Check for updates

GENOME NOTE

Long-read datasets of Kayu Bawang (*Azadirachta excelsa* (Jack) Jacobs) from 3 different identified seed stands in Bengkulu (Sumatra), Indonesia and its phylogenetic relationships

[version 1; peer review: 2 approved with reservations]

Efratenta Katherina Depari[1,2], Nurheni Wijayanto[3], Lina Karlinasari[4], Mohamad Rafi[5,6], Iskandar Zulkarnaen Siregar[3,5]

[1]Department of Forestry, Faculty of Agriculture, University of Bengkulu, Bengkulu, Bengkulu, 38371, Indonesia
[2]Tropical Silviculture Study Program, Faculty of Forestry and Environment, IPB University, Bogor, West Java, 16680, Indonesia
[3]Department of Silviculture, Faculty of Forestry and Environment, IPB University, Bogor, West Java, 16680, Indonesia
[4]Department of Forest Product, Faculty of Forestry and Environment, IPB University, Bogor, West Java, 16680, Indonesia
[5]Advanced Research Laboratory, IPB University, Bogor, West Java, 16680, Indonesia
[6]Department of Chemistry, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, West Java, 16680, Indonesia

v1 **First published:** 05 Apr 2024, **13**:254
https://doi.org/10.12688/f1000research.144155.1
Latest published: 05 Apr 2024, **13**:254
https://doi.org/10.12688/f1000research.144155.1

Abstract

Abstract*

Azadirachta excelsa (Jack) Jacobs, Kayu bawang (Meliaceae) is economically valuable and widely used by the local community in Bengkulu (Sumatra) as carpentry and construction wood because of its good durability class. However, it still has ambiguous scientific multiple names, such as *Azadirachta excelsa*, *Protium javanicum*, and *Dysoxylum mollissimum.* Additional tools such as molecular approaches can be used to verify whether it is true or not that Kayu bawang is scientifically named as *Azadirachta excelsa* based on previous morphological identification. This study aimed to construct draft chloroplast genome and verify the scientific name based on molecular identification using a single *rbc*L gene marker. Genomic DNA was extracted from bark cambium originated from three different provenances in Bengkulu, Indonesia, namely TBT-A, TBT-K, and TBT-S. MinION from Oxford Nanopore Technologies was used to sequence the samples following manufacture protocols SQK-LSK109 yielding 481.6 Mb for TBT-A, 597.4 Mb for TBT-K, and 853.1 Mb for TBT-S, respectively. Generated data were assembled and constructed, namely 58,780 bp (14 tRNAs and 47 encoding genes) for TBT-A, 142,139 bp (4 rRNAs, 24 tRNAs, and 78 encoding genes) for TBT-K, and

Open Peer Review

Approval Status ? ?

	1	2
version 1 05 Apr 2024	? view	? view

1. **Hoang Dang Khoa Do**, Nguyen Tat Thanh University, Ho Chi Minh City, Vietnam

2. **Maria Stefanie Dwiyanti**, Hokkaido University, Sapporo, Japan

Any reports and responses or comments on the article can be found at the end of the article.

84,906 bp (24 tRNAs and 53 encoding genes) for TBT-S. Based on the phylogenetic tree, *Azadirachta excelsa* from three identified tree stands were placed in the same group with other *Azadirachta excelsa* accessions.

Keywords
Azadirachta excelsa, chloroplast genome, identification, identified seed stand, long-reads, phylogenetics, gene marker

This article is included in the Plant Science gateway.

This article is included in the Genomics and Genetics gateway.

This article is included in the Nanopore Analysis gateway.

Corresponding author: Iskandar Zulkarnaen Siregar (siregar@apps.ipb.ac.id)

Author roles: Depari EK: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Visualization; **Wijayanto N**: Investigation, Supervision, Writing – Review & Editing; **Karlinasari L**: Supervision, Validation, Writing – Review & Editing; **Rafi M**: Resources, Supervision, Writing – Review & Editing; **Siregar IZ**: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: Beasiswa Pendidikan Pascasarjana Dalam Negeri (BPPDN) 2019

Copyright: © 2024 Depari EK *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Depari EK, Wijayanto N, Karlinasari L *et al.* **Long-read datasets of Kayu Bawang (*Azadirachta excelsa* (Jack) Jacobs) from 3 different identified seed stands in Bengkulu (Sumatra), Indonesia and its phylogenetic relationships [version 1; peer review: 2 approved with reservations]** F1000Research 2024, **13**:254 https://doi.org/10.12688/f1000research.144155.1

First published: 05 Apr 2024, **13**:254 https://doi.org/10.12688/f1000research.144155.1

Introduction

Azadirachta excelsa (Jack) Jacobs or known as Kayu bawang (Meliaceae) is a main local wood in Bengkulu Province that is cultivated in private forests, mainly found in North Bengkulu and Central Bengkulu Districts (Depari *et al.* 2015; Siahaan & Sumadi 2015). Kayu bawang has been used for carpentry wood. Historically, this wood has been a favorite one of the local community as the primary material for building. In addition, the wood is also known for its resistance to attack by destructive organisms (Krisdianto *et al.* 2015). Despite kayu bawang is widely used for building wood and furniture, but some references are still ambiguous with respect to its scientific naming. Some references mentioned the naming of kayu bawang tree in Bengkulu as *Protium javanicum* Burm F (Adfa *et al.* 2013). In some other references, kayu bawang is named differently such as *Dysoxylum mollissimum* Blume (Ishiguri *et al.* 2016), *Azadirachta excelsa* (Jack) M Jacobs (Premono & Lestari 2014). In addition, in the identified seed stand (TBT) in Bengkulu as registered under No. 002/BPSMBK/SSB/2016, No. 001/BPTH.I-3/SSB/2016, and No. 006/P2STH/SSB/2018, kayu bawang is named respectively as *Azadirachta exelsa*, *Dysoxylum molissimum*, and *Melia excelsa* (synonym of *Azadirachta excelsa*). Based on morphological identification on Herbarium Bogoriense under National Research and Innovation Agency (BRIN), kayu bawang has been confirmed as *Azadirachta excelsa* (Jack) Jacobs. We used three specimens from different identified seed stands, namely TBT-A, TBT-K, and TBT-S, that were collected with the specimen numbers BO1991351 (TBT-A), BO1991352;1991353;1991354 (TBT- K), and BO1991350 (TBT-S) for genome sequencing using Oxford Nanopore Technologies. The objectives of the research were to assembly long-read DNA sequences and to construct a phylogeny for verifying species identification.

Methods

Genomic DNA was extracted from kayu bawang (*Azadirachta excelsa*) cambium samples which were taken from three different identified seed stands namely Talang Boseng (TBT-A, -3.66890°S, 102.292122°E), Batu Ampar (TBT-K, -4.394333°S, 103.138056°E), and Penyangkak (TBT-S, -3.531333°S, 102.252861°E) in Bengkulu Province, Indonesia. The modified CTAB (cetyl trimethyl ammonium bromide) method was used to extract the genomic DNA from each *A. excelsa* cambium (Doyle & Doyle 1987). The DNA extract buffer contained 200 µL of 10% CTAB (CAT NO. 194004, LOT NO. QR15741, MP Biomedicals, LLC), 100 µL of 1 M Tris-HCL, 280 µL of 5M NaCl, 40 µL of 0.5 M EDTA, 100 µL of 1% PVP, 280 µL of aquades. 0.1 g cambium mashed using mortar milling was put into a 2 mL tube, 1000 µL of DNA extract buffer that had been incubated at 65°C (60 min), 40 µL of 26% PVP, and 40 µL of β-Mercaptoethanol. The mixture was homogenized using a vortex and subsequently incubated at 65°C (60 min) with rehomogenization performed every 10 min by gently flipping the tube. The cooled mixture was centrifuged for 10 min, at 10,000 rpm. The supernatant was transferred to a 1.5 mL tube, and 500 µL of chloroform: isoamyl alcohol solution (24:1) and 10 µL of phenol were added and centrifuged for 10 min, at 10,000 rpm. Separation of the mixture was done again by adding 500 µL of chloroform: isoamyl alcohol solution (24:1). The obtained supernatant was put into a 1.5 mL tube, then 500 µL cold isopropanol and 250 µL NaCl were added. The mixture was stored in a freezer at 4°C for 1 day so that the DNA precipitation process occurring. After DNA precipitation, the mixture was centrifuged for 10 min, at 10,000 rpm. The mixture was removed gradually to retrieve the DNA pellet. DNA pellet washing using 500 µL of 70% ethanol, centrifuged for 10 min, at 10,000 rpm. The DNA pellet was separated from the mixture and dried in a desiccator containing silica gel (±20 min). 50 µL of TE RNAse buffer was added to dissolve the DNA pellet. The quality of extracted genomic DNA was evaluated using 1% agarose gel electrophoresis. The quality and quantity of the extracted genomic DNA were measured using NanoPhotometer® NP80 (IMPLEN) and Qubit 1.0 Fluorometer (Thermo Fisher Scientific) with the Qubit dsDNA BR assay kit. The library preparation of long-read sequencing was performed using Oxford Nanopore Technologies (ONT) manufactured protocol SQK-LSK109 version NBE_9065_v109_revZ_14Aug2019 with the MinION R9.4.1 flow cell (FLO-MIN106D, LOT 11001352) on a MinION Mk1C sequencer (Oxford Nanopore Technologies).

The chloroplast genome assembly for TBT-A, TBT-K, and TBT-S (Accession number DRA015537; https://ddbj.nig.ac.jp/resource/bioproject/PRJDB14301) (IPB University 2023) was performed using Galaxy Server (https://usegalaxy.eu/, RRID: SCR_006281, MIT License) v.23.1.1.dev0 following long-read data with some modifications (Wang *et al.* 2018). Long-read data for each sample was quality checked using NanoPlot (https://usegalaxy.eu/, RRID:SCR_024128, MIT License) v1.41.0 (De Coster *et al.* 2018). Those clean reads data were aligned with the reference NC_023792.1 (*Azadirachta indica*) using Minimap2 (https://usegalaxy.eu/, RRID:SCR_018550, MIT License) v.2.26 (Li 2018) and SAMTools (https://usegalaxy.eu/, RRID:SCR_002105, MIT License) v.2.05 (Li *et al.* 2009). Chloroplast reads were assembled using Canu Assembler (https://usegalaxy.eu/, RRID:SCR_015880, GPLv2 and others) v.2.1.1+galaxy0 (Koren *et al.* 2017) involving parameters such as technology Nanopore, estimated genome size 160k, and min read length 1000. Medaka Consensus Pipeline (https://usegalaxy.eu/, RRID:SCR_005857, Mozilla Public License 2.0) v1.7.2 +galaxy1 was used to polish the assembly result (Oosterbroek *et al.* 2021). Geseq (https://chlorobox.mpimp-golm.mpg.de/geseq.html, RRID:SCR_017336, MIT License) (Tillich *et al.* 2017) was used to annotate the assembled chloroplast

genome using all Meliaceae reference in NCBI RefSeq and visualized using OGDRaw (https://chlorobox.mpimp-golm.mpg.de/OGDraw.html, RRID:SCR_017337, MIT License) in the CHLOROBOX (https://chlorobox.mpimp-golm.mpg.de/index.html, MIT License) (Greiner *et al.* 2019). Phylogenetic analysis was performed using MEGAX (https://www.megasoftware.net/, RRID:SCR_023471, MIT License) v10.2.2 (Kumar *et al.* 2018) with Maximum Likelihood method, Tamura-3 model and bootstrap value of 1000 replications. *rbc*L gene marker was used to construct the phylogenetic tree from TBT-A, TBT-K, and TBT-S samples (*Owenia vernicocsa* (DQ238063.1) as an outgroup) (Muellner *et al.* 2006).

Results

The library yielded 481.6 Mb for TBT-A, 597.4 Mb for TBT-K, and 853.1 Mb for TBT-S. Based on the statistics, TBT-A has 3,339 bp of N50, 3,103 bp for TBT-K, and 3,882 bp for TBT-S (Figure 1). Chloroplast draft genomes of the TBT-A sample (36% GC content) produces 58,780 bp containing 14 tRNAs and 47 encoding genes, TBT-K sample (37% GC content) produced 142,139 bp containing 4 rRNAs, 24 tRNAs, and 78 encoding genes, also TBT-S sample (36% GC

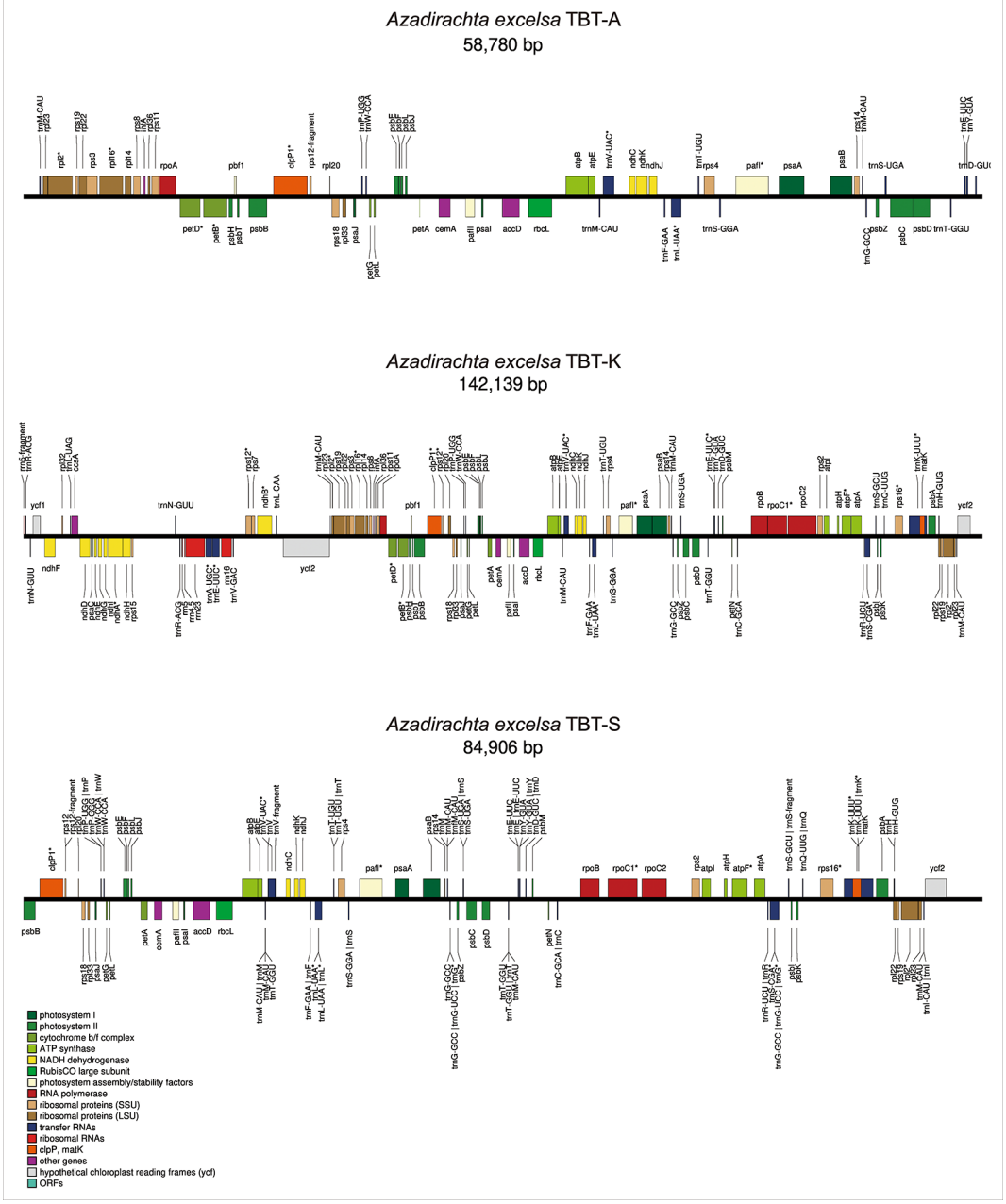


Figure 1. Chloroplast draft genome of *Azadirachta excelsa* (TBT-A, TBT-K, and TBT-S samples).

Table 1. List of genes in the chloroplast draft genomes of *Azadirachta excelsa*.

Gene groups	TBT-A genes	TBT-K genes	TBT-S genes
RNAs transfer	*trn*D-GUC, *trn*E-UUC, *trn*F-GAA, *trn*G-GCC, *trn*L-UAA*, *trn*M-CAU, *trn*P-UGG, *trn*S-GGA, *trn*S-UGA, *trn*T-GGU, *trn*T-UGU, *trn*V-UAC*, *trn*W-CCA, *trn*Y-GUA	*trn*A-UGC*, *trn*C-GCA, *trn*D-GUC, *trn*E-UUC[d]*, *trn*F-GAA, *trn*G-GCC, *trn*H-GUG, *trn*K-UUU*, *trn*L-CAA, *trn*L-UAA*, *trn*L-UAG, *trn*M-CAU, *trn*N-GUU, *trn*P-UGG, *trn*Q-UUG, *trn*R-ACG, *trn*R-UCU, *trn*S-CGA[d]*, *trn*T-GGU, *trn*T-UGU, *trn*V-GAC, *trn*V-UAC[d]*, *trn*W-CCA, *trn*Y-GUA	*trn*C-GCA, *trn*D-GUC, *trn*E-UUC[d], *trn*F-GAA, *trn*fM-CAU, *trn*G-GCC*, *trn*H-GUG, *trn*I-CAU, *trn*K-UUU*, *trn*L-UAA*, *trn*M-CAU, *trn*P-GGG, *trn*P-UGG, *trn*Q-UUG, *trn*R-UCU, *trn*S-CGA*, *trn*S-GCU, *trn*S-GGA, *trn*S-UGA, *trn*T-GGU, *trn*T-UGU[d], *trn*V-UAC*, *trn*V-fragment, *trn*W-CCA, trnY-GUA
RNAs ribosomal		*rrn*4.5, *rrn*5, *rrn*5-fragment, *rrn*16, *rrn*23	
Small subunit ribosomal proteins (SSU)	*rps*3, *rps*4, *rps*8, *rps*11, *rps*12-fragment, *rps*14, *rps*18, *rps*19	*rps*2, *rps*3, *rps*4, *rps*7, *rps*8, *rps*11, *rps*12*, *rps*14, *rps*15, *rps*16*, *rps*18, *rps*19	*rps*2, *rps*4, *rps*12, *rps*12-fragment, *rps*14, *rps*16[d]*, *rps*18, *rps*19
Large subunit ribosomal proteins (LSU)	*rpl*2*, *rpl*14, *rpl*16*, *rpl*20, *rpl*22, *rpl*23, *rpl*33, *rpl*36	*rpl*2[d]*, *rpl*14, *rpl*16*, *rpl*20, *rpl*22, *rpl*23, *rpl*32, *rpl*33, *rpl*36	*rpl*2[d]*, *rpl*20, *rpl*22, *rpl*23, *rpl*33
RNA polymerase	*rpo*A	*rpo*A, *rpo*B, *rpo*C2, *rpo*C1*	*rpo*B, *rpo*C1*, *rpo*C2
ATP synthase	*atp*B, *atp*E	*atp*A, *atp*B, *atp*E, *atp*F*, *atp*H, *atp*I	*atp*A, *atp*B, *atp*E, *atp*F[d]*, *atp*H, *atp*I
Subunit of Photosystem I	*psa*A, *psa*B, *psa*I, *psa*J	*psa*A, *psa*B, *psa*C, *psa*I, *psa*J	*psa*A, *psa*B, *psa*I, *psa*J
Subunit of Photosystem II	*psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*J, *psb*L, *psb*T, *psb*Z	*psb*A, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*H, *psb*I, *psb*J, *psb*K, *psb*L, *psb*M, *psb*T, *psb*Z	*psb*A, *psb*B, *psb*C, *psb*D, *psb*E, *psb*F, *psb*I, *psb*J, *psb*K, *psb*L, *psb*M, *psb*Z
Subunit of rubisco	*rbc*L	*rbc*L	*rbc*L
NADH dehydrogenase	*ndh*C, *ndh*J, *ndh*K	*ndh*A*, *ndh*B*, *ndh*C, *ndh*D, *ndh*E, *ndh*F, *ndh*G, *ndh*H, *ndh*I, *ndh*J, *ndh*K,	*ndh*C, *ndh*J, *ndh*K
Subunit of cytochrome b/f complex	pet*A*, *pet*B*, *pet*D*, *pet*G, *pet*L	*pet*A, *pet*B*, *pet*D*, *pet*G, *pet*L, *pet*N	*pet*A, *pet*G, *pet*L, *pet*N
Photosystem assembly factors	*paf*I**, *paf*II	*paf*II, *paf*I**	*paf*I[d]**, *paf*II
Cytochrome c synthesis		*ccs*A	
Photosystem biogenesis factor	*pbf*1	*pbf*1	
Transitional initiation factor	*inf*A	*inf*A	
ATP-dependent protease subunit P	*clp*P1**	*clp*P1**	*clp*P1**
Fatty metabolic acid	*acc*D	*acc*D	*acc*D
Carbon metabolism	*cem*A	*cem*A	*cem*A
Maturase		*mat*K	*mat*K
Conserved open reading frames		*ycf*1, *ycf*2	*ycf*2

[d]Gene duplication.
*Single intron.
**Double intron.

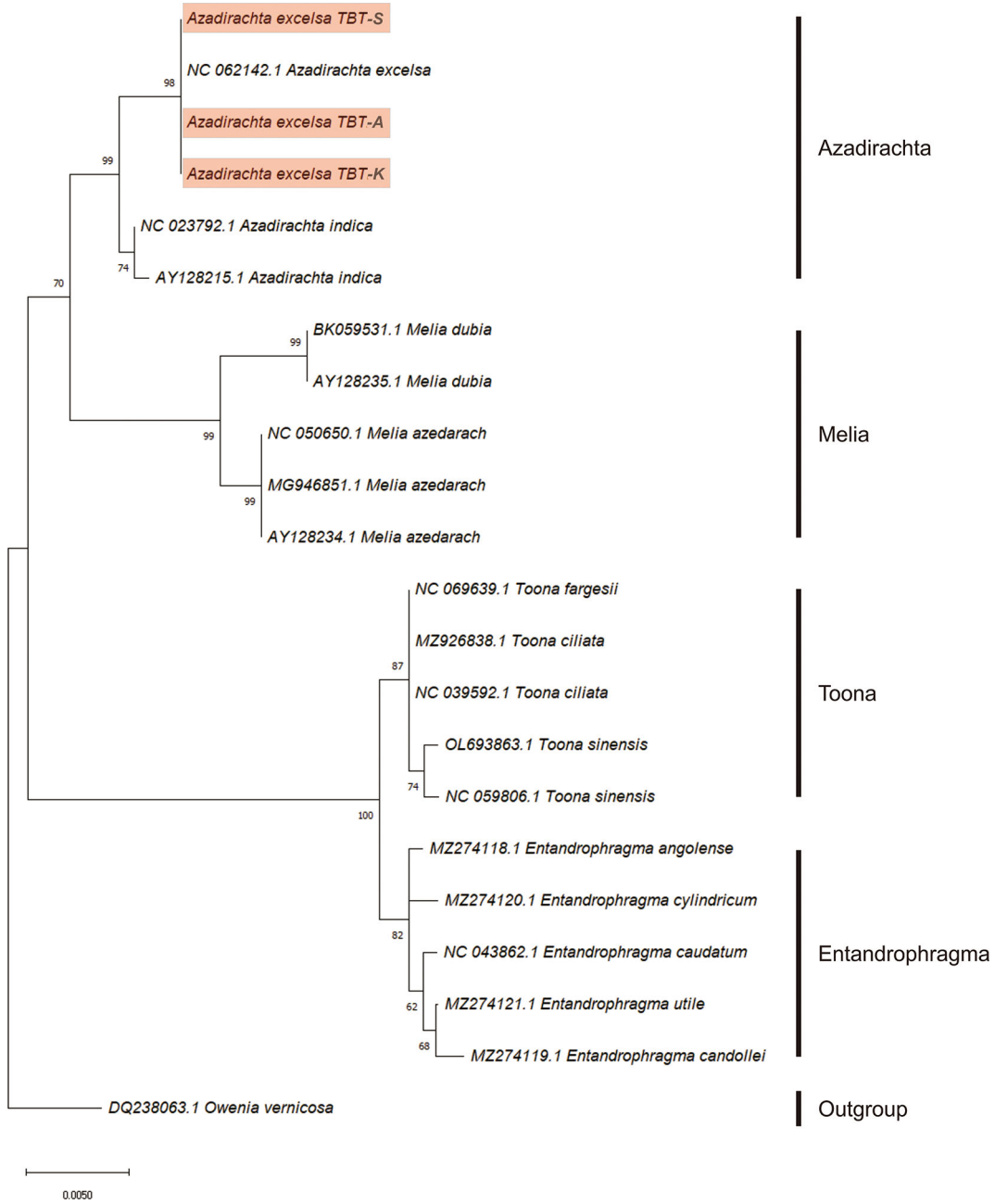


Figure 2. Phylogenetic tree based on Maximum Likelihood method. GenBank accession numbers are listed beside the scientific name. The bootstrap value of 1000 replications is shown next to the nodes. *Azadirachta excelsa* (TBT-A, TBT-K, and TBT-S samples) are marked in coloring red.

content) produced 84,906 bp containing 24 tRNAs and 53 encoding genes (Table 1). Based on the constructed phylogenetic tree, *Azadirachta excelsa* from three identified tree stands (TBT-A, TBT-K, and TBT-S) were placed in the same group with other *Azadirachta excelsa* accessions (Figure 2). We can confirm that all those samples belong to *Azadirachta excelsa.*

Author roles

Efratenta Katherina Depari: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Resources, Visualization, Writing-Original Draft;

Nurheni Wijayanto: Investigation, Supervision, Writing-Review & Editing;

Lina Karlinasari: Supervision, Validation, Writing-Review & Editing;

Rafi M: Resources, Supervision, Writing-Review & Editing;

Iskandar Zulkarnaen Siregar: Conceptualization, Funding Acquisition, Methodology, Resources, Supervision, Validation, Writing-Review & Editing.

Data availability

Underlying data

DNA Data Bank of Japan (DDBJ): Characterization of kayu bawang in agroforestry system in Bengkulu: species name, wood quality, and metabolite profile. Accession number DRA015537; https://ddbj.nig.ac.jp/resource/bioproject/PRJDB14301 (IPB University 2023).

This project contains the following underlying data:

- DRR437487 (TBT-A).
- DRR437488 (TBT-K).
- DRR437489 (TBT-S).

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Software availability statement

Galaxy server available from: https://usegalaxy.eu/ (MIT license).

CHLOROBOX available from: https://chlorobox.mpimp-golm.mpg.de/index.html (MIT license).

MEGAX available from: https://www.megasoftware.net/ (MIT license).

Acknowledgements

Special thanks to the Forest Genetics and Molecular Forestry Laboratory, Department of Silviculture, Faculty of Forestry and Environment, IPB University and Molecular Science Laboratory, Advanced Research Laboratory, IPB University, Bogor, Indonesia, for providing lab facilities during this study.

References

Adfa M, Hattori Y, Ninomiya M, *et al.*: **Chemical constituents of Indonesian plant *Protium javanicum* Burm. f. and their antifeedant activities against *Coptotermes formosanus* Shiraki.** *Nat. Prod. Res.* 2013; **27**(3): 270–273.
PubMed Abstract | Publisher Full Text

De Coster W, D'Hert S, Schultz DT, *et al.*: **NanoPack: visualizing and processing long-read sequencing data.** *Bioinformatics.* 2018; **34**(15): 2666–2669.
PubMed Abstract | Publisher Full Text | Free Full Text

Depari EK, Wiryono W, Susatya A: **Potensi tegakan kayu bawang (*Dysoxylum mollissimum* Blume) pada sistem agroforestri sederhana di Kabupaten Bengkulu Utara.** *Jurnal Hutan Tropis.* 2015; **3**(2).

Doyle JJ, Doyle JL: **A rapid DNA isolation prosedure for small quantities of fresh leaf tissue.** *Phytochemical Bulletin.* 1987; **19**(1): 11–15.

Greiner S, Lehwark P, Bock R: **OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes.** *Nucleic Acids Res.* 2019; **47**(W1): W59–W64.
PubMed Abstract | Publisher Full Text | Free Full Text

IPB University: *Characterization of kayu bawang in agroforestry system in Bengkulu: species name, wood quality, and metabolite profile.* DDBJ Sequence Read Archive. 2023.

Ishiguri F, Aiso H, Hirano M, *et al.*: **Effects of radial growth rate on anatomical characteristics and wood properties of 10-year-old (*Dysoxylum mollissimum*) trees planted in Bengkulu, Indonesia.** *Tropics.* 2016; **25**(1): 23–31.
Publisher Full Text

Koren S, Walenz BP, Berlin K, *et al.*: **Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation.** *Genome Res.* 2017; **27**(5): 722–736.
PubMed Abstract | Publisher Full Text | Free Full Text

Krisdianto, Dewi L, Muslich M: **Analisis hasil pengujian kayu yang diserang penggerek kayu di laut dengan interpretasi gambar digital.** *Jurnal Penelitian Hasil Hutan.* 2015; **33**(1): 11–18.
Publisher Full Text

Kumar S, Stecher G, Li M, *et al.*: **MEGA X: molecular evolutionary genetics analysis across Computing Platforms.** *Mol. Biol. Evol.* 2018; **35**(6): 1547–1549.
PubMed Abstract | Publisher Full Text | Free Full Text

Li H: **Minimap2: pairwise alignment for nucleotide sequences.** *Bioinformatics.* 2018; **34**(18): 3094–3100.
PubMed Abstract | Publisher Full Text | Free Full Text

Li H, Handsaker B, Wysoker A, *et al.*: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics.* 2009; **25**(16): 2078–2079.
PubMed Abstract | Publisher Full Text | Free Full Text

Muellner AN, Savolainen V, Samuel R, *et al.*: **The mahogany family "out-of-Africa": Divergence time estimation, global biogeographic**

patterns inferred from plastid *rbc*L DNA sequences, extant, and fossil distribution of diversity. *Mol. Phylogenet. Evol.* 2006; **40**(1): 236–250.
PubMed Abstract | Publisher Full Text

Oosterbroek S, Doorenspleet K, Nijland R, *et al.*: **Decona: From demultiplexing to consensus for Nanopore amplicon data.** *ARPHA ARPHA Conf. Abstr.* 2021; **4**.
Publisher Full Text

Premono B, Lestari S: **Karakteristik petani dan praktek silvikultur agroforestri kayu bawang (*Azadirachta excelsa* (Jack) M. Jacobs) di Kabupaten Bengkulu Utara.** *Jurnal Penelitian Hutan Tanaman.* 2014; **11**(3): 185–197.
Publisher Full Text

Siahaan H, Sumadi A: **Site index prediction of smallholder plantations of kayu bawang (*Disoxylum mollissimum* Blume) in Bengkulu Province.** *Indones. J. For. Res.* 2015; **2**(2): 81–92.
Publisher Full Text

Tillich M, Lehwark P, Pellizzer T, *et al.*: **GeSeq – versatile and accurate annotation of organelle genomes.** *Nucleic Acids Res.* 2017; **45**(W1): W6–W11.
PubMed Abstract | Publisher Full Text | Free Full Text

Wang W, Schalamun M, Morales-Suarez A, *et al.*: **Assembly of chloroplast genomes with long- and short-read data: a comparison of approaches using *Eucalyptus pauciflora* as a test case.** *BMC Genomics.* 2018; **19**(1): 977.
PubMed Abstract | Publisher Full Text | Free Full Text

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 09 May 2024

https://doi.org/10.5256/f1000research.157909.r269396

© **2024 Dwiyanti M.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Maria Stefanie Dwiyanti
Laboratory of Applied Plant Genome, Hokkaido University, Sapporo, Japan

This study sequenced three samples of Kayu bawang (*Azadirachta excelsa)* obtained from three stands in Bengkulu Province, Indonesia. Since there is no genomic information about *A. excelsa,* this will help science community to use the information for sample authentication and further genetic analysis. The authors have provided detailed DNA extraction, sequencing and data analysis procedures to be replicated by others. Since the authors want to prove that Kayu bawang is indeed *A. excelsa*, I will suggest to add *Protium* and *Dysoxylum* rbcL sequences in the phylogenetic analysis to show that the difference between *A. excelsa* and the other two species. In addition, the chloroplast sequence sizes obtained from the three samples are different. If there is any estimated chloroplast genome sequence size for *A. excelsa*, or closely related species, the information can be added as estimation for the completeness of *A. excelsa* chloroplast genome sequences obtained in this study.

Are the rationale for sequencing the genome and the species significance clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of the sequencing and extraction, software used, and materials provided to allow replication by others?
Yes

Are the datasets clearly presented in a usable and accessible format, and the assembly and annotation available in an appropriate subject-specific repository?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: My research area is plant genetics especially soybean and wild soybean, analysis of whole genome sequence, genotyping-by-sequencing, and SNP genotyping, GWAS and QTL analysis using SNP data.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 06 May 2024

https://doi.org/10.5256/f1000research.157909.r266729

© 2024 Do H. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Hoang Dang Khoa Do
Nguyen Tat Thanh University, Ho Chi Minh City, Ho Chi Minh, Vietnam

Dear Authors,
Thank you very much for adding new genomic data of *Azadirachta excelsa* (Jack) Jacobs, generated by long-read sequencing method of Oxford Nanopore Technologies.
The content of the manuscript is suitable for F1000research.
Please check my below comments.
1/ Why the chloroplast genomes of three samples were not completed? Please add some reason for this issue in the main text.
2/ There are several genes in the draft genomes. However, only rbcL was used for phylogenetic analysis. Please add the reason in the main text.
3/ In the Abstract, the authors mentioned that "However, it still has ambiguous scientific multiple names, such as *Azadirachta excelsa*, *Protium javanicum*, and *Dysoxylum mollissimum*.". However, in the phylogenetic analysis, the authors used only "*Azadirachta excelsa*" for comparison. It is better if the authors add the data of Protium and Dysoxylum species to make a strong and reliable conclusion for the identification of three samples in this study.
4/ The draft genomes were constructed from three samples of one species. Therefore, the authors might add information about the pairwise identity of those genomes.
Thank you.

Are the rationale for sequencing the genome and the species significance clearly described?
Yes

Are the protocols appropriate and is the work technically sound?
Yes

Are sufficient details of the sequencing and extraction, software used, and materials provided to allow replication by others?

Yes

Are the datasets clearly presented in a usable and accessible format, and the assembly and annotation available in an appropriate subject-specific repository?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Genomics, Plant systematics.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research