Quality Evaluation of Stingless Bee Honey by Physicochemical Properties and FTIR-ATR-Chemometric Classification

Siti Nor Azlina Abd Rashid[1,5]*, Muhammad Zulhelmi Nazri[1,4], Muhammad Ammar Md Diah[2], Norliza Abdul Latiff[1], Zainun Nurzahim[3] and Dayang Norulfairuz Abang Zaidel[1,4]

[1]Innovation Centre in Agritechnology for Advanced Bioprocessing (ICA), Universiti Teknologi Malaysia (UTM), 84600 Pagoh, Johor, Malaysia

[2]Politeknik Tun Syed Nasir Syed Ismail (PTSN), Eduhub Tinggi Pagoh, 84600 Pagoh, Johor, Malaysia

[3]Innovation Centre for Confectionery Technology (MANIS), Faculty of Science and Technology, Universiti Kebangsaan Malaysia (UKM) 43600 Bangi, Selangor, Malaysia

[4]Malaysia-Japan International Institute of Technology (MJIIT), Universiti Teknologi Malaysia (UTM), Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia

[5]Innovation and Commercialisation Centre (ICC), Pejabat Timbalan Naib Cancelor (Penyelidikan & Inovasi), Universiti Teknologi Malaysia (UTM), Jalan Sultan Yahya Petra, 54100 Kuala Lumpur, Malaysia

*Corresponding author (e-mail: sitinorazlina@utm.my)

Stingless bee honey (SBH) is prized for its unique composition and health benefits, but its quality varies widely due to differences in bee species, flower sources, and processing techniques. Conventional physicochemical tests provide useful clues but not sufficient for reliable classification and differentiation from *Apis mellifera* honey (BH). The integration of these tests with advanced methods, such as Fourier transform infrared-attenuated total reflectance (FTIR-ATR) spectroscopy, provides a more robust and direct approach for authentication of the honey. SBH exhibited wide variation of moisture (11.33 – 37.33%), pH (2.89 – 4.66), free acidity (10.69 – 35.27 meq/kg), and electrical conductivity (66.63 – 1080.73 µS/cm), with some samples exceeding the Malaysian Standard (MS 2683:2017). Meanwhile, the FTIR-ATR spectra revealed distinct carbohydrate and organic acid bands and chemometric analyses provided robust classification. The principal component analysis (PCA) explained up to 100% variance clearly separating SBH from BH and clustering samples according to floral and compositional differences. Agglomerative hierarchical clustering (AHC) further distinguished three major classes based on physicochemical attributes reflecting differences in mineral content and acidity. The partial least squares-discriminant analysis (PLS-DA) achieved the most effective discrimination in the 3500 – 3000 cm^{-1} region with strong model performance (R^2Y = 0962, Q^2 = 0.684) and 100% correct classification of training samples, suggesting hydroxyl-rich compounds as key discriminators. Collectively, these findings demonstrate that integrating physicochemical tests with FTIR-ATR-chemometric modelling provides a rapid, reliable, and non-destructive framework for classifying and authenticating SBH, supporting the establishment of species-specific quality standards and strengthening consumer confidence in its commercialisation.

Keywords: Authentication, Chemometric classification, FTIR-ATR spectroscopy, physicochemical properties, stingless bee honey

Received: October 2025; Accepted: December 2025

Stingless bee honey (SBH), produced by various *Meliponini* species, has gained significant attention due to its unique composition and associated health benefits [1]. Compared to *Apis mellifera* honey (BH), SBH is typically richer in organic acids, phenolic compounds, flavonoids, and other bioactive constituents that contribute to its antioxidant, antimicrobial, and anti-inflammatory properties [2]. These attributes highlight SBH's potential as a valuable functional food and nutraceutical ingredient [3]. However, the composition of SBH is highly variable, influenced by bee species, floral sources, and geographical factors, making it challenging to establish consistent quality standards [4].

Beyond its nutritional components, several physicochemical properties, such as moisture content, pH, free acidity, electrical conductivity (EC), and colour, are important indicators of SBH's quality, stability, and authenticity. SBH often has higher moisture content and acidity, as well as greater variability in EC and colour, due to its hygroscopic nature and diverse botanical origins [5]. While these characteristics contribute to its distinctive sensory profile, they can also increase its susceptibility to

fermentation and quality deterioration if not properly handled [6]. To address these challenges, Malaysia introduced MS 2683:2017, the first national standard specifically for SBH, which sets limits for moisture, pH, free acidity, and EC to ensure product stability, safety, and consumer confidence [7].

Although physicochemical parameters provide valuable information, they may not offer sufficient discriminatory power for reliable classification and authentication of SBH, given its inherent variability. Modern analytical tools such as Fourier transform infrared (FTIR) spectroscopy provide a rapid, non-destructive method for obtaining chemical fingerprints of honey by detecting vibrations associated with carbohydrates, organic acids, proteins, and minor constituents [8]. When combined with chemometric techniques, such as principal component analysis (PCA) and hierarchical clustering, FTIR data can be translated into meaningful multivariate patterns that differentiate samples according to their chemical composition [9]. This combined approach strengthens both quality assessment and authentication efforts.

Despite increasing research on SBH, many existing studies have examined physicochemical properties or FTIR characteristics separately, resulting in fragmented insights into SBH classification. To address this gap, the present study integrates physicochemical characterisation with FTIR-ATR chemometric analysis to comprehensively evaluate Malaysian SBH in comparison with BH. Specifically, the study aims to determine the physicochemical properties of SBH, identify key FTIR spectral features, and apply chemometric clustering to classify SBH according to compositional attributes. Based on these objectives, we hypothesised that SBH exhibits distinct physicochemical, spectral, and pollen-related characteristics compared to *Apis mellifera* honey, and that integrating physicochemical analysis with FTIR-ATR chemometric modelling enables effective discrimination and classification of SBH samples. By combining these methods, the study provides a more comprehensive framework for quality assessment and supports the development of robust standards and authentication strategies for SBH, thereby promoting marketing and consumer confidence.

MATERIALS AND METHODS

Materials and Chemicals

A total of 9 local SBH samples were obtained from local stingless bee farmers in Pagoh, Muar, Johor, Malaysia. The samples were labelled and coded as follows: *Apis mellifera* (BH), Commercial SBH 1 (SBHC1), Commercial SBH 2 (SBHC2), Commercial SBH 3 (SBHC3), and raw SBH samples from *Tetrigona binghami* (KP), *Heterotrigona erythrogastra* (PM), *Geniotrigona thoracica* (TH), *Homotrigona aliceae* (TP), and *Heterotrigona itama* (IT). The honey samples were collected in sterile plastic containers and transported to the laboratory under refrigerated conditions (4°C). They were then stored in their original containers at the same temperature until further analysis and all analyses were done within two weeks of sample collection. Sodium hydroxide and ethanol were from R&M Chemicals, UK. All chemicals used in this study were of analytical grade.

Sample Pre-Treatment

Before analysis, all honey samples were allowed to stand at room temperature for 2 hours. The samples were then stirred gently with a spatula for 10 minutes to dissolve the sedimented sugar crystals and obtain a homogeneous liquid. Then, the samples were used for the determination of moisture content, pH, colour, electrical conductivity, and acidity.

Physicochemical Analysis

The moisture content of the samples was done based on AOAC method (AOAC 934.01) (AOAC, 2023). Briefly, 1 g of honey sample was placed in a pre-weighed crucible and then dried at 105°C for 5 hours in a drying oven (Memmert, Germany), after which the moisture content was calculated. pH was measured using a pH meter (Mettler Toledo, USA) by preparing 10% (w/v) honey solutions. Colour was analysed using a colorimeter (Hisuc, China) using CIE colour system. The electrical conductivity (EC) was measured using a conductivity probe (Mettler Toledo SevenExcellence, USA) by preparing 10% (w/v) honey solutions. The results are expressed in microSiemens per centimetre ($\mu S\ cm^{-1}$). The free acidity (expressed in meq acid/kg honey) was determined according to Pauliuc *et al.* (2021) [8]. About 10 g of honey sample was mixed with 75 mL of free CO_2 water and titrated with 0.1 M NaOH until reaching a pH of 8.30. The free acidity was calculated using the following equation:

$$Free\ acidity = volume\ (mL)\ of\ 0.1\ M\ NaOH \times 10 \quad \text{(Eqn. 1)}$$

Pollen Analysis

The pollen analysis was performed based on the method proposed by Pauliuc *et al*. (2021) [8], with minor modification. About 10 g of honey was diluted with 40 mL of MiliQ water in a 50 mL volumetric flask and then filled to the mark. Then, the mixture was homogenised using a magnetic stirrer. The diluted honey was transferred into a 50 mL falcon tube and centrifuged using a refrigerated centrifuge (Kubota, Japan) at 4°C, 4500 rpm for 15 min, and the supernatant was removed. The remaining residue was filled up again to 50 mL with MiliQ water, vortexed for 1 min and re-centrifuged at 4°C, 4500 rpm for 15 min and the supernatant removed. The precipitate was spread in a thin layer on a microscope slide, and the pollen

grains were observed at 40× magnification using a light microscope (Olympus CX23, China) with at least 800 pollen grains counted and the pollen percent was estimated.

FTIR-ATR Analysis

The FTIR spectra were obtained using a Nicolet iS5 FTIR spectrometer (Thermo Scientific, USA) with a 4 cm^{-1} resolution and 32 scans per spectrum (4000 – 650 cm^{-1}). An ATR cell analysed a drop of the honey samples with a blank diamond ATR scan for reference. The spectra were processed using OMNIC™ software (v8.2) with ATR correction, baseline correction, and smoothing before being saved as a CSV file.

Chemometric Analysis

Dataset Pre-Processing, KMO Test and Transformation

The FTIR-ATR spectral data were compiled in comma-separated values (CSV) files and imported into XLSTAT software (2019 version, France) [10]. The transmittance data were separated into seven wavenumber ranges: 4000 – 3501 cm^{-1}, 3500 – 3001 cm^{-1}, 3000 – 2501 cm^{-1}, 2500 – 2001 cm^{-1}, 2000 – 1501 cm^{-1}, 1500 – 1001 cm^{-1}, and 1000 – 650 cm^{-1}. The important wavenumbers were then identified and the PCA was performed to determine their contribution across all the honey samples. Then, before running the PCA, the dataset's suitability was confirmed using two tests with a significance level of $\alpha = 0.05$, respectively. Firstly, the Kaiser-Meyer-Olkin (KMO) test was used to check for data adequacy. A KMO score was ranked as KMO < 0.5 = inadequate, 0.5 < KMO < 0.7 = mediocre, 0.7 < KMO < 0.8 = good, 0.8 < KMO < 0.9 = very good, and KMO > 0.9 = excellent to indicate the dataset adequacy, and only KMO > 0.5 was considered acceptable for PCA [11]. Secondly, the Shapiro-Wilk test was performed to ensure the data followed a normal distribution. To achieve a normal distribution, the datasets were processed using several transformation methods of standardised (n-1), standardised (n), centre, standard deviation-1 (n-1), standard deviation (n), and Pareto. The normality of each transformed dataset was then re-assessed using the Shapiro-Wilk at $\alpha = 0.05$. The best-performing transformation method was selected based on these results.

Principal Component Analysis (PCA)

The full FTIR-ATR spectra were extracted based on their transmittance values to generate a comprehensive dataset for PCA. Wavenumber ranges of 1000 – 650 cm^{-1}, 1500 – 1001 cm^{-1}, 2000 – 1501 cm^{-1}, 2500 – 2001 cm^{-1}, 3000 – 2501 cm^{-1}, 3500 – 3001 cm^{-1}, and 4000 – 3501 cm^{-1} were selected as they represent both the functional group and fingerprint regions of the spectra. These spectral regions were chosen due to their high potential in capturing key molecular vibrations that can effectively distinguish between different sample groups. Prior to PCA, the dataset was pre-processed using the Pareto scaling technique, which centres the data and scans each variable by the square root of its standard deviation. This method enhances moderate variations while preserving the data structure, making it particularly suitable for spectroscopic analysis. As a result of Pareto scaling, the distribution of the variables approached normality through the Gaussian distribution of the transformed data. The number of principal components (PCs) was optimised to obtain optimum differentiation among samples. The differentiation result of the samples was observed using PCA score plot. Moreover, the PCA model was evaluated using its R^2 and Q^2 values to justify the good of fitness and predictivity of the PCA model, respectively [12].

Agglomerative Hierarchical Clustering (AHC) Analysis

Agglomerative hierarchical clustering (AHC) was applied on the datasets of the honey samples in order to classify them to obtain the classification based on their physicochemical analysis using Euclidian distance and Ward's method as aggregation criterion without data pre-treatment. Noteworthy, hierarchical clustering is an important and inherent process in unsupervised machine learning. It begins with each variable signifying an individual cluster. The clusters are then subsequently merged according to their similarity. Initially, the most similar clusters (usually those with the smallest distance between them) are merged to form a new cluster at the bottom of the hierarchy. In the next step, another pair of clusters is merged and linked to a higher level of the hierarchy, and so forth [13]. Graphical representation of results is a tree graph called a dendrogram.

Partial Least Squares-Discriminant Analysis (PLS-DA)

Partial least squares-discriminant analysis (PLS-DA) was used to develop a model for identifying different types of stingless bee honey using the FTIR-ATR spectra. A 'cluster' label was added to the datasets to classify samples as either pure or SBH (stingless bee honey). The PLS-DA was performed at a significance level of $\alpha = 0.05$, projecting the spectral data onto a latent variable space optimised to distinguish between classes. The model assigned each sample to a class based on its FTIR spectrum and calculated discriminant scores for classification [14]. The model's performance was evaluated using several metrics: R^2Y (goodness-of-fit to the response), R^2X (variance explained in the predictors), and Q^2 (predictive accuracy via cross-validation). Cluster differences were assessed using Fisher distance and corresponding p-values. The model's predictive power was further validated on cross-validation and test datasets by calculating the

percentage of correctly classified samples. Based on these statistical parameters and classification results, the most effective PLS-DA model was selected [15].

Statistical Analysis

The experiments were performed in triplicate, and the results are expressed as mean ± standard deviation. Data were analysed using ANOVA with Tukey's post-hoc test in SPSS v16.0, with significance set at $p<0.05$.

RESULTS AND DISCUSSION

Physicochemical Properties of SBH

The physicochemical properties of the SBH samples in this study showed considerable variability between stingless bee and bee honey species samples (**Table 1**). The moisture content of the honey samples ranged from 11.33% (SBHC1) to 37.33% (KP). There was an inverse relationship with total solids content as higher water content is associated with lower solids content. According to the Malaysian Standard MS 2683:2017, the maximum allowable moisture content for raw SBH is 35%, while processed SBH should not exceed 22% [16]. Most raw samples in this study met the requirements for raw honey, except KP (37.33%), which exceeded the standard. SBHC1, a commercial honey, had the lowest moisture content (11.33%), which is well below the maximum value for processed honey [16]. This low value likely reflects post-harvest processing, such as controlled dehydration, which is commonly applied to commercial stingless bee honey to reduce water content, improve storage stability and minimize fermentation risks. In contrast, raw honeys such as KP, PM, and IT retained a higher moisture content, which is related to their natural hygroscopicity, storage in honey pots, and open harvesting methods [17].

The pH of the honey samples ranged from 2.89 (SBHC2) to 4.66 (SBHC1), while the free acidity ranged from 10.69 meq/kg (TH) to 35.27 meq/kg (SBHC2). According to the Malaysian Standard, the acceptable pH range for raw and processed SBH is 2.5–3.8 (Malaysian Standard, 2017); all samples except SBHC1 (4.66) met this value [16]. Lower pH and higher acidity, as observed in SBHC2 (2.89; 35.27 meq/kg), indicate higher levels of organic acids, which may enhance antimicrobial activity but may increase the risk of fermentation in honey with high moisture content [17,18]. In contrast, SBHC1 exhibited a higher pH with lower acidity, likely due to processing reducing acidity, resulting in a milder flavour but possibly lower microbial stability. Overall, pH and free acidity complement each other in determining the acidity profile of honey and, together with EC, serve as important indicators of quality, stability, and authenticity.

The EC value of the honey samples ranged from 66.63 µS/cm (SBHC1) to 1080.73 µS/cm (SBHC2), reflecting the differences in mineral and ion content. Higher EC values, such as those in SBHC2 and BH, indicate higher concentrations of minerals and organic acids, while the low EC value in SBHC1 is consistent with commercial processing and lower mineral content. The EC value is considered a reliable indicator of the authenticity and botanical origin of honey, as it is less affected by storage than moisture or pH. The typically higher EC value of SBH compared to *Apis mellifera* honey is related to the diverse floral sources and mineral-rich nectar [19]. Meanwhile, the colour of the samples ranged from very light (PM, TH) to darker shades (BH, SBHC1), with SBHC3 from the mountains showing medium values. The colour of honey is determined by the floral source, phenolic content, and pigments, which are closely related to the pollen profile [5]. Darker honeys tend to contain higher phenolic content and higher antioxidant activity, while lighter samples generally have a milder flavour and lower levels of bioactive compounds [35]. Thus, colour serves not only as a sensory and qualitative attribute, but also as a complementary indicator of botanical origin.

Compared to *Apis mellifera* honey, SBH generally contains higher moisture levels, reaching up to 37.33% compared with 20.33% in BH, and lower total solids content due to its hygroscopic nature and harvesting methods. BH exhibited moderate acidity (pH 3.33; 32.39 meq/kg), which is within the range of SBH, but showed less variability. The EC of BH (653.23 µS/cm) was comparable to that of mineral-rich SBH samples such as SBHC2 (1080.73 µS/cm) and PM (799.60 µS/cm). In terms of colour, SBH samples showed a wide range of lightness values, with several samples (such as KP and PM) appearing significantly darker than BH, while others, such as TH, displayed similar colour characteristics. These variations reflect differences in floral sources, pollen composition, and bee species [20]. Overall, SBH has distinctly different properties than BH, which emphasizes the importance of considering its unique characteristics when evaluating quality. The observed variations in moisture, acidity, EC and colour illustrate the strong influence of bee species and flower origin on the quality of stingless bee honey. Since the pollen directly reflects the botanical origin, the pollen analysis provides further information to explain these physicochemical differences.

Table 1. Physicochemical properties of SBH based on some physicochemical analyses of moisture, total solid, pH, free acidity, electrical conductivity, and colour.

Sample	Parameters							
	Moisture (%)	Total solids (%)	pH	Free acidity (meq/kg)	Electrical conductivity (μS/cm)	Colour		
						L	a	b
BH	20.33±0.58^{b}	79.67±0.58^{c}	3.33±0.01^{cd}	32.39±0.25^{e}	653.23±4.36^{e}	64.39±0.46^{g}	0.28±0.02^{a}	55.47±0.45^{d}
SBHC1	11.33±2.89^{a}	88.67±2.89^{d}	4.66±0.01^{g}	11.57±0.25^{d}	66.63±12.97^{a}	47.74±0.24^{f}	30.83±0.08^{g}	76.55±0.56^{h}
SBHC2	29.33±2.31^{c}	70.67±2.31^{b}	2.89±0.07^{a}	35.27±0.88^{e}	1080.73±2.97^{h}	35.63±0.43^{c}	34.58±0.44^{h}	60.60±0.25^{f}
SBHC3	28.00±0.00^{c}	72.00±0.00^{b}	3.37±0.01^{d}	14.96±1.33^{ab}	579.63±1.53^{d}	37.94±0.12^{d}	21.82±0.26^{e}	60.37±0.44^{f}
KP	37.33±0.58^{d}	62.67±0.58^{a}	3.23±0.02^{f}	22.92±0.48^{cd}	681.13±4.50^{f}	0.17±0.02^{a}	14.56±0.52^{c}	8.43±0.42^{b}
PM	31.00±3.46^{cd}	69.00±3.46^{ab}	3.29±0.02^{g}	26.69±3.62^{d}	799.60±6.66^{g}	0.16±0.01^{a}	4.88±0.07^{b}	1.44±0.05^{a}
TH	25.67±2.31^{bc}	74.33±2.31^{bc}	3.54±0.01^{c}	10.69±2.13^{a}	518.43±3.96^{c}	63.42±0.45^{g}	4.50±0.02^{b}	57.00±0.19^{e}
TP	26.00±0.00^{bc}	74.00±0.00^{bc}	4.50±0.00^{f}	17.70±0.34^{bc}	398.37±1.83^{b}	45.56±0.45^{d}	18.23±0.02^{d}	65.88±0.16^{g}
IT	29.67±4.73^{c}	70.33±4.73^{b}	3.49±0.01^{e}	13.62±0.39^{ab}	585.30±9.82^{d}	26.07±0.72^{b}	22.69±0.24^{f}	43.65±0.32^{c}

Remark: The results shown are mean ± standard deviation ($n=3$). Values in the same column with different subscripts are significantly different at $p<0.05$. BH: *Apis mellifera*, SBHC1: Commercial SBH 1, SBHC2: Commercial SBH 2, SBHC3: Commercial SBH C, KP: SBH of *Tetrigona binghami*, PM: SBH of *Heterotrigona erythrogastra*, TH: SBH of *Geniotrigona thoracica*, TP: SBH of *Homotrigona aliceae*, and IT: SBH of *Heterotrigona itama*.

Pollen Analysis

The pollen analysis revealed clear differences in pollen density between the honey samples (**Figure 1**). The densest pollen concentrations were observed in SBHC3, KP, PM, SBHC3, and TH, while TP and IT showed moderate densities with more scattered pollen grains. In contrast, SBHC1 and SBHC2 contained very few pollen grains, probably due to processing and filtration, while BH had larger and more distinct pollen grains. As the hives where the SBH collected were located within palm oil plantations, the observed pollen morphology was consistent with oil palm (*Elaeis guineensis*), thus it is very likely that oil palm served as the dominant floral source. Nevertheless, the overall pollen profiles suggest that the SBH retained a polyfloral character, as stingless bees typically visit multiple nectar sources. Similar trends have been reported in previous studies where raw honeys had a richer pollen content, while processing significantly reduced palynological traits [21,22]. These results confirm that pollen analysis is useful for distinguishing raw from processed SBH and assessing floral contributions, although their irregular distribution suggests that complementary approaches such as FTIR-ATR chemometrics are required for reliable authentication.

Although the pollen data obtained in this study were qualitative and based on microscopic observation, meaningful relationships could still be inferred by comparing pollen density patterns with quantitative physicochemical parameters. Samples with the highest moisture content of KP, PM, and elevated EC values (681 – 799 μS/cm) exhibited the densest pollen presence, consistent with raw, unprocessed honey that retains more botanical residues. These darker samples (L < 1.0) also had higher free acidity, suggesting that organic acids and mineral-rich nectar collected from diverse floral sources contribute to both pollen load and chemical characteristics. Conversely, commercial SBH samples, particularly SBHC1 and SBHC2, showed very low pollen density alongside low moisture (11.33%) or processing-associated reduction in mineral content (EC = 66.63 μS/cm). The lighter colour and reduced acidity of these samples further support the influence of filtration and dehydration on pollen depletion. Overall, the physicochemical data corroborate the pollen findings by demonstrating that raw SBH samples with higher moisture, EC, acidity, and darker colour tend to contain more pollen, while processed honeys show minimal pollen due to post-harvest handling.

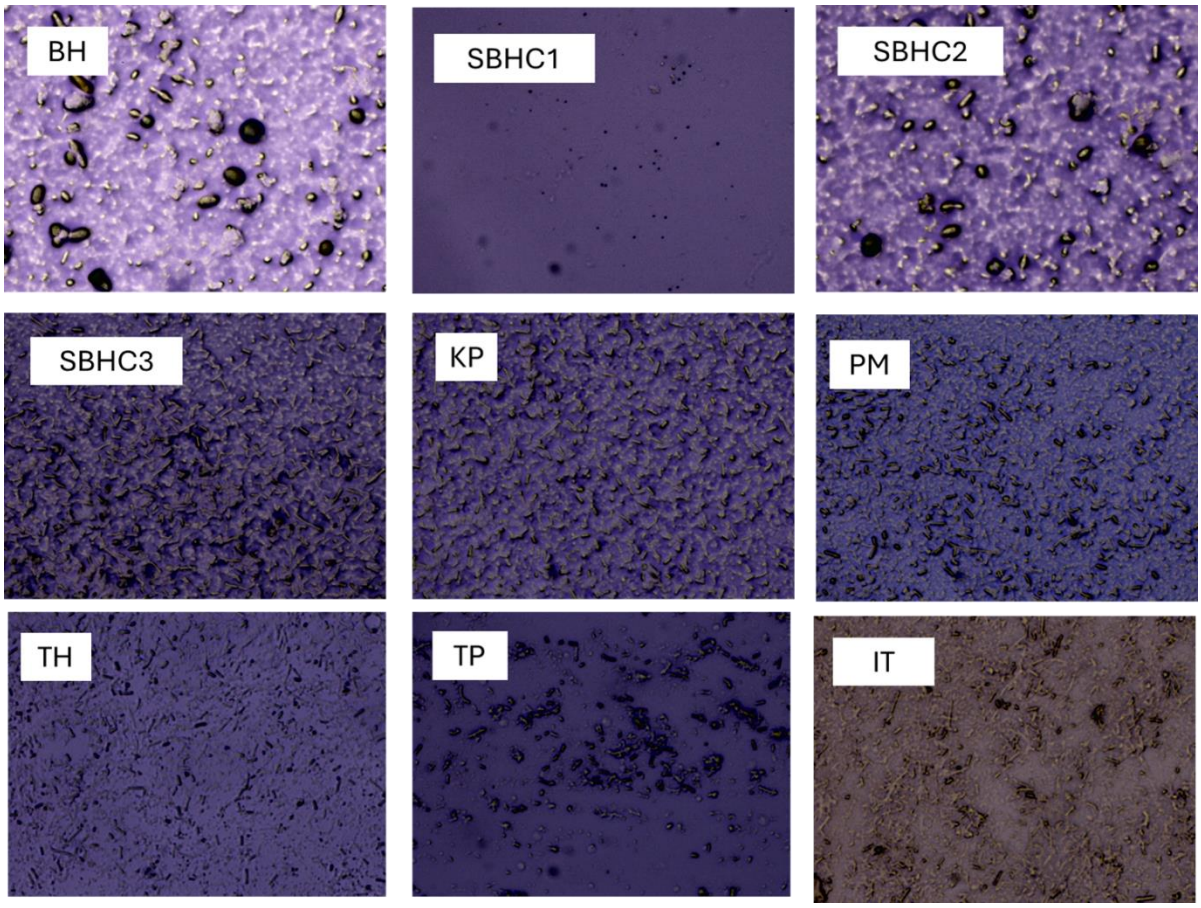

Figure 1. Morphology analysis of pollens in honey samples. BH: *Apis mellifera*, SBHC1: Commercial SBH 1, SBHC2: Commercial SBH 2, SBHC3: Commercial SBH C, KP: SBH of *Tetrigona binghami*, PM: SBH of *Heterotrigona erythrogastra*, TH: SBH of *Geniotrigona thoracica*, TP: SBH of *Homotrigona aliceae*, and IT: SBH of *Heterotrigona itama*.

FTIR-ATR Spectra of Honey

The medium IR region between 4000 – 650 cm^{-1} contains some information from molecular vibrations and is sensitive to the chemical and physical conditions of the samples, as this spectral region has aided to identify prominent samples with different composition and concentration. Moreover, this FTIR method is known to be a simple, fast, and non-destructive analysis. The FTIR spectra of the SBH samples analysed in this study are shown in **Figure 2**.

A strong and broad peak of the O–H group with a stretching vibration at ~3250 cm^{-1} was observed in the studied honey samples. A weak absorption band of the C–H stretching mainly from carboxylic acid was defined at ~2850 cm^{-1}. Moreover, a weak C–H stretching was observed at ~2300 cm^{-1} and at the lower energy region, the C=O stretching was observed at ~1600 cm^{-1}, and the characteristics of carbohydrates of C–O bond stretching of the C–O–C linkage of a strong and sharp peak at ~1100 cm^{-1}. The sweetness of SBH attributed to the chemical components of sucrose, glucose, and fructose can be observed in the spectral region of 1500 – 900 cm^{-1}, as reported by Hungerford *et al.* [23]. The FTIR spectra indicated the characteristic for saccharide configuration at 978 – 700 cm^{-1}.

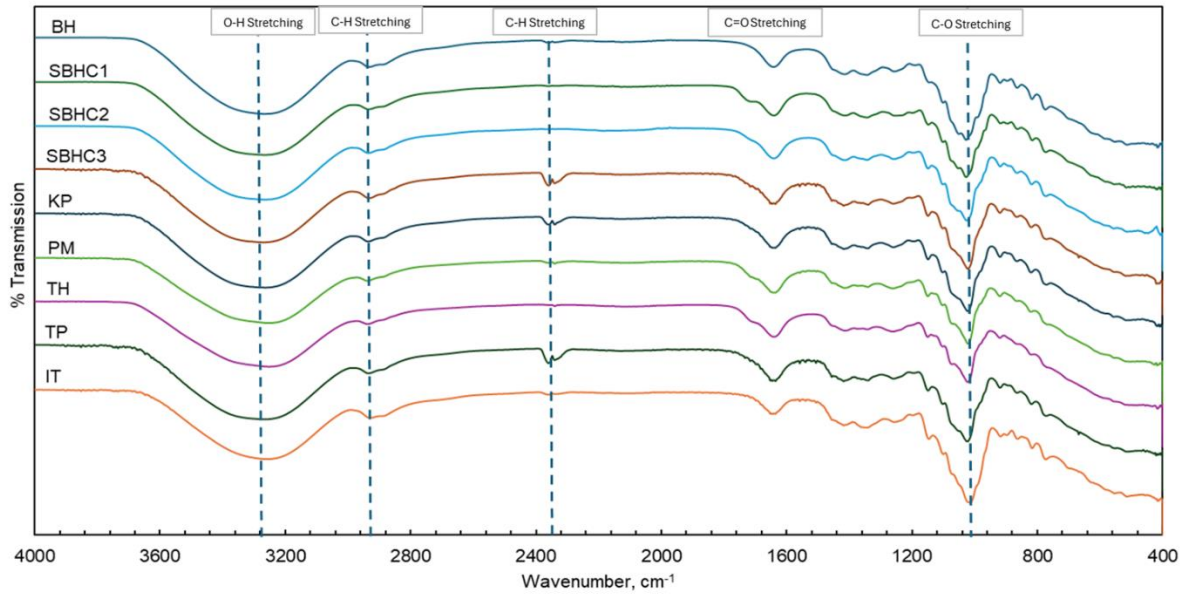

Figure 2. FTIR spectra of the honey samples. BH: *Apis mellifera*, SBHC1: Commercial SBH 1, SBHC2: Commercial SBH 2, SBHC3: Commercial SBH C, KP: SBH of *Tetrigona binghami*, PM: SBH of *Heterotrigona erythrogastra*, TH: SBH of *Geniotrigona thoracica*, TP: SBH of *Homotrigona aliceae*, and IT: SBH of *Heterotrigona itama*.

Chemometric Analysis

Principal Component Analysis (PCA)

Principal component analysis (PCA) was employed to evaluate the clustering trend and variance structure of the honey samples, providing an overview of the underlying similarities and differences. As illustrated in **Figures 3(a-g)**, the first two principal components (PC1 and PC2) accounted for a significant proportion of the total variance, ranging between 99.99% and 100% depending on the dataset with KMO test of the value 0.644 – 0.782, suggesting the data are within the acceptable range [24]. This indicates that the PCA was successful in capturing the major sources based on the compositional variation of FTIR-ATR spectral of the honey samples.

Distinct clustering was observed, where samples tended to group according to their floral, geographical origin or molecular composition of the honey [21]. For example, KP and PM formed a separate cluster along PC1 throughout the FTIR-ATR wavenumber region of 4000 – 650 cm^{-1}, reflecting differences in the chemical components, such as sugars (sucrose, glucose, and fructose), amino acids, phenolic compounds, and organic acids [25]. Other than that, SBHG and TP also formed a clear and separate cluster for the aforementioned wavenumber. The observed separation demonstrates that the PCA can discriminate honey samples based on their intrinsic chemical fingerprints without requiring prior class information. The tight grouping of each honey sample type further supports the reliability of the method, suggesting good homogeneity within groups and minimal experimental noise [26]. Conversely, the inter-group variation captured by the PCA highlights the contribution of specific compounds or functional groups that differ between honeys of different botanical or geographical sources. These findings are in line with previous reports where PCA successfully differentiated monofloral and multifloral of honey samples [27, 28].

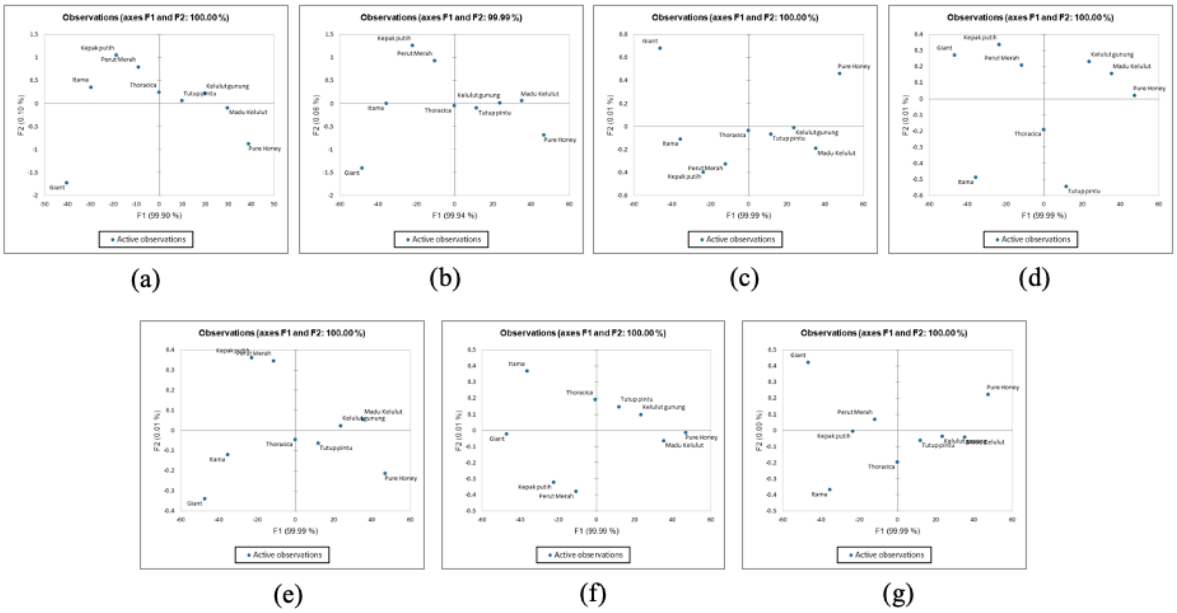

Figure 3. Principal component analysis (PCA) based on the FTIR-ATR spectra of the honey samples by different wavenumber regions: (a) 1000 – 650 cm^{-1}; (b) 1500 – 1001 cm^{-1}; (c) 2000 – 1501 cm^{-1}; (d) 2500 – 2001 cm^{-1}; (e) 3000 – 2501 cm^{-1}; (f) 3500 – 3001 cm^{-1}; and (g) 4000 – 3501 cm^{-1}.

Noteworthy, the PCA results confirm that the compositional variability of honey can be successfully captured and visualised through multivariate analysis. This highlights the utility of spectroscopic data coupled with PCA as a rapid, non-destructive, and cost-effective approach for honey authentication and quality assurance, which is increasingly important for consumer protection and regulatory compliance. Further studies should focus on identifying the chemical compounds responsible for differences in FTIR-ATR of all honey samples.

Agglomerative Hierarchical Clustering (AHC) Analysis

Agglomerative hierarchical clustering (AHC) analysis was used to visualise the classification of the honeys based on the physicochemical analysis of moisture, total solid, electroconductivity, pH, and free acidity. The AHC focused mainly on finding similarities between samples using different classification algorithms. The descriptive statistics of the honey dataset (**Table 2**) showed that all the honey samples exhibited relatively narrow ranges of moisture and total solids, but the electrical conductivity (EC), pH and free acidity showed wider variation. In particular, the EC ranged from 65.20 to 1093.30 µS/cm, reflecting strong differences in mineral content among all the honeys. Similarly, free acidity varied substantially (17.40 – 720.84 meq/kg), suggesting diverse botanical origins and fermentation states of the honeys, respectively.

Table 2. Summary of the collected dataset.

Variable	Minimum	Maximum	Mean	Std. deviation
Moisture	95.33	97.18	96.41	0.57
Total solid	2.82	4.67	3.589	0.57
EC	65.20	1093.30	595.89	265.93
pH	2.82	4.67	3.59	0.57
Free acidity	17.40	720.84	320.03	206.61

Table 3. Summary of the resultant AHC of collected dataset.

Class	1	2	3
Objects	3	15	9
Sum of weights	3	15	9
Within-class variance	28.799	18749.777	46532.973
Minimum distance to centroid	2.088	28.502	46.353
Average distance to centroid	4.065	111.311	188.357
Maximum distance to centroid	6.096	206.528	284.035

Table 4. Central objects with respect to the resultant classes.

Class	Moisture	Total solids	EC	pH	Free acidity
1 (SBHC_3)	95.340	4.660	66.000	4.660	25.124
2 (TH_2)	96.459	3.541	519.600	3.541	248.680
3 (PM_1)	96.727	3.273	810.900	3.273	568.520

Next, the agglomerative hierarchical clustering (AHC) of the dataset produced three distinct classes, as shown in **Table 3**. Class 1 consists of three samples characterised by low within-class variance of 28.799 and short distances to the centroid, indicating high internal homogeneity. Class 2 groups the largest number of samples (*n=15*) with moderate within-class variance of 18749.777, while Class 3 contains 9 samples with the highest within-class variance of 46532.973, providing more pronounced internal diversity. The distance measures also indicate that Class 1 samples were the most compact, whereas Classes 2 and 3 exhibited progressively larger dispersion.

In **Table 4**, representing each class based on the central objects provides insight into the compositional signatures based on the clusters. Class 1, which is represented by SBHC1_3, displays low EC (66 µS/cm) and low free acidity (25.12 meq/kg). Next, Class 2, represented by TH_2, exhibits intermediate characteristics with higher EC (519.6 µS/cm) and moderate free acidity (248.68 meq/kg). In contrast, Class 3, represented by PM_1, is characterised by the highest EC (810.9 µS/cm) and marked elevation of free acidity (568.2 meq/kg). These findings suggest that the AHC successfully differentiates all the honey samples according to their physicochemical profiles, with distinct clusters likely reflecting differences in floral sources, post-harvest handling, and storage [29].

Figure 4 shows the AHC of the honey samples showing three principal clusters at the selected dissimilarity threshold, reflecting clear compositional distinctions among all the honey samples. The first cluster encompasses of KP and PM, indicating a high degree of homogeneity within their chemical profiles. Moreover, the second cluster groups SBHC1 and BH, which exhibited moderately divergent characteristics relative to the first cluster and maintained internal consistency. The third cluster comprises of TP, SBHG, IT, and TH, which were markedly differentiated among all the samples. Notably, technical replicates for each sample type, for example SBHC2_1-3, KP_1-3, clustered tightly withing their respective groups, attesting to the reliability and reproducibility of the analytical approach. These clustering patterns are in agreement with previous studies that utilisied AHC, together with PCA or PLS-DA, to infer the botanical and geographical origins of honey based on physicochemical and FTIR-ATR spectral characteristics [28,30].

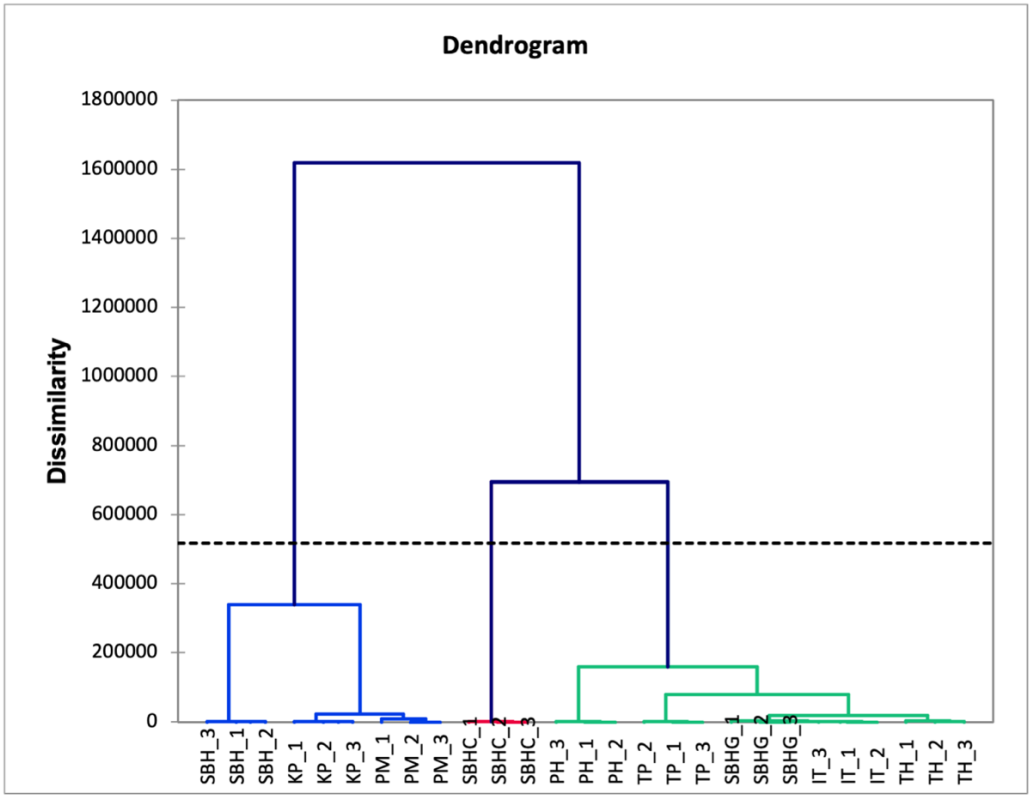

Figure 4. Agglomerative hierarchical clustering (AHC) of the honey samples (SBHC2, KP, PM, SBHC1, BH, TP, SBH3, IT, and TH) based on the honey physicochemical analysis.

Partial Least Square-Discriminant Analysis (PLS-DA)

The application of PLS-DA towards the FTIR-ATR spectral data revealed that specific spectral regions hold significant potential for discriminating between BH and SBH. The most effective classification was achieved at the 3500 – 3000 cm^{-1} spectral range, which corresponds to the O–H stretching vibrations from water, carbohydrates, and organic acids [31]. This model demonstrated strong predictive capability, as indicated by a high cumulative Q^2 value of 0.684 and R^2Y of 0.962 and it successfully classified 100% of the samples in the training set (**Table 5**). The model's predictive power was confirmed through cross-validation, which showed consistent classification trends and acceptable Q^2 values across multiple spectral ranges.

Although the present study relied on internal cross-validation rather than external validation, the agreement between the R^2Y and Q^2 parameters suggests that the model was not overfitted and maintained good internal predictive reliability. Previous studies have also demonstrated that cross-validation provides a reliable approximation of model generalizability when the available sample size is limited and external datasets are not feasible. According to Ciursă e*t al.* (2021) and Tarapoulouzi *et al.* (2024), internal cross-validation can effectively mimic the performance of independent testing by repeatedly partitioning the data into training and testing subsets, therefore reducing model bias and variance. Furthermore, this approach allows for an objective estimation of how the PLS-DA model would perform on unseen data while ensuring the optimal number of latent variables is selected to prevent overfitting, respectively. In chemometric applications involving FTIR data, particularly in food authentication studies with small sample sets, cross-validation has been shown to yield predictive accuracies comparable to those obtained from external validation, thus supporting its uses as a statistically valid alternative for evaluating model robustness.

Table 5. Classification matrix of training and validation datasets of PLS-DA of pure honey and stingless bee honey (SBH) samples.

Discriminating model wavenumber (cm^{-1})	Discriminating model quality R^2Y, R^2X and Q^2 cumulated indices and permutation test				Dataset	% correct classification
	R^2Y	R^2X	Q^2	p-value		
1000 – 650	0.298	0.999	0.143	$p<0.05$	**Training**	
					Pure	0
					SBH	100
					Validation	
					Pure	0
					SBH	100
1500 – 1001	0.295	0.999	0.191	$p<0.05$	**Training**	
					Pure	0
					SBH	100
					Validation	
					Pure	0
					SBH	100
2000 – 1501	0.592	1.000	0.039	$p<0.05$	**Training**	
					Pure	100
					SBH	100
					Validation	
					Pure	0
					SBH	100
2500 – 2001	0.307	1.000	0.140	$p<0.05$	**Training**	
					Pure	0
					SBH	100
					Validation	
					Pure	0
					SBH	100
3000 – 2501	0.303	1.000	0.141	$p<0.05$	**Training**	
					Pure	0
					SBH	100
					Validation	
					Pure	0
					SBH	100
3500 – 3001	0.962	1.000	0.684	$p<0.05$	**Training**	
					Pure	100
					SBH	100
					Validation	
					Pure	0
					SBH	0
4000 – 3501	0.294	1.000	0.137	$p<0.05$	**Training**	
					Pure	0
					SBH	100
					Validation	
					Pure	0
					SBH	100

In contrast, models developed using other spectral regions of 1000 – 650 cm^{-1}, 1500 – 1001 cm^{-1}, 2500 – 2001 cm^{-1}, 3000 – 2501 cm^{-1}, and 4000 – 3501 cm^{-1} provided poor predictive models with low Q^2 values (all below 0.2), indicating a lack of correlation between the spectral variations in these regions and the honey type. Furthermore, the model for the 2000 – 1501 cm^{-1} region, despite perfect classification in the training set was clearly overfitted evidenced by a negligible Q^2 value of 0.039 and a failure to classify pure honey in the validation set. Therefore, the significant difference observed in the 3500 – 3001 cm^{-1} region suggests that variations in the content of hydroxyl-containing compounds, primarily water and sugars, are the most reliable indicators for distinguishing between pure and SBHC2 using FTIR-ATR spectroscopy. This analysis demonstrates the critical importance of using Q^2 and an independent validation set to assess the true performance of a PLS-DA model, rather than relying solely on the training set results.

CONCLUSION

This study demonstrated that SBH exhibits substantial variation in moisture, pH, free acidity, electrical conductivity, and colour, with some raw samples exceeding the limits of MS 2683:2017, while processed SBH generally met the defined specifications. Pollen analysis qualitatively confirmed the presence of botanical residues in raw SBH, with higher pollen densities observed in samples with higher moisture, acidity, and EC, whereas commercial samples showed very low pollen due to filtration. FTIR-ATR spectra provided characteristic chemical fingerprints, and the integration of FTIR data with chemometric analyses (PCA, AHC and PLS-DA) successfully differentiated SBH from *Apis mellifera* honey and classified SBH according to compositional traits. Collectively, the combined physicochemical–spectral approach offers a rapid and robust framework for evaluating SBH quality and authenticity. The findings support the continued development of species-specific standards, such as MS 2683:2017, and highlight the need for expanded datasets with broader floral origins to further strengthen classification and authentication models.

ACKNOWLEDGEMENT

The research was financially supported by Skim Geran Pegawai Penyelidik MRUN, grant number: R.J130000.7809.5L003, and conducted in the facilities at Innovation Centre in Agritechnology for Advanced Bioprocessing (ICA), UTM Pagoh Campus, Johor, Malaysia and Pusat Inovasi Teknologi Manisan, Fakulti Sains Dan Teknologi, Universiti Kebangsaan Malaysia. The authors would like to thank Mrs. Zaheda Mohamad Azam and Mrs. Nooraina Atira Alaudin for their assistance in the project.

REFERENCES

1. Pote, C. L., Shirsat, D. V., Mahadule, P. A., Gade, K. A., Pandit, T. R., Soumia, P. S., Thangasamy, A., Kumar, S., Mahajan, V. & Karuppaiah, V. (2025) Biochemical, antioxidants, and mineral constituents of stingless bee honey. *Frontiers in Sustainable Food Systems*, **9**, 1546843.

2. Zaldivar-Ortega, A. K., Cenobio-Galindo, A. D. J., Morfin, N., Aguirre-Álvarez, G., Campos-Montiel, R. G., Esturau-Escofet, N., Garduño-García, A. & Angeles-Hernandez, J. C. (2024) The physicochemical parameters, phenolic content, and antioxidant activity of honey from stingless bees and Apis mellifera: A systematic review and meta-analysis. *Antioxidants*, **13(12)**, 1539.

3. Chen, Y. H., Chuah, W. C. & Chye, F. Y. (2021) Effect of drying on physicochemical and functional properties of stingless bee honey. *Journal of Food Processing and Preservation*, **45(4)**, 1–15.

4. Abdullah, N. A., Zullkiflee, N., Zaini, S. N. Z., Taha, H., Hashim, F. & Usman, A. (2020) Phytochemicals, mineral contents, antioxidants, and antimicrobial activities of propolis produced by Brunei stingless bees *Geniotrigona thoracica*, *Heterotrigona itama* and *Tetrigona binghami*. *Saudi Journal of Biological Sciences*, **27(11)**, 2902–2911.

5. Pauliuc, D., Dranca, F. & Oroian, M. (2020) Antioxidant activity, total phenolic content, individual phenolics and physicochemical parameters suitability for Romanian honey authentication. *Foods*, **9(3)**, 306.

6. Zhang, G. Z., Tian, J., Zhang, Y. Z., Li, S. S., Zheng, H. Q. & Hu, F. L. (2021) Investigation of the maturity evaluation indicator of honey in natural ripening process: The case of rape honey. *Foods*, **10(11)**, 2882.

7. Saidan, N. H., Roslan, N., Baharuddin, M. R. N., Ridzuan Hamil, S. M. & Krishnan, K. T. (2020) Compliance of selected stingless bee honey in Kelantan according to Malaysian Standard (MS) 2683:2017. *Malaysia Applied Biology*, **49(4)**, 187–192.

8. Pauliuc, D., Ciursă, P., Ropciuc, S., Dranca, F. & Oroian, M. (2021) Physicochemical parameters prediction and authentication of different monofloral honeys based on FTIR spectra. *Journal of Food Composition and Analysis*, **102**, 104021.

9. Shamsudin, S., Selamat, J., Sanny, M., Abd. Razak, S. B., Jambari, N. N., Mian, Z. & Khatib, A. (2019) Influence of origins and bee species on physicochemical, antioxidant properties and botanical discrimination of stingless bee honey. *International Journal of Food Properties*, **22(1)**, 239–264.

10. Addinsoft (2019) XLSTAT Statistical and Data Analysis Solution. Long Island, NY, USA. https://www.xlstat.com.

11. Rojas-Valverde, D., Pino-Ortega, J., Gómez-Carmona, C. D. & Rico-González, M. (2020) A systematic review of methods and criteria standard proposal for the use of principal component analysis in team's sports science. *International Journal of Environmental Research and Public Health*, **17(23)**, 8712.

12. Kang, M. J., Kim, K. R., Kim, K., Morrill, A. G., Jung, C., Sun, S., Lee, D. H., Suh, J. H. & Sung, J. (2023) Metabolomic analysis reveals linkage between chemical composition and sensory quality of different floral honey samples. *Food Research International*, **173**, 113454.

13. Craparo, V., Viola, E., Vella, A., Prestianni, R., Pirrone, A., Naselli, V., Amato, F., Oliva, D., Notarbartolo, G., Guzzon, R., Settanni, L., Moschetti, G., Francesca, N. & Alfonzo, A. (2024) Oenological capabilities of yeasts isolated from high-sugar matrices (manna and honey) as potential starters and co-starters for winemaking. *Beverages*, **10(3)**, 48.

14. Nazri, M. Z., Latiff, N. A., Rashid, S. N. A. A., Malik, S. A., Rahmani, A. S., Alaudin, N. A., Musa, N. F., Azam, Z. M., Embi, K. & Basar, N. (2025) Enzymatic and microbial-assisted treatments for the extraction of agarwood essential oils and authentication using FTIR-ATR spectroscopy and multivariate data analysis. *Malaysian Journal of Chemistry*, **27(2)**, 197-216.

15. de Jesus Inacio, L., Lanza, I., Merlanti, R., Contiero, B., Lucatello, L., Serva, L., Bisutti, V., Mirisola, M., Tenti, S., Segato, S. & Capolongo, F. (2020) Discriminant analysis of pyrrolizidine alkaloid contamination in bee pollen based on near-infrared data from lab-stationary and portable spectrometers. *European Food Research and Technology*, **246(12)**, 2471–2483.

16. Department of Standards Malaysia (2017) MS 2683:2017 – Kelulut (Stingless bee) honey – Specification. *Department of Standards Malaysia*.

17. Julika, W. N., Ajit, A., Sulaiman, A. Z. & Naila, A. (2019) Physicochemical and microbiological analysis of stingless bees honey collected from local market in Malaysia. *Indonesian Journal of Chemistry*, **19(2)**, 522–530.

18. Pinandita, E. P., Minarti, S., Junus, M., Eka Radiati, L., Iffany Faradilla Besari, R., Meiske Soen, J. & Ivana Br Simanjuntak, A. (2022) Potential trigona beekeeping (*Heterotrigona itama* and *Geniotrigona thoracica*) at south Labuhanbatu City, North Sumatera province, Indonesia region. *TERNAK TROPIKA Journal of Tropical Animal Production*, **23(2)**, 110–119.

19. Prica, N., Živkov-Baloš, M., Jakšić, S., Mihaljev, Ž., Kartalović, B., Babić, J. & Savić, S. (2014) Moisture and acidity as indicators of the quality of honey originating from Vojvodina region. *Arhiv Veterinarske Medicine*, **7(22)**, 99–109.

20. Rosidi Sujanto, I. S., Ramly, N. S., Abd Ghani, A., Huat, J. T. Y., Alias, N. & Ngah, N. (2021) The composition and functional properties of stingless bee honey: A review. *Malaysian Journal of Applied Sciences*, **6(1)**, 111–127.

21. Bodó, A., Radványi, L., Kőszegi, T., Csepregi, R., Nagy, D. U., Farkas, Á. & Kocsis, M. (2021) Quality evaluation of light-and dark-colored Hungarian honeys, focusing on botanical origin, antioxidant capacity, and mineral content. *Molecules*, **26(9)**, 2825.

22. Bong, Z. R., Shah, R. M., Chee, X. W., Hwang, S. S. & Ginjom, I. R. H. (2024) Simple and rapid characterization of Sarawak stingless bee honey using melissopalynological and ATR-FTIR analysis. *Food Analytical Methods*, **17(5)**, 773–786.

23. da Silva, R. N. A., Magalhães-Guedes, K. T., de Souza, C. O., de Oliveira Alves, R. M. & Umsza-Guez, M. A. (2024) Microbiological and physical-chemical characteristics of pollen and honey from stingless bees: A review. *Food Production, Processing and Nutrition*, **6(1)**, 95.

24. Hungerford, N. L., Zhang, J., Smith, T. J., Yates, H. S., Chowdhury, S. A., Carter, J. F., de Jesus, M. C. & Fletcher, M. T. (2021) Feeding sugars to stingless bees: Identifying the origin of trehalulose-rich honey composition. *Journal of Agricultural and Food Chemistry*, **69(35)**, 10292–10300.

25. Maina, M. W., Mburu, J., Maina, F. W. & Chimoita, E. L. (2024) Determinants of consumers' perceptions of honey quality attributes which

influence purchasing decisions in Nyandarua County, Kenya. *Bee World*, **101(1)**, 15–20.

26. Biluca, F. C., Braghini, F., Gonzaga, L. V., Costa, A. C. O. & Fett, R. (2016) Physicochemical profiles, minerals and bioactive compounds of stingless bee honey (*Meliponinae*). *Journal of Food Composition and Analysis*, **50**, 61–69.

27. Kek, S. P., Chin, N. L., Tan, S. W., Yusof, Y. A. & Chua, L. S. (2017) Classification of honey from its bee origin via chemical profiles and mineral content. *Food Analytical Methods*, **10(1)**, 19-30.

28. Ávila, S., Hornung, P. S., Teixeira, G. L., Malunga, L. N., Apea-Bah, F. B., Beux, M. R., Beta, T. & Ribani, R. H. (2019) Bioactive compounds and biological properties of Brazilian stingless bee honey have a strong relationship with the pollen floral origin. *Food Research International*, **123**, 1–10.

29. Shamsudin, S., Selamat, J., Sanny, M., AR, S. B., Jambari, N. N. & Khatib, A. (2019) A comparative characterization of physicochemical and anti-oxidants properties of processed *Heterotrigona itama* honey from different origins and classification by chemometrics analysis. *Molecules*, **24(21)**, 3898.

30. Manoko, M. L. K. & Mduda, C. A. (2024) Harvesting location influences the physicochemical properties, microbiological quality and sensory attributes of honey produced by an Afrotropical stingless bee, *Axestotrigona ferruginea*. *Food and Humanity*, **3**, 100433.

31. Tarapoulouzi, M., Mironescu, M., Drouza, C., Mironescu, I. D. & Agriopoulou, S. (2023) Insight into the recent application of chemometrics in quality analysis and characterization of bee honey during processing and storage. *Foods*, **12(3)**, 473.

32. Kozłowicz, K., Różyło, R., Gładyszewska, B., Matwijczuk, A., Gładyszewski, G., Chocyk, D., Samborska, K., Piekut, J. & Smolewska, M. (2020) Identification of sugars and phenolic compounds in honey powders with the use of GC-MS, FTIR spectroscopy, and X-ray diffraction. *Scientific Reports*, **10(1)**, 16269.

33. Ciursă, P., Pauliuc, D., Dranca, F., Ropciuc, S. & Oroian, M. (2021) Detection of honey adulterated with agave, corn, inverted sugar, maple and rice syrups using FTIR analysis. *Food Control*, **130**, 108266.

34. Tarapoulouzi, M., Pashalidis, I. & Theocharis, C. R. (2024) Discrimination of Cheese Products Regarding Milk Species' Origin Using FTIR, ^{1}H-NMR, and Chemometrics. *Applied Sciences*, **14(6)**, 2584.