JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Article

pubs.acs.org/JAFC

Selection of Discriminant Markers for Authentication of Asian Palm Civet Coffee (Kopi Luwak): A Metabolomics Approach

Udi Jumhawan,[†] Sastia Prama Putri,[†] Yusianto,[§] Erly Marwani,[#] Takeshi Bamba,[†] and Eiichiro Fukusaki*[,†]

[†]Department of Biotechnology, Graduate School of Engineering, Osaka University, Osaka, Japan

[§]Indonesian Coffee and Cocoa Research Institute, East Java, Indonesia

[#]Bandung Institute of Technology, West Java, Indonesia

Supporting Information

ABSTRACT: Kopi Luwak, an exotic Indonesian coffee, is made from coffee berries that have been eaten by the Asian palm civet (*Paradoxurus hermaphroditus*). Despite being known as the world's most expensive coffee, there is no reliable, standardized method for determining its authenticity. GC-MS-based multimarker profiling was employed to explore significant metabolites as discriminant markers for authentication. Extracts of 21 coffee beans (*Coffea arabica* and *Coffea canephora*) from three cultivation areas were analyzed and subjected to multivariate analyses, principal component analysis, and orthogonal projection to latent structures discriminant analysis. Citric acid, malic acid, and the inositol/pyroglutamic acid ratio were selected for further verification by evaluating their differentiating abilities against various commercial coffee products. The markers demonstrated potential application in the differentiation of original, fake Kopi Luwak, regular coffee, and coffee blend samples with 50 wt % Kopi Luwak content. This is the first report to address the selection and successful validation of discriminant markers for the authentication of Kopi Luwak.

KEYWORDS: *Kopi Luwak, Asian palm civet, GC-MS, authentication, discriminant markers*

INTRODUCTION

Coffee is one of the most popular beverages in the world. Kopi Luwak is considered to be the world's most expensive coffee, with a price tag of U.S. $150−227/pound.[1] Kopi Luwak, the Indonesian words for coffee and civet, respectively, is made from coffee berries that have been eaten by the Asian palm civet (*Paradoxurus hermaphroditus*), a small mammal native to southern and northern Asia.[2] The civet climbs coffee trees and instinctively selects coffee cherries. During digestion, the coffee pericarp is completely digested and the beans are excreted. The intact beans are then collected, cleaned, wet-fermented, sun-dried, and further processed by roasting. Kopi Luwak's high selling price is mainly attributed to its exotic and unexpected production process.[3]

Despite its profitable prospects, there is no reliable, standardized method for determining the authenticity of Kopi Luwak. Moreover, there is limited scientific information on this exotic coffee. Recently, coffee adulterated to resemble Kopi Luwak was reported in the coffee market.[4] This poses serious concern among consumers over the authenticity and quality of the products currently available in the market. Discrimination between Kopi Luwak and regular coffee has been achieved using electronic nose data.[3] However, the selection of a discriminant marker for authentication was not addressed. The methods currently employed by Kopi Luwak producers is sensory analysis including visual and olfactory testing, both of which are inadequate. For example, visual examination is only possible for green coffee beans prior to roasting, and very few trained experts can perform the highly subjective sensory analysis to discriminate Kopi Luwak.

Information flows in metabolic pathways are highly dynamic and represent the current biological states of individual cells. Hence, the metabolome has been considered as the best descriptor of physiological phenomena.[5] Metabolomics techniques can be powerful tools to elucidate variations in phenotypes imposed by perturbations such as gene modification, environmental factors, or physical stress. The "black box" process during animal digestion can be translated as physical and enzymatic consequences to the coffee bean, which presents a smoother surface and color changes after digestion.[3] Thus, a metabolomics technique was chosen to screen and select discriminant markers for the authenticity assessment of Kopi Luwak. Metabolomics techniques have been effectively applied to distinguish the phytochemical compositions of agricultural products of different origins,[6] varieties,[6,7] and cultivars[8] for quality control and breeding.

In our study, gas chromatography coupled with quadrupole mass spectrometry (GC-Q/MS)-based multimarker profiling was employed to identify discriminant markers for the differentiation of Kopi Luwak and regular coffees. A combination of gas chromatography and mass spectrometry (GC-MS) provides high sensitivity, reproducibility, and quantitation of a large number of metabolites with a single-step extraction.[9,10] Sample classification by means of chemometrics was performed using principal component analysis (PCA). Subsequently, orthogonal projection to latent struc-

Received: January 1, 2013
Revised: July 25, 2013
Accepted: July 27, 2013
Published: July 27, 2013

tures combined with discriminant analysis (OPLS-DA)[11] and significance analysis of microarrays/metabolites (SAM)[12] were used to isolate statistically significant compounds as discriminant marker candidates. The applicability of these candidates as discriminant markers was verified to differentiate various commercial coffee products.

MATERIALS AND METHODS

Materials. Samples were divided into experimental and validation coffee groups. Experimental coffees were utilized to construct the discrimination model and to select significant compounds. Verification of the applicability of the discriminant markers was carried out using the validation coffee set. In this paper, coffee that had been digested by the animal is referred to as Kopi Luwak, and the other beans are referred to as regular coffee. Kopi Luwak and regular coffee samples of two species, *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta), were used. Coffee samples were obtained from 21 sampling points in three cultivation areas in Indonesia (Java, Sumatra, and Bali). Samples were roasted in a Probat-Werke von Gimborn Maschinenfabrik GmbH model BRZ 2 (Probat, Rhein, Germany) at 205 °C for 10 min to obtain a medium degree of roasting and then were air-cooled for 5 min. Coffee beans were ground and stored in sealed 50 mL BD Falcon tubes at −30 °C with light-shielding prior to analysis. The sample descriptions are shown in Tables S1 and S2 (Supporting Information). Wet- fermentation was applied to both the Kopi Luwak and regular coffees. For regular coffees, conventional protocols were applied after harvesting, including depulping, wet fermentation, and drying.

The validation set consisted of authentic Kopi Luwak, commercial Kopi Luwak, commercial regular coffee, fake coffee, and coffee blend. Authentic Kopi Luwak was produced via controlled processing to ensure the quality of the beans pre- and postdigestion. The remaining samples were purchased commercially. The coffee blend was utilized to examine the feasibility of the method for differentiating mixed and pure coffees.

Reagents. Methanol, chloroform, distilled water, ribitol, and pyridine were purchased from Wako Pure Chemical Industries, Ltd. (Osaka, Japan). Methoxyamine hydrochloride and alkane standard solution were purchased from Sigma-Aldrich (Milwaukee, WI, USA). *N*-Methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) was purchased from GL Science, Inc. (Tokyo, Japan). The six authentic standards of the discriminant markers, 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%), and glycolic acid (Sigma-Aldrich, 99%).

Extraction. Coffee beans were put into a grinding mill container, cooled for 3 min on water ice cubes, and then ground with a Retsch ball mill (20 Hz, 3 min). Coffee bean powder (15 mg) was transferred into a 2 mL Eppendorf tube. In addition to pure samples, a 50:50 (wt %) blend of Kopi Luwak and regular coffee was used. One milliliter of a single-phase extraction solvent consisting of 2.5:1:1 (v/v/v) methanol, distilled water, and chloroform, respectively, was added to extract a wide range of metabolites. A nonspecific extraction procedure was applied to avoid limiting the target analysis to specific compounds and to comprehensively screen the components of Kopi Luwak. As an internal standard, ribitol (60 μL, diluted with deionized water to 0.2 mg/mL) was utilized. The mixture was shaken for 1 min and then centrifuged at 4 °C and 16000*g* for 3 min. The supernatant (900 μL) was transferred into a 1.5 mL Eppendorf tube and diluted with 400 μL of Milli-Q water (Wako). The mixture was then vortexed and centrifuged for 3 min. A 400 μL portion of the aqueous phase was transferred into a fresh 1.5 mL Eppendorf tube with a screw cap. The solvent was removed by vacuum centrifugation for 2 h, followed by freeze-drying overnight. All samples were analyzed in triplicate ($n = 3$).

Derivatization for GC-MS analysis. Oximation and trimethylsilylation were used for derivatization. Methoxyamine hydrochloride (100 μL, 20 mg/mL in pyridine) was added to the dried extract as the first derivatization agent. The mixture was incubated at 30 °C for 90 min. After addition of the second derivatization agent, MSTFA (50 μL), the mixture was incubated at 37 °C for 30 min.

GC-MS Analysis. GC-Q/MS analysis was performed on a GCMS-QP 2010 Ultra (Shimadzu) equipped with a CP-SIL 8 CB low-bleed column (0.25 mm × 30 m, 0.25 μm, Varian Inc., Palo Alto, CA, USA) and an AOC-20i/s (Shimadzu) as an autosampler. The mass spectrometer was tuned and calibrated prior to analysis. The derivatized sample (1 μL) was injected in split mode, 25:1 (v/v), with an injection temperature of 230 °C. The carrier gas (He) flow was 1.12 mL/min with a linear velocity of 39 cm/s. The column temperature was held at 80 °C for 2 min, increased by 15 °C/min to 330 °C, and then held for 6 min. The transfer line and ion source temperatures were 250 and 200 °C, respectively. Ions were generated by electron ionization (EI) at 0.93 kV. Spectra were recorded at 10000 u/s over the mass range m/z 85−500. A standard alkane mixture (C_8−C_{40}) was injected at the beginning and end of the analysis for tentative identification.

Identification and Quantitation of Marker Candidates. The discriminant marker candidates were identified and quantitated against six authentic standards (malic acid, citric acid, glycolic acid, pyroglutamic acid, caffeine, and inositol) at various concentrations. The final concentrations of the authentic standards were adjusted to 1, 10, 50, 100, 250, 500, 750, 1000, 1500, and 2000 μM with the extraction solvent to construct a calibration curve. For extraction, the authentic standards were processed identically to the coffee bean samples. The standards were co-injected during sample analysis. Two blank solutions were prepared by adding only extraction solvent and distilled water, respectively. The limits of detection (LOD) and quantitation (LOQ) were determined via known protocols.[13,14] The construction of the standard curve and quantitation were conducted using GC-MS Solution software (Shimadzu). No authentic standards were detected in either of the blank samples.

Data Processing. Chromatographic data were converted into ANDI files (Analytical Data Interchange Protocol, *.cdf) using the GC-MS Solution software package (Shimadzu). Peak detection, baseline correction, and peak alignment of retention times were performed on the ANDI files using the freely available software package MetAlign.[15] Spectra were normalized manually by adjusting the peak intensity against the ribitol internal standard.

Retention indices of the eluted compounds were calculated on the basis of the standard alkane mixture. By comparing the retention indices and unique mass spectra with our in-house reference library, tentative identifications were obtained. To simplify and accelerate the tentative identifications of compounds that were registered in the in-house library database, AIoutput2 (version 1.29) annotation software, developed in the authors' laboratory, was utilized.[16] For comparison with the National Institute of Standards and Technology (NIST) library, retention times were used instead.

Multivariate Data Analysis. PCA and OPLS-DA were performed. OPLS-DA with an S alphabet-like plot, or S-plot, was chosen to isolate and select statistically significant and potentially biochemically interesting compounds. The S-plot provides covariance and correlation between metabolites and the modeled class to allow easier visualization. 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.[11] A 7-fold cross-validation was carried out to assess the accuracy of the discrimination model in practice. The goodness-of-fit (R^2) and predictability (Q^2) parameters were then determined. Analysis was performed with commercial software, SIMCA-P+ version 12 (Umetrics, Umeå, Sweden). Data were Pareto-scaled to reduce the effect of noise in the chromatograms.

To confirm the selection of significant compounds by OPLS-DA, data were also subjected to significance analysis of microarrays/metabolites (SAM) using MetaboAnalyst 2.0. At first, SAM was projected to microarray analysis to assign significance genes on the basis of changes in their expression.[12] Recently, MetaboAnalyst 2.0, a Web server for metabolomics analysis, was developed for the application of SAM to metabolomics data.[17,18]

The "relative differences," $d(i)$, that is, the differences in intensity for each peak, were calculated by a formula described elsewhere.[12] To

generate the control set, data were analyzed by *n*-balanced permutations. The relative differences of the permutated data, $d_p(i)$, were determined. Next, the "expected relative differences", $d_E(i)$, were calculated as the average of $d_p(i)$ over *n*-balanced permutations. Scatter plots of the expected relative differences $(d_E(i))$ and relative differences $(d(i))$ were constructed. Insignificant compounds were identified as those that satisfied $d(i) \cong d_E(i)$. A threshold value (Δ) was assigned as a certain distance displaced from $d(i) = d_E(i)$; compounds at distances larger than the threshold value were considered significant. The higher the threshold value, the lower the false discovery rate (FDR, the percentage of falsely significant compounds) and the number of significant compounds that can be obtained. Univariate statistics, performed by the freely available software, R project,[19] was applied to compare the means of selected significant compounds using Student's *t* test. Box plots were also constructed to display the differentiation among coffee samples. Differences were considered significant when $p < 0.05$.

RESULTS AND DISCUSSION

GC-MS-Based Metabolite Profiling of Kopi Luwak. GC-Q/MS analysis was performed on aqueous extracts of coffee beans to investigate the differences in their metabolite profiles and select discriminant markers for robust authentication. In addition, this research focused on increasing the scientific information about Kopi Luwak. A quadrupole mass spectrometer was selected because of its availability as the most widely used mass analyzer. However, a conventional Q/MS can be operated only at a slow scan rate.[20] With processor improvements and high-speed data processing, newly developed GC-Q/MS instruments provide increased sensitivity at high scan speeds of up to 10000 u/s.[21]

Because of their broad cultivation areas and commercial profitability, *C. arabica* and *C. canephora*, which represent 65 and 35% of the total annual coffee trade, respectively, were utilized for metabolomics analysis.[22] A total of 182 peaks from 21 coffee beans were extracted using MetAlign. Moreover, 26 compounds were tentatively identified by comparison with our in-house library (by retention index) and the NIST library (by retention time); 6 of these were identified by co-injection with an authentic standard. Tentatively identified components consisted of organic acids, sugars, amino acids, and other compounds (Supporting Information, Table S3). Previously reported coffee bean constituents, including chlorogenic, quinic, succinic, citric, and malic acids; caffeine, one of the compounds supplying bitter taste in coffee; and sucrose, the most abundant simple carbohydrate, were identified.[23−27]

In recent research, unsupervised analysis, PCA, has been employed for data exploration and to visualize information based on sample variance.[28,29] A PCA score plot derived from the 21 coffee beans differentiated two data groups on the basis of their species, Arabica and Robusta (Figure 1), and resulted in a goodness-of-fit parameter (R^2) of 0.844. Caffeine and quinic acid were significant for the Robusta coffee data sets, whereas the Arabica data set was mainly supported by various organic acids such as malic, chlorogenic, citric, and succinic acids. The data differentiation was explained by 42.9% of variance along PC1. The results indicated that genetic diversity more strongly influenced the data separation than animal perturbation.

Because of the large variance among coffee species, sample differentiation based on the type of coffee, Kopi Luwak or regular coffee, could not be observed. Additional analyses were carried out independently for each coffee species originating from the same cultivation area. The PCA score plot revealed data separation based on the type of coffee, in which Kopi

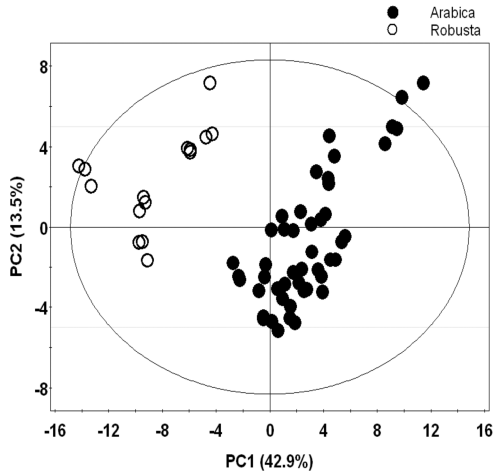

Figure 1. PCA score plot of experimental coffee set. *Coffea arabica* and *Coffea canephora* (Robusta) were clearly separated. Genetic diversity strongly influenced data separation.

Luwak and regular coffee could be clearly separated (Supporting Information, Figure S1). For the Arabica coffee data set, the separation was explained by 45.5 and 23.7% variances in PC1 and PC2, respectively. By PC2, Kopi Luwak was closely clustered in the same region, whereas regular coffees tended to separate on the basis of their cultivation areas. In the loading plot, malic and glycolic acids contributed highly to the Kopi Luwak data (data not shown). Thereby, coffee beans may possess similar profiles after animal digestion. Differences in cultivation areas were considered to have the least significance for data separation. In Robusta coffee, a clear separation between Kopi Luwak and regular coffee was observed, which was explained by 79.1% variance of PC1. Significant compounds for separation, including inositol and pyroglutamic acid for Kopi Luwak and quinic acid for regular coffee, were observed.

Discriminant Analysis To Select Candidates for Discriminant Markers. An overview of all data samples was provided by the unsupervised analysis, PCA. However, detailed information regarding compounds contributing to the data differentiation between Kopi Luwak and regular coffee remained unclear. Therefore, coffee bean data sets were subjected to supervised discriminant analysis (OPLS-DA). For analyses having two or more classes, OPLS-DA is the most suitable platform for isolating and selecting differentiation markers. Compounds with reliable, high contributions to the model may possess potentially biochemically interesting characteristics; thus, they can be selected as biomarker candidates.[11] All OPLS-DA models exhibited R^2 and Q^2 values >0.8, which would be categorized as excellent.[30] In addition, all models were in the range of validity after permutation tests using 200 variables (data not shown). The model was considered valid after permutation for those that met the following criteria: R^2Y-intercepts and Q^2-intercepts, which did not exceed 0.3−0.4 and 0.05, respectively.[31]

Potential candidates for discriminant markers can be selected via S-plots by setting the cutoff for covariance, $p[1]$, and the correlation value, $p[\mathrm{corr}]$, to > |0.2|. S-plots of the coffee data sets are shown in Figure 2B,D. In addition to cutoff values, candidates for discriminant markers were selected by variable importance in projection values (VIP). Large VIP values (>1) are more relevant for model construction.

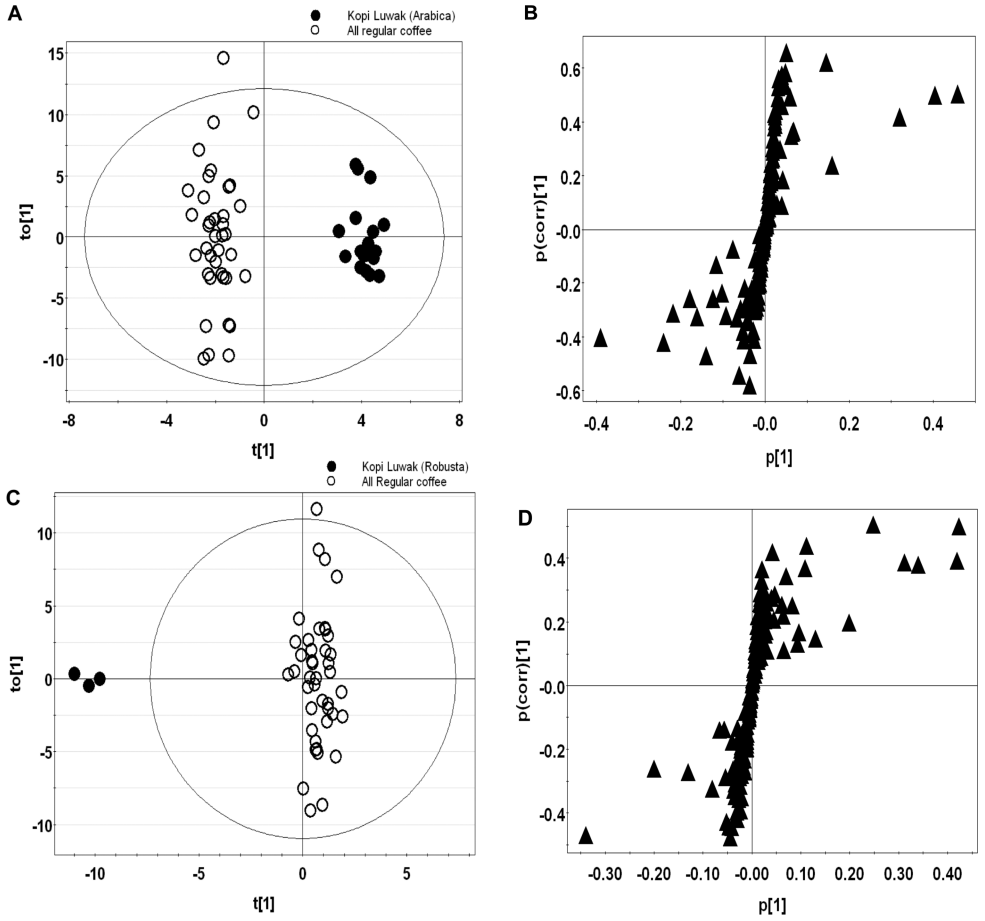

Figure 2. OPLS-DA score plots (A, C) and loadings of S-plots (B, D) derived from Arabica and Robusta coffees in the experimental coffee set. By assigning cutoff values in the S-plots, significant compounds were selected for p and p[corr] > |0.2|. Triangles indicate peaks detected by GC-MS.

Table 1. Candidates for Discriminant Markers from OPLS-DA and SAM and Analytical Parameters for Quantitation

			RSD (%) (n = 3)		linearity			
discriminant marker	RT (min)	VIP	RT (min)	area[a]	R^2	range (μM)	LOD (mg/kg)	LOQ (mg/kg)
glycolic acid	4.96	3.93	0.12	1.87	0.9999	1−1000	0.021	0.066
malic acid	9.05	5.53	0.05	2.29	0.9996	1−1000	0.043	0.132
pyroglutamic acid	9.43	1.7	0.05	3.36	0.9992	1−750	0.054	0.164
citric acid	11.61	5.6	0.04	3.29	0.9997	1−1000	0.504	1.526
caffeine	12.18	2.28	0.04	3.51	0.9961	100−2000	1.531	4.638
inositol	13.45	4.47	0.03	5.09	0.9974	1−1000	0.082	0.247

[a]At 100 μM.

The OPLS-DA score plot of Arabica coffee data sets is shown in Figure 2A. Discrimination between Kopi Luwak and regular coffee was obtained. The model was evaluated with R^2 and Q^2 values of 0.965 and 0.892, respectively. Interestingly, compounds that were uncorrelated with Kopi Luwak were quinic acid, caffeine, and caffeic acid. These compounds have been reported as contributors of bitterness as well as acidity in coffee.[23−26] In contrast, compounds that were predictive to Kopi Luwak, that is, over the cutoff value, included citric, malic, and glycolic acids. The OPLS-DA score plot of the Robusta coffee data sets (Figure 2C) was explained by R^2 and Q^2 values of 0.957 and 0.818, respectively. Caffeine, one of the bitter principles in coffee, was found to be significantly correlated with Robusta Kopi Luwak data sets. Robusta coffee has been reported to contain higher amounts of caffeine than Arabica. Thus, it tends to be bitter and flavorless, whereas Arabica coffee is considered to be milder, contains more aromatic compounds, and is more appreciated by the consumer.[32]

We employed SAM to select significant compounds as discriminant markers, as a comparison to OPLS-DA. In general, we used the same data sets, applying different multivariate data analysis. By assigning the lowest possible FDR value, a total of 12 compounds was considered as significant in the Arabica coffee data set (Figure S2 in the Supporting Information). Of these, citric acid, glycolic acid, malic acid, quinic acid, and other unidentified peaks were included. In the Robusta coffee data set, nine significant compounds were found (Supporting Information, Figure S2); inositol, caffeine, pyroglutamic acid, and six unidentified peaks exhibited high significance by SAM. Candidates for discriminant markers for the authentication assessment of Arabica and Robusta coffees are listed in Table 1. The selected marker candidates met significant criteria in both

OPLS-DA and SAM. Discriminant markers were chosen independently for the Arabica and Robusta coffee.

To confirm whether these selected markers were generated as a result of animal digestion, we investigated cause−effect relationships by quantitating the discriminant marker candidates in green and roasted coffee beans from controlled processing, pre- and post- animal digestion (samples 5 and 11, experimental sets). The results are displayed in Figure S3 of the Supporting Information. In both the raw and roasted beans, citric acid was present in higher concentration after animal digestion, exhibiting a significant value difference ($p < 0.05$) between Kopi Luwak and regular coffee. The concentration of caffeine was also increased after digestion, but the difference was insignificant ($p > 0.05$). As a result of roasting, the glycolic acid concentration increased dramatically ($p < 0.001$) from 0.8 to 25−28 mg/kg. The production of aliphatic acids, including formic, acetic, glycolic, and lactic acids, has been reported during coffee roasting.[33] Therefore, among the selected marker candidates, we confirmed citric acid as a potential marker generated by animal digestion. Passage through the civet's gut, enriched with gastric juices and microbial activity, may have contributed to the increased levels of particular organic acids. Citric acid, malic acid, quinic acid, and chlorogenic acid are the main acids in coffee, and acidity is generated by complex reactions involving these organic acids during roasting.[24] Kopi Luwak has been reported to exhibit slightly higher acidity than regular coffee.[3] However, the correlation between the increased levels of particular acids as result of animal digestion and the total acidity in coffee after roasting remains obscure and requires further investigation.

Validation of the Applicability of Discriminant Markers for Authenticity Assessment. To verify the applicability of the selected marker candidates, we analyzed a validation coffee set that included authentic Kopi Luwak, commercial Kopi Luwak, commercial regular coffee, fake coffee, and coffee blend. With the exception of the authentic coffee, the remaining samples were purchased commercially. Generally, from harvest to preroasting, samples labeled "commercial Kopi Luwak" and "commercial regular coffee" were processed similarly to the corresponding coffees in the experimental set. However, in some cases, different roasting parameters were applied. Fake coffee was processed to approximate the sensory profile of Kopi Luwak.[4] Commercial regular coffees were selected from different production areas.

To examine the effectiveness of the selected markers in differentiating pure and coffee blends, we mixed two commercial digested coffees, Kopi Luwak Golden and Kopi Luwak Wahana, with a commercial regular coffee (Wahana regular) in a 50:50 (wt %) ratio. This would also compare the applicability of the discriminant markers when coffee beans from the same and different production areas were blended. Despite being selected independently, the six marker candidates were used together for method validation.

By subjecting all detected peaks to PCA, samples were populated into four clusters. The largest variance corresponded to fake coffee, as its results were clearly separated from others (data not shown). Next, we projected the six marker candidates as an inclusion list into the PCA to obtain an overview of their applicability toward sample differentiation. Similarly to the previous results, separation of the four coffee groups was observed. The PCA was explained by 59.5 and 20.9% variances in PC1 and PC2, respectively (Figure 3). Fake coffee was clustered away by PC1. Separation was likely because of

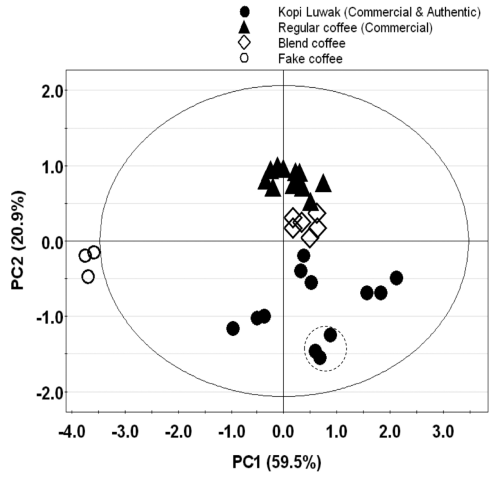

Figure 3. PCA score plot of validation coffee set. Data plotted inside dashed circle indicate authentic Kopi Luwak.

attempts by the producer to obtain a profile similar to that of Kopi Luwak. In PC2, commercial Kopi Luwak, coffee blend, and commercial regular coffee could be differentiated. Both authentic and commercial Kopi Luwak were clustered within a close distribution area. Regardless of their origins and processing (roasting) parameters, commercial regular coffee data were populated in a close area, suggesting that these factors had the least significance for data separation. From the loading plot information, citric acid, malic acid, and inositol exhibited high contribution values for the Kopi Luwak data sets. Interestingly, these three marker candidates also showed the highest VIP values for constructing the discriminant model (Table 1).

To display the applicability of the selected discriminant markers in the differentiation of samples in the validation set, box plots were constructed using the relative peak intensities of citric acid, malic acid, and inositol. The box plots of malic acid and citric acid were able to differentiate commercial Kopi Luwak (Kopi Luwak Wahana), coffee blend, commercial regular coffee (Wahana regular), and fake coffee. However, the inositol box plot failed to differentiate these samples. Hence, we selected a double marker that employed an inositol/pyroglutamic acid ratio (Figure 4). Pyroglutamic acid was selected because it had the lowest contribution toward the separation of Kopi Luwak and regular coffee (data not shown). The box plot for the other commercial Kopi Luwak (Golden Kopi Luwak) and the comparisons with its coffee blend as well as fake coffee are displayed in Figure S4 in the Supporting Information.

We confirmed the ratio of the coffee blend by quantifying the discriminant marker constituents. The analytical parameters for quantitation are shown in Table 1. All authentic standards exhibited good linearity (0.99 or higher) and good repeatability for at least seven points in the applied concentration range in which analysis could be performed. To examine the quantitation validity, the LOD and LOQ for each discriminant marker were determined. The concentrations of the discriminant marker candidates in the coffee samples were determined to be higher than the LOD and LOQ of authentic standards. The concentration ratios of the selected markers, malic acid, citric acid, and the inositol/pyroglutamic acid ratio, in all of the sample blends ranged from 47.76 to 53.73%. This result showed a relatively low error in terms of the ratios of the

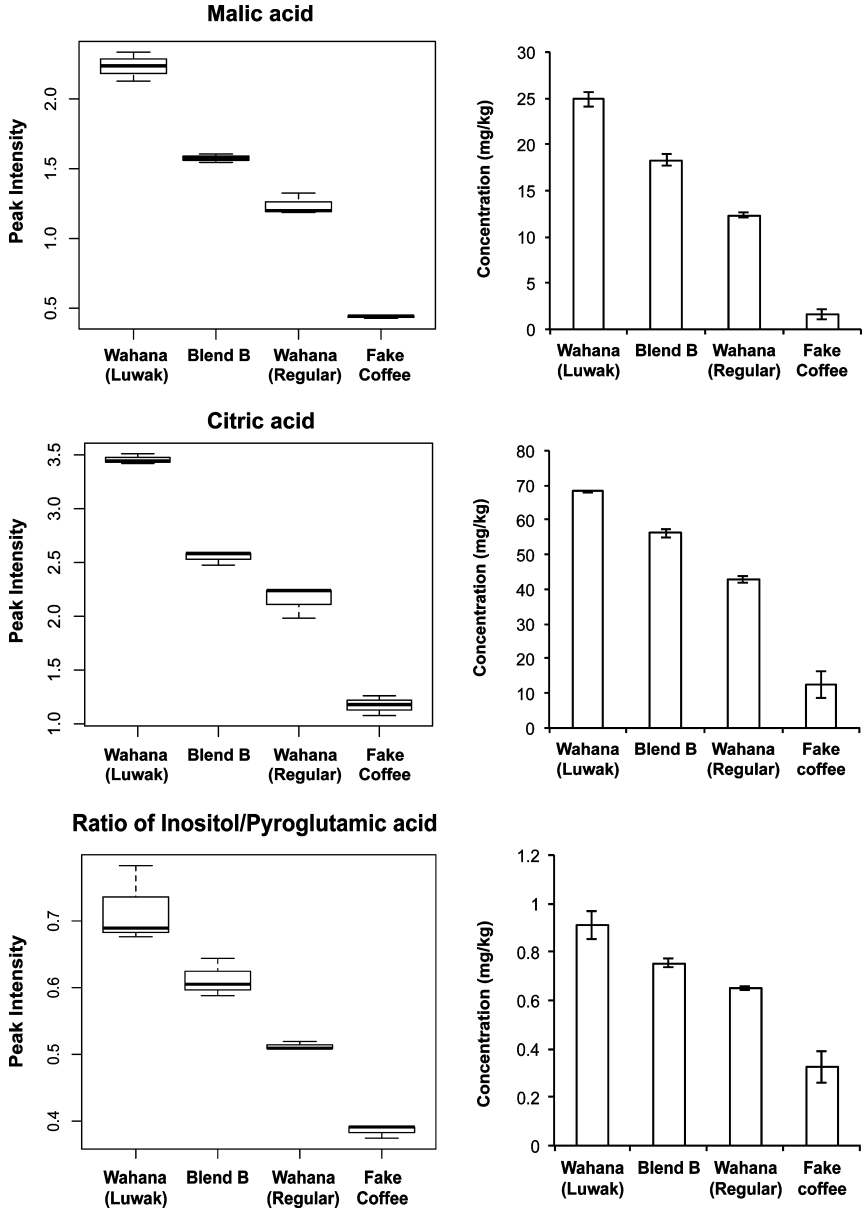

Figure 4. Box plots of peak intensity and concentrations of selected discriminant markers for validation.

discriminant markers in sample blends compared with their actual values. Moreover, the concentration of each discriminant marker corresponded well with the box plot constructed from its respective peak intensity (Figure 4 and Supporting Information, Figure S4). Hence, we confirmed the feasibility of using the proposed strategy for the robust authentication of coffee blend in a 50:50 (wt %) ratio.

In summary, to the best of our knowledge, this investigation represents the first attempt to address discriminant markers for the authentication of Kopi Luwak. Sample differentiation was greatly influenced by genetic diversity (coffee species), followed by decreasing contributions from animal perturbation and cultivation area. Because of the great variation among coffee species, candidates for the discriminant markers were selected independently for each species. The selected discriminant marker candidates were verified for the authentication of commercial coffee products. The proposed markers were able to differentiate commercial Kopi Luwak, commercial regular coffee, and fake coffee. In addition, at a certain ratio (50 wt % Kopi Luwak content), the feasibility of employing these discriminant markers to differentiate pure and mixed coffee was acceptable. Our findings highlighted the utility of metabolic profiling using GC-MS combined with multivariate analysis for the selection of discriminant markers for the authenticity assessments of valuable agricultural products. Discriminant markers are expected to perform as sole markers or in combination with sensory analysis by trained experts for the authentication of Kopi Luwak.

■ ASSOCIATED CONTENT

Supporting Information

List of coffee samples (Tables S1 and S2); tentative identification (Table S3); PCA score plot of coffee beans from same cultivation area (Figure S1); SAM analysis (Figure S2); concentration levels of discriminant markers from controlled processing (Figure S3); box plots and concentrations of discriminant markers for validation (Figure S4). This

material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*(E.F.) Mailing address: Department of Biotechnology Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan. Phone: +81-6-6879-7424. Fax: +81-6-6879-7424. E-mail: fukusaki@bio.eng.osaka-u.ac.jp.

Notes

The authors declare no competing financial interest. The study represents a portion of the dissertation submitted by U.J. to Osaka University in partial fulfillment of the requirements for his Ph.D. degree.

ACKNOWLEDGMENTS

We express our deepest appreciation to the Indonesian Coffee and Cocoa Research Institute (ICCRI) for providing and preparing coffee samples.

ABBREVIATIONS USED

GC-Q/MS, gas chromatography−quadrupole mass spectrometry; PCA, principal component analysis; OPLS-DA, orthogonal projection to latent structures-discriminant analysis; LOD, limit of detection; LOQ, limit of quantitation; PC, principal component; ANDI, analytical data interchange protocol; SAM, significance analysis of microarrays/metabolites; FDR, false discovery rate; MTSFA, methyl-*N*-(trimethylsilyl)-trifluoroacetamide

REFERENCES

(1) Onishi, N. From dung to coffee brew with no aftertaste; http://www.nytimes.com/2010/04/18/world/asia/18civetcoffee.html (most recent access April 3, 2013).

(2) Joshi, A. R.; Smith, J. L. D.; Cuthbert, F. Influence of food distribution and predation pressure on spacing behavior in palm civets. *J. Mammalogy* **1995**, *76*, 1205−1212.

(3) Marcone, M. F. Composition and properties of Indonesian palm civet coffee (Kopi Luwak) and Ethiopian civet coffee. *Food Res. Int.* **2004**, *37*, 901−912.

(4) White Koffie first low acid coffee; http://blog.kopiluwak.org/white-koffie-first-low-acid-coffee (in Indonesian language, most recent access April 3, 2012).

(5) Blow, N. Metabolomics: biochemistry's new look. *Nature* **2008**, *455*, 697−700.

(6) Choi, M.; Choi, W.; Park, J. H.; Lim, J.; Kwon, S. W. Determination of coffee origins by integrated metabolomics approach of combining multiple analytical data. *Food Chem.* **2010**, *121*, 1260−1268.

(7) Wei, F.; Furihata, K.; Koda, M.; Hu, F.; Kato, R.; Miyakawa, T.; Tanokura, M. ^{13}C NMR-based metabolomics for the classification of green coffee beans according to variety and origin. *J. Agric. Food Chem.* **2012**, *60*, 10118−10125.

(8) Fujimura, Y.; Kurihara, K.; Ida, M.; Kosaka, R.; Miura, D.; Wariishi, H.; Maeda-Yamamoto, M.; Nesumi, A.; Saito, T.; Kanda, T.; Yamada, K.; Tachibana, H. Metabolomics-driven nutraceutical evaluation of diverse green tea cultivars. *PloS One* **2011**, *6*, e23426.

(9) Pongsuwan, W.; Fukusaki, E.; Bamba, T.; Yonetani, T.; Yamahara, T.; Kobayashi, A. Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite profiling. *J. Agric. Food Chem.* **2007**, *55*, 231−236.

(10) Jonsson, P.; Johansson, E. S.; Wuolikainen, A.; Lindberg, J.; Schuppe-Koistinen, I.; Kusano, M.; Sjostrom, M.; Trygg, J.; Moritz, T.; Antti, H. Predictive metabolite profiling applying hierarchical multivariate curve resolution to GC-MS data — a potential tool for multiparametric diagnosis. *J. Proteome Res.* **2006**, *5*, 1407−1414.

(11) Wiklund, S.; Johansson, E.; Sjostrom, L.; Mellerowicz, E. J.; Edlund, U.; Shockcor, J. P.; Gottfries, J.; Moritz, T.; Trygg, J. Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal. Chem.* **2008**, *80*, 115−122.

(12) Tusher, V. G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 5116−5121.

(13) Guan, H.; Brewer, W. E.; Morgan, S. L. New approach to multiresidue pesticide determination in foods with high fat content using disposable pipette extraction (DPX) and gas chromatography-mass spectrometry (GC-MS). *J. Agric. Food Chem.* **2009**, *57*, 10532−10538.

(14) The European Agency for the Evaluation of Medicinal Products, Human Medicines Evaluation Unit, Validation of Analytical Procedures: Methodology, ICH Topic Q 2B, Nov 1996.

(15) Plant Research International. MetAlign; http://www.wageningenur.nl/nl/Expertises-Dienstverlening/Onderzoeksinstituten/rikilt/show/MetAlign.htm (most recent access April 3, 2013).

(16) Tsugawa, H.; Tsujimoto, Y.; Arita, M.; Bamba, T.; 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 Bioinf.* **2011**, *12*, 131−143.

(17) Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, *37*, 652−660.

(18) Zhang, A.; Sun, H.; Han, Y.; Yuan, Y.; Wang, P.; Song, G.; Yuan, X.; Zhang, M.; Xie, N.; Wang, X. Exploratory urinary metabolic biomarkers and pathways using UPLC-Q-TOF-HDMS coupled with pattern recognition approach. *Analyst* **2012**, *137*, 4200−4208.

(19) Institute for Statistics and Mathematics of Wirtschaftsuniversität Wien. The R project for statistical computing; http://www.r-project.org/ (most recent access April 3, 2013).

(20) Dalluge, J.; Vreuis, R. J. J.; van Iperen, D. J.; van Rijn, M.; Brinkman, U. A. T. Resistively heated gas chromatography coupled to quadrupole mass spectrometry. *J. Sep. Sci.* **2002**, *25*, 608−614.

(21) Shimadzu, Inc. GCMS-QP2010 Ultra; http://www.shimadzu.com/an/gcms/qpultra.html (most recent access April 3, 2013).

(22) Wermelinger, T.; D'Ambrosio, L.; Klopproge, B.; Yeretzian, C. Quantification of the Robusta fraction in a coffee blend via Raman spectroscopy: proof of principle. *J. Agric. Food Chem.* **2011**, *59*, 9074−9079.

(23) Moon, J. K.; Yoo, H. Y.; Shibamoto, T. Role of roasting conditions in the level of chlorogenic acid content in coffee beans: correlation with coffee acidity. *J. Agric. Food Chem.* **2009**, *57*, 5365−5369.

(24) Rodrigues, C.; Marta, L.; Maia, R.; Miranda, M.; Ribeirinho, M.; Maguas, C. Application of solid-phase extraction to brewed coffee caffeine and organic acid determination by UV/HPLC. *J. Food Compos. Anal.* **2007**, *20*, 440−448.

(25) Blumberg, S.; Frank, O.; Hofmann, T. Quantitative studies on the influence of the bean roasting parameters and hot water percolation on the concentrations of bitter compounds in coffee brew. *J. Agric. Food Chem.* **2010**, *58*, 3720−3728.

(26) Perrone, D.; Donagelo, C. M.; Adriana, F. Fast simultaneous analysis of caffeine, trigonelline, nicotinic acid, and sucrose in coffee by liquid chromatography−mass spectrometry. *Food Chem.* **2008**, *110*, 1030−1035.

(27) Wei, F.; Furihata, K.; Koda, M.; Hu, F.; Miyakawa, T.; Tanokura, M. Roasting process of coffee beans as studied by nuclear magnetic resonance: time course of changes in composition. *J. Agric. Food Chem.* **2012**, *60*, 1005−1012.

(28) Liberto, E.; Ruosi, M. R.; Cordero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B. Non-separative headspace solid phase microextraction-mass spectrometry profile as a marker to monitor coffee roasting degree. *J. Agric. Food Chem.* **2013**, *61*, 1652−1660.

(29) Ruosi, M. R.; Cordero, C.; Cagliero, C.; Rubiolo, P.; Bicchi, C.; Sgorbini, B.; Liberto, E. A further tool to monitor the coffee roasting process: aroma composition and chemical indices. *J. Agric. Food Chem.* **2012**, *60*, 11283−11291.
(30) Lundstedt, T.; Seifert, E.; Abramo, L.; Thelin, B.; Nystrom, A.; Pettersen, J.; Bergman, R. Experimental design and optimization. *Chemom. Intell. Lab. Sys.* **1998**, *42*, 3−40.
(31) Erikkson, L.; Jaworska, J.; Worth, A. W.; Cronin, M. T. D.; McDowell, R. M.; Gramatica, P. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. *Environ. Health Perspect.* **2003**, *111*, 1361−1375.
(32) Wintgens, J. N. Botany, genetics and genomics of coffee. In *Coffee: Growing, Processing, Sustainable Production a Guidebook for Growers, Processors, Traders and Researchers*, 2nd rev. ed.; Wiley VCH: Weinheim, Germany, 2009; Vol. 2, pp 25−60.
(33) Ginz, M.; Balzer, H. H.; Bradbury, A. G. W.; Maier, H. G. Formation of aliphatic acids by carbohydrate degradation during roasting of coffee. *Eur. Food Res. Technol.* **2000**, *211*, 404−410.