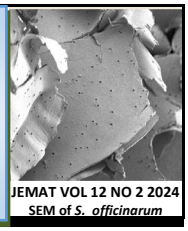
JEMAT 2024, Vol 12, No 2, 13-18
https://doi.org/10.54987/jemat.v12i2.1004

HIBISCUS PUBLISHER

JOURNAL OF ENVIRONMENTAL MICROBIOLOGY AND TOXICOLOGY

Website: http://journal.hibiscuspublisher.com/index.php/JEMAT/index



Clustering Analysis and Differentiation of Lard from Palm Oil and Soybean Oil Based on N-Alkane and Triacylglycerides Profiles Using Chemometric Analysis

Helmi Wasoh[1,2]*, Nur Ain Syaqirah Sapian[1], Ahmad Muhsin Mohammad Fatin[2], Murni Halim[2], Nor'Aini Abdul Rahman[2], Mohamad Zulfazli Mohd Sobri[2], Amalia Mohd Hashim[1,3], Mohd Termizi Yusof[3], Mohd Sabri Pak-Dek[4], Yanty Noorzianna Abdul Manaf[5], Mohd Nasir Mohd Desa[1,6] and Firdaus Mohamad Hamzah[7]

[1]Halal Products Research Institute, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[2]Department of Bioprocess Technology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[3]Department of Microbiology, Faculty of Biotechnology and Biomolecular Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[4]Department of Food Science, Faculty of Food Science and Technology, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[5]Halal Research Group, Faculty of Food Science and Nutrition, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia.
[6]Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
[7]Department of Mathematics, Faculty of Defence Science and Technology, Universiti Pertahanan Nasional Malaysia, Kem Sungai Besi, 57000 Sungai Besi, Kuala Lumpur, Malaysia.

*Corresponding author:
Helmi Wasoh,
Department of Bioprocess Technology,
Faculty of Biotechnology and Biomolecular Sciences,
Universiti Putra Malaysia,
43400 UPM Serdang,
Selangor,
Malaysia.
Email: helmi_wmi@upm.edu.my

HISTORY

Received: 18th Aug 2024
Received in revised form: 24th Nov 2024
Accepted: 14th Dec 2024

KEYWORDS

Triacylglycerides (TAGs)
n-Alkane profiling
Chemometrics techniques
Gas Chromatography-Mass Spectrometry
Multivariate analysis

ABSTRACT

This investigation focused on analyzing lard derived from pig adipose tissue and two different vegetable oils (crude palm oil and soybean oil) through triacylglycerides (TAGs) and n-alkane profiling, employing chemometrics techniques. Unsaponifiable palm and soybean oils were isolated and examined using gas chromatography-mass spectrometry (GC-MS) to assess their n-alkane profiles. The identified n-alkane profiles were verified by comparing them with n-alkane standards (nC_{08}–nC_{27}). The experimental design incorporated several multivariate techniques, including Hierarchical Clustering Analysis (HCA), Principal Component Analysis (PCA), Random Forest (RF), and Partial Least Squares-Discriminant Analysis (PLS-DA). According to PLS-DA findings, nC_{25}, nC_{27}, and PPL are suggested as key TAG and n-alkane markers for the clustering analysis of lard from crude palm and soybean oils. These results indicate that further research is necessary to refine and validate these distinctions, especially when using chemometrics techniques.

INTRODUCTION

Pig adipose tissue is rendered to produce pig fat, which is also referred to as lard. Because it is more affordable and readily available, lard is preferred by food makers in several nations as an alternative to oil [1]. Lard-rich diets have been linked to coronary heart disease and hypercholesterolemia, among other health hazards [2]. Pig-related foods can spread zoonotic infections such as tularemia, anthrax, and trichinosis because they are prone to bacterial degradation and pigs are thought to be an infection reservoir [3]. Several analytical techniques have been developed to differentiate lard from other fats, including physical and chemical methods [4]. According to Gunstone [5], triglycerides (TAG) make up most of the fats, with minor components such as mono- and diglycerides, free fatty acids (FFA), hydrocarbons (n-alkanes), phosphatides, sterols, fatty alcohols, fat-soluble vitamins, tocopherols, and waxes. Rohman et al. [4], used principal component analysis (PCA) in

conjunction with TAG composition to distinguish lard from other animal fats. Azizan et al. [6] used chemometrics in conjunction with fatty acid content to identify lard adulteration in wheat cookies.

Several analytical techniques have been developed to quantitatively analyze n-alkanes. Gas chromatography coupled with mass spectrometry (GCMS) is one of the most commonly used techniques [7]. Based on our knowledge and a thorough literature review, there are currently no published reports on the application of chemometrics for lard, palm oil, and soybean oil differentiation (which uses data from n-alkane and TAG profiles as matrix variables). Thus, this study aimed to use n-alkane and TAG profiles in conjunction with chemometric analysis to distinguish lard from palm oil and soybean oil. Furthermore, reliability and additional results can be obtained by combining n-alkane and TAG profiling with multivariate chemometric analysis.

MATERIALS AND METHODS

Sample preparation

Pork adipose tissue was used to prepare the lard samples. The other oils (palm oil and soybean oil) were purchased from a local fresh market. The equipment was supplied by the Laboratory of Halal Science Research (Halal Products Research Institute, Universiti Putra Malaysia), and the extraction procedure and storage for preservation were performed using the Soxhlet extraction method [8]. The resulting oils were filtered through a muslin cloth to eliminate contaminants and stored at 4 °C for later use.

Sample extraction

With minor adjustments, the n-alkane fractions of animal fats were produced in accordance with the method described by Troya et al. [7]. First, saponification was used to extract unsaponifiable fractions of each fat. Using a reflux condenser and heating mantle, the desired fat samples (5 g) were saponified by reflux with an ethanolic potassium hydroxide solution for 20 min (70 °C). After cooling the mixture, 25 mL of distilled water was added, and the mixture was placed in a separating funnel. Non-saponifiable materials were extracted twice using 25 mL of n-hexane.

After mixing the hexane extracts and adding them to a different separating funnel, three washing procedures were performed using a mixture of ethanol and water (12.5 mL). The extracts were dried using sodium sulfate anhydrous and were then subjected to drying over anhydrous sodium sulfate and subsequently evaporated with a rotary evaporator (30 °C under vacuum - EYELA, Japan). The resulting residues were dissolved in 2 mL of n-hexane to facilitate separation.

Glass column (internal diameter of 1.5 cm and a length of 40 cm), which was packed with 15 g of silica gel in n-hexane (as the stationary phase) was used to isolate the hydrocarbon fractions. Chromatographic elution was performed using 40 mL n-hexane as the mobile phase, and the eluate (20 mL) obtained from the column was then identified as the n-alkane fraction. The eluate was evaporated again using a rotary evaporator (EYELA, Japan) under a vacuum at 30 °C. The final sample was promptly dissolved in n-hexane (0.5 mL) and stored in a chiller at 4 °C prior to subsequent GC-MS analysis.

GC-MS

The analysis of n-alkanes in the lard and various oils was conducted using an Agilent 7890A gas chromatography system (Agilent 5975C mass spectrometry detector, Agilent, USA). The analytical capillary column employed was HP-5MS (30 m length, internal diameter of 0.25 µm and film thickness of 0.25 µm). The temperature profile of the oven commenced at 35 °C for 2 min, followed by a ramp to 250 °C at a rate of 10 °C/min and was subsequently maintained at 300 °C for 23 min at a rate of 20 °C/min.

The temperature of the detector was set at 300 °C. An injection volume of 1 µL was used, operating in split mode. Helium was used as the carrier gas and flowed at a rate of 1 mL/min. Compound identification was achieved by comparing the results with the NIST 11 mass spectral library and a standard of saturated alkanes provided by Supelco (Bellefonte, PA, USA). The area of the standard samples was recorded for each individual n-alkane, including nC_{08} – nC_{27}. Each standard or original n-alkane was administered at a concentration of 1 mg/kg. The composition of each identified n-alkane was calculated by determining the ratio of the peak area of each alkane to that of all alkanes present in the sample [9].

Statistical analysis

Data analysis was conducted by utilizing an online statistical software, MetaboAnalyst 5.0 (McGill University, Anne de Bellevue, QC, Canada; https://www.metaboanalyst.ca). A total of 17 n-alkane samples were analyzed in triplicate and were included in the study. The data for TAGs were adopted from Pacheco and Crapiste [10] and Manaf et al. [11]. The data matrix was subsequently normalized to a constant sum and subjected to range scaling to adjust the variance of the spectral data, ensuring that the peaks were equally weighted for the development of multivariate models [12]. Missing value imputations were performed using a method provided by MetaboAnalyst, where zeros or missing values were substituted with one-fifth of the minimum positive value for each variable.

Lard and other oils were clustered using unsupervised methods, including Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA). PCA was employed to provide an overview of the various patterns and clustering of lard and other oil samples based on their n-alkane and TAGs profiles. This technique summarizes the original dataset by transforming a large data matrix into a lower-dimensional space defined by axes known as principal components (PCs). It facilitated the visualization of both score plots and loading plots, illustrating the relationships among the fat/oil samples and their n-alkane and TAGs distributions, as well as the interrelationships between different types of samples [13]. HCA was executed utilizing the Euclidean distance measure alongside Ward's linkage algorithm. The results of the HCA model were illustrated through a heatmap and a dendrogram [14]. This clustering technique was employed to distinguish lard from other oils by highlighting the contribution of each variable to the respective groups or segments.

For the classification analysis of lard and other oil samples, supervised methods were utilized, specifically partial least squares-discriminant analysis (PLS-DA) and the Random Forest (RF) algorithm. RF consists of a collection of tree predictors, where each tree is generated based on the values of a randomly selected vector, which is chosen independently and follows the same distribution across all trees in the ensemble [15].

The PLS method, in conjunction with discriminant analysis, serves as a classification approach that reduces the dimensionality of the predictor variables [16]. The efficacy of the PLS-DA model was assessed through a 10-fold cross-validation, wherein 10 % of the samples were randomly excluded (training set), while the remaining samples were employed for classifying the omitted samples [17]. To further substantiate the accuracy of the model, a permutation test (1000 random permutations) was conducted by emphasizing the significant variables based on Variable Importance in Projection (VIP) scores.

The VIP score is defined as the weighted sum of squares of PLS loadings, which reflects the amount of explained y-variance for each component within the constructed model [18]. Typically, VIP scores exceeding 1 are chosen as a criterion for the iterative selection of variables, serving as potential n-alkane markers [19]. The normalized data were divided into training and testing sets (9:1) based on a random selection process. The prediction accuracy of the RF model was assessed by creating a confusion matrix.

RESULTS AND DISCUSSION

Univariate analysis

MetaboAnalyst offers one-way Analysis of Variance (ANOVA) before chemometrics analysis. ANOVA simply indicates the significance of the overall comparison. The test-determined significant differences for each variable were less than 0.05 (**Table 1**). As a result, every n-alkane and triacylglycerides (TAGs) found were deemed relevant and chosen for an additional chemometrics investigation. Fisher's Least Significant Difference (LSD) and Tukey's Honestly Significant Difference (HSD) are two popular post-hoc techniques provided by MetaboAnalyst. These univariate analyses offer a first hint of characteristics that may be important in differentiating between the circumstances under investigation.

Table 1. Univariate analysis of original dataset using One-way ANOVA.

	Compounds	F. Stat	p. value	-log10(p)
1	PPO	32242.0	0.00	12.09
2	OLL	29185.00	0.00	11.96
3	PPL	18080.00	0.00	11.34
4	SPO	13871.00	0.00	10.99
5	POO	11091.00	0.00	10.70
6	PLL	9141.70	0.00	10.45
7	PPS	4874.80	0.00	9.63
8	POL	3873.50	0.00	9.33
9	nC13	2607.40	0.00	8.82
10	OOL	2132.40	0.00	8.56
11	OOO	1445.10	0.00	8.05
12	nC21	1068.20	0.00	7.66
13	nC10	760.47	0.00	7.22
14	nC20	704.06	0.00	7.12
15	LLL	692.63	0.00	7.10
16	nC18	579.75	0.00	6.87
17	SSS	552.53	0.00	6.80
18	nC25	457.40	0.00	6.56
19	nC12	414.74	0.00	6.43
20	LLLn	393.49	0.00	6.36
21	nC27	281.30	0.00	5.93
22	SOO	136.72	0.00	5.00
23	nC14	85.83	0.00	4.41
24	nC16	70.00	0.00	4.16
25	nC15	54.32	0.00	3.84
26	nC24	36.22	0.00	3.35
27	SSO	12.38	0.01	2.13
28	nC26	6.62	0.03	1.52
29	nC08	6.55	0.03	1.51

Clustering analysis of lard and other vegetable oils based on n-alkane profiles and TAGs

To perform a clustering analysis of lard and other vegetable oils (crude palm oil and soybean oil), the multivariate data matrix of 14 alkanes and 15 distinct types of TAG molecules in fat and oils was subjected to PCA, HCA and HSL-DA analysis. The capacity of PCA, an unsupervised method, to identify patterns in a dataset is well-known [20]. The projection of samples defined by the first two principal components is depicted in the score plot (**Fig. 1**), where PC1 explained 68.9 % of the variation and PC2 accounted for 24.0 % of the variation, or 92.9 % of the total variance explained. Based on the plot, it is shown that each sample was clustered correctly into its group. The lard, crude palm oil and soybean oil were separated by the PC1 axis and there was no outlier observed.

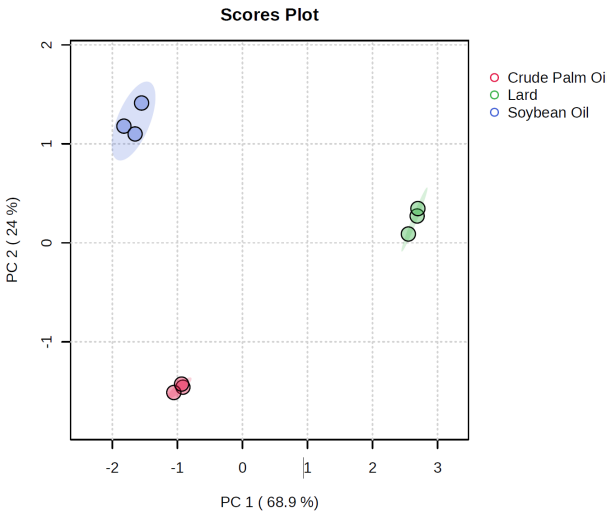


Fig. 1. Principal component analysis (PCA) score plot based on n-alkane profiles and triacylglycerides (TAGs) of lard, crude palm oil and soybean oil.

The PCA biplot displays the contribution of n-alkanes by combining the score plot and loading plot (**Fig. 2**). **Table 2** shows the PCA loading values, which indicate the influence of n-alkanes and TAGs on the clustering of samples. The further the variable is from the origin (0,0), the greater the influence of the variable on the variation of a PCA model [21]. The biplot shows that nC_{20} and nC_{25} had a strong correlation towards lard clustering. Moreover, PPL contributed to a clustering that is highly associated with soybean oil at positive scores on both PCs. Meanwhile, nC_{13}, SSO, SPO and POL contributed to the clustering of crude palm oil. The influence of the alkanes in the biplot was supported in the HCA (Heatmap) (**Fig. 3**).

In HCA, the outcome was visualised as a dendrogram, displaying clusters of samples [6], and the heatmap represents the influence of significant variables [22]. It was clearly shown in the dendrogram that the samples were clustered distinctly into three groups, i.e., lard, crude palm oil and soybean oil (**Fig. 3**). This clustering result was achieved through the combination of the Euclidean distance measure and Ward's linkage algorithm. Based on the dendrogram, all samples were grouped correctly into their sample class. This clustering analysis by HCA also clustered lard samples correctly into the isolated class of fat.

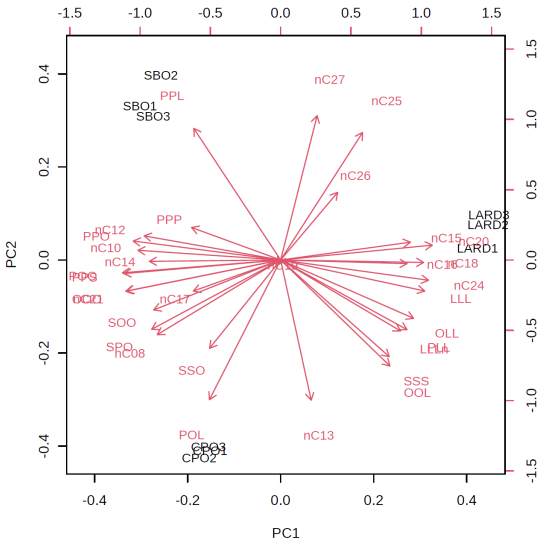


Fig. 2. PCA biplot between PC1 and PC2 that shows the strength of contribution of variables to each PC (arrows).

Table 2. PCA loading value for PC1 and PC2 that relates to lard, crude palm oil (CPO) and soybean oil (SBO) (with the PC1 or PC2 values close or more than 0.2 score).

	PC1	PC2	Samples
nC_{15}	0.1914	0.0444	Lard
nC_{13}	0.0450	-0.3505	CPO
nC_{18}	0.2110	-0.0062	Lard
nC_{20}	0.2238	0.0371	Lard
nC_{24}	0.2182	-0.0507	Lard
nC_{25}	0.1209	0.3187	Lard / SBO
nC_{27}	0.0538	0.3607	Lard / SBO
LLL	0.2126	-0.0775	Lard / CPO
OLL	0.1962	-0.1459	Lard / CPO
OOL	0.1613	-0.2651	Lard / CPO
POL	-0.1050	-0.3492	CPO
PPL	-0.1281	0.3290	SBO
PPO	-0.2174	0.0473	SBO
SPO	-0.1903	-0.1733	CPO
SSO	-0.1048	-0.2208	CPO
SSS	0.1602	-0.2418	Lard / CPO

The heatmap uses a color-coding system to show variation in the abundance of alkanes in different samples. The nC_{15}, nC_{16}, nC_{18}, nC_{20}, nC_{24} and nC_{25} were found notable and abundant in lard but low in other vegetable oils. Majority the distribution of TAGs and n-alkanes in both vegetable oils was more closely related to each other. However, POL, SSO, SPO and nC_{13} were found to be notable to crude palm oil as it was identified as being low in abundance or absent in lard and soybean oil. The PLL, OLL, LLL and nC_{25} were notable in lard. The PPL and nC_{27} were unique in soybean oil. Meanwhile, the alkanes of nC_{25} and nC_{13} were unique in lard and crude palm oil, respectively.

Classification analysis of lard and other vegetable based on n-alkane profiles and triacylglycerides

Classification analysis by supervised methods of partial least squares-discriminant analysis (PLS-DA) model and Random Forest (RF) machine learning were performed using the same data containing 14 significant alkanes and 15 TAGs. **Fig. 4** shows the classification pattern of lard, crude palm oil and soybean oil samples with variance explanation of 61.5 % and 31.4 % of PC1 and PC2, respectively. Based on the PLS-DA score plot, all samples were classified correctly into their sample class, and no outlier was observed.

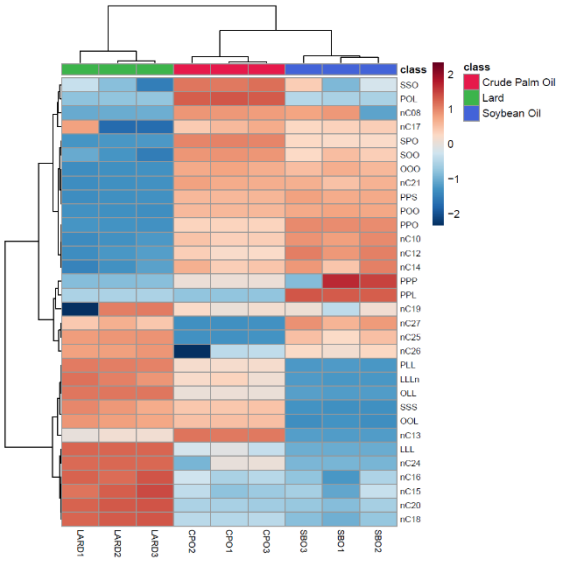


Fig. 3. Hierarchical Clustering Analysis (HCA) dendrogram (with Heatmap) representing the correlation between variables (indicated in rows) and the samples (shown in columns).

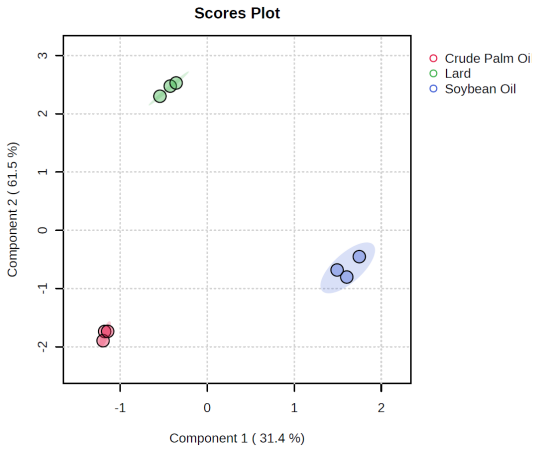


Fig. 4. Partial least squares-discriminant analysis (PLS-DA) score plot (between Component 1 and Component 2).

The PLS-DA model was then subjected to cross validation (CV) to evaluate its performance. By applying 10-fold cross-validation, the optimal number of components needed to build the PLS-DA model in this study, as well as the model predictive power (Q^2 value), were determined. As a result, the three-component model was revealed as the best classifier (**Fig. 5**), with a Q^2 value of 0.9843 (**Table** 3). According to Razali et al. [23], the difference between the R^2 and Q^2 values is used to detect outliers and model overfitting, and the value must not exceed 0.2.

The discrepancy between the R^2 and Q^2 values obtained in this study was 0.0154, showing the absence of outliers and that the PLS-DA model was not overfit. The robustness of the developed PLS-DA model was further evaluated using a response permutation test with 1000 random permutations based on prediction accuracy during training. The permutation test also confirmed that the model was not overfit since the p-value obtained was less than 0.05 (**Fig. 6**).

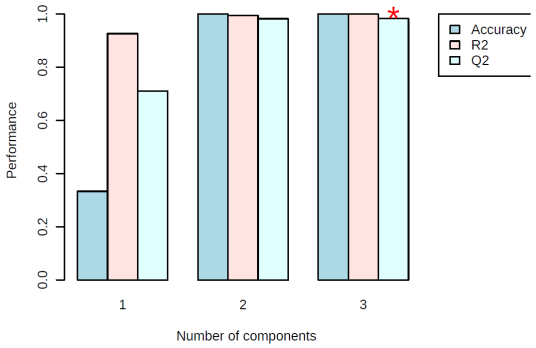


Fig. 5. 10-fold cross validation determined the three-component model as the best classifier for the PLS-DA model, indicated with asterisk (*).

Table 3. Performance of each component used in 10-fold cross validation to determine the best classifier for the PLS-DA model.

Measure	1 component	2 components	3 components
Accuracy	0.3333	1.0000	1.0000
R^2	0.9233	0.9942	0.9997
Q^2	0.7297	0.9833	0.9843

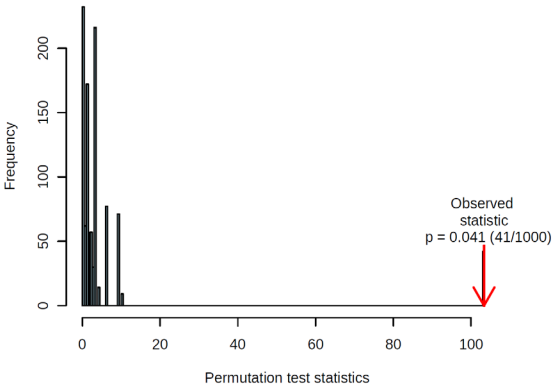


Fig. 6. PLS-DA model validation by permutation tests based on separation distance. The p value based on permutation is p = 0.041 (41/1000).

PLS-DA indicated significant variables in sample classification in terms of VIP scores. According to Cocchi et al. [24], a VIP value greater than 1.0 (i.e., greater than the average of squared VIP values) was commonly used as a threshold to exclude unnecessary variables. **Fig. 7** displays the top 20 important variables determined by the PLS-DA model in descending order of their relative significance. As a result, POL and SSO (from crude palm oil); OOL, SSS, PLL, LLLn and OLL (from lard) and PPL (from soybean oil) were highlighted as the most significant variable of TAGs in the overall sample classification.

Based on alkane number for the classification of lard, crude palm oil and soybean oil, variable nC_{27}, nC_{25} and nC_{13} were selected as the important features, respectively. For clustering analysis, nC_{25} shows high relative importance in lard and low or absent in both vegetable oils.

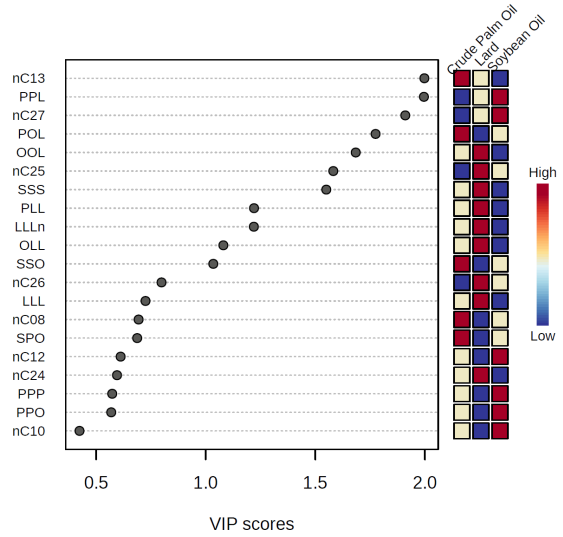


Fig. 7. Variable significance in projection (VIP) score plot for top 20 most significant variables highlighted by PLS-DA. The box represents the relative influence of the variable in the samples.

Identification of potential biomarker

Chemometric methodologies was used to compare the performance and outcome of each expected model to identify relevant indications for lard, crude palm oil and soybean oil. Combining different multivariate data handling approaches could lead to differentiation among them and allow for the extraction of more nuanced information. The model-based approach has a benefit in that it selects key variables by considering the importance calculation based on predictors as well as performance [25].

Table 4 shows the summary of loading scores of PLS-DA (VIP scores) and total abundance of selected significant n-alkanes in the differentiation of lard from crude palm oil and soybean oil. For the main focus of clustering analysis; PCA, HCA Dendogram (with Heat Map) and PLS-DA (VIP scores) highlighted several similar alkanes and TAGs in crude palm oil which were POL, SSO and nC_{13} and, for soybean oil were PPL and nC_{27}. For lard, only OLL and nC_{25} were mainly focused in all PCA, PLS-DA and HCA.

Table 4. Variable significance in projection (VIP) score for variables (score > 1) highlighted by PLS-DA for unique compounds and samples.

No.	Compounds	Vip score	Samples
1.	nC_{13}	1.9980	CPO
2.	PPL	1.9960	SBO
3.	nC_{27}	1.9108	SBO
4.	POL	1.7754	CPO
5.	OOL	1.6848	LARD
6.	nC_{25}	1.5823	LARD
7.	SSS	1.5503	LARD
8.	PLL	1.2206	LARD
9.	LLLN	1.2197	LARD
10.	OLL	1.0806	LARD
11.	SSO	1.0347	CPO

CONCLUSION

GC-MS analysis and multivariate (PCA, HCA, and PLS-DA) techniques were employed to differentiate between lard, crude palm oil and soybean oil based on their TAG profiles and minor components of n-alkanes. Using PCA and HCA, lard was clustered distinctly from other vegetable oils (crude palm oil and soybean oil). PLS-DA models were proven as predictive by providing excellent predictive power of 0.9843. It was discovered that nC_{13} is unique in crude palm oil and n-alkane nC_{25} can also be proposed as a potential biomarker in the clustering analysis of palm oil from the lard. PPL and nC_{27} appeared to be the potential TAG and alkane for the clustering analysis of lard from soybean oil. However, further study is still needed to validate its reliability.

ACKNOWLEDGEMENT

The authors are grateful to the Ministry of Higher Education, Malaysia, for providing the research grant under the Fundamental Research Grant Scheme (KPT Reference Code: FRGS/1/2020/STG01/UPM/02/14; Vote: 5540357; Project Code: 01-01-20-2225FR) along with necessary facilities to carry out the research.

REFERENCES

1. Azir M, Abbasiliasi S, Ibrahim TAT, Manaf YNA, Sazili AQ, Mustafa S. Detection of lard in cocoa butter-its fatty acid composition, triacylglycerol profiles, and thermal characteristics. Foods. 2017; 6(11).
2. Abhishek B, Divya BT. A review study on adulteration in branded packed bakery items. J Forensic Res Crime Stud. 2022;7:1–8.
3. Yardimci M. Impact of Pork Consumption on Human Health. In International European Conference on Interdisciplinary Scientific Research. 2020, July, Turkey, Ankara.
4. Rohman A, Triyana K, Erwanto Y. Differentiation of lard and other animal fats based on triacylglycerols composition and principal component analysis. Int Food Res J. 2012;19(2).
5. Gunstone FD. Production and trade of vegetable oils. In: Gunstone FD, editor. New Jersey: Wiley-Blackwell; 2011. p. 1-24.
6. Azizan NI, Mokhtar NFK, Arshad S, Sharin SN, Mohamad N, Mustafa S, Hashim AM. Detection of lard adulteration in wheat biscuits using Chemometrics-assisted GCMS and Random Forest. Food Anal Methods. 2021;14(11):2276-2287.
7. Troya F, Lerma-García MJ, Herrero-Martínez JM. Simó-Alfonso EF. Classification of vegetable oils according to their botanical origin using n-alkane profiles established by GC-MS. Food Chem. 2015;167:36-39.
8. Official Methods of Analysis of AOAC International, 22nd Edition. New York: Oxford University Press; 2023. Available from: https://doi.org/10.1093/9780197610145.001.0001. Accessed 2024.
9. Sapian NAS, Roslan AMM, Hashim AM, Manaf YNA, Desa MNM, Halim M, Sani MSA, Yusof MT, Pak-Dek MS, Wasoh H. N-alkane profiles of lard and vegetable oils, and their chemometrics differentiation. J Oil Palm Res. 2024;36(3): 524-533.
10. Pacheco C, Palla C, Crapiste GH, Carrín ME. Optimization of reaction conditions in the enzymatic interesterification of soybean oil and fully hydrogenated soybean oil to produce plastic fats. JAOCS. 2013;90(3): 391-400.
11. Manaf YN, Marikkar JMN, Abdulkarim SM. Determination of types of fat ingredient in some commercial biscuit formulations. Int Food Res J. 2014;21(1).
12. Vu T, Riekeberg E, Qiu Y, Powers R. Comparing normalization methods and the impact of noise. Metabolomics. 2018;14(108).
13. Mohd Salleh, NA, Hassan MS, Jumal J, Harun FW, Jaafar MZ. Differentiation of edible fats from selected sources after heating treatments using fourier transform infrared spectroscopy (FTIR) and multivariate analysis. AIP Conf Proc. 2018;1972(1).
14. Sharin SN, Sani MSA, Jaafar MA, Yuswan MH, Kassim NK, Manaf YN, Wasoh H, Zaki NNM, Hashim AM. Discrimination of Malaysian stingless bee honey from different entomological origins based on physicochemical properties and volatile compound profiles using chemometrics and machine learning. Food Chem. 2021;346:128654.
15. Zhang L, Li P, Sun X, Wang X, Xu B, Wang X, Ma F, Zhang Q, Ding X. Classification and adulteration detection of vegetable oils based on fatty acid profiles. J Agric Food Chem. 2014;62(34):8745-8751.
16. Van Ruth SM, Villegas B, Akkermans W, Rozijn M, van der Kamp H, Koota A. Prediction of the identity of fats and oils by their fatty acid, triacylglycerol and volatile compositions using PLS-DA. Food Chem. 2010;118(4):948-955.
17. Lee L, Liong CY. Partial Least Squares-Discriminant Analysis (PLS-DA) for classification of high-dimensional (HD) data: A review of contemporary practice strategies and knowledge gaps. The Analysis. 2018; 143.
18. Chong J, Wishart DS, Xia J. Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. Curr Protoc Bioinformatics. 2019;86(68).
19. Sengupta D, Mukherjee R, Sikdar SK. Moving to a decision point in sustainability analyses. In: Klemeš JJ, editor. Oxford: Butterworth-Heinemann; 2015. p. 87-129.
20. Brodnjak-Vončina D, Kodba ZC, Novič M. Multivariate data analysis in classification of vegetable oils characterized by the content of fatty acids. Chemometr Intell Lab Syst. 2005;75(1):31-43.
21. Cordella CB. PCA: the basic building block of chemometrics. Anal Chem. 2012;47:567.
22. Hidayah EN, Susanto FA, Joko T, Purwestri YA, Nuringtyas TR, Mousa Atoum MF, Jan A. Metabolite profiling of black rice (Oryza sativa L.) following Xanthomonas oryzae pv. oryzae infection. Jordan J Biol Sci. 2021;14(5).
23. Razali MTA, Zainal ZA, Maulidiani M, Shaari K, Zamri Z, Mohd Idrus MZ, Khatib A, Abas F, Ling YS, Rui LL, Ismail IS. Classification of raw stingless bee honeys by bee species origins using the NMR- and LC-MS-based metabolomics approach. Molecules. 2018;23(2160):1-18.
24. Cocchi M, Biancolillo A, Marini F. Chemometric methods for classification and feature selection. Compr Anal Chem. 2018;82:265-295.
25. Chen RC, Dewi C, Huang SW, Caraka RE. Selecting critical features for data classification based on machine learning methods. J Big Data. 2020;7(52).