DISCRIMINATION OF GEOGRAPHICAL ORIGINS OF COFFEES BY PAPER SPRAY MASS SPECTROMETRY

Miss Rujirat Pumbua

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Petrochemistry and Polymer Science
Field of Study of Petrochemistry and Polymer Science
FACULTY OF SCIENCE
Chulalongkorn University
Academic Year 2022

การจำแนกแหล่งที่มาทางภูมิศาสตร์ของกาแฟด้วยเปเปอร์สเปรย์แมสสเปกโตรเมทรี

น.ส.รุจิรัตน์ พุ่มบัว

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาปิโตรเคมีและวิทยาศาสตร์พอลิเมอร์ สาขาวิชาปิโตรเคมีและวิทยาศาสตร์พอลิเมอร์
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565

Thesis Title	DISCRIMINATION OF GEOGRAPHICAL ORIGINS OF COFFEES BY PAPER SPRAY MASS SPECTROMETRY
By	Miss Rujirat Pumbua
Field of Study	Petrochemistry and Polymer Science
Thesis Advisor	Associate Professor THANIT PRANEENARARAT, Ph.D.

Accepted by the FACULTY OF SCIENCE, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

..... Dean of the FACULTY OF SCIENCE
(Professor POLKIT SANGVANICH, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Professor MONGKOL SUKWATTANASINITT, Ph.D.)
..... Thesis Advisor
(Associate Professor THANIT PRANEENARARAT, Ph.D.)
..... Examiner
(Professor NISANART TRAIPHOL, Ph.D.)
..... External Examiner
(Assistant Professor Anyanee Kamkaew, Ph.D.)

รุจิรัตน์ พุ่มบัว : การจำแนกแหล่งที่มาทางภูมิศาสตร์ของกาแฟด้วยเปเปอร์สเปรย์แมสสเปกโตรเมทรี. (DISCRIMINATION OF GEOGRAPHICAL ORIGINS OF COFFEES BY PAPER SPRAY MASS SPECTROMETRY) อ.ที่ปรึกษาหลัก : รศ. ดร. ธนิษฐ์ ปราณีนรารัตน์

กาแฟคือหนึ่งในสินค้าทางการเกษตรที่สำคัญที่สุด โดยความต้องการกาแฟพิเศษและกาแฟจากแหล่งปลูกเดียวนั้นเพิ่มขึ้นอย่างมีนัยสำคัญ ซึ่งวิธีการที่สามารถแยกแหล่งกำเนิดของกาแฟได้อย่างมีประสิทธิผลจึงเป็นประโยชน์อย่างมาก งานวิจัยนี้รายงานการใช้เทคนิคเปเปอร์สเปรย์แมสสเปกโทรเมตรี (PS-MS) เป็นเครื่องมือในการเก็บข้อมูลองค์ประกอบทางเคมีของกาแฟเพื่อนำไปสู่การจำแนกแหล่งกำเนิดของกาแฟ โดยข้อมูลที่ได้นั้นถูกนำไปประมวลผลต่อด้วยการวิเคราะห์องค์ประกอบหลัก (PCA) และการวิเคราะห์การจำแนกประเภทเชิงเส้น (LDA) และให้ผลลัพธ์ที่สามารถจำแนกแหล่งกำเนิดได้อย่างประสบความสำเร็จด้วยความแม่นยำในการทำนายที่ 80-100% โดยการศึกษาการจำแนกดังกล่าว ประกอบไปด้วย การแยกระหว่างกาแฟอาราบิกากับโรบัสตา การจำแนกกาแฟอาราบิกาจากแหล่งกำเนิดทางภูมิศาสตร์ที่แตกต่างกัน การจำแนกกาแฟโรบัสตาจากแหล่งกำเนิดทางภูมิศาสตร์ที่แตกต่างกัน และแยกแหล่งกำเนิดทางภูมิศาสตร์ของกาแฟอาราบิกาตามรายจังหวัดในประเทศไทย นอกจากนี้ ยังได้มีการใช้เทคนิคเฮดสเปซโซลิดเฟสไมโครเอกซ์แทรกชัน (HS-SPME) ร่วมกับแก๊สโครมาโทกราฟี-แมสสเปกโทเมทรี (GC-MS) เพื่อเป็นวิธีทางเลือกในการเก็บข้อมูลทางเคมี ซึ่งพบว่าให้ประสิทธิภาพที่ดีในการจำแนกประเภทของกาแฟด้วยเช่นกัน กล่าวโดยสรุป ภาพรวมของกระบวนการทั้งสองวิธีนั้นเรียบง่าย มีการเตรียมตัวอย่างก่อนการวิเคราะห์เพียงเล็กน้อย แต่ยังมีประสิทธิผลในการจำแนกแหล่งกำเนิดของกาแฟได้เป็นอย่างดี

สาขาวิชา	ปิโตรเคมีและวิทยาศาสตร์พอลิเมอร์	ลายมือชื่อนิสิต .....
ปีการศึกษา	2565	ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6270154423 : MAJOR PETROCHEMISTRY AND POLYMER SCIENCE

KEYWORD: Coffee, Headspace Solid Phase Microextraction, Gas Chromatography, Chemometrics, Geographical Indication, Paper Spray – Mass Spectrometry

Rujirat Pumbua : DISCRIMINATION OF GEOGRAPHICAL ORIGINS OF COFFEES BY PAPER SPRAY MASS SPECTROMETRY. Advisor: Assoc. Prof. THANIT PRANEENARARAT, Ph.D.

Coffee is one of the most important agricultural commodities, and the demand for specialty coffee and single-origin coffees is also increasing significantly. Methods that can effectively differentiate the origins of coffee is tremendously beneficial. In this work, paper spray – mass spectrometry (PS-MS) was used as a tool for collecting chemical profiles of coffees for differentiating their origins. The acquired data was then processed by principal component analysis (PCA) and linear discriminant analysis (LDA), resulting in successful identification of the origins of coffee with prediction accuracies of 80-100%. Studies with the aforementioned performance include the discrimination of Arabica vs Robusta, Arabica coffees from different geographical origins, Robusta coffees from different geographical origins, and geographical origins of Arabica coffees within each province in Thailand. Furthermore, headspace solid-phase microextraction (HS-SPME) in combination with gas chromatography – mass spectrometry (GC-MS) was also evaluated as an alternative method for collection of chemical data. This method was found to also provide good performance for the classification of coffees. In brief, the overall workflow in both cases is simple, with little sample pre-treatment, but yet effective for distinguishing origins of coffee.

Field of Study:	Petrochemistry and Polymer Science	Student's Signature .....
Academic Year:	2022	Advisor's Signature .....

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere thanks to my thesis advisor Associate Professor Dr.Thanit Praneenararat of the Department of Chemistry, Faculty of Science, Chulalongkorn University for his kind support, guidance, and constant encouragement throughout the course of this thesis.

Besides my advisor, I would like to thank Associate Professor Dr. Kanet Wongravee (Sensor Research Unit (SRU), Department of Chemistry, Faculty of Science, Chulalongkorn University) for his support in chemometric analysis.

I wish to sincerely thank all my committee members, Professor Dr.Mongkol Sukwattanasinitt, Associate Professor Dr.Nisanart Traiphol, and Assistant Professor Dr.Anyanee Kamkaew for reviewing my thesis and giving effective suggestions and comments that further improve this thesis.

Moreover, I would like to give my thanks go to all TP lab members for their help, advice, and encouragement.

Lastly, I would like to express my thankfulness to my family who always understand and support. This accomplishment would not have been possible without them.

Rujirat Pumbua

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

$CDCl_3$	Deuterated chloroform
EtOAc	Ethyl acetate
GC-MS	Gas chromatography – mass spectrometry
HS-SPME	Headspace – solid phase microextraction
IS	Internal standard
LDA	Linear discriminant analysis
LRI	Linear retention index
MeOH	Methanol
MS	Mass spectrometry
NaOH	Sodium hydroxide
NMR	Nuclear magnetic resonance
PCA	Principal component analysis
PS-MS	Paper spray – mass spectrometry
RPA	Relative peak area
RT	Retention Time

CHAPTER I
INTRODUCTION

1.1 General information on coffee

Coffee is one of the world's most traded products with considerable economic importance on the global market. Coffee has a varieties of chemical compositions such as sugar, amino acids, alkaloids, organic acids, and phenolic compounds that contribute to flavor and other quality of coffee. Importantly, coffees from different origins can have distinct flavor profiles due to the difference in climate, soil, and other origin-related factors.[1-3] With the growing interest in specialty coffee and single-origin coffees, accurate representation of origins of coffee has become a topic of interest. This can also link to geographical indication, a type of intellectual properties that is useful in protecting and sustaining the value of local products.[4-5] Hence, methods that can efficiently distinguish origins of coffee are highly desirable, and analytical instrumentations that can unbiasedly reveal chemical profiles of food samples are a method of choice for such differentiation study.

1.2 Previous study of differentiation of origins of coffees

Several analytical techniques, in combination with chemometrics, have been utilized for coffee studies. For example, nuclear magnetic resonance (NMR) spectroscopy was used to accurately discriminate roasted Colombian coffee from other types of coffee.[6] Dicaffeoylquinic acids analysis by ultra-performance high-pressure liquid chromatography coupled with mass spectrometry (UPLC-MS) has been used to distinguish samples grown in subregions of Ethiopia.[7] Phenolic compounds have also been studied by high-performance liquid chromatography (HPLC) to separate different chemical species based on their geographical origin.[8] Gas chromatography-mass spectrometry (GC-MS) has been widely applied in the characterization of the geographical origin of coffee based on the analysis of volatile and semi-volatile compounds. Coffee samples were also prepared by a solvent extraction method followed by direct liquid injection into the GC column[9-10],

resulting in an additional process prior to sample analysis. However, some volatile and semi-volatile compounds may be evaporated during the coffee sample preparation process.

1.3 Paper spray – mass spectrometry (PS-MS)

Ambient ionization methods are a group of techniques in which ionization of the analytes from the sample occurs under atmospheric pressure and at room temperature, such as paper spray–mass spectrometry (PS-MS) (**Figure 1**).[11] In this method, it requires a high-voltage power supply to generate electrospray droplets. The ions are then passed into the mass spectrometer. The advantages of this technique are rapidness, real-time capability, high throughputness, and relatively little sample preparation. Consequently, PS-MS has been used to detect various types of compounds such as food additives[12-13] and drugs[11,14-16] in various complex matrixes. Therefore, it is expected that this technique can be used for the analysis of coffee extracts and simplify the differentiation of the geographical origins of coffees.

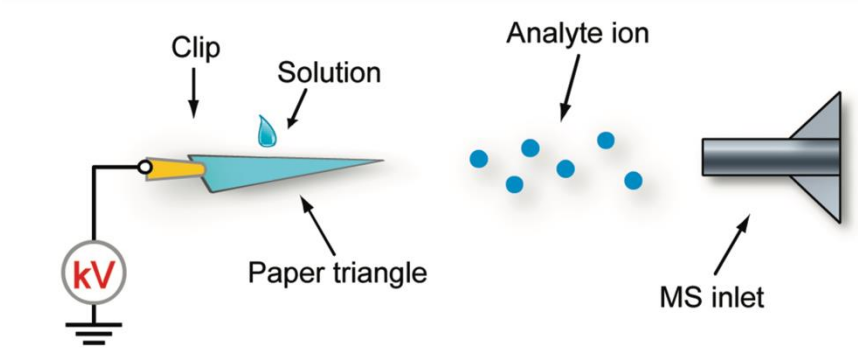


Figure 1 Illustration of paper spray mass spectrometry[11]

1.4 Gas chromatography – mass spectrometry (GC–MS)

The application of headspace solid phase microextraction (HS-SPME) in combination with gas chromatography-mass spectrometry (GC-MS) (**Figure 2**) has been widely recognized to extract volatile compounds from coffee samples[17-18] because it is a non-destructive method in the determination of volatile and semi-volatile compounds. Furthermore, it is simple, fast, and free from pre-treatment

processes. In this method, the SPME fiber is used to absorb volatile compounds from the sample, and later release them into the GC, hence acting as a pre-concentration method. This is usually done by placing SPME fiber into the headspace of a sample. After the extraction, the fiber was desorbed at the injector liner and the carrier gas transfers analytes to the GC column for analysis. MS, if available, helps identify volatile compounds by comparing their mass spectra with MS fingerprint database.[19-21]

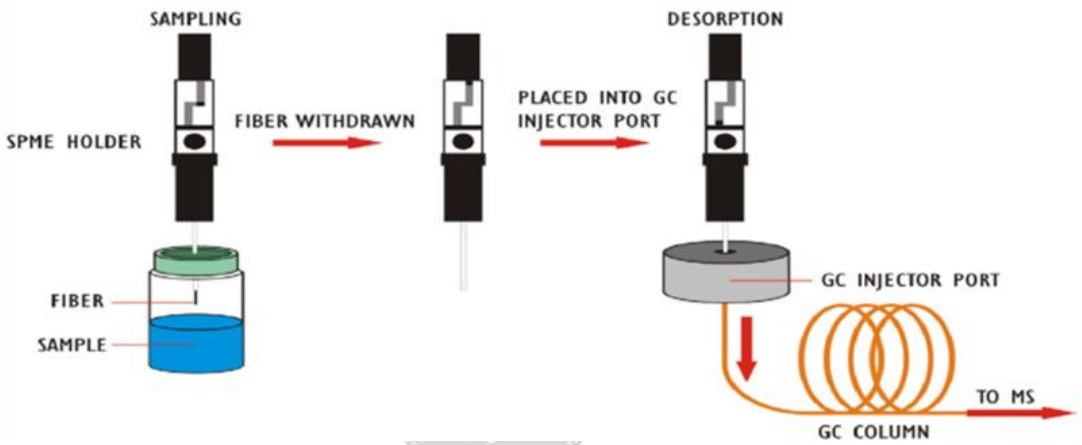


Figure 2 Diagram of analysis with solid phase microextaction-gas chromatography-mass spectrometry.[20]

Coffee analysis and classification can undoubtedly benefit from this technique in combination with chemometrics. For example, SPME/GC-MS was used to analyze the aroma profile of 10 single-origin Arabica coffees originating from eight different growing locations.[17] In this work, a total of 138 volatile compounds were tentatively identified in all samples and were successfully profiled using SPME/GC-MS combined with chemometrics.

Ongo *et al.* reported 47 volatile metabolites of Philippine Arabica and Robusta coffees (**Figure 3**) in different Philippine regions using SPME/GC-MS.[18] PCA and cluster analysis displayed a good discrimination between Arabica and Robusta coffee samples. PCA score plot is shown in **Figure 4**. The classification of volatile metabolites of Philippine Arabica and Robusta coffee was successfully carried out

using an analytical approach to outline specific volatile fingerprints through multivariate statistical tools.

Table 1
Volatile compounds detected in Philippine roasted coffee beans and their retention times (t_R).

#	t_R (min)	Volatiles (IUPAC name)	Synonyms	#	t_R (min)	Volatiles (IUPAC name)	Synonyms
1	2.14	2-Methylfuran		25	16.29	1-Pyridin-2-ylethanone	2-Acetylpyridine
2	6.20	Pyridine		26	16.42	2-(Furan-2-ylmethyl)furan	2-Furfurylfuran
3	6.30	Dodecane		27	16.60	5-Methyl-6,7-dihydro-5*H*-cyclopenta[b]pyrazine	
4	6.76	Pyrazine		28	16.74	1-Methylpyrrole-2-carbaldehyde	
5	7.21	*Unknown 1*		29	16.97	Oxolan-2-one	γ-Butyrolactone
6	8.05	2-Methylpyrazine		30	17.54	Furan-2-ylmethanol	Furfuryl alcohol
7	9.37	2,5-Dimethylpyrazine		31	18.38	1-(6-Methylpyrazin-2-yl)ethanone	2-Acetyl-6-methylpyrazin
8	9.51	2,6-Dimethylpyrazine		32	18.74	*Unknown 3*	
9	9.66	2-Ethylpyrazine		33	18.98	*Unknown 4*	
10	9.98	2,3-Dimethylpyrazine		34	20.25	*Unknown 5*	
11	10.88	2-Ethyl-6-methylpyrazine		35	20.64	*Unknown 6*	
12	11.04	2-Ethyl-5-methylpyrazine		36	20.71	*Unknown 7*	
13	11.37	2,3,5-Trimethylpyrazine		37	21.27	3-Methylcyclopentane-1,2-dione	
14	12.35	3-Ethyl-2,5-dimethylpyrazine		38	21.32	1-(Furan-2-ylmethyl)pyrrole	Furfurylpyrrole
15	12.71	Acetic acid		39	21.98	2-Methoxyphenol	Guaiacol
16	12.74	2-Ethyl-3,5-dimethylpyrazine		40	24.13	3-Hydroxy-2-methylpyran-4-one	Maltol
17	12.94	*Unknown 2*		41	24.24	1-(1*H*-Pyrrol-2-yl)ethanone	2-Acetylpyrrole
18	12.99	Furan-2-carbaldehyde	Furfural	42	24.52	2-(Furan-2-ylmethoxymethyl)furan	Furfuryl ether
19	13.53	3,5-Diethyl-2-methylpyrazine		43	24.72	*Unknown 8*	
20	13.98	1-(Furan-2-yl)ethanone	2-Acetylfuran	44	24.95	Phenol	
21	14.18	1*H*-Pyrrole	Pyrrole	45	25.34	1*H*-Pyrrole-2-carbaldehyde	2-Formylpyrrole
22	14.66	Acetic acid;furan-2-ylmethanol	Furfuryl acetate	46	25.41	4-Ethyl-2-methoxyphenol	4-Ethylguaiacol
23	15.64	5-Methyl-2-furancarbaldehyde	5-Methylfurfural	47	28.60	1-(3-Methoxyphenyl)ethanone	3-Acetylanisole
24	16.12	2-Prop-1-en-2-ylpyrazine	Isopropenylpyrazine				

Figure 3 Volatile compounds detected in Philippine roasted coffee beans and their retention times.[18]

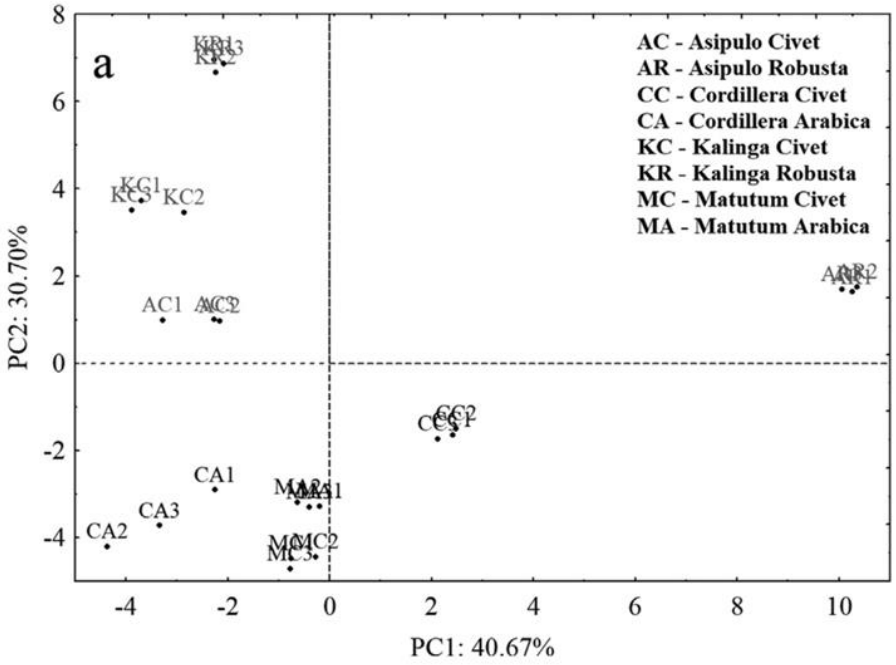


Figure 4 PCA score plot (PC1 vs. PC2) of the complete samples set.[18]

1.5 Chemometrics

The differentiation of geographical origins by analyzing chemical constituents can create a large set of chemical data. Hence, a tool is required to efficiently organize and distinguish data between samples. Chemometrics is the discipline concerned with the application of mathematical and statistical methods to improve chemical measurement and to extract useful chemical information obtained from complex chromatographic or spectroscopic profiles. Chemometrics consists of a variety of techniques, among which interesting techniques are principal component analysis (PCA) and linear discriminant analysis (LDA).[22-23]

1.5.1 Principal component analysis (PCA)

PCA is a multivariate technique for the analysis of data with the observations described by several inter-correlated dependent variables. PCA was used to reduce a large number of variables in multidimensional data sets to a smaller number of dimensions called principal components (PCs). From the PCA plot (**Figure 5**), PC1 is representing the greatest variance of the data, and PC2 accounts for the second greatest variance that is orthogonal to the PC1. Other PCs indicate smaller variability of data. PCA plot can be applied to visualize the clusters of samples, which share common influences.

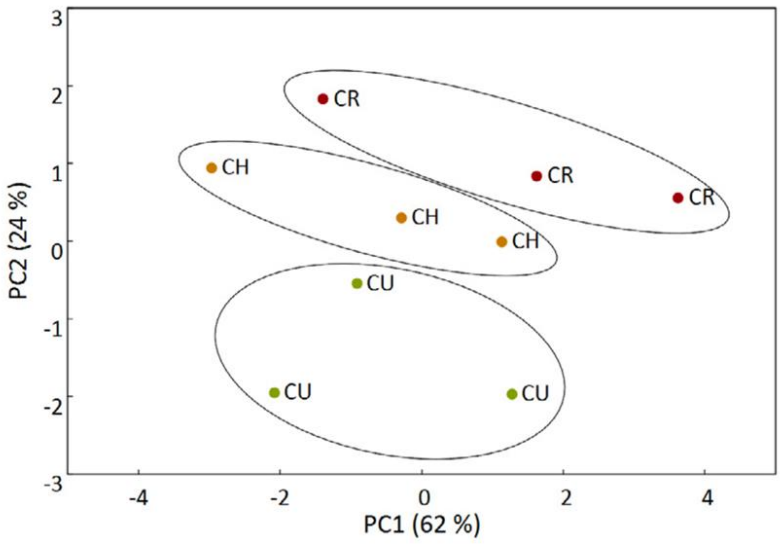


Figure 5 An example of PC score plot.[24]

1.5.2 Linear discriminant analysis (LDA)

LDA is a well-known data analysis technique to be used for class prediction. This was done by assigning linear classifiers or boundaries between classes using linear discriminant function in order to define the directions in which the separation of all determined classes. The classification was evaluated by the leave-one-out cross validation (LOOCV) approach. In the validation step, one sample is assigned as the test set, while others are in the training set. This is performed by creating a model from the training set, followed by predicting the test set. After repeating the procedure for all samples, the predictive ability is the percentage of samples belonging to the testing set that are classified correctly using the created model.

1.6 Purpose of this study

In this study, coffee samples originating from different origins were analyzed using paper spray–mass spectrometry (PS-MS) as the analytical method. The study aimed to differentiate coffees based on sub-studies including 1) the types of coffees (Arabica vs Robusta), 2) geographical origins of Arabica coffees, 3) geographical origins of Robusta coffees, and 4) geographical origins of Arabica coffees within the same province of Thailand. Furthermore, a separate study using headspace solid phase microextraction (HS-SPME) in combination with gas chromatography-mass spectrometry (GC-MS) to discriminate coffee sources was also conducted to showcase an alternative method for the classification of coffee origins. Overall, chemometric analyses were able to effectively classify coffees in both cases.

CHAPTER II

MATERIALS AND METHODS

2.1 Chemicals

Chemical reagents and solvents were purchased from Sigma-Aldrich, Merck, TCI Chemicals, and RCI Labscan. All chemicals are listed below:

- Theobromine
- Bromoethane
- Acetone
- Chloroform
- Ethyl acetate (EtOAc)
- Methanol (MeOH)
- Sodium hydroxide (NaOH)

2.2 Sources of Coffee Samples

The coffee bean samples were acquired from different regions. All samples are listed in **Table 1** and physical appearances are shown in **Figure 6**. All coffee samples were light roasted with either natural, honey, or washed processing. Samples were stored in a freezer at -20 °C until used.

Table 1 List of coffee samples from different origins.

ID	Name	Type	Processing	Location
C1	Doi Pha Hee	Arabica	Washed	Chiang Rai, Thailand
C2	Doi Pang Khon	Arabica	Washed	Chiang Rai, Thailand
C3	Robusta Chiang Rai	Robusta	Washed	Chiang Rai, Thailand
C4	Doi Chang	Arabica	Washed	Chiang Rai, Thailand
C5	Mae Ton Luang	Arabica	Natural	Chiang Mai, Thailand
C6	Om koi	Arabica	Honey	Chiang Mai, Thailand
C7	Thep Sadet	Arabica	Natural	Chiang Mai, Thailand
C8	Manipruek	Arabica	Natural	Nan, Thailand

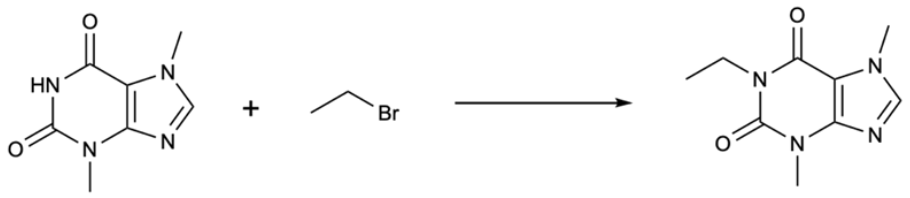
Table 1 (continued).

ID	Name	Type	Processing	Location
C9	Doi Sakad	Arabica	Washed	Nan, Thailand
C10	Huai Wai	Arabica	Natural	Mae Hong Son, Thailand
C11	Dulapur	Arabica	Natural	Mae Hong Son, Thailand
C12	Robusta Chumphon	Robusta	Washed	Chumphon, Thailand
C13	Robusta Krabi	Robusta	Washed	Krabi, Thailand
C14	Vietnam	Robusta	Washed	Vietnam
C15	Ethiopia	Arabica	Washed	Ethiopia
C16	Brazil	Arabica	Washed	Brazil
C17	Colombia	Arabica	Washed	Colombia
C18	Guatemala	Arabica	Washed	Guatemala

Figure 6 Images of representative coffee bean samples as outlined in Table 1.

2.3 Paper Spray Mass Spectrometry Experiments

2.3.1 Synthesis and Characterization of 1-ethyltheobromine (the internal standard, IS)



theobromine bromoethane 1-ethyltheobromine

A solution of theobromine (502.9 mg) in 1:1 acetone:water (25 mL) was added successively with 25 mL of 0.5 M NaOH solution and 250 μL of bromoethane. Then, the solution was added with two additional aliquots of bromoethane (250 μL) after 1 and 2 hours. The stirring was continued for 2 days at room temperature, the solvent was evaporated by a rotary evaporator and then extracted with water (50 mL). Finally, the resulting aqueous solution was extracted with chloroform (50 mL, 3x). The combined organic phase was evaporated by a rotary evaporator. The crude product was purified by column chromatography using 80:20 ethyl acetate: methanol to afford as a white solid in 63% yield.; ^{1}H NMR (500 MHz, $CDCl_3$) was acquired from a JEOL JNM-ECZ500R/S1 NMR spectrometer to give the following chemical shifts : δ 7.50 (s), 4.08 (q, *J* = 7.1 Hz), 3.99 (s), 3.58 (s), 1.25 (t, *J* = 7.0 Hz). The exact mass of the product was determined by high-resolution mass spectrometry (HRMS) on JEOL AccuTOF^{TM}-DART Mass Spectrometer (calcd for $[C_9H_{13}N_4O_2]^+$: 209.1033, observe $[M+H]^+$: 209.1035).

2.3.2 Triangular paper preparation

Whatman 1Chr chromatography paper was cut into triangular papers with the dimensions of 6 mm base width and 12 mm height. The obtained paper pieces were washed with deionized water for 5 minutes, followed by 5 minutes of methanol with shaking at 160 rpm to remove adsorbed impurities, and then dried at room temperature.

2.3.3 Sample preparation

The coffee bean (1 g) was finely ground by a handheld grinder, and the ground coffee was extracted with 20 mL of boiling Milli-Q water for 5 minutes. Thereafter, the resulting solution (10 mL) was diluted with methanol (40 mL). One mL of 5,000 ppm IS was added to 39 mL of the aforementioned diluted solution. Thereafter, an aliquot of 500 μL of sample was transferred into each of five 0.6-mL microfuge tube. In each tube, a triangular paper was immersed into the solution for 30 minutes with shaking at 160 rpm. The paper was dried at room temperature for at least 30 minutes. The preparation of each source of coffee was repeated in triplicate, totaling 15 repetitions (3 sets of coffee extract × 5 pieces of paper).

2.3.4 Mass spectrometric analysis

The mass spectrometer in this experiment was a Thermo-scientific TSQ Quantum Ultra EMR triple quadrupole mass spectrometer. TSQ Tune software was used to control the mass spectrometer. All data processing was performed with the Xcalibur software. Mass spectra were acquired over the *m/z* 100-500 range. The triangular paper adsorbed with compounds from the coffee solution was held by a copper clip. The tip of the paper was placed approximately 0.5 cm from the inlet of the mass spectrometer. A 20 μL of methanol was deposited onto the paper to elute the analyte. The spraying condition was achieved using a DC power supply (3B scientific model 1019234) that was set at 3.5 kV which operated in the positive mode with a capillary temperature of 300 °C. The experimental setup is shown in **Figure 7**.

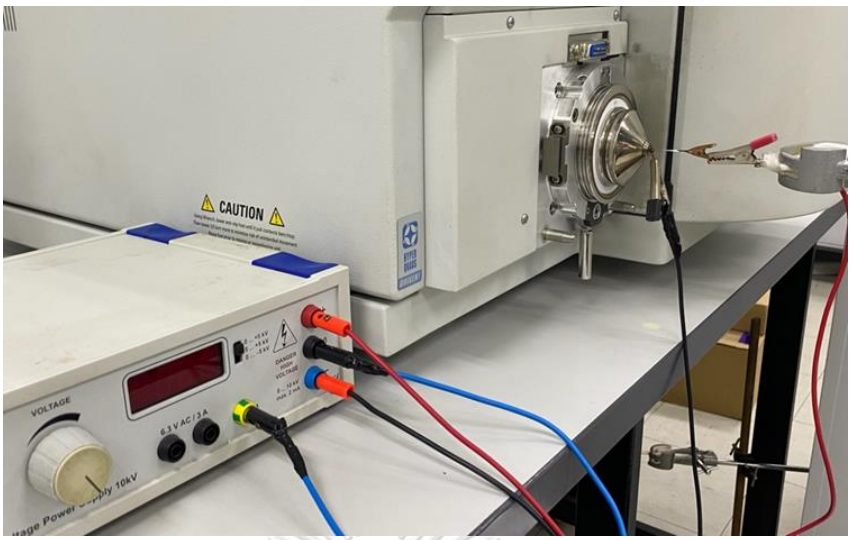


Figure 7 Experiment setup for paper spray mass spectrometry (PS-MS).

2.4 Gas chromatography-mass spectrometry Experiments

2.4.1 Sample preparation

Three grams of each source of the coffee bean was ground by a handheld grinder and placed in a 20 mL headspace vial, which was then closed tightly by an aluminum cap lined with a PTFE/silicone septum. The preparation of each coffee sample was repeated in triplicate.

2.4.2 Gas chromatography-mass spectrometric analysis

The sample vials were placed in the sample tray of a 7697A headspace autosampler connected to a 7890B GC system and a 5977B mass spectrometer (Agilent Technologies, USA). The GC-MS system was controlled by Agilent MassHunter GC-MS Acquisition software, version B.07.04. The sample vials were heated at 70 °C for 10 minutes to reach sample headspace equilibrium. The volatile compounds were extracted using a SPME fiber (DVB/C-WR/PDMS). The fiber was inserted into the vial and exposed to the headspace above the coffee sample for 30 minutes at 70 °C. After the extraction, the fiber was desorbed into the GC injection port for 5 minutes. GC-MS analysis was performed in split ratio mode at 10:1. The oven temperature program was set as 40 °C, held for 2 minutes, increased to 240 °C at a rate of 5 °C/min, and finally held for 5 minutes.

2.4.3 Identification of Volatile Compounds

Data acquisition and peak integration were identified using Agilent MassHunter Qualitative Analysis software (version B.07.02). The tentative identification of the volatile compounds in the coffee samples were identified according to a comparison of their mass spectra and the linear retention index (LRI) with those present in the NIST 14 database. The identification criteria were selected with a difference of ≤20 units between the calculated LRI and the database values. The LRI value was determined from *n*-alkane retention time data, which was obtained from the analysis of *n*-alkane standards (C8 to C20) using the same experimental conditions as with the one for real samples. LRI values were calculated according to[25]:

$$LRI = 100n + 100\left(\frac{RT_{(i)} - RT_{(n)}}{RT_{(n+1)} - RT_{(n)}}\right)$$

where $RT_{(i)}$ is the retention time of the unknown compound *i*

$RT_{(n)}$ is the retention time of the *n*-alkane eluted before the unknown compound *i*

$RT_{(n+1)}$ is the retention time of the *n*-alkane eluted after the unknown compound *i*

2.5 Chemometrics

2.5.1 Pre-processing of data

For PS-MS, the pre-processing was done as follows. First, the intensity of each *m/z* data point from the blank was subtracted from the data point of the sample. Second, all negative values after blank subtraction were changed to zero. Lastly, all data points were normalized by calculating the intensity of the sample data divided by the intensity of the IS data (*m/z* 209). It should be noted that the data was centering over all samples prior to Principal Component Analysis (PCA).

In the case of GC-MS, the relative peak area (%RPA) from GC-MS analysis were calculated prior to chemometric processing. The %RPA was calculated by the peak area divided by the total peak area of all identified peaks in each chromatogram.

2.5.2 Data analysis

In this study, chemometric analyses were performed using MATLAB software, version R2022a. PCA was employed to visualize the clusters of samples, using PCs with the maximum variances (PC1- PC3) or the PCs with the best discrimination power. To obtain the classification accuracy, Linear Discriminant Analysis (LDA) was used for class prediction. The LDA model was generated by creating a model boundary between classes using linear discrimination function. The distance between an unknown sample to the centroid of each class was calculated and the class of the unknown was estimated by the class given the smallest distance. In the case of the data containing the number of variables larger than the number of samples, the inverse of the variance-covariance matrix in LDA model could not be computed. To prevent this problem, the classification feature of the PCA-LDA were used to generate the classification model. The model was done by applying PCA to reduce the dimension of the data matrix into the desired PCs (1-20 PCs in the case), and the extracted score matrix was subsequently projected to LDA to create the classifier for class predictions. The classification performance of the PCA-LDA is validated by using "leave-one-out" cross validation (LOOCV). In the validation step, the mean centering was performed on the training set, while the test sample was centered by using parameters (mean values) obtained from the training set. The details of validation process of PCA-LDA and the LOOCV were discussed in detail elsewhere[26]. Besides, the discrimination power of each PCs was calculated by PCA-LDA of each PCs to reveal those with the best discrimination power.

CHAPTER III

RESULTS AND DISCUSSION

3.1 Paper Spray Mass Spectrometry Experiments

3.1.1 MS profile of coffee extracts analyzed by PS-MS

Coffee solutions from various regions were analyzed in the positive mode by PS-MS. All representative spectrums for each source are shown in **Table 2**. The major ion is *m/z* 195 associated with the protonated caffeine. Besides, other ions observed were caffeine ($[M+K]^+$ at *m/z* 233), trigonelline ($[M+H]^+$ and $[M+K]^+$ at *m/z* 138 and 176, respectively), choline ($[M+H]^+$ at *m/z* 104), coumaric acid ($[M+H]^+$ at *m/z* 165), and linolenic acid ($[M+K]^+$ at *m/z* 317)[27-28]. In this study, 1-ethyltheobromine was used as an internal standard, which gave signals of the protonated ion at *m/z* 209 along with the potassium adduct at *m/z* 247. The use of IS was to compensate for signal fluctuation and accurate quantification of data. Overall, the MS profiles of coffee samples provided enough profiles for effective differentiation.

Table 2 Representative MS spectra of all coffee sources in the *m/z* range of 100–500.

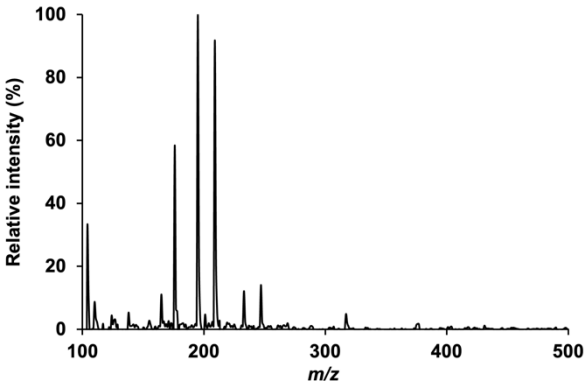
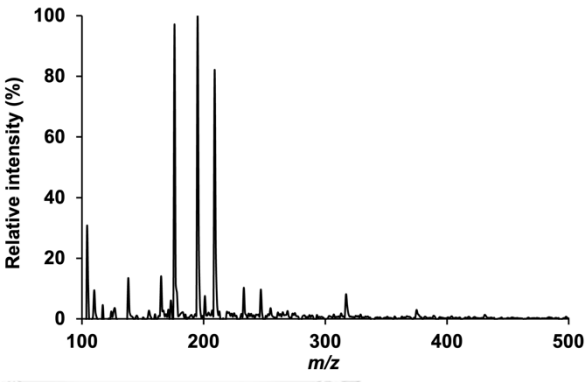
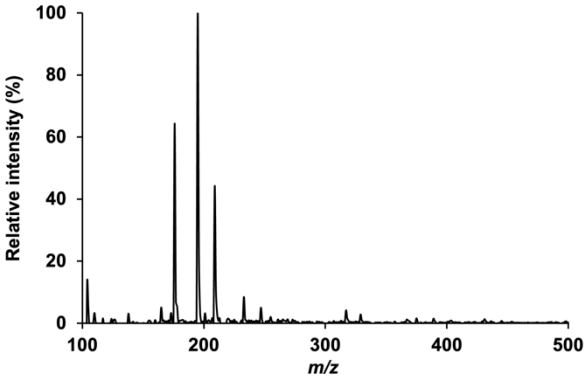
Sample	Spectrum
C1	
C2	
C3	

Table 2 *(continued).*

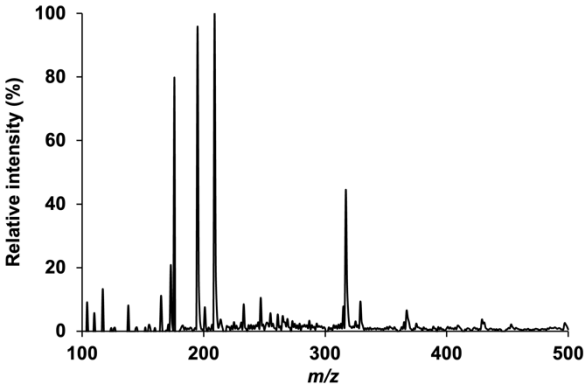
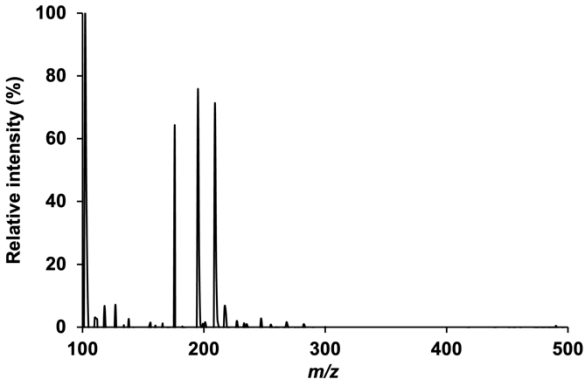
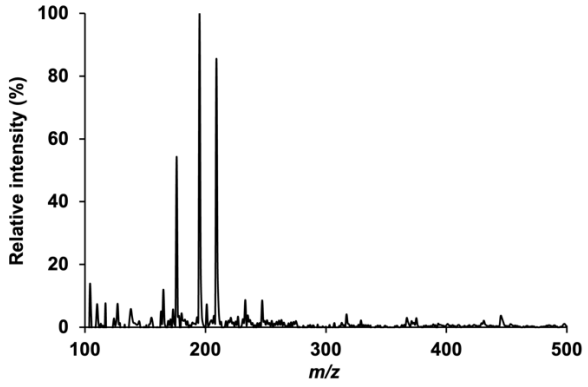
Sample	Spectrum
C4	
C5	
C6	

Table 2 *(continued).*

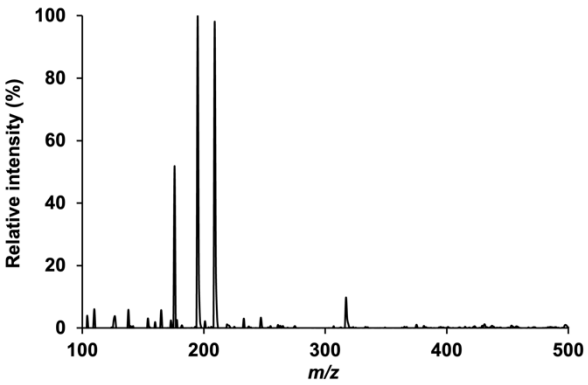
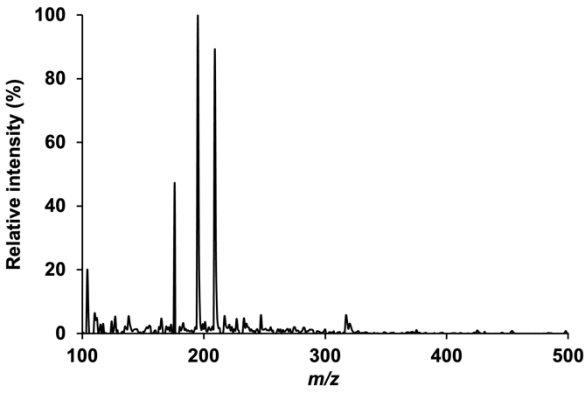
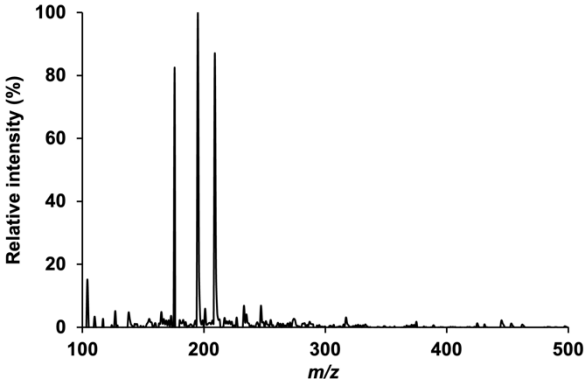
Sample	Spectrum
C7	
C8	
C9	

Table 2 (continued).

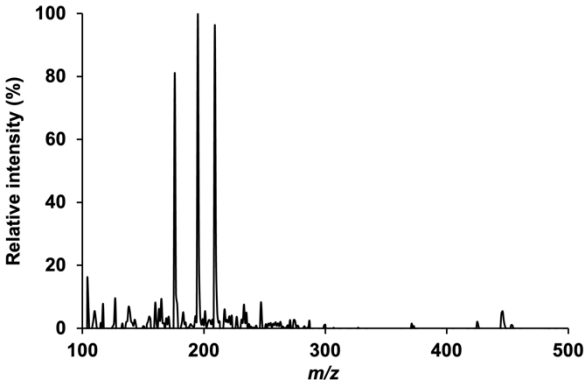
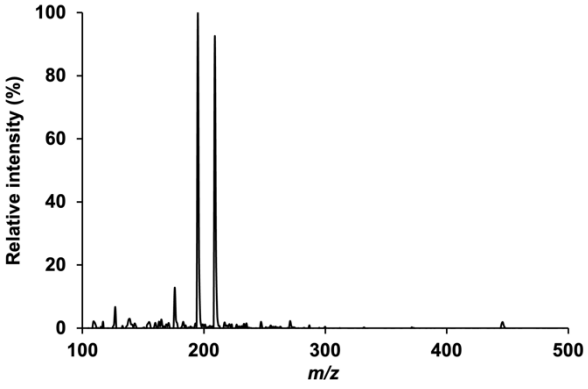
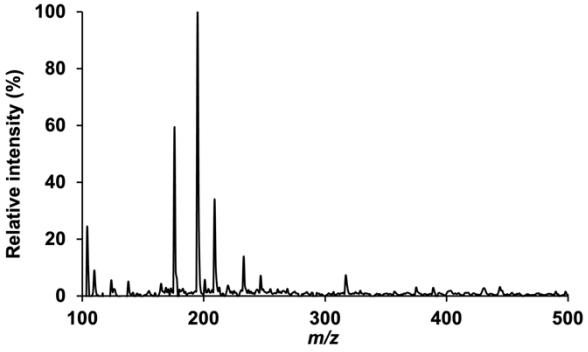
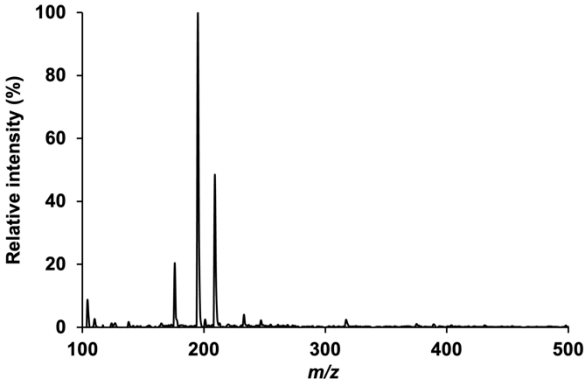
Sample	Spectrum
C10	
C11	
C12	
C13	

Table 2 *(continued).*

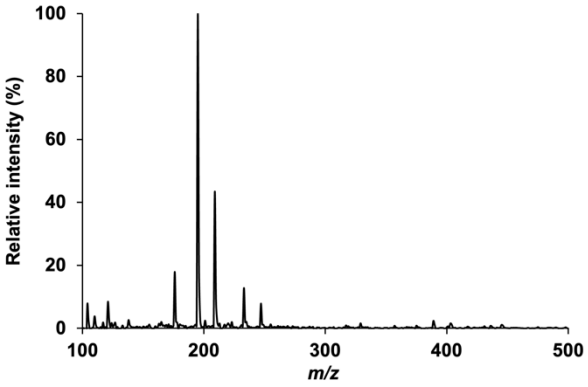
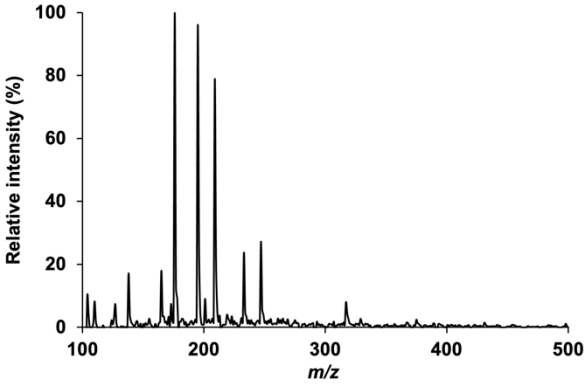
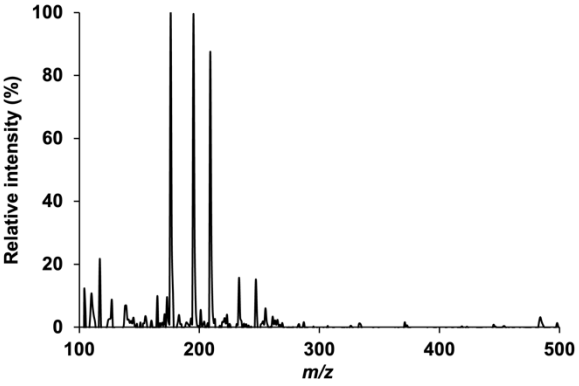
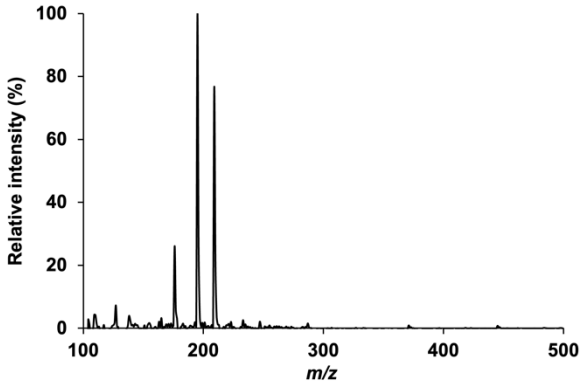
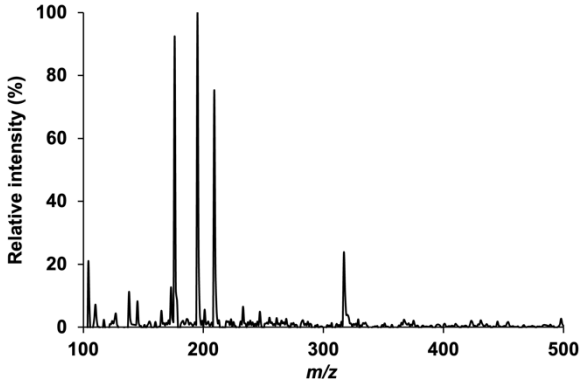
Sample	Spectrum
C14	
C15	
C16	

Table 2 *(continued).*

Sample	Spectrum
C17	
C18	

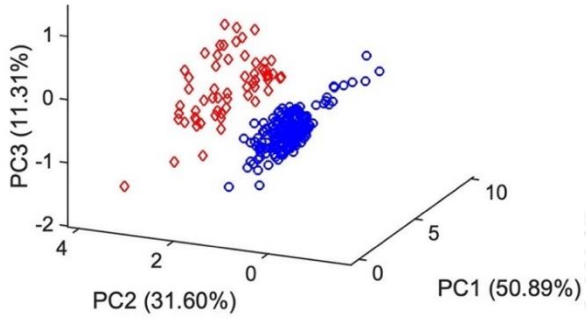
3.1.2 Chemometric analysis

To allow for chemometric analysis, the numerical data in the form of intensity values were collected from all coffee sources in the range of *m/z* 100-500. Some data pre-treatment was then followed. First, the intensity of each *m/z* value from the blank sample was subtracted from each sample to be analyzed. Second, all negative values after subtraction were changed to zero. Lastly, all data points were normalized against the IS peak at *m/z* 209 to account for any variation during the experimental process. This treated data set was then subject to different chemometric analysis.

3.1.2.1 Comparison of chemical profiles in Arabica and Robusta coffees

Since Arabica and Robusta coffees are known to have some distinct attributes, chemometric analysis was first performed to group these two classes of

coffees. As shown in **Figure 8**, principal component (PC) plot of the first 3 PCs clearly distinguish these two groups of coffees with no ambiguity. This is also reflected in the classification accuracy by LDA, where all coffee sources were correctly predicted to be their respective type (100% correct classification). Hence, this effective classification is the first confirmation that the method may show promise as a tool for more complex studies, which is the identification of geographical origins of coffees.



geographical origin	% correct classification
Arabica	100
Robusta	100
Overall	100

Figure 8 PC plots based on two sets of data including Arabica and Robusta coffee sources, along with percentage of correct classifications from LDA.

3.1.2.2 Discrimination of Coffees based on Geographical Origins

After successfully distinguishing two types of coffees, differentiation based on geographical origin was attempted. This task is expected to be more challenging due to the fact that they likely share more similarity within the same type of either Arabica or Robusta coffee. First, we analyzed the case of Arabica coffees by grouping all Arabica coffee sources in Thailand to be a single group, making this study to be a classification based on countries of origin. Since classifying the same type of coffees is likely more challenging, a more thorough analysis was performed as follows. This was done by first evaluating each PC for its prediction accuracy using LDA. In the case of Arabica coffees, it was revealed that PC3, PC4, and PC5 were the best discriminators (**Figure 9(A)**). While surprising at first glance, the fact that the best discriminating PCs were not necessarily the first three PCs (PC1-3) can be explained by considering that all Arabica coffees should contain a similar set of “major”

chemical constituents. It was rather the minor multivariate patterns of variance (including both major and minor components) that should be different between one source from another. Also, a PC plot based on these aforementioned PCs is shown in **Figure 9(B)**, which gives visual suggestion that this method can significantly separate Arabica coffees from different countries. To further clarify the efficiency of differentiation, prediction accuracies were also evaluated based on the combination of PCs. As shown in **Figure 9(C)**, the combination of more numbers of PCs can drastically change the discrimination performance, with 5 PCs being the first point of significant improvement. The performance reached a plateau around 10 PCs, where all sources were correctly predicted at > 85% except in one case (**Figure 9(D)**).

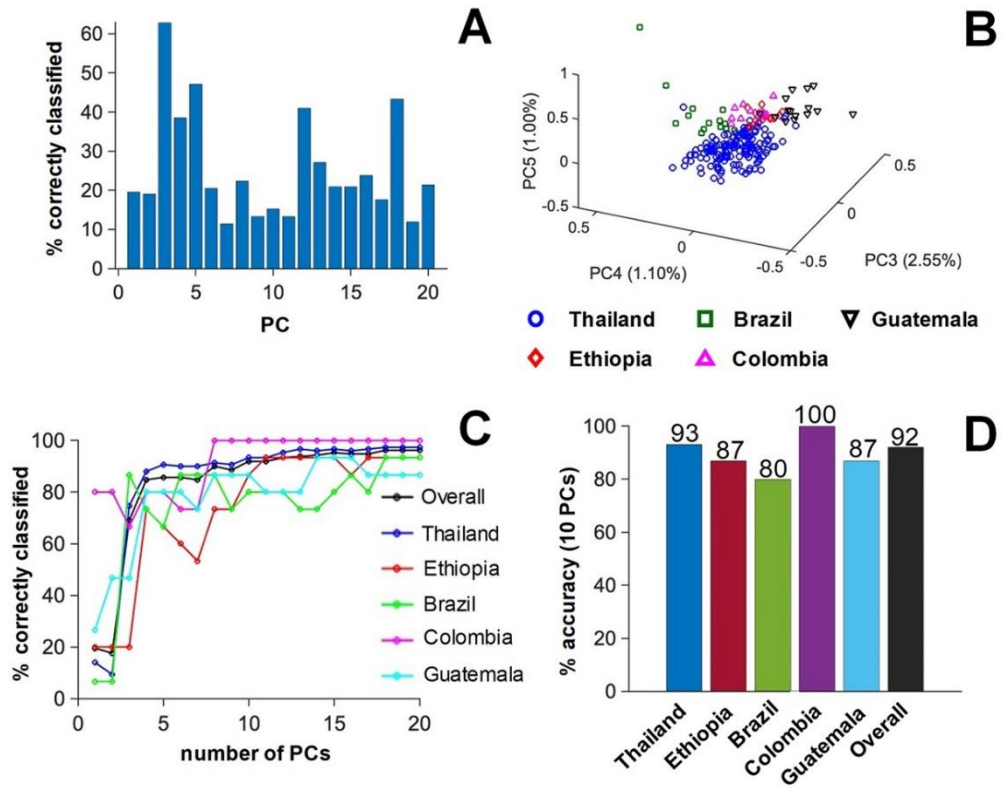


Figure 9 A) Percentage of correct classifications (referenced to the validation sample in cross validation) of each PC (PC1-20) in distinguishing Arabica coffees; B) A PC plot using the best three PCs (PC3-5); C) The relationships between the number of PCs and the percentage of correct classification in all Arabica coffee sources; D) Percentage of prediction accuracies in each Arabica coffee source using first 10 PCs.

The case of Robusta coffees also showed good utility with some superior outcome. First, the best performing PCs were PC2-4 (**Figure 10(A)**), which showed quite clear separation in PCA (**Figure 10(B)**). The combination of PCs also increased the differentiation performance, but with slightly greater overall rate of improvement around the first 5 PCs (**Figure 10(C)**). The first 10 PCs gave great prediction efficiencies, with all being >93 % except in one case (**Figure 10(D)**). Given that three sources (except Vietnam) were actually from the same country, this result confirms the outstanding performance of PS-MS in distinguishing coffee sources with similar profiles.

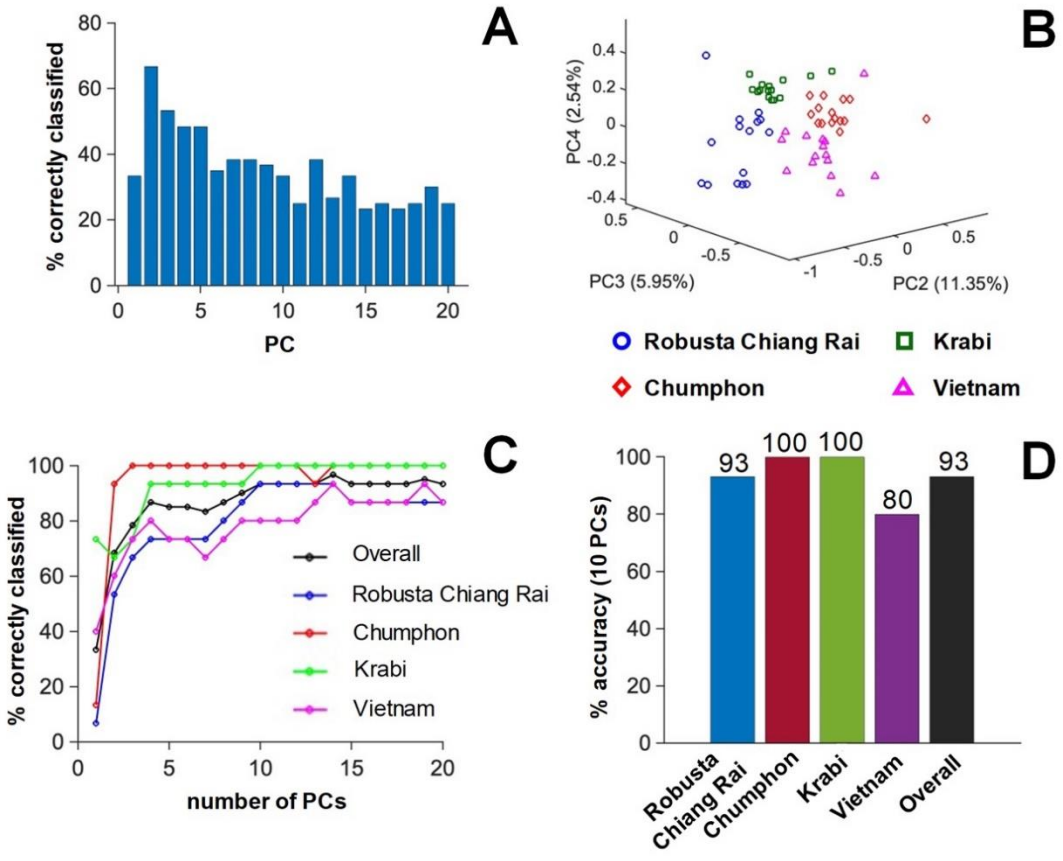
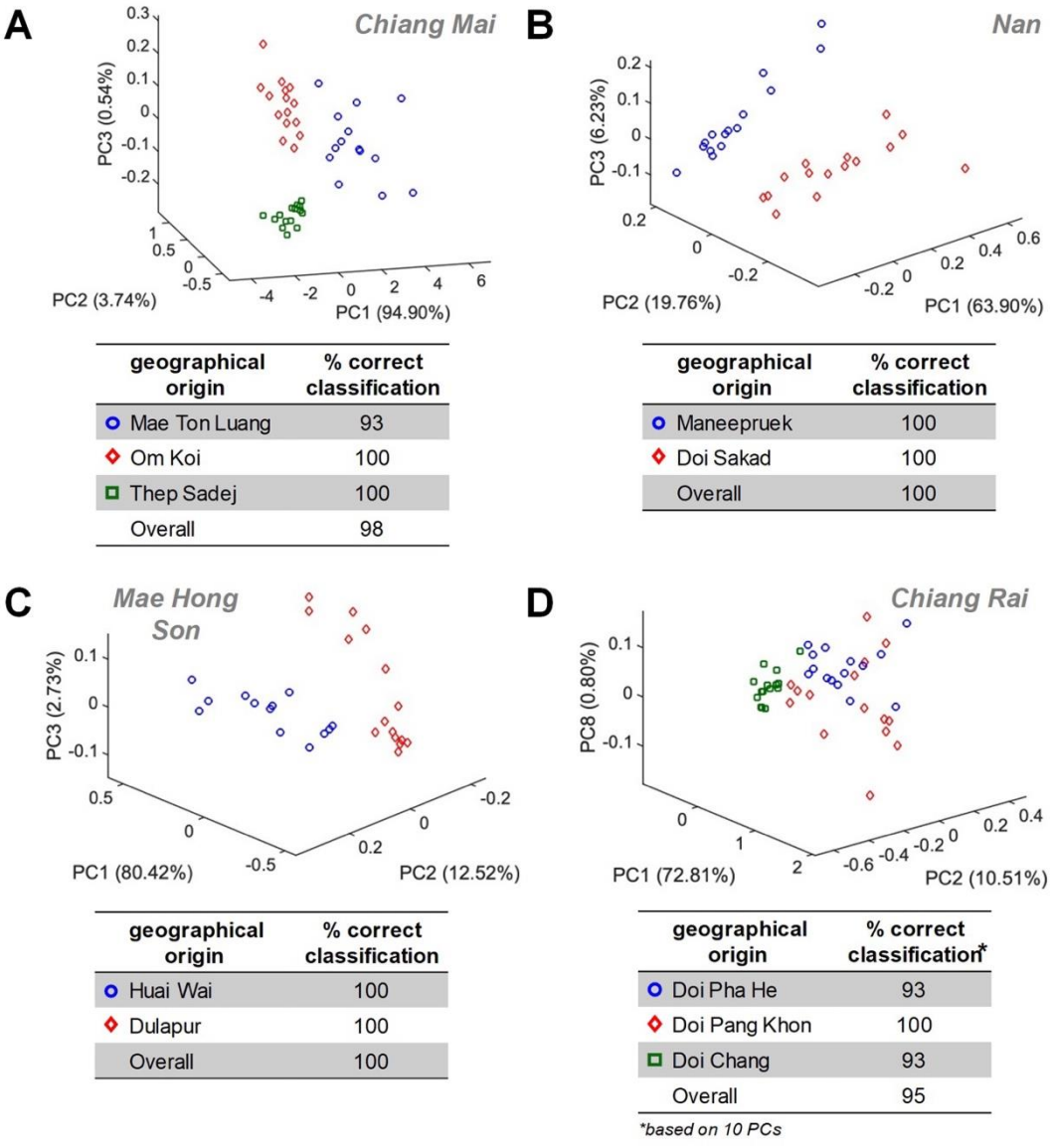


Figure 10 A) Percentage of correct classifications (referenced to the validation sample in cross validation) of each PC (PC1-20) in distinguishing Robusta coffees; B) A PC plot using the best three PCs (PC2-4); C) The relationships between the number of PCs and the percentage of correct classification in all Robusta coffee sources; D) Percentage of prediction accuracies in each Robusta coffee source using first 10 PCs.

With encouraging results as discussed above, the differentiation of the sources of Arabica coffee within the same province was also investigated. The results clearly showed that sources within the provinces of Chiang Mai, Nan, and Mae Hong Son could be very well differentiated using only the first three PCs, with most of the prediction accuracies reaching 100% (**Figure 11(A-C)**). On the other hand, the sources from Chiang Rai were not well separated by the first three PCs, but the combination of more PCs such as 10 PCs can drastically improve the prediction accuracy (**Figure 11(D)** and **Figure 12** for the relationship of percentage of classification accuracy and the number of PCs). This again indicates that chemometric analysis on the data obtained from PS-MS was robust enough for various types of studies including those samples with potentially similar attributes.



geographical origin	% correct classification
Mae Ton Luang	93
Om Koi	100
Thep Sadej	100
Overall	98

geographical origin	% correct classification
Maneepruek	100
Doi Sakad	100
Overall	100

geographical origin	% correct classification
Huai Wai	100
Dulapur	100
Overall	100

geographical origin	% correct classification*
Doi Pha He	93
Doi Pang Khon	100
Doi Chang	93
Overall	95

*based on 10 PCs

Figure 11 PC score plots and the classification accuracies of coffees by sources within the provinces of A) Chiang Mai, B) Nan, C) Mae Hong Son, and D) Chiang Rai.

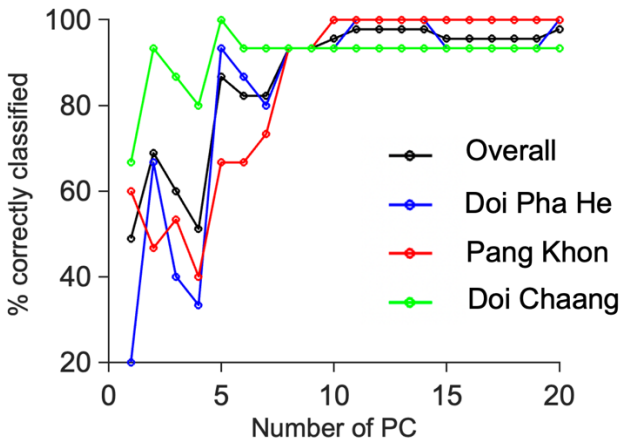


Figure 12 The relationship of percent classification accuracy and the number of PCs used in the classification of coffee sources from the Chiang Rai province in Thailand.

3.2 Gas chromatography-mass spectrometry Experiments

After successfully PS-MS was used to uncover chemical data that can discriminate geographical origins of coffees, an additional technique, HS-SPME/GC-MS, was also used as an alternative method. This method is well-known for the analysis of volatile compounds in coffee samples and may lead to complementary sets of data that are indicative of geographical origins.

3.2.1 Chemical Profile of Volatiles in Coffee Beans

In this study, the grounded coffees from various regions were analyzed by HS-SPME/GC-MS technique. All the coffee samples revealed similar volatile profiles with differing intensities. A representative set of total ion chromatograms (TIC) from different sources of coffee is shown in **Table 3**. The tentative identification of the volatile compounds in the coffee samples were identified according to a comparison of their mass spectra and the linear retention index (LRI) with those present in the NIST 14 database. The criteria for compound identification required an LRI difference of ≤20 units between the calculated LRI and the database values. Using these criteria, the list of 17 volatile metabolites were identified (**Table 4**). Several classes of organic compounds were observed, including notably aromatic heterocycles (furans, pyrazines, pyridine), furanones, aldehyde, phenol, aliphatic hydrocarbons, and carboxylic acid. All compounds detected in the GC-MS chromatograms have also

been reported by several research groups that analyzed coffee by HS/GC-MS.[29-30] According to the %RPAs, the major compounds found in coffee samples were 2-Furanmethanol and 2-Methoxy-4-vinylphenol, which are labeled as peaks no. 7 and 15, respectively. Overall, the volatile profiles of coffee samples provide enough features that may lead to effective differentiation of coffee origins.

Table 3 Representative total ion chromatogram (TIC) from different sources of coffee.

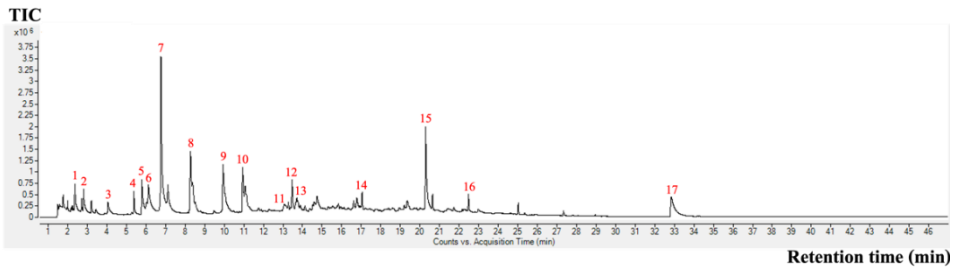
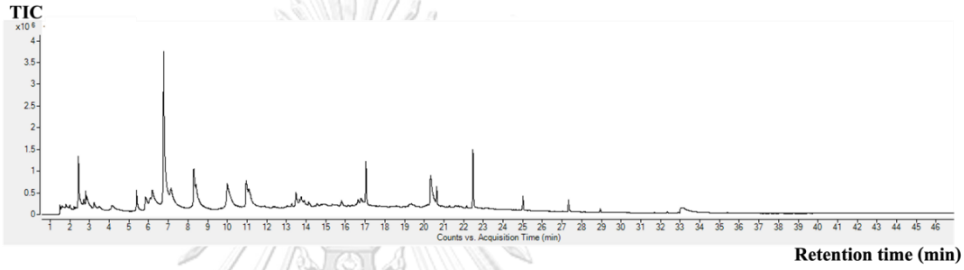
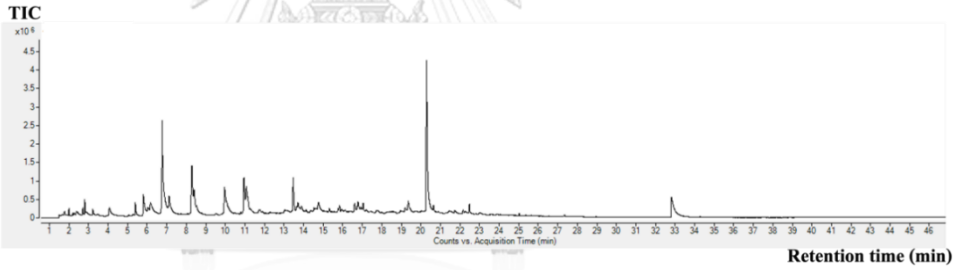
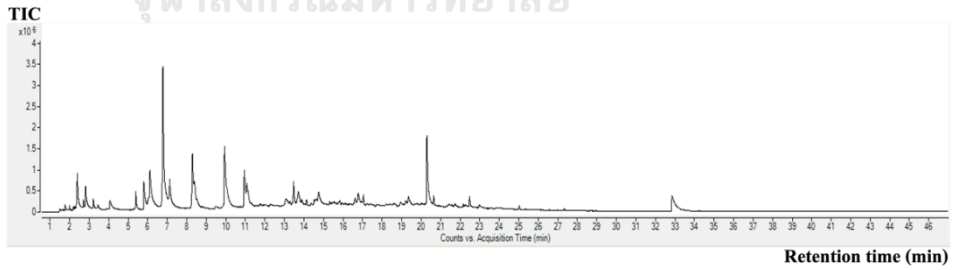
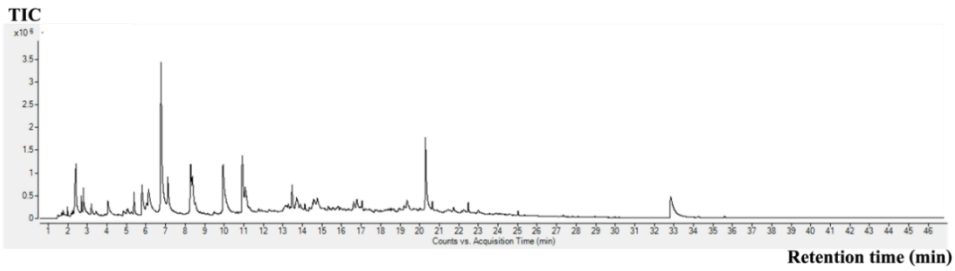
Sample	Chromatogram
C1	
C2	
C3	
C4	
C5	

Table 3 (continued).

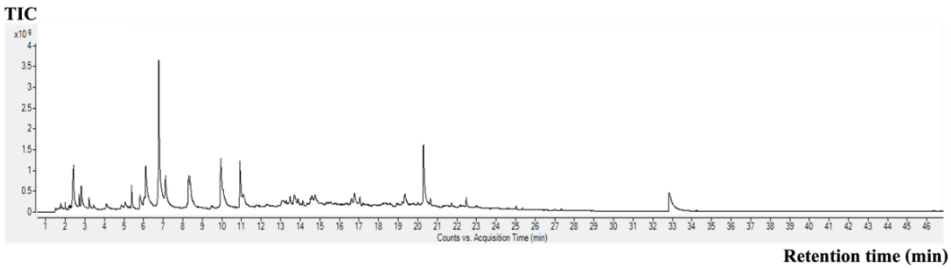
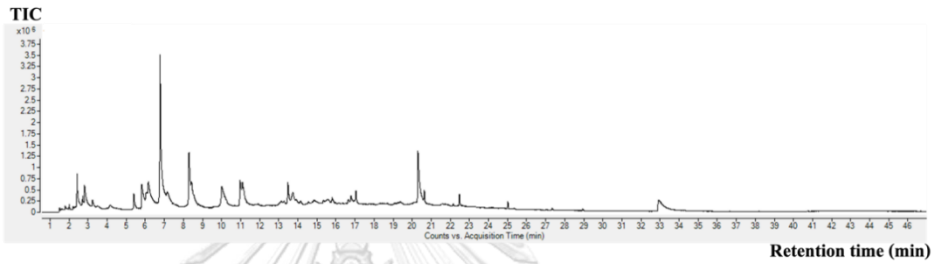
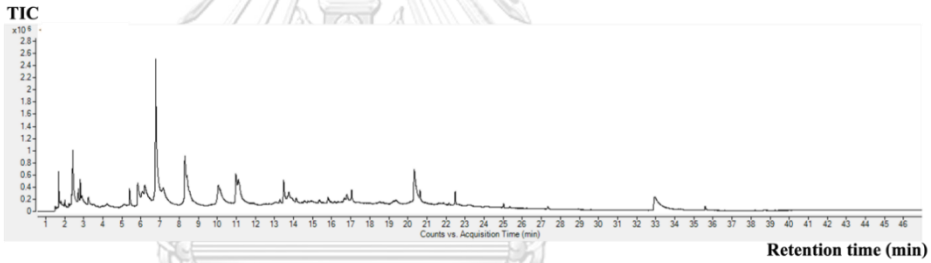
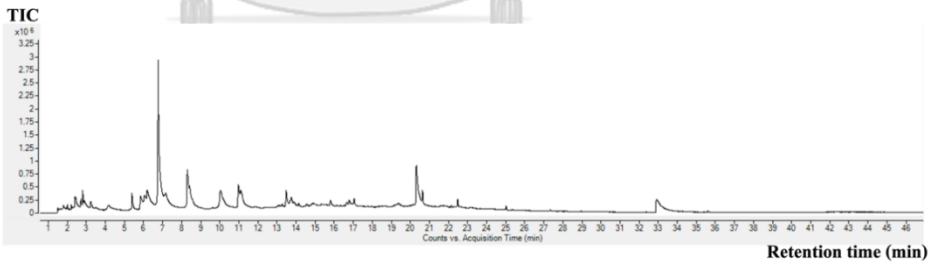
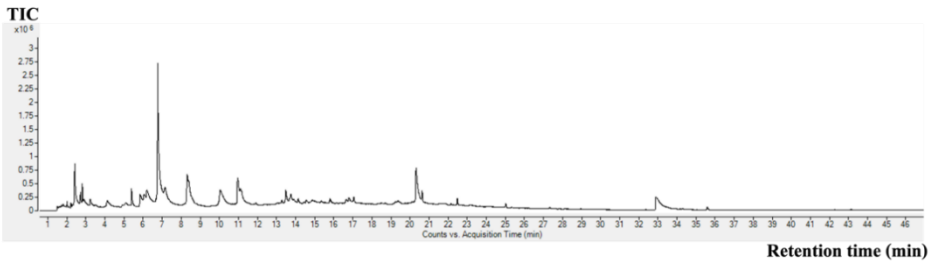
Sample	Chromatogram
C6	
C7	
C8	
C9	
C10	

Table 3 *(continued).*

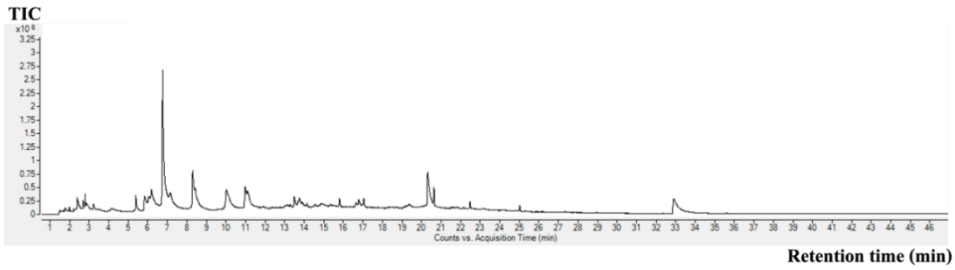
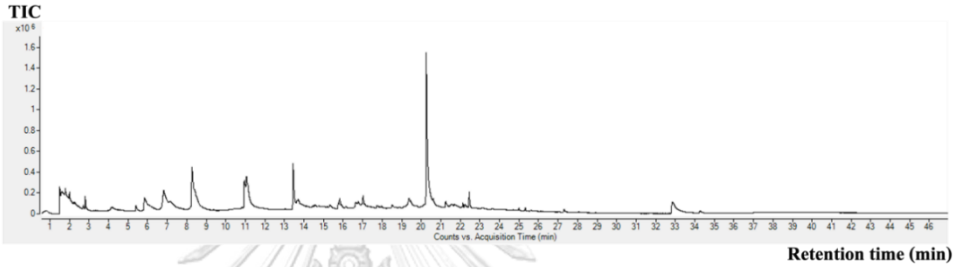
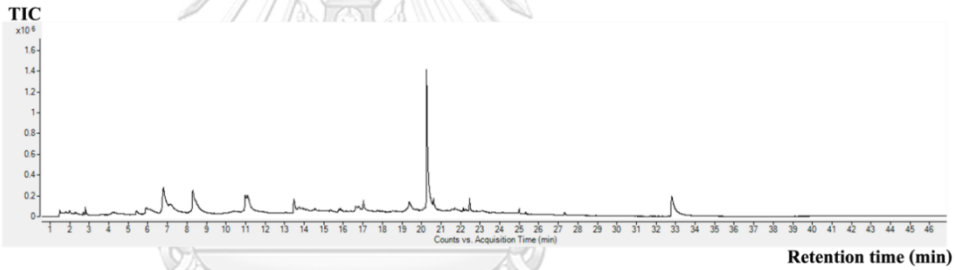
Sample	Chromatogram
C11	
C12	
C13	

Table 4 Volatile compounds of Doi Pha Hee coffee (C1) obtained by HS-SPME GC-MS.

Peak No.	RT (min)	Compound	CAS No.	LRI	
				Exp[a]	Database[b]
1	2.375	Acetic acid	64-19-7	<800	610±10
2	2.806	2-Methylbutanal	96-17-3	<800	662±8
3	4.055	Pyridine	110-86-1	<800	746±7
4	5.378	Dihydro-2-methyl-3(2H)-furanone	3188-00-9	825	809±3
5	5.783	2-Methylpyrazine	109-08-0	836	831±7
6	6.120	Furfural	98-01-1	846	833±4
7	6.765	2-Furanmethanol	98-00-0	865	859±6
8	8.280	2,6-Dimethylpyrazine	108-50-9	910	917±7
9	9.933	5-Methyl-2-furancarboxaldehyde	620-02-0	964	965±5
10	10.941	Furfuryl acetate	623-17-6	996	995±4
11	13.070	2,5-Dimethyl-4-hydroxy-3(2H)-furanone	3658-77-3	1065	1070±10
12	13.464	2,5-Dimethyl-3-ethylpyrazine	13360-65-1	1085	1082±3
13	13.704	2-Ethyl-3,5-dimethylpyrazine	13925-07-0	1078	1084±3
14	17.043	Dodecane	112-40-3	1196	1200
15	20.290	2-Methoxy-4-vinylphenol	7786-61-0	1311	1317±5
16	22.477	Tetradecane	629-59-4	1392	1400
17	32.848	Caffeine	58-08-2	1855	1835±7

[a]Exp : Experimental linear retention indices calculated using *n*-alkane standards.

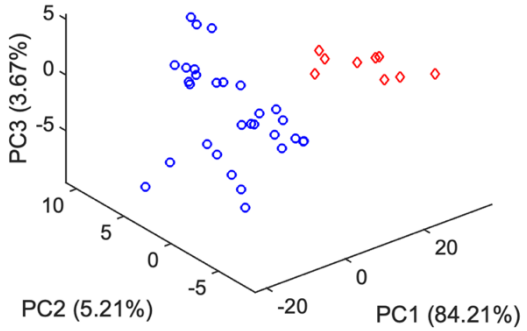
[b]Database : Linear retention indices obtained from NIST 14 database.

3.2.2 Chemometric analysis

To allow for chemometric analysis, the %RPA of 17 volatile compounds (variables) from 39 coffee samples (13 coffee sources with 3 replicates) were used as an input dataset for calculation. The dataset was then subject to different chemometric analysis. PCA, a popular unsupervised technique was applied to reduce the dimensionality in a data set for better visual presentations. To obtain the classification performance, LDA was used for class prediction[31].

3.2.2.1 Comparison of Volatile Compounds in Arabica and Robusta coffees

Classification between Arabica and Robusta was the first to be studied. Since variation between coffee species is expected to give clear differences in chemical properties of coffees,[32] this study was a prerequisite to more complex studies. The dataset was subjected to an unsupervised analysis by PCA to evaluate the possibility of discriminating Arabica and Robusta coffees. The PC plot of the first 3 PCs is shown in **Figure 13**. The result indicated a clear separation between the coffee species. This is also reflected in the classification accuracy by LDA, where the overall predictive ability of the constructed model was 92%. Hence, the result from this initial study confirms that this method has potential to classify coffee samples based on their origins.



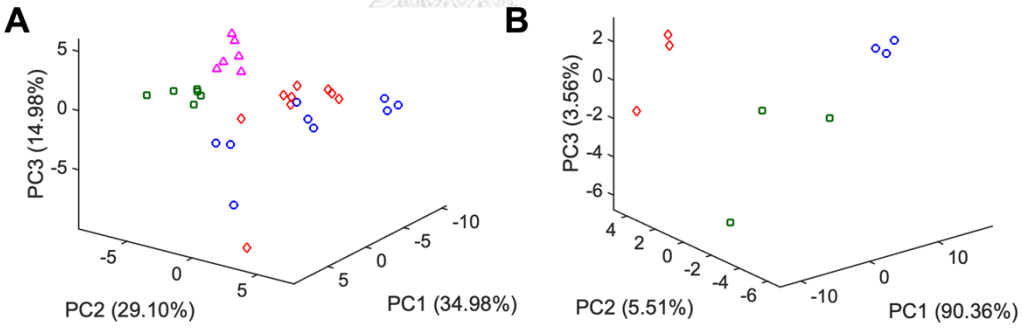
geographical origin	% correct classification
Arabica	100
Robusta	67
Overall	92

Figure 13 PC score plot of the first 3 principal components to visualize the cluster relationship of types of coffees, along with the percentage of correct classifications from LDA.

3.2.2.2 Discrimination of Coffees based on Geographical Origins

Encouraged by the initial result, we also studied the differentiation of coffees based on geographical origins. First, we analyzed different sources of Arabica coffee within certain provinces of Thailand. The resulting PCA score plot is shown in **Figure 14 (A)**. The first 3 PCs (PC1-PC3) were chosen in order to represent the data objects with the highest variation (34.98%, 29.10%, and 14.98% of the variation). The results showed that these coffee samples can be separated into four groups according to the provinces of origin. The PCA grouping result was in agreement with that obtained using LDA, with classification rates of 56%, 67%, 83%, 67%, and 67% from Chiang Rai, Chiang Mai, Nan, Mae Hong Son, and overall, respectively. Interestingly, lower classification accuracies among the provinces of Chiang Rai, Chiang Mai, and Mae Hong Son may probably due to their closer proximities in comparison to Nan.

The case of Robusta coffees also showed good differentiation performance with the first 3 PCs (**Figure 14 (B)**). The first 3 PCs gave great prediction accuracies by reaching 100% accuracy in all cases.

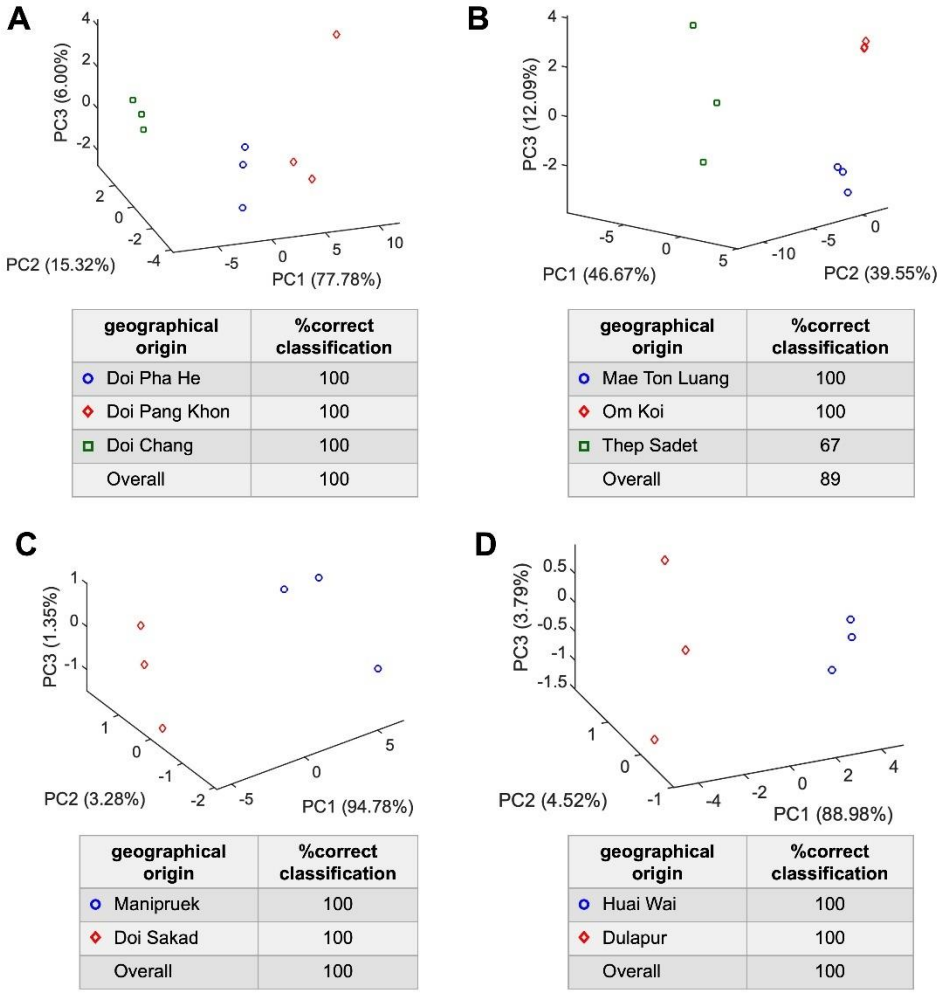


geographical origin	% correct classification
Chiang Rai	56
Chiang Mai	67
Nan	83
Mae Hong Son	67
Overall	67

geographical origin	% correct classification
Robusta Chiang Rai	100
Chumphon	100
Krabi	100
Overall	100

Figure 14 A) PC score plots of the first 3 principal components in all Arabica coffee sources and B) PC score plots of the first 3 principal components in all Robusta coffee sources.

After successfully distinguishing coffees based on their provinces of origin as discussed above, we studied the differentiation of the sources of Arabica coffee within the same province. This study was more challenging because the chemical profiles are expected to be more similar due to their closer locations of origin. The PCA score plot of all sources within the provinces of Chiang Rai, Chiang Mai, Nan, and Mae Hong Son indicated successful discrimination of origins using only the first 3 PCs, with most of the prediction accuracies reaching 100% (except in one case) (**Figure 15**). Thus, HS-SPME/GC-MS, which requires no solvent and free from pre-treatment processes, is an attractive and convenient method to uncover geographical indications of coffees.



geographical origin	%correct classification
Doi Pha He	100
Doi Pang Khon	100
Doi Chang	100
Overall	100

geographical origin	%correct classification
Mae Ton Luang	100
Om Koi	100
Thep Sadet	67
Overall	89

geographical origin	%correct classification
Manipruek	100
Doi Sakad	100
Overall	100

geographical origin	%correct classification
Huai Wai	100
Dulapur	100
Overall	100

Figure 15 PC score plots of Arabica coffees within the same province of A) Chiang Rai, B) Chiang Mai, C) Nan, and D) Mae Hong Son.

CHAPTER IV
CONCLUSION

In this work, PS–MS was used to obtain MS data from aqueous extracts of roasted coffees from a variety of sources with little sample pre-treatment. In contrast, profiles of volatiles in a variety of coffee samples were obtained by HS-SPME/GC-MS. The resulting data were used for PCA and LDA, which were able to discriminate a variety of coffees based on 1) the coffee species, 2) geographical origins of Arabica coffees, 3) geographical origins of Robusta coffees, and 4) geographical origins of Arabica coffees within the same province of Thailand. In the case of PS-MS, classification accuracies were in the range of 80-100%. Hence, this study clearly illustrated that PS-MS with chemometric analysis is effective in discriminating coffee samples based on their origins, which can be further adopted as a tool that aids the registration of GI.

In addition, HS-SPME/GC-MS is an alternative method for the discrimination of coffee origins. PCA and LDA suggested that this method is slightly less effective (in the range of 65-100% accuracies except in one case) than PS-MS. Overall, both methods can be viable tools for differentiating coffee origins.

REFERENCES

1. Sharma, H. A Detail Chemistry of Coffee and Its Analysis. In *Coffee*; Dalyse Toledo, C., IntechOpen, 2020.
2. Cheng, B.; Furtado, A.; Smyth, H. E.; Henry, R. J. Influence of genotype and environment on coffee quality. *Trends in Food Science & Technology* **2016**, *57*, 20–30.
3. Sunarharum, W. B.; Williams, D. J.; Smyth, H. E. Complexity of coffee flavor: A compositional and sensory perspective. *Food Research International* **2014**, *62*, 315–325.
4. Katerinopoulou, K.; Kontogeorgos, A.; Salmas, C. E.; Patakas, A.; Ladavos, A. Geographical Origin Authentication of Agri-Food Products: A Review. *Foods* **2020**, *9* (4).
5. Teuber, R. Geographical Indications of Origin as a Tool of Product Differentiation: The Case of Coffee. *Journal of International Food & Agribusiness Marketing* **2010**, *22* (3-4), 277–298.
6. Arana, V. A.; Medina, J.; Alarcon, R.; Moreno, E.; Heintz, L.; Schäfer, H.; Wist, J. Coffee's country of origin determined by NMR: The Colombian case. *Food Chemistry* **2015**, *175*, 500–506.
7. Mehari, B.; Redi-Abshiro, M.; Chandravanshi, B. S.; Combrinck, S.; Atlabachew, M.; McCrindle, R. Profiling of phenolic compounds using UPLC–MS for determining the geographical origin of green coffee beans from Ethiopia. *Journal of Food Composition and Analysis* **2016**, *45*, 16–25.
8. Babova, O.; Occhipinti, A.; Maffei, M. E. Chemical partitioning and antioxidant capacity of green coffee (Coffea arabica and Coffea canephora) of different geographical origin. *Phytochemistry* **2016**, *123*, 33–39.
9. Mehari, B.; Redi-Abshiro, M.; Chandravanshi, B. S.; Combrinck, S.; McCrindle, R.; Atlabachew, M. GC-MS profiling of fatty acids in green coffee (Coffea arabica L.) beans and chemometric modeling for tracing geographical origins from Ethiopia. *Journal of the Science of Food and Agriculture* **2019**, *99* (8), 3811–3823.

10. Putri, S. P.; Irifune, T.; Yusianto; Fukusaki, E. GC/MS based metabolite profiling of Indonesian specialty coffee from different species and geographical origin. *Metabolomics* **2019**, *15* (10), 126.
11. Liu, J.; Wang, H.; Manicke, N. E.; Lin, J.-M.; Cooks, R. G.; Ouyang, Z. Development, Characterization, and Application of Paper Spray Ionization. *Analytical Chemistry* **2010**, *82* (6), 2463–2471.
12. Di Donna, L.; Taverna, D.; Indelicato, S.; Napoli, A.; Sindona, G.; Mazzotti, F. Rapid assay of resveratrol in red wine by paper spray tandem mass spectrometry and isotope dilution. *Food Chemistry* **2017**, *229*, 354–357.
13. Yu, M.; Wen, R.; Jiang, L.; Huang, S.; Fang, Z.; Chen, B.; Wang, L. Rapid analysis of benzoic acid and vitamin C in beverages by paper spray mass spectrometry. *Food Chemistry* **2018**, *268*, 411–415.
14. Liu, J.; He, Y.; Chen, S.; Ma, M.; Yao, S.; Chen, B. New urea-modified paper substrate for enhanced analytical performance of negative ion mode paper spray mass spectrometry. *Talanta* **2017**, *166*, 306–314.
15. Wang, H.; Liu, J.; Cooks, R. G.; Ouyang, Z. Paper Spray for Direct Analysis of Complex Mixtures Using Mass Spectrometry. *Angewandte Chemie International Edition* **2010**, *49* (5), 877–880.
16. Yang, Y.; Wu, J.; Deng, J.; Yuan, K.; Chen, X.; Liu, N.; Wang, X.; Luan, T. Rapid and on-site analysis of amphetamine-type illicit drugs in whole blood and raw urine by slug-flow microextraction coupled with paper spray mass spectrometry. *Analytica Chimica Acta* **2018**, *1032*, 75–82.
17. Zakidou, P.; Plati, F.; Matsakidou, A.; Varka, E.-M.; Blekas, G.; Paraskevopoulou, A. Single Origin Coffee Aroma: From Optimized Flavor Protocols and Coffee Customization to Instrumental Volatile Characterization and Chemometrics. *Molecules* **2021**, *26* (15), 4609.
18. Ongo, E. A.; Montevecchi, G.; Antonelli, A.; Sberveglieri, V.; Sevilla Iii, F. Metabolomics fingerprint of Philippine coffee by SPME-GC-MS for geographical and varietal classification. *Food Research International* **2020**, *134*, 109227.

19. Sparkman, O. D.; Penton, Z. E.; Kitson, F. G. Chapter 2 – Gas Chromatography. In *Gas Chromatography and Mass Spectrometry*, 2 ed.; Sparkman, O. D., Penton, Z. E., Kitson, F. G., Eds.; Academic Press, 2011; pp 15–83.
20. Schmidt, K.; Podmore, I. Solid Phase Microextraction (SPME) Method Development in Analysis of Volatile Organic Compounds (VOCS) as Potential Biomarkers of Cancer. *Journal of Molecular Biomarkers & Diagnosis* **2015**, *06*.
21. Hussain, S. Z.; Maqbool, K. GC-MS: Principle, Technique and its application in Food Science. *International Journal of Current Science* **2014**, *13*, 116–126.
22. Brown, S. D.; Sum, S. T.; Despagne, F.; Lavine, B. K. Chemometrics. *Analytical Chemistry* **1996**, *68* (12), 21–62.
23. Brereton, R. G. *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*; John Wiley & Sons, Ltd., 2003.
24. Smrke, S.; Kroslakova, I.; Gloess, A. N.; Yeretzian, C. Differentiation of degrees of ripeness of Catuai and Tipica green coffee by chromatographical and statistical techniques. *Food Chemistry* **2015**, *174*, 637–642.
25. Bianchi, F.; Careri, M.; Mangia, A.; Musci, M. Retention indices in the analysis of food aroma volatile compounds in temperature-programmed gas chromatography: Database creation and evaluation of precision and robustness. *Journal of Separation Science* **2007**, *30* (4), 563–572.
26. Charoensumran, P.; Rauytanapanit, M.; Sricharoen, N.; Smith, B. L.; Wongravee, K.; Maher, S.; Praneenararat, T. Rapid geographical indication of peppercorn seeds using corona discharge mass spectrometry. *Scientific Reports* **2021**, *11* (1), 16089.
27. Amorim, A.; Hovell, A.; Pinto, A.; Eberlin, M.; Arruda, N.; Pereira, E.; Bizzo, H.; Catharino, R.; Filho, Z.; Rezende, C. Green and Roasted Arabica Coffees Differentiated by Ripeness, Process and Cup Quality via Electrospray Ionization Mass Spectrometry Fingerprinting. *Journal of the Brazilian Chemical Society* **2008**, *20*, 313–321.
28. Garrett, R.; Rezende, C. M.; Ifa, D. R. Coffee origin discrimination by paper spray mass spectrometry and direct coffee spray analysis. *Analytical Methods* **2013**, *5* (21), 5944–5948.

29. Amanpour, A.; Selli, S. Differentiation of Volatile Profiles and Odor Activity Values of Turkish Coffee and French Press Coffee. *Journal of Food Processing and Preservation* **2016**, *40* (5), 1116–1124.
30. Ochiai, N.; Tsunokawa, J.; Sasamoto, K.; Hoffmann, A. Multi-volatile method for aroma analysis using sequential dynamic headspace sampling with an application to brewed coffee. *Journal of Chromatography A* **2014**, *1371*, 65–73.
31. Giacomino, A.; Abollino, O.; Malandrino, M.; Mentasti, E. The role of chemometrics in single and sequential extraction assays: A Review. Part II. Cluster analysis, multiple linear regression, mixture resolution, experimental design and other techniques. *Analytica Chimica Acta* **2011**, *688* (2), 122–139.
32. Clifford, M. N. Chemical and Physical Aspects of Green Coffee and Coffee Products. In *Coffee: Botany, Biochemistry and Production of Beans and Beverage*, Clifford, M. N., Willson, K. C. Eds.; Springer US, 1985; pp 305–374.

VITA

NAME	Rujirat Pumbua
DATE OF BIRTH	19 February 1997
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Department of Petrochemistry and Polymer Science, Faculty of Science, Chulalongkorn University
HOME ADDRESS	10/205 Moo 3 Suan Phrik Thai, Muang Pathum Thani, Pathum Thani 12000