PHENETIC AND UNSUPERVISED MULTIVARIATE ANALYSIS
SYZYGIUM POLYANTHUM (WIGHT) WALP

Abdulrahman & et al.

SYZYGIUM POLYANTHUM (WIGHT) WALP

Mahmoud Dogara Abdulrahman1,2† Hasain Nudin Nur Fatihah1 Mohammed Abdul Manaf Ali1 Nashriyah Mat1 Moneruzzaman Khandaker1†

1Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin, Tembila Campus, 22200 Besut, Terengganu, Malaysia
2Department of Biological Sciences, Faculty of Science, Kaduna State University, Kaduna. Nigeria
*Corresponding author: abdulrahmanjj@kasu.edu.ng; moneruzzaman@unisza.edu.my

ABSTRACT

Myrtaceae family is widely distributed in Asia has been the largest group of plant; mainly trees and few shrubs. Distributed all over the world in tropical and subtropical areas. Syzygium is the largest genus with economical value found all over the Malaysian Peninsular. Evolutionary relationships within the Syzygium is unclear and there are currently no reliable criteria to divide the genus into manageable entities for systematic study. Species of Syzygium is the richest genus of woody plants in South East Asia with approximately 1000 or more species but little is known about the genus. Syzygium polyanthum Wight is one of the favourites Ulam that have been consumed for ages in Peninsular Malaysia and also as herbal medicine. The species is widely misunderstood due to extreme morphological variability, similarity in aroma and flavor. The species is substituted or adulterated with several other species. The study was aimed to construct phenetic tree and unsupervised multivariate analysis from morphological and anatomical the data matrix. Phenetic analysis, Principal component and hierarchical cluster analysis revealed they are two different cultivars species but inter variation exited among cultivars of same species. The above documented information has added new taxonomic information with regard to the identification of the cultivars in Peninsular Malaysia. The study recommends further study on de novo sequence of Serai kayu and Serai kayu hutan.

Keywords: Phylogenetic, Morphology, Anatomy, Phenetic

Received:2/2/2020, Accepted:22/5/2020
INTRODUCTION
Myrtaceae family is a tropical trees and shrubs with nearly about 55,000 species, grouped into two sub-family, 17 tribes and 142 genera. They are mainly found in large amount of number in Central America, South America, Australia and southern hemisphere (Harrington & Gadek, 2004; Craven & Biffin, 2010). Generally, the leaves of Myrtaceae family are in opposite direction, phloem internal, ever green and woody and abundance of oil glands in most of the members (Abdulrahman et al., 2018a). They are bisexual, poly stemonous, fully inferior or partial inferior ovaries mostly with nectariferous hypanthium and actinomorphic flower (Wilson et al., 2001; Gamage et al., 2003). The fruit of the family are normally divided as either fleshy (berry) or dry (capsular or nut-like), but numerous other fruit types, such as drupes, and various intermediates, also occur within the group (Wilson 201). Syzygium is a large genus of Myrtaceae, occurring from Africa eastwards to the Hawaiian Islands and from India and southern China southwards to south eastern Australia and New Zealand (Van Wyk, 1985). In terms of species richness, the genus is centered in Malaysia but in terms of its basic evolutionary diversity it appears to be centered in the Melanesian-Australian region (Hussin et al., 1992). Syzygium are densely foliaceous and large ever green plants with thick bark and grayish brown utilised in woody industry (Abdulrahman et al., 2018a). Wood from Syzygium is durable, close grained and whitish (Soh & Parnell 2011). They have leathery leaves, obovate elliptic or oblong ovate shape with six to twelve centimeters (Retamales et al., 2015). There is wide variation in shape, presence of primary, secondary and tertiary vein, shining, smooth and the tips of the leaves are less acuminate and broad (Abdulrahman et al., 2018a). From the branch lets below the leaves panicles arise which are four to six centimeters and being terminal or axillary (Retamales et al., 2015). The flower from Syzygium is in cluster oblong or round shape, greenish white and scented. The calyx is about four millimeters long with funnel shape and toothed (Abdulrahman et al., 2018b). The petals cohere looks like a small disk (Gamage et al., 2003). Numerous stamens about four millimeters long but differ in size and colour based on the species and environtmental factor (Gamage et al., 2003). Most members of the genus their fruits are oblong ranging from 0.5 to 3.5 centimeters long which are black or dark purple, edible and flesh and almost all species bear seed within the fruits (Gamage et al., 2003; Retamales et al., 2015). Most members their fruits are sweet, flavour and mildly sour (Retamales et al., 2015). Syzygium polyanthum W., is a tree or shrub ranging between 22m tall with glabrous or greenish brown bark. Opposite leaves with eliptic apex angle and cuneate base shape. S polyanthum grow low land areas, widely distributed in tropical, subtropical region of the world. Syzygium polyanthum native to Malaysia, china, Thailand and Indonesia. Known by Malay ethnic group as Serai kayu and Serai kayu hutan and locally identified by Indonesian as laurel Indonesia or bay leaf, widely used in their daily diet as Ulam in order to enhance their appetite (Abdulrahman et al., 2018b). Traditionally, the plants are known for the treatment of diabetes, malarial fever, gastrointestinal infection, ulcer, endometriosis and also health improvement through postpartum (Abdulrahman et al., 2018b). Despite the medicinal history or record of Syzygium polyanthum in Peninsular Malaysia the cultivars are widely misidentified or misnamed in ecological inventories or surveys which cause serious problems for biodiversity management. The taxonomy of Syzygium polyanthum is still controversial due to the extreme morphological variability. Multivariate data was utilised (qualitative and quantitative) to determine the relationship of closely related species. The present study aims at adding taxonomic information to the genus Syzygium; through phenetic and cluster analysis of Syzygium polyanthum cultivar (Serai kayu and Serai kayu hutan) based on morphology and anatomical data set.

MATERIALS AND METHODS
Plant materials
Samples of Serai kayu were collected from Terengganu, Kelantan, Johor (Batu Pahat area) and Penang; while Serai kayu hutan was collected from Terengganu, Peninsular Malaysia, and were taken to the Forest Research Institute of Malaysia (KEP),
Universiti Kebangsaan Malaysia (UKMB) and University of Malaya (KLU) for identification. The plant samples of *Serai kayu* and *Serai kayu hutan* were mounted, pressed and deposited at the Herbarium of Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin (UniSZA) with voucher numbers of 395 and 396, respectively.

**Morpho anatomical measurement and character coding**

A total of twenty morphological characteristics were observed, measured and coded, the characteristics were all derived from the leaf parts of the plants (leaf length, width and shape) because they are taxonomically important (Table 1). The following plants features can be observed and coded from herbarium specimen or live plants (Fatihah et al., 2012; Monteiro et al., 2015). Table 2 present fourteen anatomical characteristics were observed, measured and coded (Fatihah et al., 2012). *Syzygium mytifolium* was used as outgroup to determine the relationship and closeness of the cultivars of *Syzygium polyanthum* (*Serai kayu* and *Serai kayu hutan*) in Peninsular Malaysia. Phenetic analysis was carried out using Phylogenetic Analysis Using Parsimony (PAUP) software (V. 4.0b1; Swofford) was used to generate trees (Fatihah et al., 2012).

### Table 1. Morphological Characters and Character States Used

<table>
<thead>
<tr>
<th>S/N</th>
<th>Characters</th>
<th>Character coding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Leaf length</td>
<td>0=less or equal to 9 cm; 1= less or equal to 18 cm</td>
</tr>
<tr>
<td>2</td>
<td>Leaf width</td>
<td>0= less or equal to 2 cm; 1= less or equal to 4 cm</td>
</tr>
<tr>
<td>3</td>
<td>Petiole length</td>
<td>0= less or equal to 0.5 cm; 1= less or equal to 1 cm</td>
</tr>
<tr>
<td>4</td>
<td>Petiole morphology</td>
<td>0= base swollen; 1= pulvinate</td>
</tr>
<tr>
<td>5</td>
<td>Petiole attachment</td>
<td>0=marginal; 1=peltate central</td>
</tr>
<tr>
<td>6</td>
<td>Leaf shape</td>
<td>0= eliptic; 1= obovate</td>
</tr>
<tr>
<td>7</td>
<td>Leaf symmetry</td>
<td>0= symmetrical; 1= base asymmetrical</td>
</tr>
<tr>
<td>8</td>
<td>Leaf base angle</td>
<td>0=acute; 1= obtuse</td>
</tr>
<tr>
<td>9</td>
<td>Leaf apex angle</td>
<td>0=acute; 1= acute to acuminate</td>
</tr>
<tr>
<td>10</td>
<td>Leaf apex shape</td>
<td>0=straight; 1=convex</td>
</tr>
<tr>
<td>11</td>
<td>Leaf base shape</td>
<td>0= cuneate; 1=convex</td>
</tr>
<tr>
<td>12</td>
<td>Leaf margin</td>
<td>0=entire; 1= serrate</td>
</tr>
<tr>
<td>13</td>
<td>Leaf lobation</td>
<td>0=unlobed; 1=bilobed</td>
</tr>
<tr>
<td>14</td>
<td>Leaf colour</td>
<td>0=evergreen; 1=partial</td>
</tr>
<tr>
<td>15</td>
<td>Leaf odour</td>
<td>0= nil; 1= aromatic</td>
</tr>
<tr>
<td>16</td>
<td>Vein type</td>
<td>0=pinnate; 1=basal</td>
</tr>
<tr>
<td>17</td>
<td>Vein agrophic</td>
<td>0=simple agrophic; 1=compound agrophic</td>
</tr>
<tr>
<td>18</td>
<td>Vein spacing</td>
<td>0=uniform; 1=irregular</td>
</tr>
<tr>
<td>19</td>
<td>Vein category</td>
<td>0=opposite percurrent; 1=random reticulate</td>
</tr>
<tr>
<td>20</td>
<td>Tertiary vein</td>
<td>0= present; 1= absent</td>
</tr>
</tbody>
</table>

### Table 2. Anatomical Characters and Character States Used

<table>
<thead>
<tr>
<th>S/N</th>
<th>Anatomical characters</th>
<th>Character state</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Type of stomata</td>
<td>0=anisocytic; 1=anomocytic</td>
</tr>
</tbody>
</table>
data matrix derived from their morphological and anatomical features respectively. *Syzygium mytfolium* was used as outgroup throughout the study. Phenetic construction using the morphological and anatomical data matrix (Figure 1) was found to have Consistency index (CI) of 0.8500, Retention index (RI) of 0.9524 and homoplasy index (HI) 0.1500; the phylogenetic tree constructed from the combined data divided into three major clades with bootstrap value 90%, the first clade contain only the outgroup species *Syzygium mytfolium* (Figure 1). The second major clades divide into three subclades which contain all the five species of *Serai kayu* accounting to bootstrap value of 64, 61% and the last clades collapse. The last clade with a bootstrap value of 100% contain all the five species of *Serai kayu hutan* bases on the combined data matrix (Figure 1).

Figure 1. Phenetic Tree of Morphological and Anatomical Features of *Serai kayu* and *Serai kayu hutan*. Under line numbers are percentage of bootstrap value. Branches without values have less than 50% bootstrap value
Discrimination and classification of the plants species based on multivariate analysis was carried out using unsupervised analysis. This is due to the fact of unbiasedness of the methods PCA and HCA. The following study adopted the unsupervised pattern to discriminate and classified *Syzygium polyanthum* cultivars. Spectra filter model was obtained from the combination of morphological and anatomical data with X matrix with highest variation ($R^2_X$ (cum), 0.974) and the highest predictive power (fitness of the model) ($Q^2_{(cum)}$.0.913). The PCA constructed from combination of morphological and anatomical data; established the relationship (similarities and discrimination) that existed between the two cultivars *Serai kayu* and *Serai kayu hutan* (Figure 2). The PCA clearly discriminate the two species into two cultivar along the PC1 but no clear discrimination was found along the PC2 (Figure 2). Along the PC1 two main clusters were formed in *Serai kayu* and three main clusters in *Serai kayu hutan*. Similarly, in PC2 three main clusters were formed which comprises of both the *Serai kayu* and *Serai kayu hutan* along the negative score plots (Figure 2). While in the positive score plots of the PC2 two main clusters were formed of both *Serai kayu* and *Serai kayu hutan*. The PCA established relationship existed between the two cultivars as no discrimination was found in the PC2. Variation also existed within the cultivar of same species as it can be seen in *Serai kayu* and *Serai kayu hutan* (Figure 2). Loading plots display the correlation structure of the variables. Loading plots revealed the variables responsible for the discrimination between the two cultivars. The score and loading plots are complementing each other. The position of objects in a given direction in a score plot is influenced by variables lying in the same direction in the loading plot.

![Figure 2. PCA Score Plot of morphological and anatomical features of *Syzygium polyanthum* (SK; *Serai kayu* and SKH; *Serai kayu hutan*)](image-url)
Figure 3. Loading line of PC1 morphological and anatomical features of *Syzygium polyanthum* (SK; Serai kayu and SKH; Serai kayu hutan)

Biplots was constructed augments and allow us to further understand the score plots and loading line relationship and discrimination to display clearly the variables responsible for the relationship among the examined cultivars (Figure 5).

Figure 4. Loading line of PC1 morphological and anatomical features of *Syzygium polyanthum* (SK; Serai kayu and SKH; Serai kayu hutan)

Figure 5. Biplots of the morphological and anatomical features of *Syzygium polyanthum* (Serai kayu and Serai kayu hutan)
Hierarchical cluster analysis (HCA) was constructed from the combination of morphological and anatomical data matrix; resulted in to two major clades, with first major clade divided into two subclades, the first subclade contain two species of *Serai kayu* and the second subclade contain three species of *Serai kayu*. The second major clade was also divided into two subclade, the first subclade contain species of *Serai kayu hutan* (Figure 6). Second subclade further divided into two mini clades. The first mini clade contain only one species of *Serai kayu hutan* and the second mini clade contain two species of *Serai kayu hutan*.

![Dendrogram of the HCA on the morphology and anatomy of *Syzygium polyanthum* cultivars (*Serai kayu* and *Serai kayu hutan*)](image)

**Figure 6.** Dendrogram of the HCA on the morphology and anatomy of *Syzygium polyanthum* cultivars (*Serai kayu* and *Serai kayu hutan*)

As a result of diverse challenges attached to the molecular identification of plants, wrong identification happened with consequences in the wrong deposit of sequence in the NCBI (gene bank). In the process of extracting DNA from plant sample, you might end of extracting bacterial or fungal DNA attached to the plant sample (Klingenberg & Gidaszewski, 2010). Looking at this numerous challenges; there is quite a number of reasons why phenetic analysis of morphological data is of paramount importance (Baum & Smith, 2011). There are some taxa that are presently extinct, which make it very difficult for molecular studies (Lawing & Polly, 2010). For extinct species, you can only establish or traced their fossil record from phenetic analysis of morphological data which can be obtained from the literatures (Lawing & Polly, 2010). Quite a number of studies have been carried out in order to identify plants material to overcome adulteration and this add to the taxonomic identification of the plants (Peters et al., 2014). In studying taxonomic relationship in plants, morphological characters are still consider valuable (Brazeau et al., 2017). Though molecular sequence data is trying to outpace it but it remains indispensable in order to established ancestral phenotypic conditions. The study first documents the phenetic analysis of the genus *Syzygium* based on the morphological and anatomical data matrix; with the case study of *Syzygium polyanthum* cultivars as most of the previous study carried out on the genu are purely on qualitative morphological and anatomical features. Analogous situation was previously reported by Fatihah et al. (2012) on their studies phylogenetic analysis of seven varieties of *Ficus deltoidea* Jack in Peninsular Malaysia. previously genus *Syzygium* were classified in peninsular Malaysia based on physical observation of morphological data (Hussain et al., 1992). Hussain et al. (1992)
also reported the species in (1901) Malaysia Syzygium divided into two species known as Eugenia Syzygium and Jambosa group based on observable features of floral. Further divide them in to three groups Syzygium, Jamboselle and Jambosa. That is how it continues until late 90s where the genus was classified based on the shape of vascular bundle (Hussain et al., 1991). The following study used numerical analysis to further investigation on the taxonomic description of the genus Syzygium. The phenetic trees constructed from the combination of morphology and anatomy were found to have lower consistency index and higher retention index. The CI and RI show how good the constructed phenetic tree was. It has been previously reported by Fatihah et al. (2012) and Monteiro et al. (2015) that; lower values of CI and higher in RI revealed how good and robust the data matrix is in revealing the relationship that existed between the examined species. In order to further validate the results of phenetic analysis unsupervised recognition pattern is used, from which resulted in with how a particular plants sample differed from the others. The outcome of the analysis would bring the idea on what features they are sharing or what makes them different from each other (Gad et al., 2013). PCA and HCA are the two most popular and widely used unsupervised pattern analysis (Lagunin et al., 2014; Skov et al., 2014). Principal component analysis (PCA) is a technique that is unsupervised and provides useful visual result classification of the multivariate analysis (Gad et al., 2013; Lagunin et al., 2014; Matera et al., 2014; Song et al., 2014; Ballabio, 2015). The technique (PCA) has pull out the variation within the data set and displayed them according to their relationships or similarities, and at the same time making the data set very readable. Results from the PCA revealed the to two cultivars exhibit relationship but still differences existed between the two cultivar species. Where variation can clearly be seen in cultivar of same species. Hierarchical cluster analysis is used in multivariate data analysis (Okada et al., 2010). The technique presents its results in a form of tree (dendrogram) that explains the relationships between the study samples (Okada et al., 2010). HCA has been widely employed in other to discriminate plants according to their closest allies in terms metabolomic (Gad et al., 2013; N. Kumar et al., 2014). HCA employed in the present study found there is a strong relationship between the study cultivars. But variation existed even within the same species of same cultivars. Previously Yaradua et al. (2018) has also utilised numerical data for descrimination of genus Crotalaria. Chemical and bioactive descrimination of Ficus deltoidea was carried out in Penisular Malaysia (Yunusa et al., 2018). Raj et al. (2001) reported numerical analysis were believed to be robust in plant species identification toward similarity and differences they exhibited. The present study finds the two cultivars of Syzygium polyanthum based on morphological and anatomical data matrix; the cultivars were monophyletic, exhibited a very strong relationship, but nevertheless, there is variation between and within the cultivars. The relationship and variation documented in the present study might be as a result of mutation, cross pollination and sexual recombination. Therefore, this nictitated the morphometric analysis of medicinal plants with respect to their environment to avoid adulteration of the plants either intentionally or unintentionally. The study agrees with Fatihah et al. (2012) on the phylogenetic analysis of Ficus in Terengganu state, Peninsular Malaysia.

CONCLUSION
The phenetic analysis of the two cultivars based on morphometric data have delimited the two group into two distinct cultivars with intra variation within species of same cultivar. The study has added taxonomic information to the genus Syzygium. The study recommend further studies should be carried out on de novo sequence of the two cultivars; as the were no complete genome sequence of the species in National Centre for Biotechnology Information NCBI data base.

ACKNOWLEDGEMENTS
The authors wish to acknowledge Universiti Sultan Zainal Abidin (Unisza) Malaysia and Kaduna State University (KASU) Nigeria.
### REFERENCES


