ADULTERATION DETECTION OF STINGLESS BEE HONEY USING UNTARGETED ^{1}H-NMR METABOLOMICS IN ADDRESSING THE LIMITATION OF STABLE CARBON ISOTOPE RATIO ANALYSIS

YONG CHIN HONG

UNIVERSITI SAINS MALAYSIA

2022

ADULTERATION DETECTION OF STINGLESS BEE HONEY USING UNTARGETED ^{1}H-NMR METABOLOMICS IN ADDRESSING THE LIMITATION OF STABLE CARBON ISOTOPE RATIO ANALYSIS

by

YONG CHIN HONG

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

September 2022

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my main supervisor, Ts. ChM. Dr. Syahidah Akmal Muhammad for her invaluable advice, guidance, patience, understanding and encouragement throughout the course of this research. She has been supportive and giving me much independence, time flexibility and trust in doing my research. I would like to give my utmost appreciation to my co-supervisors, Dr. Fatimatuzzahra Abdul Aziz and Ts. Dr. Seow Eng Keng for their time, suggestions, guidance and advices during the hard time in my research.

I would like to thank Puan Ainolsyakira Mohd Rodhi and Puan Fatin Ilyani Nasir from Analytical Biochemistry Research Centre for their time and willingness to help and guide me during the operation of instruments which I had used throughout this research. I would also like to thank my friend and colleague Ng Jing Sheng as we worked together, advised each other and helped each other whether or not it was about our research studies.

I would give my acknowledgement of the research fund granted by the Ministry of Higher Education, Malaysia under the Fundamental Research Grant Scheme (FRGS) with the grant number: FRGS/1/2017/WAB01/USM/02/2 and the financial aid by the Kuok Foundation for the MSM Fellowship Award provided. Their kind offerings have enables me to focus on my study without financial worries.

I would like to thank my parents and my brothers for allowing me to pursue my study away from home. I have been known by them as the ‘black sheep’ who lives in

a ‘different world’ and understanding, and thus, they do not agree with my decision and may not understand the things I do nor how I think. We do argue a lot yet, they still welcome me home whenever I get back.

Most importantly, I would like to dedicate my gratefulness to my dearest mates: Gui Kang En, Tan Kah Aik Jack, Lee Sze Chi, Choong Yee Yaw, Azieda Abdul Talib and Low Yan Kitt. Thank you for having my back throughout the tough time I had. We teased each other, complained, talked about everything and we really had our fun time all these years. I am really glad that I meet all of you and I appreciate all the wonderful memories we had together.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

‰	Parts per thousand or per mil
P_A	Percentage of protein extracted from honey after lyophilization
P_B	Percentage of protein extracted from honey before lyophilization
Q^2	Predictive ability
R^2	Fitness of data
$R^2(Y)$	Total sum variation
R_{sample}	Isotope ratio ($^{13}C/^{12}C$) of the sample
$R_{standard}$	Isotopic ratio of the international reference materials
t[1]	Principal component one
t[2]	Principal component two
W%	Percentage of moisture content
$\delta^{13}C$	Stable isotopes ratio of ^{13}C:^{12}C
$\delta^{13}C_{(honey)}$	$\delta^{13}C$ of honey
$\delta^{13}C_{(protein)}$	$\delta^{13}C$ of protein extracted from honey
$\delta^{13}C_{(sweetener)}$	$\delta^{13}C$ of sweetener

LIST OF ABBREVIATIONS

ACV	Apple cider vinegar
AOAC	Association of Official Analytical Chemists
CV-ANOVA	Cross-validated analysis of variation
D1	Relaxation delay
D_2O	Deuterated water/heavy water
DSS	4,4-dimethyl-4-silapentane-1-sulfonic acid
EA-IRMS	Elemental analyser – isotope ratio mass spectrometer
FTIR	Fourier-transform infrared spectroscopy
GCMS	Gas chromatography – mass spectrometer
H_2O	Distilled water
H_2SO_4	Sulphuric acid
HBH	Honey bee honey
HFCS	High fructose corn syrup
HILDA	Honey Interlinked Dehydration and Dispenser Apparatus
HMF	5-hydroxymethylfurfural
HPLC	High-performance liquid chromatography
HPLC-DAD	High-performance liquid chromatography – diode array detection
ISCIRA	Internal standard isotope ratio analysis
LCS	Light corn syrup
MpS	Maple syrup
MS	Mass spectrometry
Na_2HPO_4	Sodium hydrogen phosphate
NaH_2PO_4	Sodium dihydrogen phosphate
NIR	Near-infrared spectroscopy
NMR	Nuclear magnetic resonance
NOESY	^{1}H-nuclear Overhauser effect spectroscopy
ns	Number of scans
O1	Transmitter frequency offset
OAA	4-carbon compound oxaloacetic acid
OPLS-DA	Orthogonal partial least square – discriminant analysis
PC	Principal component

PCA	Principal component analysis
PGA	3-carbon compound phosphoglyceric acid
PS	Palm sugar
PTFE	Polytetrafluoroethylene
R.I.	Refractive index
RG	Receiver gain
SBH	Stingless bee honey
SCIRA	Stable carbon isotope ratios analysis
SIMCA	Soft independent modelling of class analogy
TSP	3-(Trimethylsilyl)propanoic acid
USGS	United States Geological Survey
USM	Universiti Sains Malaysia
UV-Vis	Ultraviolet – visible spectroscopy
VPDB	Vienna Pee Dee Belemnite

LIST OF APPENDICES

PENGESANAN PENGADUKAN MADU KELULUT MENGGUNAKAN METABOLOMIK ^{1}H-NMR TIDAK TERTUMPU UNTUK MENANGANI LIMITASI ANALISIS NISBAH KARBON ISOTOP STABIL

ABSTRAK

Madu kelulut yang kini semakin popular dalam pasaran Malaysia, ia terdedah kepada masalah pengadukan. Beberapa kajian telah mendapati bahawa bahan nutrien madu kelulut adalah sama ataupun lebih tinggi berbanding dengan madu lebah. Harga madu kelulut yang lebih tinggi berbanding dengan madu lebah telah menyebabkan penggunaan bahan campurpalsu yang jarang digunakan seperti cuka, termasuk madu lebah untuk meniru rasa dan penampilan madu kelulut yang unik. Untuk menentukan ketulenan madu kelulut, kaedah semasa yang digunakan dan diterima secara antarabangsa adalah kaedah 'Association of Official Analytical Chemists' (AOAC) Standard Dalaman Analisis Isotop Stabil Karbon (ISCIRA). Kaedah ini memerlukan protein diekstrak daripada madu kelulut. Namun begitu, kandungan lembapan yang tinggi dalam madu kelulut menyebabkan jumlah protein diekstrak yang rendah bagi setiap gram. Selain itu, kaedah semasa AOAC ISCIRA ini gagal untuk mengesan bahan campurpalsu yang diperolehi daripada tumbuhan C_3 memandangkan madu kelulut yang dicampur dengan bahan campurpalsu ini mempunyai nilai julat δ^{13}C yang sama. Oleh itu, resonans magnetik nuklear proton (^{1}H-NMR) metabolomik tidak bersasar dengan kemometri telah dicadangkan untuk menubuhkan profil metabolik madu kelulut untuk mengesan madu kelulut yang dicampur dengan bahan campurpalsu yang diperolehi daripada tumbuhan C_3 dan C_4. Untuk menentukan ketulenan madu kelulut sebelum mencipta profil metabolik, kaedah AOAC ISCIRA telah digunakan. Proses liofilisasi telah diperkenalkan untuk mengurangkan

kandungan lembapan madu kelulut. Didapati bahawa apabila kandungan lembapan sampel madu kelulut dikurangkan ke tahap kurang daripada 20%, peratusan kenaikan protein yang diekstrak daripada sampel adalah antara 6 hingga 385% berbanding dengan protein yang diekstrak daripada madu kelulut sebelum liofilisasi. Dengan itu, sembilan daripada 22 sampel madu kelulut yang digunakan dalam kajian ini didapati tidak tulen. Setelah ketulenan setiap sampel madu kelulut ditentukan, kesemua sampel madu kelulut dianalisis dengan ^{1}H-NMR metabolomik dengan kemometri. Analisis komponen utama tanpa pengawasan menggunakan profil ^{1}H-NMR madu kelulut dapat membezakan madu kelulut yang tulen daripada madu kelulut campurankecuali dua sampel. Analisis diskriminasi dengan pengawasan dua-kelas menunjukkan keupayaan meramal yang baik (79.20%) dan tiga-kelas menunjukkan keupayaan meramal sederhana (58.90%) bagi membezakan madu kelulut yang tulen daripada madu kelulut campuran, termasuk diskriminasi madu kelulut yang telah dicampur dengan bahan campurpalsu diperolehi daripada tumbuhan C_3 dan C_4 berbeza. Kesimpulannya, proses penyingkiran air dari madu kelulut perlu dilakukan dengan waspada untuk mengelakkan perbezaan nilai isotop karbon berlaku. Bagi profil metabolik madu kelulut, setiap bahagian ^{1}H-NMR metabolik berpotensi menjadi penanda yang penting mengesan sebarang bahan campurpalsu yang baharu.

ADULTERATION DETECTION OF STINGLESS BEE HONEY USING UNTARGETED ^{1}H-NMR METABOLOMICS IN ADDRESSING THE LIMITATION OF STABLE CARBON ISOTOPE RATIO ANALYSIS

ABSTRACT

As stingless bee honey (SBH) is gaining in popularity in the Malaysian market, it is now prone to adulteration. Some studies have found that the nutritional values of SBH may be similar if not more than honey bee honey (HBH), The higher price of SBH as compared to HBH has led to the use of unusual adulterants such as vinegar and even HBH to mimic the unique taste and appearance of SBH. To determine the authenticity of honey in general, the current internationally accepted method used is the Association of Official Analytical Chemists (AOAC) Internal Standard Carbon Isotope Ratio Analysis (ISCIRA) method. This method requires protein to be extracted from the honey. However, the high moisture content found in SBH lead to the lesser amount of protein to be extracted per gram. Additionally, the AOAC ISCIRA method fails to detect adulterants derived from C_3 plants since SBH adulterated with these adulterants will be in the same $\delta^{13}C$ value range. Hence, untargeted proton-nuclear magnetic resonance (^{1}H-NMR) metabolomic with chemometrics was developed to create metabolic fingerprints of SBHs to detect SBH adulterated with adulterants derived from both C_3- and C_4 plants. To determine the authenticity of SBH prior to creating the metabolic fingerprints, the AOAC ISCIRA method was used. Lyophilization process was introduced to decrease the moisture content of SBH. It was found that once moisture content of SBH samples was decreased to a level below 20%, the percentage increment of protein extracted from the samples varied between 6 to 385% in relative to protein extracted from SBH before lyophilization. With that, nine

of the 22 SBH samples used in this study were found to be adulterated. Once the authenticity of each SBH was determined, the SBH samples were analysed with ^{1}H-NMR metabolomic with chemometrics. The unsupervised principal component analysis of SBH ^{1}H-NMR fingerprints was able to distinguish commercial authentic SBHs from adulterated ones with two misclustered samples. Both two-class and three-class supervised discriminant analysis showed good (79.20%) and moderate (58.90%) predictive ability, respectively in distinguishing authentic SBHs from adulterated ones, including discriminating SBHs adulterated with different adulterants derived from C_3 and C_4 plants with 100% accuracy. In conclusion, the removal of water from SBH is required but caution is necessary as carbon isotopic shifts were observed. Meanwhile, any ^{1}H-NMR metabolic fingerprint region of SBH could potentially be crucial as markers for any emerging adulterants that can be used to adulterate SBHs.

CHAPTER 1

INTRODUCTION

1.1 Background

Stingless bee honey (SBH), produced by stingless bees (*Melipona* tribe known for its unique sour taste and odour, unlike the honey bee honey (HBH) produced by honey bees (*Apis mellifera*) (Chan et al., 2017). SBH, which is less murky as compared to HBH, is attributed to have a much higher moisture content (Chan et al., 2017; Fatima et al., 2018). SBH is stored naturally in egg-shaped pots formed from beeswax and plant resins (cerumen), which may contribute to its high nutritional value which some studies have found to be similar if not better than that of common HBH (Amin et al., 2018; Kek et al., 2016; Vit et al., 2011). SBH has been proven to have anti-inflammatory, antibacterial and anticancer properties on top of many more benefits to health (Moniruzzaman et al., 2013; Rao et al., 2016). Having that in mind, consumers may consume honey in the hope of managing their disease or keeping a healthy lifestyle, unbeknownst of the authenticity of the honey. Consuming non-authentic honeys could potentially cause more health issues rather than managing them. Adulteration of honey is not a new global issue and efforts to combat it were made even up to recent years (Awram, 2020; García, 2018; Phipps, 2020).

To combat the adulteration of honeys, many techniques have been implemented. These include measuring the physicochemical and the chemical composition of the honey. The physicochemical property is the intrinsic physical and chemical characteristics of a substance such as measurement of the moisture content and sugar content while chemical composition is the chemical compounds that make up a particular substance which requires the use of analytical techniques and instruments. According to the Codex Standard for Honey (Codex Alimentarius

Commission, 2001), the moisture content for honey should be less than 20.00% while the sum of fructose and glucose content should be less than 60.00%. However, the moisture content of SBH have been found to range from 21.40 to 33.70% (Chan et al., 2017; M. M. Ismail & Ismail, 2018), while the total sugar content of SBH is in the range of 51.00 ± 21.00% (Ávila et al., 2018). Analytical technique was used whereby more targeted compounds such as the amount of 5-hydroxymethylfurfural (HMF) values and sugar derivatives were quantified and measured as a potential marker for adulterated honey (Ávila et al., 2018; Se et al., 2019). However, these compounds may not be present if unknown arising adulterants were used. The current internationally approved method for honey adulteration detection, which is the Association of Official Analytical Chemists (AOAC) internal standard carbon isotope ratio analysis (ISCIRA) method is highly reliable in detecting honey adulterated with adulterants derived from C_4 plants, and yet, this method was not able to detect honey adulterated with adulterants derived from C_3 plants (AOAC, 2005; Tosun, 2013; Zábrodská & Vorlová, 2015). Thus, untargeted proton nuclear magnetic resonance (^{1}H-NMR) metabolomic with chemometrics is proposed whereby the metabolic fingerprints of both authentic and adulterated SBH will be collated first before building classification and prediction models using chemometrics.

1.2 Problem statement

The current internationally approved method for honey adulteration detection, which is the AOAC internal standard carbon isotope ratio analysis (ISCIRA) is not able to detect honeys adulterated by ingredients or syrup derived from C_3 plants (AOAC, 2005; Tosun, 2013; Zábrodská & Vorlová, 2015). This is due to the natural isotope carbon-13 value of the adulterants derived from C_3 plants whereby δ^{13}C of

these adulterants are in the range of $\delta^{13}C$ value for authentic honey. Moreover, SBH has high moisture content which contributed to the total mass of honey, thus decreasing the percentage of protein extracted per gram to be analysed using ISCIRA method (AOAC, 2005).

The physicochemical characteristics of SBH can be ambiguous at times due to external factors such as environmental factor and geographical origin which may affect the authenticity of honey. Thus, this has open up options for food fraudsters to adulterate SBH but escaping the conventional methods of detection such as measuring the physicochemical properties of SBH or even chromatography techniques such as gas chromatography or liquid chromatography. Other available methods targeted on specific compounds as markers for adulterated honeys do not account for external factors that may affect the authenticity of honey, let alone detect any possible arising adulterants used in future adulteration activities of SBH (Se et al., 2019).

1.3 Objectives

The main objective of this research work is to detect adulteration of SBH using untargeted ^{1}H-NMR metabolomic with chemometrics to address the limitation of current available methods. Within the broad overall aim of the study, the following specific objectives were identified:

i. To determine the effect of lyophilization on the $\delta^{13}C$ value and the adulteration percentage of SBH using the AOAC ISCIRA method.

ii. To develop untargeted ^{1}H-NMR metabolomic method to fingerprint authentic and adulterated SBHs.

iii. To differentiate authentic SBHs from SBHs adulterated with adulterants derived from C_3 and C_4 plants using ^{1}H-NMR metabolomic with chemometrics.

1.4 Structure of subsequence chapters

There are five chapters in this thesis, including the introduction in Chapter 1. Chapter 2 discussed on the literature reviews available. This includes the literature search on the current issue with the adulteration of honey and the available method to detect the adulteration of honey. Then, the theory of lyophilization is highlighted briefly before introducing NMR metabolomic and chemometrics.

Chapter 3 highlights the materials used in this study and the overall method to determine the authenticity of the SBH used. This includes the layout of how lyophilization and protein extraction are conducted for prior to the AOAC ISCIRA method. Following this, the sample extraction method development and the metabolite extraction of SBH were outline prior to the description on the NMR instruments optimization and chemometrics.

In Chapter 4, the results on the authentication of SBH used in this study was first laid out. Then, the discussion is separated into two major parts. The first major part of the discussion is the discussion on the effect of lyophilization on the SBH in terms of the moisture content, the mass of protein extracted, the SCIRA and ISCIRA values before discussing the final effect of lyophilization on the adulteration percentage of SBH. Following this, the findings of ^{1}H-NMR metabolomic fingerprints of the SBH and the chemometrics predictive models built are discussed at length and thoroughly. Chapter 5 presents the overall conclusions arising from this study and suggestions for future work that could be carried out.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Honey bee honey (HBH) produced by the honey bees (*Apis* sp.) has long been sold in Malaysia as a premium food supplement. In general, honey can be categorized into blossom honey (also known as floral or nectar honey) and honeydew honey (Codex Alimentarius Commission, 2001; Pita-Calvo & Vázquez, 2017). Blossom honey is the honey which comes from nectars of plants while honeydew honey is the honey which comes mainly from excretions of plant sucking insects (Hemiptera) on the living parts of plants or secretions of living parts of plants (Codex Alimentarius Commission, 2001). Blossom honey can be further categorised into monofloral honey and unifloral honey. Monofloral honey is honey originating from a single type of flora while multifloral honey is honey originating from multiple floral sources (Ismail et al., 2021; Lim et al., 2019; Moniruzzaman et al., 2014). The few common HBHs produced in this tropical country are Acacia honey (monofloral), pineapple honey (monofloral honey), Tualang honey (multifloral) and Gelam honey (multifloral) (Chan et al., 2017; Ismail et al., 2021; Moniruzzaman et al., 2014).

Recently, stingless bee honey (SBH) is gaining is becoming popular among Malaysian consumers (Saludin et al., 2019). . There are 17 species of stingless bees found in various location of Peninsular Malaysia, 12 species in Sabah, and 5 or more different species in Sarawak (Eltz et al., 2003; Kelly et al., 2014; Nagamitsu & Inoue, 2002; Salim et al., 2012). Nevertheless, stingless bees or locally known as ‘kelulut’, are commonly classified into two genera, namely, the *Melipona* sp. and the *Trigona* sp.

In terms of physicochemical properties of honey, SBH can be very different from that of the HBH. The moisture content of SBH can be higher as compared to that

of HBH. The moisture content of SBH ranged from 13.30% to 56.30% depending on the geographical origins, as compared to that of HBH which ranged from 15.00% to 21.00% (Ávila et al., 2018; da Silva et al., 2016; B. Souza et al., 2006). While the sugar content of HBH ranged from 80.00% to 83.00% on average (Santos-Buelga & González-Paramás, 2017), the sugar content SBHs varies from 50.00% to 98.80% (Ávila et al., 2018). The protein content of SBHs ranged from 1.20% to 3.10% (Ávila et al., 2018), whereas the protein content of HBH ranged from 0.20% to 3.30% (da Silva et al., 2016; Santos-Buelga & González-Paramás, 2017). The high variation of these physicochemical properties in SBH as compared to HBH can be dependent on the species of the bee and the environmental factors (Ávila et al., 2018; Mohammed, 2020). Commonly, HBHs are mostly obtained from honey produced by honey bee of two *Apis* genera particularly the *Apis mellifera* and *A. cerana* (Alvarez-Suarez, 2017; da Silva et al., 2016; Mohammed, 2020). On the contrary, SBHs are obtained from honey produced by various genera of stingless bee such as *Heterotrigona* sp., *Melipona* sp., *Tetragonula* sp. and the *Trigona* sp., depending on the geographical origin of the SBH (Ávila et al., 2018). Thus, it is possible that SBHs are more likely to be affected by environmental factors since SBHs are obtained from a wider range of stingless bee species. These environmental factors including the humidity, floral source and geographical origin. The moisture content of honeys, in general, tends to be higher in area with high humidity or during rainy season (Chan et al., 2017; da Silva et al., 2016; Fatima et al., 2018; Mohammed, 2020). The sugar profile and the protein content of the honeys varies depending on the flora and vegetation that predominates in that region (Ávila et al., 2018; Nordin et al., 2018; Se et al., 2018).

Unlike the HBH produced by honey bees which is stored in honeycomb hives, SBH is stored naturally in egg-shaped pots formed from beeswax and plant resins

(cerumen) (Amin et al., 2018; Kek et al., 2016; Vit et al., 2011). The unique way of storing the SBH by the stingless bee may contribute to its high nutritional value which some studies have found to be similar, if not better than that of common HBH (Amin et al., 2018; Kek et al., 2016; Vit et al., 2011). These unique properties have also given additional farming advantage of stingless bee whereby other by-products such as beebread and propolis can be harvested and sold (Ismail & Ismail, 2018). Like HBH, SBH has been proven to have anti-inflammatory, antibacterial and anticancer properties on top of many more benefits to human health (Moniruzzaman et al., 2013). This has prompted the Malaysian government to promote SBH as a superfood due to its nutritional value, and efforts are being made for it to be exported and marketed overseas (Hamid, 2019). The price of SBH can be as high as US$100/kg; in comparison, HBH (excluding manuka honey) is being sold at $20–40/kg on average (Shadan et al., 2018).

2.2 Setting the quality of honey

To ensure the good quality of honey in the international market, there are a few standards to comply with, namely Codex Standard for Honey (Codex Alimentarius Commission, 2001), International Honey Commission (IHC) (Bogdanov, 2009) and the European Union (EU) Standard for Honey (European Commission, 2002). These standards are very similar to each other. These includes the measurement of moisture content, the sugar content and the (HMF) content of honey.

According to the Codex Standard for Honey (Codex Alimentarius Commission, 2001), the IHC (Bogdanov, 2009) and the EU Standard for Honey, the moisture content for honey should be less than 20.00%. The moisture content of honey can be determined by using refractometric method (Bogdanov, 2009; Codex Alimentarius Commission, 2001; European Commission, 2002). In terms of sugar content in honey, the amount of

the sugar can be determined by using either high-performance liquid chromatography (HPLC) or gas chromatography (GC) (Bogdanov, 2009; Codex Alimentarius Commission, 2001; European Commission, 2002). The sum of both fructose and glucose should not be less than 60.00% whereas the amount of sucrose should not be more than 5.00% (Bogdanov, 2009; Codex Alimentarius Commission, 2001; European Commission, 2002). The amount of HMF in honey is also determined using HPLC, whereby the amount of HMF should not exceed 40 mg/kg (Bogdanov, 2009; Codex Alimentarius Commission, 2001; European Commission, 2002).

However, SBH may not comply to these parameter set by the Codex Standard for Honey (Codex Alimentarius Commission, 2001), the IHC (Bogdanov, 2009) and the EU Standard for Honey (European Commission, 2002). The moisture content of SBH ranged from 13.30% to 56.30% while the sugar content SBHs varies from 50.00% to 98.80% (Ávila et al., 2018; Nordin et al., 2018). The HMF values also varies in SBH where the highest HMF value recorded in Brazil was 51.38 mg/kg (Ávila et al., 2018). These parameters may be affected by environmental factors such as humidity and temperature (Chan et al., 2017; da Silva et al., 2016; Fatima et al., 2018; Mohammed, 2020). In addition, improper storage of the honey may also cause increase in the HMF value where the exposure of honey to heat may cause the decomposition of fructose to HMF (Fatima et al., 2018; Kek et al., 2016). With that, SBH are often subjected to accusation of adulteration due to their high HMF content (Nordin et al., 2018).

The physicochemical properties is the intrinsic physical and chemical characteristics of a substance such as measurement of the moisture and sugar contents (Mohammed, 2020). Looking at the SBH produced locally in Malaysia (Chan et al., 2017; Ngah, 2016), the physicochemical properties of Malaysian SBH may not comply with the standards set by the Codex Standard for Honey (Codex Alimentarius

Commission, 2001), the IHC (Bogdanov, 2009) and the EU Standard for Honey (European Commission, 2002). Taking that into consideration, Malaysia government has set a Malaysian Standard specifically for SBH (MS 2683: 2017). According to the standard, the few quality requirements for SBH is that the moisture content should not be more than 35.00%, the total sugar content (sum of fructose and glucose) should not be more than 85.00% and the HMF value should not exceed 30.0 mg/kg. Based on the work done by Abu Bakar et al. (2017), Chan et al. (2017), Fatima et al. (2018) and Ranneh et al. (2018), Malaysian SBHs do comply with the Malaysian Standard for SBH (MS 2683: 2017). The moisture content of Malaysian SBH ranged from 21.40% to 33.70% (Chan et al., 2017; Fatima et al., 2018). The total sugar content of SBH ranged from 68.10% to 73.01% (Abu Bakar et al., 2017; Ranneh et al., 2018). The HMF values of Malaysia SBH may varies from 0.08 mg/kg to 3.42 mg/kg (Fatima et al., 2018).

As SBH is gaining its traction in different country's market, many research has been done to set a new parameter for SBH since its properties are different from the HBH (Ávila et al., 2018; Nordin et al., 2018). Having said that, by measuring these physicochemical parameters alone does not fully determine the authenticity of SBH (Chan et al., 2017; Fatima et al., 2018).

2.3 Methods to decrease the moisture content of honey

According to the Malaysian Standard for SBH (MS 2683: 2017), the allowable moisture content for SBH is not more 35.00%. However, the limitation for moisture content of honey set in the Codex Standard for Honey (Codex Alimentarius Commission, 2001) is that the moisture content does not exceed 20.00%, and thus, it is still a requirement for SBH to fulfil the standard if it was to be marketed internationally. Additionally, the high moisture content of SBH may decrease the amount of protein

extracted for the honey authentication using ISCIRA method as the presence of water contributes unnecessary weight to the sample prior to protein extraction. Therefore, there is a need to reduce the moisture content in SBH.

Various technologies have been developed to reduce the moisture content of HBH instead. The purpose of the developed technologies is to reduce the moisture content of HBH so that the HBH fulfil the Codex Standard for Honey in terms of moisture content (Bogdanov, 2016; Singh & Singh, 2018). These include heating at lower temperatures using water bath or on electric plates (Büdel & Grziwa, 1959a, 1959b), heating at higher temperatures (for pasteurisation of honey) using the application of different kind of waves (Ivanov & Ivanova, 1995; Kaloyereas & Oertel, 1958; Liebl, 1978) and dehumidification whereby the hive was placed in a warm room together with a dehumidifier before harvesting the honey (Bogdanov, 2016). To reduce the moisture content for larger amount of honey, technologies such as 'rotating discs type honey moisture reduction system' (Platt & Ellis, 1984), 'rotating cone type honey moisture reduction' (Wakhle et al., 1988) and 'desiccant honey dehydrator system' (Singh et al., 2011) have been introduced. All these technologies have in common is that these technologies utilized heat to reduce the moisture content of HBH (Singh & Singh, 2018). However, introducing heat at different temperatures during the processing of honey may or may not affect the parameter of the honey, if the process is not controlled (Bogdanov, 2016; Eshete & Eshete, 2018; Singh & Singh, 2018). There are a few technologies known to reduce the moisture content of SBH. These includes Honey Interlinked Dehydration and Dispenser Apparatus (HILDA) system (Mustafa et al., 2018), low temperature vacuum drying (Ramli et al., 2017) and using a food dehydrator (Yap et al., 2019). The low temperature vacuum drying method used vacuum and heat from water at 30ºC to reduce the moisture content of SBH (Ramli et al., 2017). HILDA

system utilizes a semi-automated platform combining honey harvesting, dehydration and dispenser in a close system (Mustafa et al., 2018). It has integrated elements of working bench, bottle drying rack and honey suction pump with honey dehydration component and dispenser unit, whereby the reduction of SBH's moisture content was done at the temperature below 36°C (Mustafa et al., 2018). However, there is still pending for the patent of this system.

To avoid the use of heat, lyophilization which is commonly known as freeze drying was introduced to reduce the moisture content from a frozen material (Bhambere et al., 2015; Meirowitz, 2019). Lyophilization is the process whereby water is removed from a product after it is frozen and placed under a vacuum, allowing the ice to change directly from solid to vapor without passing through a liquid phase (sublimation). Lyophilization consists of three stages; i.e. freezing, primary drying and secondary drying (Meirowitz, 2019). The first stage of lyophilization which is freezing is the stage where a liquid sample must be frozen completely before proceeding to the next stage (Bhambere et al., 2015; Meirowitz, 2019). In the second stage of lyophilization, the primary drying begins when the pressure in the freeze dryer starts to decrease, subliming the frozen moisture in the sample (Meirowitz, 2019). Primary drying is completed when there is no free water in the sample, and the sample will look dry. However, residue of moisture is still bound in the sample. So, once primary drying is completed, lyophilization process will proceed to the third stage of drying whereby secondary drying will begin. Secondary drying is the process where the bound water molecules will be removed completely (Meirowitz, 2019). By the end of secondary drying, the final product will usually appear as powdered form (Bhambere et al., 2015; Meirowitz, 2019).

The main advantage of lyophilization is that this process does not involve high heat, therefore will not cause denaturation of proteins in SBHs and at the same time, able to remove the moisture content in SBHs (Bhambere et al., 2015). Hence, a water removal step using lyophilization is proposed with a dual purpose 1) to decrease the moisture content to a level below 20.00%; and 2) to increase the amount of protein extracted for the authentication of SBH using the ISCIRA method as the presence of water contributes unnecessary weight to the sample prior to protein extraction. Since the objective of this study is to reduce the moisture content of SBH down to 20.00%, the duration of lyophilization needs to be controlled so that the SBH will not be overdried or to the point of secondary drying.

2.4 Current issues on adulteration activity of honey

According to the International Honey Market report by the *American Bee Journal*, the adulteration of honey is not a new international issue and efforts to combat it were made as early as 1889 by the US Department of Agriculture (Awram, 2020; Phipps, 2020). This major issue occurs not only in America and European countries, but also in countries in Asia including Malaysia (García, 2018; Phipps, 2020). The amount of honey exported is more than the honey produced and the possibility of this 'never-ending supply' of honey may be due to production of adulterated honey (Phipps, 2020). Locally in Malaysia, it has been reported that around 80.00% of honey sold has been adulterated (Osman et al., 2017). These adulterated honeys, also known as 'diabetic honey' in Malaysia, are detrimental to health, especially to those who have, or have had diabetes (Osman et al., 2017). Sugar syrups derived from corn and sugarcane are commonly used as adulterants. These adulterants are cheaper and have even less nutritional value in comparison to the honeys sold in Malaysia.

There are common beliefs that the ways to differentiate adulterated honeys from authentic ones are by looking at the appearance of the honey, smelling its odour and tasting it. Authentic honey is believed to have a flower-like odour, a natural sweet taste and more murky appearance (due to its active compounds) in comparison to adulterated honey which has a more transparent appearance and sugar-like odour with a very sweet taste (Sahib, 2020). However, these beliefs are not applicable to SBHs. SBH which is produced by stingless bees (*Melipona* tribe and *Trigona* tribe) is known for its unique sour taste and odour, unlike the HBH produced by honey bees (*Apis mellifera*) (Chan et al., 2017). SBH is less murky than HBH, as though it was adulterated, which can be attributed to the higher moisture content of SBH as compared to HBH (Chan et al., 2017; Fatima et al., 2018). Thus, due to the different physicochemical characteristics of SBH, the common belief that adulterated honeys can be differentiated from authentic ones by looking at the physical characteristics is incorrect; the characteristics may be superficial. This has motivated unscrupulous food fraudsters to resort to adulterating SBH in order to make quick profits. The physicochemical characteristics of SBH can be ambiguous, which has provided options for food fraudsters to adulterate SBH. The most common adulterants used in SBH are glucose syrup, fructose syrup and HBH (Tosun, 2013). Vinegar has also been used as an additive to cheaper HBH to mimic the sour taste of SBH, and this may be sold as counterfeit SBH or used to adulterate SBH to increase volume with the cheaper adulterant (Mail et al., 2019). Selling these adulterated SBHs is not only deceitful and unfair to the consumers at large, but can cause health issues to some as well.

2.5 Current techniques on adulteration detection of honey

To address the issue on the adulteration of honey, there are a few techniques introduced to detect the adulteration of honey. These includes the AOAC internal standard carbon isotope ratio analysis, hyphenated chromatographic techniques and spectroscopic techniques.

2.5.1 Stable carbon isotope ratio analysis (SCIRA) and internal standard carbon isotope ratio analysis (ISCIRA)

The current internationally accepted method used to detect honey adulteration is the AOAC internal standard carbon isotope ratio analysis (ISCIRA) (AOAC, 2005). This method was introduced by Doner and White (1977) and has been modified over recent years (White & Winters, 1989; White et al., 1998; Rogers et al., 2013). Honey adulteration is detected by measuring the $\delta^{13}C$ value of the protein extracted from the whole honey using the ISCIRA method and comparing it with the $\delta^{13}C$ value of the whole honey measured by stable carbon isotope ratio analysis (SCIRA) (Padovan et al., 2003, 2007; White & Winters, 1989). To understand how this method works, it is necessary to look into the carbon fixation metabolic pathway of plants during photosynthesis.

During photosynthesis, plants will produce glucose using either the Calvin Cycle (C_3 cycle) or the Hatch and Slack Pathway (C_4 cycle), excluding the crassulacean acid metabolism (CAM) pathway. Most of the flowering plants such as sunflower produce glucose using the C_3 cycle while most plants are non-flowering such as corn or sugarcane produce glucose using the C_4 cycle. The first stable product in C_3 cycle is a 3-carbon compound phosphoglyceric acid (PGA) and hence the name C_3 plants, while for the C_4 cycle, the first stable product is a 4-carbon compound oxaloacetic acid (OAA)

and hence the name C_4 plants. The glucose produced by C_3 plants have the $\delta^{13}C_{(C3\ sugar)}$ value ranged from -22‰ to -33‰ while that of the C_4 plants have the $\delta^{13}C_{(C4\ sugar)}$ value ranged from -10‰ to -20‰ (Padovan et al., 2007).

When bees forage the sugar (in the form of nectar) of C_3 flowers and metabolise it to honey, the $\delta^{13}C_{(honey)}$ value, remains in the same range of C_3 plants which ranged from -22‰ to -33‰. Since honey is foraged from flower nectar, this value affects both the sugar and the proteins in the honey. So, once the protein is extracted from the honey, the $\delta^{13}C_{(protein)}$ value is in the same range of $\delta^{13}C_{(honey)}$. If a honey is adulterated with sugar syrup, where the common sugar syrup is made of C_4 plants such as sugarcane or corn syrup, the $\delta^{13}C_{(honey)}$ will be affected and it will be in the range close to $\delta^{13}C_{(C4\ sugar)}$. Since the sugar syrup mainly contains only sugars, the protein in the honey is not affected by these adulterants. Hence, the $\delta^{13}C_{(protein)}$ value will remain in the range of $\delta^{13}C_{(C3\ sugar)}$. By comparing the $\delta^{13}C_{(honey)}$ with $\delta^{13}C_{(protein)}$ using the Equation 1.1 developed by White & Winters, (1989), the adulteration percentage of honey can be determined.

$$\text{Adulteration (\%)} = \frac{\delta^{13}C_{(protein)} - \delta^{13}C_{(honey)}}{\delta^{13}C_{(protein)} - \delta^{13}C_{(sweetener)}} \text{ x } 100 \qquad \text{Eq. 1.1}$$

Since this calculation has been used to obtain adulteration percentage for HBH whereby the $\delta^{13}C_{(honey)}$ range which had been determined to be between -22‰ to -33‰ was used in the calculation, the $\delta^{13}C_{(honey)}$ of Malaysian SBH range needed also be determined in order to assess the applicability of the said calculation to detect adulteration activities in our locally produced SBHs. Hence, SBHs from different origins in Malaysia were collected between 2014-2015 to determine the range and the $\delta^{13}C$ of pure SBHs were found to be in the range of -26.33‰ to -28.29‰ (Asmadi, 2016; Ngah, 2016) as shown in Table 2.1. Additionally, the protein extracted from these

pure SBHs using AOAC ISCIRA method had $\delta^{13}C$ ranged from -26.27‰ to -28.93‰ (Asmadi, 2016; Ngah, 2016). Once the Equation 1.1 is used, the adulteration percentage were found to be less than 7%, which is in accordance to the threshold limit of authentic honey set by AOAC (2005). Therefore, this calculation can be applied to determine the authenticity of SBH as well.

Table 2.1 $\delta^{13}C$(honey) and $\delta^{13}C$(protein) values of pure SBHs adapted from Asmadi (2016) and Ngah (2016).

Sample origin	**Source**	**Honey Quality**	**$\delta^{13}C_{(honey)}$ ± Std. Dev. (‰)**	**$\delta^{13}C_{(protein)}$ ± Std. Dev. (‰)**	**Adulteration percentage (%)**	**Reference**
Segamat	SBH Farm	Pure	-26.51 ± 0.12	-26.69 ± 0.18	1.06	(Asmadi, 2016)
Paloh	SBH Farm	Pure	-26.33 ± 0.07	-26.70 ± 0.49	2.18	(Asmadi, 2016)
Merbok	SBH Farm	Pure	-26.67 ± 0.21	-26.27 ± 0.28	-2.41	(Asmadi, 2016)
Pekan Nenas	SBH Farm	Pure	-26.34 ± 0.13	-26.58 ± 0.10	1.42	(Asmadi, 2016)
Pontian	SBH Farm	Pure	-28.06 ± 0.04	-28.33 ± 0.06	1.45	(Asmadi, 2016)
Kemaman	SBH Farm	Pure	-28.29 ± 0.72	-28.93 ± 0.35	3.33	(Asmadi, 2016)
Segamat	SBH Farm	Pure	-27.19 ± 0.25	-26.68 ± 0.22	-3.00	(Ngah, 2016)
Paloh	SBH Farm	Pure	-26.90 ± 0.08	-26.70 ± 0.08	-1.18	(Ngah, 2016)
Merbok	SBH Farm	Pure	-27.60 ± 0.10	-26.27 ± 0.27	-8.03	(Ngah, 2016)

The drawback of this technique is that it fails to detect honeys adulterated with ingredients or syrup derived from C_3 plants such as beet sugar or, in the case of SBH, cider vinegars and HBH (Bogdanov & Martin, 2002; Mail et al., 2019; Thomas & Jamin, 2009; Tosun, 2013; Zábrodská & Vorlová, 2015). This is because when syrup derived from C_3 plants is used, the affected $\delta^{13}C_{(honey)}$ will still be in the range of $\delta^{13}C_{(C3\text{-}sugar)}$ (Tosun, 2013). Eventually, the adulteration percentage will be close to zero when comparison of $\delta^{13}C_{(honey)}$ with $\delta^{13}C_{(protein)}$ were made since the $\delta^{13}C_{(protein)}$ will also be in the range of $\delta^{13}C_{(C3\text{-}sugar)}$. Therefore, this method has been used effectively to detect honey adulterated by sugars refined from C_4 plants due to the difference in the natural abundances of carbon-13 and carbon-12 in C_3 and C_4 plants and the resulting differences in their ratio when C_4 sugars are added to natural honey but fails when C_3 sugars are added as adulterants.

Apart from that, it has been reported that SBH may lacks the protein to be analysed using the ISCIRA method (Abu Bakar et al., 2017). Abu Bakar et al. (2017) reported that the protein content of SBH determined by Kjeldahl method ranges from 0.09% to 0.31%, unlike the protein content in normal HBH which ranges from 0.10% to 3.30%, depending on the species of the bee, as reported by da Silva et al., (2016). This may be due to the high moisture content in SBH. It has been reported that the moisture content of SBH in Malaysia ranges from 21.40% - 33.70% (Chan et al., 2017; Fatima et al., 2018). The moisture content of SBH does not fulfil the allowable moisture content of honey set by the Codex Standard for Honey which is below 20.00%, unlike the common HBH (Chan et al., 2017; Codex Alimentarius Commission, 2001; Fatima et al., 2018). The high moisture content of SBH not only contributes unnecessary weight to the sample prior to protein extraction, it also tends to cause the spoilage of honey by

fermentation (Akbulut et al., 2009; Chirife et al., 2006; Mustafa et al., 2018). The yeast responsible for fermentation that occurs naturally in honey is *Saccharomyces* spp., which represents the dominant yeast found and other genera used water in the honey for fermentation (Chirife et al., 2006; Snowdon & Cliver, 1996). In other words, the higher moisture content contributes to a higher extent of honey fermentation (Zuccato et al., 2017). Since SBH has high moisture content (%), it is more susceptible to fermentation and thus causing the honey to taste sour naturally (Zuccato et al., 2017). Therefore, reducing the moisture content may increase the amount of protein extracted for the honey authentication using ISCIRA as the presence of water contributes unnecessary weight to the sample prior to protein extraction, and at the same time limits the fermentation of SBH.

2.5.2 Hyphenated chromatographic techniques

As technology advance, more research has been conducted on different techniques to combat with the adulteration of honey. Since the AOAC ISCIRA method is only limited to detect honey adulterated with C_4 sugars, different hyphenated chromatographic techniques are also being explored to detect the adulteration of honey with different type of sugars (Se et al., 2019).

Hyphenated chromatography techniques such as GCMS and HPLC coupled with diode array detection (HPLC-DAD) system are highly versatile technique for detecting sugar adulterants in honey (Se et al., 2019). A method using GCMS has been developed by Ruiz-Matute et al. (2010) to detect honey adulterated with high fructose inulin syrups. Since inulin syrups have different degrees of polymerization of sugar derivatives, it is found that inulotriose provided the marker of honey adulteration with as low as 5.00% (w/w) high fructose inulin syrups (Ruiz-Matute et al., 2010). On the

other hand, the HPLC-DAD system developed by Xue et al. (2013) was able to detect honey adulterated with rice syrup (a sugar syrup derived from C_3 plant) down to 10.00% adulteration level. Based on the work by Xue et al. (2013), it was found that honey adulterated with rice syrup contained 2-acetylfuran-3-glucopyranoside, thus having it as a potential marker.

Although both of these techniques highly sensitive and able to detect honey adulterated with sugar syrup derived from C_3 plants, these techniques are laborious and consume high amount of standard analytes (Se et al., 2019). In addition to that, both of these techniques focused only on specific type of adulterants used on HBH only and therefore, focusing only in the sugar profile of HBH.

2.5.3 Spectroscopic techniques

To attune the high pace of the honey international market, rapid detection honey adulteration through spectroscopic techniques coupled with chemometrics has been developed. These includes Fourier-transform infrared spectroscopy (FTIR), near-infrared spectroscopy (NIR), ultraviolet–visible spectroscopy (UV-Vis) and NMR (Se et al., 2019).

A UV-Vis method coupled with chemometrics developed by de Souza et al. (2021) was able to detect adulterations of HBH with corn syrup, agave syrup and sugarcane molasses. Meanwhile, Valinger et al. (2021) combined both UV-Vis and NIR, coupled with chemometrics, showed the potential of the combined methods in detecting HBH adulterated with corn syrup. On the other hand, NMR has been used in the work by Spiteri et al. (2015) and it has been able to detect HBH adulterated with several industrial sugar syrups of various types and sources. Yet again, the work done by de Souza et al. (2021), Valinger et al. (2021) and Spiteri et al. (2015) still focused

only on the adulteration of HBH using sugar syrup. Looking at the more recent work on adulteration of SBH instead of HBH, a rapid detection and quantification of adulterants in SBH using FTIR combined with chemometrics was successfully developed by Se et al. (2018). This method was successfully detect SBH adulterated with corn syrup above 8.00% (w/w) and cane sugar over 2.00% (w/w) (Se et al., 2018).

In short, most of the work on honey adulteration focus heavily on HBH instead of SBH. At the same time, only sugar-derived compounds have been used as adulterants in most of the research work, thus, limiting the scope of potential markers only within sugar compounds, whereby most of these adulterants can be detected using the AOAC ISCIRA methods.

2.6 NMR metabolomics

Metabolomics is an emerging method for food analysis and authentication (Cubero-Leon et al., 2014; Wishart, 2008)(Cubero-Leon et al., 2014; Wishart, 2008). The method is a scientific field to identify and quantify cellular metabolites by combined strategies using sophisticated analytical technologies with the application of statistical and multivariate methods (chemometrics) for information extraction and data interpretation (Roessner & Bowne, 2009; Roussel et al., 2014). Metabolomics has been widely used in food science, nutrition research and food authentication (Cubero-Leon et al., 2014; Wishart, 2008). Currently, mass spectrometry (MS) and NMR are the go-to instruments in metabolomics. Although MS has its own advantages over NMR, NMR is non-destructive, unlike MS. The data produced by NMR are highly reproducible and quantitative over a wide range. NMR spectroscopy has been used extensively for multivariate metabolite profiling and metabolic fingerprinting (Beckonert et al., 2007; Hong et al., 2017; Markley et al., 2017). Metabolomics using NMR requires little effort

for sample preparation and is able to measure as many metabolites as possible without tampering with their integrity (Esslinger et al., 2014).

Metabolomic analyses are generally classified into targeted or untargeted method. Targeted metabolomic method measures a specified list of metabolites, focusing on one or more related pathways of interest (Cevallos-Cevallos et al., 2009; Patti et al., 2012). Targeted method has been used as one of the methods to determine the authenticity of honey. For instance, HMF, amino acid and ethanol are the common targeted metabolites as indicators to determine the authenticity of honey using NMR. Usually, authentic honey shows low amounts of HMF, lactic acid and ethanol as they may be formed from the natural fermentation of honey, and high amounts of amino acids will be present (Boffo et al., 2012; da Silva et al., 2016; Spiteri et al., 2015). Prolonged exposure to high temperature may also cause the decomposition of fructose in the presence of acid, forming HMF, thus acting as an indicator to the freshness of the honey (da Silva et al., 2016; Kek et al., 2016). As for adulterated honeys, high amount of HMF will be present as it is formed from the decomposition of high amount of adulterated sugars while amino acids may be absent or present in small amounts (Boffo et al., 2012; da Silva et al., 2016; Kek et al., 2016). Ethanol may present and citric acid may also be added intentionally, probably to act as the so-called 'antioxidant' (Boffo et al., 2012). However, if non-sugar-based adulterant is used, then HMF will not form. On the other hand, untargeted metabolomic method simultaneously measures as many metabolites as possible from biological samples without bias, forming a metabolite fingerprints of the samples (Cevallos-Cevallos et al., 2009; Erban et al., 2019; Esslinger et al., 2014; Patti et al., 2012). In other words, untargeted metabolomics does not exclude any metabolites which can be crucial to detect potential markers for any emerging adulterants used. Untargeted metabolomics approaches to obtain metabolite

fingerprinting are expected to become potent tools for food authentication and discovery of adulteration in food (Erban et al., 2019).

Since untargeted metabolomic approach simultaneously measures as many metabolites as possible from biological samples, sample preparation step is done under a strict and consistent method. Once the biofluid sample is completely dissolved in a suitable solvent, the sample pre-treatment is done such as controlling the pH value and introducing filtration step to remove any insoluble solids before transferring the filtrate to the NMR tube prior to analysis by NMR instrument (Smolinska et al., 2012). One of the main advantages of using the NMR metabolomic is that this method uses the ^{1}H-nuclear Overhauser effect spectroscopy (NOESY) experiment to reduce the solvent's ^{1}H signal in the NMR, thus providing a reproducible and easy-to-implement experiment for the recording of the ^{1}H spectrum of biological samples with a good water suppression (Alonso et al., 2015; Smolinska et al., 2012). Since NMR-based metabolomics typically involves the collection of dozens to hundred (even thousands) of spectra, consistency across experiments is a key factor that must be considered when implementing solvent suppression (Smolinska et al., 2012). Once the spectra have been acquired, the spectra will be pre-processed to obtain the metabolite fingerprints by correcting the baseline, realign all the samples' spectra to a consistent manner, and exclude certain signals such as internal standard signals. Prior to data analysis which includes the use of statistical software or better known as chemometrics to interpret the chemical data, the spectra are subjected to binning (or bucketing) to reduce the dimensionality of data prior to being normalised and transformed to prevent the overt contribution of certain signals with high intensity (Alonso et al., 2015; Erban et al., 2019; Smolinska et al., 2012).

2.7 Chemometrics for interpretation of chemical data

Chemometrics, a technique that uses mathematical and statistical methods to interpret chemical data, has been used with various analytical instruments as part of the data analysis. In chemometrics, analysis begins with an unsupervised pattern recognition technique and the commonly used unsupervised technique which is principal component analysis (PCA), is used to explore possible hidden patterns and relationships of data. Then, the analysis will be followed up by a supervised technique such as orthogonal partial least squares discriminant analysis (OPLS-DA) where each sample is given a classification and the data are projected into this new space, showing the relation between samples and variables (Abdi & Williams, 2010; Granato et al., 2018; Smolinska et al., 2012).

PCA extracts important information through dimension reduction of the dataset and represents the variation present in the dataset using a small number of factors, namely principal components (PCs) (Abdi & Williams, 2010; Granato et al., 2018; Smolinska et al., 2012). The PCs are calculated iteratively to hold as much variation from the original data set as possible, where the first PC (denoted as PC1) explains the greatest data variation followed by the second PC (PC2). In PC2, part of the data is taken out to perform cross-validation with a different number of PCs on the rest of the data. Dimension reduction of the dataset is repeated in PC2 to produce the third PC (PC3) where PC2 explains more data variation than PC3, and the calculation will be repeated until every piece of data has been reduced (Granato et al., 2018). Since PCA only projects or displays the dataset under investigation, it does not create a 'mathematical model' for classification and authentication purposes and thus grouping of samples needs to be identified by the user (Granato et al., 2018).