectra, indicating as another major compound in the roasted coffees. Three methyl groups of caffeine were clearly recorded as singlet peaks in the aliphatic region at δ 3.28, 3.45 and 3.88 ppm, assigned as H-11, H-10, and H-12, respectively. Meanwhile, the aromatic proton of caffeine was observed as singlet peak at δ 7.77 ppm correspondence to H-8.

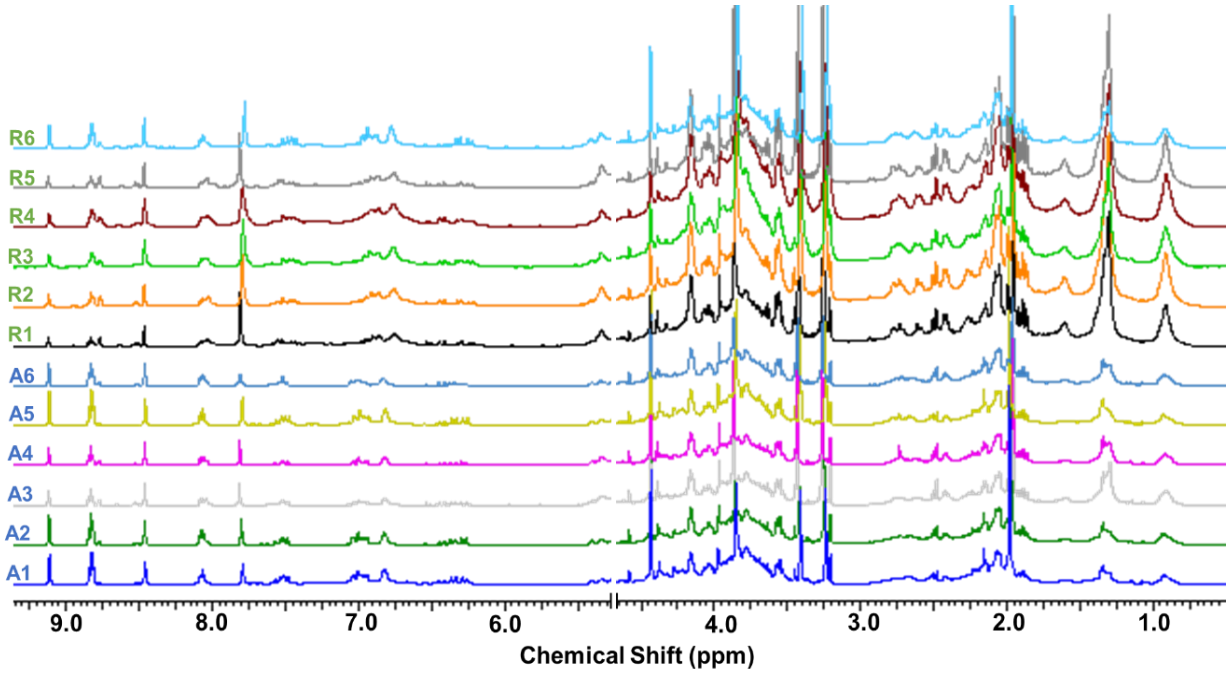



**Figure 1.** $^1$H NMR spectra of the roasted Gayo coffee samples consisting of arabica (A1-A6) and robusta (R1-R6) coffees.

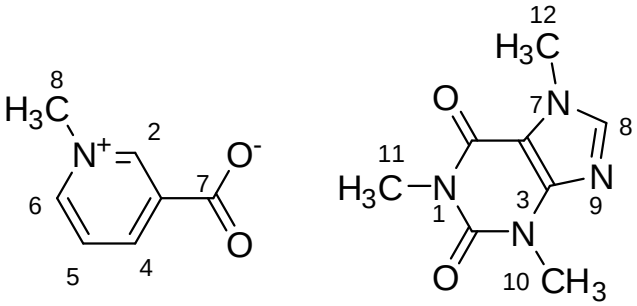



**Figure 2.** Molecular structures of trigonelline and caffeine identified in the $^1$H NMR spectra of the roasted coffee samples.

**Table 2.** $^1$H NMR signals of the identified compounds in the Gayo roasted coffee samples.

| Compound | Chemical shift (ppm) |
|---|---|
| 3-arabinose | 4.27 (br s), 5.25 (br s) |
| 5-arabinose | 4.21 (br s), 5.10 (br s) |
| 3-caffeoylquinic acid | 2.04 (m), 2.16 (m), 5.40 (m), 6.35 (d), 6.82 (br s), 6.96 (m), 7.02 (m), 7. 51 (m) |
| 4-caffeoylquinic acid | 2.04 (m), 2.16 (m), 4.92 (m), 6.35 (d), 6.82 (br s), 6.96 (m), 7.02 (m), 7.53 (m) |
| 5-caffeoylquinic acid | 2.04 (m), 2.16 (m), 5.33 (m), 6.27 (d), 6.82 (br s), 6.96 (m), 7.02 (m), 7.48 (m) |
| 2-furylmethanol | 4.58 (s), 6.43 (m), 7.57 (br s) |
| 5-(hydroxymethyl)furfural | 9.49 (s) |
| 3-galactose | 3. 65 (m), 4.62 (d) |
| 6-galactose | 3.73 (m), 4.44 (br s) |
| alpha-linoleic acid | 0.98 (m), 1.31(m), 1.65 (m), 2.04 (m), 2.35 (m), 2.78 (m), 5.32 (m) |
| Acetic acid | 1.98 (s) |
| Caffeine | 3.26 (s), 3.45 (s), 3.88 (s), 7.77(s) |
| Catechol | 6.65-6.70 (d) |
| Choline | 3.20 (s) |
| Citric acid | 2.61 (d), 2.74 (d) |
| Formic acid | 8.46 (s) |
| Lactic acid | 1.35 (d) |
| linoleic acid | 0.92 (m), 1.32 (m), 1.59 (m), 2.08 (m), 2.34 (m), 2.76 (m), 5.32 (m) |
| Lipids | 0.92 (m), 1.30 (m) |
| Malic acid | 2.36 (m), 2.68 (m) |
| Mannose | 3.55 (m), 3.82 (m), 3.93 (m), 5.17 (br s) |
| Inositol | 3.27 (t), 3.52 (m), 3.62 (m), 4.06 (m) |
| *N*-methyl-pyridine | 4.37 (s), 8.02 (m), 8.52 (t) and 8.77 (d) |
| Palmitic acid | 0.88 (m), 1.29 (m), 1.64 (m), 2.36 (m) |
| Quinic acid | 1.89 (m), 1.96 (m), 2.06 (m), 3.56 (m), 4.03 (m), 4.16 (m) |
| γ-quinide | 1.95 (m), 2.14 (m), 2.41 (m), 2.49 (m), 3.89 (m), 4.06 (m), 4.91 (m) |
| Stearic acid | 1.01 (m), 1.43 (m), 1.74 (m), 2.48 (m) |
| Trigonelline | 4.43 (s), 8.09 (t), 8.82 (m), 8.84 (m), 9.12 (s) |

The chemical shifts of these strong singlet peaks of caffeine were changeable since forming complex structure with chlorogenic acids as reported in previous work (D'Amelio, Fontanive, Uggeri, Suggi-Liverani, & Navarini, 2009). The molecular structure of caffeine was described in **Figure 2**. The other intense singlet peaks were also detected at δ 1.95 ppm, assigned as the proton signal of acetic acid, and at δ 3.20 ppm correlated to methyl groups of choline. Acetic acid was sucrose degradation product caused by the roasting process. Fortunately, the signals of other compounds yielded from sucrose degradation were also successfully identified in the spectra, including formic acid at δ 8.46 ppm, lactic acid at δ 1.35 ppm, 2-furyl-methanol at δ 4.58, 6.43, and 7.57 ppm, and 5-(hydroxymethyl)-furfural at δ 9.49 ppm.

Caffeoylquinic acids (CQA) or well-known with chlorogenic acids, are other major metabolites in the coffee sample. Three CQA isomers, including 3-caffeoylquinic acid (3-CQA), 4-caffeoylquinic acid (4-CQA) and 5-caffeoylquinic acid (5-CQA) were successfully identified in the $^1$H NMR spectra of the roasted coffees. The detail proton chemical shifts of these metabolites were described in **Table 2**. Quinic acid, a precursor of CQAs, together with its ester cyclic form, γ-quinide, was recorded in the spectra as well. The proton signals at δ 1.89, 1.96, 2.06, 3.56, 4.03, and 4.16 ppm were assigned as quinic acid protons, while the signals at δ 1.95, 2.14, 2.41, 2.49, 3.89, 4.06, and 4.91 ppm were related to γ-quinide protons.

Further analysis in the aliphatic regions successfully identified sugar compounds including arabinose, galactose, and mannose units. Signals correlated to arabinose units were observed at δ 4.27 and 5.25 ppm assigned as 3-arabinose protons, and at δ 4.21 and 5.10 ppm corresponded to 5-arabinose protons. Signals at δ 3.65 and 4.62 ppm were designed to 3-galactose unit, while at δ 3.73 and 4.44 ppm were correlated to 6-galactose unit. Meanwhile proton signals of mannose unit were detected at δ 3.55, 3.82, 3.93, and 5.17 ppm. Proton signals of *myo*-inositol, another identified sugar compound, were also successfully detected at δ 3.27, 3.52, 3.62, and 4.06 ppm.

Proton signals of fatty acids were recorded as well, especially in the $^1$H NMR spectra of robusta coffee. The signals at δ 0.88 (methyl), 1.29 (methylene), 1.64 (methylene closer to carboxylic acid group), and 2.36 ppm (methylene closest to carboxylic acid group), were assigned as the protons of palmitic acid. The signals of stearic acid were observed at δ 1.01 ppm designed as methyl proton, δ 1.43 ppm corresponded to methylene, δ 1.74 ppm assigned as the methylene closer to carboxyl, and δ 2.48 ppm correlated to methylene the closest to carboxyl. The signals belong to alpha-linoleic acid, and linoleic acid were depicted in **Table 2**. Meanwhile, the other identified metabolites in the $^1$H NMR spectra of coffee samples were catechol at δ 6.67 ppm, citric acid at δ 2.61 and 2.74 ppm, and malic acid at δ 2.36 and 2.68 ppm.

## Discrimination of Arabica and Robusta Roasted Gayo Coffees

In this work, metabolite profiles of roasted arabica and robusta coffees originated from Gayo (Aceh), were compared. The cleaned dataset extracted from $^1$H NMR spectra, were further evaluated with multivariate data analysis using SIMCA-P version 12.0 (Umetrics, Umeå, Sweden). Initially, the data was evaluated with PCA, unsupervised approach. PCA is a data reduction method describing the multivariate data in a low-dimensional space (Happyana, Muntendam, & Kayser, 2012). PCA able to reveal the relationships and variances in the data, produce a model of how chemical system behave, and separate an underlaying systematic data from noise (Wold, Esbensen, & Geladi, 1987).

The resulted PCA model possessed 3 new principal components (PCs) explaining 87.5% of cumulative variances ($R^2X$) in which the first 2 PCs described 80.8% of variation in the spectral data. This model had 75.1% of cross validation coefficient ($Q^2$) indicating a good predictive ability. The PCA score plot built by combining PC1 (66.0%) and PC2 (14.8%), provided a clear cluster separation within the roasted coffee samples as described in **Figure 3a**. PC1 could almost separate Gayo roasted arabica coffee from the counterpart. Gayo arabica coffee samples were explicitly grouped on the left while almost all Gayo robusta coffee samples located on the right of the PC1 axis.

Loading column plot of PC1 was evaluated to reveal the responsible buckets in the coffee sample separation on the PCA score plot. In this loading column plot, the buckets contributed on the Gayo roasted robusta coffee were on the positive PC1 axis, while the correlated buckets with arabica coffee samples, presented on the negative PC1 axis. As seen in **Figure 3b**, the roasted robusta coffee were characterized with the buckets belong to fatty acids, including alpha-linoleic acid, linoleic acid, palmitic acid, and stearic acid. Overlapped signals in the buckets at δ 0.88-0.94, 0.94-1.00, 1.27-1.29, 1.29-1.35, 1.35-1.41, 1.57-1.63, and 1.63-1.69 ppm corresponded to these fatty acids, were observed contributing positively to robusta coffee samples. Meanwhile, buckets at δ 4.41-4.47, 8.02-8.08, 8.08-8.14, 8.43-8.49, 8.79-8.95, and 9.09-9.15 ppm attributed to trigonelline signals, were found as important characteristic buckets for roasted arabica Gayo coffees. Buckets at δ 3.76-3.80 and 3.92-4.00 ppm related to galactoses and mannose, respectively, were also detected contributing to the arabica coffee samples. The other bucket responsible for arabica coffee was the bucket at δ 1.94-2.00 ppm assigned to acetic acid.

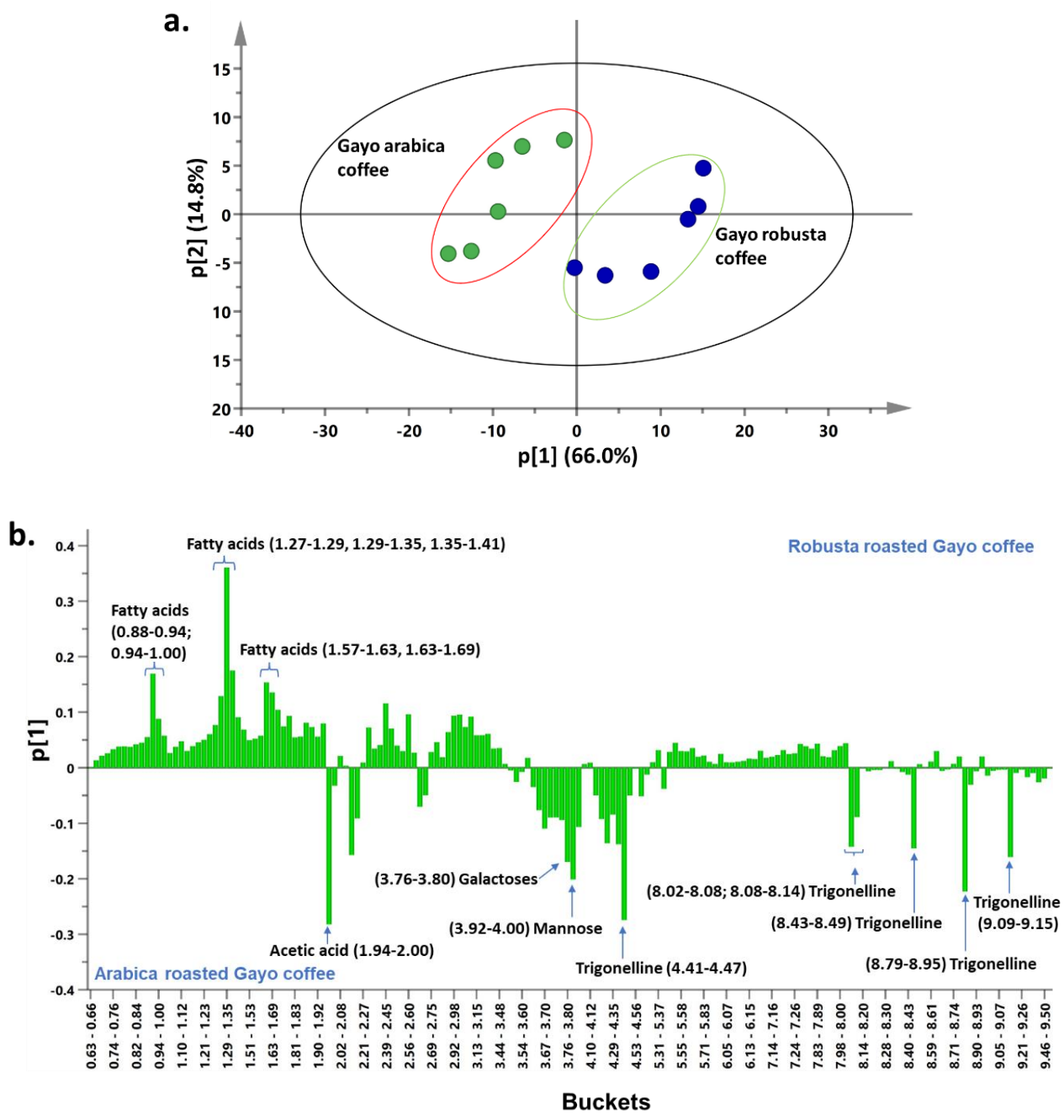
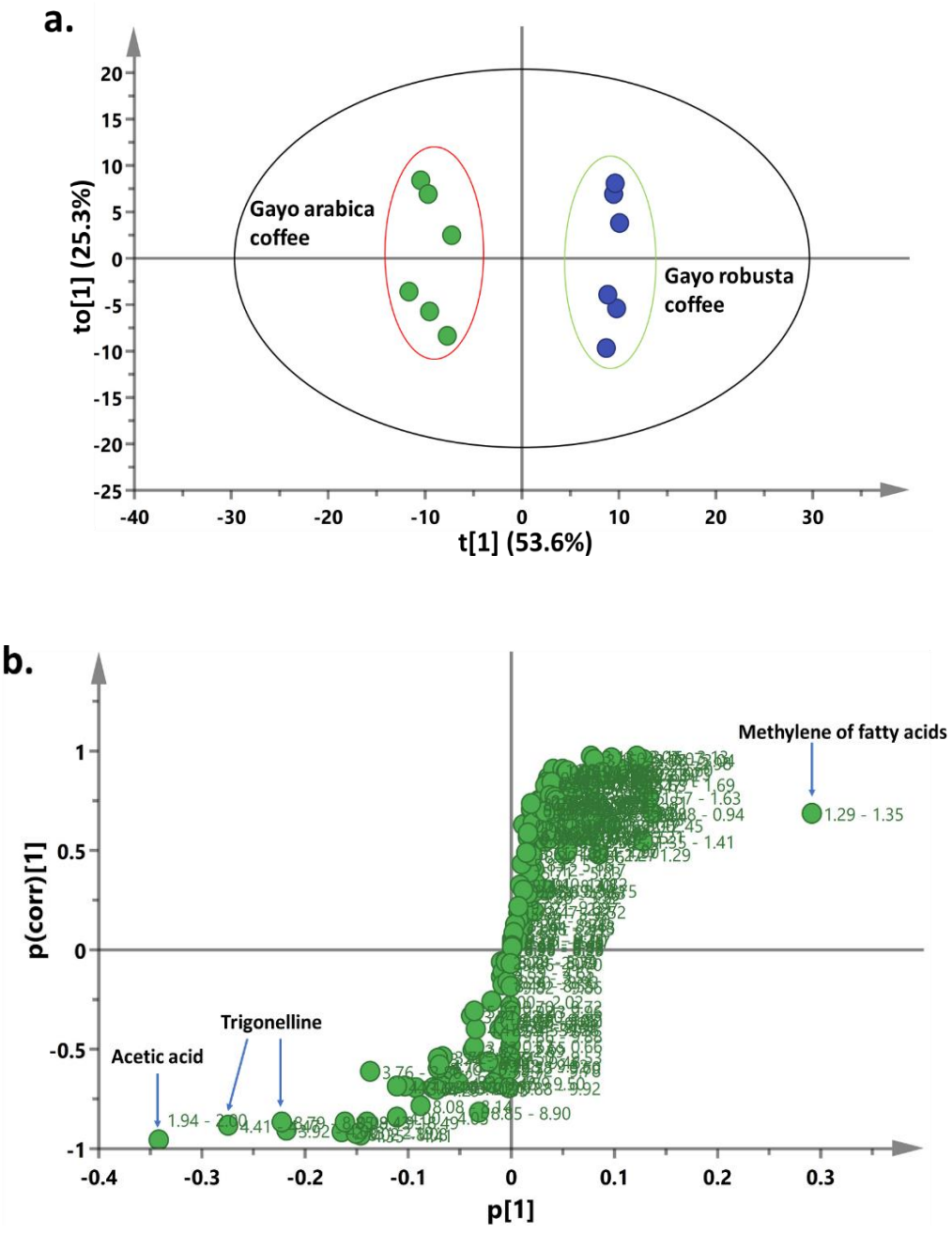



**Figure 3.** Score (**a**) and column loading (**b**) plots computed for PCA model of Gayo roasted arabica and robusta coffees.

To obtain better class separation, OPLSD model was created. OPLSDA is a supervised approach combining the advantages of partial least square discriminant analysis (PLSDA) and soft independent modelling of class analogy (SIMCA) classification (Bylesjo et al., 2006). This model had 3 components and described 84.0% and 98.6% of cumulative variations ($R^2X$ and $R^2Y$, respectively). This OPLSDA model possessed a good predictability with 86.3% of $Q^2$. Furthermore, it was validated with the permutation test using 200 iterations. As the result, the $Q^2$ regression lines crossed the y-axis at point below zero [$Q^2$ = (0.00, -0.73); $R^2$ = (0.00, 0.67)] indicating the validity of the model statistically. As described in **Figure 4a**, the OPLSDA score plot provided a better class separation between arabica and robusta coffees when compared to the PCA score plot. The buckets facilitated to the coffee discrimination in the PCA model, also contributed to the sample separation in the OPLSDA model as revealed by its loading plot (data not shown).

S-plot of the OPLSDA model (**Figure 4b**) was evaluated to discover the most important bucket for each coffee sample. As the results, bucket at δ 1.29-1.35 ppm was found as the most discriminant bucket for robusta coffee since located at the edge of the S-plot. This bucket was attributed to the methylene signal of fatty acids, including alpha-linoleic acid, linoleic

acid, palmitic acid, and stearic acid. Thus, this result indicated that the content of fatty acids was higher in the Gayo robusta coffee than in its counterpart. The levels of fatty acids in arabica and robusta were diverse depending on the fatty acid types, the varieties and the origins (Alves et al., 2003). However, the dispersion of fatty acid contents tended to be higher for the robusta than for arabica coffees (Alves et al., 2003). Meanwhile, the bucket at δ 1.94-2.00 ppm assigned to acetic acid was found as the most characteristic bucket for Gayo roasted arabica coffees. The higher concentration of acetic acid in the roasted arabica coffee, confirmed that the amount of sucrose (the precursor of acetic acid) was more abundance in the green bean of arabica coffee than in its counterpart. Buckets corresponded to trigonelline were other important buckets for Gayo roasted arabica coffee as depicted in **Figure 4b**. It indicated concentration of trigonelline was higher in the arabica than in the robusta coffees. Interestingly, these results were in accordance with reported works in the literature (Casal et al., 2000; Ky et al., 2001). The result of S-plot was supported by the VIP plot (**Figure 4c),** indicated that acetic acid, fatty acids and trigonelline were the most contributing buckets in discriminating metabolite profiles of Gayo roasted arabica and robusta coffees.



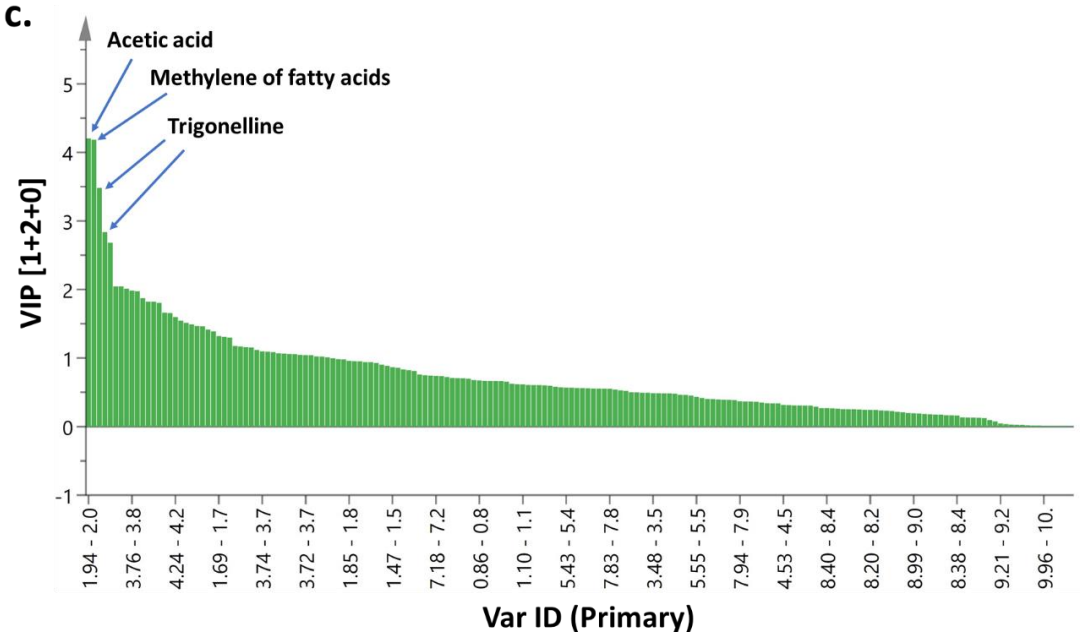



**Figure 4.** Score plot (a), S-plot (b), and VIP plot (c) of OPLSDA model of Gayo roasted arabica and robusta coffees.

## CONCLUSIONS

The differences in the metabolite profiles between arabica and robusta coffees originated from Gayo, were successfully evaluated with $^{1}$H NMR spectroscopy method combined with multivariate data analysis. In total, 28 compounds were successfully identified in both Gayo roasted coffees. Gayo roasted arabica coffee was characterized with the higher concentrations of acetic acid and trigonelline. Meanwhile, the robusta coffee was discriminated with higher concentration of fatty acids. This work confirmed the uniqueness of Gayo coffee either its arabica or robusta.

## ACKNOWLEDGMENTS

This research was funded by RISET ITB 2021 program from the Institute for Research and Community Services (LPPM), Bandung Institute of Technology (ITB).

## REFERENCES

Alves, M. R., Casal, S., Oliveira, M. B. P. P., & Ferreira, M. A. (2003). Contribution of FA profile obtained by high-resolution GC/chemometric techniques to the authenticity of green and roasted coffee varieties. *Journal of the American Oil Chemists' Society, 80*(6), 511-517.

Alves, R. C., Casal, S., Alves, M. R., & Oliveira, M. B. (2009). Discrimination between arabica and robusta coffee species on the basis of their tocopherol profiles. *Food Chemistry, 114*, 295–299.

Ayelign, A., & Sabally, K. (2013). Determination of chlorogenic acids (CGA) in coffee beans using HPLC. *American Journal of Research Communication, 1*, 78-91.

Bicchi, C. P., Binello, A. E., Pellegrino, G. M., & Vanni, A. C. (1995). Characterization of green and roasted coffees through the chlorogenic acid fraction by HPLC-UV and principal component analysis. *Journal of Agricultural and Food Chemistry, 43*, 1549-1555.

Burton, I. W., Farina, C. F. M., Ragupathy, S., T., A., Newmaster, S., & Berrué, F. (2020). Quantitative NMR methodology for the authentication of roasted coffee and prediction of blends. *Journal of Agricultural and Food Chemistry, 68*(49), 14643–14651.

Bylesjo, M., Rantalainen, M., Cloarec, O., Nicholson, J. K., Holmes, E., & Trygg, J. (2006). OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *Journal of Chemometrics, 20*(8-10), 341-351. doi: Doi 10.1002/Cem.1006

Casal, S., Alves, M. R., Mendes, E., Oliveira, M. B. P. P., & Ferreira, M. A. (2003). Discrimination between arabica and robusta coffee species on the basis of their amino acid enantiomers. *Journal of Agricultural and Food Chemistry, 51*, 6495–6501.

Casal, S., Oliveira, M. B. P. P., Alves, M. R., & Ferreira, M. A. (2000). Discriminate analysis of roasted coffee varieties for trigonelline, nicotinic acid, and caffeine content. *Journal of Agricultural and Food Chemistry, 48*, 3420-3424.

Consonni, R., Cagliani, L. R., & Cogliati, C. (2012). NMR based geographical characterization of roasted coffee. *Talanta, 88*, 420– 426.

Correia, M. R., Loureiro, L. B., Rodrigues, R. R. T., Costa, H. B., Oliveira, B. G., Filgueiras, P. R., Romão, W. (2016). Chemical profiles of robusta and arabica coffee by ESI(-)FT-ICR MS

and ATR-FTIR: a quantitative approach. *Analytical Methods, 8*, 7678-7688.

Crops, D. G. o. E. (2018). Tree crop estate statistics of Indonesia 2017-2019: coffee. Jakarta: Secretariate of Directorate General of Estate Crops, Ministry of Agriculture, Republic of Indonesia.

D'Amelio, N., Fontanive, L., Uggeri, F., Suggi-Liverani, F., & Navarini, L. (2009). NMR reinvestigation of the caffeine–chlorogenate complex in aqueous solution and in coffee Brews. *Food Biophysics, 4*, 321–330.

EU. (2016). Kopi arabika Gayo EU No: PGI-ID-02115 – 26.1.2016 *Official Journal of the European Union, C 403*(07).

Happyana, N., Hermawati, E., Syah, Y. M., & Hakim, H. H. (2020a). Discrimination of the Indonesian roasted arabica coffees using 1H NMR-based metabolomics. *Current Research in Nutrition and Food Science, 8*(2), 479-488.

Happyana, N., Hermawati, E., Syah, Y. M., & Hakim, H. H. (2020b). Metabolite profile evaluation of Indonesian roasted robusta coffees by $^{1}$H NMR technique and chemometrics. *Indonesian Journal of Chemistry, 20*(4), 850-857.

Happyana, N., Muntendam, R., & Kayser, O. (2012). Metabolomics as bioanalytical tool for characterization of medicinal plants and their phytomedical preparations. In O. Kayser & H. Warzecha (Eds.), *Pharmaceutical Biology*. Weinheim: Wiley-Blackwell.

ICO. (2019). Coffee development report 2019: International Coffee Organization.

Konieczka, P. P., Aliaño-González, M. J., Ferreiro-González, M., Barbero, G. F., & Palma, M. (2020). Characterization of arabica and robusta coffees by ion mobility sum spectrum. *Sensors, 20*, 3123.

Ky, C. L., Louarn, J., 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 Chemistry, 75*, 223–230.

Medina, J., Rodriguez, D. C., Arana, V. A., Bernal, A., Esseiva, P., & Wist, J. (2017). Comparison of attenuated total reflectance mid-infrared, near infrared, and $^{1}$H-nuclear magnetic resonance spectroscopies for the determination of coffee's geographical origin. *International Journal of Analytical Chemistry*, 7210463.

Nuńez, N., Collado, X., Martínez, C., Saurina, J., & Nuńez, O. (2020). Authentication of the origin, variety and roasting degree of coffee samples by non-targeted HPLC-UV fingerprinting and chemometrics. Application to the tetection and quantitation of adulterated coffee samples. *Foods, 9*, 378.

Procida, G., Lagazio, C., Cateni, F., Zacchigna, M., & Cichelli, A. (2020). Characterization of arabica and robusta volatile coffees composition by reverse carrier gas headspace gas chromatography–mass spectrometry based on a statistical approach. *Food science and biotechnology, 29*(10), 1319–1330.

Statistics, S. D. o. E. C. (2018). Indonesian coffee statistics 2018. In B.-S. Indonesia (Ed.). Jakarta: BPS-Statistics Indonesia.

Toci, A. T., Ribeiro, M. V. M., Toledo, P. R. A. B., Boralle, N., Pezza, H. R., & Pezza, L. (2017). Fingerprint and authenticity roasted coffees by 1H-NMR: the Brazilian coffee case. *Food Science and Biotechnology, 27*(1), 19-26.

Wahyudi, T., & Jati, W. (2012). *Challenges of Sustainable Coffee Certification in Indonesia*. Paper presented at the Seminar on the Economic, Social and Environmental Impact of Sertification on the Coffee Supply Chain, Internationale Coffee Council C109$^{th}$ London.

Wei, F., Furihata, K., Hu, F., Miyakawa, T., & Tanokura, M. (2011). Two-dimensional $^{1}$H-$^{13}$C nuclear magnetic resonance (NMR)-based comprehensive analysis of roasted coffee bean extract. *Journal of Agriculture and Food Chemistry, 59*, 9065–9073.

Wei, F., Furihata, K., Koda, M., Hu, F., Miyakawa, T., & Tanokura, M. (2012). Roasting process of coffee beans as studied by nuclear magnetic resonance: time course of changes in composition. *Journal of Agriculture and Food Chemistry, 60*, 1005−1012.

Wei, F., Furihata, K., Miyakawa, T., & Tanokura, M. (2014). A pilot study of NMR-based sensory prediction of roasted coffee bean extracts. *Food Chemistry, 152*, 363–369.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2*(1-3), 37-52.