SINERGI Vol. 27, No. 1, February 2023: 23-30
http://publikasi.mercubuana.ac.id/index.php/sinergi
http://doi.org/10.22441/sinergi. 2023.1.004

Implementation of Bayesian inference MCMC algorithm in phylogenetic analysis of Dipterocarpaceae family

Mirna Yunita[1], Rachmat Muwardi[2*], Zendi Iklima[2]
[1]School of Computer Science and Technology, Beijing Institute of Technology, China
[2]Department of Electrical Engineering, Faculty of Engineering, Universitas Mercu Buana, Indonesia

Abstract

Dipterocarpaceae is one of the most prominent plant families, with more than 500 members of species. This family mostly used timber plants for housing, making ships, decking, and primary materials for making furniture. In Indonesia, many Dipterocarpaceae species have morphological similarities and are challenging to recognize in the field. As a result, the classification process becomes difficult and even results are inconsistent when viewed only from the morphology. This research will analyze the phylogenetic tree of Dipterocarpaceae based on the chloroplast matK gene. The aim of the research is to classify the phylogenetics tree of Dipterocarpaceae family using Bayesian inference algorithm. This research used the chloroplast gene instead of morphological characters which has more accurate. The analysis steps are collecting data, modifying the structure sequence name, sequence alignment, constructing tree by using Markov Chain Monte Carlo (MCMC) from Bayesian Inference, and evaluating and analyzing the phylogenetic tree. The results showed that the tree constructed based on the gene is different from the tree based on morphology. Based on the morphological, Dipterocarpus should be in the Dipterocarpeae tribe but based on the similarity of its genes, Dipterocarpus is more similar to the Shoreae tribe.

Keywords:
Bayesian inference;
Dipterocarpaceae;
Markov Chain Monte Carlo;
Phylogenetic;

Article History:
Received: March 22, 2022
Revised: May 24, 2022
Accepted: July 8, 2022
Published: February 2, 2023

Corresponding Author:
Rachmat Muwardi
Electrical Engineering Department, Universitas Mercu Buana, Indonesia
Email:
rachmat.muwardi@mercubuana.ac.id

INTRODUCTION

Dipterocarpaceae or Meranti group is one of the leading timber plants from tropical forests [1][2]. The family member has big size and height until 70-85m. *Dipterocarpaceae* consist of three subfamilies: *Dipterocarpoideae*, *Pakaraimoideae*, and *Monotoideae*. This family also has 13 genera and 470 species [3].

Indonesia's tropical forests have at least nine genera from the *Dipterocarpaceae* tribe. Among the many genera of *Dipterocarpaceae*, the genera found in Indonesia are Anisoptera (Mersawa), *Cotylelobium*, *Dipterocarpus* (Keruing), *Dryobalanops* (Kapur), *Hopea* (Giam), *Parashorea*, *Shorea* (Meranti), *Vatica* (Resak), and *Upuna*.

Dipterocarpaceae are known as closely related species. It is sometimes morphologically tricky to identify the member of the *Dipterocarps* family because of its similarity [4]. Each type has a different sequence of DNA bases, but in one type, the sequence is almost unchanged to be used as a unique marker of a particular type. A phylogenetic tree based on morphology of Dipterocarpaceae is shown in Figure 1.

A new way in modern taxonomy is using DNA barcoding. The role of DNA barcoding is considered as a powerful and promising tool for plant biodiversity analysis [5][6]. A short-standardized sequence (no more than 800 bp) that is able to delimit species is called DNA barcoding [7].

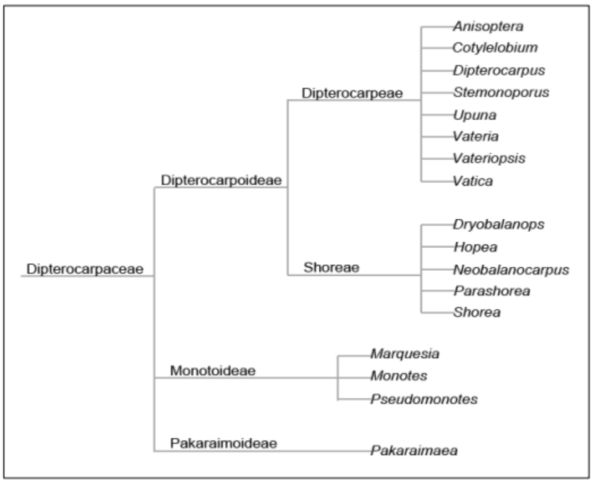

Figure 1. *Dipterocarpaceae* phylogenetic tree based on morphology

DNA barcoding is short-standardized DNA, no more than 800 bp, used as a new tool to identify species. Moreover, DNA barcoding is about amplifying and sequencing the informative site of the DNA sequences.

The closely related organisms have similar characters that are difficult to differentiate [8]. *Dipterocarpaceae* is well known for its closely related species, so these family members are difficult to identify by morphological characters. Chloroplast gene such as *mat*K is commonly used to learn the relationship among plant species [9].

Bayesian inference is one of the very popular used in inferring phylogenetic relationships among species [10]. Bayesian estimation of phylogeny has become widely used by evolutionary biologists [10, 11, 12, 13, 14]. This method applies the Markov Chain Monte Carlo (MCMC) analysis to develop the best tree based on prior and posterior probability to strengthen the relationship between the posterior trees [15].

Maulida [16] conducted a phylogenetic study of the *Dipterocarpaceae* family based on the *mat*K chloroplast gene. She constructed a phylogenetic tree of the *Dipterocarpaceae* family using another tree construction algorithm, the Neighbor-joining algorithm.

Another research was also conducted by [17]. Researchers evaluated the performance of SMC phylogenetic annealing by reviewing and comparing with other computational Bayesian phylogenetic methods. Researchers exploited the MCMC chain tree motion and benefited from the large inventory of such movements available in the literature.

In other respects, MCMC is also known as a good method for clustering, another research was also conducted by [18]. The researcher develops a Monte Carlo coupling algorithm, describes a procedure for pairing a Markov Chain targeting the posterior distribution over a phylogenetic tree space with age, scalar parameters, and latent variables in a phylogenetic study.

Based on the description above, the initial problem identified was that the grouping of the *Dipterocarpaceae* could not be done based on morphology alone because the morphological form of the tribe had a high degree of similarity between its sub-tribes. It is necessary to analyze the phylogenetic relationship between the *Dipterocarpaceae* family to find out which species are closely related to the *Dipterocarpaceae* based on the phylogenetic tree that has been formed.

This study proposes the construction of a more accurate gene-based phylogenetic tree. The Bayesian inference method with the MCMC algorithm was chosen because it has good accuracy in forming a phylogenetic tree by taking into account the probability of each tree formation. The results of the accuracy of the tree can be seen from the score of the tree branches formed.

METHOD

In terms of getting the best tree, the analytical steps carried out include data collection, modification of sequence structure names, alignment of sequences, tree creation using the MCMC method from Bayesian Inference, as well as evaluation and analysis of phylogenetic trees. The illustration of the work scheme can be seen in Figure 2. The stages of the research work will be explained in the next subchapter.

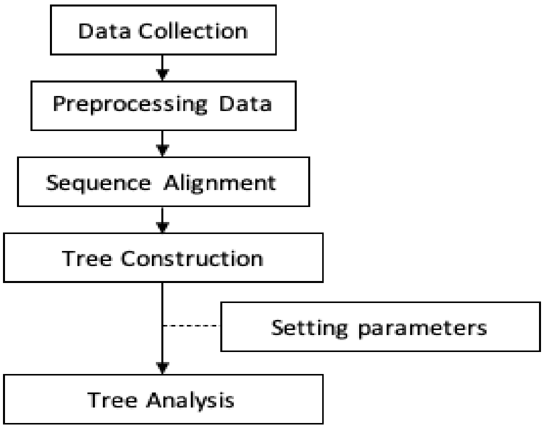

Figure 2. Work schematic illustration

Material

The data in this study were collected from the NCBI site in the nucleotide database from the *Dipterocarpaceae* sub-tribe, *Dipterocarpoideae*, which consists of the *Dipterocarpaeae* clan which has eight genera (*Anisoptera, Cotylobeum, Dipterocarpus, Stemonoporus, Upuna, Ateria, Vateriopsis*, and Vatica) and the *Shorea* family, which has five genera (*Dryobalan, Hopea, Neobalanocarpus, Parashorea, and Shorea*). One other species was taken from a different clan used as an outgroup, namely the *Monotoidea*.

We search for *mat*K sequence data of the *Dipterocarpaceae* family from the NCBI nucleotide database using tool BLAST (Basic Local Alignment System Tools) by entering one of the tribes from the *Dipterocarpaceae* tribe, *Anisoptera*, with the search query 'Anisoptera *mat*K'. The query entered must be one of the family of the *Dipterocarpaceae* family in order to produce a number of species that are indicated to have a relationship with the query entered.

The results of this search found 100 sequences similar to the *Anisoptera* sequence. Next, several types are selected to be used as data. There were 38 *Dipterocarpeae* sub-family sequences were downloaded from the NCBI database. We put *Monotes glaber* from subfamily *Monotoideae* as an outgroup.

Phylogenetic Tree Construction

The phylogenetic tree was constructed using the sequence data. The sequence alignment process is carried out to see similarities between two or more homologous sequences [19]. Sequence alignment was done to get the homolog DNA using Clustal-W [20]. At the time of sequence alignment, gap processes occur, including substitution, deletion, and insertion, which result in sequence separation or what is known as a gap. Clustal-W program can determine the quantity and quality of transcription of a gene that can indicate which genes are active under certain circumstances [21]. Then we develop the phylogenetic tree using the alignment sequences by the Bayesian inference algorithm.

The branch length analysis is then calculated using the MCMC analysis method by setting important parameters to support the analysis process.

Bayesian Inference

The Bayesian method is generated from conditional probabilities, with two random parameters (y and θ). Both are random variables with the combined density function $h(\theta|y)$. All information obtained in an event can be calculated based on the Bayesian formula in (1).

$$h(\theta|y) = g(\theta|y)f(y) = f(y|\theta)g(\theta) \qquad (1)$$

The sample y observation comes from the density function $f(\theta|y)$, identical to the likelihood function $l(\theta|y)$ in the equation. The likelihood function $l(\theta|y)$ contains all sample information about θ. The combined density function $h(\theta|y)$ Posterior is considered a new possibility that occurs accidentally when taking the probability $p(\theta|y)$. However, the posterior is still considered the original condition [15].

So that it is obtained:

$$\begin{aligned} g(\theta|y) &= \frac{f(y|\theta)g(\theta)}{f(y)} \qquad (2) \\ &= l(\theta|y)g(\theta)/f(y) \end{aligned}$$

The Bayesian method has the advantage of updating the probability distribution to be close to the truth. The Bayesian method will find the best tree based on the posterior probability distribution among other trees. Example tree of Bayesian method shown as in Figure 3.

In Figure 3, the first branch (on the left) is one of the two original conditions for a probability that occurs, while the branch that leaves the first branch (on the right) is one of the four conditions for the possibility of the first branch occurring. In constructing a phylogenetic tree for the *Dipterocarpaceae* family, the Bayes method was used to look for new possibilities that could strengthen the possible relationship of 38 types of results from multiple sequence alignment.

MCMC

The MCMC algorithm is a class for sampling probability distribution by building a Markov chain in a particular stationary distribution [22]. This method will calculate the likelihood value in a phylogenetic tree. MCMC integrates the random variable t and the prior parameter. The integration of this method will be more efficient when compared to other numerical integration methods as an integral dimension [23].

The search process in the MCMC method uses the concept regarding the data set, then the trees that are allowed to be in one prospective term will converge (towards the same point) with each other. In other words, the selected group of trees will have a higher probability of explaining the phylogeny of the dataset than the other trees. In the MCMC context, a new tree is instead drawn from a suitable probability distribution over the candidate trees [24].

In the first stage, the Markov chain starts from a tree along with the estimated evolution parameters. The tree can be chosen randomly or made with estimates that describe the data well. In the second stage, a new tree τ' is proposed to replace τ. In the third stage, the proposed new tree will be accepted with probabilities in (3).

$$
\begin{aligned}
R &= m\left[1, \frac{\Pr(\tau'|X)}{P(\tau|X)} x \frac{P(\tau|\tau')}{P(\tau'|\tau)}\right] \\
&= m\left[1, \frac{P(X|\tau')P(\tau')/P(X)}{P(X|\tau)P(\tau)/P(X)} x \frac{P(\tau|\tau')}{P(\tau'|\tau)}\right] \qquad (3) \\
&= m\left[1, \frac{P(X|\tau')}{P(X|\tau)} x \frac{P(\tau')}{P(\tau)} x \frac{P(\tau|\tau')}{P(\tau'|\tau)}\right]
\end{aligned}
$$

Where $P(X|\tau')/P(X|\tau)$ is the likelihood ratio, $P(\tau')/P(\tau) P(\tau')/P(\tau)$ is the prior ratio, and $P(\tau|\tau')/P(\tau'|\tau) P(\tau|\tau')/P(\tau'|\tau)$ is the proposal ratio. The results of the calculation in this third stage determine whether the posterior probability of the new status is better (uphill) or worse (downhill) than the current status.

As the MCMC chain moves following each proposal leading to areas with a high posterior probability value, the likelihood value increases rapidly, as shown in Figure 4. The initial phase of the MCMC chain with a low likelihood value is referred to as the burn-in phase. Samples in this burn-in phase are generally not used in the analysis or removed because the likelihood value is strongly influenced by the starting point of the MCMC chain.

The burn-in phase marked with faded colors accounted for about 10% of the total posterior sample and was not included in the statistical analysis.

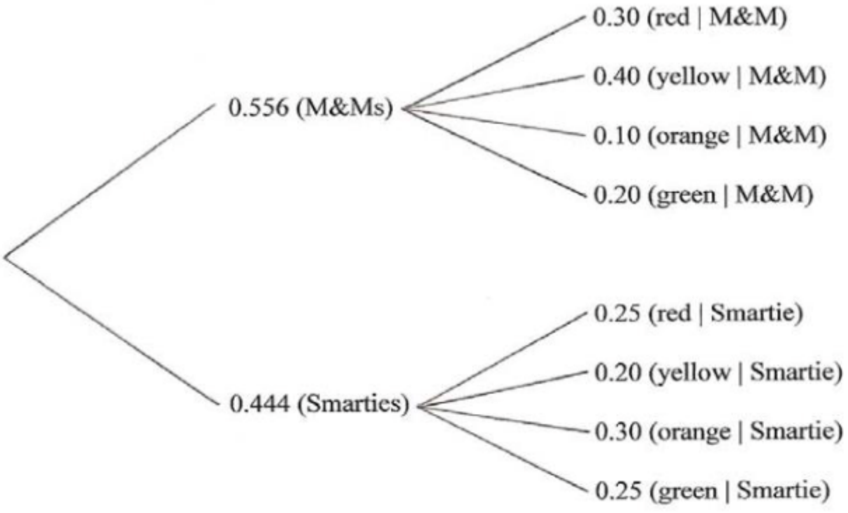

Figure 3. Example tree of Bayesian method

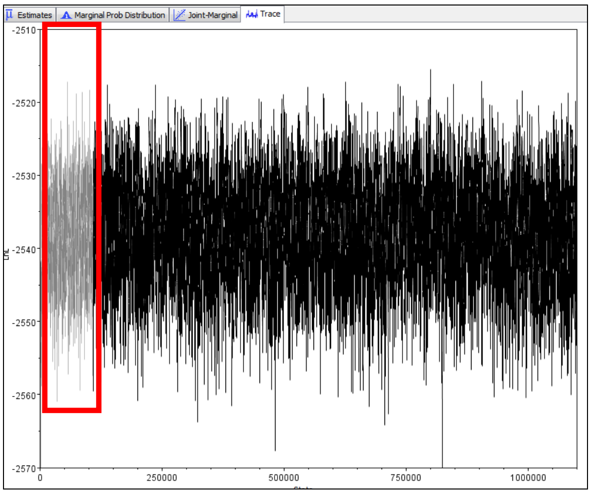

Figure 4. The burn-in phase marked

Trace plot graph of the MCMC chain visualized with the TRACER program [25]. The burn-in value that was not included was 25% of the total posterior sample.

RESULTS AND DISCUSSION

This work use Mr. Bayes software [26][27] to generate the tree. Bayesian methods used the accuracy analysis by counting the branch length with MCMC analysis. According to [28], MCMC in the Bayesian method estimates the phylogenetic tree posterior based on the dataset. While using this method, the essential parameters are arranged to determine the prior and posterior and how many generations will be done based on iteration number.

The tree was validated by determining several important parameter values such as likelihood and prior parameters. In addition to setting the likelihood of the evolution model from the prior settings that will be carried out next, such as substitution between nucleotides. The Bayesian method uses the general time reversible (gtr) nucleotide substitution evolution model. We also set the number of generations to be performed from the number of iterations of 100,000, this is done to find the best tree.

The MCMC method will stop analyzing until the review of the standard deviation value approaches the number 0.

Evolutionary relationship of *Dipterocarpaceae* using Bayesian inference algorithm

The tree construction using the Bayesian algorithm is divided into two monophyletic groups. One group is for tribe *Shoreae*, and one group is for *Dipterocarpeae*. However, from the cladogram, we see that the *Dipterocarpus* genus does not belong to the group of the *Dipterocarpeae* tribe, it is the basal group of the *Dipterocarpeae* tribe. Moreover, the *Dipterocarpus* genus is known as the basal of the *Dipterocarpaceae* family therefore the characters in this genus are very specific, and the name *Dipterocarpaceae* has come from this genus.

Our tree result is slightly different from the the morphological characters tree. Where, in morphological character, *Dipterocarpus* is allied with the *Dipterocarpeae* tribe. According to [29], Dipterocarpus belong to the *Dipterocarpeae* tribe, which has chromosome number 11.

The tree was validated using Trace software by determining several important parameter values such as likelihood parameters, prior parameters and iterations. The result shown a branch length value which is defined as the evolutionary value of the organism. The tree accuracy using the Bayesian algorithm can be seen from the branch length value. The construction time depends on the number of generations arranged from the Markov chain in the MCMC process. The longer the branches, the farther level of similarity with the closest organism. The construction results using the Bayesian algorithm can be seen in Figure 5.

Figure 5 describe the results of the construction form groupings that are connected to each other. The results of the construction using the Bayesian algorithm formed three groups consisting of the *Shoreae*, the *Dipterocarpeae* and the Dipterocarpus tribe. The genus *Dipterocarpus* is grouped outside of the *Dipterocarpeae* tribe but also not within the *Shoreae* tribe. Dipterocarpus forms a separate group outside the sub-tribe *Dipterocarpoideae*. This is not in accordance with the morphological grouping, in which the *Dipterocarpus* genus should be in the *Dipterocarpeae* family group.

However, there is one *Dipterocarpus* group that does not join other Dipterocarpus groups but belongs to the *Shoreae* tribe. This is thought to be because the sequences of these species have different genes in the same group, or it could also be caused by misclassification based on morphology.

The tree that was formed also saw an outgroup separation (*Monotes glaber*) with the subgroup of the *Dipterocarpoideae* tribe, but in the results using the Bayesian algorithm, the *Dryobalanops* group was not included in the *Shoreae* group, but this group also became an outgroup that came out of the *Dipterocarpoideae* subgroup.

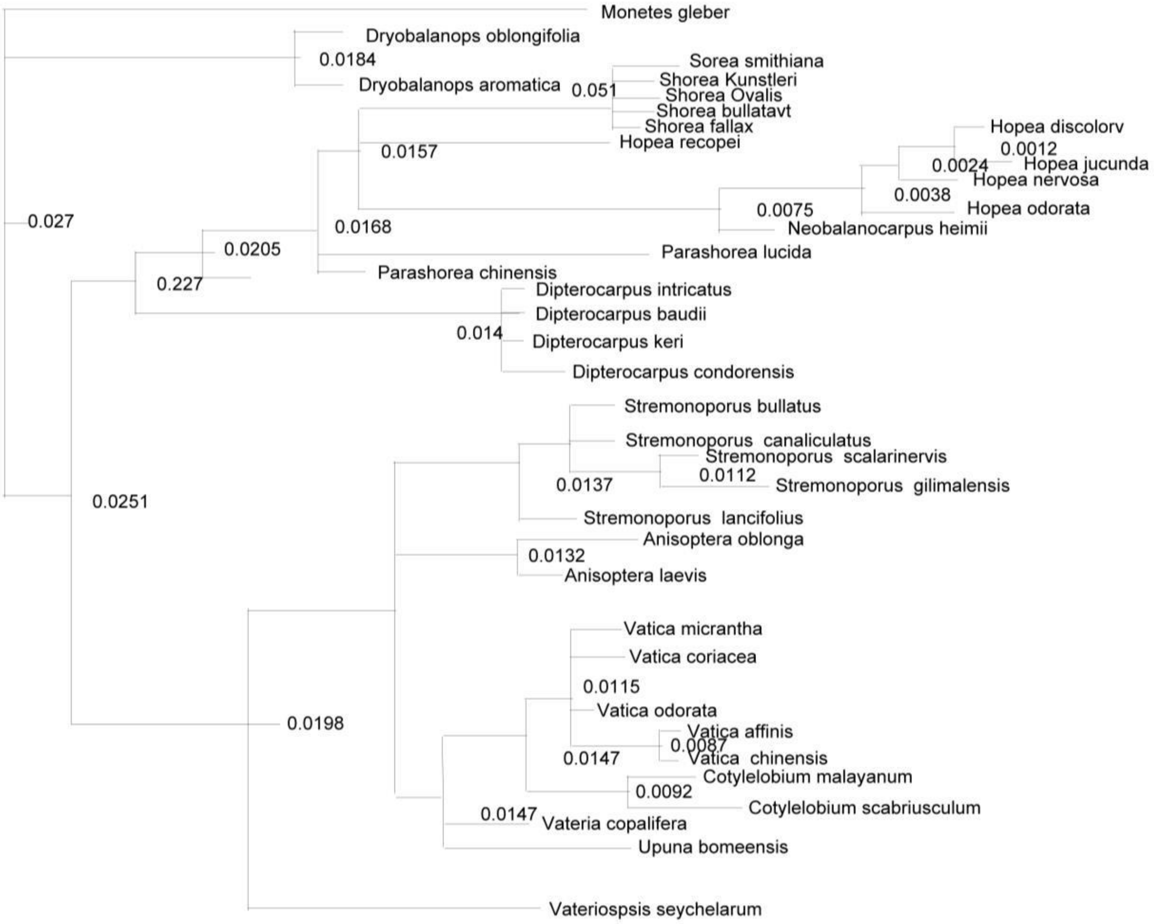

Figure 5. Phylogenetic tree with Bayesian Inference algorithm

This is also not in accordance with the morphological grouping. This happens because the possibility of these groups did not have similar genes so that they are not thought to have come from the same group.

CONCLUSION

The conclusion of this study is that the analysis using the Bayesian Inference algorithm shows that there is a difference with the tree based on morphology. This can be due to several reasons such as misclassification, or because of evolution that causes the genes to change. One of the visible differences is that the *Dipterocarpus* genus belongs to the *Shoreae* family group which is different from the tree constructed based on morphology. This proves that evolution can also cause gene changes in plants. The accuracy of the constructed tree can also be seen from the value of the branch length. The most extended branch was for the *Hopea* genus, and the closest branch length is for the *Dipterocarpus* genus. This is different from the results of phylogenetic trees based on morphology. The *Dipterocarpus* group does not belong to the *Dipterocarpeae* genus in bayesian trees.

ACKNOWLEDGMENT

The first special thanks go to Universitas Mercu Buana, Rachmat Muwardi and Zendi Iklima for their help and cooperation during this research. Providing unfailing support and continuous encouragement. Hopefully, the collaboration in publishing papers will continue in further research.

REFERENCES

[1] R. Djarwanto et al., *Pengelompokan Jenis Kayu Perdagangan Indonesia*, Bogor: Publisher: FORDA Press, 2017.

[2] R. Hegnauer, *Chemotaxonomie Der Pflanzen: Eine Ubersicht Uber Die Verbreitung Und Die Systematische Bedeutung Der Pflanzenstoffe - Band 4 - Dicotyledoneae: Daphniphyllaceae-Lythraceae*, Birkhauser Verlag AG, 1966.

[3] D. C. Zemp et al., "Tree performance in a biodiversity enrichment experiment in an oil

palm landscape," *Journal of Applied Ecology,* vol. 56, no. 10, pp. 2340-2352, 2019, doi: 10.1111/1365-2664.13460

[4] A. B. Rangkuti, A. Susilowati, H. Arinah, L. Irmayanti, I. Z. Siregar and U. J. Siregar, "Identification of Shorea species based on leaf morphological character with multivariate analysis," *IOP Conf. Series: Earth and Environmental Science*, vol. 912, no. 1, p. 012084, 2021.

[5] K. A. Mosa, S. Gairola, R. Jamdade, A. El-Keblawy, K. I. A. Shaer, E. K. A. Harthi, H. A. Shabana and T. Mahmoud, "The Promise of Molecular and Genomic Techniques for Biodiversity Research and DNA Barcoding of the Arabian Peninsula Flora," *Frontiers in Plant Science*, vol. 9, 2019, doi: 10.3389/fpls.2018.01929

[6] J.-J. Wilson, K.-W. Sing and N. Jaturas, "DNA Barcoding: Bioinformatics Workflows for Beginners," *Reference Module in Life Sciences*, vol. 3, pp. 985-995, 2019, doi: 10.1016/B978-0-12-809633-8.20468-8

[7] C. L. Schoch, S. Ciufo, M. Domrachev, C. L. Hotton, S. Kannan, R. Khovanskaya, D. Leipe, R. Mcveigh, K. O'Neill, B. Robbertse, S. Sharma, V. Soussov, J. P. Sullivan, L. Sun, S. Turner and I. Karsch-Mizrachi, "NCBI Taxonomy: a comprehensive update on curation, resources and tools," *Database (Oxford)*, vol. 2020, 2020, doi: 10.1093/database/baaa062

[8] J. DiFrisco, G. P. Wagner and A. C. Love, "Reframing research on evolutionary novelty and co-option: Character identity mechanisms versus deep homology," *Seminars in Cell & Developmental Biology*, 2022, doi: 10.1016/j.semcdb.2022.03.030

[9] C. C. d. M. Moura, F. Brambach, K. J. H. Bado, K. V. Krutovsky, H. Kreft, S. S. Tjitrosoedirdjo, I. Z. Siregar and O. Gailing, "Integrating DNA Barcoding and Traditional Taxonomy for the Identification of Dipterocarps in Remnant Lowland Forests of Sumatra," *Plants*, vol. 8, no. 11, p. 461, 2019, doi: 10.3390/plants8110461

[10] F. F. Nascimento, M. d. Reis and Z. Yang, "A biologist's guide to Bayesian phylogenetic analysis," *Nature Ecology & Evolution*, vol. 1, no. 10, pp. 1446-1454, 2017, doi: 10.1038/s41559-017-0280-x

[11] G. W. Hassler, A. F. Magee, Z. Zhang, G. Baele, P. Lemey, X. Ji, M. Fourment and M. A. Suchard, "Data Integration in Bayesian Phylogenetics," *Annual Review of Statistics and Its Application*, vol. 10, 2023, doi: 10.1146/annurev-statistics-033021-112532

[12] R. Bharti and D. G. Grimm, "Current challenges and best-practice protocols for microbiome analysis," *Briefings in Bioinformatics*, vol. 22, no. 1, pp. 178-193, 2021, doi: 10.1093/bib/bbz155

[13] J. Kelleher, Y. Wong, A. W. Wohns, C. Fadil, P. K. Albers and G. McVean, "Inferring whole-genome histories in large population datasets," *Nature Genetics*, vol. 51, no. 9, pp. 1330-1338, 2019, doi: 10.1038/s41588-019-0483-y

[14] P. Kapli, Z. Yang and M. J. Telford, "Phylogenetic tree building in the genomic age," *Nature Reviews Genetics*, vol. 21, no. 7, pp. 428-444., 2020, doi: 10.1038/s41576-020-0233-0

[15] B. J. Whalen, J. Ma and P. V. Balachandran, "A Bayesian Approach to the Eagar–Tsai Model for Melt Pool Geometry Prediction with Implications in Additive Manufacturing of Metals," *Integrating Materials and Manufacturing Innovation*, vol. 10, no. 4, pp. 597-609, 2021, doi: 10.1007/s40192-021-00238-z

[16] I. G. A. S. Wangiyana, "Comparation of Dendrogram and Cladogram Topology of Gyrinops versteegii and Others Gyrinops Member for Polyphasic Taxonomy," *Jurnal Silva Samalas*, vol. 2, no. 1, pp. 13-18, 2019.

[17] L. J. Kelly, R. J. Ryder and G. Clarté, "Lagged couplings diagnose Markov chain Monte Carlo phylogenetic inference," *Statistics*, arXiv, vol. 2108, p. 13328, 2021.

[18] L. Wang, S. Wang and A. Bouchard-Côté, "An Annealed Sequential Monte Carlo Method for Bayesian Phylogenetics," *Systematic Biology*, vol. 69, no. 1, p. 155–183, 2020, doi: 10.1093/sysbio/syz028

[19] X. Xia, *Sequence alignment*, in Bioinformatics and the Cell, Springer, 2018, pp. 33-75, doi: 10.1007/978-3-319-90684-3_2

[20] P. Wu, Y. Mo, M. Peng, T. Tang, Y. Zhong, X. Deng, F. Xiong, C. Guo, X. Wu, Y. Li, X. Li, G. Li, Z. Zeng and W. Xiong, "Emerging role of tumor-related functional peptides encoded by lncRNA and circRNA," *Molecular Cancer*, vol. 19, no. 1, pp. 1-14, 2020, doi: 10.1186/s12943-020-1147-3

[21] H. F. Abou-Shaara, "How can Sequence Alignment Method Affect the Phylogenetic Relationship Results?" *Scientific Papers: Animal Science and Biotechnologies*, vol. 53, no. 1, 2020.

[22] S. Saleh, K. C. Fatemeh Dorri, F. Kabeer, N. Rusk, T. Funnell, M. J. Williams, D. Lai, M. Andronescu, K. R. Campbell, A. McPherson,

S. Aparicio, A. Roth, S. Shah and A. Bouchard-Côte, "Cancer phylogenetic tree inference at scale from 1000s of single cell genomes," *bioRxiv*, 2020, doi: 10.1101/2020.05.06.058180

[23] M. Fourment, A. F. Magee, C. Whidden, A. Bilge, F. A. Matsen and V. N. Minin, "19 dubious ways to compute the marginal likelihood of a phylogenetic tree topology," *Systematic Biology,* vol. 69, no. vol. 2, pp. 209-220, 2020, doi: 10.1093/sysbio/syz046

[24] C. Zhang, J. P. Huelsenbeck and F. Ronquist, "Using Parsimony-Guided Tree Proposals to Accelerate Convergence in Bayesian Phylogenetic Inference," *Systematic Biology,* vol. 69, no. 5, pp. 1016–1032, 2020, doi: 10.1093/sysbio/syaa002

[25] A. Rambaut, A. J. Drummond, D. Xie, G. Baele and M. A. Suchard, "Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7," *Systematic Biology*, vol. 67, no. 5, p. 901–904, 2018, doi: 10.1093/sysbio/syy032

[26] S. H. Jayady and H. Antong, “Theme Identification using Machine Learning Techniques,” *Journal of Integrated and Advanced Engineering (JIAE)*, vol. 1, no. 2, pp. 123-134, 2021, doi: 10.51662/jiae.v1i2.24

[27] D. Romahadi, F. Anggara, R. P. Youlia, H. L. Habibullah, and H. Xiong, “Bayesian networks approach on intelligent system design for the diagnosis of heat exchanger,“ *SINERGI*, vol. 26, no. 2, pp. 127-136, 2020, doi: 10.22441/sinergi.2022.2.001

[28] M. A. Naranjo-Ortiz and T. Gabaldón, "Fungal evolution: major ecological adaptations and evolutionary transitions," *Biological Reviews*, vol. 94, no. 4, pp. 1443-1476, 2019, doi: 10.1111/brv.12510

[29] C. C. d. M. Moura, F. Brambach, K. J. H. Bado, K. V. Krutovsky, H. Kreft, S. S. Tjitrosoedirdjo, I. Z. Siregar and O. Gailing, "Integrating DNA Barcoding and Traditional Taxonomy for the Identification of Dipterocarps in Remnant Lowland Forests of Sumatra," *Plants*, vol. 8, no. 11, p. 461, 2019, doi: 10.3390/plants8110461