Malaysian Journal of Chemistry, 2021, Vol. 23(2), 183-198
(Special Issue: *IKM Pahang Branch Online Symposium 2021 – Chemistry for Sustainable World*)

Combining Chemometrics, Sensory Analysis and Chromatographic Fingerprint of Volatile, and Phenolic Compositions for Systematic Classification of Pineapple (*Ananas comosus* L.)

Syaidatul Faraha Zainuddin[1], Rozita Osman[1]*, Hafizan Juahir[2] and Siti Raihan Zakaria[3]
[1]Faculty of Applied Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2]East Coast Environmental Research Institute, Universiti Sultan Zainal Abidin,
21300 Kampus Gong Badak, Kuala Terengganu, Malaysia
[3]Faculty of Applied Sciences, Universiti Teknologi MARA Pahang, 26400 Jengka, Pahang, Malaysia
*Corresponding author (e-mail: rozit471@uitm.edu.my)

Classification and quality control of fruits in Malaysia is based on the morphological traits manual carried out by the agricultural officer. As this approach is based on human perception and judgment, it may be biased and inconsistent. The aroma of pineapple is made up of a wide range of volatile and non-volatile compounds depending on the varieties and maturity stages of the fruits and thus, can be valuable factors in classifying the fruits. The chromatographic fingerprints of volatile and phenolic compounds from pineapple pulp were obtained by using Gas Chromatography-Mass Spectrometry Detector (GC-MSD) and 2-Dimensional-Liquid Chromatography – Diode Array Detector (2D-LC-DAD), respectively. Sensory profiles, conducted using quantitative descriptive analysis (QDA), showed that the fruity aroma of pineapple were not able to differentiate between the pineapple varieties with scales of 4.13±2.07, 5.33±2.58, 3.87±2.07, and 3.00±0.00 for the Morris, Josephine, MD2, and Sarawak varieties, respectively. Thus, sensory analysis alone could be biased and unreliable in discriminating pineapple varieties. Chemometric techniques based on unsupervised (principal component analysis (PCA) and hierarchical cluster analysis (HCA)), and supervised (discriminant analysis (DA) and partial least squares discriminant analysis (PLS-DA) using 13 sensory attributes, 10 selected phenolic compounds, and 35 volatile compounds allowed the discrimination of these four pineapple varieties.

Key words: Chemometrics; volatile compounds; phenolic compounds; chromatographic fingerprint; pineapple

Received: November 2020; Accepted: January 2021

Pineapple (*Ananas comosus* L.) is one of the commercially important fruit crops in Malaysia and is listed as one of the special projects in the National Key Economic Agenda (NKEA) under the agriculture sector. Pineapple draws attention not only because of its nutritional value as it is rich in vitamin C, minerals, and fiber but also due to its exotic aroma and taste. Despite its profitable prospects, there is no reliable, standardized method for the quality control of this fruit. Classification and quality control of fruits in Malaysia are based on morphological traits using manual inspection carried out by the agricultural officer [1]. This approach is not systematic as human perception and judgment are often biased and varied over time. The aroma of pineapple is made up of a wide range of chemical compounds including volatile and non-volatile compounds and varied qualitatively and quantitatively based on the maturity stages of the fruits. Therefore, the maturity index of the fruits can be an informative and valuable tool for the quality assessment and control of the fruit. Besides, bioactive compounds such as phenolic compounds are widely used as markers for authentication of fruit-based products because their presence in the fruit can differentiate it from other kind of fruit. Pineapples in Malaysia are of different varieties, namely Morris, N36, Sarawak, Morris Gajah, Gandul, Yankee, Josephine, Masapine, and most recently MD2 [2]. Among these varieties, Sarawak, Morris, MD2, and Josephine are the typical pineapple varieties available in the local fresh fruit market.

The evaluation of internal and external qualities of fruits in the past has been difficult and time-consuming due to the subjective form of assessment. Most aspects of the evaluation have always been based on the visual appearance, morphological traits, and defects of the fruit which could lead to inaccuracies in the results [3]. Further, assessment by manual sorting depends very much on human labor and time management that is liable to subjectivity [4]. Thus, since the demand for good-

†Paper presented at the IKM Pahang Branch Online Symposium 2021 – Chemistry for Sustainable World

quality fruit with accurate information on cultivar, origin, and maturity index is increasing, an automated system is desirable. Fruit quality is a consequence of many biochemical processes that result in changes of its intrinsic properties such as color, texture, flavor, and aroma, together with the exterior appearance (size, color and shape) and nutritional value [5]. Sensory attributes greatly depend on the composition of non-volatile and volatile compounds. The content of non-volatile components, such as phenolic compounds stimulate the taste receptors perceived as bitterness, astringency, and pungency while the volatile compounds stimulate the olfactory receptors responsible for the aroma [6]. Quantitative descriptive analysis (QDA) permits the determination of the most significant sensory attributes, which is important to the overall acceptance together with the use of reference standards [7].

The volatile components of pineapples have been studied extensively. More than 280 compounds are known to be involved in generating the characteristic pineapple flavor [8]. Pineapple varieties [9], areas where the pineapple crop is grown, ripening stages, and storage conditions are several factors known to affect the volatiles and phenolic profiles of pineapples [3]. Among these factors, the degree of ripeness and the nature of cultivar is reported to have a pronounced influence on the volatiles and phenolic compositions. Although much work has been done, to the best of our knowledge, there is still no study on a comprehensive system that utilizes volatiles profile, phenolic compounds, sensory analysis and the contribution of each of these constituents to the fresh, sweet aroma of pineapples grown in Malaysia. Thus, a comprehensive study on our own pineapple varieties is a must to make sure our national agenda to export this special fruit is successful.

Dealing with the chemical compounds for the quality control of fruits often involves a large set of data that would require an effective statistical tool like chemometric analysis. Chemometric methods, such as principal component analyses (PCA), neural networks, discriminant analyses (DA), hierarchical cluster analysis (HCA), and partial least square regression (PLSR), combined with flavor fingerprint data would enable the visualization of the fingerprints information [10; 11; 12] and can be considered as a convenient visual aid in creating a reasonable differtiation of the minor differences among the similar chromatograms [13]. PCA and PLS-DA are common examples of chemometric methods that have been applied in the identification of marker compounds in herbal plants and fruit varieties for quality assessments [14; 12]. Oliveri and Simonetti [15] suggested that chemometrics be classified into supervised and unsupervised methods. The most common supervised methods are soft independent modeling of class analogy (SIMCA) and linear or quadratic discriminant analysis (LDA or QDA), k-nearest neighbors (k-NN), partial least squares discriminant analysis (PLS-DA), and fuzzy rule building expert system (FuRES). Unsupervised methods are applied to study the data structure, identify similarities between data sets, and also to (determine?) access the outliers that may exist in the data set [16]. The most common unsupervised methods are PCA and HCA, and their applications are increasing in the food and chemistry fields for providing both sub-classes visualization and agglomerative algorithms, respectively [17; 18].

EXPERIMENTAL

1. Materials and Apparatus

Gallic acid, epicatechin, caffeic acid, ferulic acid, quercetin, and 2,2-diphenyl-1-picrylhydrazyl (DPPH) were purchased from Sigma Aldrich. HPLC–grade methanol, acetonitrile, sodium chloride were purchased from Merck (Darmstadt, Germany). Headspace solid-phase microextraction (HS-SPME) was conducted using a manual solid-phase microextraction holder. SPME fiber assembly was equipped with crosslinked phase, 65 µm polydimethylsiloxane–divinylbenzene (DVB/PDMS) supplied by (Supelco, Bellefonte, USA). The fiber was thermally conditioned as recommended by the manufacturer. Extraction vials (15 mL) with silicon septa were purchased from Supelco (Bellefonte, USA)

2. Sample Collection

Four varieties of Malaysian pineapple including Josephine (*Ananas comosus* L. var. comosus cv. Josephine), Morris (*Ananas comosus* L. var. comosus cv. Morris), Sarawak (*Ananas comosus* L. var. comosus cv. Sarawak), and MD2 (*Ananas comosus* L. var. comosus cv. MD2) were purchased from local orchards in Selangor, Malaysia. For each variety, 20 samples were analyzed to provide enough input data for chemometric analysis.

3. Sample Analysis

3.1. Sensory Analysis by Quantitative Descriptive Analysis (QDA)

The sensory analysis involved descriptive sensory analysis by trained panelists. Identification of descriptors and definition of the reference materials were performed according to DIN 10967-1 [19]. The identified attributes and corresponding aqueous reference solutions are as shown in Table 1.

Table 1. Sensory attributes and reference standards used by the descriptive analysis panel to evaluate pineapple varieties.

Attributes	Reference Standard	Evaluation Procedure
1. Aroma		
a. Sweaty/rancid	96.25 mg/L butanoic acid	Cover the sample container about three quarters, hold it close to the nose and inhale. Evaluate all perceived aroma.
b. Fruity	ethyl butanoate; 1.52 mg/L in 0.02% ethanol	
c. Floral	linalool; 17 mg/L in 0.01% ethanol	
2. Texture		
a. Firmness	1 cm unripe melon	Place the pineapple titbits between incisors, bite evenly through it
b. Juiciness	1 cm ripe melon	Place the pineapple titbits between molars. Assess after repeated chewing to evaluate the juiciness
c. Fibrousness	0.5 cm of raw celery slice	Chew the pineapple tidbits and evaluate the number of stringy fibers present
d. Chewiness	Gummy bear	Chewing with back teeth
e. Crunchiness	1.5 cm of Kitkat bar	Crunchiness assess after the first bite
3. Gustatory/Flavor		
a. Sweet	1 mL pineapple nectar	Allow the pineapple titbits to reach across all tasting zone of your tongue. Locate the flavor and the way pineapple titbits feel in your mouth.
b. Sour /Acidic	0.1% citric acid (food grade) solution	
c. Astringency	0.1% tannic acid water solution	
4. Aftertaste	-	Note the weight of taste remaining on the palate after the removal of the sample
5. Overall Acceptability	-	Chew until the mouthful is ready to swallow, evaluate the overall stimuli perceived retro-nasally

3.2. Analysis of Phenolic Compositions by Pressurised Liquid Extraction Two Dimensional Liquid Chromatography (PLE-2D-LC)

The pulp of pineapples was oven-dried overnight at 60 °C. The dried sample was mixed with diatomaceous earth and packed into the PLE extraction cell. The optimum PLE conditions (pressure of 1500 psi, the temperature of 105 ºC for 20 minutes) [20] were used for the extraction of phenolic compositions. Phenolic compositions were determined using selected standards (gallic acid, catechin, epicatechin, caffeic acid, rutin, *p*-coumaric acid, ferulic acid, quercetin, naringenin, and bromelain) and analyzed by 2D-LC-DAD.

3.3. Analysis of VOCs by Headspace Solid Phase Microextraction (HS-SPME) and Gas Chromatography-Mass Spectrometry (GC-MS)

Fresh pineapple was sliced and homogenized using a wet blender. The optimum HS-SPME conditions (temperature of 30 °C, time of 29 min, and salt addition of 1 g [21] were employed for the extraction of VOCs from pineapple. SPME fiber (65 µm PDMS/DVB) was inserted into the vial containing the sample for the extraction of VOCs. To diminish any carryover between each sample, SPME fiber was pre-conditioned at 250 °C for 7 min before the next analysis as suggested by Chmiel *et al.*[22].

GC-MS conditions applied in this study were similar to the study conducted by Zakaria *et al*. [23]. The SPME fiber containing the volatile compounds was thermally desorbed for 5 minutes at 280 ºC in split mode (1:10). The composition of VOCs from the pineapple pulp was analyzed on an Agilent 6890 gas chromatograph coupled to an Agilent 5975N mass spectrometer (Agilent, Santa Clara, CA, USA). The carrier gas, helium was applied at a constant flow of 1.2 mL/min and the separation of volatile compounds was carried out using a HP-5MS (Agilent) capillary column (30 m x 0.25 mm I.D x 0.25 µm film thickness).

3.4. Application of Chemometrics

The volatile and phenolic compositions were subjected to a chemometric method for extracting information and classification for the identification of marker compounds. XLSTAT software (XLSTAT v. 2016 (Addinsoft, Newyork, NY, USA) was used to compute the chemometric analyses (HCA, PCA, DA, and PLS-DA).

RESULTS AND DISCUSSION

The characteristic and sensory attributes (aroma, texture, flavor, aftertaste, and overall acceptability) were used by panelists to evaluate the acceptability and preferences of the 4 pineapple varieties (Josephine, MD2, Morris, and Sarawak). A formal sensory evaluation was conducted with samples in a random and coded manner after a consensus set of sensory descriptors was established as suggested by Bell *et al*. [24]. Sensory panelists were trained to evaluate 13 sensory properties of the four different pineapple varieties. Only sweaty aroma, firmness (texture), and sour (flavor) out of 13 sensory properties showed a significant difference between the four different pineapple varieties while the rest of the sensory properties were recorded as insignificant. Different alphabets (Figure 1) were used to represent the significant differences between the samples. The floral aroma was rated at low intensity (< 2 hedonic scores), whereas fruity aroma was noted at a level of 3 – 5 score for all varieties. The sweaty aroma showed significant differences between all pineapple varieties with a score of 4.40±1.55, 3.60±1.55, 8.80±0.00, and 1.00±0.00 for Morris, Josephine, MD2, and Sarawak varieties, respectively. From QDA analysis, the fruity aroma of pineapple was found to be insignificant for differentiating between the pineapple varieties with a scale of 4.13±2.07, 5.33±2.58, 3.87±2.07, and 3.00±0.00 for Morris, Josephine, MD2, and Sarawak varieties, respectively.

Sensory analysis has proved that human perception and judgment may be biased and inconsistent. As reported by Sinelli *et al*. [25], this approach has several drawbacks such as the lack of reference standards, experts(?), subjectivity, and variability of responses over time. Therefore, it is crucial to have an accurate, comprehensive, and systematic classification that can give more reliable information about this fruit. Further classification of pineapple varieties is needed by considering the aroma profile of the pineapple varieties. Several approaches have been introduced as alternatives to human perception and judgment using instrumental techniques that correlates sensory attributes to specific responses of chemical compositions such as phenolic and volatile compounds [6].

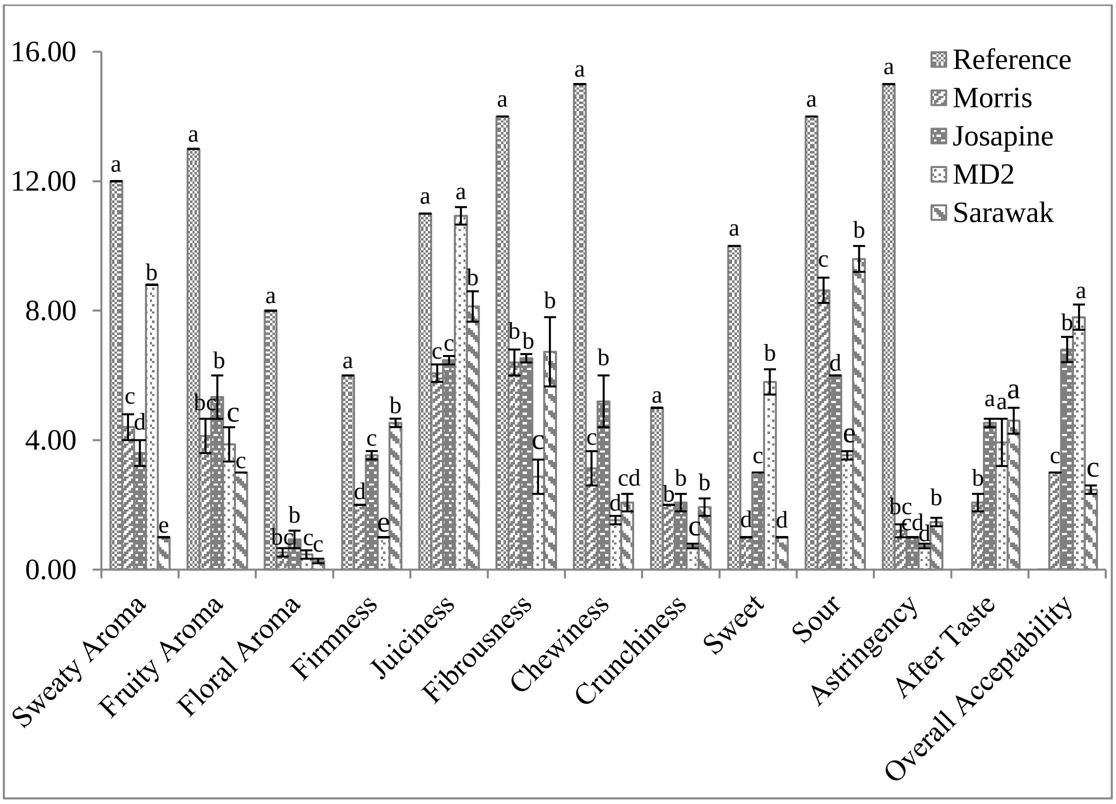

Figure 1. Bar Chart Diagram for the 13 Different Attributes from Different Pineapple Varieties, Evaluated Using a 15-Point Hedonic Sensory Evaluation.

(Note: Values are expressed as means ± standard deviation. Different letters (a-e) indicate significant difference at $p < 0.05$.)

Figure 2 shows the chromatographic fingerprints of pineapple varieties, showing the response intensity of selected phenolic compounds found in different pineapple varieties. Previous study conducted by Khan *et al*., [20] reported seven selected bioactive compounds namely catechin, epicatechin, chlorogenic acid, ferulic acid, myricetin, quercetin, and bromelain in MD2, Morris and Josephine pineapples. In their study, ferulic acid, quercetin, and myricetin were found to be absent in Morris pineapple. However, in this study, ferulic acid was detected in Morris pineapple albeit at only a 1% composition level. Similar to the study by Khan *et al*., [20] catechin, epicatechin, ferulic acid, quercetin, and bromelain were also the bioactive compounds found in MD2 and Josephine pineapples in this study. Ten phenolic compounds (gallic acid, catechin, epicatechin, caffeic acid, rutin, p-coumaric acid, ferulic acid, quercetin, naringenin, and bromelain) were selected for use in this study. The order for elution of compounds were the same for all fruit samples, except for differences in peak areas and peak heights.

For volatile compositions, the identification of VOCs of pineapples was conducted using GC-MS. The chromatograms of four pineapples varieties are shown in Figure 3. All chromatograms showeded slightly similar patterns with a few common intense compounds observed in all the samples. Thirty-five VOCs were selected for further study due to their high frequency of occurrences and consistent presence (repeatability) in all samples. Table 2 tabulates all the VOCs extracted using the optimized SPME conditions including 25 esters, four hydrocarbons, two monoterpenes, two aldehydes, one carboxylic acid, and one imine. MD2 pineapple contained the highest number of VOCs with a total of 26 compounds, followed by the Sarawak variety with 25 compounds. Josephine and Morris varieties had the least VOCs identified with 24 compounds. A study done by Lasekan and Hussein [26] reported a total of 59 volatile compounds in six different pineapple varieties, with eight compounds, namely methyl-2-methyl butanoate, methyl hexanoate, methyl-3-(methylthiol)-propanoate, methyl octanoate, 2,5-dimethyl-4-methoxy-3(2H)-furanone, δ-octalactone, 2-methoxy-4-vinyl phenol, and δ-undecalactone, being greatly responsible for aroma quality. In the present study, some of these same compounds such as methyl-2-methylbutanoate and methyl hexanoate were found in all the four pineapple varieties.

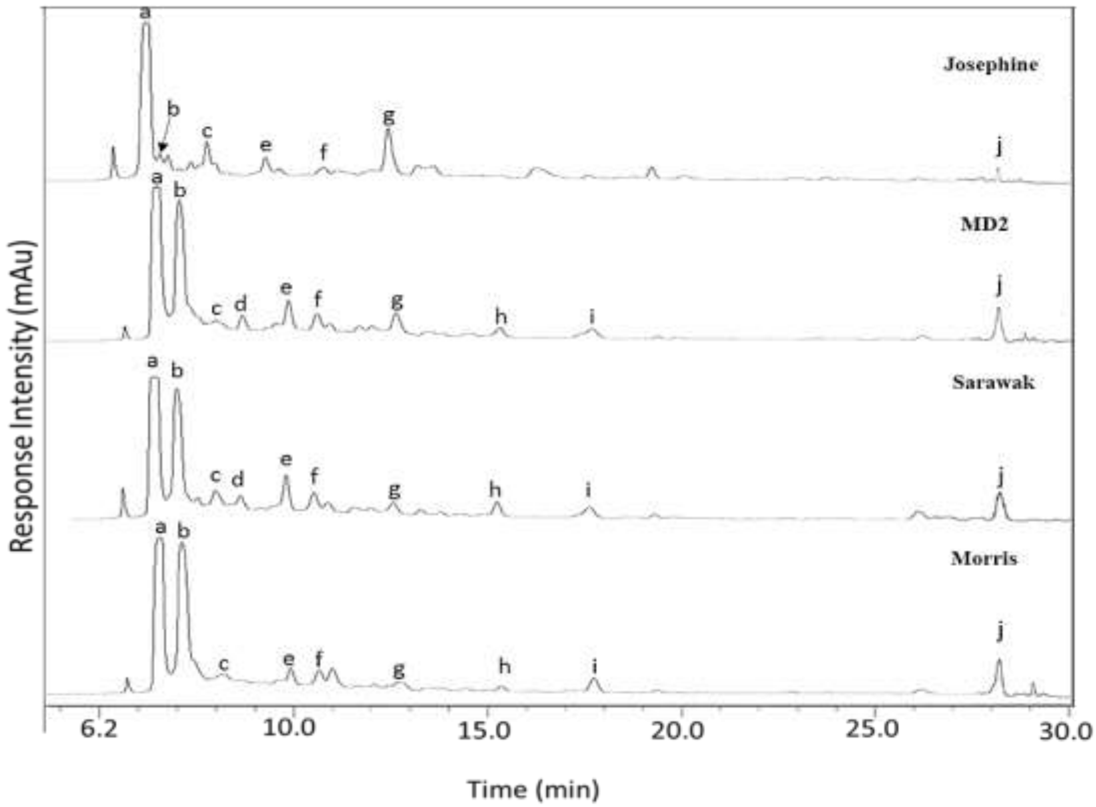

Figure 2. A representative chromatographic fingerprint of phenolic compounds; (a) gallic acid, (b) catechin (c) epicatechin (d) caffeic acid (e) rutin (f) p-coumaric acid (g) ferulic acid (h) quercetin (i) naringenin and (j) bromelain of four pineapples obtained at 280 nm

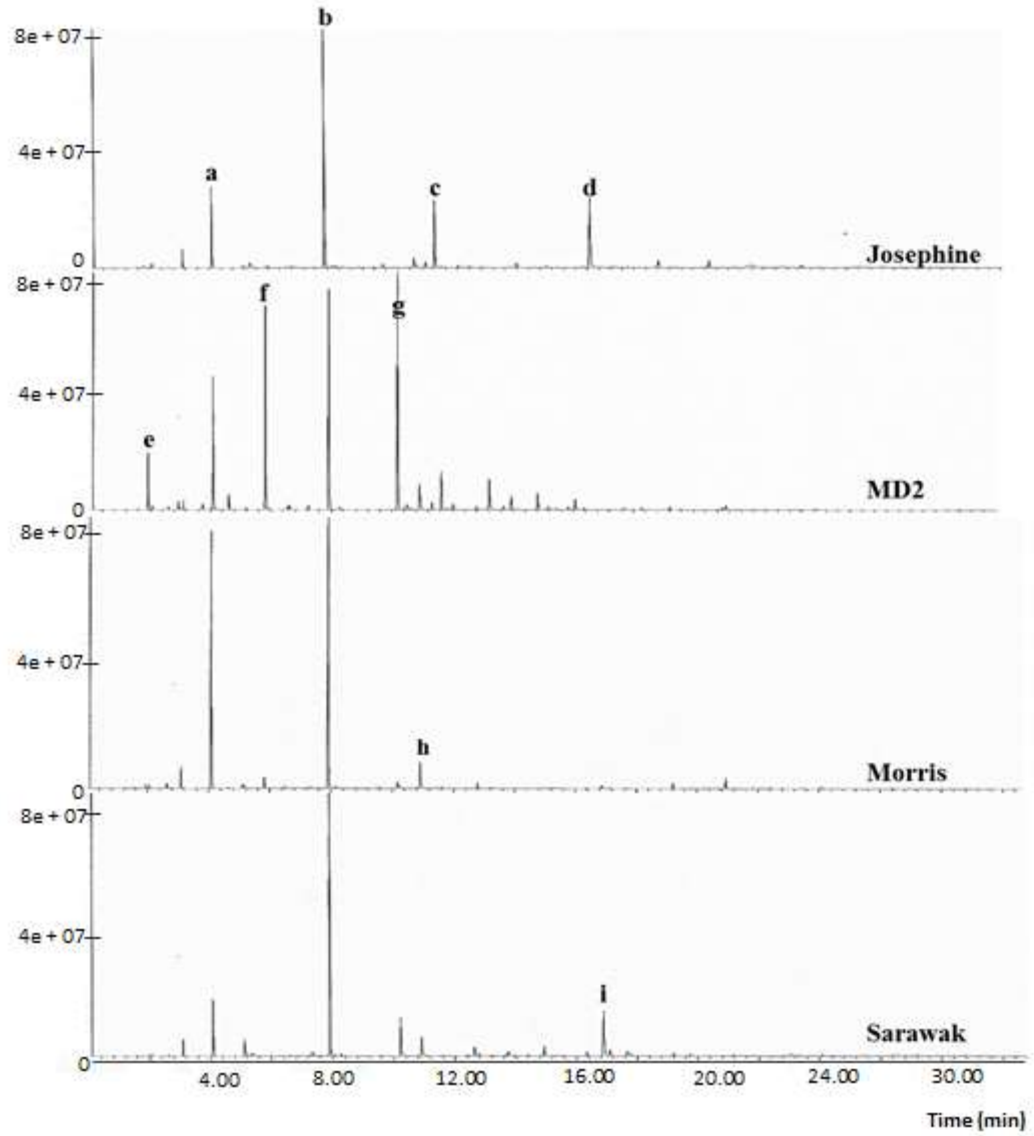

Figure 3. Chromatograms of volatile organic compounds found in pineapple varieties (a) Methyl 2-methylbutanoate (b) methyl hexanoate (c) methyl 2,4-hexadienoate (d) 5-methylfuran-2-carbaldehyde (e) ethyl ethanoate (f) ethyl 2-methylbutanoate (g) ethyl hexanoate (h) methyl octanoate and (i) ethyl octanoate separated by GC-MS using optimized SPME conditions.

Chemometric analysis was performed due to the large amount of complex data generated from chromatographic and sensory analyses. Data on phenolic compositions and volatile organic compounds in pineapple varieties were subjected to chemometric analysis to provide a more meaningful interpretation of the results. In this study, HCA using Euclidian distance and Ward's method as similarity criterion were carried out to cluster the homogeneous samples based on sensory, volatile, and phenolic compositions dat sets of the four different pineapple varieties. The result of a hierarchical clustering procedure is displayed graphically using a tree diagram known as a dendrogram (Figure 4), with four major clusters observed. Figure 4a represents the clustering of phenolic compounds into Sarawak (Cluster 1), Josephine (Cluster 2), Morris (Cluster 3), and MD2 (Cluster 4). Using VOCs data, the pineapple varieties were also successfully grouped into four distinct clusters as shown in Figure 4b for the MD2 (Cluster 1), Sarawak (Cluster 2), Josephine (Cluster 3), and Morris (Cluster 4) pineapples.

Table 2. Volatile organic compounds in four pineapple varieties

Coded Compounds	Retention time (min)	Volatile Organic Compounds (VOCs)	Odour Impression
C1	1.288	Cycloocta-1,3,5,7-tetraene	fruity
C2	1.882	Ethyl ethanoate	ether-like
C3	1.964	Trichloromethane	fruity, rum-like
C4	2.068	Methyl propanoate	ethereal, diffusive, fruity floral
C5	2.615	Methyl methylpropanoate	acrid odor,
C6	2.767	Ethyl 2-propenoate	pineapple-like
C7	2.851	Ethyl propanoate	sweet
C8	2.945	Methyl 2-methylpropanoate	fruity notes resembling apples or pineapples
C9	3.106	Methyl butanoate	fruity
C10	3.748	Ethyl 2-methylpropanoate	fruity or floral smell
C11	4.054	2-Methylpropyl ethanoate	ethereal ester fruity apple
C12	4.082	Methyl 2-methylbutanoate	fruity odor resembling pineapple
C13	4.63	Ethyl butanoate	fruity odor, ethereal
C14	4.953	Butyl ethanoate	sweet
C15	5.132	Tris(trimethylsilyl)arsane	n.d
C16	5.175	Methyl pentanoate	sweet, almond, and baked
C17	5.363	Furan-2-carbaldehyde	sweet, apple, and fruity
C18	5.851	Ethyl 2-methylbutanoate	pungent, gassy
C19	5.924	Ethyl 3-methylbutanoate	sweet, banana, and bitter
C20	6.563	3-Methylbutyl ethanoate	fruity and juicy
C21	6.639	2-Methylbutyl acetate	sweet smell, balsamic, and floral.
C22	6.865	Ethenylbenzene	sweet, apple and fruity
C23	7.241	Ethyl pentanoate	sweet, fruity
C24	7.362	Oxime-,methoxy-phenyl	n.d
C25	7.553	Methyl 5-hexenoate	sweet,
C26	7.901	Methyl hexanoate	apple-like odor
C27	8.075	Dimethyl propanedioate	sweet, camphor, earthy
C28	8.117	1R-α-Pinene	fruity, sweet, and earthy
C29	8.156	Methyl (*Z*)-3-hexenoate	almond, burnt sugar, and caramel
C30	9.063	5-methylfuran-2-carbaldehyde	floral
C31	8.29	Methyl (*E*)-3-hexenoate	fatty, green, fruity, and reminiscent of pineapple
C32	9.15	Methyl (*E*)-2-hexenoate	fatty, cheesy, and waxy
C33	9.58	Hexanoic acid	earthy, fruity, and clove-like (pungent in high concentration)
C34	9.88	ß-myrcene	sweet
C35	10.205	Ethyl hexanoate	sweet smell (like pear drops) and fruity

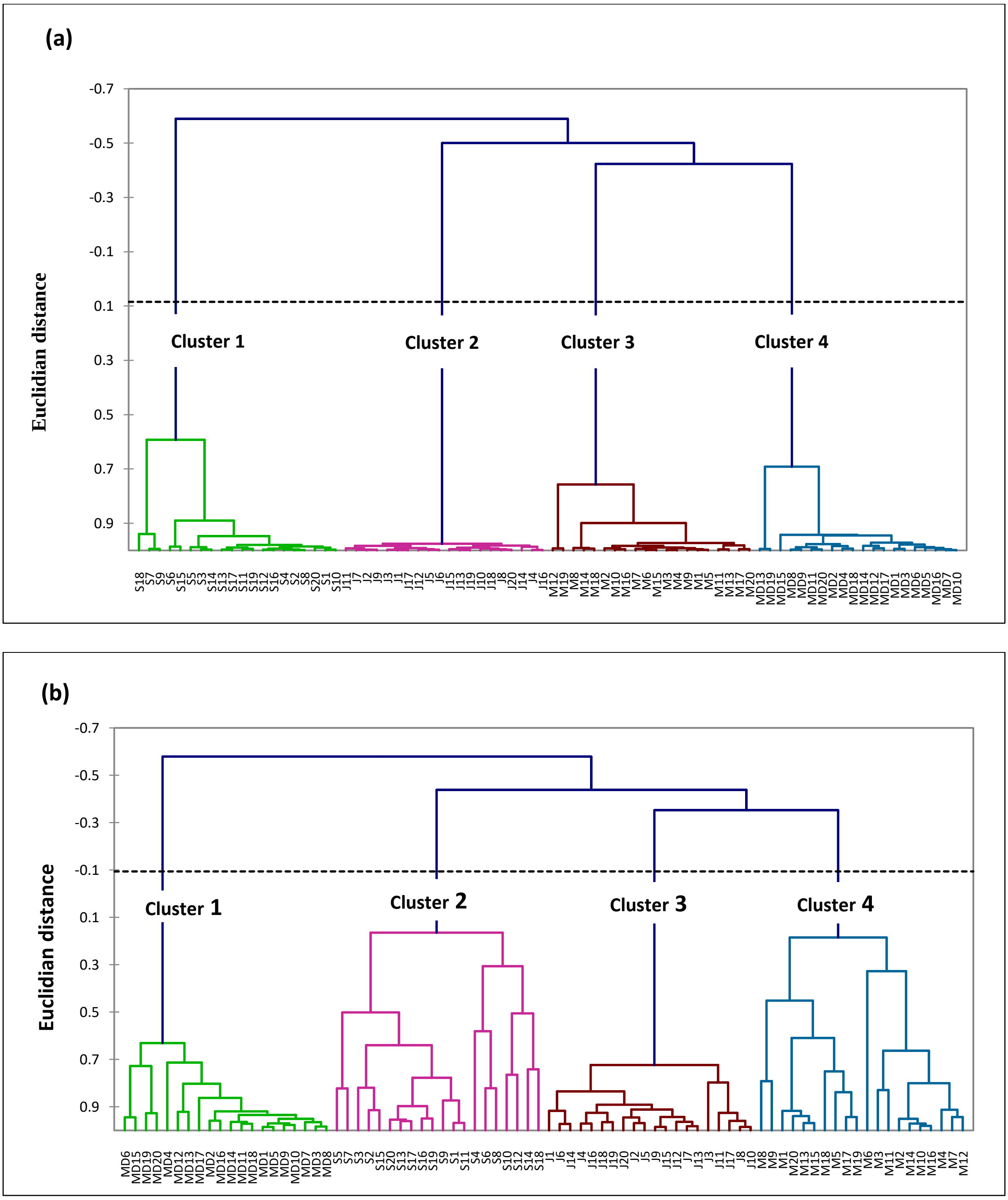

Figure 4. Hierarchical Dendrograms Showing Clusters for (a) Phenolic Compounds and (b) VOCs of the Pineapple Varieties based on 80 pineapple samples

PCA is the most powerful pattern recognition technique that is usually coupled with HCA. PCA and HCA are pattern recognition methods that generate potent visualization tools (dendrogram for HCA and score plots for PCA). In this study, HCA was used for classifying pineapple samples into its varieties and PCA was used primarily to find principal components which describe each variable [27]. PCA analyses were performed on the same data sets to determine which parameter explains how each sample was different from one another [28]. Four distinctive clusters were observed (Figure 5), each one corresponding to different pineapple varieties (Josephine, Sarawak, Morris, and MD2) according to their selected phenolic compounds and were very similar to those clustered by HCA.

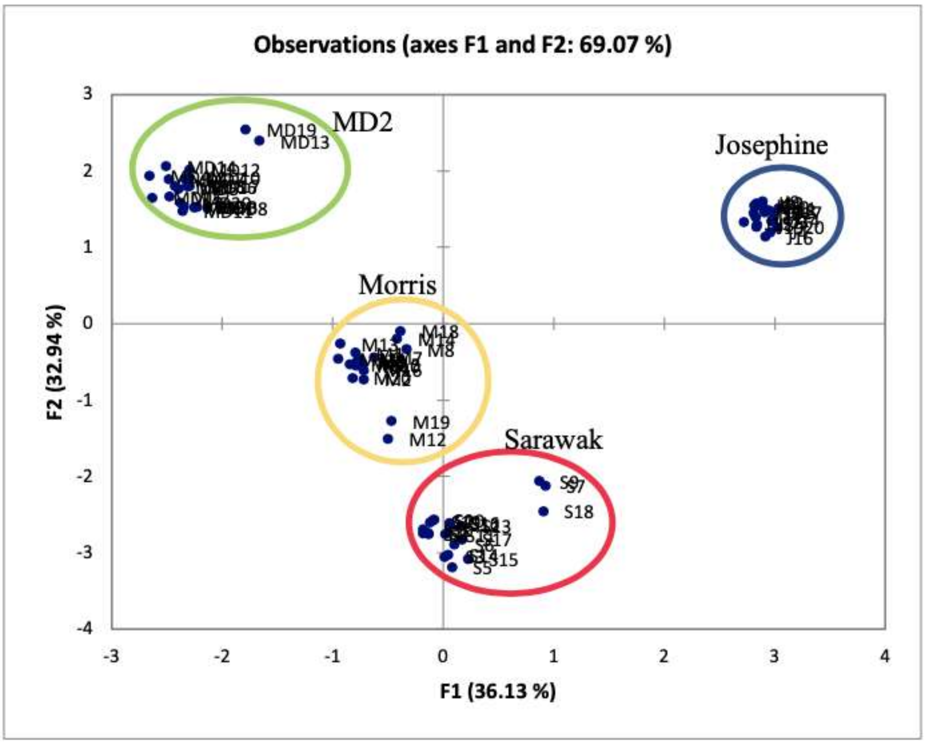

Figure 5. PCA scores of the first two principal components (axes) of phenolic compounds in pineapple varieties

Table 3. Factor Loadings of Selected Phenolic Compounds in Pineapple Varieties

Parameters	PC1	PC2	PC3
Gallic Acid	-0.461	*0.589*	*-0.646*
Catechin	*-0.706*	-0.356	*0.585*
Epicatechin	*-0.515*	**0.848**	-0.015
p-coumaric acid	*-0.749*	0.289	*-0.588*
Caffeic acid	**-0.833**	-0.199	0.488
Rutin	0.141	*-0.700*	*-0.687*
Ferulic acid	**0.875**	0.418	0.209
Quercetin	-0.249	**0.924**	0.113
Naringenin	*-0.646*	-0.477	0.397
Bromelain	0.340	0.481	**0.801**
Eigenvalue	3.613	3.294	2.672
Variability (%)	36.127	32.944	26.724
Cumulative (%)	36.127	69.071	95.795

Note: Strong loadings (> 0.75) are shown in bold; moderate loading (0.5-0.75) in italic

Table 3 tabulates the factor loadings of phenolic compounds. Parameters with high loading on PC1 were assigned for phenolic acid (ferulic acid and caffeic acid), known as the most common cinnamic acid derivatives. Catechin, epicatechin, *p*-coumaric acid, and naringenin show moderate loading in this PC with a total variance of 36.127%. PC2 which accounted for 32.9% variance was positively related to epicatechin and quercetin, while moderate loadings of gallic acid and rutin were recorded. Particularly, high positive loadings on PC2, corresponding to epicatechin and quercetinm belong to the flavonoids, the phenolic compounds that contribute to the body's defense against degenerative diseases such as cancer and cardiovascular diseases. PC3 (26.7% variance) is associated primarily with strong positive loading of bromelain and correlated moderate loadings of gallic acid, catechin, p-coumaric acid, and rutin. The phenolic compounds extracted from three principal components contribute to the clear discrimination of four different pineapple varieties.

PCA analysis of volatile organic compounds was developed using 80 samples which results in four separated groups for the four different pineapple varieties: Josephine, Sarawak, Morris, and MD2, respectively. The first four principal components with eigenvalue >1 explained 83.92% of data variance. Table 4 tabulates the factor loadings of volatile organic compounds where four factors were extracted having a total variance of 83.16%, with contributions of 41.33%, 22.01%, 17.12%, and 3.50% respectively. The highest total variance (41.33%) of the first principal component, PC1 showed strong positive loadings for C2, C5, C7, C8, C11, C12, C13, C16, C18, C20, C21, C23, C27, and C34. The second principal component (PC2) with a total variance of 22.01% recorded the highest positive loadings for C1, C4, C9, C19, C28, and C33. The third principal component (PC3) contributed 17.12% of the total variance showed a strong positive loading of C3 and C22 and positive loadings for C6 and C29. The fourth principal component (PC4) accouting for 3.50% of the variance showed high positive loadings for C31.

Table 4. Factor Loadings for Volatile Organic Compounds in Pineapple Varieties

Coded Parameters	PC1	PC2	PC3	PC4
C1	-0.491	**0.817**	-0.143	0.039
C2	**0.965**	0.145	-0.152	0.014
C3	-0.297	*-0.650*	*-0.615*	0.044
C4	-0.285	**0.902**	-0.242	0.043
C5	**0.926**	-0.045	-0.257	-0.121
C6	-0.370	0.071	**0.811**	-0.025
C7	**0.960**	0.166	-0.128	0.018
C8	**0.957**	0.165	-0.128	0.019
C9	-0.154	**0.877**	-0.390	0.035
C10	-0.285	*-0.613*	*0.583*	-0.013
C11	**0.959**	0.164	-0.130	0.012
C12	**0.891**	-0.188	0.067	0.020
C13	**0.964**	0.070	-0.230	0.012
C14	-0.398	*-0.711*	-0.400	-0.358
C15	0.465	-0.191	**0.794**	-0.038
C16	**0.829**	0.100	0.255	0.038
C17	-0.271	*-0.600*	-0.105	0.230
C18	**0.891**	0.196	-0.139	0.022
C19	*-0.624*	*0.585*	-0.393	-0.087
C20	**0.952**	0.161	0.158	0.011
C21	**0.832**	0.121	0.433	-0.001
C22	-0.411	-0.470	*-0.655*	0.010
C23	**0.882**	0.155	-0.108	0.029
C24	-0.236	-0.482	-0.458	0.289
C25	-0.043	**0.913**	0.289	0.045
C26	-0.012	-0.290	**0.778**	-0.004
C27	**0.896**	-0.154	0.285	0.062
C28	-0.432	*0.732*	-0.118	0.033
C29	-0.427	0.360	*0.726*	0.025
C30	-0.213	-0.491	*0.555*	0.096
C31	-0.060	-0.095	-0.056	**0.943**
C32	-0.296	-0.002	*0.690*	0.008
C33	-0.483	**0.784**	0.080	0.004
C34	**0.929**	0.161	-0.113	0.015
C35	0.465	-0.443	-0.411	-0.069
Eigenvalue	14.466	7.703	5.993	1.211
Variability (%)	41.331	22.007	17.122	3.459
Cumulative (%)	41.331	63.338	80.460	83.919

Note: Strong loadings (> 0.75) are shown in bold; moderate loading (0.5-0.75) in italic

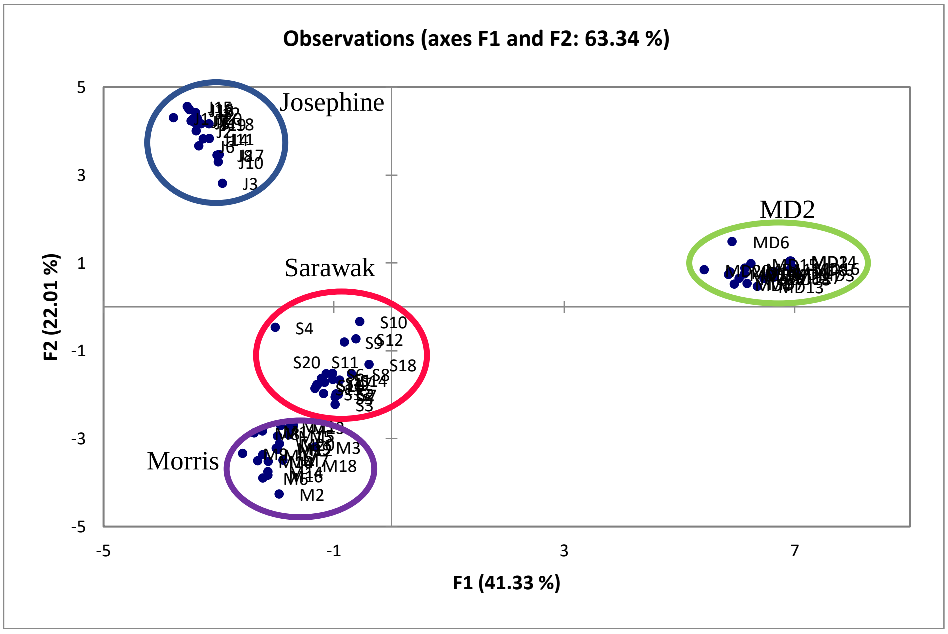

Figure 6. PCA scores of the first two principal components (axes) of volatile organic compounds in pineapple varieties

Performing PCA on the VOC data revealed that 63.34 % of the variation was explained by PC1 and PC2, representing the validity of the VOC profiles to develop robust PCA models. The score plot (Figure 6) shows the clustering of pineapple varieties according to their volatile organic compounds. PCA provides a clear visual relationship between phenolic and volatile compounds and pineapple varieties. 10 phenolic compounds and 35 VOCs successfully discriminated the four different pineapples varieties according to their varieties.

The same data set was applied to discriminant analysis. Discriminant analysis (DA) was performed using a set of observations for which the classes are known. The pineapple varieties were treated as dependent variables, while the 10 selected phenolic compounds and 35 VOCs were considered as independent variables. DA was carried out via three modes which were standard, forward stepwise, and backward stepwise modes.

In DA samples sharing common properties will be grouped together in the same group with a high correct percentage. Effective DA will classify according to correct and incorrect yields. The classification matrix with 100% correct assignation for standard DA mode is shown in Table 5. The results of DA in forward and backward stepwise mode also recorded 100% correct classification of the phenolic and VOCs according to the pineapple varieties. Discriminant plot rendered clear distinction between pineapple varieties for selected phenolic compounds (Figure 7a) and VOCs (Figure 7b) thus, it can be used to classify pineapple according to its variety.

As mentioned earlier, the phenolic and volatile compositions could be more useful in explaining the sensory quality of fruits. The correlation loadings are displayed as vectors in the PLS-DA biplots (Figures 8 and 9), which illustrate the importance of these variables as a discriminative potential marker for each pineapple variety. Potential marker compounds for explaining the dependent *Y*-variables (pineapple varieties) were determined from VIP (variable importance in the projection values) [29]. VIP represents a degree of how much each variable contributes to the PLS-DA model and its importance in class separation of which, variables with VIP > 0.8 were significant for the overall model followed by moderately significant VIP (0.8 - 0.5) and VIP < 0.5 indicated insignificant variables [30]. Variables with VIP higher than 0.80 for phenolic compounds and volatile organic compounds are summarized in Table 6 and 7, respectively.

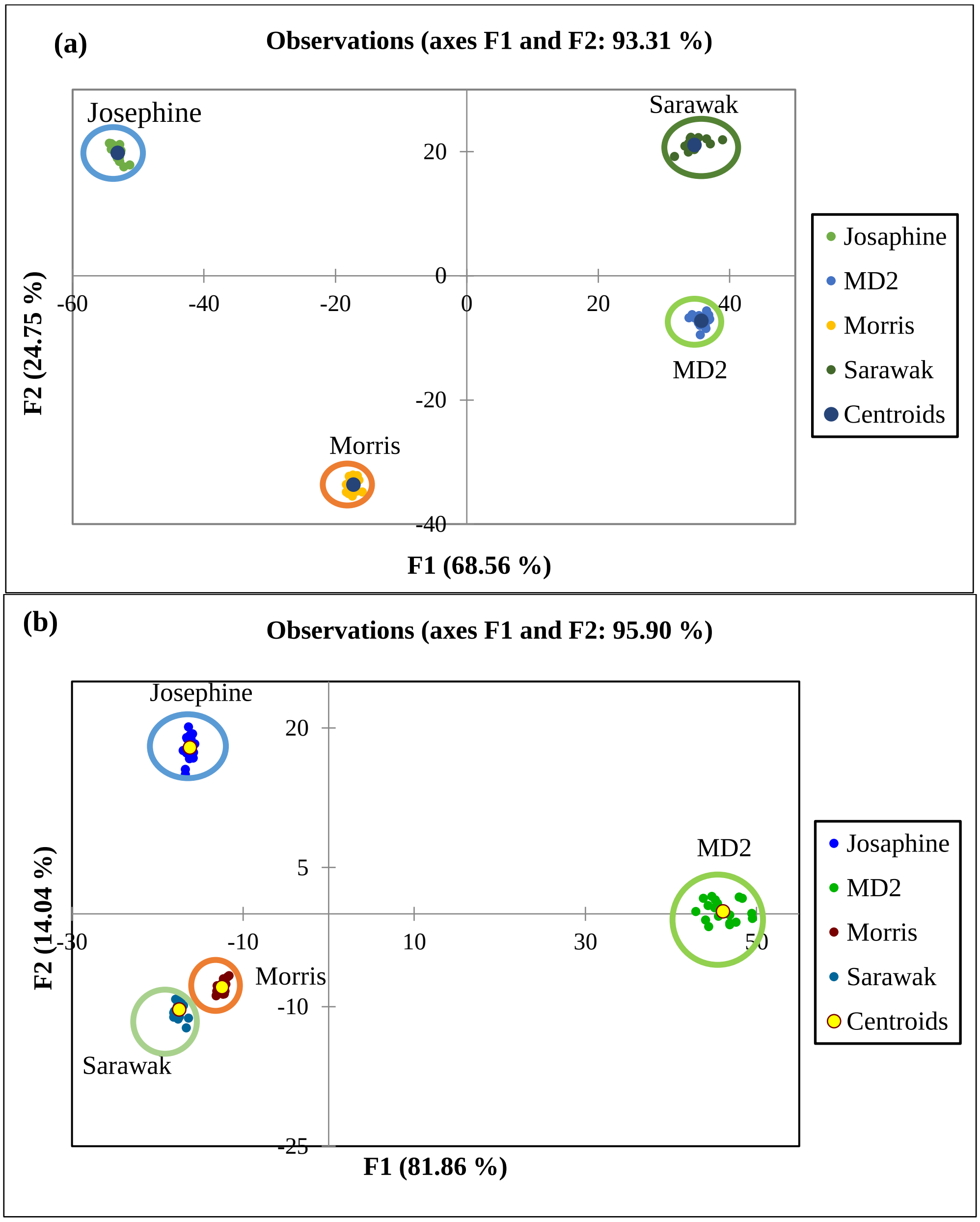

Figure 7. The plot of a discriminant function of (a) phenolic compounds and (b) volatile organic compounds for pineapple varieties (Sarawak, Morris, MD2, and Josephine)

Chewiness, floral and fruity are the discriminating variables of Josephine with ferulic acid as the discriminative phenolic compounds. Juiciness, sweet, sweaty aroma are the sensory attributes while, epicatechin, and caffeic acid are the marker compounds of MD2 pineapple. Morris and Sarawak pineapple are located in the same region with sour and astringency sensory attributes with rutin as the discriminating variable (Figure 8). PLS-DA biplots based on the composition of VOCs and sensory attributes can be used to identify the marker compounds of each pineapple variety. Morris and Sarawak were separated with C26 (VIP = 0.837) as the discriminating variable of Sarawak. Morris pineapple significantly correlated with C6, C32, C22, C31, and C3 as the marker compounds. Volatile compounds (C11, C18, C23, C8, C7, C34, C2, C13, C20, C16, C21, C12, C5, and C27) were able to discriminate MD2 pineapple. Discriminative volatile compounds for Josephine include C9, C4, C1, C33, C28, and C19. A comprehensive and systematic classification of the four pineapple varieties was established by combining the 13 sensory attributes, 10 selected phenolic compounds and 35 VOCs.

Table 5. Classification matrix (CM) of Standard DA Mode for Pineapple Varieties based on Phenolic Compounds and Volatile Organic Compounds (VOCs)

Pineapple varieties	% Correct	Pineapple varieties assigned by DA			
		Josephine	MD2	Morris	Sarawak
Phenolic compounds					
Josephine	100	20	0	0	0
MD2	100	0	20	0	0
Morris	100	0	0	20	0
Sarawak	100	0	0	0	20
Total	100	20	20	20	20
Volatile organic compounds (VOCs)					
Josephine	100	20	0	0	0
MD2	100	0	20	0	0
Morris	100	0	0	20	0
Sarawak	100	0	0	0	20
Total	100	20	20	20	20

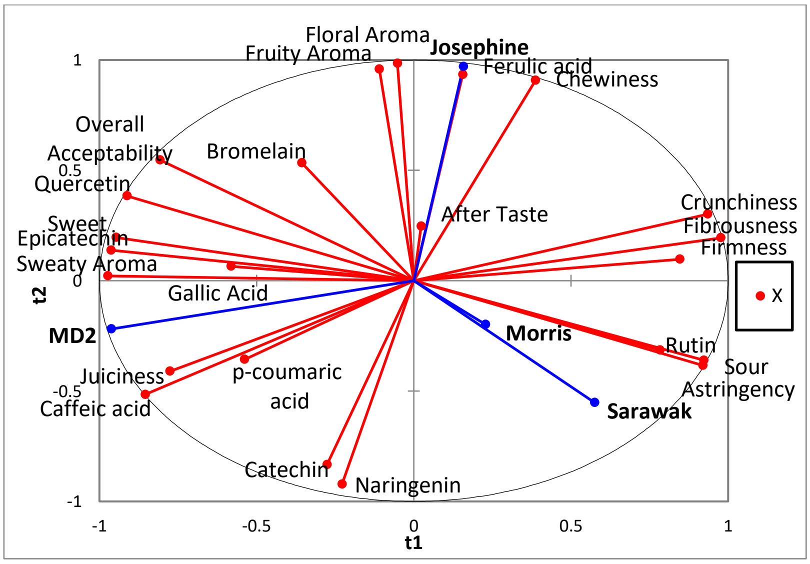

Figure 8. PLS-DA biplots based on the composition of phenolic compounds and sensory attributes of four pineapple varieties (Josephine, Morris, Sarawak, and MD2)

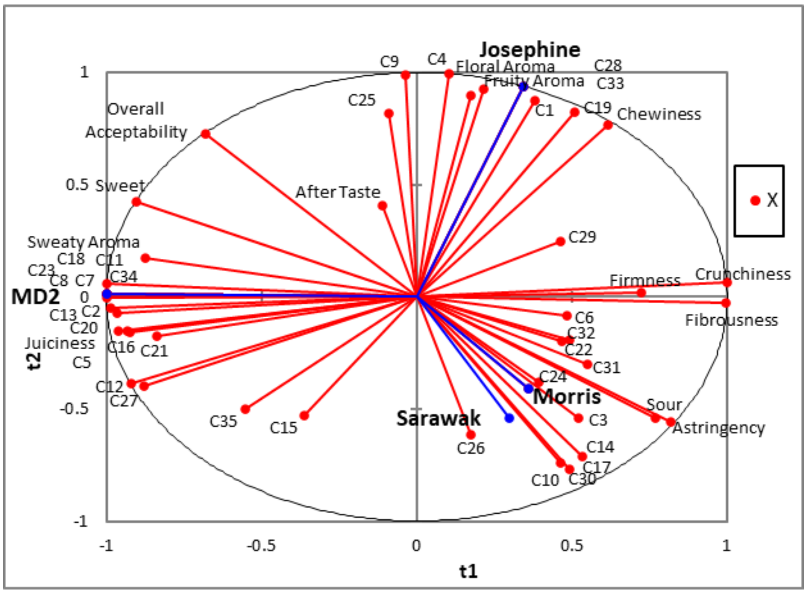

Figure 9. PLS-DA biplots based on the composition of volatile compounds and sensory attributes of four pineapple varieties (Josephine, Morris, Sarawak, and MD2)

Table 6. Potential marker compounds (VIP > 0.80) based on sensory attributes and phenolic compounds identified via PLS-DA to differentiate between four different pineapple varieties

Variable	VIP Value	Variable	VIP Value
Floral Aroma	1.258	Sour	1.031
Fruity Aroma	1.229	Crunchiness	1.011
Chewiness	1.221	Fibrousness	1.007
Ferulic acid	1.202	Sweet	0.979
Naringenin	1.198	Epicatechin	0.978
Catechin	1.099	Sweaty Aroma	0.973
Caffeic acid	1.078	Juiciness	0.936
Overall Acceptability	1.067	Rutin	0.879
Astringency	1.042	Firmness	0.855
Quercetin	1.035		

Table 7. Potential marker compounds (VIP > 0.80) based on sensory attributes and volatile organic compounds identified via PLS-DA to differentiate between four different pineapple varieties

Variable	VIP Value	Variable	VIP Value
C4	1.333	C18	1.005
C9	1.326	C34	1.004
C1	1.297	C7	1.004
C28	1.297	C8	1.004
Floral Aroma	1.259	C11	1.004
C33	1.231	C23	1.004
C19	1.216	C2	1.002
Fruity Aroma	1.212	Fibrousness	0.999
Chewiness	1.197	C13	0.999
Overall Acceptability	1.192	C5	0.988
C10	1.143	C20	0.976
Sour	1.110	Juiciness	0.957
C25	1.099	C16	0.953
C17	1.093	Sweaty Aroma	0.911
C30	1.091	C14	0.894
Sweet	1.073	C21	0.872
C12	1.059	C35	0.868
Astringency	1.058	C26	0.837
C27	1.031	C15	0.800
Crunchiness			

CONCLUSION

The combination of sensory attributes, phenolic, and volatile compounds provided useful information in classifying pineapple samples according to their varieties. Due to the large and complex data generated, chemometric techniques were applied. Although the chromatographic profiling of the four pineapple varieties was almost similar, systematic classification was obtained by combining the data of volatile compositions, phenolic compounds, and sensory with that of chemometric techniques. The application of chemometric approaches together with a PLS-DA showed the potential marker compounds for each pineapple variety, permitting the unambiguous distinction between Morris, Josephine, MD2, and Sarawak pineapples.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support obtained from the Ministry of Education Malaysia for this project (Project number: FRGS 1/2018/STG01/UiTM/02/8) and Universiti Teknologi MARA.

REFERENCES

1. Saad, R., Pohyeen, T., Khan, J., Wenji, L., Sultan, S. and Abdul Hameed, J. (2014) Phytochemical screening and antioxidant activity of different parts from five Malaysian herbs. *International Journal of Sscience and Technology*, **19(2)**, 1336–1347.

2. Thalip, A. A., Tong, P. S. and Ng, C. (2015) The MD2 “Super Sweet” Pineapple (*Ananas comosus*). *Utar Agriculture Science Journal*, **1**, 15–17.

3. Silva, B. M., Andrade, P. B., Valentão, P., Mendes, G. C., Seabra, R. M. and Ferreira, M. A. (2000) Phenolic profile in the evaluation of commercial quince jellies authenticity. *Food Chemistry*, **71(2)**, 281–285.

4. Hashim, M., Hidekazu, Y., Takeshi, M. and Ming, Y. (2014) Application case study of AP1000 automatic depressurization system (ADS) for reliability evaluation by GO-FLOW methodology. *Nuclear Engineering and Design*, **278**, 209–221.

5. López, A., Sánchez, A. H., Cortés-Delgado, A., de Castro, A. and Montaño, A. (2018) Relating sensory analysis with SPME-GC-MS data for Spanish-style green table olive aroma profiling. *LWT*, **89**, 725–734.

6. Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L., Calvo, A. and Busto, O. (2016) Olive oil sensory defects classification with data fusion of instrumental techniques and multivariate analysis (PLS-DA). *Food Chemistry*, **203**, 314–322.

7. Lawless, H. T. and Heymann, H. (2010) Descriptive Analysis in Sensory Evaluation of Food. *Springer, New York, NY*.

8. Steingass, C. B., Grauwet, T. and Carle, R. (2014) Influence of harvest maturity and fruit logistics on pineapple (*Ananas comosus* [L.] Merr.) volatiles assessed by headspace solid phase microextraction and gas chromatography–mass spectrometry (HS-SPME-GC/MS). *Food Chemistry*, **150**, 382–391.

9. Zheng, L. Y., Sun, G. M., Liu, Y. G., Lv, L. L., Yang, W. X., Zhao, W. F. and Wei, C. B. (2012) Aroma volatile compounds from two fresh pineapple varieties in China. *International Journal of Molecular Sciences*, **13(6)**, 7383–7392.

10. Gan, F. and Ye, R. (2006) New approach on similarity analysis of chromatographic fingerprint of herbal medicine. *Journal of Chromatography A*, **1104(1-2)**, 100–105.

11. Lee, H., Grosse, R., Ranganath, R. and Ng, A. Y. (2009) Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. In *Proceedings of the 26th Annual International Conference on Machine Learning*, 609–616.

12. Chun, M. H., Kim, E. K., Yu, S. M., Oh, M. S., Moon, K. Y., Jung, J. H. and Hong, J. (2011) GC/MS combined with chemometrics methods for quality control of *Schizonepeta tenuifolia* Briq: Determination of essential oils. *Microchemical Journal*, **97(2)**, 274–281.

13. Tistaert, C., Dejaegher, B. and Vander Heyden, Y. (2011) Chromatographic separation techniques and data handling methods for herbal fingerprints: A review. *Analytica Chimica Acta*, **690(2)**, 148–161.

14. Huang, Y., Wu, Z., Su, R., Ruan, G., Du, F. and Li, G. (2016) Current application of chemometrics in traditional Chinese herbal medicine research. *Journal of Chromatography B*, **1026**, 27–35.

15. Oliveri, P. and Simonetti, R. (2016) Chemometrics for food authenticity applications. In *Advances in food authenticity testing*,701–728. Woodhead Publishing.

16. Liu, N., Koot, A., Hettinga, K., De Jong, J. and van Ruth, S. M. (2018) Portraying and tracing the impact of different production systems on the

volatile organic compound composition of milk by PTR-(Quad) MS and PTR-(ToF) MS. *Food Chemistry*, **239**, 201–207.

17. Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L. and Maggio, R. M. (2018) Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends in Food Science & Technology*, **72**, 83–90.

18. Wang, X., Zeng, Q., del Mar Contreras, M. and Wang, L. (2017) Profiling and quantification of phenolic compounds in Camellia seed oils: Natural tea polyphenols in vegetable oil. *Food Research International*, **102**, 184–194.

19. DIN 10967-1. (1999) Sensory analysis for Investigation of profiles-Part 1: Conventional profiling.

20. Khan, A. A. M., Saim, N. and Hamid, R. D. (2017) Optimization of pressurized liquid extraction of bioactive compounds from *Ananas comosus* (pineapple) fruit. *Pertanika J. Sci. Technol*, **25**, 175–182.

21. Zainuddin, S. F., Zakaria, S. R., Saim, N., Dol Hamid, R. and Osman, R. (2020) Optimization of Headspace Solid Phase Microextraction (HS-SPME) for the Extraction of Volatile Organic Compounds (VOCs) in MD2 Pineapple. *Science Letters*, **14(2)**, 58–70.

22. Chmiel, T., Kupska, M., Wardencki, W. and Namieśnik, J. (2017) Application of response surface methodology to optimize solid-phase microextraction procedure for chromatographic determination of aroma-active monoterpenes in berries. *Food Chemistry*, **221**, 1041–1056.

23. Zakaria, S. R., Tajuddin, R., Osman, R., Saim, N. and Saaid, M. (2017) Optimization of Headspace Solid Phase Microextraction (HS-SPME) for the Extraction of Volatile Organic Compounds (VOCs) in Mangoes (*Harumanis* cv.) Using 2 Stages Multivariate Analysis. *Pertanika Journal of Science & Technology*, **25**:167–174.

24. Bell, L., Methven, L., Signore, A., Oruna-Concha, M. J. and Wagstaff, C. (2017) Analysis of seven salad rocket (*Eruca sativa*) accessions: The relationships between sensory attributes and volatile and non-volatile compounds. *Food Chemistry*, **218**, 181–191.

25. Sinelli, N., Cerretani, L., Di Egidio, V., Bendini, A. and Casiraghi, E. (2010) Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the basis of fruity attribute intensity. *Food research international*, **43(1)**, 369–375.

26. Lasekan, O. and Hussein, F. K. (2018) Classification of different pineapple varieties grown in Malaysia based on volatile fingerprinting and sensory analysis. *Chemistry Central Journal*, **12(1)**, 140.

27. Ferreira, J. E., da Costa, C. H., de Miranda, R. M., Figueiredo, A. F. and Ginarte, Y. M. (2014) An interdisciplinary study on monosubstituted benzene involving computation, statistics and chemistry. *Educación Química*, ***25*(4),** 418–424.

28. Sánchez-Salcedo, E. M., Tassotti, M., Del Rio, D., Hernández, F., Martínez, J. J. and Mena, P. (2016) (Poly) phenolic fingerprint and chemometric analysis of white (*Morus alba* L.) and black (*Morus nigra* L.) mulberry leaves by using a non-targeted UHPLC–MS approach. *Food Chemistry*,**212**, 250–255.

29. Jumtee, K., Komura, H., Bamba, T. and Fukusaki, E. (2011) Predication of Japanese green tea (Sen-cha) ranking by volatile profiling using gas chromatography-mass spectrometry and multivariate analysis. *Journal of bioscience and bioengineering*, **112(3)**, 252–255.

30. Jayadev, C., Hulley, P., Swales, C., Snelling, S., Collins, G., Taylor, P. and Price, A. (2020) Synovial fluid fingerprinting in end-stage knee osteoarthritis: a novel biomarker concept. *Bone & Joint Research*, ***9*(9)**, 623–632.