Journal of Bioscience and Bioengineering
VOL. 120 No. 5, 555–561, 2015

ELSEVIER

www.elsevier.com/locate/jbiosc

Application of gas chromatography/flame ionization detector-based metabolite fingerprinting for authentication of Asian palm civet coffee (Kopi Luwak)

Udi Jumhawan,[1] Sastia Prama Putri,[1] Yusianto,[2] Takeshi Bamba,[1] and Eiichiro Fukusaki[1,*]

Department of Biotechnology, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan[1] *and Indonesian Coffee and Cocoa Research Institute, Jl. PB. Sudirman 90, Jember, East Java 68118, Indonesia*[2]

Received 17 January 2015; accepted 3 March 2015
Available online 23 April 2015

Development of authenticity screening for Asian palm civet coffee, the world-renowned priciest coffee, was previously reported using metabolite profiling through gas chromatography/mass spectrometry (GC/MS). However, a major drawback of this approach is the high cost of the instrument and maintenance. Therefore, an alternative method is needed for quality and authenticity evaluation of civet coffee. A rapid, reliable and cost-effective analysis employing a universal detector, GC coupled with flame ionization detector (FID), and metabolite fingerprinting has been established for discrimination analysis of 37 commercial and non-commercial coffee beans extracts. gas chromatography/flame ionization detector (GC/FID) provided higher sensitivity over a similar range of detected compounds than GC/MS. In combination with multivariate analysis, GC/FID could successfully reproduce quality prediction from GC/MS for differentiation of commercial civet coffee, regular coffee and coffee blend with 50 wt % civet coffee content without prior metabolite details. Our study demonstrated that GC/FID-based metabolite fingerprinting can be effectively actualized as an alternative method for coffee authenticity screening in industries.

[**Key words:** Civet coffee; Kopi Luwak; Gas chromatography/flame ionization detector; Metabolite fingerprinting; Industrial routine]

Coffee is one of the most traded agriculture products worldwide. Up to 475 million bags of coffee have been exported during November 2014 (International Coffee Organization, http://www.ico.org/prices/m3.htm). Among many varieties that are grown across the globe, a few, such as *Jamaican Blue Mountain* and *Hawaiian Kona*, have been recognized for its remarkable flavor thus, they are branded as premium or specialty coffee (1). Of these premium coffees, Kopi Luwak (Asian palm civet coffee) is considered to be the most expensive. Kopi Luwak, the Indonesian phrase for civet coffee, is a kind of coffee that had been eaten by the Asian palm civet (*Paradoxurus hermaphroditus*). The actions of civet's gastrointestinal tract and digestive enzymes modify the chemical composition of these coffees, thus yielding a unique flavor (1). Thus, its exotic process and short supply are the principal reasons for its high price, which is approximately 2–3 times higher than other premium coffees (2). Consequently, a protocol for authentication involving detection of potential frauds, such as illegal blending of cheaper coffee into civet coffee, is indispensable in the future.

Metabolomics, the comprehensive study of metabolome, provides a snapshot of dynamics in metabolic pathways. Unlike the genome and transcriptome, metabolome has been considered as the best descriptor of an organism's phenotype (3,4). Dramatic changes within a cell due to various perturbations are presumed to be reflected in the metabolite profiles. Many analytical systems with various degrees of sensitivity and specificity have been developed and widely applied to metabolomics. Particularly for coffee metabolomics, several studies were reported utilizing mass spectrometry (MS) (5–8), inductively coupled plasma atomic emission spectrometry (ICP-AES) (9), Fourier transform infrared (FTIR) spectroscopy (10), Raman spectroscopy (11,12), and nuclear magnetic resonance (NMR) (13,14). MS has gained popularity in the past decade because of its superior sensitivity, thus it is extensively used for metabolomics studies. Coupling MS with available separation techniques such as gas chromatography (GC) and liquid chromatography (LC) facilitates the selectivity of a wide range of compounds (15–18). Previously, we reported a metabolite profiling strategy employing GC/MS to pinpoint potential discrimination marker candidates for civet coffee and regular coffee (2). A combination of GC and MS provides straightforward analyses within a single-step extraction with good reproducibility. However, the cost for maintenance and the machine itself is relatively expensive. Therefore, it is necessary to develop a protocol that is robust, sensitive and cost-effective for quality and authenticity evaluation of civet coffee.

Gas chromatography with a universal detector like flame ionization detector (FID), has established itself as an inexpensive analytical system which can cover high-throughput analysis of carbon containing compounds (19). A few applications of gas chromatography/flame ionization detector (GC/FID) in metabolomics studies were reported for quality control of herbal plants (20,21), green tea (19), and pine wood tree (22). The availability of GC/FID in most small coffee factories and research institutes favors its application for routine analysis to evaluate authenticity of civet coffee. In this study, we developed a robust authentication technique for civet coffee using GC/FID-based metabolite fingerprinting.

* Corresponding author. Tel./fax: +81 6 6879 7424.
E-mail address: fukusaki@bio.eng.osaka-u.ac.jp (E. Fukusaki).

1389-1723/$ – see front matter
http://dx.doi.org/10.1016/j.jbiosc.2015.03.005

Metabolite fingerprinting presents rapid sample classification according to biological background rather than focusing on the small set of individual compounds thus it is a suitable approach for large size sample screening. At first, GC/FID was employed for fingerprinting of metabolites extracted from the first set of coffee samples collected from different production areas. A discrimination model of the coffee's metabolite profiles, constructed by orthogonal projection to latent structures-discriminant analysis (OPLS-DA), was then compared to the previously reported GC/MS data to verify repeatability of the established protocol. Lastly, a second set of coffee samples composed of commercial coffees and coffee blends, were analyzed to confirm the validity of the method.

MATERIALS AND METHODS

Coffee beans Samples were divided into experimental and validation coffee sets. The first set included twenty coffee beans that were collected from several cultivation areas in Indonesia. It consisted of civet coffee (no. 1–6, Arabica) that had been digested by civet, and undigested beans referred to as regular coffees (no. 7–20, Arabica and Robusta). All coffee samples were treated identically for post harvesting. Coffee samples were roasted in Probat-Werke von Gimborn Maschinenfabrik GmbH model BRZ 2 (Probat, Rhein, Germany) at 205°C for 10 min and followed by immediate air-cooling for 5 min. Roasted coffee beans were kept in sealed Falcon tubes at −30°C until use.

The second set of coffee samples included 3 civet coffees and 3 regular coffees which were bought commercially and 2 additional authentic civet coffees from the Indonesian Coffee and Cocoa Research Institute. In addition, each civet coffee and regular coffee was mixed in equal proportions (50:50, wt %) to obtain representative coffee blends. A total of 17 coffee samples, 8 pure and 9 coffee blends, were then analyzed by GC/MS to verify the established protocol for coffee authentication. All coffee samples were measured in triplicates. The sample descriptions are shown in Tables S1 and S2.

Chemicals All chemicals used in this study were analytical grade. Methanol, chloroform and distilled water used for the preparation of extraction solvent and ribitol used as internal standard were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Pyridine (Wako), methoxyamine hydrochloride (Sigma Aldrich, Milwaukee, WI, USA) and *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) (GL Science Inc., Tokyo, Japan) were applied for derivatization. Chemical standards for confirmation of compound identification, their providers, and purities were as follows: citric acid (Nacalai-Tesque, Kyoto, Japan, 99.5%), malic acid (Nacalai-Tesque, 99%), pyroglutamic acid (ICN Biomedicals, OH, USA, 99.5%), caffeine (Sigma Aldrich, 98.5%), inositol (Wako, 99%), quinic acid (Sigma Aldrich, 98%) and glycolic acid (Sigma Aldrich, 99%).

Extraction and derivatization To produce fine powder for extraction, coffee beans were ground with a Retsch ball mill (20 Hz, 3 min). Fifteen mg of coffee powder was extracted with 1 mL MeOH/$CHCl_3$/H_2O (5/2/2) and added with 60 μL of ribitol (0.2 mg/mL) as internal standard. The samples were centrifuged at 16,000 g for 3 min at 4°C. Nine hundred microliter of the supernatant was then transferred into 1.5 mL Eppendorf tube and added with 400 μL Milli-Q water. After re-centrifugation, 400 μL aqueous layer was transferred into a new tube with a pierced cap. The extract was evaporated by vacuum centrifugation for 2 h and freeze-drying overnight. The dried extract was mixed with 100 μL of methoxyamine hydrochloride (20 mg/mL in pyridine) and subsequently incubated at 30°C for 90 min. The second agent, 50 μL MSTFA was added to the mixture and re-incubated at 37°C for 90 min.

GC/FID and GC/MS analysis The typical workflow for GC/MS analysis of aqueous coffee bean extract has been described in detail elsewhere (2). GC/FID was conducted on a GC-2010 (Shimadzu, Kyoto, Japan) installed with an AOC-20s autosampler and AOC-20i autoinjector. One microliter of each derivatized sample was injected in split mode, 25/1 (v/v). To establish proper comparison and validation with the reported GC/MS data, the same type of column, CP-SIL 8 CB low bleed column (0.25 mm × 30 m, 0.25 μm, Varian Inc., Palo Alto, CA, USA), and identical temperature program were applied to GC/FID analysis. The carrier gas (He) was maintained at a constant velocity of 45 cm/s. The injector and FID temperature were set at 230 and 320°C, respectively.

Data preprocessing and multivariate analysis Raw chromatographic data of GC/FID were converted into CDF format using GCMS Solution software package (Shimadzu, Kyoto, Japan). The converted files were subjected to baseline correction, normalization and alignment of retention times using the in-house software, PiroTran ver 1.41 (GL Sciences, Tokyo, Japan). The retention time of internal standard ribitol was confirmed with co-injection of authentic chemical standard before being utilized as reference for normalization and retention time alignment. To reduce the run-to-run variation, the threshold for peak intensity (RSD) was set to <20%, in each measurement replicate. To construct the data matrix, in which each row and column represent the samples and relative peak intensity at certain retention time, respectively, the outcome data were imported into Pirouette ver 4.0 (Infometrix, Inc., Woodinville, WA, USA). The data matrix was then subjected to multivariate analysis.

Multivariate analysis was carried out using SIMCA-P+ ver. 13 (Umetrics, Umeå, Sweden) to reduce dimensionality of the huge MS data and extract biological interpretation. In this study, PCA and OPLS-DA were employed. OPLS-DA was used to decipher the relationships between two data matrices, *X* (predicted variables), and *Y* (observed variables) (23). Here, the chromatographic GC/FID data were used as *X* and for *Y*, the binary vector of 0 and 1 was assigned for civet coffee and regular coffee, respectively. The data were Pareto scaled prior to analysis without any transformation.

RESULTS AND DISCUSSION

Chromatographic data of GC/FID Representative GC/FID and GC/MS chromatograms of aqueous coffee extracts are shown in Fig. S1. We compared the chromatogram obtained from GC/FID with the one from GC/MS analysis reported previously (2) using the same coffee extract and column type. The chromatographic data of GC/FID and GC/MS gave similar metabolite patterns, which contained the peaks from diverse metabolites, i.e., glycolic acid (peak no. 1), malic acid (peak no. 2), pyroglutamic acid (peak no. 3), citric acid, (peak no. 4) quinic acid (peak no. 5), inositol (peak no. 6), sucrose (peak no. 7) and chlorogenic acid (peak no. 8). A total of 678 peaks were obtained from GC/FID, compared to 182 peaks from GC/MS analysis.

For metabolite fingerprinting, it is not necessary to determine the individual information of every peak (24). Nonetheless to confirm the overall quality of GC/FID analysis, peak confirmation of the GC/FID chromatogram was performed by comparing to the identified peaks in the GC/MS data and co-injection of authentic chemical standards. Whilst most of the detected peaks that represented key coffee metabolites were identical between GC/FID and GC/MS, we also observed a shift in their retention times, such as in glycolic acid (5.02 and 4.96 min), malic acid (9.11 and 9.05 min), and citric acid (11.68 and 11.61 min), respectively. Since metabolomics data are often subject to unwanted variations (25), the retention time shift reported here, albeit not severe, may be due to experimental variation between analytical instruments.

As shown in Fig. S1, the overall chromatographic profiles between GC/FID and GC/MS were similar. However, it is noticeable that GC/FID analysis provided higher relative peak intensity than GC/MS for almost all detected peaks. The higher relative peak intensity often implies higher sensitivity as GC/FID analysis has been described to generate higher sensitivity compared to the mass detector which frequently operated in a full-scan mode for gathering entire profiles of biological samples (19,20,24). Measurement of total ions over mass range resulted in the limitation of sensitivity for the mass detector. The efficient reduction of relative intensity for detected peaks within the range of 4.2 and 6 min was also observable for GC/FID analysis. The peaks were confirmed by comparison with the NIST library and identified as siloxane, common peak contaminants from injector and vial septa (Fig. S1A). The result was explicable since FID primarily responds to a wide variety of carbon-containing organic compounds whereas a mass detector relies on the recognition of the entire ionized and fragmented molecules. The results suggested the practicability of using GC/FID for metabolite fingerprinting of coffee beans as it provided higher sensitivity over a similar range of detected compounds than GC/MS analysis.

Multivariate data analysis of coffee bean extracts by GC/FID-based metabolite fingerprinting Metabolite fingerprinting of coffee extracts for the development of rapid assessment method was done using GC coupled with a universal detector, FID. Metabolite identification has been reported as tedious work and a major challenge in the metabolomics workflow (24,26–28). However, determination of the individual level and identity for each metabolite is not a key requirement for metabolite fingerprinting

strategy therefore, metabolite fingerprinting is suitable for quality screening of large number of coffee samples.

We carried out a comparison of the multivariate analyses obtained from GC/MS analysis with that of GC/FID in order to evaluate the performance quality of the latter platform. The quality of PCA and OPLS-DA models constructed from GC/MS data has been published previously (2). Principal component analysis was performed as unsupervised and non-biased method to reduce the dimensionality of multivariate data and visualize the differences in the metabolite fingerprints of coffee extracts. The score plot of PCA derived from GC/FID analysis, where the first two components (PC1 and PC2) accounted for 39.2% and 16.9% of the total variance, revealed distinct separation between samples due to their genetic trait, Arabica and Robusta (Fig. S2A). Arabica and Robusta were analyzed in this study to represent the two major coffee species traded annually (11). The Arabica coffee data set, comprised of both civet and regular coffee, clearly separated from Robusta coffee. Results indicate that genetic variability between species gave greater impact for data separation in comparison to perturbation during animal digestion and cultivation area (along PC1-axis). The PCA obtained from GC/FID had better quality in terms of goodness-of-fit ($R^2X = 0.88$) compared to the GC/MS data ($R^2X = 0.84$) (Fig. S2).

We employed OPLS-DA modeling to indicate the role of animal perturbation in coffee sample separation by setting a binary vector with the value 0 for the civet coffee class and 1 for the regular coffee class. A discrimination model was then built from the total data set, in which the chromatographic data were used as predicted variable (X) and the binary vector as the observed variable (Y). The OPLS-DA score plot of the total data set of GC/FID indicated the apparent clustering between samples on the basis of animal perturbation (Fig. 1A). Civet coffee and regular coffee samples were clearly separated in the predictive component ($t[1]$) (Fig. 1A–B). Separation of Arabica and Robusta coffee remained observable in the right cluster occupied by regular coffee samples in the orthogonal component ($to[1]$). In GC/FID, the OPLS-DA model was built with an R^2Y value of 0.996 and a Q^2 value of 0.78 (Fig. 1A). The correlation coefficient (R^2Y), describes how a model fit a set of predicted data set related to class separation. The model derived from GC/FID data is less robust in terms of model predictability (Q^2) compared to the model from GC/MS data ($R^2Y = 0.965$ and $Q^2 = 0.892$, Fig. 1B). The lower Q^2 value in GC/FID may signify an overfitted model which resulted from the use of all confirmed peaks from data processing to construct the discrimination model in GC/FID. The use of irrelevant components increases the risk of overfitting and eventually create poor predictive precision (29–31). However, the Q^2 value ≥ 0.5 was widely considered as good and acceptable for a model derived from biological samples (20,32,33).

Permutation tests were performed in the PLS-DA model to confirm the quality of OPLS-DA model created from GC/FID. According to Setoyama *et al*, if the OPLS-DA model is overfitted, the R^2Y and Q^2 values would not virtually change after permutation (5). Both parameters were in the range of the requirement for a reliable model; R^2Y-intercepts fluctuated between 0.3–0.4 and Q^2-intercept was below 0.05, respectively (Fig. S3) (34,35). These denoted that there was a change in the values of the two parameters.

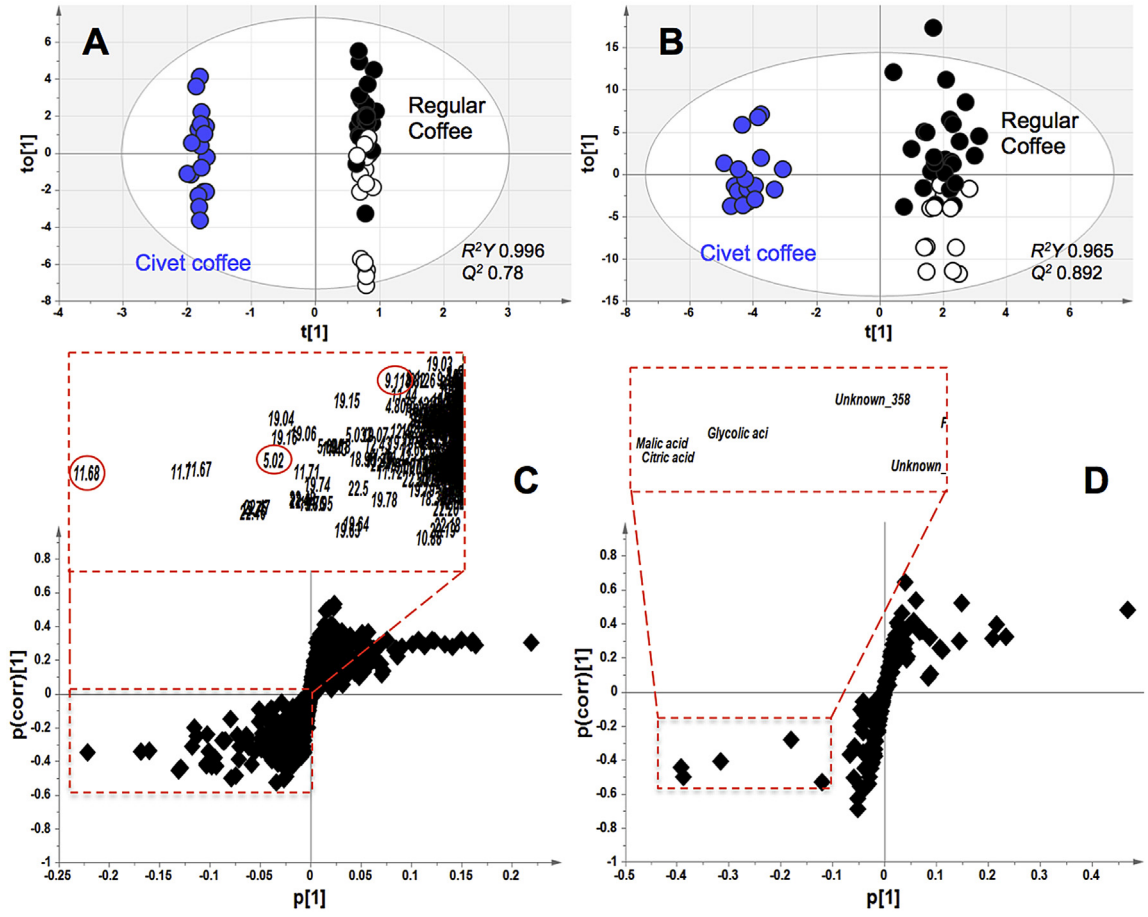

FIG. 1. OPLS-DA score plots and S-plots based on (A, C) GC/FID and (B, D) GC/MS chromatograms of 20 coffee bean extracts. The blue-filled circles, open and closed circles represent civet coffee, regular coffee (Robusta) and regular coffee (Arabica), respectively. The S-plot displayed the covariance p against correlation p(corr) of the variables to the model class designation. The closed diamonds represent each variable (detected peak) used for model construction; identities of variables with high reliability to civet coffee are given in the inset figure.

Furthermore, external validation was then performed to estimate the accuracy of the performance of the discrimination model in practice. We prepared and analyzed the GC/FID data of three commercial samples of civet coffee and three commercial samples of regular coffee in separate days. The samples were then projected onto the constructed OPLS-DA model as external validation. Fig. 2 showed that commercial civet coffee and commercial regular coffee fit into the discriminant model; if those samples were classified into the same class with prior civet and regular coffee data ($R^2 = 0.982$ and $Q^2 = 0.741$). Additionally, the six samples from each class were left randomly (test set) and the OPLS-DA was performed three times for the remaining samples only (training set). The R^2Y and Q^2 values after cross validation were simultaneously calculated and the values obtained were in the range of 0.907–0.957 and 0.616–0.667, respectively. Although a drop in R^2Y and Q^2 values was seen, the quality of the model was still acceptable. To visualize how accurately external validation will perform to differentiate samples, we set the cut off of the prediction at 0.5. As a result, almost all test samples were correctly classified except for one civet coffee sample (sample ID 24_3) (Fig. 3). Taken together, we could confirm the practicability of the GC/FID coupled to metabolite fingerprinting strategy for rapid discrimination and prediction of new samples with statistical significance. The results also suggested that the combined techniques could effectively minimize variability, *i.e.* error from day-to-day measurements.

Statistically significant variables contributing to the differentiation of civet coffee and regular coffee were selected from the S-plot of the OPLS-DA. S-plot combines both covariance (contribution or magnitude) and correlation (reliability between the variables (metabolites) with the model class designation) (23). Consequently, on the basis of their contribution and reliability, the variables that changed significantly are plotted at the top and bottom of the S-plot, and those that do not significantly contribute are plotted in the middle. Three variables were highlighted in the S-plot of GC/FID, variable 5.02_, 9.11_, and 11.68_ (labeled after their retention time). To confirm the reproducibility of the metabolite fingerprinting strategy, selection of variables from GC/FID was done for those with identical retention time to that of the biomarker from the reported GC/MS analysis and later confirmed with co-injection of the chemical standard for giving a valid identification (Fig. 1C, D). In GC/FID, the S-plot shows that variable 11.68_ had the highest contribution p, which means a high correlation p(corr), for class separation, followed by variable 5.02_ (Fig. 1C). Variable 9.11_ was plotted in the middle region with ambiguous significance level. The contribution of those variables in the model projection could also be explained using variable important in the projection (VIP). The average of 95% confidence interval VIP is equal to 1.0 (19); therefore large VIP values (>1) are often considered relevant for explaining the OPLS-DA model. Table 1 displays the comparison between the three variables extracted from the GC/FID data with corresponding biomarkers from GC/MS analysis. The VIP score of each variable represents high contribution to the model. Lower VIP score for variable 9.11_ corroborated with the S-plot result. It is implied that passage through civet's digestive tract may enhance the level of particular organic acids (1,2). However, to understand the underlying biological meaning of these biomarkers as a result of animal digestion and its correlation to the sensory profile, further investigation is needed.

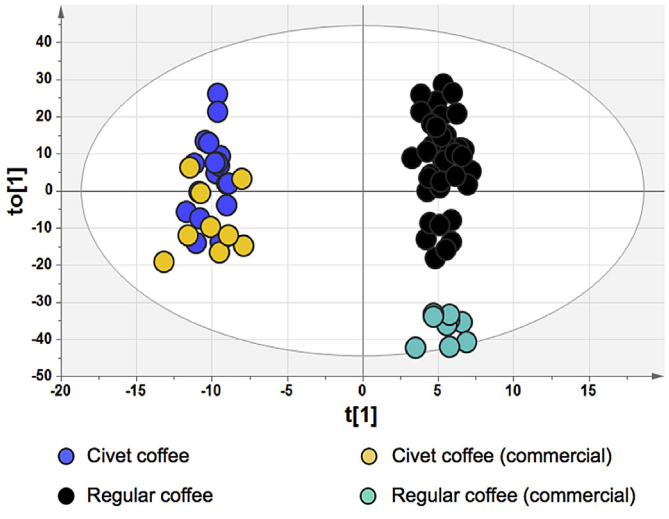

FIG. 2. OPLS-DA score plot of validation test. We projected commercial civet coffee (orange filled-circles) and commercial regular coffee (light blue-filled circles) samples into the same class with civet and regular coffee samples from experimental set, respectively.

Validation of metabolite fingerprinting strategy to confirm authenticity of commercial samples A set of commercial samples from the coffee market has been analyzed to provide scientific evidence of the GC/FID application in the coffee industry. Since processing commercial samples is based on the customers' preference, the roasting temperature may vary from experimental coffee. Commercially available regular and civet coffee were selected from different production areas. To set the validation threshold, we acquired two authentic civet coffees from different production years as benchmark samples. Furthermore, a total of 9 coffee blends were prepared from the combination of each commercial sample with mixing ratio of 50:50 (wt %). These four differentiation parameters, occurrence of perturbation, production area, roasting parameter and mixing ratio, would present comprehensive coverage for validation.

PCA modeling was applied to the validation coffee data set. As shown in Fig. 4, commercial coffee samples were distinguished from one another on account of perturbation and mixing ratio by 47.5% variance in PC1 ($R^2X = 0.838$). Both civet coffees, commercial and authentic, were populated in a wide margin at the right region of the intercept (0,0). Four civet coffee samples (2 commercial and 2 authentic coffee) from neighboring production areas were clustered within close range in PC2. It is suggested that the production area may play a significant role for separation of civet coffee. One commercial civet coffee (no. 5) was located outside the 95% confidence interval of the Hotelling's T^2, thus marking the possibility of an outlier.

Coffee blends were populated within the area around the intercept. This is probably linked to the equal mixing ratio between civet and regular coffee. Interestingly, regular coffee samples were populated in a close area, only partially separated owing to production area (foreign and local). According to the loading scatter plot (data not shown), some significant variables, including variables 11.68_, 9.11_ and 5.02_, were captured and highlighted as responsible for the differentiation of commercial samples in the PCA score plot.

Fig. 5 summarizes the box plot construction to visualize performance of significant variables captured by PCA and OPLS-DA for sample classification. Relative peak intensities of significant variables were used to create a whisker-box plot. Generally, the box plot was divided into three groups on the account of relative intensity level; civet coffee, coffee blends, regular coffee. An overlap box plot was spotted only between civet coffee and coffee blend clusters due to equal mixing ratio, but not for civet coffee and regular coffee. The box plot clustering was clearly observed for variables 9.11_ and 11.68_, corresponding to malic and citric acid, respectively, in GC/MS analysis. Although represented with high VIP and correlation, such result could not be reproduced for variable 5.02_ (Fig. S4). Variable 5.02_ corresponds to glycolytic acid in

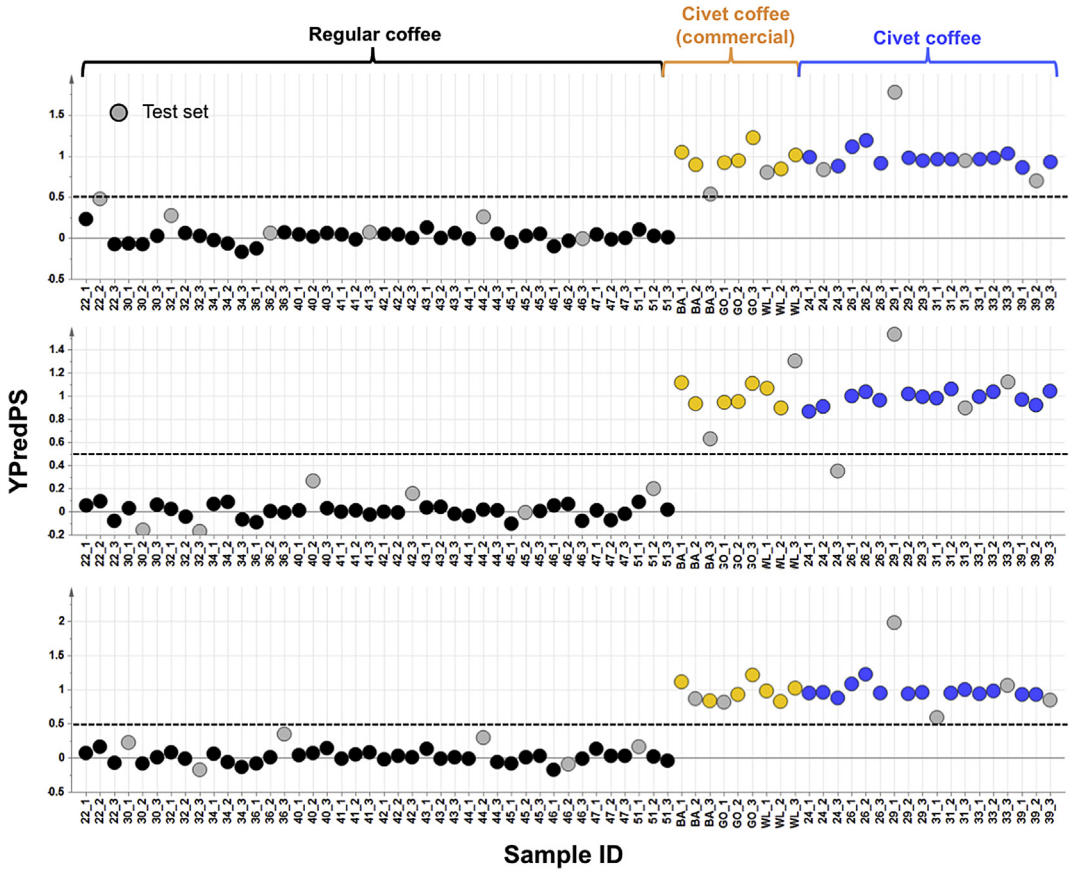

FIG. 3. Validation test of OPLS-DA model derived from GC/FID. The six samples from each class were left randomly (test set, gray filled-circles) and the OPLS-DA was performed three times for only remaining samples (training set, colored filled-circles). The cut off of the prediction (dashed line) was set at 0.5.

GC/MS analysis, a biomarker candidate that has been shown to have poor predictive performance for authentication (2). The box plot pattern of variable 5.02_ conflicted with the other two variables, indicating a dependency to regular coffee. As shown in Fig. 5A and B, the commercial civet coffee no. 5 gave very low relative intensities for variable 9.11_ and 11.68_, therefore suffered from overlap with the other clusters. This outlier can also be found in PCA score plot in Fig. 4, suggesting a possible counterfeit or a difference in coffee species used for analysis, *i.e.* Robusta, since the rest of the civet coffee samples were Arabica coffee and display relevant margin in their relative intensities (regardless of roasting temperature and production area). If sample no. 5 is indeed civet coffee (Arabica), its relative intensity for those significant variables should be much higher. These results verified the feasibility of employing the significant variables obtained from GC/FID for practical use (for authentication). Collectively, our results demonstrate in principle that GC/FID coupled with metabolite fingerprinting is a good complementary and cost effective analysis platform for quality assessment of civet coffee.

TABLE 1. Comparison of significant variables extracted from GC/FID and GC/MS.

No.	GC/MS			GC/FID		
	Variable ID[a]	RT (min)	VIP[c]	Variable ID[b]	RT (min)	VIP[c]
1	Glycolic acid	4.96	3.93	5.02_	5.02	4.44
2	Malic acid	9.05	5.53	9.11_	9.11	1.48
3	Citric acid	11.61	5.6	11.68_	11.68	8.39

[a] Confirmed with co-injection with chemical standard and comparison of mass fragment spectra with NIST database.
[b] Selected on the account of RT comparison with GC/MS data and co-injection of chemical standard.
[c] Variable Importance in the Projection, extracted from S-plot of OPLS-DA.

Previously, we have demonstrated a metabolite profiling strategy for civet coffee quality assessment through GC/MS. In this method, automated procedures for tentative identification of unknown peaks by matching the information derived from mass spectral fragmentation patterns with chemical databases are necessary (2). However, the unambiguous and exhaustive metabolite identification in a biological system has been well documented as challenging (24,26–28). In contrast, an alternative strategy using GC/FID analysis paired with metabolite fingerprinting provided rapid classification of coffee samples without prior metabolite details. The significant compounds contributing to civet coffee quality

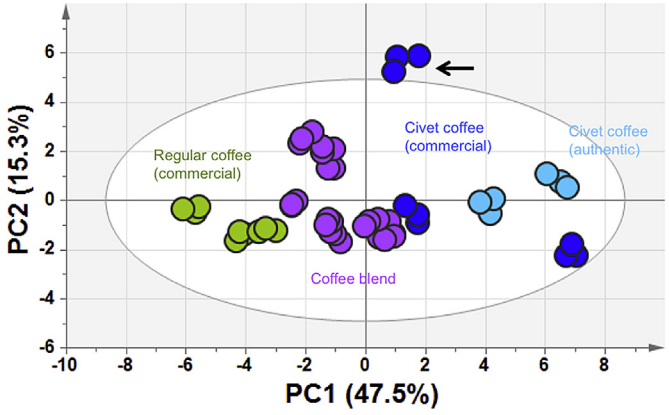

FIG. 4. PCA score plot derived from GC/FID chromatograms of 17 coffee samples in validation set (n = 3). Coffee samples were clustered according to occurrence of perturbation (both animal digestion and blending) along PC1 axis. Arrow indicates possible outlier, sample no. 5.

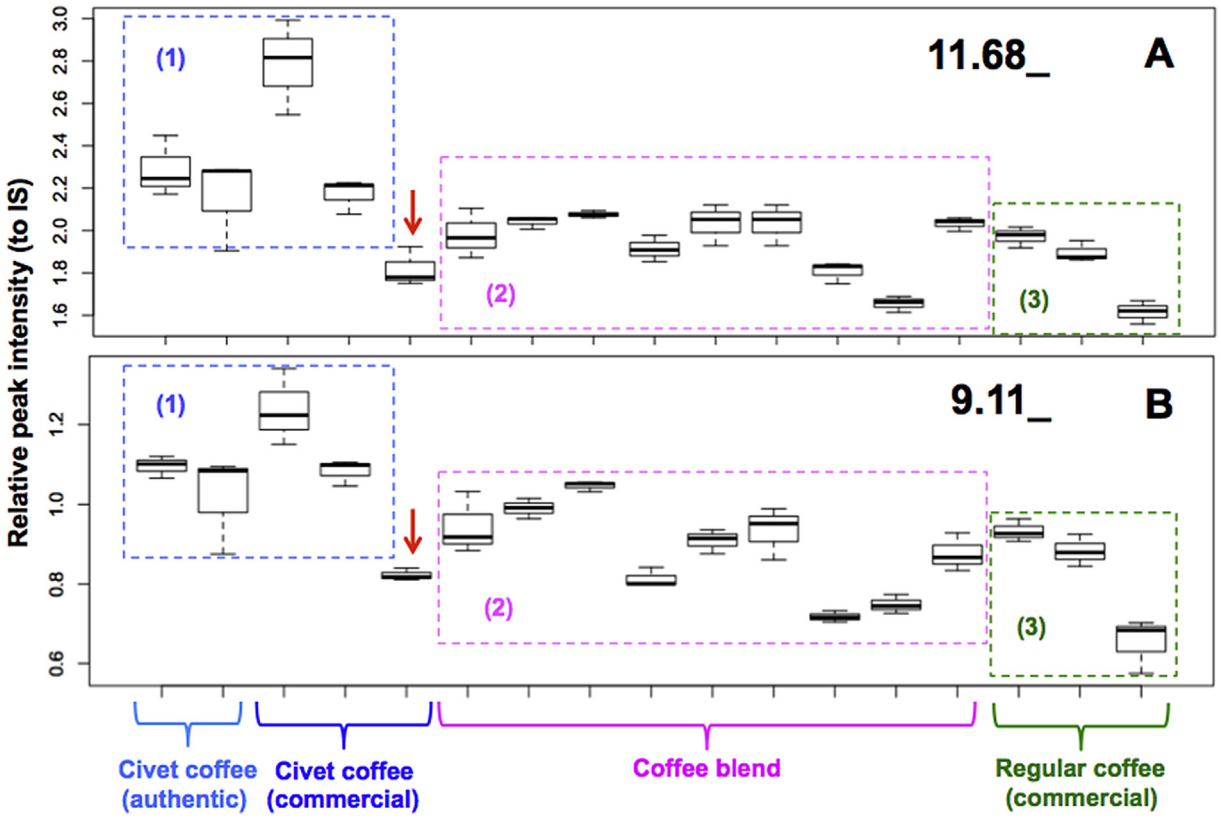

FIG. 5. Box plot of significantly different variables between samples in validation set from result of OPLS-DA, (A) 11.68_ and (B) 9.11_. Both variables successfully differentiated group of commercial samples from one another, except one commercial civet coffee, sample no. 5 (red arrow) that was spotted as potential outlier in Fig. 4.

assessment, such as malic and citric acid, showed better sensitivity in FID compared to an MS detector. By employing multivariate data analysis such as OPLS-DA, construction of a good prediction model to confirm authenticity of commercial samples was accomplished. Our study demonstrated that metabolite fingerprinting through GC/FID could effectively reproduce coffee quality prediction from the previous technique (2). With elimination of the tedious identification step, the GC/FID system offered high-speed analysis for coffee quality assessment. This advantage can be beneficial to manufacturers for quality control, especially for authentication of commercial coffee and other agricultural products in industrial scale.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jbiosc.2015.03.005

ACKNOWLEDGMENTS

The authors thank Dr. Teguh Wahyudi, Dr. Surip Mawardi and Dr. Misnawi from Indonesian Coffee and Cocoa Research Institute, for the samples and discussions, Dr. Yasumune Nakayama and Dr. Hiroshi Tsugawa for data analysis and Dr. Walter Alvarez Lavina for manuscript improvement.

References

1. **Marcone, M. F.:** Composition and properties of Indonesian palm civet coffee (Kopi Luwak) and Ethiopian civet coffee, Food Res. Int., **37**, 901–912 (2004).
2. **Jumhawan, U., Putri, S. P., Yusianto, Marwani, E., Bamba, T., and Fukusaki, E.:** Selection of discriminant markers for authentication of asian Palm civet coffee (Kopi Luwak): a metabolomics approach, J. Agric. Food Chem., **61**, 7994–8001 (2013).
3. **Ibáñez, A. J., Fagerer, S. R., Schmidt, A. M., Urban, P. L., Jefimovs, K., Geiger, P., Dechant, R., Heinemann, M., and Zenobi, R.:** Mass spectrometry-based metabolomics of single yeast cells, Proc. Natl. Acad. Sci. USA, **110**, 8790–8794 (2013).
4. **Zenobi, R.:** Single-cell metabolomics: analytical and biological perspectives, Science, **342**, 6163 (2013).
5. **Setoyama, D., Iwasa, K., Seta, H., Shimizu, H., Fujimura, Y., Miura, D., Wariishi, H., Nagai, C., and Nakahara, K.:** High-throughput metabolic profiling of diverse green *Coffea arabica* beans identified tryptophan as a universal discrimination factor for immature beans, PLoS One, **8**, e70098 (2013).
6. **Garrett, R., Vaz, B. G., Hovell, A. M. C., Eberlin, M. N., and Rezende, C. M.:** Arabica and robusta coffees: identification of major polar compounds and quantification of blends by direct-infusion electrospray ionization – mass spectrometry, J. Agric. Food Chem., **60**, 4255–4258 (2012).
7. **Mancha Agresti, P. D. C., Franca, A. S., Oliveira, L. S., and Augusti, R.:** Discrimination between defective and non-defective Brazilian coffee beans by their volatile profile, Food Chem., **106**, 787–796 (2008).
8. **Mendonça, J. C. F., Franca, A. S., Oliveira, L. S., and Nunes, M.:** Chemical characterisation of non-defective and defective green arabica and robusta coffees by electrospray ionization-mass spectrometry (ESI-MS), Food Chem., **111**, 490–497 (2008).
9. **Anderson, K. A. and Smith, B. W.:** Chemical profiling to differentiate geographic growing origins of coffee, J. Agric. Food Chem., **50**, 2068–2075 (2002).
10. **Briandet, R., Kemsley, E. K., and Wilson, R. H.:** Discrimination of arabica and robusta in instant coffee by Fourier transform infrared spectroscopy and chemometrics, J. Agric. Food Chem., **44**, 170–174 (1996).
11. **Wermelinger, T., D'Ambrosio, L., Klopprogge, B., and Yeretzian, C.:** Quantification of the robusta fraction in a coffee blend via Raman spectroscopy: proof of principle, J. Agric. Food Chem., **59**, 9074–9079 (2011).
12. **El-Abassy, R. M., Donfack, P., and Materny, A.:** Discrimination between Arabica and Robusta green coffee using visible micro Raman spectroscopy and chemometric analysis, Food Chem., **126**, 1443–1448 (2011).
13. **Wei, F., Furihata, K., Koda, M., Hu, F., Kato, R., Miyakawa, T., and Tanokura, M.:** C NMR-based metabolomics for the classification of Green coffee beans according to variety and origin, J. Agric. Food Chem., **60**, 10118–19125 (2012).
14. **Wei, F., Furihata, K., Koda, M., Hu, F., Miyakawa, T., and Tanokura, M.:** Roasting process of coffee beans as studied by nuclear magnetic, J. Agric. Food Chem., **60**, 1005–1012 (2012).
15. **Shiga, K., Yamamoto, S., Nakajima, A., Kodama, Y., Imamura, M., Sato, T., Uchida, R., Obata, A., Bamba, T., and Fukusaki, E.:** Metabolic profiling

approach to explore compounds related to the umami intensity of soy sauce, J. Agric. Food Chem., **62**, 7317−7322 (2014).
16. **Yamamoto, S., Shiga, K., Kodama, Y., Imamura, M., Uchida, R., Obata, A., Bamba, T., and Fukusaki, E.:** Analysis of the correlation between dipeptides and taste differences among soy sauces by using metabolomics-based component profiling, J. Biosci. Bioeng., **118**, 56−63 (2014).
17. **Mimura, N., Isogai, A., Iwashita, K., Bamba, T., and Fukusaki, E.:** Gas chromatography/mass spectrometry based component profiling and quality prediction for Japanese sake, J. Biosci. Bioeng., **118**, 406−414 (2014).
18. **Taguchi, K., Fukusaki, E., and Bamba, T.:** Simultaneous analysis for water- and fat-soluble vitamins by a novel single chromatography technique unifying supercritical fluid chromatography and liquid chromatography, J. Chromatogr. A, **1362**, 270−277 (2014).
19. **Jumtee, K., Bamba, T., and Fukusaki, E.:** Fast GC-FID based metabolic fingerprinting of Japanese green tea leaf for its quality ranking prediction, J. Sep. Sci., **32**, 2296−2304 (2009).
20. **Kobayashi, S., Putri, S. P., Yamamoto, Y., Donghyo, K., Bamba, T., and Fukusaki, E.:** Gas chromatography-mass spectrometry based metabolic profiling for the identification of discrimination markers of *Angelicae Radix* and its application to gas chromatography-flame ionization detector system, J. Biosci. Bioeng., **114**, 232−236 (2012).
21. **Ma, C., Wang, H., Lu, X., Xu, G., and Liu, B.:** Metabolic fingerprinting investigation of *Artemisia annua* L. in different stages of development by gas chromatography and gas chromatography-mass spectrometry, J. Chromatogr. A, **1186**, 412−419 (2008).
22. **Yeh, T., Morris, C. R., Goldfarb, B., Chang, H., and Kadla, J. F.:** Utilization of polar metabolite profiling in the comparison of juvenile wood and compression wood in loblolly pine (*Pinus taeda*), Tree Physiol., **26**, 1497−1503 (2006).
23. **Wiklund, S., Johansson, E., Sjostrom, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., Gottfries, J., Moritz, T., and Trygg, J.:** Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models, Anal. Chem., **80**, 115−122 (2008).
24. **Fiehn, O.:** Metabolomics−the link between genotypes and phenotypes, Plant Mol. Biol., **48**, 155−171 (2002).
25. **De Livera, A. M., Dias, D. A., De Souzza, D., Rupasinghe, T., Pyke, J., Tull, D., Roessner, U., McConville, M., and Speed, T. P.:** Normalizing and integrating metabolomics data, Anal. Chem., **84**, 10768−10776 (2012).
26. **Dieterle, F., Ross, A., Schlotterbeck, G., and Senn, H.:** Metabolite projection analysis for fast identification of metabolites in metabonomics. Application in an amiodarone study, Anal. Chem., **78**, 3551−3561 (2006).
27. **Moco, S., Vervoort, J., Bino, R. J., De Vos, R. C. H., and Bino, R.:** Metabolomics technologies and metabolite identification, TrAC Trends Anal. Chem., **26**, 855−866 (2007).
28. **Li, S., Park, Y., Duraisingham, S., Strobel, F. H., Khan, N., Soltow, Q. A., Jones, D. P., and Pulendran, B.:** Predicting network activity from high throughput metabolomics, PLoS Comput. Biol., **9**, e1003123 (2013).
29. **Wang, Z., Zhang, M., and Harrington, P. D. B.:** Comparison of three algorithms for the baseline correction of hyphenated data objects, Anal. Chem., **86**, 9050−9057 (2014).
30. **Eriksson, L., Trygg, J., and Wold, S.:** A chemometrics toolbox based on projections and latent variables, J. Chemom., **28**, 332−346 (2014).
31. **Hawkins, D. M.:** The problem of overfitting, J. Chem. Inf. Comput. Sci., **44**, 1−12 (2004).
32. **Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nystrom, A., Pettersen, J., and Bergman, R.:** Experimental design and optimization, Chemom. Intell. Lab. Syst., **42**, 3−40 (1998).
33. **Kobayashi, S., Nagasawa, S., Yamamoto, Y., Donghyo, K., Bamba, T., and Fukusaki, E.:** Metabolic profiling and identification of the genetic varieties and agricultural origin of *Cnidium officinale* and *Ligusticum chuanxiong*, J. Biosci. Bioeng., **114**, 86−91 (2012).
34. **Eriksson, L., Jaworska, J., Worth, A. P., Cronin, M. T. D., McDowell, R. M., and Gramatica, P.:** Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs, Environ. Health Perspect., **111**, 1361−1375 (2003).
35. **Campo, E., Ferreira, V., Escudero, A., Marqués, J. C., and Cacho, J.:** Quantitative gas chromatography−olfactometry and chemical quantitative study of the aroma of four Madeira wines, Anal. Chim. Acta, **563**, 180−187 (2006).