Morphological Multivariate Discrimination of *Parlatoria cakiloidea* Boissm

Mahmoud Dogara Abdulrahman[1], Muhammad Usman[2] and Abdulkadir Yunusa[3]
[1]Department of Biology, Faculty of Education, Tishk International University, Erbil/Iraq
[2]Department of Biotechnology, Nigerian Defence Academy, Kaduna, Nigeria
[3]Department of Basic and Applied Science, Federal Polytechnic Mubi, Adamawa State, Nigeria

abdulrahman.mahmud@tiu.edu.iq

Abstract

Brassicaceae, previously Cruciferae, is a flowering plant family that includes 338 genera and 3,700 species. Members of the family are economically important but domestically altered by humans. The taxonomic information of *Parlatoria cakiloidea* is ambiguous. Presently, no standard criteria to identify the species into simple entities for taxonomic, conservation and pharmacognostic studies. The study aimed to taxonomically identified *Parlatoria cakiloidea*. Fourteen herbarium specimens of *P. cakiloidea* were examined from Salahaddin University using Unsupervised multivariate analysis. Principal component and hierarchical cluster analysis were constructed using software SIMCA-P (V.14.1 Umetrics Sweden). The multivariate analysis established the relationship (similarities and discrimination) that existed. *Parlatoria cakiloidea* were found to be divided into four major groups in Kurdistan region of Iraq. The information provided will serve as the basis for standardization for drug exploration from *P. cakiloidea* in Kurdistan. The study recommends further study on the anatomical features, de novo sequence and the chemical constituents of the species.

Keywords
Principal component analysis, Hierarchical cluster, Multivariate

1. Introduction

Medicinal plants are natural resources that produce valuable herbal products that are frequently used in the treatment of a variety of ailments (Bjørklund et al., 2018). The therapeutic, preventative, and curative characteristics of the medicinal plants under inquiry have been widely described in Ayurveda and have been employed since ancient times (Süntar, 2019). Because quite diverse species can be found under the same vernacular name, great attention should be devoted to the precise botanical identification of the materials and the thorough collecting of voucher specimens (Mahmoud et al., 2020). Chemometrics is used to analyze plants in a comprehensive and simultaneous manner. Chemometrics' novel technique have aided in the advancement of plant identification studies for herbal and medication formulation (Chiriac et al., 2021). Chemometrics is an effective approach in plant science for analyzing and identifying various plants and its component (De Carvalho et al., 2018). The number of plant species is estimated to exceed 1,600,000 all over the world. The method is computational, and it employs multivariate analysis to statistically process numerical or metabolite data (Abdulrahman et al. , 2021). To make it easier to identify them, divide them into groups. Principal component analysis (PCA), orthogonal projections to latent structure discriminant analysis (OPLS-DA), and hierarchical cluster analysis (HCA) are examples of discrimination analysis (Bajer et al., 2020). Multivariate analyses generate score plots and loading plots, also known as discrimination maps, to reduce data and interpret it visually (Madiona et al., 2018). Presently, taxonomic classification is aided by numerical taxonomy (Winston, 2018). Multivariate approaches are commonly used to determine the evolutionary relationships between different species (Tucker et al., 2018). Many authors have examined the interrelationships among flowering plant genera and families using various numerical taxonomic methods. A medicinal plant complex is a group of species that share common names, morphological features, and medicinal uses (Haider, 2018). A medicinal plant complex is a group of species that share common names, morphological features, and medicinal uses (Abdulrahman et al., 2018). Traditionally, *P. cakiloidea* is used in treating various ailments but our surprised no scientific investigation on the said plants in Kurdistan and other part of the world to the best of our literature search. To protect consumers from contamination, it is critical to have good quality control for medicinal herbs (Abdulrahman et al., 2018; Fibigr et al., 2018; Steinhoff, 2019). Traditional medicine has a significant impact on a population's overall health. Due to the misidentification of the plant and it is mixtures can pose a health risk. The study aimed at numerical taxonomic

identification of *P. cakiloidea* in Kurdistan region of Iraq to provide a baseline for the herbal product formulation from its parts.

1.1 Objectives
The study aimed to taxonomically identified *Parlatoria cakiloidea***.**

2. Material and Methods
2.1 Morphological Measurement and Character Coding
Thirty different morphological traits were seen, measured, and coded, all of which were derived from the leaf. plant components (leaf length, width, and thickness) since they are taxonomically distinct) significance (Table 2) (Abdulrahman et al., 2021). The plants listed below From there, features can be examined and coded (Table 1). live plants or herbarium specimens (Table 1).

Table 1: List of herbarium species used in the study

S/N	Species name	Collector	Location	Date	Voucher number
1	*Parlatoria cakiloidea*	Al-Shehbaz and Weinert	Sharanish	14/05/1978	3475
2	*Parlatoria cakiloidea*	Al-Shehbaz and Weinert	Penjwin	11/05/1978	2502
3	*Parlatoria cakiloidea*	Al-Shehbaz and Y. Faris	Azmer	02/04/1979	00838
4	*Parlatoria cakiloidea*	Al-Shehbaz and A.R. Mayah	Sinjar	05/1977	21975
5	*Parlatoria cakiloidea*	Al-Shehbaz and A.R. M. Hilli	Sharanish	4/051978	3477
6	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	15/04/1981	5189
7	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	14/04/1981	5185
8	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	13/04/1981	5187
9	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	16/04/1981	5191
10	*Parlatoria cakiloidea*	Al-Shehbaz and A.R. M. Hilli	Sinar	03/05/1978	3328
11	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	16/04/1981	5190
12	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	03/05/1981	5197
13	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	03/05/1981	5196
14	*Parlatoria cakiloidea*	Y. Faris	Piramagrun	29/04/1981	5192

2.2 Multivariate Analyses
Hierarchical Cluster analysis (HCA) and principal component analysis (PCA) were used to conduct multivariate analyses using a SIMCA-P (V.14.1 Umetrics Sweden) with an imputed data (Table 1).

2.3 Principal Component Analysis (PCA)
The correlation matrix of individual studies through the PCA. Its aim is to look at linearly related variables. Distinguishing of closely related species and identification of character state responsible for the relationship is determine through PCA (Abdulrahman et al., 2021).

2.4 Hierarchical Cluster analysis
Hierarchical cluster analysis is a multivariate analysis with an aim to discriminate date based on the available inputted information (Abdulrahman et al., 2021). By creating a similarity matrix between the studied individuals, HCA groups individuals who are very similar in one cluster and shows similarities and differences among and

within the clusters (Abdulrahman et al., 2021). Species clustered in one clan are considered to be the most closely related . To see how well the cluster analysis matches the distance matrix, the cophenetic correlation coefficient between the distance matrix and the tree matrix was determined.

TABLE 2. Character Codes and Character States Used

S/N	Characters	Character coding
1	Leaf length	0=less or equal to 9 cm; 1= less or equal to 18 cm; 2= greater than 18
2	Leaf width	0= less or equal to 3 cm; 1= less or equal to 6 cm; 2= greater than 6
3	Angle of the forked midrib	0= not forked or forked less than 45 degrees, 1= forked more than 45 degrees.
4	Primary midrib	0= yes; 1= no
5	Petiole length	0= less or equal to 2 cm; 1= less or equal to 4 cm; 2= greater than 4
6	Petiole morphology	0= base swollen; 1= pulvinate
7	Petiole attachment	0=marginal; 1=peltate central
8	Leaf attachment	0= opposite; 1= alternate; 2= decussate or other
9	Leaf organisation	0= simple; 1= palmately compound: 2= ternate or other
10	Lamina shape	0= elliptic; 1= obovate; 2= oblong and other
11	Leaf shape	0= eliptic; 1= obovate
11	Leaf reolation	0= lacking 1= poorly developed 2= moderately developed
12	Leaf symmetry	0=symmetrical; 1=base asymmetrical
13	Leaf base angle	0=acute; 1= obtuse
14	Leaf apex angle	0=acute;1= obtuse 2= odd-lobed acute apex
15	Leaf apex shape	0=straight; 1=convex
16	Leaf base shape	0= cuneate; 1=convex
17	Leaf margin	0=entire; 1=serrate 2= dentate
18	Leaf lobation	0=unlobed; 1=bilobed
19	Leaf colour	0=evergreen; 1=partial
20	Leaf colour (dried)	0=evergreen; 1= changed 2= partial
21	Leaf odour	0= nil; 1= aromatic
22	Marginal Ultimate	0= incomplete loops 1= complete loops
23	Vein type	0=pinnate; 1=basal
24	Vein agrophic	0=simple agrophic; 1=compound agrophic
25	Vein spacing	0=uniform; 1=irregular
26	Vein category	0=opposite percurrent; 1=random reticulate
26	Tertiary vein	0= present; 1= absent
27	Flower type	0= bunch ; 1= singles
28	Flower colour	0 = white; 1= different colour; 2= another colour
29	Fruits	0= edible; 1= nonedible
30	Leaf margin when dried	0=serrate, 1=entire

3. Results and Discussion

Natural plant medicines are becoming increasingly popular all over the world. Building a sophisticated and comprehensive quality evaluation system is the foundation for medicinal plant authenticity, safety, and efficacy. Specific species identification is required for the safety of medicinal plants (Liu, Guo, & Liu, 2018). Authenticating medicinal plants is a complex undertaking when botanical and traditional microscopy identification is impossible due to the drying and processing process (Pant, Singh, & Marwah, 2021). Because of their chemical diversity, poor availability, and variability, analyzing active compounds in medicinal plants is a complex undertaking (Pant et al., 2021). On the other hand, using sophisticated analytical equipment generates a vast amount of measurement data (Abdulrahman et al. , 2021). Researchers exploring therapeutic plants are now faced with a new challenge: gathering sufficient chemical data and interpreting the results (Ragupathy et al., 2019). To address these challenges, chemometrics is required. Multivariate analysis is commonly used in this

methodology to statistically analyse the vast amount of data (Yun, Li, Deng, & Cao, 2019). The use of chemometrics allows medicinal plants to be evaluated based on their similarities and differences (Ng, Samsuri et al., 2020). Because of its ease of use and quick results, Numerical taxonomy is an excellent method for determining phylogenetic relationships between flowering plants (Heylen, Debortoli, Marescaux, & Olofsson, 2021).Spectra filter model obtained from the coded character states (Table 3) with X matrix with highest variation (R^2X (cum). 0.916) and the highest predictive power (fitness of the model) (Q^2 (cum).0.965). The fitness of the model revealed how good the data sets are. The model is therefore fit for further analysis of the PCA and HCA model (figure 1 and 3). If the model is not fit the data will not run (Fatihah et al., 2012; Fatihah et al., 2013). The matrix data demonstrating the data's robustness in forecasting the study species' relationship (Abdulrahman et al.; Fatihah et al., 2012). As previously reported, morphological data combined with chemometrics has the potential to be used as rapid methods to discriminate between plants of the same cultivar or different species (Abdulrahman et al., 2021; Yaradua et al., 2018). Unsupervised analysis was used to discriminate and classify plant species based on multivariate analysis (Abdulrahman et al., 2021; Yaradua et al., 2018). This is due to the unbiased of the PCA and HCA methods. The unsupervised pattern was also employed in the following investigation to differentiate and classify *P. cakiloidea* species in the Kurdistan region. The PCA constructed from the character states coded from the leaves. Established the relationship (similarities and discrimination) that existed between the species collected from different part of Kurdistan. Along the Y axes two clusters were constructed in the second PC 2 (figure 1 and 2). Along the X axis also two clusters are formed in the PC 1(intra species variation). The X axes of the PC1 with variation accounting for 37% (R^2X.0.37) while the PC2 accounting for 29% (R^2X.0.29) (figure 1 and 2). PCA discriminate the fourteen species into four major groups. According to Principal Component Analysis (PCA), there were significant phenetic similarities and differences across the study species, as well as intra-group variance, with no evident distinction in any group along the PC1 and PC2 axes. This further prompted the construction of HCA.

Table 3: Results of the character state

Species	Morphological Coded data matrix																												
PK 1	1	0	0	0	1	0	1	0	1	1	1	0	0	0	1	0	1	0	1	1	1	1	0	1	0	0	1	1	1	1
PK2	1	0	0	0	1	0	1	1	1	0	0	0	0	0	1	0	0	0	1	1	1	1	0	1	0	0	0	1	1	1
PK3	1	0	1	0	1	0	1	1	1	2	0	0	1	0	0	0	0	0	?	1	1	1	1	1	1	1	1	0	0	0
PK4	1	1	1	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	1	1	1	0	1	1	1	1	0	0
PK5	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0
PK6	1	1	0	0	1	0	0	0	0	1	0	0	1	0	2	0	1	0	0	0	1	1	1	0	1	1	1	0	0	0
PK7	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1	0	0	1	1	1	1
PK8	1	0	0	0	1	0	1	0	1	1	1	0	0	0	1	0	0	0	1	1	1	1	0	1	0	0	1	1	1	1
PK9	1	0	1	0	1	2	1	0	1	1	1	0	0	0	1	0	?	0	1	1	1	1	0	1	0	0	0	1	1	1
PK10	1	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	0	1	0	1	1	1	0	1	0	1	0	0	0
PK11	1	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0
PK12	1	0	0	0	1	0	1	0	1	1	1	0		0	1	0	1	0	1	1	2	1	0	1	0	1	1	1	1	1
PK13	0	0	1	1	1	1	1	0	0	0	1	1	0	0	?	0	0	0	0	1	1	1	1	0	1	0	1	0	0	0
PK14	1	1	0	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	1	0	1	0	0	0

Note: PK= *Parlatoria cakiloidea*

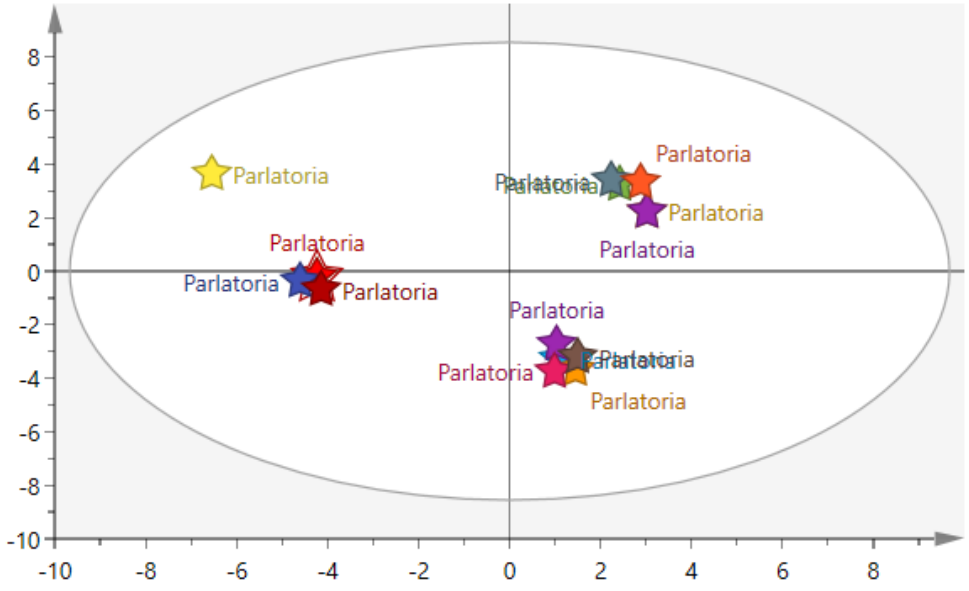

Figure 1: PCA Score Plot of morphological character state data of *P. cakiloidea*

PCA score plots are a good model for numerically classifying or identifying plants into small groups based on their relationship (Zhao et al., 2017). To determine their relationships, dissimilarities, and further categorize them into various classes.

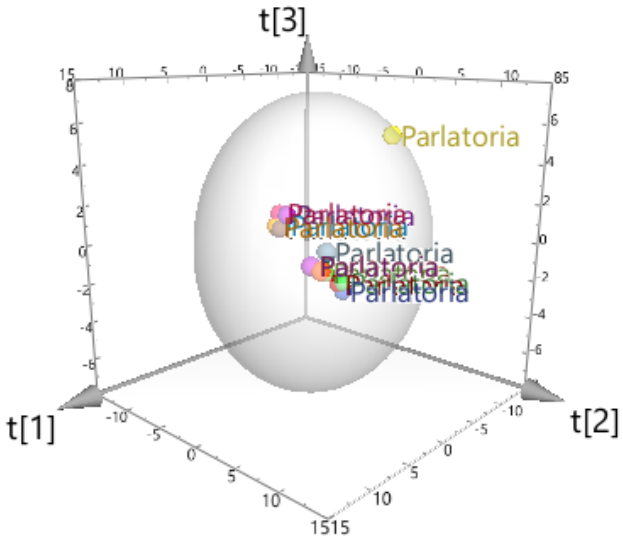

Figure 2: 3D Score Plot of morphological character state data of *P. cakiloidea*

The data set was utilized to create hierarchical cluster analysis (HCA) to classify them into various clades based on their relationship. HCA was constructed from the coded character state and two major clades were constructed (figure 3). The first major clades divided into two subclades with the first clades having only one species of *P. cakiloidea.* The second sub clades having three species of *P. cakiloidea.* The other major clades further subdivided into two major clades with each of the subclades containing five species of the *P. cakiloidea* (figure 3). HCA further discriminate the species based on their similarities into to four major clades. The HCA's dendrogram further supported the Relationship (similarities and dissimilarities) that the PCA had already established. The models revealed a clear separation of the species. Climate change has led to changes in species distributions, ecosystem changes, rare species impacts, and invasive species impacts (Ashcroft et al., 2017). Population genetic studies are required because they determine the degree of diversity and how it is spread across regions and people. Few research concentrated on individual variety. According to the following research, chemometrics coupled with multivariate analysis can be an effective approach for comprehensive identification of plants into small entities.

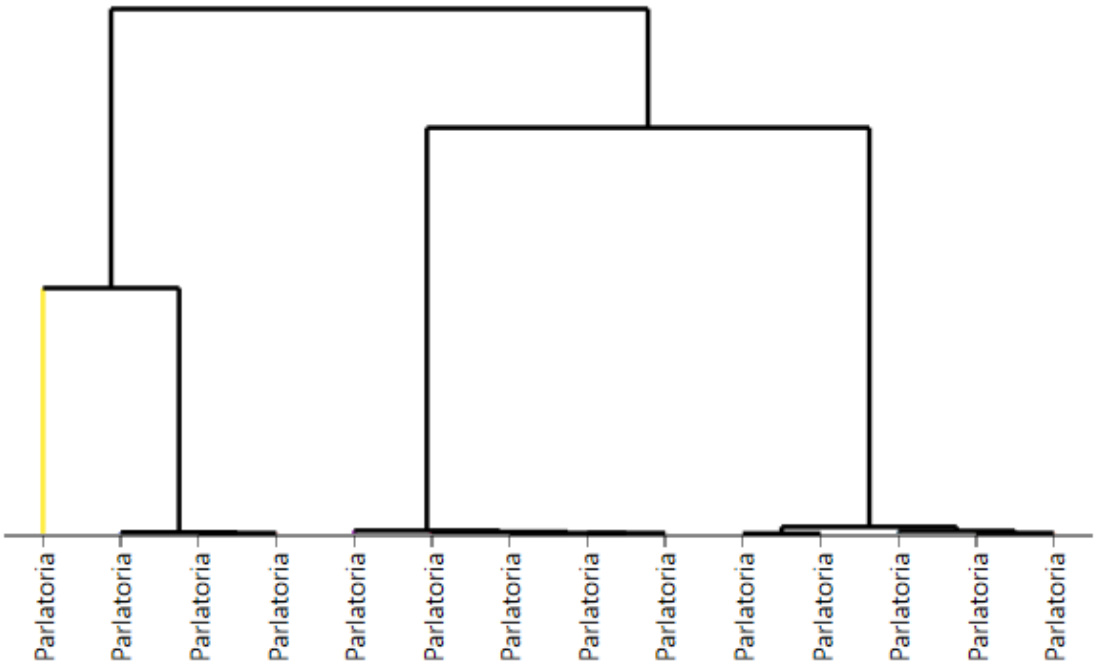

Figure 3: HCA display relationship of *P. cakiloidea*

4. Conclusion

Based on the phenetic, the researchers were able to fully differentiate *P. cakiloidea*. The study proved that there is a strong association depending on the species' location, but it also showed that there is intra variance within the same species gathered at the same area, despite their associations.

5. Acknowledgments

The researchers wish to acknowledge the support of Salahaddin University-Erbil herbarium and Tishk International University, Erbil/Iraq and all traditional herbalist that have contributed to the field collection.

References

Abdulrahman MD, Ali AM, Fatihah H, Khandaker MM, Mat N. Traditional medicinal knowledge of Malays in Terengganu, Peninsular Malaysia. *Malayan Nature Journal, vol. 70, no.* 3, pp. 349-364, 2018.

Abdulrahman, MD, Fatihah HNN, M. Abdul, M. Ali, N. Mat and M. Khandaker. Phenetic And Unsupervised Multivariate Analysis *Syzygium Polyanthum* (Wight) Walp. *Iraqi Journal of Agriculture, vol.* 52, no. 1, pp. 249-258, 2021.

Ashcroft MB, King DH, Raymond B, Turnbull JD, Wasley J, Robinson SA. Moving beyond presence and absence when examining changes in species distributions. *Global Change Biology, vol. 23, no.* 8, 2929-2940, 2017.

Bajer T, Hill M, Ventura K, Bajerová P. Authentification of fruit spirits using HS-SPME/GC-FID and OPLS methods. *Scientific Reports, vol. 10, no.* 1, 1-8, 2020.

Bjørklund G, Dadar M, Martins N, Chirumbolo S, Goh BH, Smetanina K, Lysiuk R. Brief Challenges on Medicinal Plants: An Eye-Opening Look at Ageing-Related Disorders. *Basic & Clinical Pharmacology & Toxicology, vol. 122, no.* 6, pp. 539-558, 2018.

Chiriac ER, Chiţescu CL, Geanăm EI, Gird CE, Socoteanu RP, Boscencu R. Advanced Analytical Approaches for the Analysis of Polyphenols in Plants Matrices—A Review. *Separations, vol. 8,* no. 5, 65-76, 2021.

De Carvalho LD, Neto AJS. Classification and authentication of plants by chemometric analysis of spectral data. In *Comprehensive Analytical Chemistry*, vol. 80, pp. 105-125, 2018.

Fatihah HN, Mat N, Zaimah AR, Zuhailah MN, Norhaslinda H, Khairil M, Ali AM. Morphological phylogenetic analysis of seven varieties of Ficus deltoidea Jack from the Malay Peninsula of Malaysia. *PLoS One, vol. 7, no.* 12, pp. 124-141, 2012.

Fatihah HN, Mat N, Zaimah AR, Zuhailah MN, Norhaslinda H, Khairil M, Ali AM. Correction: Morphological Phylogenetic Analysis of Seven Varieties of Ficus deltoidea Jack from the Malay Peninsula of Malaysia. *Plos one, vol. 8, no.* 8, pp. 171-189, 2013.

Fibigr J, Šatínský D, Solich P. Current trends in the analysis and quality control of food supplements based on plant extracts. *Analytica Chimica Acta, vol. 10, no. 36*, pp. 1-15, 2018.

Haider N. A Brief Review on Plant Taxonomy and its Components. *The Journal of Plant Science Research, vol. 34, no.* 2, pp. 277-292, 2018.

Heylen OC, Debortoli N, Marescaux J, Olofsson JK. A Revised Phylogeny of the Mentha spicata Clade Reveals Cryptic Species. *Plants, vol. 10, no.* 4, pp. 819-831, 2021.

Liu C, Guo Da, Liu L. Quality transitivity and traceability system of herbal medicine products based on quality markers. *Phytomedicine, vol. 44*, pp. 247-257, 2018.

Madiona RM, Welch NG, Russell SB, Winkler DA, Scoble JA, Muir BW, Pigram PJ. Multivariate analysis of ToF-SIMS data using mass segmented peak lists. *Surface and Interface Analysis, vol. 50, no.* 7, pp. 713-728, 2018.

Mahmoud AD, Fatihah HNN, Khandaker MM, Ali AM, Mat N. Ethnobotany of Syzygium polyanthum (Wight) Walp In Terengganu, Peninsular Malaysia. *Journal Of Agrobiotechnology, vol. 11, no.* 2, pp. 39-47, 2020.

Ng ZX, Samsuri SN, Yong PH. The antioxidant index and chemometric analysis of tannin, flavonoid, and total phenolic extracted from medicinal plant foods with the solvents of different polarities. *Journal of Food Processing and Preservation, vol. 44, no.* 9, pp. 46- 61, 2020.

Pant J, Singh RM, Marwah H. Methods employed in quality control in ayurveda. *Journal of Medicinal Plants, vol. 9, no.* 3, pp. 141-155, 2021.

Ragupathy S, Faller AC, Shanmughanandhan D, Kesanakurti P, Shaanker RU, Ravikanth G, Han J. Exploring DNA quantity and quality from raw materials to botanical extracts. *Heliyon, vol. 5, no.* 6, pp. 193-215, 2019.

Steinhoff B. Quality of herbal medicinal products: State of the art of purity assessment. *Phytomedicine, vol. 60*, pp. 153-168, 2019.

Süntar I. Importance of ethnopharmacological studies in drug discovery: role of medicinal plants. *Phytochemistry Reviews*, pp. 1-11 2019.

Tucker CM, Davies TJ, Cadotte MW, Pearse WD. On the relationship between phylogenetic diversity and trait diversity. *Ecology, vol. 99,* no. 6, pp. 1473-1479, 2018.

Winston JE. Twenty-first century biological nomenclature—the enduring power of names. *Integrative and Comparative Biology, vol. 58, no.* 6, pp. 1122-1131, 2018.

Yaradua SS, Alzahrani DA, Bello A. Numerical taxonomic study of the genus Crotalaria L.(Crotalarieae, Fabaceae) in Nigeria. *Biodiversity Research and Conservation, vol. 50*, no. 1, pp. 25-32, 2018.

Yun YH, Li HD, Deng BC, Cao DS. An overview of variable selection methods in multivariate analysis of near-infrared spectra. *TRAC Trends in Analytical Chemistry, vol. 113*, pp. 102-115, 2019.

Zhao SY, Liu ZL, Shu YS, Wang ML, He D, Song ZQ, Zeng HL, Ning ZC, Lu C, Lu AP. Chemotaxonomic Classification Applied to the Identification of two closely-related citrus TCMs using UPLC-Q-TOF-MS-based metabolomics. *Molecules, vol. 22, no.* 10, pp. 17-29, 2017.

Mahmoud Dogara Abdulrahman. Department of Biology, Faculty of Education, Tishk International University, Erbil/Iraq. With Research Area in Plant Systematics, Plant Taxonomy and Plant Anatomy. Published papers at Web of Science, Scopus index and some other non-index research articles.

Muhammad Usman. Department of Biotechnology, Nigerian Defence Academy, Kaduna, Nigeria is a researcher in the National Board for Arabic and Islamic studies. Published in the index and non index journals.

Abdulkadir Yunusa. Is a lecturer with the [3]Department of Basic and Applied Science, Federal Polytechnic Mubi, Adamawa State, Nigeria.