Application of GC/MS and GC/FID-based metabolomics for authentication of Asian palm civet coffee (Kopi...

Article · June 2015

CITATIONS
0

READS
233

3 authors:

Sastia Prama Putri
Osaka University
29 PUBLICATIONS **283** CITATIONS
SEE PROFILE

Udi Jumhawan
12 PUBLICATIONS **40** CITATIONS
SEE PROFILE

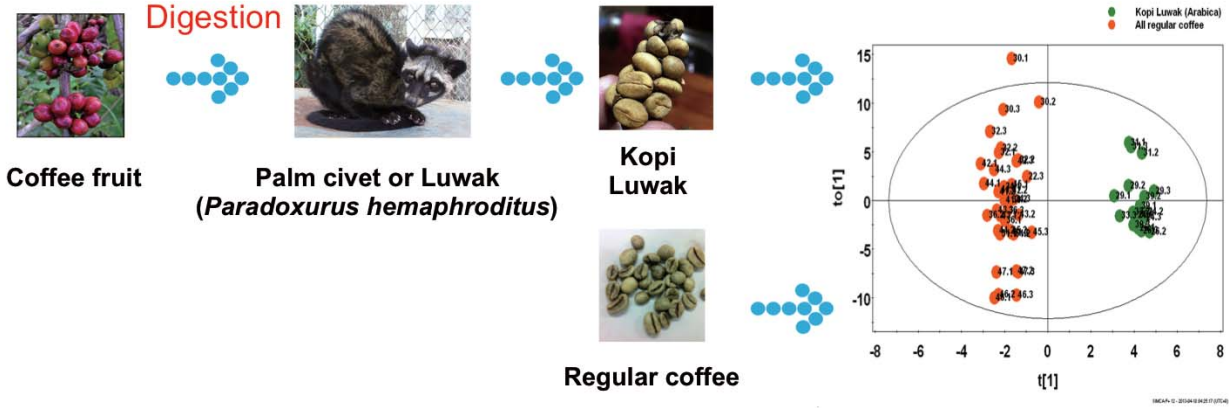
Eiichiro Fukusaki
Osaka University
376 PUBLICATIONS **5,967** CITATIONS
SEE PROFILE

Application of GC/MS and GC/FID-based metabolomics for authentication of Asian palm civet coffee (Kopi Luwak)

Sastia P. Putri[1], Udi Jumhawan[1] and Eiichiro Fukusaki[1]
[1]Osaka University, Osaka, Japan

Abstract

Kopi Luwak, world's most expensive coffee firstly originated from Indonesia, is made from coffee berries that have been digested by the animal Asian palm civet (*Paradoxurus hemaphroditus*). Despite its profitable prospect, there is no reliable and standardized method for determining its authenticity. Therefore, research on the development of robust method for authentication of Kopi Luwak is urgently needed, particularly to prevent fraud in market worldwide. Twenty-one coffee samples (*Coffea arabica, Coffea canephora*) of three cultivation areas were firstly analyzed by GC-Q/MS. Principal Component Analysis (PCA), Orthogonal Projection to Latent Structures-Discriminant Analysis (OPLS-DA) and Significance Analysis of Microarrays/Metabolites (SAM) of Kopi Luwak was employed to select discrimination marker candidates for authentication of Kopi Luwak. Applicability of discriminant marker candidates was verified using commercial samples. We further established a rapid, reliable and cost-effective analysis employing a universal detector, GC coupled with flame ionization detector (FID) for discrimination analysis of 37 commercial and non-commercial coffee beans extracts. Our study demonstrated that GC/FID-based metabolite fingerprinting can be effectively actualized as an alternative method for coffee authenticity screening in industries.



Keywords: Metabolomics; Coffee; Kopi Luwak; Discriminant marker; Authentication; GCMS-QP 2010 Ultra; GC-FID; GC-2010

1. Introduction

The world's most expensive and unique coffee is Kopi Luwak (Kopi is an Indonesian word for coffee and luwak is the animal producing it), one of Indonesia's exotic agricultural products[3]. Marketed with a price tag of 300-400 USD per kg[4], Kopi Luwak's high selling price is attributed to its rarity, unique flavor, and interesting production process. Kopi Luwak is made from coffee cherries that have been eaten by common palm civets (Luwak or *Paradoxurus hemaphroditus*), which use their keen sense of smell to select the best and ripest beans. Kopi Luwak's high quality is mainly contributed by two factors: natural selection of the best coffee cherries by luwak and changes that occur in the digestive track of luwak, which yield an aromatic coffee with flavor described as earthy, syrupy, musty, smooth, and rich with chocolate undertones[5].

Despite its profitable prospect, there is no reliable and standardized method for determining originality of Kopi Luwak. The limited availability of the authentic product and its increasing popularity has opened the possibility of adulteration by blending Kopi Luwak with regular coffee or *in vitro* enzymatic treatment of regular coffee beans to mimic the fermentation by civet in order to increase the production and to meet market demand. This poses serious concern among consumers over the authenticity and the quality of the products currently available in the market. Therefore, research on the development of an unbiased and reproducible method for Kopi Luwak authentication and discrimination is urgently needed, particularly to prevent fraud in Kopi Luwak market worldwide.

Discrimination of Kopi Luwak and regular coffee has been achieved using electronic nose data[5]. However, selection of discriminant marker for authentication was not addressed. The method currently employed by Kopi Luwak producers is by visual and organoleptic testing. Both methods are inadequate because visual examination would only be possible for green coffee bean prior to roasting. As for organoleptic testing, trained experts that could discriminate Kopi Luwak are very few and the test tends to be highly subjective.

Information flows in metabolic pathways are highly dynamic and represent current biological state of individual cells. Hence, metabolome has been considered as the best descriptor of physiological phenomenon[6]. With this capability, metabolomics technique can be a powerful tool to elucidate variations in phenotype imposed by any perturbations such as gene modification, environmental factor, and physical stress. Processes inside animal digestive tract could be translated as physical and enzymatic stress to coffee bean, as it reported to pose smoother surface and color changes after digestion[3]. Thus, metabolomics technique was selected to seek and select discriminant marker for authenticity assessment of Kopi Luwak. Metabolomics technique has been effectively applied to distinguish phytochemical compositions of agricultural products among different origins, varieties, and cultivars for quality control and breeding[7-8].

In this study, gas chromatography coupled with quadrupole mass spectrometry (GC-Q/MS) based metabolic profiling was employed to identify discriminant marker for differentiation of Kopi Luwak and regular coffee. A combination of gas chromatography and mass spectrometry (GC/MS) has demonstrated as an effective analytical platform as it provides high sensitivity, reproducibility and quantitation of large amount of metabolites within a single step extraction[8-9]. Samples classification by means of chemometrics was performed using principal component analysis (PCA). Subsequently, orthogonal projection to latent structures combined with discriminant analysis (OPLS-DA) and significance analysis of microarrays/metabolites (SAM) to identify statistically significant compounds as discriminant marker candidates, was utilized[10-11]. The applicability of discriminant marker candidates was verified to determine authenticity of commercial coffee.

However, a major drawback of this GC-MS-based 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. 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[2].

2. Experimental

Samples and chemicals

Samples were divided into experimental and validation coffee sets. The first set included twenty one coffee beans that were collected from several cultivation areas in Indonesia. Kopi Luwak and regular coffee samples of two species, *Coffea arabica* (Arabica) and *Coffea canephora* (Robusta), were utilized. Coffee samples were obtained from 21 sampling points of three cultivation areas in Indonesia (Java, Sumatera, Bali) as shown in Table 1. In the second experiment for GC-FID application, one Robusta civet coffee sample (Sample no. 17 in Table 1) is removed from analysis resulting in only 20 coffee bean samples. The experimental coffee sets include 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 validation coffee sets. The first validation set consists of authentic Kopi Luwak, commercial Kopi Luwak, commercial regular coffee, imitation/adulteration coffee, and blend coffee (Table 2a). The second validation set consists of 3 civet coffees and 3 regular coffees were bought commercially and 2 additional authentic civet coffees from the Indonesian Coffee and Cocoa Research Institute (Table 2b). 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. All chemicals used in this study were analytical grade.

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 16000 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.

Table 1 List of coffee samples (Experimental coffee set)

No.	Origin [a]	Species	Type [b]	Harvest year	Samples extraction [c] period	Samples [d] code
1	Sumatra	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	29
2	Sumatra	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	31
3	Sumatra	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	39
4	Bali	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	33
5	Java	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	24
6	Java	*Coffea arabica*	Kopi Luwak	2011	Sept 2011	26
7	Sumatra	*Coffea arabica*	Regular coffee	2011	Sept 2011	30
8	Sumatra	*Coffea arabica*	Regular coffee	2011	Sept 2011	32
9	Sumatra	*Coffea arabica*	Regular coffee	2011	Sept 2011	40
10	Bali	*Coffea arabica*	Regular coffee	2010	Sept 2011	34
11	Java	*Coffea arabica*	Regular coffee	2011	Sept 2011	22
12	Java	*Coffea arabica*	Regular coffee	2011	Sept 2011	41
13	Java	*Coffea arabica*	Regular coffee	2010	Sept 2011	42
14	Java	*Coffea arabica*	Regular coffee	2011	Sept 2011	45
15	Bali	*Coffea arabica*	Regular coffee	2011	Sept 2011	43
16	Bali	*Coffea arabica*	Regular coffee	2011	Sept 2011	44
17	Sumatra	*Coffea canephora*	Kopi Luwak	2010	Sept 2011	35
18	Sumatra	*Coffea canephora*	Regular coffee	2010	Sept 2011	36
19	Java	*Coffea canephora*	Regular coffee	2011	Sept 2011	46
20	Java	*Coffea canephora*	Regular coffee	2011	Sept 2011	47
21	Java	*Coffea canephora*	Regular coffee	2010	Sept 2011	51

[a] Samples no.1 – 5 and 7 – 11 were originated from same cultivation area
[b] Processing: regular coffee (wet fermentation after harvested), Kopi Luwak (after secreted by animal, washed, and then wet-fermentation)
[c] Samples were obtained from ICCRI during harvest period (around June 2011). Samples from 2010 are ICCRI's coffee collection. All samples were stored at 4°C and extracted per September 2011
[d] For multivariate analysis

Table 2a List of coffee samples (Validation coffee set) for the GC/MS-based metabolite profiling to seek for discriminant markers

No.	Brand [a]	Origin	Species	Species	Production year	Samples extraction period
1	Kopi Luwak Andungsari	East Java, Indonesia	*Arabica*	(Authentic) Kopi Luwak	2012	Feb 2013
2	Golden Kopi Luwak	West Java, Indonesia	*Arabica*	(Commercial) Kopi Luwak	2012	Feb 2013
3	Kopi Luwak Wahana	Sumatra, Indonesia	*Arabica*	(Commercial) Kopi Luwak	2012	Feb 2013
4	Kopi Luwak Bali	Bali, Indonesia	*Unknown*	(Commercial) Kopi Luwak	2011	Feb 2013
5	Kopi Wahana	Sumatra, Indonesia	*Arabica*	(Commercial) regular coffee	2012	Feb 2013
6	Kona coffee	Hawaii, USA	*Arabica*	(Commercial) regular coffee	2012	Feb 2013
7	Cerrado Chapado coffee	Brazil	*Arabica*	(Commercial) regular coffee	2012	Feb 2013
8	Aceh Special	Sumatra, Indonesia	*Arabica*	(Commercial) regular coffee	2012	Feb 2013
9	White Koffie	Java, Indonesia	*Arabica*	adulteration coffee	2012	Feb 2013
10	Blend coffee A [b]					
11	Blend coffee B [c]					

[a] Provider 1 ICCRI
2 CV. Kopi Luwak
3 Wahana-Mandheling Kopi
4 Hema-Wiwi Bali
5 Wahana-Mandheling Kopi
6 Hiro coffee
7 Hiro coffee
8 Hiro coffee
9 PT. Java Prima Abadi

[b] Blend between Golden Kopi Luwak (No.2) and Kopi Wahana (No.5) with ratio 50:50 (w/w)

[c] Blend between Kopi Luwak Wahana (No.3) and Kopi Wahana (No.5) with ratio 50:50 (w/w)

Table 2b List of coffee samples (Validation coffee set) for the GC/FID-based metabolite fingerprinting for validation of method for routine analysis

No.	Brand [a]	Production area	Classification	Production year	Samples ID [b]
1	Andungsari	Java, Indonesia	Civet coffee (authentic)	2012	Au_1
2	Andungsari	Java, Indonesia	Civet coffee (authentic)	2013	Au_2
3	Golden	Java, Indonesia	Civet coffee (commercial)	2012	GO
4	Wahana	Sumatra, Indonesia	Civet coffee (commercial)	2012	WL
5	Bali	Bali, Indonesia	Civet coffee (commercial)	2011	BA
6	Wahana	Sumatra, Indonesia	Regular coffee (commercial)	2012	WR
7	Kona	Hawaii, USA	Regular coffee (commercial)	2012	KO
8	Cerrado Chapado	Brazil, Brazil	Regular coffee (commercial)	2012	CE
9	Coffee blends (9 samples)	Combination of 3 civet coffee (commercial) & 3 regular coffee (commercial)			

[a] Provider 1, 2 ICCRI, Indonesia
3 CV. Kopi Luwak, Indonesia
4, 5 Wahana-Mandheling, Indonesia
6 Hema-Wiwi Bali, Indonesia
7, 8 Hiro coffee, Japan

[b] For multivariate analysis

GC/FID and GC/MS analysis

GC/MS analysis was carried out using GC-Q/MS, GCMS-QP 2010 Ultra (Shimadzu, Kyoto, Japan), installed with CP-SIL 8 CB low bleed column; 0.25 mm × 30 m, 0.25 µm (Varian Inc., Palo Alto, California, USA) and AOC-20i/s autoinjector (Shimadzu, Kyoto, Japan) as an autosampler. Mass spectrometer was tuned and calibrated prior to analysis. A 1 µL of derivatized sample was injected in split mode, 25/1 (v/v), with injection temperature was set to 230°C. The carrier gas flow (Helium) was 1.12 mL/minute with linear velocity, 39 cm/sec. The column temperature was held at 80°C for 2 minute, raised by 15°C/minute to 330°C, and then held for 6 minute. Transfer line and ion source temperature was set to 250°C and 200°C, respectively. A 0.93 kV of electron ionization (EI) was applied to generate ion. Spectra were recorded by 10000 u/sec within mass range 85 - 500 *m/z*. Standard alkane mixture (C_8-C_{40}) was injected at the beginning and end of analysis for tentative identification.

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), 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

Chromatographic data from GC/MS were converted into ANDI files (Analytical Data Interchange Protocol, *.cdf). This feature include in GC/MS Solution software (Shimadzu, Kyoto, Japan). These ANDI files were subjected to freely available software, MetAlign version 041011, to perform peak detection, baseline correction and peak alignment of retention times[15]. Spectra were normalized manually by adjusting the peak intensity of each sample with internal standard, ribitol. Retention indexes of eluted compounds were calculated based on standard alkane mixture. By comparing its retention indexes and unique mass spectra with in-house reference library constructed from 500 authentic standard chemicals, tentative identification was performed. For comparison with NIST library, retention time was used instead. To simplify and accelerate tentative-identification with compounds that registered in in-house library database, Aloutput2 version 1.29, annotation software developed by authors' laboratory, was utilized[12]. 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, Washington, USA). The data matrix was then subjected to multivariate analysis.

Multivariate data analysis

The details on discriminant marker identification for Kopi Luwak authentication by means of multivariate analyses, namely PCA and OPLS-DA, have been described elsewhere[1]. Briefly, the coffee bean data sets were subjected to supervised discriminant analysis, Orthogonal projection to latent structures-discriminant analysis (OPLS-DA). OPLS-DA was selected to seek and select statistically significant discriminant markers for Kopi Luwak authentication. To confirm selection of significant compounds by OPLS-DA, data were also subjected to MetaboAnalyst 2.0 to perform signifincance analysis of microarrays/metabolites (SAM).
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. PCA and OPLS-DA were used to decipher the relationships between two data matrices, *X* (predicted variables), and *Y* (observed variables)[10]. 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.

3. Results and Discussion

GC/MS-based metabolite profiling of Kopi Luwak

GC-Q/MS analysis was performed on aqueous extracts of Kopi Luwak and regular coffee bean to investigate the differences in their metabolite profiles to select discriminant marker for robust authentication. Quadrupole mass spectrometer (Q/MS) was selected due to its availability as the most widely used mass analyzer. Therefore, the application of GC/Q-MS is expected to meet with

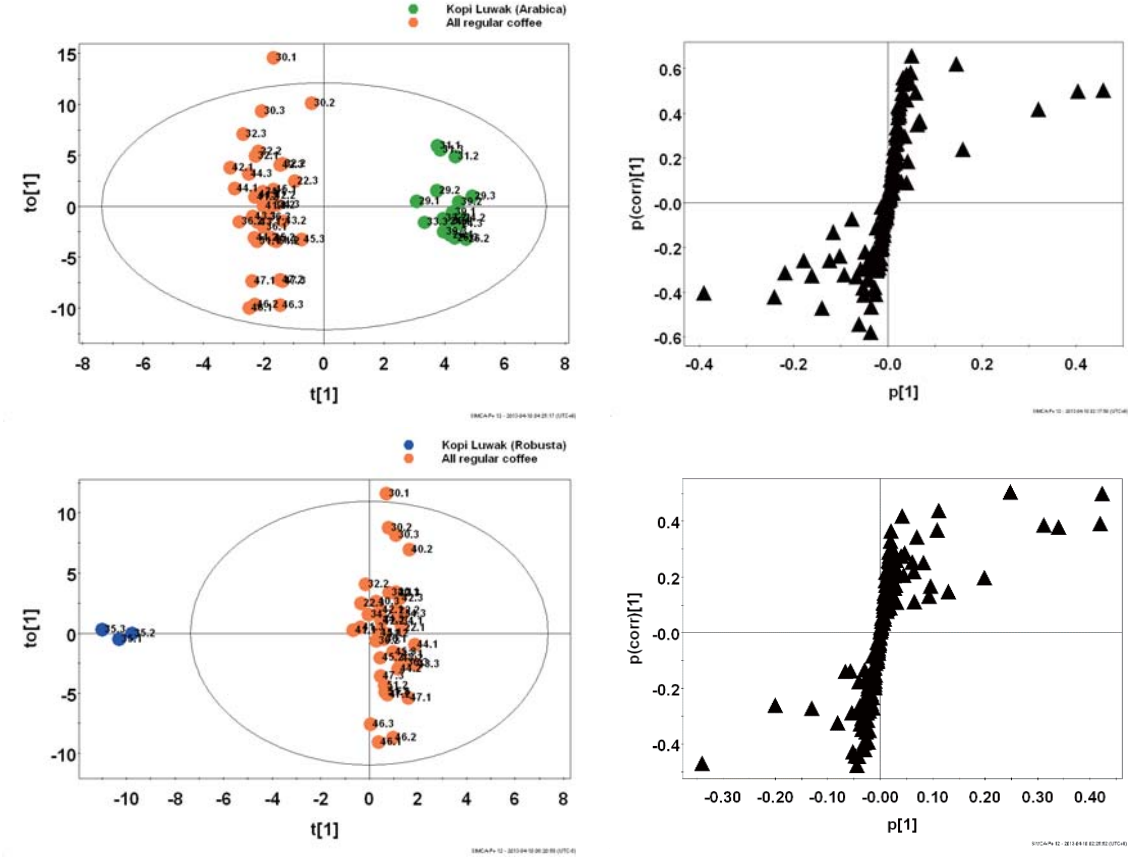


Fig. 1 PCA score plot of Kopi Luwak and regular coffee from same cultivation area, Arabica (right) and Robusta (left).

research objectives. However, the conventional Q/MS can be operated only at slow scan rate[13]. With the improvement of processor and high-speed data processing, newly developed GC-Q/MS provides increased sensitivity in high scan speed up to 10,000 u/sec[14].
Total of 182 peaks from 21 coffee bean samples were extracted using freely available software, MetAlign. Twenty compounds were tentatively identified by comparison with in-house library (retention index) and NIST library (retention time) and six compounds were identified by co-injecting authentic standards. Tentatively identified compounds consist of organic acids, sugars, amino acids, and other compounds. Compounds that have been reported previously in research of coffee bean, including chlorogenic acid, quinic acid, succinic acid, citric acid, and malic acid, caffeine, one of compounds for bitter taste in coffee, and sucrose, the most abundant simple carbohydrates in coffee beans, were identified[1].
To explore an overview of all samples and to obtain general information on sample variances, unsupervised multivariate analysis, PCA, was selected. In previous research, PCA score plot derived from 21 coffee bean displayed differentiation of two data groups based on their species, Arabica and Robusta. Due to large variance between Arabica and Robusta coffee, samples differentiation based on type of coffee (Kopi Luwak and regular coffee) could not be observed. PCA score plot showed data separation based on type of coffee, in which Kopi Luwak (digested by animal) and regular coffee (not digested) can be clearly separated only when the analyses were carried out independently for each coffee species originated from same cultivation area (data not shown).

Discriminant analysis to select candidates of discriminant marker

An overview of all data samples was provided by unsupervised analysis, PCA. However, the detail information regarding contributed compounds for the data differentiation of Kopi Luwak and regular coffee remain unclear. Therefore, coffee bean data sets were subjected to supervised discriminant analysis, OPLS-DA. For analysis having two or more classes, OPLS-DA is the most suitable analysis platform to isolate and select differentiation marker. The analysis provides visualization of the covariance and correlation between compounds and the constructed discrimination model. Compounds that are highly contribute to the model and their reliability may possess potentially biochemically interesting characteristic, thus it can be selected as biomarker candidates[10]. The S-plot featured in OPLS-DA visualizes covariance (contribution or magnitude) and correlation (reliability) between metabolites and modelled clas[1]. Potential candidates of discriminant markers for authenticity assessment of Kopi Luwak can be selected via S-plot by setting the cut off for covariance, *p[1]*, and correlation value, *p[corr]*, > |0.2|. S-plot of coffee data sets was shown in Fig. 1 In addition to cut off value, compounds for candidates of discriminant markers were selected by its VIP (variable importance in projection) value. Large VIP (> 1) values are more relevant for model construction.
OPLS-DA score plot of Arabica coffee data sets were shown in Fig. 1A. Discrimination between Kopi Luwak and regular coffee was obtained. The model explained the goodness-of-fit parameter (R^2) and the predictability parameter (Q^2), 0.965 and 0.892, respectively (0.936 and 0.829 after cross validation). This model was considered good based on the above criteria. Significant compounds for samples separation were plotted at the top and bottom of S-plot. Interestingly, compounds uncorrelated with Kopi Luwak were quinic acid, caffeine and caffeic acid. As for predictive compounds (component to correlate with Kopi Luwak) in above cut-off value, consist of citric acid, malic acid and glycolic acid. OPLS-DA score plot of Robusta coffee data sets (Fig. 1B) were explained by R^2 and Q^2, 0.957 and 0.818, respectively (0.957 and 0.833 after cross validation). Inositol and caffeine (p, p[corr] exceeded cut-off value), were selected as discriminant marker candidates from OPLS-DA.
We employed significance analysis of microarrays/metabolites (SAM) to select significant compounds for discriminant marker as comparison to OPLS-DA. SAM indicates a total of 12 compounds considered as significant in Arabica coffee data set (Fig. 2). Of those compounds,

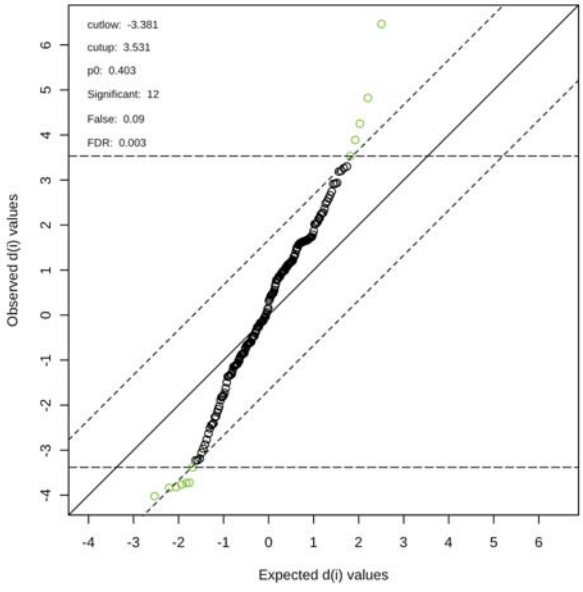


Fig. 2 SAM analysis derived from Kopi Luwak (Arabica) and all regular coffee data sets of experimental coffee set. Twelve detected peaks (highlighted by green color) were considered significant. The list including glycolic acid, malic acid, citric acid, quinic acid and unidentified peaks.

Table 3 Tentative identification

Name	Retention time (tR) in minute	RSD of tR (n=3)	Identification [a] method
Pyruvate	4.69	0.094	a
Lactic acid	4.79	0.086	a, b
Glycolic acid	4.99	0.079	a, b, c
Oxalate	5.62	0.074	a
Succinic acid	7.54	0.615	a, b
Glyceric acid	7.65	0.051	a, b
Fumaric acid	7.89	0.051	a, b
Malic acid	9.09	0.039	a, b, c
Aspartic acid	9.37	2.16	a, b
Pyroglutamic acid	9.47	0.04	a, b, c
Pyrogallol	9.59	0.041	a
Arabinose	10.47	0.035	a, b
Arabitol	10.9	0.034	a, b
Shikimic acid	11.57	0.032	a
Citric acid	11.68	0.031	a, b, c
Quinic acid	11.93	0.032	a
Fructose	12	0.031	a, b
Glucose	12.15	0.041	a, b
Caffeine	12.18	0.038	a, b, c
Inositol	13.5	0.028	a, b, c
Caffeic acid	13.85	0.025	a, b
Tryptophan	14.36	0.602	a, b
Octadecanoic acid	14.49	0.025	a, b
Sucrose	16.49	0.023	a, b
Melibiose	17.33	0.025	a, b
Chlorogenic acid	18.76	0.025	a

[a] (a = in-house library, b = NIST library, c = authentic standard)

citric acid, glycolic acid, and malic acid, identified from OPLS-DA, and other unidentified peaks were included. In Robusta coffee data set, 9 significant compounds were found. The result showed that inositol, caffeine and pyroglutamic acid and six unidentified peaks possessed high significance in SAM[1]. Candidates of discriminant marker for authentication assessment for Arabica and Robusta coffee were listed in Table 3. Marker candidates were selected for those met with significant criteria in both OPLS-DA and SAM. Discriminant markers were then selected independently for Arabica and Robusta coffee due to great variances among coffee species.

Validation of the applicability of discriminant marker for authenticity assessment

To verify the applicability of selected marker candidates, we have conducted analysis of validation set including authentic Kopi Luwak, commercial Kopi Luwak, commercial regular coffee, adulteration coffee, and blend coffee. Processing of authentic Kopi Luwak was controlled comprehensively. To provide unbiased analysis, the rest of samples were purchased commercially. In general, from harvest to pre-roasting, samples labelled as "commercial Kopi Luwak" and "commercial regular coffee", were processed in similar way to produce Kopi Luwak and regular coffee in the experimental set, respectively. However during roasting, their respective providers often to apply different parameter. Imitation coffee was processed to reduce its acidity in order to obtain characteristic close to Kopi Luwak[4]. Commercial regular coffees were selected from different cultivation area. To examine feasibility the selected marker to differentiate pure and blend coffee, we mixed two commercial Kopi Luwak, Golden Kopi Luwak and Kopi Luwak Wahana, and commercial regular coffee (Kopi Wahana) with ratio 50:50 (w/w), respectively, to compare applicability of discriminant marker to perform when blending was carried out by coffee beans from same and different cultivation area.
By employing all detected peaks to PCA, samples were populated into four clusters with the largest variance correspond to imitation coffee as it clearly separated from others (data not shown). Next, we projected six marker candidates as inclusion list into PCA to obtain overview of applicability of marker candidates for samples differentiation. Similar to prior strategy, separation of those four groups coffee was observed. The PCA was explained by 59.5% and 20.9% variance in PC1 and PC2, respectively (Fig. 3). Imitation coffee was populated separately by PC1. Separation was likely due to attempt by producer in order to obtain close characteristic of Kopi Luwak. In PC2, commercial Kopi Luwak, blend coffee, and commercial regular coffee could be differentiated. Both authentic and commercial Kopi Luwak was clustered within near area. In spite of originated from different country and processed with different parameter, commercial regular coffee were clustered in near region, suggesting these factors have least significance for data separation. From the loading plot information, citric acid, malic acid and inositol exhibited high contribution value for Kopi Luwak data sets. Interestingly, these three marker candidates also showed highest VIP value for constructing discriminant model (Table 1).
To display the applicability of selected discriminant markers to differentiate samples in validation set, box plot was constructed using relative peak intensity of citric acid, malic acid and inositol. Box plot of malic acid and citric acid were able to differentiate commercial Kopi Luwak (Kopi Luwak Wahana), blend coffee, commercial regular coffee (Kopi Wahana) and adulteration coffee. However, box plot of inositol failed to differentiate these samples. Hence, we selected double marker ratio, inositol/pyroglutamic acid (Fig. 4). Pyroglutamic acid was selected for having lowest contribution for separation of Kopi Luwak and regular coffee. We confirmed ratio of blend coffee by quantitate the constituent of discriminant marker. Analytical parameter for quantitation was shown in Table 3. All authentic standards exhibited good linearity of 0.99 or more for at least seven points in the applied concentration range. To examine the validity of quantitation, limit of detection (LOD) and limit of quantitation for each discriminant marker were also determined. The amount of six discriminant marker candidates in coffee sample was quantitated higher than the LOD and LOQ of authentic standards. Concentration of selected marker (malic acid, citric acid, and inositol/pyroglutamic acid) in all blend samples was in range of 48.5 ± 0.02 to 52.3 ± 0.75% (Fig. 4). The result corresponded well with box plot of peak intensity for each discriminant marker. We confirmed feasibility of the proposed strategy for robust authentication of Kopi Luwak in pure and blend coffee for ratio of 50:50.

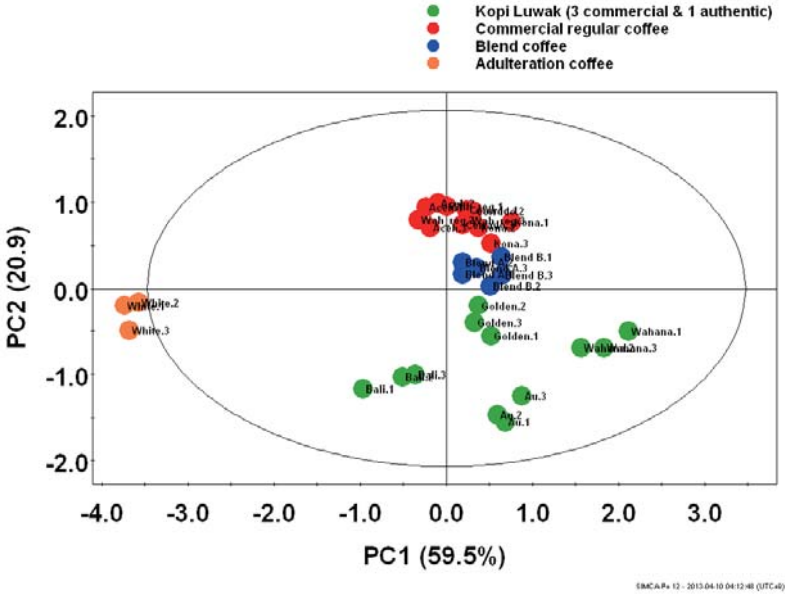


Fig. 3 PCA score plot of validation coffee set. Separation with adulteration coffee was obtained in PC1, while commercial Kopi Luwak, blend coffee and commercial regular coffee can be differentiated in PC2

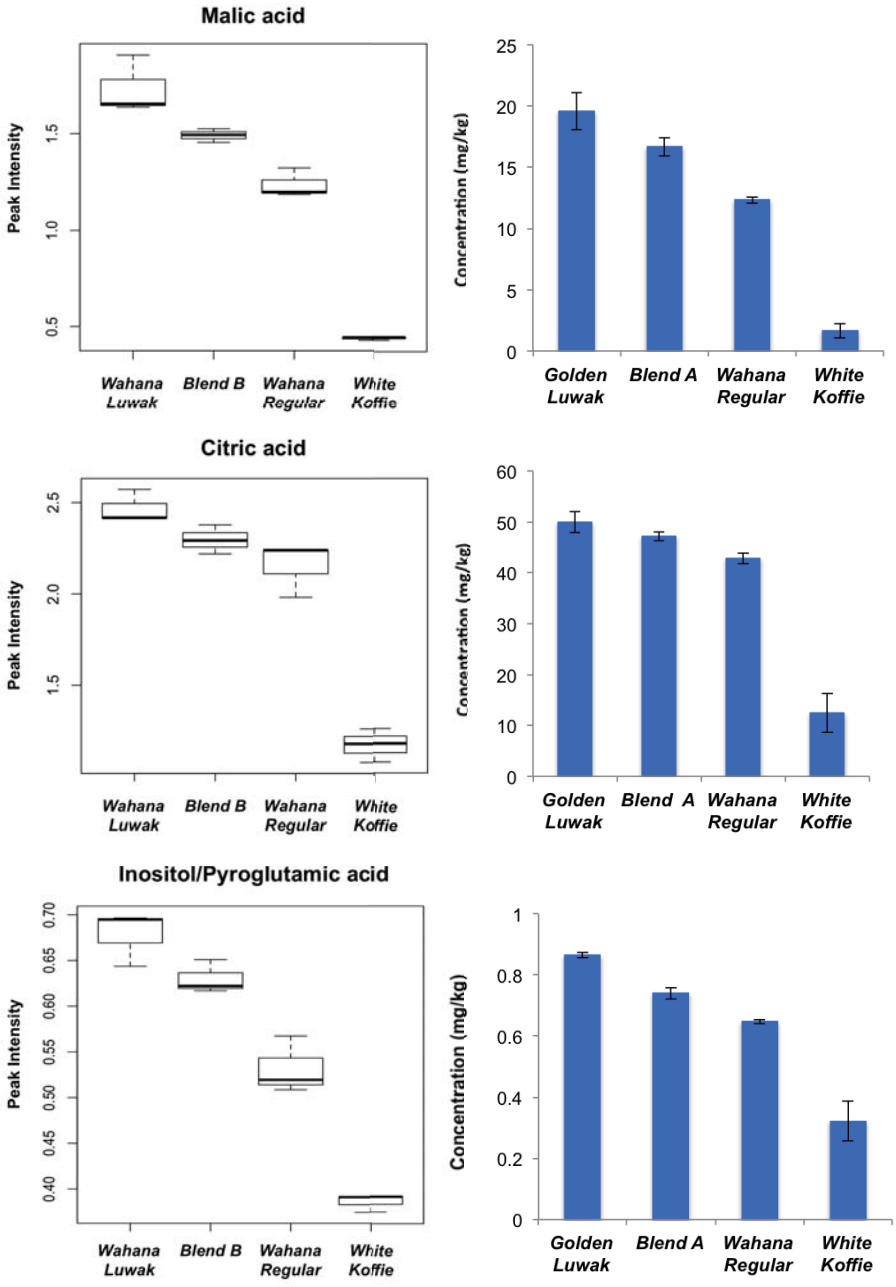


Fig. 4 Box plots of peak intensity and concentration for selected discriminant markers of validation coffee sets (for Blend coffee A)

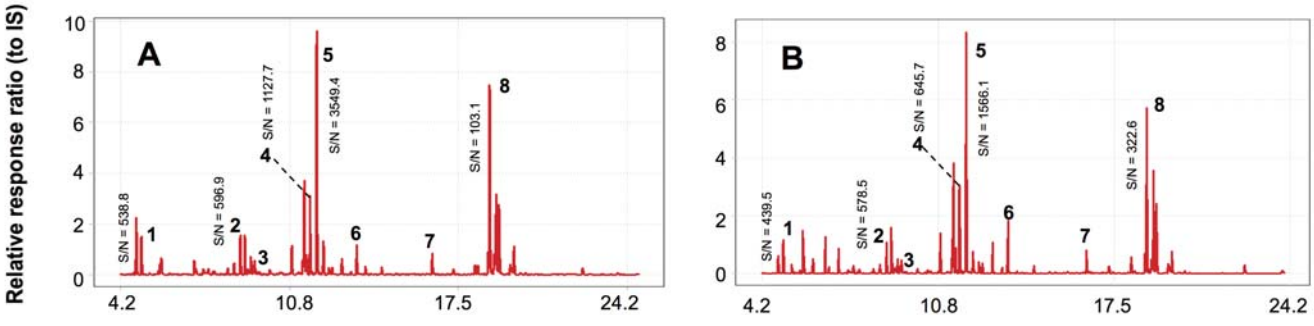


Fig. 5 Gas chromatograms of representative coffee bean extracts obtained from (A) GC/FID; and (B) GC/MS analysis. Both analyses used same column, CP-SIL 8 CB low bleed. Peak tentative identification: (1) glycolic acid, (2) malic acid, (3) pyroglutamic acid, (4) citric acid, (5) quinic acid, (6) inositol, (7) sucrose, and (8) chlorogenic acid. S/N, signal-to-noise ratio.

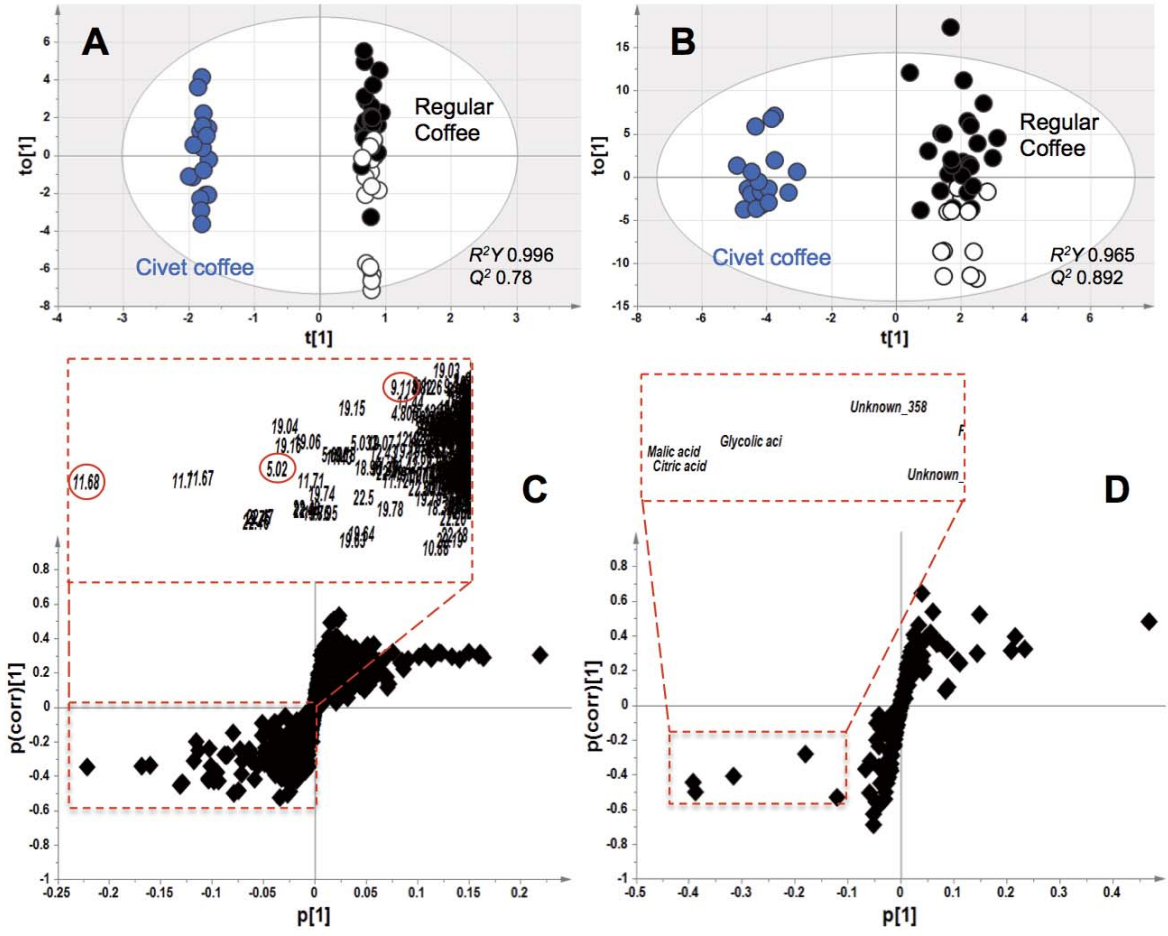


Fig. 6 OPLS-DA score plots and S-plots based on (A,C) GC/FID and (B, D) GC/MS chromatograms of 20 coffee bean extracts. The S-plot displayed the covariance *p* against correlation *p*(corr) of the variables to the model class designation. The closed diamonds represent each variable (detected peak) used for model construction; identities of variables with high reliability to civet coffee are given in the inset figure.

GC/FID-based metabolite fingerprinting for Kopi Luwak authentication

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 as shown in Fig. 5, 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[16]. 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[17], the retention time shift reported here, albeit not severe, may be due to experimental variation between analytical instruments.

Although the overall chromatographic profiles between GC/FID and GC/MS were similar, 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 sample[2]. 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. 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.

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[2]. Based on the results of multivariate analyses, 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 (Fig. 6).

Finally, 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. The validation experiment was successful to verify the feasibility of employing the significant variables obtained from GC/FID for practical use in authentication of Kopi Luwak.

Conclusion

Our findings highlighted the utility of metabolic profiling using GC/MS combined with multivariate analysis for selection of discriminant marker for authenticity assessment of valuable agricultural products. Once the discriminant markers have been identified, we propose the utility of GC/FID coupled with metabolite fingerprinting as a good complementary and cost effective analysis platform for quality assessment of civet coffee. 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.

References

[1] 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*. **2013**, *61*, 7994–8001.

[2] Jumhawan, U.; Putri, S. P.; Yusianto, Bamba, T.; and Fukusaki, E. Application of GC/FID-based metabolite fingerprinting for authentication of Asian palm civet coffee (Kopi Luwak), *J. Biosci. Bioeng*. **2015**. Available online 23 Apr 2015.

[3] http://www.nytimes.com/2010/04/18/world/asia/18civetcoffee.html (Last access: April 3, 2013)

[4] Joshi, A. R.; Smith, J. L. D.; Cuthbert, F. Influence of food distribution and predation pressure on spacing behavior in palm civets. *J. Mammal*. **1995**, *76*, 1205-1212.

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

[6] Blow, N. Metabolomics: biochemistry's new look. *Nature*. **2008**, *455*, 697-700.

[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] 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.

[9] 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 multi-parametric diagnosis. *J. Prot. Res*. **2006**, 5, 1407–1414.

[10] 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.

[11] Tusher, V.G.; Tibshirani, R.; Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl Acad. Sci. USA*. **2001**, *98*, 5116-5121.

[12] 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.

[13] 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.

[14] http://www.shimadzu.com/an/gcms/qpultra.html (Last access: May 7, 2015)

[15] 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.

[16] Fiehn, O.; Metabolomics--the link between genotypes and phenotypes, Plant Mol. Biol. **2002**, 155–71

[17] 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*. **2012**, *84*, 10768–76.