PAPER • OPEN ACCESS

Characterization of the chloroplast genome sequence of *Calophyllum inophyllum*, a bioenergy tree species, using Oxford Nanopore Technologies

To cite this article: A A E Kristianti et al 2024 IOP Conf. Ser.: Earth Environ. Sci. 1315 012077

View the article online for updates and enhancements.

You may also like

- Morphological characterization of Calophyllum Inophyllum as a biodiesel E Yuniastuti, I Anjani, Nandariyah et al.
- Optimized Conversion of Nyamplung Seeds Oil to Biodiesel Using Box-Behnken Response Surface Methodology (RSM) Z Ilham and M R A Mansor
- Effect of solvent polarity levels on separation of xanthone and coumarin from <u>Calophyllum inophyllum leaves extract</u>
 D F Susanto, S Hapsari, Z Trilutfiani et al.





DISCOVER how sustainability intersects with electrochemistry & solid state science research



This content was downloaded from IP address 3.20.238.187 on 03/05/2024 at 03:43

IOP Conf. Series: Earth and Environmental Science 1315 (2024) 012077

Characterization of the chloroplast genome sequence of Calophyllum inophyllum, a bioenergy tree species, using Oxford **Nanopore Technologies**

A A E Kristianti¹, F G Dwiyanti^{1,2}, R Pratama^{2,3}, Iskandar Z Siregar^{1,2*},

¹Department of Silviculture, Faculty of Forestry and Environment, IPB University, IPB Dramaga Campus, Dramaga, Bogor, West Java, 16680 Indonesia

² Molecular Science Research Group, Advanced Research Laboratory, IPB University, IPB Dramaga Campus, Dramaga, Bogor, West Java, 16680 Indonesia

³ Department of Biochemistry, Faculty of Mathematics and Natural Science, IPB University, IPB Dramaga Campus, Dramaga, Bogor, West Java, 16680 Indonesia

*Corresponding author: siregar@apps.ipb.ac.id

Abstract. Calophyllum inophyllum is a forest tree species that offers significant renewable energy benefits. However, the genetic information of these bioenergy tree species for the study of tree improvement is still limited. This study aimed to assemble the draft chloroplast genome using MinION Oxford Nanopore Technologies (ONT) and analyze the genetic relationship based on molecular markers for C. inophyllum. The study involved several steps: DNA extraction and isolation using the CTAB method, DNA library preparation, DNA sequencing using MinION Oxford Nanopore Technologies (ONT), and phylogenetic analysis based on selected DNA markers. The results showed that 115.1 Mb of high-quality DNA was successfully generated using ONT long-read sequencing, and 114.708 bp was annotated into partial genome chloroplast. The genetic relationship analysis using two selected DNA barcodes, namely rbcL and accD showed that the *accD* gene is more recommended for constructing a phylogenetic tree for the genus *Calophyllum* than the *rbcL* gene as it has a high bootstrap value (93%). Furthermore, the accD gene can also be considered a potential marker for further genetic analysis of C. inophyllum.

1. Introduction

The energy demand is increasing yearly in line with the increasing economic growth rate, population growth, and industrial development. Final energy consumption in Indonesia from 2000 to 2014 was still dominated by petroleum oil, petro solar oil, kerosene, and aviation fuel [1]. Thus, the national final energy demand by KS scenario will reach 433 MTOE in 2050 or increase with an average annual growth rate of 3.7% especially fossil energy demand [2]. In addition, Indonesia's oil reserves decreased by 3,530.43 MMSTB during 2015-2019, which shows a scarcity of fossil energy [3]. These data show that fuel consumption tends to increase and is continually needed, but resources are minimal, so it is important to have alternative fuel, such as biofuels.

Biodiesel can be produced as one of the alternative renewable energy sources to achieve national energy security and reduce greenhouse gas emissions. Calophyllum inophyllum known as tamanu or nyamplung in Indonesia is a non-edible oil seed producing tree that is naturally distributed in Southeast

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under licence by IOP Publishing Ltd 1

IOP Conf. Series: Earth and Environmental Science 1315 (2024) 012077

Asia, Australia, the South Pacific, Southern coastal India, and eastern Africa [4]. *C. inophyllum* grows along the coast and in lowland forests throughout Indonesia. The seed kernel of *C. inophyllum* can be used as raw material for biodiesel production, where the oil content is relatively high compared to other plants (*Schleichera oleosa* 55-70%, *Jatropha curcas* 40-70%, *Elaeis guineensis* 46-54%, and *Calophyllum* inophyllum 40-73%) [5]. *C. inophyllum* seed oil is a potentially environmentally friendly energy resource with a higher selling value than kerosene and castor oil without competing with food needs [6].

Genomic-wide selection (GS) is a new strategy utilizing high-density molecular markers developed for breeding selection [7]. The genomic estimated breeding value (GEBV) can be obtained from chromosome fragments or all marker effect values in the whole genome accumulated by GS. The breeding value of an individual can be estimated by combining marker and phenotypic information. Early prediction and selection of individuals can be carried out based on the breeding values to shorten generation intervals, improve selection accuracy, and save costs [8]. For GS to be applied to the genetic improvement of target species, reference genome sequencing is needed at an early stage. Long-read sequencing approaches with Oxford Nanopore Technologies (ONT) can assemble the chloroplast genome that contains a 10-30 kb pair of long inverted repeats, which can confound sequencing and assembling efforts when using short-read technologies [9].

The research of genomic sequencing is needed to study the breeding selection and establishment of new forests, genetic variation, and gene expression, and to construct the evolutionary development of organisms through phylogenetic analysis. Therefore, this study aimed to assemble the chloroplast genome draft using the MiniON sequencer Oxford Nanopore Technologies (ONT) and analyze the molecular marker-based genetic relationship for *C. inophyllum*. The data produced from this study is expected to be a reference for compiling the whole chloroplast genome through the genome resequencing process and reveal the molecular marker or gene potentially selected as the DNA barcode for *C. inophyllum*.

2. Materials and method

2.1. Plant material collection

Fresh leaves collection was carried out in November 2021 from one *C. inophyllum* individual mature tree growing in the Faculty of Forestry and Environment area of IPB University Dramaga Campus, Bogor, West Java, Indonesia (latitude -6.560452, longitude 106.728187). Leaves were preserved in zip plastic containing silica gel and then stored at -20°C in the freezer.

2.2. DNA extraction and quality test

The genomic DNA of *C. inophyllum* silica gel-dried leaves was extracted by using the modified cetyl trimethyl ammonium bromide (CTAB) method [10]. 100 mg of silica gel-dried leaves was ground manually into a fine powder using mortar, pestle, and liquid nitrogen. The fine leaf powder was then placed into a 2 mL microtube with an additional 1000 μ L of preheated CTAB buffer. Samples were then homogenized with the help of a vortex, followed by incubation at 65°C for 30 min using a water bath. The microtubes containing the sample was inverted every 10 min to mix evenly. When finished, 500 μ L of chloroform: isoamyl alcohol (CIA) 24:1 and 10 μ L of phenol were added into the microtube. The mixture was then centrifuged for 10 min at 10,000 rpm.

To separate and form supernatant layers, organic matter, and chloroform, the supernatant layer in the uppermost was transferred to a new 1.5 mL microtube using a micropipette. This process was repeated twice without phenol addition for the second repetition. Cold isopropanol was added to the supernatant (1:1) and 5 M NaCl with a supernatant ratio of 1:4. The mixture was stored in a freezer at -20°C overnight. The mixture was then centrifuged for 10 min at 10,000 rpm until precipitate formed. The mixture (aqueous solution) was discarded leaving only the DNA pellet and 500 μ L of 70% ethanol was subsequently added before centrifuging for 5 min at 10,000 rpm. The 70% ethanol was then discarded. This process was carried out twice. The DNA pellet was dried in a desiccator for 10 min so that the tube

The 4th International Conference on Tropical Silviculture		IOP Publishing
IOP Conf. Series: Earth and Environmental Science	1315 (2024) 012077	doi:10.1088/1755-1315/1315/1/01207

was completely clean of ethanol. In the final step, the DNA pellet was preserved by 50 μ L of TE buffer in the microtube and stored in the freezer at -20°C.

DNA quality test was carried out by 1% (w/v) agarose gel electrophoresis method at 100 V for 20 min. Moreover, the DNA concentration was also measured using a Qubit 1.0 Fluorometer.

2.3. DNA library preparation and sequencing

DNA library was prepared following the Nanopore Native Barcoding Genomic DNA (EXP-NBD104, EXP-NBD114, *and* SQK-LSK109) version NBE_9065_v109_revAK_14Aug2019 (https://community.nanoporetech.com/protocols/native-barcoding-genomic-

dna/checklist_example.pdf). The DNA library was inserted into the R9.4 MinION Flowcell, and sequencing was carried out using the MinKnow software developed by ONT.

2.4. Data analysis

2.4.1. Sequence raw data analysis. The output of MinION sequencing was raw Fast5 data, which was then converted into FastQ files using base-calling. This conversion involved translating the ionic signal into nucleotide sequences using Guppy Basecaller v4.2.3+8aca2af8 [11]. The FastQ data were further analyzed using NanoStat v1.2.1 to evaluate the quality and statistics of the reads. Additionally, NanoPlot v1.31.0 was employed to construct a distribution plot illustrating the read length and quality score [12][13]. All reads' quality was filtered using the NanoFilt v2.7.1 tool to obtain data with a Q score more than 7 and a base length greater than 500 bp [14]. The filtered reads were continued in the assembly process to compile reads into contigs using the Rebaler program combination (v0.2.0). The resulting contigs were polished using the Medaka program to obtain contigs with a higher degree of accuracy [15]. The Quast program was used to assess the corrected contig statistics, by referring to the chloroplast genome. Afterward, the polished contigs were annotated using the cloud server ChloroBox with the GeSeq tool [16].

2.4.2. Chloroplast marker analysis. The GenBank annotation produced by GeSeq was visualized using SnapGene v6.2.1 to identify the potential molecular marker [17]. Two coding gene sequences (*rbcL* and *accD*) were chosen for further analysis. These sequences were subjected to a Blastn search on the NCBI website, specifically targeting the *Calophyllum* genus. 50 homologous sequences of the *rbcL* marker and 12 homologous sequences of the *accD* marker were obtained and saved in a Fasta file for phylogenetic tree construction.

2.4.3. Phylogenetic tree construction. The MEGA-X program was utilized to create phylogenetic trees [18][19]. The process of aligning homolous sequences and markers was performed using ClustalW. The phylogenetic trees were generated by applying the Maximum-Likelihood algorithm to the aligned sequences. A bootstrap value of 1000 repetitions was used to assess the reliability of the phylogenetic tree. The completed phylogenetic tree was displayed in a circular format using iToL and modified using Inkscape.

3. Result and discussion

3.1. Genome sequencing and assembly

The initial stage of long-read sequencing involved conducting a quality check (QC) on the FastQ file using the Galaxy Cloud software. This was done after the base-calling procedure to assess the read length and quality of reads in raw data [20]. Based on the distribution (Figure 1), the read length values for *C. inophyllum* were generally spread from 0 bp to 10000 bp. The average quality of *C. inophyllum* sequences is in the range of Q8-Q15. The histogram also showed the most extended read length value of *C. inophyllum* around 103,794 bp. The highest quality reads (Q scores) were also obtained for both species, namely > Q25, with relatively few numbers. The data required for subsequent analysis requires a minimum read length of 500 bp and read quality > Q7 [21][22]. Following the filtering process, over 99.9% of reads successfully met the quality control criteria (30,794 reads) with a read length N50 of

12,503 bp and a total base of 115.1 Mb. Hence, the *C. inophyllum* sequence produced in this study is suitable for long-read sequencing.



Figure 1. Histogram of read length distribution data and average read quality for *Calophyllum inophyllum* sequences.

3.2. Chloroplast genome annotation

The purpose of genome annotation was to seek out functional genes across the genome sequence [23]. The length of *Calophyllum inophyllum* partial chloroplast genome was 114,708 bp or 114 kb, which was constructed from contig 57 (Figure 2). A total of 93 genes were identified throughout the genome, with 4 of these genes being duplicated in the invert repeat regions, resulting in a total number of 97 genes. The genetic material consisted of 1 rRNA gene, 25 tRNA genes, and 71 protein-coding genes. Several genes potentially used as DNA barcoding in cpDNA are *rbcL* and *accD* [24]. The requirement for a gene to be used as a suitable DNA barcoding is to have low intra-specific and high inter-specific variations [25]. Contig 57 *C. inophyllum* annotation results showed the presence of an *rbcL* gene of 1,428 bp in length and 43% GC, and an *accD* gene of 1,488 bp in length and 35% GC. This is following the statement of CBoL [26] that the length of the *rbcL* gene sequence is around 1,400 bp.

doi:10.1088/1755-1315/1315/1/012077

IOP Conf. Series: Earth and Environmental Science 1315 (2024) 012077



Figure 2. Partial assembly of the *Calophyllum inophyllum* chloroplast genome.

3.3. Phylogenetic Analysis

In this study, maximum likelihood phylogenetic approaches were used to construct the phylogenetic tree because this method typically includes trade-offs between accuracy and computing demand, with more accurate tree reconstruction necessitating deeper and time-consuming tree space exploration[27]. The maximum likelihood technique was adopted for multiple sequence alignment and cladogram tree visualization. The statistical probability of phylogram topology was established using bootstrap analysis, with the value computed based on 1000 bootstrap replications.

Two molecular markers namely *rbcL* and *accD* genes were selected for constructing phylogenetic tree of *C. inophyllum*. The Consortium for the Barcode of Life recommends the use of the *rbcL* gene as a reliable and efficient method for evaluating the genetic condition of a species with thoroughness, precision, and promptness [28]. The *rbcL* gene in cpDNA encodes RuBisCo and exhibits a high degree of species similarity due to its low mutation rate [29]. The length of the *rbcL* gene is around 1400 kb [30]. The *accD* gene is found across plants, including the shortened chloroplast genomes of parasitic and non-photosynthetic plants. The expression of *accD* was thought to be crucial throughout the embryo development stage in Arabidopsis and increased the total ACCase levels in the plastid. Acetyl-CoA carboxylase (ACCase) facilitates the conversion of acetyl-CoA into malonyl-CoA and is thought to regulate de novo fatty acid production [31][32].

A phylogenetic tree was constructed using the *rbc*L gene sequences from the chloroplast genome of *C. inophyllum* and 50 additional members of the *Calophyllum* genus, as reported in NCBI. The comparison of these sequences allowed for the construction of a phylogenetic tree (figure 3). In this study, the results of genetic relationship analysis using the *rbc*L gene with the maximum likelihood method of 1000 bootstraps on *C. inophyllum* showed very weak topological validity (39% in bootstrap value). The highest bootstrap value obtained by the *rbcL* gene for constructing the genus *Calophyllum* was only 59%. Typically, the bootstrap value was classified into four categories: very weak (<50%), weak (50-69%), moderate (70-85%), and high (>85%) [33]. The *rbcL* gene exhibits a high efficacy in amplifying gene fragments, but it also possesses significant limitations, such as limited accuracy in distinguishing between multiple closely related species [34]. Consequently, the *rbcL* gene proved to be inadequate in differentiating across genera.

IOP Conf. Series: Earth and Environmental Science 1315 (2024) 012077 doi:10.1088/1755-1315/1315/1/012077



Figure 3. Phylogenetic tree of *Calophyllum inophyllum* based on *rbc*L gene sequences.

The *rbcL* gene was not good enough in classifying variations between species in the *Calophyllum* genera. Due to the lack of study of chloroplast genome sequencing on the *Calophyllum* genus, the homology level of the nucleotide sequences obtained is not optimal. Therefore, other markers were needed that could be used as reference sequences and were available in the NCBI database. The *accD* marker was then chosen to construct the *C. inophyllum* phylogenetic tree because it was available in NCBI and had a high value of homology proportions.

The *acc*D gene is one of the genes in the chloroplast region that codes for the β -carboxyl transferase subunit of acetyl co-enzyme A carboxylase [35]. This gene plays a vital role in photosynthesis and affects plant leaf growth [36]. The *acc*D gene belongs to a more conservative coding region than the non-coding region. Other studies showed that *acc*D markers have regions that vary greatly in sequence and gene length [37][38].

IOP Conf. Series: Earth and Environmental Science 1315 (2024) 012077 doi:



Figure 4. Phylogenetic tree of Calophyllum inophyllum on accD gene sequences.

The phylogenetic tree of *C. inophyllum* based on *acc*D gene sequences showed that the studied *C. inophyllum* was in the same clade as the other *Calophyllum inophyllum* (Figure 4). The *C. inophyllum* sample used in this study was *C. inophyllum* native from Java Island. *C. inophyllum* from Java forms a monophyletic clade with sister group *C. inophyllum* from Tahiti Avera (KT369681.1 and KT369680.1) with a bootstrap value of 93%. *C. inophyllum* from Java also forms a monophyletic clade with *C. inophyllum* from Tahiti Patio (KT384377.1). The formation of a monophyletic clade in a cladogram indicates the presence of genetic material and the same characters from a common ancestor [39]. Several other studies have shown that *Calophyllum* is a genus of 190 species that is widely dispersed, primarily along coastlines in the Indo-Malesian region, Micronesia, Melanesia, northern Australia, Central and South America[40]. The presence of *C. inophyllum* in both Java and Tahiti (South Pacific) [24] suggest that the distribution of *C. inophyllum* in Java is likely due to a shared ancestor with similar chemical-genetic characteristics or structures, as indicated by its natural occurrence in both regions.

4. Conclusion

The genome of *Calophyllum inophyllum* was successfully sequenced using long-read DNA sequencing performed by MinION Oxford Nanopore Technologies (ONT), with a total base of 115.1 Mb and generated a 114,708 bp of partial chloroplast genome. In this study, the *rbcL* gene was not recommended for phylogenetic tree construction in the genus of *Calophyllum* considering its low bootstrap value. However, the phylogenetic tree based on the *accD* gene revealed that *C. inophyllum* in Java was closely related to *C. inophyllum* in the South Pacific with a high number of bootstrap values (93%). Furthermore, the *accD* gene can be considered a potential marker for further genetic analysis of *C. inophyllum*.

Acknowledgments

This study was supported by The National Competitive Basic Research 2022-2023 (Skema Penelitian Dasar Kompetitif Nasional 2023) No: 001/E5/PG.02.00PT/2022 and 001/E5/PG.02.00.PL/2023 entitled

" Pilot Sequencing of 100 Native Forest Tree Genomes to Support Ecosystem Restoration (*Rintisan Sekuensing 100 Genom Pohon Hutan Asli untuk Mendukung Restorasi Ekosistem*)" awarded by Ministry of Education, Culture, Research and Technology of the Republic of Indonesia to the corresponding author.

References

- Prasodjo E, Nurzaman H, Walujanto, Rosdiana D, Ismutadi P, Malik C, Santosa J, Nurrohim A, Widiastuti K D, Pambudi S H, Wibowo J L and Sauqi A 2016 *Indonesia Energy Outlook 2016* (Jakarta: Secretariat General of the National Energy Council of Republic of Indonesia)
- [2] Walujanto, Suharyati, Kristanto N, Widiastuti KD, Pambudi SH, Wibowo JL and Sauqi A 2017 Indonesia Energy Outlook (Jakarta: Secretariat General of the National Energy Council of Republic of Indonesia)
- [3] Adriawan A M, Hartanto R R, Novizar A, Aryani D, Syarifudin S, Saputra M Y, Asep H, Alida T, Ovan P D, Suciati, Ramona S E, Djamalulael, Septiana A, Ali A, Wiwin H, Imron M, Edward G, Bakhri A G, Andriani N, Laras R L, Jungjungan M, Alsa S, Fahmi M, Nastiti S, Beni H, Wisnu R, Taffaquh F, Laila N, Firman S, Apriyani D, Allia A V, Rozi F F, Rendhatya P, Octavia S A, Koesnobroto S and Prasetyo A 2019 *Statistik Minyak dan Gas Bumi 2019 Oil and Gas Statistics 2019* (Jakarta: Directorate General of Oil and Gas, Ministry of Energy and Mineral Resources of Republic of Indonesia)
- [4] Ansel J L, Lupo E, Mijouin L, Guillot S, Butaud J F, Ho R, Lecellier G, Raharivelomanana P and Pichon C 2016 Biological activity of Polynesian *Calophyllum inophyllum* oil extract on human skin cells *Planta Med* 82 p 961–966
- [5] Qadariyah L, Bhuana D S, Selaksa R, Shodiq J A and Mahfud M 2018 Biodiesel production from *Calophyllum inophyllum* using base lewis catalyst *ASEAN Journal of Chemical Engineering* 18 p 53–59
- [6] Susila I W W 2018 Nyamplung tanaman multifungsi potensi penyebarannya di NTB dan Bali (Yogyakarta: PT Kanisius)
- [7] Meuwissen T H, Solberg T R, Shepherd R and Woolliams J A 2009 A fast algorithm for BayesB type of prediction of genome-wide estimates of genetic value *Genetics Selection Evolution* 41 p 1–10
- [8] Jonas E and De Koning D-J 2013 Does genomic selection have a future in plant breeding? *Trends in Biotechnology* 31 (9) p 1–8
- [9] Wang W, Schalamun M, Morales-Suarez A, Kainer D, Schwessinger B and Lanfear R 2018 Assembly of chloroplast genomes with long- and short-read data: A comparison of approaches using *Eucalyptus pauciflora* as a test case *BMC Genomics* 19 (977) p 1-15
- [10] Doyle J and Hortorium B 1991 DNA protocols for plants CTAB total DNA isolation *Molecular Techniques in Taxonomy* vol 57, ed G M Hewitt, A W B Johnston and J P W Young (Springer Berlin, Heidelberg) p 283–284
- [11] Amarasinghe S L, Su S, Dong X, Zappia L, Ritchie M E and Gouil Q 2020 Opportunities and challenges in long-read sequencing data analysis *Genome Biol* **21** (30) p 1-16
- [12] Urban L, Holzer A, Baronas J J, Hall M, Braeuninger-Weimer P, Scherm M J, Kunz D J, Perera S N, Martin-Herranz D E, Tipper E T, Salter S J and Stammnitz M R 2021 Freshwater monitoring by nanopore sequencing *Elife* 10 p 1–27
- [13] De Coster W, D'Hert S, Schultz D T, Cruts M and Van Broeckhoven C 2018 NanoPack: Visualizing and processing long-read sequencing data *Bioinformatics* **34** p 2666–2669

- sequences using Oxford Nanopore Technology from Diospyros celebica genomic DNA BMC Res Notes 14 p 1–4
- [15] Oxford Nanopore Technologies 2018 Medaka: Sequence correction provided by ONT Research
- [16] Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones E S, Fischer A, Bock R and Greiner S 2017 GeSeq - Versatile and accurate annotation of organelle genomes *Nucleic Acids Res* 45 W6–11
- [17] Wahyuni F D, Saraswati H and Dewi K S 2020 In-silico analysis for cryI gene amplification from Bacillus thuringiensis Bioedukasi 18 p 8–14
- [18] Hall B G 2013 Building phylogenetic trees from molecular data with MEGA Mol Biol Evol 30 p 1229–1235
- [19] Kumar S, Stecher G, Li M, Knyaz C and Tamura K 2018 MEGA X: Molecular evolutionary genetics analysis across computing platforms *Mol Biol Evol* **35** p 1547–1549
- [20] Blankenberg D and Hillman-Jackson J 2014 Analysis of next-generation sequencing data using galaxy *Methods in Molecular Biology* 1150 p 21–43
- [21] Delahaye C and Nicolas J 2021 Sequencing DNA with nanopores: Troubles and biases *PLoS One* **16** (10) p 1-29
- [22] Tyler A D, Mataseje L, Urfano C J, Schmidt L, Antonation K S, Mulvey M R and Corbett C R 2018 Evaluation of Oxford Nanopore's MinION sequencing device for microbial whole genome sequencing applications *Sci Rep* 8 p 1–12
- [23] Indayati I and Purwanto H 2021 Controlling Culex quinquefasciatus Say, 1823 (Diptera: Culicidae) using several Lysinibacillus sphaericus isolates endogenic to Indonesia Jurnal Biologi Tropis 21 p 298–304
- [24] Ginigini J, Lecellier G J, Nicolas M, Nour M, Hnawia E, Lebouvier N, Herbette G, Lockhart P and Raharivelomanana P 2019 Chemodiversity of *Calophyllum inophyllum* L. oil bioactive components related to their specific geographical distribution in the South Pacific region *PeerJ* 7 (e6896) p 1–21
- [25] Hebert P D N, Cywinska A, Ball S L and DeWaard J R 2003 Biological identifications through DNA barcodes *Proceedings of the Royal Society B: Biological Sciences* 270 p 313–321
- [26] CBOL Plant Working Group 2009 A DNA barcode for land plants Proceedings of The National Academy of Science 106 p 12794–12797
- [27] De Maio N, Kalaghatgi P, Turakhia Y, Corbett-Detig R, Minh B Q and Goldman N 2023 Maximum likelihood pandemic-scale phylogenetics *Nat Genet* 55 p 746–752
- [28] Hidayatul Mursyidin D and Irfan Makruf M 2020 Keanekaragaman dan kekerabatan genetik Artocarpus berdasarkan penanda DNA kloroplas matK & rbcL: kajian in silico Floribunda 6 p 188–194
- [29] Dobrogojski J, Adamiec M and Luciński R 2020 The chloroplast genome: a review Acta Physiol Plant 42 p 1–13
- [30] Nurhasanah, Sundari and Papuangan N 2019 Amplification and analysis of *rbcL* Gene (ribulose-1,5-bisphosphate carboxylase) of clove in Ternate Island *IOP Conference Series: Earth and Environmental Science* vol 276 p 1–7
- [31] Konishi T and Sasaki Y 1994 Compartmentalization of two forms of acetyl-CoA carboxylase in plants and the origin of their tolerance toward herbicides (acctyl-CoA carboxylase loafzatlon/blotln enzyme/eukaryote form/gr odde/prokayote form) *Proceedings of the National Academy of Sciences* 91 p 3598–3601

- [32] Sasaki Y and Nagano Y 2004 Plant acetyl-CoA carboxylase: structure, biosynthesis, regulation, and gene manipulation for plant breeding *Biosci Biotechnol Biochem* **68** p 1175–1184
- [33] Lestari D A, Azrianingsih R and Hendrian 2018 Filogenetik jenis-jenis Annonaceae dari Jawa Timur koleksi Kebun Raya Purwodadi berdasarkan coding dan non-coding sekuen DNA J Trop Biodivers Biotechnol 3 p 1–7
- [34] Cabelin V L D and Alejandro G J D 2016 Efficiency of matK, rbcL, trnH-psbA, and trnL-F (cpDNA) to molecularly authenticate Philippine ethnomedicinal Apocynaceae through DNA barcoding Pharmacogn Mag 12 S384–388
- [35] Wakasugi T, Tsudzuki T and Sugiura M 2001 The genomics of land plant chloroplasts: gene content and alteration of genomic information by RNA editing *Photosynth Res* **70** p 107–118
- [36] Li J, Su Y and Wang T 2018 The repeat sequences and elevated substitution rates of the chloroplast *accD* gene in *Cupressophytes Front Plant Sci* **9** (533) p 1–11
- [37] Liu Y, Huo N, Dong L, Wang Y, Zhang S, Young H A, Feng X and Gu Y Q 2013 Complete Chloroplast Genome Sequences of Mongolia Medicine Artemisia frigida and Phylogenetic Relationships with Other Plants *PLoS One* 8 (2) p 1–14
- [38] Chen J, Hao Z, Xu H, Yang L, Liu G, Sheng Y, Zheng C, Zheng W, Cheng T and Shi J 2015 The complete chloroplast genome sequence of the relict woody plant *Metasequoia glyptostroboides* Hu et Cheng *Front Plant Sci* 6 (447) p 1–11
- [39] Hidayat T, Kusumawaty D, Kusdianti, Yati D D, Muchtar A A and Mariana D 2008 Analisis Filogenetik Molekuler pada Phyllanthus niruri L. (Euphorbiaceae) Menggunakan Urutan Basa DNA Daerah Internal Transcribed Spacer (ITS) Jurnal Matematika dan Sains 13 p 16–21
- [40] Nurtjahjaningsih I L G, Prihatini I, Qiptiyah M and Rimbawanto A 2020 Genetic Structure of Calophylluminophyllum L. In Dompu Natural Forest IOP Conference Series: Earth and Environmental Science vol 522