ASIAN J NAT PROD BIOCHEM
Volume 24, Number 1, June 2026
Pages: 14

E-ISSN: 2580-2550
https://doi.org/10.13057/biofar/f240101

Chemometric-assisted Headspace Solid-Phase Microextraction Gas Chromatography-Mass Spectrometry (HS-SPME-GC-MS) for agarwood quality assessment

THOMAS ZUO CHEN KUEK, LEONG SENG WONG, BENEDICT SAMLING, SIONG FONG SIM♥
Faculty of Resource Science and Technology, Universiti Malaysia Sarawak. 94300 Kota Samarahan, Sarawak, Malaysia.
Tel.: +60-82582995, ♥email: sfsim@unimas.my

Manuscript received: 7 August 2025. Revision accepted: 20 February 2026.

Abstract. *Kuek TZC, Wong LS, Samling B, Sim SF. 2026. Chemometric-assisted Headspace Solid-Phase Microextraction Gas Chromatography-Mass Spectrometry (HS-SPME-GC-MS) for agarwood quality assessment. Asian J Nat Prod Biochem 24 (1): f240101. https://doi.org/10.13057/biofar/f240101.* Agarwood, the rare and fragrant resinous heartwood of *Aquilaria* species, is highly valued for medicinal, cultural, and perfumery uses, yet its quality is typically assessed through subjective sensory evaluation. This study aimed to establish an objective, data-driven approach for agarwood quality differentiation using large-scale chemical profiling combined with chemometric analysis. In this study, 304 agarwood samples from a local collector were analyzed using Headspace Solid-Phase Microextraction Gas Chromatography-Mass Spectrometry (HS-SPME-GC/MS). The resulting large-scale profiling detected 1,036 chemical compounds, providing a database on the chemical characteristics of wild agarwood. Hierarchical clustering, supported by the Calinski-Harabasz Index and silhouette analysis, identified three optimal groups. After filtering for prevalent compounds, Principal Component Analysis (PCA) showed partial overlap among groups, while Partial Least Squares Discriminant Analysis (PLS-DA) achieved an average of 90.31% classification accuracy across 100 training-test splits. Variable Importance in Projection (VIP) scores highlighted key discriminatory compounds, including α-agarofuran, γ-eudesmol, (-)-aristolene, allo-khusiol, β-maaliene, and β-dihydroagarofuran. These findings demonstrate that integrating chemical profiling with multivariate analysis enables objective differentiation of agarwood samples. Although detailed species identity and provenance information were unavailable due to reliance on private collections, the results demonstrate that integrating HS-SPME-GC-MS with multivariate analysis enables reliable, relative classification of agarwood based on chemical composition. The study also establishes a valuable chemical reference for wild agarwood, offering insights relevant to cultivated agarwood production and aiding efforts to optimize induction methods and improve industry practices.

Keywords: Agarwood quality assessment, hierarchical clustering dendrogram, PCA, PLS-DA

INTRODUCTION

Aquilaria species, evergreen trees of the Thymelaeaceae family, are found mainly in Malaysia, Southern China, Vietnam, Indonesia, Cambodia, Brunei, India, Laos, and Thailand (Adam et al. 2017; Wang et al. 2019). These trees produce agarwood as a defence mechanism against damage caused by lightning strikes, animals, and insect attacks (Kalra and Kaushik 2017). When the wounded tissue becomes infected by fungi, the tree secretes resinous materials collectively known as agarwood (Naziz et al. 2019; Ma et al. 2021).

Agarwood has long been valued in traditional medicines and for religious practices, and its distinctive aroma makes it highly sought after in the perfumery and incense industries (Tan et al. 2019). However, over-exploitation and excessive logging have drastically reduced wild *Aquilaria* populations. As a result, agarwood-producing species are listed under Appendix II of the Convention on International Trade in Endangered Species, restricting their global trade (Chhipa et al. 2017).

Agarwood quality is traditionally judged by color, density, and burning scent, with darker pieces indicating higher resin content (Ismail et al. 2017; Villareal et al. 2022). Prices vary widely, from USD 10,000-20,000 per kg for wild agarwood to USD 100-300 per kg for cultivated material (Abdulah et al. 2022), making it one of the most expensive natural products.

Despite its high value, agarwood grading remains largely subjective, relying on expert evaluation (Wang et al. 2021). To reduce this subjectivity, analytical techniques such as Gas Chromatography Mass Spectrometry (GCMS), High Performance Liquid Chromatography (HPLC), and High-Performance Thin Layer Chromatography (HPTLC) are increasingly applied, with GCMS being the most common (Sen et al. 2017; Kao et al. 2018). Within GCMS methods, headspace solid-phase microextraction (HS-SPME-GCMS) is especially effective for limited agarwood samples, producing rich chemical data while minimizing sample preparation time (Hung et al. 2014). Over the past 18 years, 64 ScienceDirect-indexed studies have applied GCMS to agarwood, highlighting growing scientific interest (Figure 1).

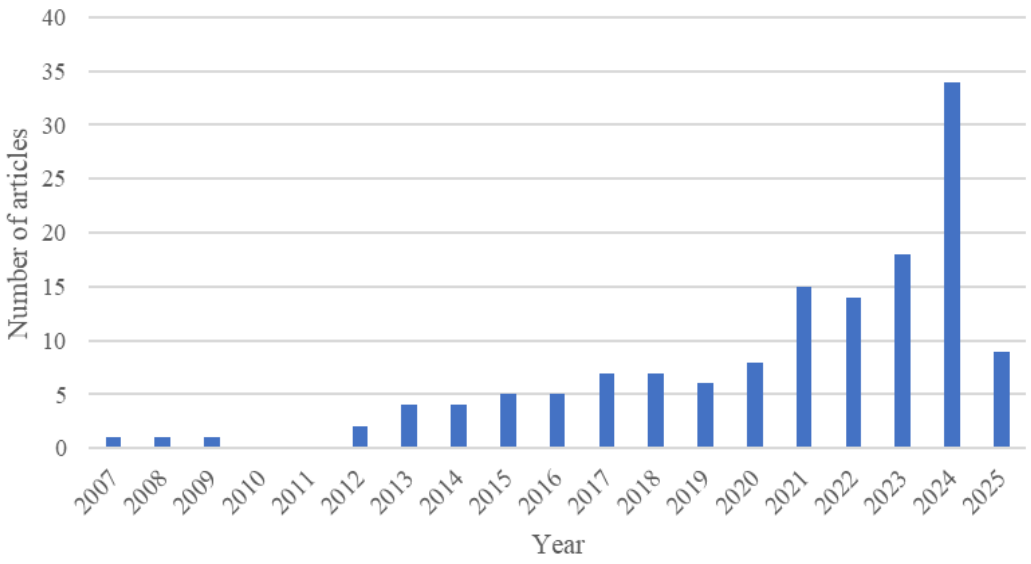

Figure 1. Number of papers published under ScienceDirect discusses the application of GC-MS in agarwood analysis

Agarwood is both rare and costly, rendering large-scale scientific studies impractical. For example, essential oil extraction requires 200-500 g of material. Since agarwood quality is traditionally assessed by its smoke, which can be scientifically interpreted as the release of volatile organic compounds, SPME provides an ideal alternative for trapping these compounds for GCMS analysis. This technique enables the analysis of large sample sets using minimal material and without the need for solvents.

Scientific evaluation increasingly integrates large-scale chemical profiling with multivariate statistics. Techniques such as hierarchical clustering and PCA have been widely applied to interpret agarwood's complex matrix (Tufariello et al. 2022), and Hidayat et al. (2010) showed that electronic nose data can be classified effectively using these methods. PLS-DA has also been used to identify indicator components and marker metabolites (Qian et al. 2025). However, many GC-MS studies remain constrained by small sample sizes.

This study addresses this gap by analyzing the volatile profiles of more than 300 agarwood samples using HS-SPME-GC-MS, generating one of the most extensive datasets to date. The resulting chemical database provides a comprehensive reference for understanding the compositional diversity of wild agarwood and offers valuable insights to support the cultivated agarwood industry in refining plantation practices and induction methods. By integrating large-scale HS-SPME-GC-MS data with multivariate analyses, this work establishes a robust foundation for future research, comparison, and standardization efforts. Accordingly, this study hypothesizes that HS-SPME-GC-MS-based chemical profiling integrated with chemometric analysis can objectively differentiate agarwood samples into distinct quality groups based on their volatile composition.

MATERIALS AND METHODS

Samples and sample preparation

A total of 304 agarwood samples, sourced from a local collector in Sarawak, were obtained for analysis. The collector confirmed that the samples are wild agarwood originating from Borneo. The agarwood samples were part of the personal collection of the collector, displayed in his art gallery. A research permit application has been submitted, and approval is recommended (SORAS Application ID 2025109). There is a lack of detailed information on the precise origin of the samples, and their current state does not allow for species-level identification. The absence of verified species identity and provenance, therefore, limits this study. Despite this limitation, the study focuses on chemical profiling and pattern recognition of agarwood compounds, rather than ecological or species-level inference. The consistent sourcing from a single collector offers the advantage of reduced variability associated with differences in storage handling and collection methods, providing a uniform set of samples for chemical analysis. A piece from each agarwood sample was obtained, stored in a ziplock bag, and labelled with a unique code. The samples were transported to the laboratory and cut into small pieces (2-5 mm) for further examination. An example of an agarwood sample, including the collected and processed sample, is shown in Figure 2.

Figure 2. An example of an agarwood sample - the collected and processed sample

Headspace Solid-Phase Microextraction (HS-SPME)

HS-SPME extraction was performed using an SPME fiber with a 50/30 µm divinylbenzene/carboxen/polydimethylsiloxane (DVB/CAR/PDMS) coating, followed by an SPME manual holder. The SPME fiber and manual holder were purchased from Supelco. The fiber was preconditioned in the GC injection port according to the manufacturer's recommendations before initial use. The agarwood sample was weighed to 50 mg and placed in a 2 mL clear vial, which was sealed with a PTFE/silicone septum screw cap. The sealed vial was pre-heated and incubated at 80°C in an oven for 40 minutes. After 40 minutes, the SPME needle was inserted through the septum, exposing the fiber to the sample headspace for 10 minutes at 80°C. The fiber was then injected into the GC-MS and desorbed for 2 minutes. After each analysis, the SPME fiber underwent a blank run in the GC-MS to check for any carryover.

Gas Chromatography-Mass Spectrometry (GC-MS)

The volatile organic profile was analyzed using a GC-MS system (Agilent 7890B) equipped with a mass-selective detector (Agilent 5977B). Chromatographic separation was performed using an HP-5ms Ultra Inert Column (30 m × 0.25 mm inner diameter, 0.25 µm film thickness). The injection port was set to 250°C for sample analysis and blank runs in splitless mode. Pure helium (99.99%) was used as the carrier gas at a flow rate of 1.0 mL/min. Operating conditions included an initial temperature of 40°C with a 1-minute equilibration time, followed by a ramp to 280°C at a rate of 6°C/min, held for 5 minutes. The interface temperature was maintained at 280°C. Mass spectra were acquired in scan mode by electron ionization at 70 eV within the m/z 33-500 amu range, with the ion source temperature set to 230°C. The temperatures of the sample and transfer lines were set to 150°C and 160°C, respectively. Data acquisition and processing were carried out using Agilent Mass Hunter B.08.00. Chemical compounds in the agarwood samples were identified by matching their mass spectra with the NIST-17 mass spectral library. The data files, containing the names of compounds detected along with their corresponding retention times, matching factors, CAS numbers, and peak areas, were saved in CSV format. Only peaks with a mass spectral similarity greater than 75% to the NIST library were reported. All samples were analyzed using the same instrument, column, and operating parameters to minimize run-to-run variability. To further reduce batch effects, the chromatograms were baseline corrected.

Chromatogram analysis

An in-house algorithm developed in MATLAB R2022a was employed to ensure reproducible processing of a large number of chromatograms. The algorithm compiled all detected compounds into a master list containing the sample ID, retention time, CAS number, and peak area. Compound alignment across samples was performed within a predefined retention time window (*ws*=0.5 min) based on CAS number matching. To verify reproducibility, the time window was varied from 0.5 to 2.0 min, during which the size of the resulting peak table remained unchanged - confirming consistent peak alignment and reliable data processing.

Hierarchical clustering dendrogram

The resultant peak table was normalized with a z-score to standardize the data across variables. The Calinski-Harabasz Index (CHI) value was used to determine the optimal number of clusters. The silhouette analysis was performed using squared Euclidean distance as the dissimilarity metric for further evaluation of cluster quality. Hierarchical clustering was then performed using Ward's method, a variance-minimizing agglomerative method which iteratively grouped samples in a manner that reduced within-cluster heterogeneity (Ward 1963). Sample cluster assignments were determined based on hierarchical relationships among the samples, which were then visualized through a hierarchical clustering dendrogram (Charikar et al. 2019; Fei et al. 2021; Randriamihamison et al. 2021). Note that the clustering was derived solely from unsupervised chemometric analysis. No independent physical, sensory, or commercial quality references were available to calibrate these clusters.

Principal Component Analysis (PCA)

The peak table obtained was square-rooted, scaled to one, standardized, and subjected to PCA. To reduce noise and remove insignificant compounds, an algorithm was implemented in which a loop iteratively excluded compounds occurring between 10 and 220 samples at intervals of 10. Silhouette analysis was then performed to assess clustering quality, and the peak table yielding the best clustering result was selected for variable selection and classification.

Partial Least Squares Discriminant Analysis (PLS-DA)

The refined peak table was normalized using z-score normalization. A PLS-DA model was built by fitting the standardized dataset using two latent components, which

were selected to capture the maximum covariance between the chemical compounds and the predefined cluster assignments. The dataset was partitioned into training and test sets for classification using PLS-DA, with 70% of the samples allocated for training and the remaining 30% for testing. To prevent information leakage, standardization of the test set employed the mean and standard deviation computed solely from the training set. Model performance was evaluated over 100 iterations, each involving a new training-test split. Classification accuracy was calculated as the percentage of correctly classified test samples, defined as the ratio of correctly predicted samples to the total number of test samples. The mean classification accuracy across the 100 iterations was then reported. Variable Importance in Projection (VIP) scores were calculated to evaluate the contribution of each chemical compound to the PLS-DA model. The compounds were selected based on VIP scores greater than 1 and visualized through a bar plot. The analysis was conducted using MATLAB R2022a. The process of analysis is illustrated in the flow chart in Figure 3.

RESULTS AND DISCUSSION

Chemical profile and chemical compositions

A peak table with dimensions 304×1036 was generated from the chromatogram analysis. Among the samples, 16 contained more than 70 compounds, 109 contained 60-70 compounds, 129 contained 50-60 compounds, and the remaining 50 contained fewer than 50. A heat map (Figure 4) was generated to illustrate variability in compound profiles, with red indicating higher concentration percentages and blue indicating lower ones. Such chemical variation is influenced by species differences, geographic origin, microbial infection, and environmental conditions (Chen and Rao 2022; Zhang et al. 2022).

The complexity of agarwood's chemical composition arises from its broad spectrum of secondary metabolites, which may also vary in stereochemistry, adding further diversity (Shao et al. 2016; Gutiérrez et al. 2024). The peak table was evaluated, and the major compound groups with their characteristic aromas are summarized in Table 1. Detailed information on the abundance and frequency of the compounds within each group is provided in Table 4. The four most abundant compounds, allo-khusiol, guaiol, γ-eudesmol, and α-agarofuran, are illustrated with representative chromatograms in Figure 5. Most compounds eluted after 17 minutes and were primarily sesquiterpenes, consistent with previous reports using HS-SPME from agarwood chips, incense smoke, and essential oils (Ahmaed et al. 2017; Ahmaed et al. 2022; Sundaraj et al. 2023).

Quality indicator assessment - Frequently detected compounds

The peak table was analyzed to identify prominent chemical features of agarwood as potential quality indicators. Agarwood quality is largely determined by resin content, with higher resin content generally commanding higher prices. Aroma profiles also contribute to value, as distinctive aromas are often associated with premium pricing. However, no standardized approach currently exists for assessing agarwood quality based on chemical composition. We postulate that compounds present consistently in agarwood samples may serve as typical chemical characteristics. Table 2 lists compounds detected in over 50% of samples. Aristolene was the most frequent compound, detected in 258 samples (84.87%), followed by γ-eudesmol (76.64%), β-elemene (76.32%), β-dihydroagarofuran (72.37%), and 2-butanone, 4-phenyl (72.37%). Some consistently detected compounds occurred at lower concentrations, including propanoic acid, 2-methyl-, 3-hydroxy-2,2,4-trimethylpentyl ester (71.05%), β-guaiene (65.46%), and eremophila-1,11-dien-9-one (59.87%).

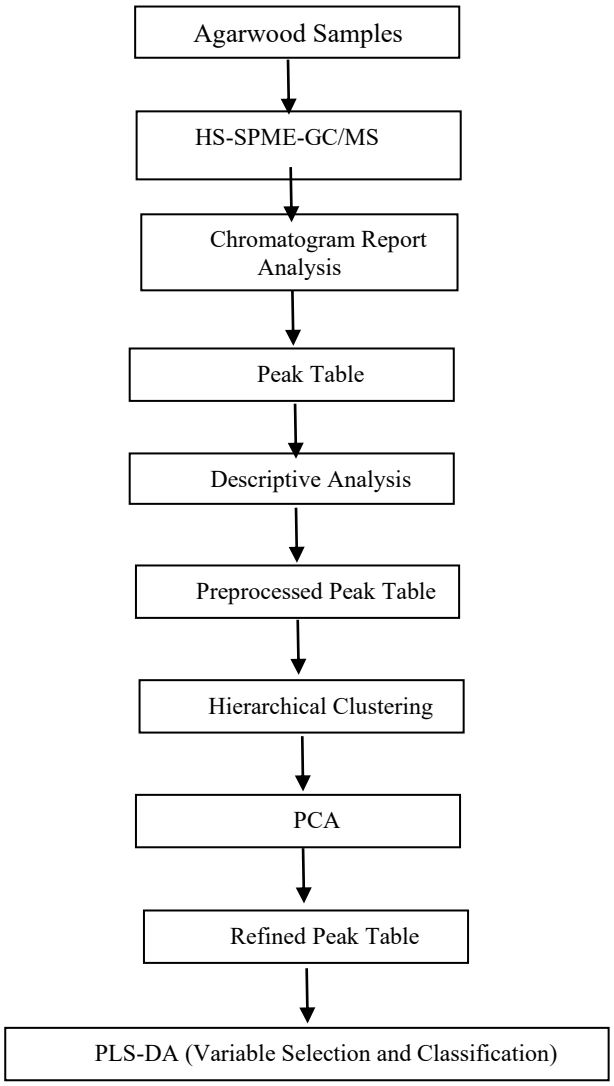

Figure 3. Flow chart of the data analysis process for the agarwood dataset

Table 1. Major groups of compounds present and their characteristics aroma

Category	Aroma profile
Alcohols and Esters	Sweet and fruity with subtle green, creamy, and floral undertones.
Alkanes and Alkenes	Neutral and waxy, with subtle green, herbal, and oily notes, with faint earthy, marine, and citrus fragrance
Ketones and Chromones	Warm and woody with balsamic sweetness, complemented by mild floral, spicy, resinous, and slightly smoky undertones.
Sesquiterpenes	Woody and herbal, with complementary notes of spicy, earthy, balsamic, floral, and subtle smoky or bitter undertones
Sesquiterpene Alcohols	Woody and warm, with complementary floral, balsamic, earthy, and creamy notes, along with subtle musky, green, camphoraceous, and minty undertones.
Sesquiterpenoids	Woody and earthy, with subtle balsamic, resinous, peppery, and warm spicy notes.
Aromatics and Phenolics	Phenolic, smoky, medicinal
Phenylpropanoids	Spicy, slightly sweet with floral citrusy nuances

Table 2. The volatile organic compounds frequently detected in agarwood samples

Compound name	Frequency	% total of sample	Mean % peak area	Range
Aristolene	258	84.87	5.33±3.08	0.08-15.40
γ-Eudesmol	233	76.64	13.30±9.77	0.09-38.74
2-Naphthalenemethanol, 1,2,3,4,4a,5,6,7-octahydro-.alpha.,.alpha.,4a,8-tetramethyl-, (2R-cis)-				
β-Elemene	232	76.32	0.79±1.72	0.02-23.55
Cyclohexane, 1-ethenyl-1-methyl-2,4-bis(1-methylethenyl)-, [1S-(1.alpha.,2.beta.,4.beta.)]-				
β-Dihydroagarofuran	221	72.70	2.52±1.87	0.08-13.70
2H-3,9a-Methano-1-benzoxepin, octahydro-2,2,5a,9-tetramethyl-, [3R-(3.alpha.,5a.alpha.,9.alpha.,9a.alpha.)]-				
2-Butanone, 4-phenyl-	220	72.37	1.63±1.79	0.004-12.72
Guaiol	219	72.04	5.23±5.15	0.13-41.32
α-Guaiene	217	71.38	0.85±1.39	0.02-11.08
4,6,6-Trimethyl-2-(3-methylbuta-1,3-dienyl)-3-oxatricyclo[5.1.0.0(2,4)]octane	216	71.05	6.17±5.72	0.06-41.32
Propanoic acid, 2-methyl-, 3-hydroxy-2,2,4-trimethylpentyl ester	216	71.05	3.50±3.37	0.11-22.37
cis-Thujopsene	206	67.76	0.18±0.17	0.01-1.30
β-Costol	202	66.45	1.05±0.89	0.01-6.63
2-((2R,4aR,8aS)-4a-Methyl-8-methylenedecahydronaphthalen-2-yl)prop-2-en-1-ol				
β-Guaiene	199	65.46	2.19±1.87	0.04-18.04
Humulene	198	65.13	1.65±2.22	0.04-16.15
cis-Z-α-Bisabolene epoxide	191	62.83	1.72±1.51	0.01-10.38
Eremophila-1,11-dien-9-one	182	59.87	3.87±3.10	0.06-17.21
(3S,4aR,5S,8aS)-4a,5-Dimethyl-3-(prop-1-en-2-yl)-2,3,4,4a,5,6-hexahydronaphthalen-1(8aH)-one				
Caryophyllene	181	59.54	1.18±1.32	0.01-11.53
Cedrol	179	58.88	0.91±0.67	0.02-3.71
Viridiflorol	175	57.57	6.58±3.85	0.02-15.36
1H-Cycloprop[e]azulen-4-ol, decahydro-1,1,4,7-tetramethyl-, [1aR-(1a.alpha.,4.beta.,4a.beta.,7.alpha.,7a.beta.,7b.alpha.)]-				
(8R,8aS)-8,8a-Dimethyl-2-(propan-2-ylidene)-1,2,3,7,8,8a-hexahydronaphthalene	175	57.57	0.95±0.77	0.001-4.43
1-Hexanol, 2-ethyl-	173	56.91	1.68±1.94	0.03-13.54

Quality indicator assessment - High-concentration compounds

High-quality indicators can be postulated to be agarwood-containing compounds at significant concentrations, which are likely to produce highly distinctive aroma profiles. Table 3 summarizes compounds present in agarwood at relatively high abundances compared to other compounds. Allo-khusiol (prezizaan-7-ol) was found at an average concentration of 24.79±9.40% despite being detected in only 44 out of the 304 samples. α-Agarofuran and agarospirol were two other prominent compounds accounting for 10.46±6.24% and 7.43±3.37% of the total peak area, respectively. These compounds have also been regarded as potential quality indicators in literature. As presented in Table 4, the identified compounds are associated with distinct aroma characteristics.

Table 3. The volatile organic compounds present in agarwood at high concentrations but low frequency

Compound name	Frequency	Mean % peak area	Range
allo-khusiol or prezizaan-7-ol	44	24.79±9.40	6.32-48.75
(3S,3aS,6R,7S,8aS)-3,7,8,8-Tetramethyloctahydro-1H-3a,6-methanoazulen-7-ol			
Cyclodecane	8	15.13±2.67	0.66-22.42
α-Agarofuran	133	10.46±6.24	0.45-23.85
(3R,5aS,9aR)-2,2,5a,9-Tetramethyl-3,4,5,5a,6,7-hexahydro-2H-3,9a-methanobenzo[b]oxepine			
5-epi-7-epi-α-Eudesmol	1	10.38±0.60	-
2-(4a,8-Dimethyl-2,3,4,5,6,8a-hexahydro-1H-naphthalen-2-yl)propan-2-ol			
Cyclododecane	14	8.95±2.08	2.90-21.20
Maaliol	1	8.40±0.48	-
(1aR,3aS,7S,7aS,7bR)-1,1,3a,7-Tetramethyldecahydro-1H-cyclopropa[a]naphthalen-7-ol			
Hexadecane, 1,16-dichloro-	7	8.29±1.32	3.41-11.96
1-Dodecanol	8	8.01±2.57	0.05-42.95
Hedycaryol	4	7.89±0.90	7.67-8.08
3,7-Cyclodecadiene-1-methanol, .alpha.,.alpha.,4,8-tetramethyl-, [s-(Z,Z)]			
6-Tridecene, (Z)-	3	7.87±1.24	0.71-21.60
Agarospirol	46	7.43±3.37	0.68-25.61
Cyclododeca[b]furan-3-carbonitrile, 2-amino-4,5,6,7,8,9,10,11,12,13-decahydro-	13	6.86±1.49	3.02-12.45
α-Costol	73	6.49±3.25	0.002-12.80
2-((2R,4aR,8aR)-4a,8-Dimethyl-1,2,3,4,4a,5,6,8a-octahydronaphthalen-2-yl)prop-2-en-1-ol			
Prezizaene	69	6.25±3.64	0.10-28.95
(3S,3aS,6R,8aS)-3,8,8-Trimethyl-7-methyleneoctahydro-1H-3a,6-methanoazulene			
Taylorione	17	5.64±1.61	0.17-12.88
(-)-Globulol	17	5.50±1.47	0.66-12.27
Epiglobulol	34	5.50±3.83	0.08-40.56
(1aR,4S,4aR,7R,7aS,7bS)-1,1,4,7-Tetramethyldecahydro-1H-cyclopropa[e]azulen-4-ol			

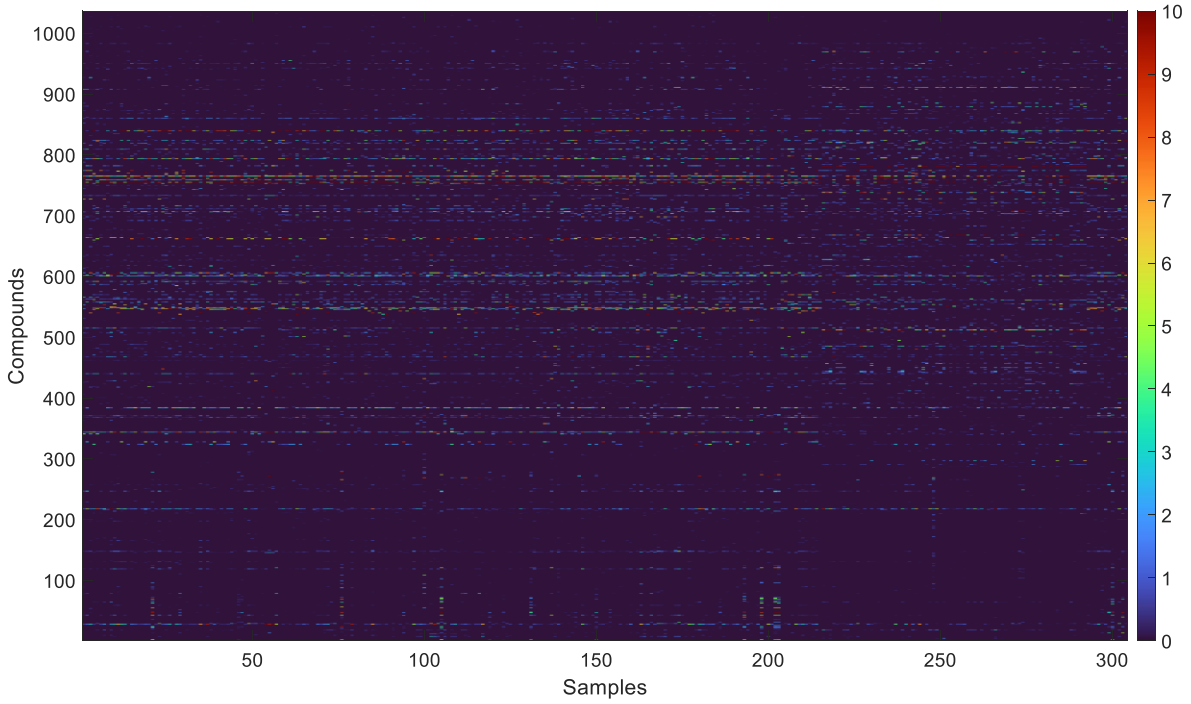

Figure 4. Variability in chemical composition among the samples was visualized using a heat map generated from the peak area matrix

Table 4. Major compounds present in significant abundance and their frequency of detection across 304 agarwood samples, categorized by functional groups

Category	Compounds	RT (min)	Aroma profile	Frequency	Average abundance (%)
Alcohols and Esters	Propanoic acid, 2-methyl-, 3-hydroxy-2,2,4-trimethylpentyl ester	17.00	Sweet, fruity, slightly creamy	216	3.50±3.37
	1-Hexanol, 2-ethyl-	8.90	Green, fresh, slightly fruity	173	1.68±1.94
	1-Dodecanol	13.97	Waxy, slightly floral, fatty	8	8.01±2.57
Alkanes and Alkenes	Tetradecane	17.49	Neutral, waxy, slightly marine-like	145	0.77±1.09
	(4S,4aR,6R)-4,4a-Dimethyl-6-(prop-1-en-2-yl)-1,2,3,4,4a,5,6,7-octahydronaphthalene	19.07	Green, herbal with a hint of woody freshness	140	3.11±2.66
	Decane, 2,6,7-trimethyl-	9.68	Oily, slightly musty	31	4.34±2.23
	Octane, 2,6-dimethyl-	9.44	Neutral, light waxy	12	2.22±0.73
	Tridecane	15.30	Mild, waxy, slightly earthy	15	2.55±1.02
	Cyclododecane	19.06	Waxy, neutral	14	8.95±2.08
	Hexadecane, 1,16-dichloro-	19.05	Oily, faintly chemical	7	8.29±1.32
	1-Undecene, 7-methyl-	15.62	Light, waxy, faintly citrus	6	2.94±0.60
Ketones and Chromones	2-Butanone, 4-phenyl-	14.13	Sweet, balsamic, mild floral	220	1.63±1.79
	Eremophila-1,11-dien-9-one	23.18	Spicy, warm, and slightly earthy, resinous	182	3.87±3.1
	α-Costal	23.38	Woody, balsamic, slightly earthy, resinous	122	2.84±3.18
	Valerenal	22.85	Warm, woody with a mild balsamic sweetness	76	3.99±3.19
	Taylorione	22.81	Deep woody, slightly smoky	17	5.64±1.61
Sesquiterpenes	Aristolenes	22.53	Subtle peppery or herbal nuances	258	5.33±3.08
	β-Elemene	17.41	Floral aroma	232	0.79±1.72
	α-Guaiene	18.40	Earthy, woody, and spicy	219	0.85±1.39
	β-Guaiene	17.62	Dry, woody, balsamic, spicy	199	2.19±1.87
	Humulene	18.76	Woody, herbal, slightly bitter	198	1.65±2.22
	Caryophyllene	18.04	spicy, woody, and slightly sweet	181	1.18±1.32
	Prezizaene	18.74	Deep herbal, woody, slightly smoky	69	6.25±3.64
	α-Bulnesene	19.80	Warm, earthy, herbal	159	2.33±2.71
Sesquiterpene Alcohols	Agarospirol	22.43	Warm, floral, balsamic, slightly musky	46	7.43±3.37
	cis-α-Santalol	21.30	Creamy, soft sandalwood, slightly powdery	57	1.70±1.11
	α-Santalol	21.31	Warm, deep sandalwood, velvety	14	3.24±1.33
	Guaiol	22.40	Sweet, woody, floral	219	5.23±5.15
	(-)-Spathulenol	19.17	Warm, woody, earthy	30	2.96±1.37
	α-Costol	20.65	Earthy, woody, and slightly sweet	73	6.49±3.25
	γ-Eudesmol	22.15	Woody, slightly sweet with earthy and balsamic undertones	233	13.30±9.77
	Allo-khusiol	22.98	Woody with mild green and herbal nuances	44	24.79±9.40
	Viridiflorol	19.08	Woody, green, slightly camphoraceous	175	6.58±3.85
	Epiglobulol	18.55	Earthy, camphoraceous, slightly minty	34	5.50±3.83
Sesquiterpenoids	β-Dihydroagarofuran	19.76	woody, earthy, and slightly balsamic	221	2.52±1.87
	4,6,6-Trimethyl-2-(3-methylbuta-1,3-dienyl)-3-oxatricyclo[5.1.0.0(2,4)]octane	23.73	-	216	6.17±5.72
	Isoaromadendrene epoxide	21.31	Woody, slightly peppery, resinous	111	2.93±2.47
	α-Agarofuran	20.71	Deeply woody with warm, slightly spicy, and earthy facets.	133	10.46±6.24
Aromatics and Phenolics	Phenol, 3,5-bis(1,1-dimethylethyl)-	19.84	Phenolic, smoky, medicinal	31	3.47±1.41
Phenylpropanoids	Elemicin	20.80	Spicy, slightly sweet with floral citrusy nuances	89	4.35±7.31

Hierarchical clustering dendrogram

Limited prior information was available on the samples, so clustering was performed and evaluated using the Calinski-Harabasz Index (CHI), which measures the ratio of between-cluster to within-cluster dispersion (Karna and Gibert 2022). The CHI identified three as the optimal number of clusters by maximizing separation and compactness (Rachwał et al. 2023). CHI values for 1-10 clusters are shown in Figure 6.

In addition to the Calinski-Harabasz Index, clustering quality was evaluated using the silhouette method, which measures how well a sample fits within its cluster compared to others (Ogbuabor and Ugwoke 2018; Januzaj et al. 2023). Silhouette scores range from -1 to 1, with higher positive values indicating better clustering. Scores were calculated for 2-5 clusters (Figure 7). The two-cluster solution showed mostly positive values but suggested possible subgroups. The three-cluster solution exhibited consistently positive scores with minimal overlap, while four- and five-cluster solutions showed negative values in some clusters, indicating misclassification or over-segmentation. Combining CHI and silhouette results, three clusters were determined as optimal. The samples were classified into three groups based solely on unsupervised analysis and were not benchmarked against established market grades. Figure 8 presents the hierarchical clustering dendrogram of the agarwood samples, illustrating their separation into three distinct clusters. The proposed "quality grading system" therefore represents a relative, data-driven classification based on chemical composition, rather than a formal commercial or market-based grading scheme. Future studies may incorporate validated quality indicators - such as sensory evaluation, resin content analysis, and market price data - to correlate the chemometric clusters with recognized quality standards and thereby strengthen the practical applicability of the proposed grading system.

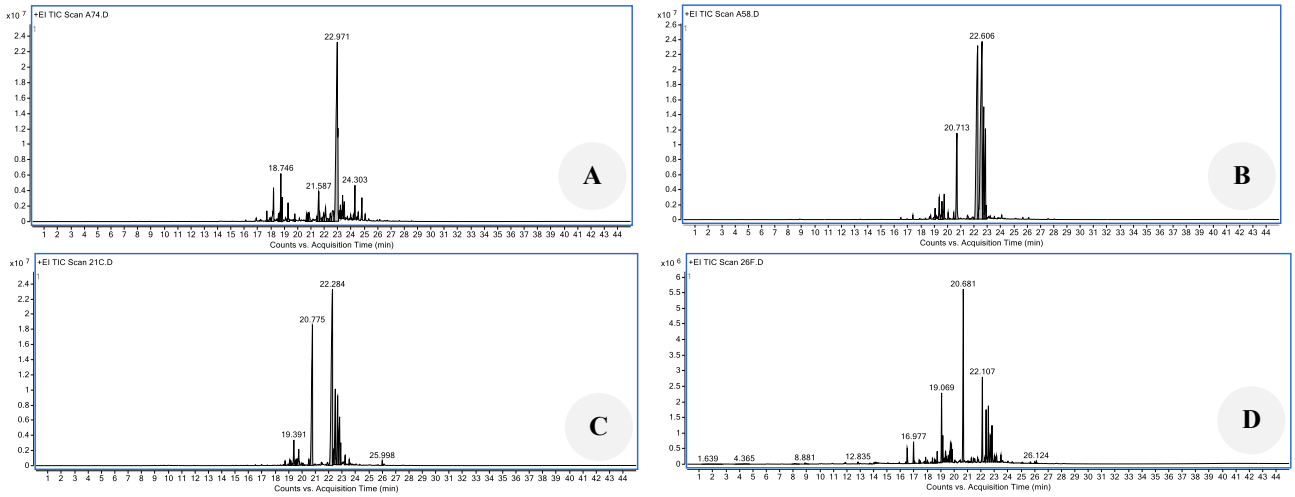

Figure 5. Example chromatograms predominated by four major compounds. A. Allo-khusiol (22.971 min), B. Guaiol (22.606 min), C. γ-eudesmol (22.284 min), and D. α-agarofuran (20.681 min), which were detected at the highest concentrations in selected agarwood samples.

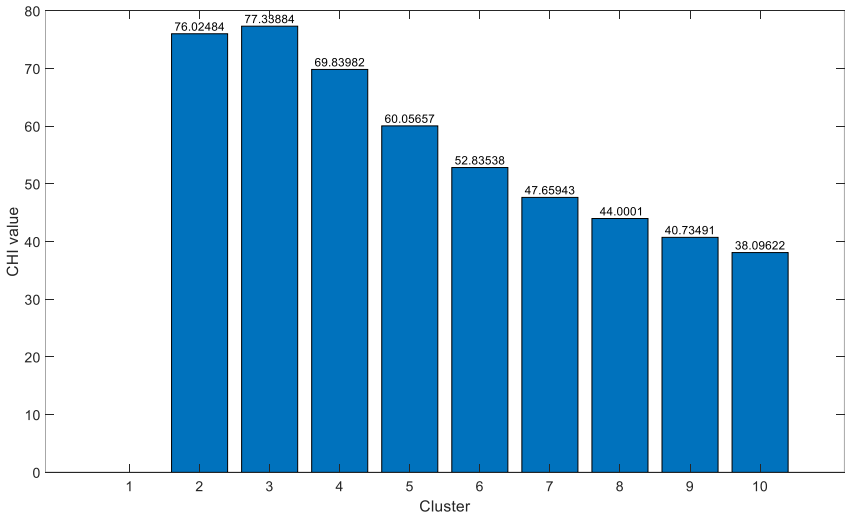

Figure 6. The bar chart of CHI values from 1 to 10 clusters

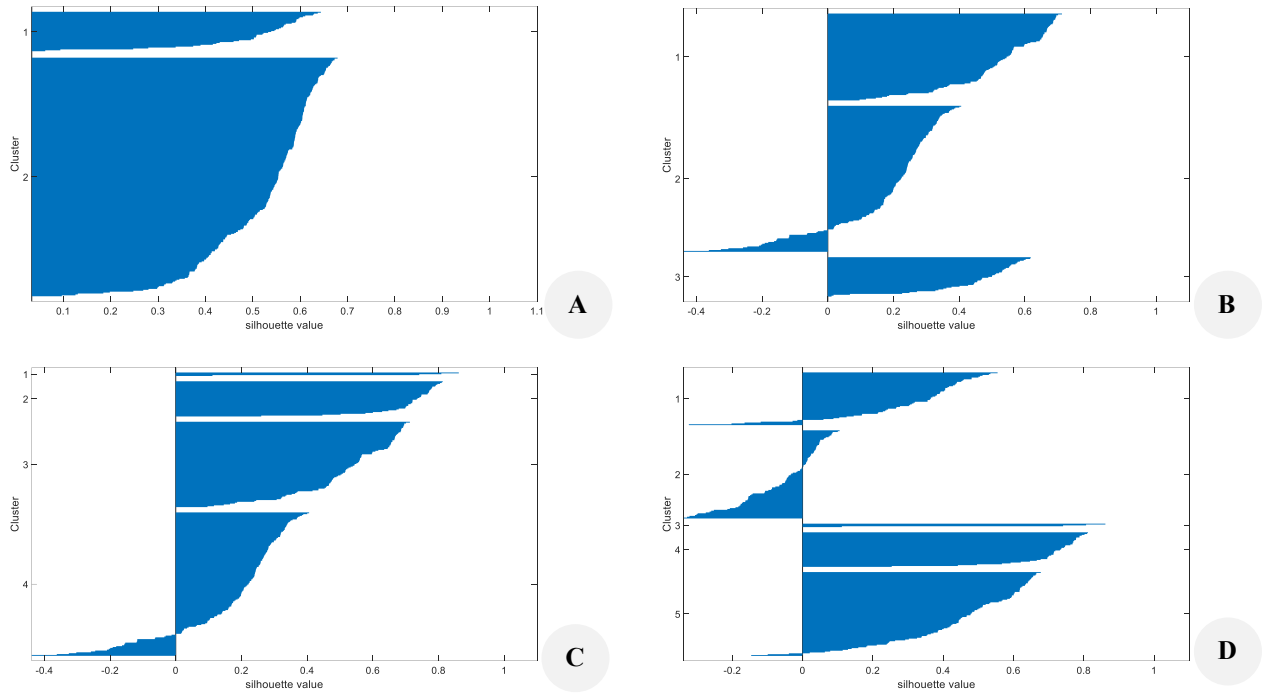

Figure 7. Silhouette plots of clustering models ranging from: A. 2 cluster, B. 3 cluster, C. 4 cluster, D. 5 cluster

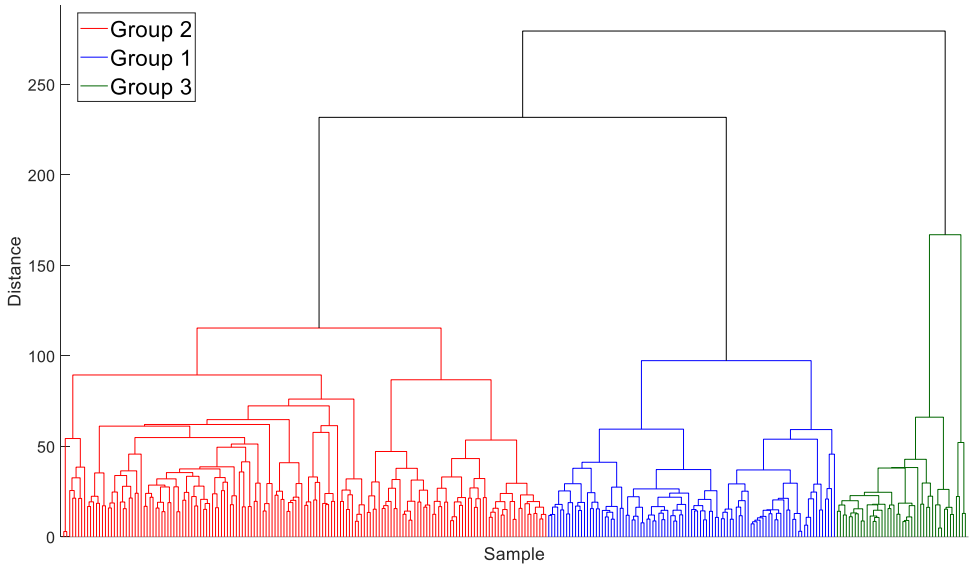

Figure 8. Hierarchical clustering dendrogram of agarwood samples based on the peak area table (304×1,036)

Principal Component Analysis (PCA)

The original peak table (304×1036) was subjected to PCA, and the scores plot (Figure 9.A) shows samples grouped according to their assigned classes (Jolliffe and Cadima 2016; Ma 2024). Groups 1 and 2 largely overlap, while Group 3 is partially separated along PC1. This limited distinction likely arises from noise and redundancy introduced by numerous low-frequency compounds, which may not consistently reflect the core chemical profile of agarwood. To address this, compounds detected in fewer than 40 samples were excluded to minimize noise while retaining features critical for group separation. This reduced the peak table to 304×110. PCA of this filtered dataset (Figure 9.B) revealed three distinct clusters: Group 1 formed a compact cluster indicative of strong internal consistency (Lever et al. 2017), Group 2 displayed moderate variability, and Group 3 was clearly separated along negative PC1. Although PC1 and PC2 account for

only 8.39% and 6.93% of the variance, they capture the most discriminating patterns among groups. This observation aligns with Hidayat et al. (2010), who demonstrated reliable differentiation of agarwood oil groups based on electronic nose profiles. Low explained variance is expected for high-dimensional GC-MS datasets, as the large number of correlated variables distributes the total variance across multiple components.

Partial Least Squares Discriminant Analysis (PLS-DA)

PLS-DA is a supervised method that maximizes separation between predefined groups using chemical profiles (Peiró-Vila et al. 2024). Unlike unsupervised PCA, PLS-DA leverages class labels to enhance discrimination (Bi et al. 2021), projecting the original data onto latent components (Lasalvia et al. 2022). The PLS-DA scores plot (Figure 10) confirmed the clustering observed in PCA, with Groups 1 and 3 being chemically distinct and Group 2 positioned as intermediate.

The loadings plot identifies compounds driving class separation. Group 1 is associated with aristolene, γ-eudesmol, α-agarofuran, β-dihydroagarofuran, β-maaliene, guaiol, and β-cadinene. Group 2 is linked to propanoic acid 2-methyl-3-hydroxy-2,2,4-trimethylpentyl ester, cedrol, cis-thujopsene, viridiflorol, eremophila-1,11-dien-9-one, 1,3-pentanediol 2,2,4-trimethyl-, 2,4-di-tert-butylphenol, and 1-hexanol 2-ethyl-. Group 3 corresponds to isolongifolol, allo-khusiol, prezizaene, alloaromadendrene oxide-(1), 1-oxaspiro[2.5]octane 5,5-dimethyl-4-(3-methyl-1,3-butadienyl)-, α-curcumene, β-patchoulene, squamulosone, α-cedrene, and aristolediene.

VIP scores were calculated to identify the most significant discriminant compounds (Shi et al. 2023), with scores >1 shown in Figure 11. Among them, α-agarofuran and γ-eudesmol emerged as the most influential markers having been reported to exhibit antioxidant, antimicrobial, and insecticidal properties (Ismail et al. 2013; Monggoot et al. 2017; Xu et al. 2025). Zhang et al. (2024) suggested that agarofuran and agarospirol may serve as significant markers for assessing agarwood quality. Likewise, a self-organizing map - based grading model identified β-agarofuran, α-agarofuran, and 10-epi-γ-eudesmol as classification markers (Haron 2020). Other sesquiterpenes such as (-)-aristolene , β-maaliene, allo-khusiol, and isolongifolol, which demonstrated significant differentiating ability, have also been known to contribute bioactivities ranging from vasodilation to neuroprotection (Choudhary et al. 2005; Tian et al. 2021; Ahmaed et al. 2022; Xue et al. 2023). Other high-VIP compounds are widely recognized for their biological activities, which are summarized in Table 5. Collectively, these compounds contribute to the characteristic aromas used in traditional sensory grading while demonstrating pharmacological significance, highlighting a strong link between chemical markers and established perceptions of agarwood quality. The functional roles discussed are derived exclusively from published literature; no bioassays were performed in the present study.

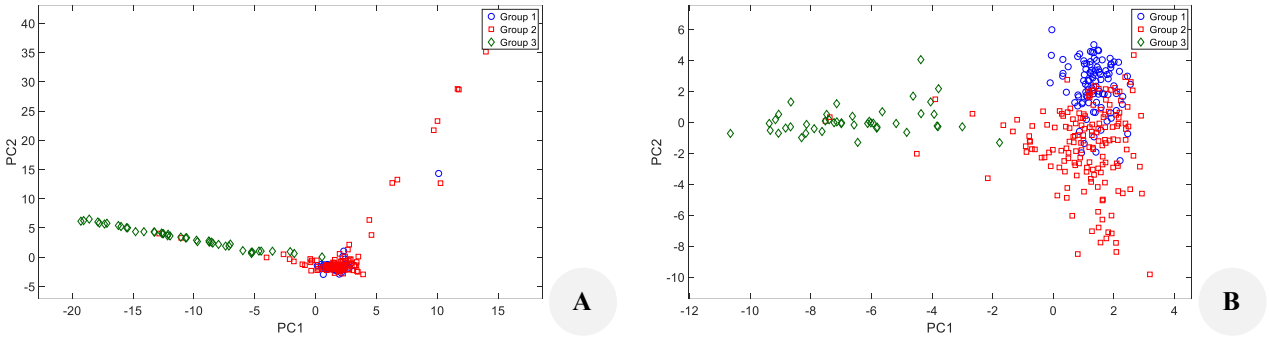

Figure 9. PCA scores plots based on: A. The full peak table (304×1,036) and B. The filtered table including compounds detected in >40 samples (304×110)

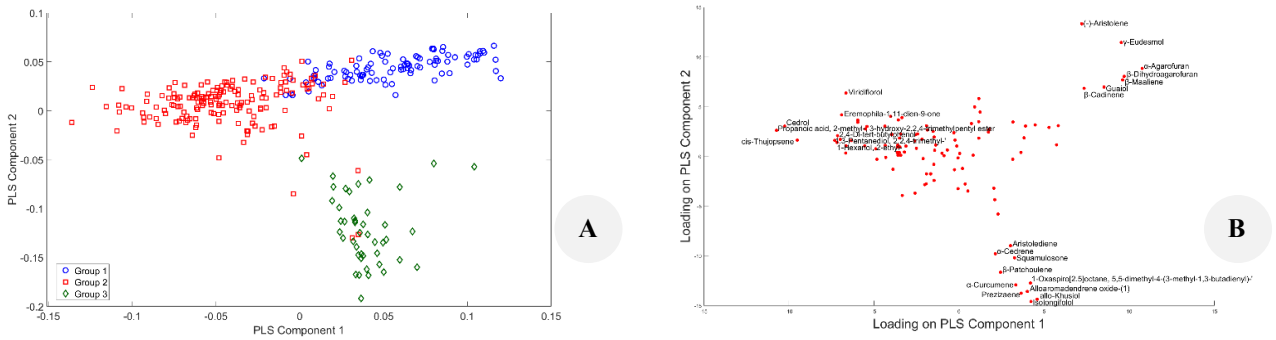

Figure 10. A. Scores plot and B. Loadings plot of PLS-DA based von peak table (304 × 110)

Table 5. Significant compounds identified and their corresponding biological activities according to literature

Compound	Biology activities	References
Guaiol	Antioxidant and antimicrobial	Tian et al. 2021
α-costol	Antioxidant and antimicrobial	Tian et al. 2021
viridiflorol	Antifungal	Ntana et al. 2021
β-vetivenene	Antioxidant and anticancer	Zhang et al. 2025
β-costol	Antioxidant and anticancer	Zhang et al. 2025
β-cadinene	Antibacterial and antifungal	Shalini et al. 2024
α-curcumene	Antifungal and antibacterial	da Silva et al. 2015
cedrol	Cardioprotective effects	Rastegar-Moghaddam et al. 2024
cis-Z-α-bisabolene	Antioxidant and antimicrobial	Obasi and Ogugua 2021
squamulosone	Insecticidal	Korada et al. 2010
Propanoic acid, 2-methyl-, 3-hydroxy-2,2,4-trimethylpentyl ester	Potential biomarker for lung cancer	Rubio-Sánchez et al. 2023
β-patchoulene	Protect against hepatic ischemia/reperfusion injury	Tao et al. 2022
2,4-di-tert-butylphenol	Antifungal, antioxidant, and anticancer	Varsha et al. 2015
α-cedrene	Antioxidant, antimicrobial, and antibiofilm	Bellioua et al. 2024
2-ethyl-1-hexanol	Antifungal	Marcomini et al. 2025
4,6,6-trimethyl-2-(3-methylbuta-1,3-dienyl)-3-oxatricyclo[5.1.0.0(2,4)]octane	Antidiabetic and antimicrobial	Kandsi et al. 2024

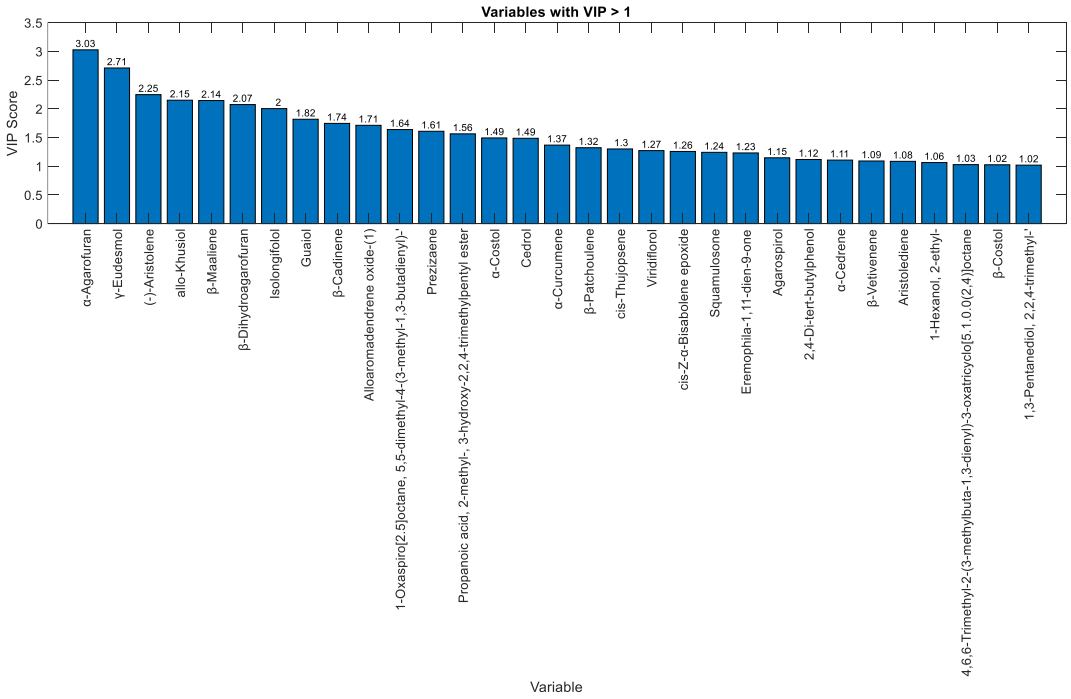

Figure 11. VIP score plot based on the PLS-DA model

The peak areas of chemical compounds with substantial VIP scores, identified as key discriminant markers, are illustrated in the bar chart (Figure 12). Among these, α-agarofuran is the most discriminant, exhibiting high average intensity in Group 1, moderate levels in Group 2, and minimal or absent levels in Group 3. A similar decreasing trend from Group 1 to Group 3 was observed for (-)-aristolene, β-maaliene, and β-dihydroagarofuran. In contrast, allo-khusiol is predominantly abundant in Group 3, with only trace amounts in Groups 1 and 2. These bar chart patterns align with the PLS-DA score and loading plots, reinforcing the reliability of the group separation and compound discrimination.

In another study, Qian et al. (2025) applied GC-MS to analyze agarwood oil, identifying a total of 127 volatile compounds, with sesquiterpenes and aromatic compounds being predominant. The study also highlighted 26 key markers with VIP scores greater than 1, including α-gurjunene, agarospirol, guaiol, γ-eudesmol, and 2-phenylethyl-4H-chromen-4-one, some of which were also

observed in the present study. Using a supervised learning approach, the authors were able to assess agarwood quality based on species and geographic origin. Hung et al. (2014), on the other hand, performed PCA to differentiate five agarwood samples from Indonesia and Vietnam, identifying compounds that distinguished the samples by source (α-cedrene) and price (α-copaene, caryophyllene, guaiene), without incorporating supervised classification. In this study, information on the origin and species of the samples was unavailable, preventing classification based on the source or price.

The agarwood samples were distinctly separated into three groups, each defined by characteristic compounds that allowed reliable classification. The peak table was divided into training and test sets for classification using PLS-DA. The model shows strong predictive performance, achieving a mean accuracy of 90.31±2.62% across 100 iterations. Figure 13 presents the accuracy distribution and confusion matrix. Three Class 1 samples were misclassified as Class 2, and five Class 2 samples were misclassified as Class 1. Class 3 achieved perfect classification with no errors.

The findings of this study suggest that common characteristics of agarwood include aristolene, γ-eudesmol, β-elemene, β-dihydroagarofuran, and 2-butanone, 4-phenyl. Compounds contributing to the distinctive aroma of agarwood, which can be used to differentiate samples, include α-agarofuran, γ-eudesmol, and allo-khusiol. While the models demonstrate stable clustering and reliable identification of discriminant compounds, their application in industry should be interpreted cautiously. The dataset represents chemically diverse wild agarwood samples without confirmed species or geographic origin. External validation using independent commercial datasets and standardized analytical protocols is necessary before implementation in large-scale quality assessment or grading system.

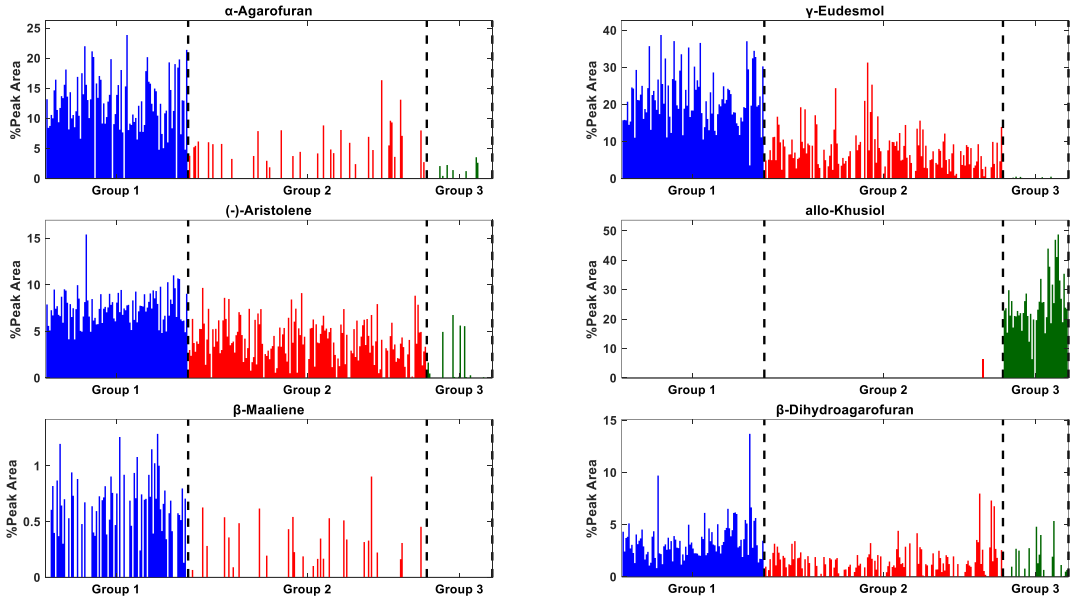

Figure 12. Bar chart of the percentage peak area of key discriminating compounds

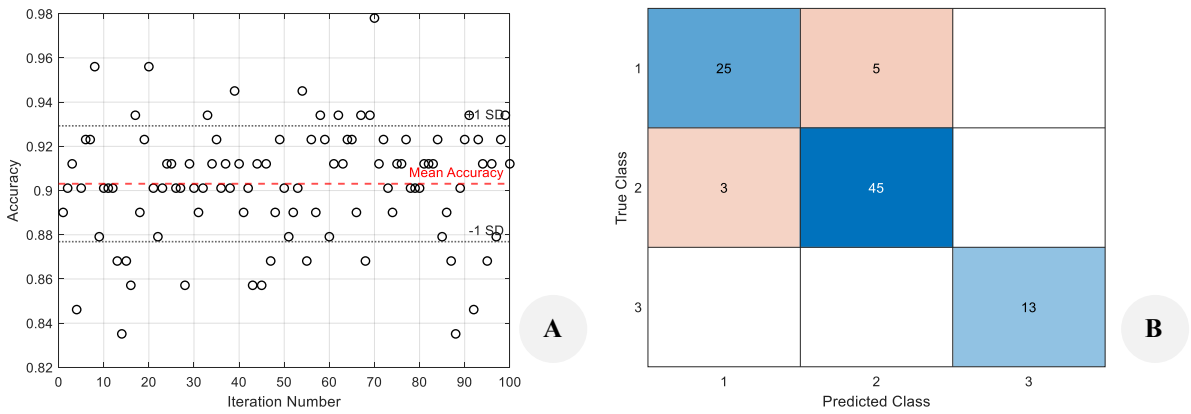

Figure 13. A. PLS-DA classification accuracy over 100 iterations, and B. The confusion matrix

In conclusion, this study tested the hypothesis that volatile chemical profiles can objectively differentiate agarwood quality using HS-SPME-GC-MS combined with chemometric analysis. The results support this hypothesis: analysis of 304 wild agarwood samples identified 1,036 volatile compounds and consistently classified them into three chemometric groups. While PCA revealed partial overlap due to high chemical complexity, PLS-DA provided strong discrimination, achieving a mean classification accuracy of 90.31±2.62% across 100 iterations. Key discriminatory compounds-α-agarofuran, γ-eudesmol, (-)-aristolene, allo-khusiol, β-maaliene, and β-dihydroagarofuran-were identified as reliable chemical indicators of relative quality. Although species identity and provenance were unavailable, these findings establish a robust chemical reference for wild agarwood and demonstrate that chemometric-assisted profiling can complement traditional sensory evaluation with reproducible, data-driven classification. The future work should aim at validating chemometric models using independent commercial and geographically verified agarwood samples, investigating species-specific volatile markers to improve quality prediction and integrating sensory evaluation with chemical profiling to develop standardized grading guidelines.

ACKNOWLEDGEMENTS

The authors thank Swee Ho Construction Works for funding this project (IRG/F07/SHCW/86107/2023). The authors also thank Universiti Malaysia Sarawak, Malaysia, for supporting this work.

REFERENCES

Abdulah L, Susanti R, Rahajoe JS et al. 2022. Feasibility of agarwood cultivation in Indonesia: Dynamic system modeling approach. Forests 13 (11): 1869. https://doi.org/10.3390/f13111869.

Adam AZ, Lee SY, Mohamed R. 2017. Pharmacological properties of agarwood tea derived from *Aquilaria* (Thymelaeaceae) leaves: An emerging contemporary herbal drink. J Herb Med 10: 37-44. https://doi.org/10.1016/j.hermed.2017.06.002.

Ahmaed DT, Mohammed M, Masaad AM, Tajuddin SN. 2017. Investigation of agarwood compounds in *Aquilaria malaccensis* and *Aquilaria rostrata* chipwood by using solid phase microextraction. Biomed J Sci Tech Res 1 (6): 1609-1616. https://doi.org/10.26717/BJSTR.2017.01.000499.

Ahmaed DT, Osman FAY, Masaad AM, Tajuddin SN. 2022. Phytochemical screening and characterization of agarwood (*Aquilaria malaccensis*) chips wood grade as incense headspace volatile compounds by GC-MS. Ms. Q. TOF, SPME. Omdurman J Pharm Sci 2: 85-104. https://doi.org/10.52981/ojps.v2i2.2206.

Bellioua S, Polito F, Dilagui I, Benrazzouk K, De Feo V, Bekkouche K, Larhsini M, Markouk M. 2024. Chemical composition, antioxidant, antimicrobial, and antibiofilm activities of essential oil from the Moroccan endemic plant, *Calendula maroccana* (Ball) BD Jacks. J Essent Oil Bear Plants 27 (3): 678-692. https://doi.org/10.1080/0972060X.2024.2338174.

Bi Y, Wu J, Zhai X, Shen S, Tang L, Huang K, Zhang D. 2021. Application of partial least squares-discriminate analysis model based on water chemical compositions in identifying water inrush sources from multiple aquifers in mines. Geofluids 2021: 6663827. https://doi.org/10.1155/2021/6663827.

Charikar M, Chatziafratis V, Niazadeh R, Yaroslavtsev G. 2019. Hierarchical clustering for Euclidean data. PMLR 89: 2721-2730.

Chen ST, Rao YK. 2022. An overview of agarwood, phytochemical constituents, pharmacological activities, and analyses. Tradit Med 3 (1): 1-71. https://doi.org/10.35702/Trad.10008.

Chhipa H, Chowdhary K, Kaushik N. 2017. Artificial production of agarwood oil in *Aquilaria* sp. by fungi: A review. Phytochem Rev 16: 835-860. https://doi.org/10.1007/s11101-017-9492-6.

Choudhary MI, Musharraf SG, Nawaz SA, Anjum S, Parvez M, Fun HK. 2005. Microbial transformation of (-)-isolongifolol and butyrylcholinesterase inhibitory activity of transformed products. Bioorg Med Chem 13: 1939-1944. https://doi.org/10.1016/j.bmc.2005.01.015.

da Silva GS, Pozzatti P, Rigatti F, Hörner R, Alves SH, Mallmann CA, Heinzmann BM. 2015. Antimicrobial evaluation of sesquiterpene α-curcumene and its synergism with imipenem. J Microbiol Biotechnol Food Sci 4 (5): 434-436. https://doi.org/10.15414/jmbfs.2015.4.5.434-436.

Fei N, Gao Y, Lu Z, Xiang T. 2021. Z-Score Normalization, Hubness, and Few-Shot Learning. Proc IEEE Intl Conf Comput Vision 2021: 142-151. https://doi.org/10.1109/ICCV48922.2021.00021.

Gutiérrez S, Overmans S, Wellman GB, Samaras VG, Oviedo C, Gede M, Szekely G, Lauersen KJ. 2024. A synthetic biology and green bioprocess approach to recreate agarwood sesquiterpenoid mixtures. Green Chem 26: 2577-2591. https://doi.org/10.1039/D3GC03708H.

Haron MH. 2020. Agarwood Oil Quality Grading Model using Self-Organizing Map (SOM). [Dissertation]. Universiti Teknologi MARA, Shah Alam.

Hidayat W, Shakaff AYM, Ahmad MN, Adom AH. 2010. Classification of agarwood oil using an electronic nose. Sensors 10 (5): 4675-4685. https://doi.org/10.3390/s100504675.

Hung CH, Lee CY, Yang CL, Lee MR. 2014. Classification and differentiation of agarwoods by using non-targeted HS-SPME-GC/MS and multivariate analysis. Anal Method 6: 7449-7456. https://doi.org/10.1039/C4AY01151A.

Ismail N, Azah MAN, Jamil M, Rahiman MHF, Tajuddin SN, Taib MN. 2013. Analysis of high-quality agarwood oil chemical compounds by means of SPME/GC-MS and Z-score technique. Malays J Anal Sci 17: 403-413

Ismail SN, Maulidiani M, Akhtar MT, Abas F, Ismail IS, Khatib A, Ali NAM, Shaari K. 2017. Discriminative analysis of different grades of gaharu (*Aquilaria malaccensis* Lamk.) via ^{1}H-NMR-based metabolomics using PLS-DA and random forests classification models. Molecules 22 (10): 1612. https://doi.org/10.3390/molecules22101612.

Januzaj Y, Beqiri E, Luma A. 2023. Determining the optimal number of clusters using the silhouette score as a data mining technique. Intl J Online Biomed Eng 19: 174-182. https://doi.org/10.3991/ijoe.v19i04.37059.

Jolliffe IT, Cadima J. 2016. Principal component analysis: A review and recent developments. Philos Trans A Math Phys Eng Sci 374: 20150202. https://doi.org/10.1098/rsta.2015.0202.

Kalra R, Kaushik N. 2017. A review of chemistry, quality, and analysis of the infected agarwood tree (*Aquilaria* sp.). Phytochem Rev 16: 1045-1079. https://doi.org/10.1007/s11101-017-9518-0.

Kandsi F, Abdnim R, Benkhaira N, Zahra Lafdil F, Bnouham M, Yamani B, Naceiri Mrabti H, Wondmie GF, Bin Jardan YA, Ibenmoussa S, El Hachlafi N. 2024. Integrated assessment of phytochemicals, antilipase, hemoglobin antiglycation, antihyperglycemic, antifungal, and antibacterial properties of *Vetiveria zizanioides* (L.) Nash. Intl J Food Prop 27 (1): 1150-1166. https://doi.org/10.1080/10942912.2024.2387435.

Kao WY, Hsiang CY, Ho SC, Ho TY, Lee KT. 2018. Chemical profiles of incense smoke ingredients from agarwood by headspace gas chromatography-tandem mass spectrometry. Molecules 23 (11): 2969. https://doi.org/10.3390/molecules23112969.

Karna A, Gibert K. 2022. Automatic identification of the number of clusters in hierarchical clustering. Neural Comput Appl 34: 119-134. https://doi.org/10.1007/s00521-021-05873-3.

Korada RR, Naskar S, Mukherjee A, Jayaprakas CA. 2010. Management of sweet potato weevil, *Cylas formicarius*: A world review. J Root Crop 36: 14-26.

Kristanti AN, Tanjung M, Aminah NS. 2018. Secondary metabolites of *Aquilaria*, a Thymelaeaceae genus. Mini-Rev Org Chem 15: 36-55. https://doi.org/10.2174/1570193X14666170721143041.

Lasalvia M, Capozzi V, Perna G. 2022. A comparison of PCA-LDA and PLS-DA techniques for classification of vibrational spectra. Appl Sci 12: 5345. https://doi.org/10.3390/app12115345.

Lever J, Krzywinski M, Altman N. 2017. Points of significance: Principal component analysis. Nat Method 14: 641-642. https://doi.org/10.1038/nmeth.4346.

Ma S, Fu Y, Li Y, Wei P, Liu Z. 2021. The formation and quality evaluation of agarwood induced by the fungi in *Aquilaria sinensis*. Ind Crop Prod 173: 114129. https://doi.org/10.1016/j.indcrop.2021.114129.

Ma X. 2024. Utilizing principal component analysis to enhance machine learning in bankruptcy prediction: A comparative investigation. Appl Comput Eng 68: 304-310. https://doi.org/10.54254/2755-2721/68/20241500.

Marcomini EK, Rezende PS, Gomes J, Leite CR, de Oliveira KM, Pomini AM, Svidzinski TI, Negri M. 2025. 2-ethyl-1-hexanol: A promising molecule against priority fungal pathogens. J Appl Microbiol 136 (7): lxaf165. https://doi.org/10.1093/jambio/lxaf165.

Monggoot S, Popluechai S, Gentekaki E, Pripdeevech P. 2017. Fungal endophytes: An alternative source for production of volatile compounds from agarwood oil of *Aquilaria subintegra*. Microb Ecol 74: 54-61. https://doi.org/10.1007/s00248-016-0908-4.

Naziz PS, Das R, Sen S. 2019. The scent of stress: Evidence from the unique fragrance of agarwood. Front Plant Sci 10: 840. https://doi.org/10.3389/fpls.2019.00840.

Ntana F, Bhat WW, Johnson SR, Jørgensen HJ, Collinge DB, Jensen B, Hamberger B. 2021. A sesquiterpene synthase from the endophytic fungus *Serendipita indica* catalyzes the formation of viridiflorol. Biomolecules 11 (6): 898. https://doi.org/10.3390/biom11060898.

Obasi DC, Ogugua VN. 2021. GC-MS analysis, pH, and antioxidant effect of Ruzu herbal bitters on alloxan-induced diabetic rats. Biochem Biophys Rep 27: 101057. https://doi.org/10.1016/j.bbrep.2021.101057.

Ogbuabor G, Ugwoke FN. 2018. Clustering algorithm for a healthcare dataset using the silhouette score value. Intl J Comput Sci Inf Technol 10: 27-37. https://doi.org/10.56899/151.05.07.

Peiró-Vila P, Pérez-Gracia C, Baeza-Baeza JJ, García-Alvarez-Coque MC, Torres-Lapasió JR. 2024. Analysis and classification of tea varieties using high-performance liquid chromatography and global retention models. J Chromatogr A 1730: 465128. https://doi.org/10.1016/j.chroma.2024.465128.

Qian SZ, Jiang YM, Yan QL, Wu DH, Zhang WX, Chung JP. 2025. Visualization of OPLS class models of GC-MS-based metabolomics data for identifying agarwood essential oil extracted by hydro-distillation. Sci Rep 15: 5421. https://doi.org/10.1038/s41598-025-85976-2.

Rachwał A, Popławska E, Gorgol I, Cieplak T, Pliszczuk D, Skowron Ł, Rymarczyk T. 2023. Determining the quality of a dataset in clustering terms. Appl Sci 13: 2942. https://doi.org/10.3390/app13052942.

Randriamihamison N, Vialaneix N, Neuvial P. 2021. Applicability and interpretability of Ward's hierarchical agglomerative clustering with or without contiguity constraints. J Classif 38: 363-389. https://doi.org/10.1007/s00357-020-09377-y.

Rastegar-Moghaddam SH, Amirahmadi S, Akbarian M, Sharizina M, Beheshti F, Rajabian A, Ghalibaf MH, Azimi M, Mahmoudabady M, Hosseini M. 2024. Cardioprotective effect of cedrol in a systemic inflammation model induced by lipopolysaccharide: Biochemical and histological verification. J Cardiovasc Thorac Res 16 (2): 120-128. https://doi.org/10.34172/jcvtr.33112.

Rubio-Sánchez R, Ríos-Reina R, Ubeda C. 2023. Effect of chemotherapy on urinary volatile biomarkers for lung cancer by HS-SPME-GC-MS and chemometrics. Thorac Cancer 14 (36): 3522-3529. https://doi.org/10.1111/1759-7714.15154.

Sen S, Dehingia M, Talukdar NC, Khan M. 2017. Chemometric analysis reveals links in the formation of fragrant bio-molecules during agarwood (*Aquilaria malaccensis*) and fungal interactions. Sci Rep 7: 44406. https://doi.org/10.1038/srep44406.

Shalini K, Guleria S, Salaria D, Rolta R, Fadare OA, Mehta J, Awofisayo O, Mandyal P, Shandilya P, Kaushik N, Choi EH, Chandel SR, Kaushik NK. 2024. Antimicrobial potential of phytocompounds of *Acorus calamus*: An in silico approach. J Biomol Struct Dyn 42 (5): 2726-2737. https://doi.org/10.1080/07391102.2023.2209653.

Shao H, Mei WL, Kong FD, Dong WH, Gai CJ, Li W, Zhu GP, Dai HF. 2016. Sesquiterpenes of agarwood from *Gyrinops salicifolia*. Fitoterapia 113: 182-187. https://doi.org/10.1016/j.fitote.2016.07.015.

Shi X, Song J, Wang H, Lv X, Zhu Y, Zhang W, Bu W, Zeng L. 2023. Improving soil organic matter estimation accuracy by combining optimal spectral preprocessing and feature selection methods based on pXRF and vis-NIR data fusion. Geoderma 430: 116301. https://doi.org/10.1016/j.geoderma.2022.116301.

Sundaraj S, Mediani A, Rodriques KF, Baharum SN. 2023. GC-MS olfactometry reveals that sesquiterpenes α-humulene and δ-cadinene significantly influence the aroma of treated *Aquilaria malaccensis* essential oil. Aust J Crop Sci 17: 893-901. https://doi.org/10.21475/ajcs.23.17.12.p3916.

Tan CS, Isa NM, Ismail I, Zainal Z. 2019. Agarwood induction: Current developments and future perspectives. Front Plant Sci 10: 122. https://doi.org/10.3389/fpls.2019.00122.

Tao T, Ye B, Xu Y, Wang Y, Zhu Y, Tian Y. 2022. β-Patchoulene preconditioning protects mice against hepatic ischemia-reperfusion injury by regulating Nrf2/HO-1 signaling pathway. J Surg Res 275: 161-171. https://doi.org/10.1016/j.jss.2022.02.001.

Tian CP, Yao XD, Lu JH, Shen LQ, Wu AQ. 2021. GC-MS fingerprints of essential oils from agarwood grown in wild and artificial environments. Trees 35: 2105-2117. https://doi.org/10.1007/s00468-021-02177-w.

Tufariello M, Pati S, Palombi L, Grieco F, Losito I. 2022. Use of multivariate statistics in the processing of data on wine volatile compounds obtained by HS-SPME-GC-MS. Foods 11 (7): 910. https://doi.org/10.3390/foods11070910.

Varsha KK, Devendra L, Shilpa G, Priya S, Pandey A, Nampoothiri KM. 2015. 2, 4-Di-tert-butyl phenol as the antifungal, antioxidant bioactive purified from a newly isolated *Lactococcus* sp. Intl J Food Microbiol 211: 44-50. https://doi.org/10.1016/j.ijfoodmicro.2015.06.025.

Villareal JF, Abasolo WP, Mendoza RC. 2022. Fiber morphology and extractive content of *Aquilaria cumingiana* (Decne.) Ridl. wood from Davao Oriental, Philippines. Philipp J Sci 151 (5): 1623-1631. https://doi.org/10.56899/151.05.07.

Wang Y, Hussain M, Jiang Z, Wang Z, Gao J, Ye F, Mao R, Li H. 2021. *Aquilaria* species (Thymelaeaceae) distribution, volatile and non-volatile phytochemicals, pharmacological uses, agarwood grading system, and induction methods. Molecules 26 (24): 7708. https://doi.org/10.3390/molecules26247708.

Wang Y, Lin S, Zhao L, Sun X, He W, Zhang Y, Dai YC. 2019. *Lasiodiplodia* spp. associated with *Aquilaria crassna* in Laos. Mycol Prog 18: 683-701. https://doi.org/10.1007/s11557-019-01481-7.

Ward Jr JH. 1963. Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58: 236-244. https://doi.org/10.2307/2282967.

Xu Z, Fang X, Zhi Y, Xiao X, Yang J, Hu J, Zhu H, Chen F, Cheng W, Liu T, Lu L. 2025. Systematic identification of terpene synthases from sacred lotus (*Nelumbo nucifera*) and heterologous biosynthesis of the insecticidal and antimicrobial compound γ-eudesmol. Hortic Res 12 (10): uhaf191. https://doi.org/10.1093/hr/uhaf191.

Xue BX, Liu SX, Oduro PK, Mireku-Gyimah NA, Zhang LH, Wang Q, Wu HH. 2023. Vasodilatory constituents of essential oil from *Nardostachys jatamansi* DC.: Virtual screening, experimental validation, and the potential molecular mechanisms. Arab J Chem 16: 104911. https://doi.org/10.1016/j.arabjc.2023.104911.

Zhang SY, Tibpromma S, Karunarathna SC, Ye S, Mapook A, Wang YH, Xu JC. 2022. Endophytic fungi of agarwood and their chemical compounds: A review. Fungal Biotech 2: 16-35.

Zhang X, Wang LX, Hao R, Huang JJ, Zargar M, Chen MX, Zhu FY, Dai HF. 2024. Sesquiterpenoids in agarwood: Biosynthesis, microbial induction, and pharmacological activities. J Agric Food Chem 72: 23039-23052. https://doi.org/10.1021/acs.jafc.4c06383.

Zhang Y, Wu S, Zhang B, Zhou X, Zhou W, Zhang W, Gao X, Chen X. 2025. Determination of antitumor active ingredients in agarwood essential oil by Gas Chromatography-Mass Spectrometry (GC-MS) and grey relational analysis. Anal Lett 58 (2): 341-363. https://doi.org/10.1080/00032719.2024.2325570.