

## Article

# Metabolomics-Driven Discovery of an Introduced Species and Two Malaysian *Piper betle* L. Variants

Muhamad Faris Osman <sup>1,2</sup>, Soo Yee Lee <sup>1</sup>, Shahrul Razid Sarbini <sup>3</sup>, Siti Munirah Mohd Faudzi <sup>1,4</sup>, Shamsul Khamis <sup>5</sup>, Badrul Hisyam Zainudin <sup>6</sup> and Khozirah Shaari <sup>1,\*</sup>

- <sup>1</sup> Natural Medicines and Products Research Laboratory (NaturMeds), Institute of Bioscience, Universiti Putra Malaysia, UPM, Serdang 43400, Selangor, Malaysia; farisosman@iium.edu.my (M.F.O.); daphne.leesooyee@gmail.com (S.Y.L.); sitimunirah@upm.edu.my (S.M.M.F.)
  - <sup>2</sup> Department of Pharmaceutical Chemistry, Kulliyah of Pharmacy, International Islamic University Malaysia, Kuantan 25200, Pahang, Malaysia
  - <sup>3</sup> Department of Crop Science, Faculty of Agricultural Science and Forestry, Universiti Putra Malaysia, Bintulu 97008, Sarawak, Malaysia; shahrulrazid@upm.edu.my
  - <sup>4</sup> Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, UPM, Serdang 43400, Selangor, Malaysia
  - <sup>5</sup> Department of Biological Sciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, UKM, Bangi 43600, Selangor, Malaysia; shamsulk@ukm.edu.my
  - <sup>6</sup> Analytical Services Laboratory, Chemistry and Technology Division, Malaysian Cocoa Board, Cocoa Innovation and Technology Centre, Lot 12621 Kawasan Perindustrian Nilai, Nilai 71800, Negeri Sembilan, Malaysia; badrul@koko.gov.my
- \* Correspondence: khozirah@yahoo.com.my; Tel.: +60-13-3420686



**Citation:** Osman, M.F.; Lee, S.Y.; Sarbini, S.R.; Mohd Faudzi, S.M.; Khamis, S.; Zainudin, B.H.; Shaari, K. Metabolomics-Driven Discovery of an Introduced Species and Two Malaysian *Piper betle* L. Variants. *Plants* **2021**, *10*, 2510. <https://doi.org/10.3390/plants10112510>

Academic Editors: Daniel Potter and Manuel B. Crespo

Received: 30 September 2021  
Accepted: 1 November 2021  
Published: 19 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The differences in pungency of “sirih” imply the probable occurrence of several variants of *Piper betle* L. in Malaysia. However, the metabolite profiles underlying the pungency of the different variants remain a subject of further research. The differences in metabolite profiles of selected Malaysian *P. betle* variants were thus investigated; specifically, the leaf aqueous methanolic extracts and essential oils were analyzed via <sup>1</sup>H-NMR and GC-MS metabolomics, respectively. Principal component analysis (PCA) of the <sup>1</sup>H-NMR spectral data showed quantitative differences in the metabolite profiles of “sirih melayu” and “sirih india” and revealed an ambiguous group of samples with low acetic acid content, which was identified as *Piper rubro-venosum* hort. ex Rodigas based on DNA sequences of the internal transcribed spacer 2 (ITS2) region. The finding was supported by PCA of two GC-MS datasets of *P. betle* samples obtained from several states in Peninsular Malaysia, which displayed clustering of the samples into “sirih melayu” and “sirih india” groups. Higher abundance of chavicol acetate was consistently found to be characteristic of “sirih melayu”. The present research has provided preliminary evidence supporting the notion of occurrence of two *P. betle* variants in Malaysia based on chemical profiles, which may be related to the different genders of *P. betle*.

**Keywords:** *Piper betle* L.; *Piper rubro-venosum* hort. ex Rodigas; <sup>1</sup>H-NMR; GC-MS; PCA; essential oils; acetic acid; chavicol acetate; chavibetol; chavibetol acetate

## 1. Introduction

In many Asian countries, chewing the leaves of betel vine (*Piper betle* L.) is a common social practice, largely embedded in the tradition, custom, or ritual of indigenous populations. The betel leaves are usually chewed together with areca nut and slaked lime, along with selected spices. The practice is believed to be beneficial in fighting bad breath (halitosis), while the leaves are also a traditional remedy for treating nosebleeds (epistaxis), and promoting pus removal [1,2]. Phylogenetic analyses have provided strong support for the placement of *P. betle* within the tropical Asian clade of *Piper*, specifically the west of Wallace’s line (WWL) clade. The WWL is a region that includes the Malay Peninsula, Sumatra, Java, and Borneo [3–5]. Locally known as “sirih” in Malay language, *P. betle* is

an important plant in the Malaysian herbal industry [6,7]. Occurrence of several “sirih” variants, namely “sirih melayu”, “sirih cina”, and “sirih udang”, has been recorded in a transliterated manuscript of traditional Malay medicine [8].

Meanwhile, in India, there are more than 100 different variants of *P. betle*, which are termed landraces. A landrace is defined as a cultivated variety of plant, which has evolved in a geographical region as a result of its adaptation to the region’s edaphic and climatic conditions, and to the traditional management and consumption [9]. Since *P. betle* is a dioecious plant, the landrace is identified either as male or female based on its floral morphology [10]. A macroscopic evaluation of eight eastern Indian *P. betle* landraces emphasized that the variation between the landraces was apparent in terms of color and size of the leaves. The color of the leaves ranged from greenish yellow to dark green and the landraces with dark green leaves were larger in size compared to the landraces with greenish yellow leaves [11].

Previous mass spectrometry (MS) metabolomic studies of various Indian *P. betle* landraces have revealed clustering of leaf samples into several major groups based on the variation of metabolite profiles in intact plant materials, hexane extracts, and essential oils of dried and fresh leaves. Despite inconsistency in the reporting of one of the major metabolites that was identified either as eugenol or its isomer chavibetol, the studies reported several phenylpropenes, namely allylpyrocatechol, chavicol, and their acetate esters as the chemical markers of the resolved groups [10,12,13]. Phenylpropenes and their acetate esters are the characteristic constituents of *P. betle* leaf extracts that impart pungency upon chewing of the leaves [14].

While the difference between *P. betle* variants was physically noticeable from the color of the leaves, the variation was described by the individuals we communicated with as more pronounced in terms of their pungency. Since the notion of occurrence of different variants is solely based on human perception, a scientific investigation was deemed necessary to further investigate the association between varying pungency of the variants and their metabolite profiles. Hence, the present research employed a combination of  $^1\text{H-NMR}$  and GC-MS metabolomics to explore the variation in metabolite profiles of selected Malaysian *P. betle* variants and identify the chemical markers that may be utilized to discriminate between the different variants.

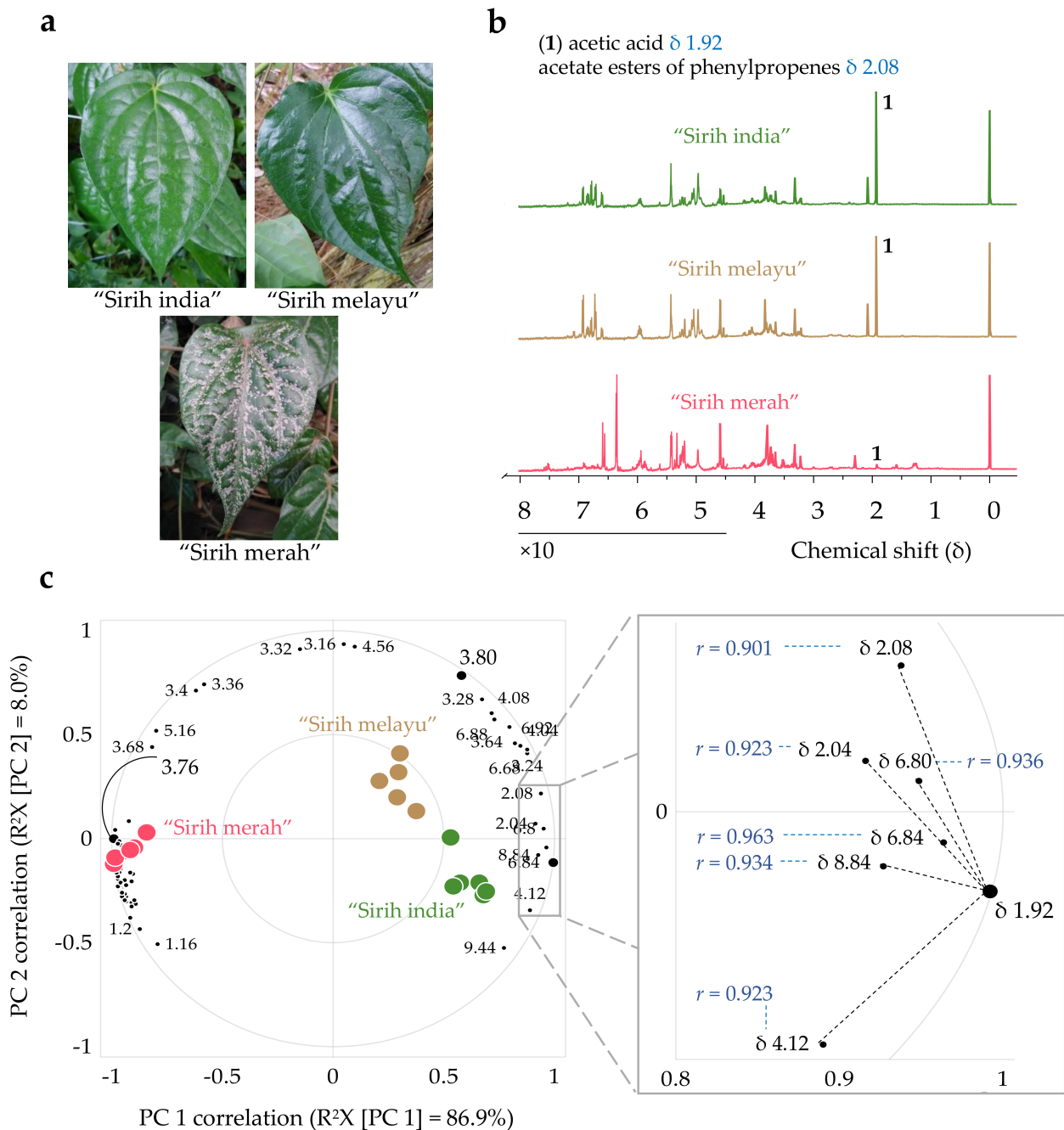
## 2. Results and Discussion

### 2.1. Principal Components Analysis of $^1\text{H-NMR}$ Spectra of Three “Sirih” Variants

The investigation was initiated with a pilot experiment. Three “sirih” plants, growing in close vicinity to each other, were conveniently sampled from the herbal garden of Universiti Putra Malaysia (UPM) Agriculture Park. The plants were locally known as “sirih melayu” ( $n = 1$ ), “sirih india” ( $n = 1$ ), and “sirih merah” ( $n = 1$ ) (Figure 1a). Six mature leaves were sampled from each plant.  $^1\text{H-NMR}$  spectra were acquired for the aqueous methanol-D<sub>4</sub> (50% *v/v*) extracts of the leaf samples, followed by principal component analysis (PCA) of the spectral data. Representative  $^1\text{H-NMR}$  spectra of the samples are shown in Figure 1b and Figure S1. The intensity of the NMR signals at  $\delta$  1.92 and  $\delta$  2.08 for “sirih melayu” and “sirih india” was remarkably higher than those for “sirih merah”, and there were noticeable differences in the intensity of the signals in the sugar ( $\delta$  3.00–5.00) and aromatic ( $\delta$  6.00–8.00) regions for all three groups of samples.

Prior to the PCA, the  $^1\text{H-NMR}$  spectra were converted to an  $18 \times 221$  data matrix (18 samples of 221 variables) and the data matrix was preprocessed using three preprocessing methods commonly employed in metabolomics; i.e., unit variance (UV) scaling, Pareto scaling, and mean centering, to find the method that explained the highest percentage of variation in the first two principal components (PC) of the data matrix. Mean centering of the variables enabled the first two PCs to explain the highest percentage of variation in the data matrix ( $R^2\text{X}$  [cumulative up to PC 2] = 94.9%), with predicted percentage of variation estimated by leave-one-out cross-validation ( $Q^2\text{X}$  [cumulative up to PC 2]) of 87.8% (Figure S2). Mean centering the variables is advantageous because it removes an offset from the data matrix, reduces

the model rank and improves the model fit [15]. Hence, the PCA model of mean-centered data matrix was selected for the interpretation of discriminating variables (Figure 1c).



**Figure 1.** Principal component analysis (PCA) of  $^1\text{H}$ -NMR spectra of "sirih" leaf aqueous methanolic extracts. (a) Leaf of "sirih india", "sirih melayu", and "sirih merah" sampled in the pilot experiment. (b) Representative  $^1\text{H}$ -NMR spectra of the three "sirih" variants showing that acetic acid (1) and acetate esters of phenylpropenes are the major metabolites in the leaf aqueous methanolic extracts of "sirih india" and "sirih melayu" but remarkably low in "sirih merah". The spectra were normalized to TSP peak at  $\delta$  0.00 (intensity of TSP peak = 100). The intensity at  $\delta$  4.50–8.00 was increased tenfold to help visualize the signals in the olefinic and aromatic regions. The full spectra are shown in Figure S1. (c) Correlation-scaled PCA biplot of mean-centered data matrix showing only well-modeled variables (cumulative percentage of variation of the variable predicted by the first two PCs, as estimated by cross-validation ( $Q^2VX$  [cumulative up to PC 2])  $\geq 80\%$ ). The zoomed-in region indicates the variables that have strong positive correlation with the  $\delta$  1.92 variable ( $r > 0.900$ ). Complete labeling of characteristic variables of "sirih merah" can be viewed in Figure S3.

Based on directions and magnitudes of the loading vectors in the first two PCs, PC scores of the samples were found to be heavily influenced by the  $\delta$  1.92, 3.76, and 3.80 variables (Figure S2a). The  $\delta$  3.76 and 3.80 variables belong to methoxy protons. However, identification of the exact metabolites giving rise to the  $^1\text{H-NMR}$  signals binned into these variables was not trivial, as the methoxy protons may have belonged to any metabolite in the groups of *O*-methylated phenylpropanoids [16–18]. In contrast, the  $\delta$  1.92 variable, the intensity of which was remarkably high in “sirih india” and “sirih melayu”, was unequivocally identified as acetic acid (1), based on perfect automatic fit of the singlet at  $\delta$  1.92 to acetic acid in the Chenomx Profiler software library. The fitting also allowed quantification of acetic acid concentration in the samples relative to the concentration of internal standard 3-(trimethylsilyl)propionic-2,2,3,3-D4 acid sodium salt (TSP).

For the three groups of samples, at least one group had a statistically significantly different mean concentration of acetic acid (Welch’s *F* [2, 6.90] = 2805.50,  $p < 0.001$ , estimated  $\omega^2 = 1.00$ ). A Games–Howell post hoc test indicated that the mean concentration (mM) of acetic acid in the leaves of “sirih merah” (mean = 0.36 [95% CI = 0.30, 0.43]) was significantly lower than that of the leaves of “sirih india” (mean = 10.94 [95% CI = 10.54, 11.34], mean difference = 10.57 [95% CI = –11.08, –10.07],  $p < 0.001$ ) and “sirih melayu” (mean = 10.51 [95% CI = 9.88, 11.13], mean difference = 10.14 [95% CI = –10.93, –9.36],  $p < 0.001$ ). However, the difference in the mean concentration of acetic acid in the leaves of “sirih india” and “sirih melayu” was not statistically significant (mean difference = 0.43 [95% CI = –0.38, 1.24],  $p = 0.339$ ). Thus, relying only on the concentration of acetic acid to discriminate between the different variants of *P. betle* leaves was insufficient, because the mean concentration of acetic acid in “sirih india” was not statistically significantly higher than that of “sirih melayu”.

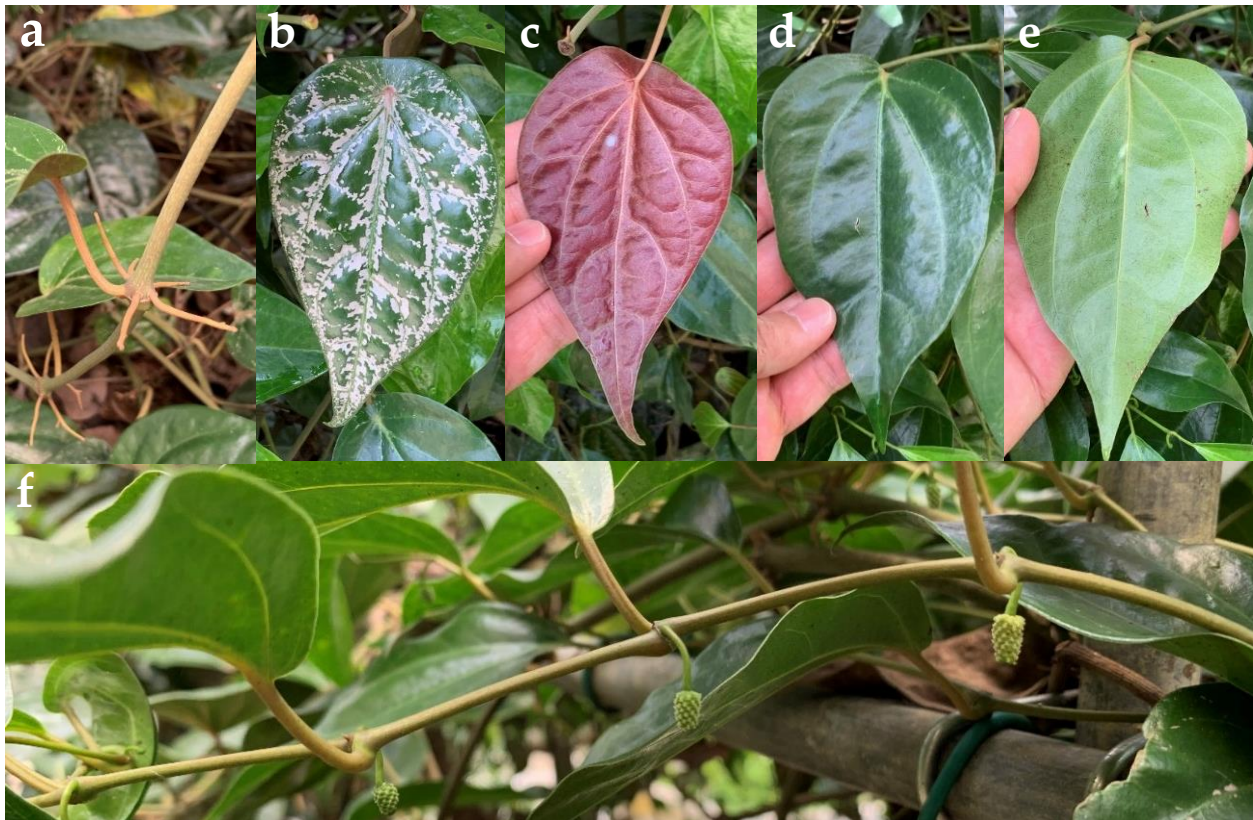
As shown in Figure 1c, acetic acid has strong positive correlation with variables in the aromatic region of the  $^1\text{H-NMR}$  spectra ( $\delta$  6.80 and 6.84) and variables that represent acetate protons of phenylpropene acetate esters ( $\delta$  2.04 and 2.08) [19]. In the biplot, these four variables are close to each other, thus providing support that they may belong to acetate esters of phenylpropenes, which are secondary metabolites commonly reported as constituents of *P. betle* leaf essential oil [20]. Meanwhile, the PC scores of “sirih merah” are heavily influenced by slightly different variables in the aromatic region ( $\delta$  6.28, 6.32, and 6.36, Figure S3), showing an abundance of a group of phenolic metabolites with different substituents on the benzene ring [21]. If the second PC of the biplot shows infraspecific variation within *P. betle*, then the first PC can be hypothesized to be an indication of a variation that is almost ten times greater than the infraspecific variation—infrageneric variation. While “sirih india” and “sirih melayu” were identified as *P. betle* based on morphology of the leaf and climbing habit of the plants, “sirih merah” was initially assumed to be a variant of *P. betle* because of similarities in climbing habit and shape of the leaves. Nevertheless, the findings from this pilot experiment revealed that “sirih merah” may be a different species of *Piper*, thus requiring further verification to confirm its identity.

Besides, further investigation of the variation of phenylpropenes and their acetate esters in the different *P. betle* variants requires more representative samples and specific analytical instruments of higher sensitivity and selectivity. For these reasons, the sample size was increased by sampling *P. betle* variants from selected locations in Peninsular Malaysia and the investigation was continued by GC-MS metabolomics of leaf essential oils of the different variants, thus allowing more focused analysis on the phenylpropenes and their acetate esters.

## 2.2. Identification of “Sirih Merah” as *P. rubro-venosum* hort. ex Rodigas

A literature search using the term “sirih merah” showed that the plant was identified either as *P. crocatum* Ruiz and Pav., *P. ornatum* N.E.Br., or *P. rubro-venosum* hort. ex Rodigas in Indonesia and Thailand [4,17,21–27]. Close examination of voucher specimens of the plant by our local botanist led to identification of the plant as *P. cf. fragile* Benth. Another botanist identified the plant as *P. fragile* and *P. ornatum* complex that displayed several leaf forms,

and stated that the plant is often misidentified as *P. crocatum*, which is a shrubby species from Peru [28]. Other than that, previous research has reported color change of the leaves of the plant when it was planted in a new location [29]. Variations in the morphological characteristics have made the exact species identification of “sirih merah” difficult. In this respect, DNA barcoding, which is a method that takes advantage of the stability of DNA, could be a more effective method to differentiate and resolve closely related species [30]. Based on high similarity (99.39%) of the nuclear ribosomal DNA sequences of the internal transcribed spacer 2 (ITS2) region of “sirih merah” to a reference sequence in GenBank, the variant was identified as *P. rubro-venosum* hort. ex Rodigas (Figure 2 and Figure S4), which is an introduced species in Peninsular Malaysia and Singapore [31]. Despite being identified as a different *Piper* species, GC-MS data of *P. rubro-venosum* leaf essential oils were also included in GC-MS metabolomics to incorporate infrageneric variation into the PCA model. Only two samples of *P. rubro-venosum* leaf essential oils that were obtained from one plant ( $n = 1$ ) were included as there was only one plant available at the time of sample collection and the number of leaves that fulfilled the selection criteria (healthy mature leaves of similar texture and color) allowed us to obtain at maximum only two samples of essential oils (see Figures S5 and S6, and Tables S1 and S2).



**Figure 2.** Photographs of *Piper rubro-venosum* hort. ex Rodigas. (a) Nodal roots; (b) adaxial surface of silver pink leaf, showing silver pink stripes mostly along primary veins; (c) maroon abaxial surface of leaf (b); (d) adaxial surface of plain green leaf; (e) green abaxial surface of leaf (d); (f) leaf-opposed inflorescences. (b,c) are designated as RM whereas (d,e) are designated as RG.

### 2.3. Leaf Essential Oil Profiles of *P. betle* Variants and *P. rubro-venosum*

Further analysis was carried out using GC-MS to identify the leaf essential oil constituents of *P. betle* variants that were sampled in the present research, namely “sirih melayu”, “sirih india”, “sirih” unknown variant, and “sirih manis” (henceforth labeled as *P. betle* ‘melayu’, *P. betle* ‘india’, *P. betle* unknown variant, and *P. betle* ‘manis’) (Figure 3a) and *P. rubro-venosum* followed by PCA of the GC-MS data. In terms of data acquired for analysis of complex

plant extracts, GC-MS provides three-dimensional (3D) data that include retention time and mass spectral information of metabolites detected in the extracts, whereas  $^1\text{H-NMR}$  provides 1D data, which is a combination of all  $^1\text{H-NMR}$  spectra of metabolites detected in the extracts. However, in electron ionization (EI) GC-MS, since the match scores, known as similarity index (SI) or reverse similarity index (RSI), against reference spectra in EI mass spectral libraries depend on (i) algorithms of the library search which vary between different manufacturers, and (ii) reproducibility of analysis, the correct metabolite does not necessarily occupy the first position in the list of hits [32,33]. Hence, relying only on high match scores may lead to misidentification of metabolites.

Examination of both match scores and retention indices of other possible candidates in the list of hits is a powerful filtering method to reduce misidentification of metabolites in GC-MS [34]. Usually match scores of  $\text{SI} \geq 80$  and a linear retention index (LRI) that is within  $\pm 5$  deviation from values reported in the literature are the guiding criteria for identification of metabolites. As another filtering method to reduce misidentification of metabolites, we also analyzed the leaf essential oil samples on two different GC-MS systems and checked for consistency of the metabolite profiles generated by the two systems. The basis for analyzing the samples using two GC-MS systems is that slightly different chromatograms may be acquired when a sample is analyzed using two different GC-MS systems, even though columns of similar dimension and polarity are used. This is due to the difference in a multitude of parameters of the GC-MS systems. For this reason, any generalization regarding metabolite profiles of the sample shall consider the metabolites that are consistently identified in both GC-MS systems.

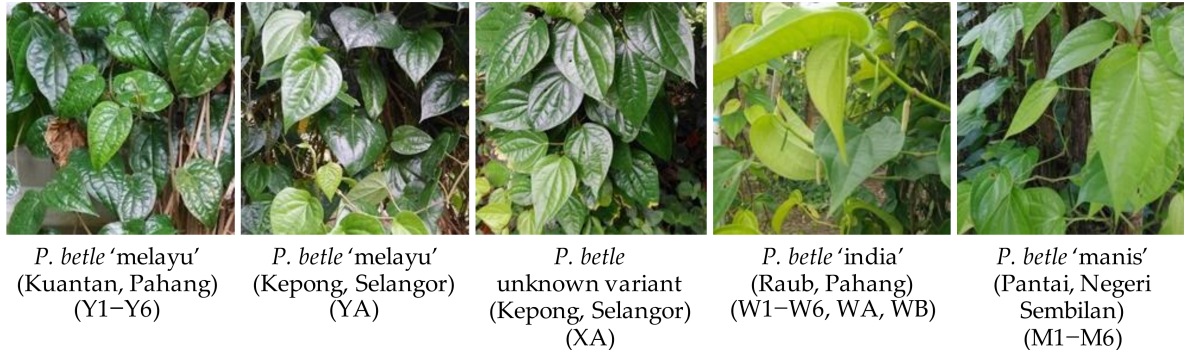
The first GC-MS analysis (GC-MS analysis I) was carried out to examine whether the findings of the pilot experiment were also reflected in the leaf essential oil profiles of the *P. betle* 'melayu', 'india', and unknown variant, and *P. rubro-venosum* that were collected from six different locations in Peninsular Malaysia (Figures S5–S7, and Tables S1 and S2). Based on the results of GC-MS analysis I, the GC-MS parameters for the second analysis (GC-MS analysis II) were improved, which successfully detected a higher number of metabolites. The number of samples was increased by the inclusion of *P. betle* 'manis' in GC-MS analysis II. According to the locals, *P. betle* 'manis' has the least pungency among the variants that were sampled in our research. The zoomed-in base peak chromatograms (BPC) of leaf essential oils of *P. betle* variants and *P. rubro-venosum* are provided in Figures S8–S11 and the identified essential oil constituents are listed in Tables S3–S6 [35–48].

Unlike the pungent aroma of *P. betle* leaf essential oils, the aroma of *P. rubro-venosum* leaf essential oils was notably non-pungent, with higher sweet notes. Representative BPCs for the essential oils of *P. rubro-venosum* leaves with green underside (RG) and maroon underside (RM), *P. betle* 'melayu' (Y1–Y5 and YA), *P. betle* 'india' (W1–W6, WA and WB), and *P. betle* unknown variant (XA) obtained in GC-MS analysis I are shown in Figure 3b. The BPCs show higher intensities of oxygenated monoterpenes linalool (2) and  $\alpha$ -terpineol (3) peaks in the *P. rubro-venosum* leaf essential oil (RG) compared to the *P. betle* leaf essential oils. A high intensity of the chavicol acetate (4) peak was consistently measured only in *P. betle* 'melayu' in both GC-MS analyses, differentiating it from *P. betle* unknown variant, 'india', and 'manis'. The BPCs also indicate that the latter three variants have similar leaf essential oil profiles, varying only in peak intensities. Chavibetol (5), chavibetol acetate (7), and 4-allyl-1,2-diacetoxybenzene (8) were found to be characteristic *O*-methylated phenylpropene and acetate esters of phenylpropenes in all *P. betle* leaf essential oils (Figure 3b,c).

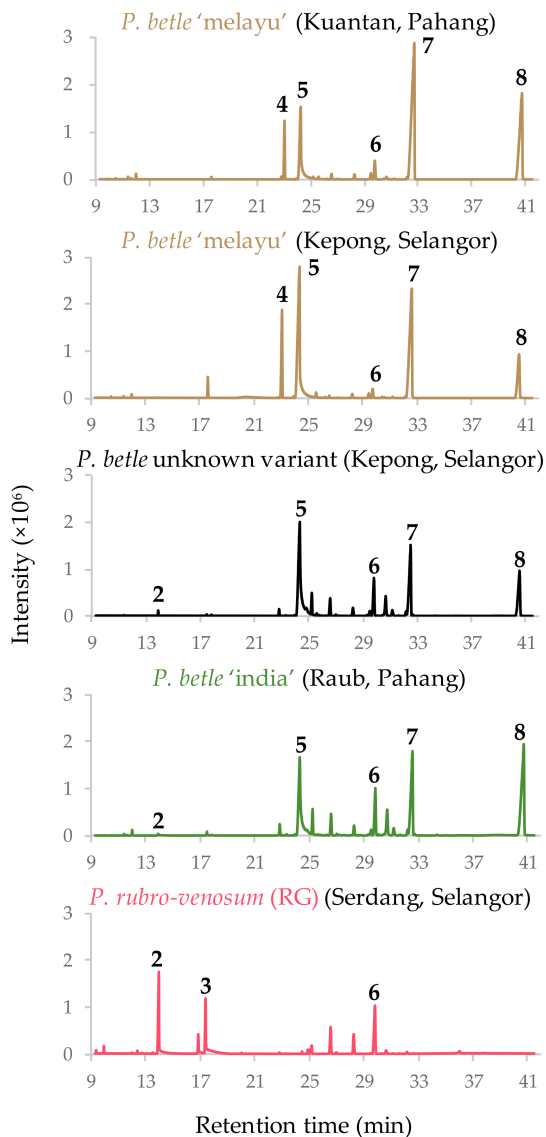
Despite having similar sesquiterpene hydrocarbon germacrene D (6) peak intensities, the *P. rubro-venosum* leaf essential oils could be clearly distinguished from *P. betle* by the absence of chavibetol (5), chavibetol acetate (7), and 4-allyl-1,2-diacetoxybenzene (8) peaks. Of the 68 identified constituents, GC-MS analysis IIa of *P. rubro-venosum* leaf essential oils revealed the presence of only one acetate ester, namely (*E*)-chrysanthenyl acetate (Table S4 and Figure S9). Very low intensities of the signals at  $\delta$  2.08 and  $\delta$  6.80–6.92 in the  $^1\text{H-NMR}$  spectrum of *P. rubro-venosum* aqueous methanolic leaf extracts in the pilot experiment (Figure 1b)

also reflected the lack/absence of the acetate esters of phenylpropenes in *P. rubro-venosum* leaves.

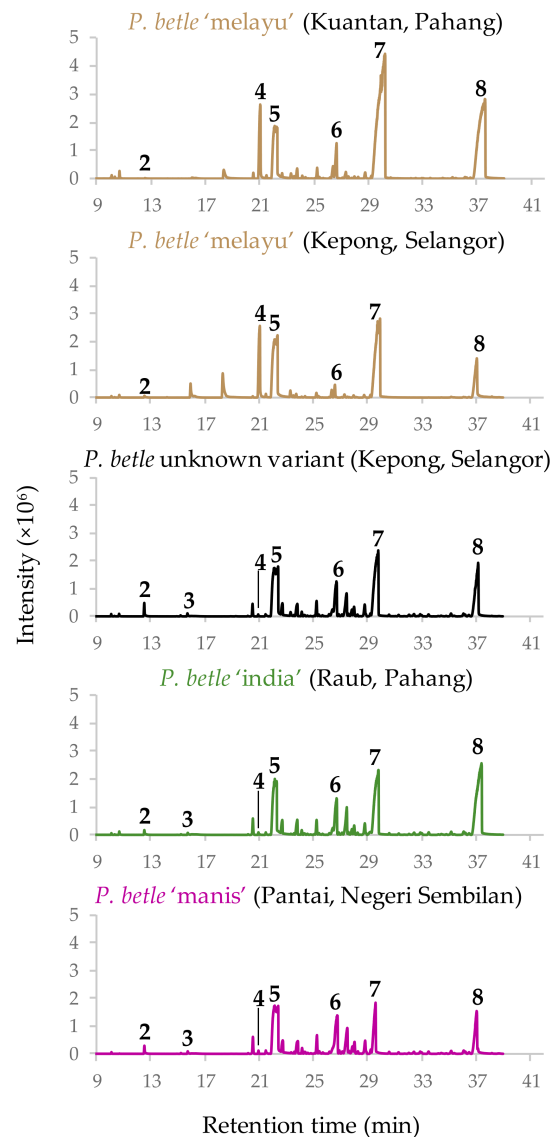
**a**



**b**



**c**



**Figure 3.** GC–MS analysis of *Piper betle* and *Piper rubro-venosum* leaf essential oils. (a) Photographs of *P. betle* 'melayu', unknown variant, 'india', and 'manis'. (b) Representative base peak chromatograms (BPC) of leaf essential oils in GC–MS analysis I and (c) GC–MS analysis II; (2) linalool, (3)  $\alpha$ -terpineol, (4) chavicol acetate, (5) chavibetol, (6) germacrene D, (7) chavibetol acetate, (8) 4-allyl-1,2-diacetoxybenzene.

In both GC-MS analyses, chavibetol (5) (LRI  $1374.0 \pm 1.4$ ), instead of eugenol, was identified as one of the major constituents of *P. betle* leaf essential oils. Eugenol (LRI  $1361.7 \pm 0.5$ ) and (Z)-isoeugenol ( $1391.8 \pm 0.5$ ), which were detected only in GC-MS analysis II, were identified as minor constituents. Another derivative of eugenol, methyl eugenol (LRI  $1410.3 \pm 0.5$ ), was detected as a minor constituent in both GC-MS analyses (Table S6 and Figure S11). Initially, for the peak with LRI  $1374.0 \pm 1.4$ , eugenol, 3-allyl-2-methoxyphenol, (Z)-2-methoxy-4-(prop-1-en-1-yl)phenol, and (Z)-2-methoxy-4-(prop-1-en-1-yl)-phenol were the top candidates in the list of hits with match scores ranging from 90 to 95. However, the LRIs of these constituents were found to deviate by more than  $\pm 5$  when compared to the retention indices reported in the main references [35,36]. Moreover, chavibetol or 5-allyl-2-methoxyphenol was not among the top candidates of the hit list. Further examination of other top candidates in MassHunter's list of hits located chavibetol, with a match score of 85 and LRI that matched literature data [46]. In the literature, either eugenol or its isomer chavibetol were reported as major constituents and chemical markers of *P. betle* leaf extracts [19,20,47–54]. Apparently, some studies that reported eugenol as the major constituent of *P. betle* leaf extracts based the identification solely on high match score or comparison with literature data without application of any other filtering method [12,19,49,54].

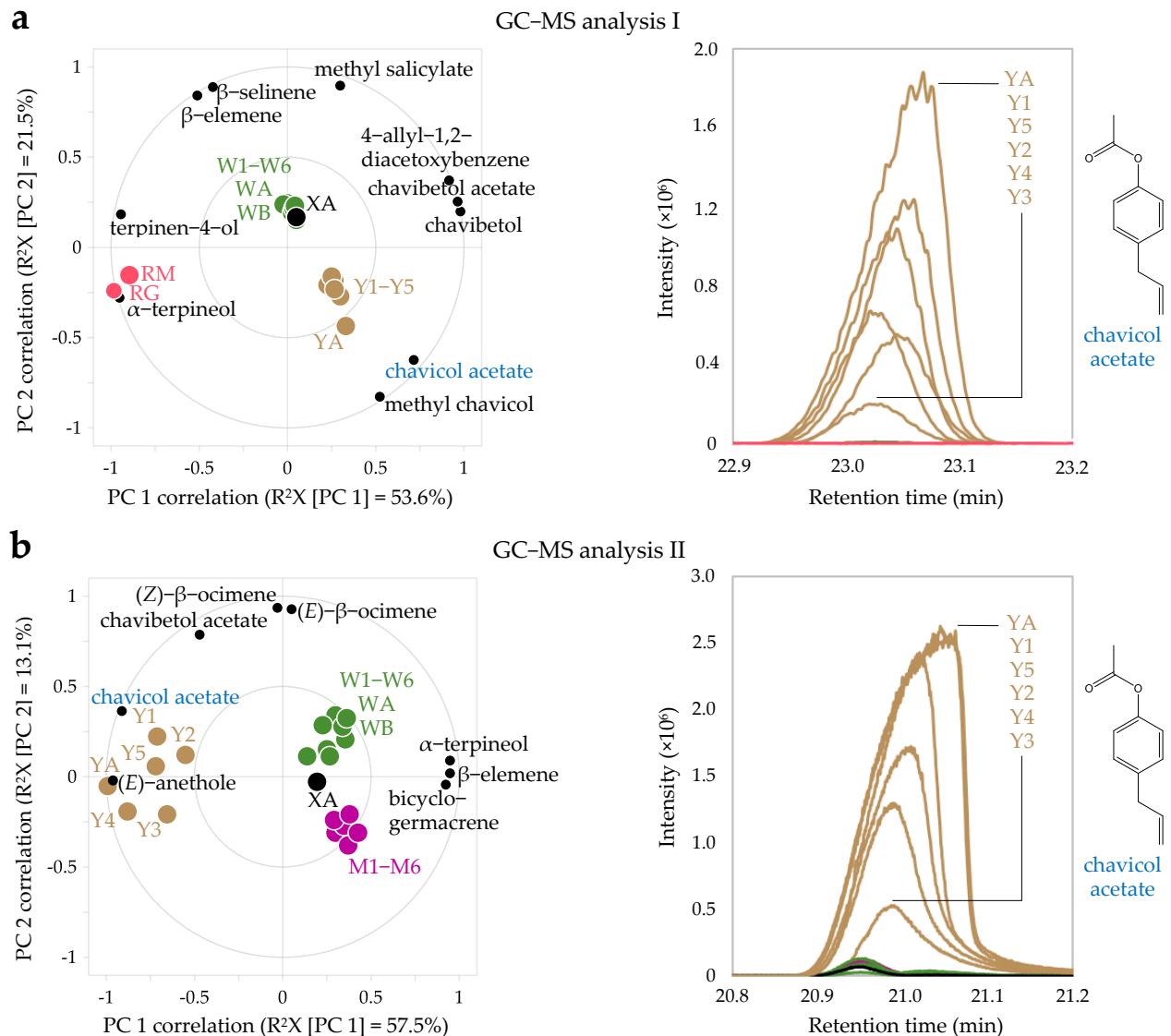
In the *Malaysian Herbal Monograph*, eugenol is listed as the chemical marker of *P. betle* leaf, and an HPLC method has been developed to confirm its presence in leaf extracts of the herb [7]. However, eugenol and its derivatives are also the chemical markers of several other spices, such as dried barks of *Cinnamomum verum* and *C. cassia* (Lauraceae), dried flower buds of *Syzygium aromaticum* (Myrtaceae), and leaves of *Ocimum sanctum* (Lamiaceae) [55–57], to which *P. betle* is distantly genetically related. This raises some doubts on the suitability of using eugenol as the chemical marker of *P. betle* leaf. Moreover, previous research that used both HPLC and GC-MS has shown that the difference in retention times between the eugenol and chavibetol standard is more apparent when analyzed using GC-MS (0.3 versus 0.6 min), and the retention time of the major peak in the GC-MS chromatogram of *P. betle* leaf extract also showed a better match with that of chavibetol standard, rather than eugenol, which appeared as a small peak preceding that of chavibetol [58]. Similarly, in our research, eugenol was shown to be a minor constituent of *P. betle* leaf essential oils, appearing as a small peak (peak number 15 in Figure S11) immediately before the much higher intensity peak of chavibetol in the GC-MS chromatogram, thus supporting our finding that chavibetol is the major constituent of *P. betle* leaf essential oils, while eugenol is a minor constituent. The same situation applies for the identification of chavibetol acetate (7) (LRI  $1532.5 \pm 0.7$ ), which was identified as another major constituent instead of eugenol acetate (see Figures S10 and S11 and Tables S5 and S6). Based on our findings, we propose that chavibetol (5) and chavibetol acetate (7) are more suitable for use in pharmacognostical analysis of *P. betle* leaves.

#### 2.4. Principal Components Analysis of GC-MS Data of *P. betle* Leaf Essential Oils

Despite our efforts to identify as many constituents as possible using two different GC-MS systems, several small peaks could not be detected or identified by the integrators in the GC-MS data processing software. Hence, for PCA of the GC-MS data, all BPCs from both GC-MS analyses were matricized to avoid any exclusion of undetected or unidentified peaks. This method is different from those used in previous studies, which performed PCA only on peak areas or peak intensities of constituents that were identified in *P. betle* samples [10,12,59]. Figure 4a shows that the *P. betle* of unknown variant (sample XA) is clustered together with *P. betle* 'india' (sample W1–W6, WA, and WB) in the PCA biplot of GC-MS analysis I data. The first PC shows that chavibetol (5), chavibetol acetate (7), and 4-allyl-1,2-diacetoxybenzene (8) are the discriminant factors for *P. betle* leaf essential oils, whereas  $\alpha$ -terpineol (3) and terpinen-4-ol are the discriminant factors for *P. rubrovenosum* leaf essential oils. Linalool (2) was not identified as a discriminant factor for *P.*



*rubro-venosum* as the intensity of the linalool peak was found to be high only in RG and not RM.



**Figure 4.** Principal component analysis (PCA) of matricized, logarithm (base 10)–transformed, and mean–centered base peak chromatograms (BPC) of *Piper betle* and *Piper rubro-venosum* leaf essential oils. *P. rubro-venosum* leaves with green underside (RG) and maroon underside (RM), *P. betle* ‘melayu’ (Y1–Y5 and YA), *P. betle* ‘india’ (W1–W6, WA, and WB), *P. betle* unknown variant (XA), and *P. betle* ‘manis’ (M1–M6). Only well–modeled variables are shown ( $Q^2VX$  [cumulative up to PC 2]  $\geq 80\%$ ). (a) Correlation–scaled PCA biplot of GC–MS analysis I data ( $Q^2X$  [cumulative up to PC 2] = 61.8%) and chavicol acetate peak in the BPC of GC–MS analysis I. (b) Correlation–scaled PCA biplot of GC–MS analysis II data ( $Q^2X$  [cumulative up to PC 2] = 63.4%) and chavicol acetate peak in the BPC of GC–MS analysis II. Considerations for selection of PCA models can be viewed in Figures S12 and S13.

The second PC of GC-MS analysis I data shows that the higher intensity of methyl salicylate is characteristic of *P. betle* ‘india’ and higher intensity of chavicol acetate (4) and methyl chavicol peak are characteristic of *P. betle* ‘melayu’ (samples Y1–Y5 and YA). However, in the PCA of GC-MS analysis II data (Figure 4b) only chavicol acetate (LRI  $1347.0 \pm 1.4$ ) was identified as the discriminant factor for *P. betle* ‘melayu’. In GC-MS analysis I, chavicol acetate was not detected in *P. betle* ‘india’ and unknown variant, but in GC-MS analysis II chavicol acetate was detected in low intensities in both variants and *P. betle* ‘manis’. These differences in detection capability between the analyses could be

attributed to the use of split and splitless injections in GC-MS analyses I and II, respectively, and higher sensitivity of the GC-MS system used in GC-MS analysis II. In the first PC of GC-MS analysis II data, the essential oil sample of *P. betle* 'melayu' leaves collected from Kepong, Selangor (sample YA) clustered together with the essential oil samples of *P. betle* 'melayu' leaves collected from Kuantan, Pahang (sample Y1–Y5), indicating the high loading of chavicol acetate in the first PC scores of all *P. betle* 'melayu' samples, regardless of their geographical origins. Consequently, other *P. betle* samples with low loadings of chavicol acetate clustered together, as in the case of *P. betle* 'india', unknown variant, and 'manis'. On the other hand, the second PC of GC-MS analysis II data could be related to the variation in intensities of the constituent peaks in the BPCs as illustrated in Figure 3c, where it can be observed that the intensities of constituent peaks in *P. betle* 'india' are generally higher than those in *P. betle* 'manis'. Considering the distribution of the second PC scores of all *P. betle* 'melayu' samples in the biplot of GC-MS analysis II data, it is reasonable to group the other samples (W1–W6, WA, WB, XA, and M1–M6) as *P. betle* 'india'.

### 2.5. Hypotheses

The PCA of  $^1\text{H-NMR}$  spectra and GC-MS data revealed a paucity of acetic acid and acetate esters of phenylpropenes in *P. rubro-venosum* leaves, respectively. Other than *P. rubro-venosum* leaves, analysis of  $^1\text{H-NMR}$  spectra of *P. sarmentosum* aqueous methanolic leaf extracts, which were morphologically similar to *P. betle* leaves, also showed the paucity of acetic acid and acetate esters of phenylpropenes [60]. In contrast, the abundance of acetic acid in aqueous methanolic leaf extracts and *O*-methylated phenylpropene chavibetol and acetate esters of phenylpropenes, namely chavicol acetate, chavibetol acetate, and 4-allyl-1,2-diacetoxybenzene, in the essential oils were found to be characteristic of *P. betle* leaves, regardless of the variants. Leaf essential oils of several species in the genus *Pimenta* (Myrtaceae), of which some are used as spices, were reported to contain chavibetol as major constituents. However, identification of acetate esters of phenylpropenes in the leaf essential oils of these species has never been reported in the literature [37,61–63]. Acetic acid is a precursor of acetyl-CoA, which is a metabolic intermediate responsible for the biosynthesis of acetylated metabolites and isoprenoids in plants [64,65]. Hence, we hypothesized that biosynthesis of acetate esters of phenylpropenes is associated with pooling of acetic acid, which is a part of specialized metabolism in *P. betle* leaves.

In general, consumption of *P. betle* 'india' is preferred over 'melayu' due to its less pungent flavor, which may be partly attributed to the remarkably lower level of chavicol acetate compared to *P. betle* 'melayu'. Identification of chavicol acetate as a chemical marker of a cluster of *P. betle* samples in PC scores space is parallel with the findings of previous research, which showed clustering of several Indian *P. betle* landraces into two major groups based on the intensity of the chavicol acetate peak [59]. Other than that, analyses using DNA markers of *P. betle* samples have revealed two major clusters based on gender [66,67]. Macroscopic examination for gender distinction of *P. betle* can be made only after flowering, based on morphology of the inflorescences (termed spikes) [10,66]. Male spikes are generally longer than female spikes [66]. As exemplified in our sampling, gender distinction of *P. betle* remains a challenging task because only *P. betle* 'melayu' collected from Kuantan, Pahang, and *P. betle* 'india' collected from Raub, Pahang bore the spikes (Figure S7). The spikes of *P. betle* 'melayu' and 'india' approximately matched the description of morphological characteristics of male and female spikes, respectively [66]. Both variants were resolved into two major groups in the first PC due to higher abundance of chavicol acetate in *P. betle* 'melayu' leaf essential oils. For these reasons, it was hypothesized that *P. betle* 'melayu' is the male plant whereas *P. betle* 'india' is the female plant, and chavicol acetate may be used as a chemical marker to differentiate between the two Malaysian *P. betle* variants and the two genders of this dioecious perianthless Piperalea plant.

### 3. Materials and Methods

#### 3.1. Chemicals and Disposables

Methanol-D4 (99.8% minimum deuteration degree), deuterium oxide (99.9% minimum deuteration degree) and 3-(trimethylsilyl)propionic-2,2,3,3-D4 acid sodium salt (98% minimum deuteration degree) purchased from Merck KGaA (Darmstadt, Germany) were used in the preparation of extracts in <sup>1</sup>H-NMR metabolomics. Liquid chromatography grade *n*-hexane ( $\geq 98.0\%$  purity) and *n*-nonane ( $99.3 \pm 0.1\%$  purity) were purchased from Fisher Scientific, Leicestershire, United Kingdom and CPChem, Stara Zagora, Bulgaria, respectively. Alkanes standard (C7–C33) was purchased from Restek Corporation, Bellefonte, PA, USA. PTFE syringe filters (13 mm, 0.22  $\mu$ m) were purchased from Orioner Hightech, Kuala Lumpur, Malaysia.

#### 3.2. <sup>1</sup>H-NMR Metabolomics of “Sirih” Leaf Extracts

##### 3.2.1. Collection of Plant Materials

In the present research, the term ‘variant’ is used instead of landraces to indicate that the names are not based on the specific geographical region where the plants were grown. Firstly, fresh leaves with intact petioles of three local variants of “sirih”, known as “sirih merah”, “sirih india”, and “sirih melayu”, were collected on 5 July 2018 from the herbal garden within the University Agriculture Park, Universiti Putra Malaysia, Serdang, Selangor, Malaysia. Coordinates of sampling locations are listed in Table 1. For each variant, a total of six leaves were collected from a plant to represent six replicates. The six leaves collected were similar in size and color and grew at the 11th or 12th node of different stems [12]. Collection of the leaves was carried out between 10:15 and 10:55 a.m. at a temperature of 35 °C and relative air humidity of 76%. Voucher specimens were deposited in the herbarium of Universiti Kebangsaan Malaysia (UKMB).

**Table 1.** Coordinates of sampling locations of the “sirih” variants.

Sample	Voucher Specimen No.	Coordinates	Altitude (m)
“sirih merah”	UKMB404018	2°59′21.4″ N, 101°42′30.3″ E	50
“sirih india”	UKMB404019	2°59′19.4″ N, 101°42′30.0″ E	51
“sirih melayu”	UKMB404020	2°59′19.6″ N, 101°42′29.7″ E	49

##### 3.2.2. Preparation of Samples and Extracts

Methods of sample preparation and extraction were adapted from the protocol of plant NMR metabolomics [68]. Immediately after harvesting, the leaves were wiped clean with dry tissue paper and the petioles were removed manually. The leaf blades (without midvein) were then quenched in liquid nitrogen and ground into powder form using a mortar and pestle. The mortar and pestle were wiped clean between samples to minimize cross contamination. The powdered samples were then transferred into 2 mL centrifuge tubes, frozen at  $-80$  °C for 24 h and lyophilized for 24 h. The lyophilized samples were kept in a desiccator, away from light, prior to extraction.

Thirty (30) mg of lyophilized sample was extracted with 1.2 mL deuterated solvent made up of 600  $\mu$ L methanol-D4 and 600  $\mu$ L phosphate buffer (90 mM, pH 6.0) in deuterium oxide containing 0.1% TSP as internal standard, using a vortex mixer for 1 min. Metabolites in the sample were further extracted by sonication for 15 min at 35 kHz and room temperature in an ultrasonicator (LC 60/H, Elma Schmidbauer GmbH, Singen, Germany) followed by centrifugation at 13,000 rpm for 15 min (Hettich® MIKRO 20, Andreas Hettich GmbH & Co. KG, Tuttlingen, Germany). The supernatant was filtered through a 0.22  $\mu$ m PTFE syringe filter and 600  $\mu$ L was transferred into 5 mm NMR tubes [69].

### 3.2.3. Acquisition of <sup>1</sup>H-NMR Spectra

Firstly, all extracts were left at room temperature for 30 min. Next, NMR spectral data were acquired at 24 °C on an Agilent VNMR500 spectrometer system equipped with a narrow bore premium shielded magnet operating at a frequency of 500 MHz, with a 5 mm inverse detection pulsed field gradient (ID/PFG) probe, and VnmrJ 6.1C software (Agilent Technologies, Santa Clara, CA, USA). Acquisition parameters are described herein: solvent lock = deuterium oxide, spectral width = −2.00–14.00 ppm, number of scans = 8, relaxation delay = 1 s, pulse angle = 45° and receiver gain = 32 dB. Presaturation pulse sequence (PRESAT) was utilized to suppress the residual water signal with the following parameters: number of scans = 64 and presaturation delay = 2 s. The FID was transformed to NMR spectrum using Fourier transform with weighting function to enhance the resolution (command wft), followed by automatic phasing of both zero- and first-order terms (command aph). A total of 18 extracts were analyzed.

### 3.2.4. Preprocessing of <sup>1</sup>H-NMR Spectra

NMR spectra of all samples were superimposed to identify meaningful regions ( $\delta$  0.50–9.70) and residual solvent signals (methanol:  $\delta$  3.27–3.33 and water:  $\delta$  4.70–5.07) using MestReNova version 6.0.2-5475 (Mestrelab Research S.L., Santiago de Compostela, Spain). The NMR spectra were further processed using Chenomx Processor version 8.2 (Chenomx, Edmonton, AB, Canada) with the following parameters: chemical shape indicator (CSI) = TSP, TSP concentration = 0.2902 mM, sample pH = 6.0  $\pm$  0.5, and automatic phase, baseline (cubic spline), and shim correction. Identification and quantification of metabolites were carried out by fitting the NMR spectra to reference spectra in Chenomx Spectral Reference Library in Chenomx Profiler version 8.2. Identity and estimated concentration of a metabolite was considered valid only in case of all NMR signals of the metabolite fitting perfectly to all signals of the metabolite in the reference library and no overlapping with signals of other metabolites. The NMR spectra were normalized to total area and the residual solvent signals were removed. Next, the meaningful regions were discretized to buckets of  $\delta$  0.04 width, producing a data matrix dimension of 18  $\times$  221 in tsv file format.

### 3.3. DNA Barcoding of “Sirih Merah”

The sequence of ITS2 region was selected for the DNA barcoding analysis due to its length, which is shorter than that of full-length ITS (average length = 233 versus 634 base pairs), thus less prone to amplification and sequencing errors for rapid identification of plant species [70]. DNA barcoding was performed following the protocol described in previous research [30]. Briefly, whole cell DNA was extracted twice using the conventional cetyltrimethylammonium bromide (CTAB) procedure and purified using a High Pure PCR template preparation kit (Roche Molecular Systems, Inc., Pleasanton, CA, USA). DNA barcodes were generated using the primer pairs ITS2-2F and ITS2-3R [71] and amplified in 10  $\mu$ L reaction mixture, comprising 0.3  $\mu$ M of each primer, 10 ng of template DNA, and one Type-it Multiplex PCR master mix (Qiagen, Germantown, MD, USA) using GeneAmp PCR System 9700 (Thermo Fisher Scientific, Waltham, MA, USA). Next, the PCR products were bidirectionally sequenced using a BigDye Terminator v3.1 cycle sequencing kit (Thermo Fisher Scientific, Waltham, MA, USA) based on the Sanger chain-termination method, purified using a BigDye XTerminator purification kit (Thermo Fisher Scientific, Waltham, MA, USA) and analyzed using ABI3130xl Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). ITS2 sequence data was edited using Sequencher version 5.0 (Gene Codes Corporation, Ann Arbor, MI, USA) and matched against reference sequences in the NCBI GenBank database using Basic Local Alignment Search Tool (BLAST) and a neighbor-joining (NJ) tree. The NJ tree was constructed based on Kimura two-parameter (K2P) divergences and 1000 bootstrap replications in MEGA6 software [72].

### 3.4. GC-MS Metabolomics of Leaf Essential Oils

#### 3.4.1. Collection of Plant Materials

Our efforts to collect samples of the three *P. betle* variants described in the transliterated manuscript of traditional Malay medicine [8] were only partly successful. From the three variants, we managed to obtain only *P. betle* 'melayu' after communications with several Malaysian locals and *P. betle* farmers. Apart from *P. betle* 'melayu', we collected samples locally known as *P. betle* 'india', 'manis', and an unknown variant that morphologically resembled *P. betle* 'india'. Mature leaves of similar texture and color with intact petioles of *P. betle* 'melayu' (sample code Y1–Y5 and YA), 'india' (W1–W5, WA and WB), 'manis' (M1–M6), unknown variant (XA), and *P. rubro-venosum* (RM and RG) were collected from six different locations, comprising three states in Peninsular Malaysia (Figure S6). Details on collection date and time, sampling locations, growing conditions, and occurrence of inflorescence are listed in Table S1. Following the local practice of maintaining the freshness of harvested *P. betle* leaves, immediately after harvesting, the leaves were transported to the laboratory. The leaves were stored in plastic bags at room temperature, lightly sprayed with tap water daily, and processed within five days after harvesting.

#### 3.4.2. Hydrodistillation of Plant Materials

Prior to hydrodistillation, the leaves were examined for undesired physical attributes, which might be inflicted upon collection, transportation, or storage. Leaves of unsatisfactory quality were excluded (Figure S5). Next, the petioles of the leaves were removed, and the leaf blades were cut in slices using clean scissors. Depending on the number of good quality leaves, varying amounts of sliced leaf blades were loaded into a 10 L round-bottomed flask filled with 7 L tap water (Table S2). Hydrodistillation was carried out for 2 h, where the starting time ( $t_0$ ) was the moment the heater of the apparatus being switched on [73]. Two samples were hydrodistilled per day, with thorough washing of the Clevenger-type apparatus between the samples. Subsequently, 10 mL HPLC grade *n*-hexane was added to the distillate in a separatory funnel. The *n*-hexane layer was collected after 18 h of partitioning, followed by evaporation of the *n*-hexane in a fume hood. Next, the essential oils obtained were weighed and stored in tightly sealed containers at 4 °C until further analysis.

#### 3.4.3. GC-MS Analyses

GC-MS analyses were performed using two gas chromatography systems manufactured by different manufacturers, two columns (dimethylsilicone stationary phase with 5% phenyl groups) by two manufacturers, and two sets of electron ionization mass spectral libraries. The analytical parameters for analysis I were an improvisation of standard procedure for analysis of essential oils in our laboratory while the parameters for analysis II were optimized from the findings of GC-MS analysis I.

##### GC-MS Analysis I

The gas chromatography system used for GC-MS I was GCMS-QP2010 Ultra (Shimadzu, Kyoto, Japan). The mass spectrometer was operated in electron ionization mode at 70 eV, ion source temperature = 200 °C, scanning mode = full scan, scan rate = 10 scans/s, mass range = 40–700 Da, and solvent delay = 2.55 min. The inlet temperature was set at 250 °C. Firstly, 10 µL of the essential oil was diluted in 2 mL LC grade *n*-hexane. One (1) µL of the diluted essential oil was injected in split mode (10:1) into a Rxi-5ms column of 30.00 m length × 0.25 mm internal diameter × 0.25 µm film thickness (Restek Corporation, Bellefonte, PA, USA). Helium flow rate was constant at 0.80 mL/min.

The oven temperature was programmed as the following: 50 °C (1.0 min), 4 °C/min to 130 °C (0.0 min), 1 °C/min to 170 °C (0.0 min), 5 °C/min to 240 °C (5.0 min), making a total run time of 80 min. The order of run sequence was not randomized, and solvent blank (*n*-hexane) was injected at the beginning of the run sequence and after every injection of sample to ensure no carryover from the previous sample. Alkanes standard was injected

after the first solvent blank injection for calculation of linear retention indices (LRI) using the H. Van den Dool and P. Dec Kratz equation [74].

For the identification of essential oil constituents, EI mass spectra were matched against reference spectra in Wiley Registry 9, National Institute of Standards and Technology (NIST) 2011 and Wiley's Flavors and Fragrances of Natural and Synthetic Compounds (FFNSC) 1.3 EI mass spectral libraries, and the threshold of similarity index (SI) between measured and reference spectra was set at 80. A hit list of top 10 best candidates for each constituent was generated in pdf file format using GCMSsolution 2.72 software (Shimadzu, Kyoto, Japan). For both GC-MS analyses, the difference between calculated LRI and literature data was considered to minimize false positives. Hence, only constituents with a LRI within  $\pm 5$  deviation from the LRI of literature data were reported [75]. The main references for the identification of essential oil constituents were compilations of LRI of frequently reported essential oil constituents [35,36] and only in case where the LRI did not match the LRI reported in the main references were other studies referred to [37–48].

#### GC-MS Analysis II

The gas chromatograph used for GC-MS II was the 7890A Gas Chromatograph equipped with 7000 Series Triple Quadrupole Mass Spectrometer (Agilent Technologies, Santa Clara, CA, USA). The mass spectrometer was operated in EI mode at 70 eV and at ion source temperature of 200 °C. The following parameters were set for data acquisition: mode = full scan, scan rate = 20 scans/s, mass range = 40–250 Da and solvent delay = 5 min. The inlet temperature was set at 250 °C. An aliquot of 2.5  $\mu$ L essential oil was diluted in 1 mL LC grade *n*-hexane containing 5 ppm *n*-nonane as internal standard. A pooled quality control (QC) sample was prepared by mixing equal aliquots of each sample. One (1)  $\mu$ L of the diluted essential oil was injected in splitless mode into a HP-5MS column of 30.00 m length  $\times$  0.25 mm internal diameter  $\times$  0.25  $\mu$ m film thickness (Agilent Technologies, Santa Clara, CA, USA). Helium flow rate was maintained at 0.80 mL/min.

The oven temperature was programmed as the following: 50 °C (1.0 min), 4 °C/min to 126 °C (0.0 min), 1 °C/min to 148 °C (0.0 min), 100 °C/min to 250 °C (0.0 min), making a total run time of 43.02 min. Post-run column temperature was set at 310 °C for a period of 1 min, to elute out remaining sample prior to injection of the next sample. A solvent blank (*n*-hexane containing 5 ppm *n*-nonane) was injected at the beginning of the run sequence. The alkanes standard was injected after the first solvent blank injection and the end of the run sequence for calculation of LRI and examination of retention time shift [76]. The order of run sequence was randomized manually by drawing cards of sample codes from a bowl. After injection of four to six samples, one QC sample was injected throughout the run sequence.

Identification of essential oil constituents was carried out using MassHunter Workstation Qualitative Analysis Workflows B.08.00 software (Agilent Technologies, Santa Clara, CA, USA). The following parameters were set: identification algorithm = peak integration, integrator = Agile 2 and minimum integrated area = 0.05% of largest peak in the total ion current (TIC) chromatogram. Mass spectra of constituents were matched against 15 best candidates in NIST 2014 and Wiley's FFNSC 3.0 EI mass spectral libraries with minimum reverse similarity index (RSI) threshold set at 80. Only constituents with a LRI within  $\pm 5$  deviation from the LRI of literature data were reported [75].

#### 3.4.4. Preprocessing of GC-MS Data

Raw data in Shimadzu QGD (GC-MS analysis I) and Agilent D (GC-MS analysis II) file format were converted to netCDF using GCMSsolution 2.72 (Shimadzu, Kyoto, Japan) and GC/MS Translator (Agilent Technologies, Santa Clara, CA, USA), respectively. The netCDF files were exported to MZmine version 2.53 software for visualization of chromatograms [77]. Two- and three-dimensional base peak chromatograms (BPC) were visualized to aid the input of parameters in the MS-DIAL version 4.24 software [78].

Trial and error was used to optimize the parameters in the MS-DIAL. The optimized parameters for peak detection, deconvolution, alignment, and filtering, including parameters selected for results export, are listed in Table S7. Aligned peak lists for both GC-MS analyses were normalized to TIC, producing data matrices in txt file format. Subsequently, the data matrices were exported to Microsoft Office Excel 2016. Next, preparation of the data matrices (exclusion of variables, transposing, and labeling) for principal component analysis (PCA) was carried out using Microsoft Office Excel 2016. All variables in the data matrix of GC-MS I were retained while only variables with less than 30% relative standard deviation (<30% RSD) in all four QC samples of GC-MS II were retained [79].

### 3.5. Welch's Analysis of Variance (ANOVA) and Principal Components Analysis (PCA)

Welch's ANOVA with Games–Howell post hoc test was performed using SPSS version 26 software (International Business Machines Corporation, New York, NY, USA). PCA was carried out using SIMCA-P 14.1 software (Sartorius Stedim Data Analytics AB, Umeå, Sweden). Firstly, normal distribution of variables was assessed based on skewness and ratio of minimum–maximum values of the variables. In case of skewed distribution of most variables in a data matrix (as in the GC-MS data matrices), the whole data matrix was log-transformed [80]. Variables that could not be log-transformed were treated as missing data, thus giving no influence on the final PCA model. Unit variance (UV) and Pareto scaling and mean centering were performed on the data matrices to find an operation that allowed the capture of the highest percentage of variation explained by the first two principal components (highest  $R^2X$  [cumulative up to PC 2]) [81]. Leave-one-out cross validation procedure was used for the  $^1\text{H-NMR}$  and GC-MS analysis I data matrices whereas SIMCA-P's default seven-groups cross validation procedure was used for the GC-MS analysis II data matrix [82,83]. To facilitate interpretation of PCA loadings, only well-modeled variables were made visible in loading plots and biplots. Well-modeled variables were defined as the variables that were well-predicted by the first two principal components, as estimated by cross-validation ( $Q^2VX$  [cumulative up to PC 2]  $\geq 80\%$ ).

## 4. Conclusions

Utilization of 1D  $^1\text{H-NMR}$  metabolomics at the initial stage of the present research resolved the ambiguity in identification of Malaysian *P. betle* variants, where the significantly lower concentration of acetic acid and intensity of  $^1\text{H-NMR}$  signals for acetate esters of phenylpropenes in a group of samples necessitated its verification using DNA barcoding of ITS2 sequences. The group was identified as *P. rubro-venosum*, which originates from Indonesia and is considered as an introduced species in Malaysia. Results from the principal component analysis of the  $^1\text{H-NMR}$  spectral data directed further detailed analysis using GC-MS metabolomics. Using two 3D datasets, GC-MS metabolomics confirmed the absence of acetate esters of phenylpropenes in *P. rubro-venosum* leaf essential oils and revealed the occurrence of two *P. betle* variants, namely *P. betle* 'melayu' and 'india', which could be distinguished based on the abundance of chavicol acetate in their leaf essential oils. The abundance of acetic acid and acetate esters of phenylpropenes in the leaves were hypothesized to be a part of a specialized metabolism in *P. betle*, and that *P. betle* 'melayu' and 'india' are the male and female *P. betle*, respectively. Nevertheless, more detailed studies involving larger sample sizes and wider coverage of geographical distribution, as well as controlled experimentation, shall be carried out to test the hypotheses and provide better insights into the essential oil biochemistry of this medicinal plant.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/plants10112510/s1>, Figure S1: Full representative  $^1\text{H-NMR}$  spectra of three "sirih", Figure S2: PCA score and loading plots of preprocessed  $^1\text{H-NMR}$  spectral data, Figure S3: Correlation-scaled PCA biplot for "sirih merah", Figure S4: Excerpt from DNA barcoding report on identification of "sirih merah" as *P. rubro-venosum*, Figure S5: Leaf samples included in or excluded from GC-MS metabolomics, Figure S6: Six sampling locations of *P. betle* variants and *P. rubro-venosum*, Figure S7: Photographs of *P. betle* inflorescences (spikes), Figure S8: BPC of *P. rubro-venosum* leaf essential oils

in GC-MS analysis I, Figure S9: BPC of *P. rubro-venosum* leaf essential oils in GC-MS analysis IIa, Figure S10: BPC of *P. betle* leaf essential oils in GC-MS analysis I, Figure S11: BPC of *P. betle* leaf essential oils in GC-MS analysis II, Figure S12: Score plots when different reference files were used in MS-DIAL (GC-MS analysis I), Figure S13: Score plots when different reference files were used in MS-DIAL (GC-MS analysis II), Table S1: Description of *P. betle* and *P. rubro-venosum* studied in GC-MS metabolomics, Table S2: Yield percentage of *Piper* leaf essential oils, Table S3: Constituents of *P. rubro-venosum* leaf essential oils identified in GC-MS analysis I, Table S4: Constituents of *P. rubro-venosum* leaf essential oils identified in GC-MS analysis IIa, Table S5: Constituents of *P. betle* leaf essential oils identified in GC-MS analysis I, Table S6: Constituents of *P. betle* leaf essential oils identified in GC-MS analysis II, Table S7: Parameter settings in MS-DIAL version 4.24.

**Author Contributions:** Conceptualization, K.S. and S.Y.L.; methodology, K.S., S.Y.L. and M.F.O.; validation, S.K. and B.H.Z.; formal analysis and data curation, M.F.O. and B.H.Z.; writing—original draft preparation, M.F.O.; writing—review and editing, K.S., S.Y.L., S.R.S., S.M.M.F., S.K. and B.H.Z.; visualization, M.F.O.; funding acquisition, K.S. and S.R.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Universiti Putra Malaysia, grant number GP-IPS/2018/9664600. M.F.O. is an academic trainee at the Department of Pharmaceutical Chemistry, Kulliyah of Pharmacy, International Islamic University Malaysia, and a Ph.D. candidate at Universiti Putra Malaysia, funded by the Ministry of Higher Education Malaysia via Skim Latihan Akademik Bumiputera (SLAB) and International Islamic University Malaysia via IIUM Postgraduate Fellowship Scheme.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Metabolomics data have been deposited in the EMBL-EBI Metabo-Lights database with identifier number MTBLS2585. ITS2 sequence of “sirih merah” was deposited in the NCBI GenBank database with accession number OK422240.

**Acknowledgments:** The authors would like to thank Rani Asmarayani (Herbarium Bogoriense (BO), Research Center for Biology, Indonesian Institute of Sciences) for her research on phylogeny of Malesian-Pacific *Piper* [4] and discussion on the identity of “sirih merah”. We also would like to thank Tnah Lee Hong, Kevin Ng Kit Siong, and Lee Soon Leong from Genetics Laboratory Forest Research Institute Malaysia (FRIM) who helped us in DNA barcoding of “sirih merah”.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the research; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Patra, A.P.; Mukherjee, A.K.; Acharya, L. Comparative study of RAPD and ISSR markers to assess the genetic diversity of betel vine (*Piper betle* L.) in Orissa, India. *Am. J. Biochem. Mol. Biol.* **2011**, *1*, 200–211. [[CrossRef](#)]
2. Ong, H.C.; Nordiana, M. Malay ethno-medico botany in Machang, Kelantan, Malaysia. *Fitoterapia* **1999**, *70*, 502–513. [[CrossRef](#)]
3. Jaramillo, M.A.; Callejas, R.; Davidson, C.; Smith, J.F.; Stevens, A.C.; Tepe, E.J. A phylogeny of the tropical genus *Piper* using ITS and the chloroplast intron psbJ-petA. *Syst. Bot.* **2008**, *33*, 647–660. [[CrossRef](#)]
4. Asmarayani, R. Phylogenetic relationships in Malesian-Pacific *Piper* (Piperaceae) and their implications for systematics. *Taxon* **2018**, *67*, 693–724. [[CrossRef](#)]
5. van Welzen, P.C.; Parnell, J.A.N.; Slik, J.W.F. Wallace’s line and plant distributions: Two or three phytogeographical areas and where to group Java? *Biol. J. Linn. Soc.* **2011**, *103*, 531–545. [[CrossRef](#)]
6. Adam, N.S.; Abd Rahman, R.; Abu Bakar, A.F.; Zainal, N.F.; Safie, S.Z.; Mansor, M. Distribution of herbs cultivation in Peninsular Malaysia. In Proceedings of the Prosiding Persidangan Industri Herba (Proceedings of Herbal Industry Conference), Putrajaya, Malaysia, 3–5 November 2015; Osman, A., Lim, H.F., Eds.; Forest Research Institute Malaysia: Kepong, Malaysia, 2015; pp. 318–323.
7. Malaysian Herbal Monograph Committee. *Piper betle* L. (Leaves). In *Malaysian Herbal Monograph*; Institute for Medical Research: Kuala Lumpur, Malaysia, 2015; pp. 416–427.
8. Mat Piah, H.; Mustapha, N.M. Glosari II: Istilah botani dan saintifik. In *Kitab Tib MSS 1292 PNM edisi dan Suntingan Teks*; Forest Research Institute Malaysia (FRIM): Kepong, Malaysia, 2019; p. 242.
9. Casañas, F.; Simó, J.; Casals, J.; Prohens, J. Toward an evolved concept of landrace. *Front. Plant Sci.* **2017**, *8*, 1–7. [[CrossRef](#)] [[PubMed](#)]
10. Bajpai, V.; Pandey, R.; Negi, M.P.S.; Bindu, K.; Kumar, N.; Kumar, B. Characteristic differences in metabolite profile in male and female plants of dioecious *Piper betle* L. *J. Biosci.* **2012**, *37*, 1061–1066. [[CrossRef](#)] [[PubMed](#)]



11. Khan, A.A.; Bhatnagar, S.P.; Sinha, B.N.; Lal, U.R. Pharmacognostic specifications of eight cultivars of *Piper betle* from eastern region of India. *Pharmacogn. J.* **2013**, *5*, 176–183. [[CrossRef](#)]
12. Karak, S.; Bhattacharya, P.; Nandy, A.; Saha, A.; De, B. Metabolite profiling and chemometric study for varietal difference in *Piper betle* L. leaf. *Curr. Metab.* **2016**, *4*, 129–140. [[CrossRef](#)]
13. Islam, M.A.; Ryu, K.Y.; Khan, N.; Song, O.Y.; Jeong, J.Y.; Son, J.H.; Jamila, N.; Kim, K.S. Determination of the volatile compounds in five varieties of *Piper betle* L. from Bangladesh using simultaneous distillation extraction and gas chromatography/mass spectrometry (SDE-GC/MS). *Anal. Lett.* **2020**, *53*, 2413–2430. [[CrossRef](#)]
14. Guha, P.; Nandi, S. Essential oil of betel leaf (*Piper betle* L.): A novel addition to the world food sector. In *Essential Oil Research Trends in Biosynthesis, Analytics, Industrial Applications and Biotechnological Production*; Malik, S., Ed.; Springer Nature Switzerland AG: Cham, Switzerland, 2019; pp. 149–196. [[CrossRef](#)]
15. Bro, R.; Smilde, A.K. Centering and scaling in component analysis. *J. Chemom.* **2003**, *17*, 16–33. [[CrossRef](#)]
16. Hartini, Y.S.; Nugroho, L. The accumulation of two neolignan in the leaves, stems, and flower of red betel (*Piper crocatum* Ruiz & Pav.). *J. Phys. Conf. Ser.* **2017**, *835*, 2–6. [[CrossRef](#)]
17. Li, H.X.; Widowati, W.; Azis, R.; Yang, S.Y.; Kim, Y.H.; Li, W. Chemical constituents of the *Piper crocatum* leaves and their chemotaxonomic significance. *Biochem. Syst. Ecol.* **2019**, *86*, 103905. [[CrossRef](#)]
18. Pandey, R.; Chandra, P.; Srivastva, M.; Arya, K.R.; Shukla, P.K.; Kumar, B. A rapid analytical method for characterization and simultaneous quantitative determination of phytoconstituents in *Piper betle* landraces using UPLC-ESI-MS/MS. *Anal. Methods* **2014**, *6*, 7349–7360. [[CrossRef](#)]
19. Kemprai, P.; Bora, P.K.; Mahanta, B.P.; Sut, D.; Saikia, S.P.; Banik, D.; Haldar, S. *Piper betleoides* C. DC.: Edible source of betel-scented sesquiterpene-rich essential oil. *Flavour Fragr. J.* **2020**, *35*, 70–78. [[CrossRef](#)]
20. Kemprai, P.; Protim Mahanta, B.; Kumar Bora, P.; Jyoti Das, D.; Lakshmi Hati Boruah, J.; Proteem Saikia, S.; Haldar, S. A <sup>1</sup>H NMR spectroscopic method for the quantification of propenylbenzenes in the essential oils: Evaluation of key odorants, antioxidants and post-harvest drying techniques for *Piper betle* L. *Food Chem.* **2020**, *331*, 127278. [[CrossRef](#)] [[PubMed](#)]
21. Chai, Y.J.; Go, Y.; Zhou, H.Q.; Li, H.X.; Lee, S.J.; Park, Y.J.; Widawatib, W.; Rizal, R.; Kim, Y.H.; Yang, S.Y.; et al. Unusual bicyclo[3.2.1]octanoid neolignans from leaves of *Piper crocatum* and their effect on pyruvate dehydrogenase activity. *Plants* **2021**, *10*, 1855. [[CrossRef](#)] [[PubMed](#)]
22. Chaveerach, A.; Mookkamul, P.; Sudmoon, R.; Tanee, T. Ethnobotany of the genus *Piper* (Piperaceae) in Thailand. *Ethnobot. Res. Appl.* **2006**, *4*, 223–231. [[CrossRef](#)]
23. Chaveerach, A.; Mookkamul, P.; Sudmoon, R.; Tanee, T. Two new species of *Piper* (Piperaceae) from Malay Peninsula. *Taiwania* **2007**, *52*, 210–215. [[CrossRef](#)]
24. Li, H.X.; Yang, S.Y.; Kim, Y.H.; Li, W. Isolation of two new compounds and other constituents from leaves of *Piper crocatum* and study of their soluble epoxide hydrolase activities. *Molecules* **2019**, *24*, 489. [[CrossRef](#)]
25. Rajudin, E.; Fernando, A.; Yulindari, R.; Rullah, K.; Indrayani, N.R.; Susanty, A.; Yerti, R.; Ahmad, F.; Mohd Sirat, H.; Arbain, D. Cytotoxic activities of fractions and two isolated compounds from sirih merah (Indonesian red betel), *Piper crocatum* Ruiz & Pav. *Procedia Chem.* **2014**, *13*, 79–84. [[CrossRef](#)]
26. Sanubol, A.; Chaveerach, A.; Sudmoon, R.; Tanee, T.; Noikotr, K.; Chuachan, C. Betel-like-scented *Piper* plants as diverse sources of industrial and medicinal aromatic chemicals. *Chiang Mai J. Sci.* **2014**, *41*, 1171–1181. Available online: <https://epg.science.cmu.ac.th/ejournal/journal-detail.php?id=5278> (accessed on 3 August 2021).
27. Sanubol, A.; Chaveerach, A.; Tanee, T.; Sudmoon, R. Pre-clinical evaluation of extracts and essential oils from betel-like scent *Piper* species identified potential cancer treatment. *Afr. J. Tradit. Complement. Altern. Med.* **2017**, *14*, 89–102. [[CrossRef](#)]
28. Vertical Garden Patrick Blanc: Leaf Colors and Structures 2. Available online: <https://www.verticalgardenpatrickblanc.com/inspiration/leaf-colors-and-structures-2?page=3> (accessed on 17 February 2021).
29. Sulianti, S.B.; Chairul, D. Perbandingan komponen kimia penyusun minyak atsiri sirih liar (*Piper ornatum*) yang berasal dari Sulawesi Selatan dan Pulau Seram dengan sirih biasa (*Piper betle*). *Ber. Biologi* **2002**, *6*, 493–499. [[CrossRef](#)]
30. Tnah, L.H.; Lee, S.L.; Tan, A.L.; Lee, C.T.; Ng, K.K.S.; Ng, C.H.; Zakaria, N.F. DNA barcode database of common herbal plants in the tropics: A resource for herbal product authentication. *Food Control* **2019**, *95*, 318–326. [[CrossRef](#)]
31. Turner, I.M. A catalogue of the vascular plants of Malaya (part 2 of 2). *Gard. Bull. Singapore* **1995**, *47*, 399–403. Available online: <https://www.biodiversitylibrary.org/page/43660944#page/67/mode/1up> (accessed on 25 April 2019).
32. Matyushin, D.D.; Sholokhova, A.Y.; Karnaeva, A.E.; Buryak, A.K. Various aspects of retention index usage for GC-MS library search: A statistical investigation using a diverse data set. *Chemometr. Intell. Lab. Syst.* **2020**, *202*, 104042. [[CrossRef](#)]
33. Samokhin, A.; Sotnezova, K.; Lashin, V.; Revelsky, I. Evaluation of mass spectral library search algorithms implemented in commercial software. *J. Mass Spectrom.* **2015**, *50*, 820–825. [[CrossRef](#)] [[PubMed](#)]
34. Babushok, V.I.; Andriamaharavo, N.R. Use of large retention index database for filtering of GC-MS false positive identifications of compounds. *Chromatographia* **2012**, *75*, 685–692. [[CrossRef](#)]
35. Babushok, V.I.; Linstrom, P.J.; Zenkevich, I.G. Retention indices for frequently reported compounds of plant essential oils. *J. Phys. Chem. Ref. Data* **2011**, *40*, 043101. [[CrossRef](#)]
36. Adams, R.P. *Identification of Essential Oil Components by Gas Chromatography/Mass Spectrometry*, 4th ed.; Allured Publishing: Carol Stream, IL, USA, 2017.

37. García, D.; Alvarez, A.; Tornos, P.; Fernandez, A.; Sáenz, T. Gas chromatographic-mass spectrometry study of the essential oils of *Pimenta racemosa* var. *terebinthina* and *P. racemosa* var. *grisea*. *Z. Naturforsch. C* **2002**, *57*, 449–451. [[CrossRef](#)]
38. Cardeal, Z.L.; Gomes da Silva, M.D.R.; Marriott, P.J. Comprehensive two-dimensional gas chromatography/ mass spectrometric analysis of pepper volatiles. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2823–2836. [[CrossRef](#)] [[PubMed](#)]
39. Yassaa, N.; Meklati, B.Y.; Cecinato, A. Evaluation of monoterpenic biogenic volatile organic compounds in ambient air around *Eucalyptus globulus*, *Pinus halepensis* and *Cedrus atlantica* trees growing in Algiers city area by chiral and achiral capillary gas chromatography. *Atmos. Environ.* **2000**, *34*, 2809–2816. [[CrossRef](#)]
40. Perez-Cacho, P.R.; Mahattanatawee, K.; Smoot, J.M.; Rouseff, R. Identification of sulfur volatiles in canned orange juices lacking orange flavor. *J. Agric. Food Chem.* **2007**, *55*, 5761–5767. [[CrossRef](#)]
41. Radulović, N.; Lazarević, J.; Ristić, N.; Palić, R. Chemotaxonomic significance of the volatiles in the genus *Stachys* (Lamiaceae): Essential oil composition of four Balkan *Stachys* species. *Biochem. Syst. Ecol.* **2007**, *35*, 196–208. [[CrossRef](#)]
42. Bortolomeazzi, R.; Berno, P.; Pizzale, L.; Conte, L.S. Sesquiterpene, alkene, and alkane hydrocarbons in virgin olive oils of different varieties and geographical origins. *J. Agric. Food Chem.* **2001**, *49*, 3278–3283. [[CrossRef](#)] [[PubMed](#)]
43. Lo Presti, M.; Sciarrone, D.; Crupi, M.L.; Costa, R.; Ragusa, S.; Dugo, G.; Mondello, L. Evaluation of the volatile and chiral composition in *Pistacia lentiscus* L. essential oil. *Flavour Fragr. J.* **2008**, *23*, 249–257. [[CrossRef](#)]
44. Weyerstahl, P.; Marschall, H.; Thefeld, K.; Subba, G.C. Constituents of the essential oil from the rhizomes of *Hedychium gardnerianum* Roscoe. *Flavour Fragr. J.* **1998**, *13*, 377–388. [[CrossRef](#)]
45. NIST Chemistry WebBook, SRD 69. Available online: <https://rb.gy/zy2yw9> (accessed on 20 August 2019).
46. dos Santos, B.C.B.; da Silva, J.C.T.; Guerrero, P.G.; Leitão, G.G.; Barata, L.E.S. Isolation of chavibetol from essential oil of *Pimenta pseudocaryophyllus* leaf by high-speed counter-current chromatography. *J. Chromatogr. A* **2009**, *1216*, 4303–4306. [[CrossRef](#)] [[PubMed](#)]
47. Mohottalage, S.; Tabacchi, R.; Guerin, P.M. Components from Sri Lankan *Piper betle* L. leaf oil and their analogues showing toxicity against the housefly, *Musca domestica*. *Flavour Fragr. J.* **2007**, *22*, 130–138. [[CrossRef](#)]
48. Saroglou, V.; Dorizas, N.; Kypriotakis, Z.; Skaltsa, H.D. Analysis of the essential oil composition of eight *Anthemis* species from Greece. *J. Chromatogr. A* **2006**, *1104*, 313–322. [[CrossRef](#)]
49. Karak, S.; Acharya, J.; Begum, S.; Mazumdar, I.; Kundu, R.; De, B. Essential oil of *Piper betle* L. leaves: Chemical composition, anti-acetylcholinesterase, anti- $\beta$ -glucuronidase and cytotoxic properties. *J. Appl. Res. Med. Aromat. Plants* **2018**, *10*, 85–92. [[CrossRef](#)]
50. Syahidah, A.; Saad, C.R.; Hassan, M.D.; Rukayadi, Y.; Norazian, M.H.; Kamarudin, M.S. Phytochemical analysis, identification and quantification of antibacterial active compounds in betel leaves, *Piper betle* methanolic extract. *Pak. J. Biol. Sci.* **2017**, *20*, 70–81. [[CrossRef](#)] [[PubMed](#)]
51. Pin, K.Y.; Luqman Chuah, A.; Abdull Rashih, A.; Rasadah, M.A.; Law, C.L.; Choong, T.S.Y. Solid-liquid extraction of betel leaves (*Piper betle* L.). *J. Food Process Eng.* **2011**, *34*, 549–565. [[CrossRef](#)]
52. Yusoff, Z.; Mahmud, Z.; Saleh, S.H.; Esa, Y.M. Study on the chemical constituents of *Piper betle* L. in relation to their possible insect attractant property. *Malays. J. Sci.* **2005**, *24*, 143–147.
53. Jantan, I.; Ahmad, A.R.; Ahmad, A.S.; Ali, N.A.M. A comparative study of the essential oils of five *Piper* species from Peninsular Malaysia. *Flavour Fragr. J.* **1994**, *9*, 339–342. [[CrossRef](#)]
54. Zamakshshari, N.; Ahmed, I.A.; Nasharuddin, M.N.A.; Mohd Hashim, N.; Mustafa, M.R.; Othman, R.; Noordin, M.I. Effect of extraction procedure on the yield and biological activities of hydroxychavicol from *Piper betle* L. leaves. *J. Appl. Res. Med. Aromat. Plants* **2021**, *24*, 100320. [[CrossRef](#)]
55. World Health Organization. Cortex Cinnamomi. In *WHO Monographs on Selected Medicinal Plants*; World Health Organization: Geneva, Switzerland, 1999; Volume 1, pp. 99–100.
56. World Health Organization. Flos Caryophylli. In *WHO Monographs on Selected Medicinal Plants*; World Health Organization: Geneva, Switzerland, 2002; Volume 2, p. 48.
57. World Health Organization. Folium Ocimi Sancti. In *WHO Monographs on Selected Medicinal Plants*; World Health Organization: Geneva, Switzerland, 2002; Volume 2, p. 209.
58. Mohamad, N.N. Isolation of Chavibetol and Hydroxychavicol and Selected Anti-obesity Studies of Standardized *Piper betle* Linn. Leaf Extracts. Master's Thesis, Universiti Sains Malaysia, Pulau Pinang, Malaysia, September 2016.
59. Bajpai, V.; Sharma, D.; Kumar, B.; Madhusudanan, K.P. Profiling of *Piper betle* Linn. cultivars by direct analysis in real time mass spectrometric technique. *Biomed. Chromatogr.* **2010**, *24*, 1283–1286. [[CrossRef](#)]
60. Adib, A.M.; Abdullah, F.; Kasim, N.; Kiong, L.S.; Ismail, N.H.  $^1\text{H}$  NMR Profiling of *Piper sarmentosum* Leaves: Optimisation for Sample Preparation (poster presentation). In *Book of Abstracts Asian Symposium on Medicinal Plants and Spices XVII (2021)-Virtual Conference, Malaysia, 18–19 August*; Ab. Ghani, N., Rasol, N.E., Eds.; Malaysian Natural Products Society (MNPS): Puncak Alam, Selangor, Malaysia, 2021.
61. Contreras-Moreno, B.Z. Chemical composition of essential oil of genus *Pimenta* (Myrtaceae): Review. In *Potential of Essential Oils*; El-Shemy, H., Ed.; IntechOpen Limited: London, UK, 2018; pp. 21–39. [[CrossRef](#)]
62. Marques, F.A.; Wendler, E.P.; Baroni, A.C.M.; De Oliveira, P.R.; Sasaki, B.S.; Guerrero, P.G. Leaf essential oil composition of *Pimenta pseudocaryophyllus* (Gomes) L. R. Landrum native from Brazil. *J. Essent. Oil Res.* **2010**, *22*, 150–152. [[CrossRef](#)]

63. Niculau, E.d.S.; Ribeiro, L.d.P.; Ansante, T.F.; Fernandes, J.B.; Forim, M.R.; Vieira, P.C.; Vendramim, J.D.; Da Silva, M.F.d.G.F. Isolation of chavibetol and methyl eugenol from essential oil of *Pimenta pseudocaryophyllus* by high performance liquid chromatography. *Molecules* **2018**, *23*, 2909. [CrossRef]
64. de Souza, L.P.; Garbowicz, K.; Brotman, Y.; Tohge, T.; Fernie, A.R. The acetate pathway supports flavonoid and lipid biosynthesis in *Arabidopsis*. *Plant Physiol.* **2020**, *182*, 857–869. [CrossRef] [PubMed]
65. Fu, X.; Yang, H.; Pangestu, F.; Nikolau, B.J. Failure to maintain acetate homeostasis by acetate-activating enzymes impacts plant development. *Plant Physiol.* **2020**, *182*, 1256–1271. [CrossRef] [PubMed]
66. Sanubol, A.; Chaveerach, A.; Sudmoon, R.; Tanee, T.; Liehr, T. Verification of selected *Piper* species (Piperaceae) using morphological characters, molecular data, and chemical constituents. *Malay. Nat. J.* **2014**, *66*, 60–81. Available online: <https://rb.gy/r00c7r> (accessed on 7 November 2020).
67. Verma, A.; Kumar, N.; Ranade, S.A. Genetic diversity amongst landraces of a dioecious vegetatively propagated plant, betel vine (*Piper betle* L.). *J. Biosci.* **2004**, *29*, 319–328. [CrossRef]
68. Kim, H.K.; Choi, Y.H.; Verpoorte, R. NMR-based metabolomic analysis of plants. *Nat. Protoc.* **2010**, *5*, 536–549. [CrossRef] [PubMed]
69. Oh, T.-J.; Hyun, S.-H.; Lee, S.-G.; Chun, Y.-J.; Sung, G.-H.; Choi, H.K. NMR and GC-MS based metabolic profiling and free-radical scavenging activities of *Cordyceps pruinosa* mycelia cultivated under different media and light conditions. *PLoS ONE* **2014**, *9*, e90823. [CrossRef]
70. Han, J.; Zhu, Y.; Chen, X.; Liao, B.; Yao, H.; Song, J.; Chen, S.; Meng, F. The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. *BioMed Res. Int.* **2013**, *2013*, 741476. [CrossRef]
71. Chen, S.; Yao, H.; Han, J.; Liu, C.; Song, J.; Shi, L.; Zhu, Y.; Ma, X.; Gao, T.; Pang, X.; et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* **2010**, *5*, e8613. [CrossRef]
72. Tamura, K.; Stecher, G.; Peterson, D.; Filipiński, A.; Kumar, S. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* **2013**, *30*, 2725–2729. [CrossRef]
73. Papadatou, M.; Argyropoulou, C.; Grigoriadou, C.; Maloupa, E.; Skaltsa, H. Essential oil content of cultivated *Satureja* spp. in Northern Greece. *Nat. Vol. Essent. Oils* **2015**, *2*, 37–48.
74. Van den Dool, H.; Kratz, P.D. A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography. *J. Chromatogr. A* **1963**, *11*, 463–471. [CrossRef]
75. Bicchì, C.; Chaintreau, A.; Joulain, D. Identification of flavour and fragrance constituents. *Flavour Fragr. J.* **2018**, *33*, 201–202. [CrossRef]
76. Jumhawan, U.; Putri, S.P.; Yusianto, Y.; Marwani, E.; Bamba, T.; Fukusaki, E. Selection of discriminant markers for authentication of asian palm civet coffee (Kopi luwak): A metabolomics approach. *J. Agric. Food Chem.* **2013**, *61*, 7994–8001. [CrossRef] [PubMed]
77. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **2010**, *11*, 395. [CrossRef] [PubMed]
78. Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. MS-DIAL: Data independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat. Methods* **2015**, *12*, 523–526. [CrossRef]
79. Broadhurst, D.; Goodacre, R.; Reinke, S.N.; Kuligowski, J.; Wilson, I.D.; Lewis, M.R.; Dunn, W.B. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* **2018**, *14*, 1–17. [CrossRef] [PubMed]
80. Stettin, D.; Poulin, R.X.; Pohnert, G. Metabolomics benefits from orbitrap GC-MS—Comparison of low- and high-resolution GC-MS. *Metabolites* **2020**, *10*, 143. [CrossRef] [PubMed]
81. Lever, J.; Krzywinski, M.; Altman, N. Points of significance: Principal component analysis. *Nat. Methods* **2017**, *14*, 641–642. [CrossRef]
82. Bro, R.; Kjeldahl, K.; Smilde, A.K.; Kiers, H.A.L. Cross-validation of component models: A critical look at current methods. *Anal. Bioanal. Chem.* **2008**, *390*, 1241–1251. [CrossRef]
83. Bro, R.; Smilde, A.K. Principal component analysis. *Anal. Methods* **2014**, *6*, 2812–2831. [CrossRef]