MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 4, 541–545
https://doi.org/10.1080/23802359.2024.2345773

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome of *Pandanus amaryllifolius* Roxb. ex Lindl. (Pandanaceae) and its phylogenetic relationship

Jun Han Chin[a], Xue Jing Wong[a], Teek Foh Chong[b], Piemjit Muangkot[d], Aik Teng Heng[b], Tawatchai Tanee[c,d] and Shiou Yih Lee[a,e]

[a]Faculty of Health and Life Sciences, INTI International University, Nilai, Malaysia; [b]School of Health Management, INTI International University, Nilai, Malaysia; [c]Faculty of Environment and Resource Studies, Mahasarakham University, Maha Sarakham, Thailand; [d]One Health Research Unit, Mahasarakham University, Maha Sarakham, Thailand; [e]Centre for Health, Well-being, and Environmental Sustainability, INTI International University, Nilai, Malaysia

ABSTRACT

Pandanus amaryllifolius of Pandanaceae, a plant native to Southeast Asia, has been domesticated for its health benefits and aromatic leaves. It is also used for phytoremediation and soil rehabilitation. However, genetic studies of this species are limited. This study aims to expand its genomic information by assembling and characterizing the complete chloroplast genome of *P. amaryllifolius*. The chloroplast genome, which was 157,839 bp long, contains a total of 133 genes, including 87 protein-coding (CDS), 38 tRNA, and eight rRNA genes. The overall G/C content was 37.7%. A phylogenetic analysis using 79 shared unique CDS revealed a monophyletic relationship in Pandanales. Based on the limited sampling size, *Pandanus amaryllifolius* was the first to diverge in Pandanaceae. The genomic data will be useful for future phylogenetic and evolutionary studies of Pandanaceae.

ARTICLE HISTORY
Received 22 January 2024
Accepted 16 April 2024

KEYWORDS
Aromatic plant; genetic resources; next-generation sequencing; Pandanales; phylogenomics

Introduction

Pandanus amaryllifolius Roxb. ex Lindl. 1829 of Pandanaceae is a tropical plant species found in various tropical peninsular countries (Wakte et al. 2012). Approximately 600 species of *Pandanus* are recorded at present (POWO 2024); however, only *P. amaryllifolius* has scented leaves (Bhuyan and Sonowal 2021). The plant is easily propagated and cultivated as a spice, with medicinal properties such as antiviral, antioxidant, anticancer, hemagglutination, and antimicrobial activities (Dewanti and Sofian 2017). The scent of the leaf is also useful in aromatherapy applications (Pradopo et al. 2017). As the plant could mitigate nutrient pollution in phytoremediation systems (Han et al. 2014), performs biosorption of heavy metals, and improves soil microbial homeostasis (Ngadi et al. 2014; Zhong et al. 2022), farmers in Thailand cultivate *P. amaryllifolius* to reduce metal accumulation and generate income (Nakbanpote et al. 2024).

Research on the genetic background of *P. amaryllifolius* is limited, with only one report using DNA fingerprinting techniques (Wakte et al. 2012). A phylogenetic analysis with *P. amaryllifolius* included, using the concatenated chloroplast DNA dataset, i.e. *mat*K, *trn*L-*trn*F, and *trn*Q-*rps*16, revealed an unresolved relationship in *Pandanus* (Buerki et al. 2012). The lack of informative sites in the dataset failed to reveal the interspecific relationship in *Pandanus*. As the complete genome sequence data was proposed to be able to provide a better resolution in the phylogenetic relationship of complex plant genera when compared to short gene sequences (Lee et al. 2022), to provide information on the genetic identity and molecular of this species, we sequenced and assembled the complete chloroplast genome of *P. amaryllifolius*.

Materials and methods

Fresh leaves of *P. amaryllifolius* were obtained from an herb collection planted in the TCM herb garden of the Faculty of Health and Life Sciences, INTI International University (2°48′41″N, 101°45′29″E) (Figure 1). A voucher specimen of the sample has been deposited in the Molecular Science Lab of INTI International University under the collection number LSY006 (https://newinti.edu.my, Dr. Lee Shiou Yih, shiouyih.lee@newinti.edu.my).

Total genomic DNA extraction was carried out using the FavorPrep™ Plant Genomic DNA Extraction Mini Kit (Favorgen, China), according to the manufacturer's protocol. A genomic library was prepared with an insert size of approximately 350 bp using the TruSeq DNA Sample Prep Kit (Illumina, USA). Next-generation sequencing was carried out

CONTACT Tawatchai Tanee tawatchai5@hotmail.com Faculty of Environment and Resource Studies, Mahasarakham University, Maha Sarakham 44150, Thailand; Shiou Yih Lee shiouyih.lee@newinti.edu.my Faculty of Health and Life Sciences, INTI International University, Nilai, Negeri Sembilan 71800, Malaysia

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2345773.

Figure 1. *Pandanus amaryllifolius*. (A) whole plant with long, narrow leaves growing in a spiral arrangement; (B) aerial roots emerging from the nodes at the lower portion of the stem and extend downwards. Photo taken by S.Y. Lee.

on the Illumina Novaseq (Illumina, USA) platform, which 150-bp paired-end raw reads were generated. The raw data were assembled using NOVOWrap v1.20 (Wu et al. 2021), with the *rbc*L sequence of *P. amaryllifolius* (GenBank accession number: MZ422731) selected as the seed sequence. The depth coverage analysis was carried out by mapping the raw NGS data against the assembled genome sequence using Geneious Prime v.2022.0.2 (Kearse et al. 2012). Gene annotation was conducted using GeSeq v2.03 (Tillich et al. 2017) by activating the BLAT search and Chloë v.1.1.0 functions. No genome reference was included. The result was manually checked for errors. The annotated chloroplast genome and the structure of the genes that are difficult to annotate, including the cis-splicing and trans-splicing genes, were visualized using CPGView (Liu et al. 2023).

To reveal the phylogenetic placement of *P. amaryllifolius*, based on the availability of genome data in the GenBank database, the 79 shared unique CDS sequences extracted from the complete chloroplast genome sequence of 10 taxa of Pandanales were used for phylogenetic reconstruction. Two outgroup taxa were selected, including *Dioscorea glabra* (Diocoreaceae, Dioscoreales; GenBank accession number: OL638497; Wonok et al. 2023) and *Lilium souliei* (Liliaceae, Liliales; GenBank accession number: ON409199; Li et al. 2022). The CDS genes were aligned prior to concatenation. Phylogenetic analysis was conducted based on the maximum likelihood (ML) and Bayesian inference (BI) methods, using the RAxML v8 (Stamatakis 2014) and the MrBayes v3.2 (Ronquist et al. 2012) pipelines that are available in the CIPRES Portal v3.1 (Miller et al. 2010), respectively. For ML, the tree was constructed using the general-time reversible (GTR) with gamma distribution (+G) (=GTR + G) nucleotide substitution model, coupled with 1000 bootstrap replicates. For BI, a total of 2,000,000 generations were used in the Markov chain Monte Carlo analysis, and reads were sampled at every 100 cycles. The results were visualized using FigTree v1.4.4 (Rambaut 2018).

Results

With a minimum read mapping depth of 161× and an average read mapping depth of 1464.9× (Supplementary Figure 1), the chloroplast genome sequence of *P. amaryllifolius* was 157,839 bp in length (GenBank accession number: ON920703). The chloroplast genome comes in a typical quadripartite structure that contains two inverted repeats (IRs) of 26,951 bp, a large single copy (LSC) region of 85,746 bp, and a small single-copy (SSC) region of 18,191 bp. A total of 133 genes were annotated, including 87 protein-coding, 38 tRNA, and eight rRNA genes (Figure 2). Among them, 12 CDS were cis-splicing genes, of which one contained two introns and 11 contained one intron (Supplementary Figure 2(A)). The gene structure of the trans-splicing gene, *rps*12, was also identified (Supplementary Figure 2(B)). The overall GC content of the chloroplast genome was 37.7%.

Both the ML and BI trees revealed identical topology, so only the ML tree was displayed (Figure 3). The phylogenetic relationship was well-resolved with bootstrap support (BS) and posterior probabilities (PP) greater than 75% and 0.95, respectively. In Pandanales, a monophyletic relationship was observed among the four selected families; Pandanaceae is closely related to Cyclanthaceae. Based on the current sampling, in the Pandanceae clade, *P. amaryllifolius* was the first to diverge, followed by *Benstonea copelandii*, *P. odorifer*, and *P. tectorius*.

Discussion

The size of the chloroplast genome of *P. amaryllifolius* is similar to that of *P. odorifer* (GenBank accession number: OK648473; Darshetkar et al. 2022), but smaller when compared to *P. tectorius* (GenBank accession numbers: MH748568; Tan et al. 2019). The overall GC content of the chloroplast genome of *P. amaryllifolius* is similar to the

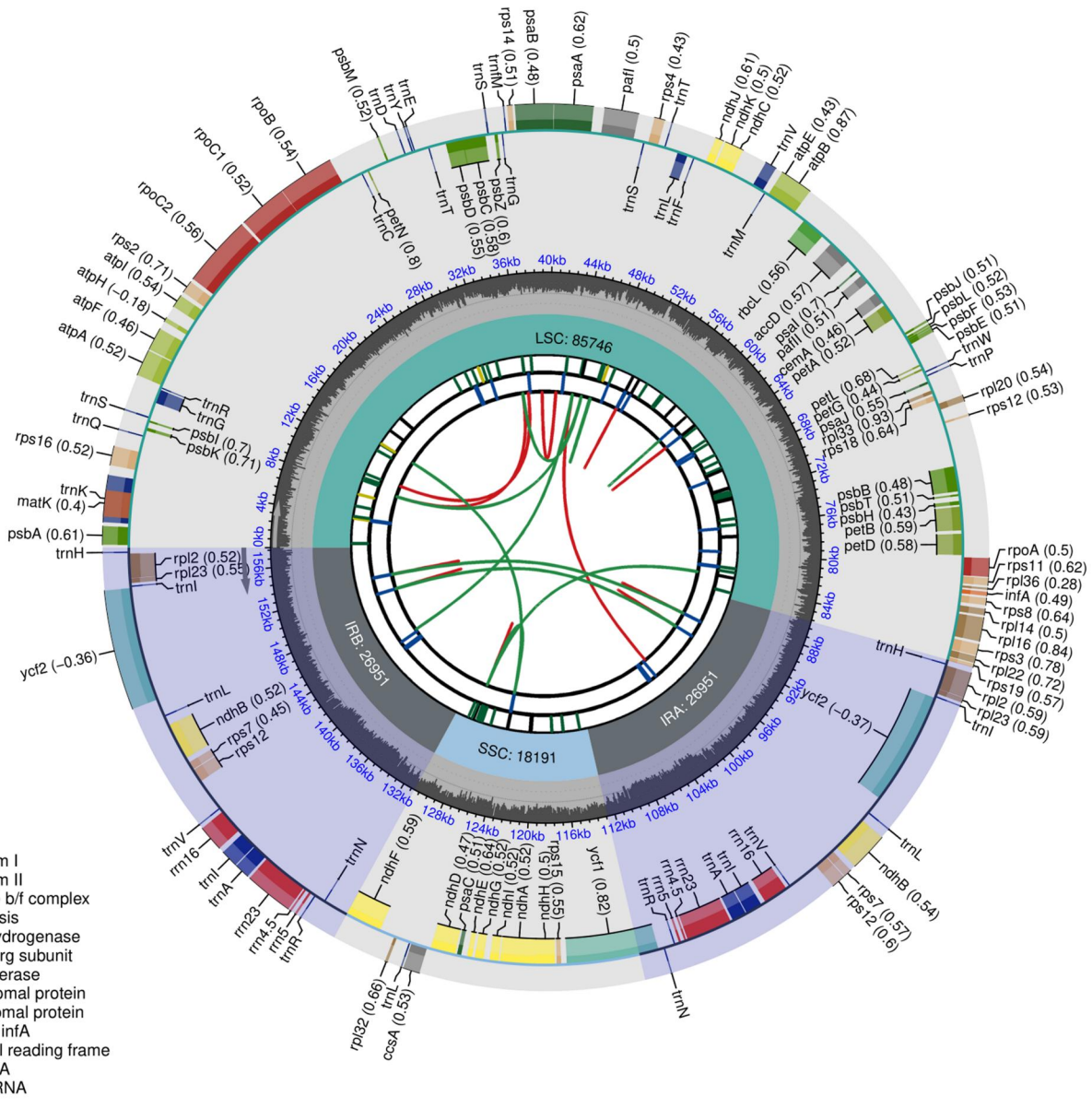

Figure 2. Complete chloroplast genome map of *Pandanus amaryllifolius*. From the center outward, the first track shows the dispersed repeats, in which the forward (F) and palindromic (P) repeats are connected with red and green arcs. The second track shows the long tandem repeats as short blue bars. The third track shows the short tandem repeats or microsatellite sequences as short bars with different colors that correspond to their repeat unit size: Black: complex repeat; green: repeat unit size = 1; yellow: repeat unit size = 2; purple: repeat unit size = 3; blue: repeat unit size = 4; orange: repeat unit size = 5; red: repeat unit size = 6. The small single-copy, inverted repeat, and large single-copy regions are shown on the fourth track. The GC content along the genome is plotted on the fifth track. The genes are shown on the sixth track, while the optional codon usage bias is displayed in the parenthesis after the gene name. Genes are color-coded by their functional classification (bottom left corner), while the transcription directions for the inner and outer genes are clockwise and anticlockwise, respectively.

chloroplast genome of *P. odorifer*, but greater than that of *P. tectorius* (37.6%) (Tan et al. 2019; Darshetkar et al. 2022).

To date, there are only four complete chloroplast genome sequences of *Pandanus* that are publicly available. The current sampling size suggests *Pandanus* is paraphyletic, contradicting previous findings using short chloroplast gene sequence datasets (Callmander et al. 2012; Buerki et al. 2016). The complex relationship between *Benstonea* and *Pandanus* has been explored, with distinct morphology and genetic variations; many of the species of *Pandanus* were transferred to *Benstonea* (Callmander et al. 2012, 2013, 2016). Despite molecular placement of *P. amaryllifolius* suggesting it might be a member of *Benstonea*, the small sample size used in this study may not have allowed for an exact genetic relationship (Hillis et al. 2003). As a resolved phylogenetic tree was reconstructed in this study, it reveals the potential of the complete chloroplast genome sequences in resolving the complex relationship in *Pandanus*. Therefore, to accurately reveal the molecular placement of *P. amaryllifolius*, it is suggested to increase the sampling size to better represent these complicated genera.

Conclusion

The study characterized the complete chloroplast genome of *P. amaryllifolius* and performed phylogenetic analysis to determine its molecular position. Despite limited reports on the chloroplast genome of Pandanaceae, the findings suggest

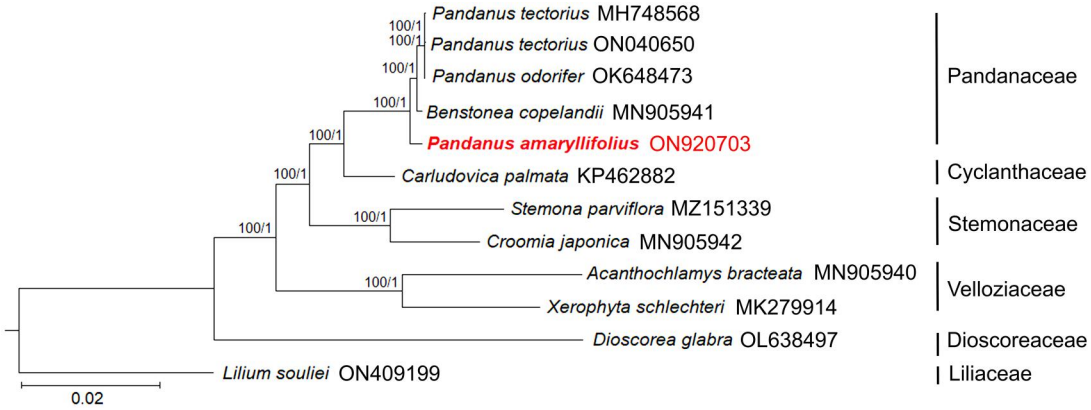

Figure 3. Phylogenetic tree based on the 79 shared unique CDS of 10 selected taxa of Pandanales, with *Dioscorea glabra* (Diocoreaceae, Dioscoreales; GenBank accession number: OL638497; Wonok et al. 2023) and *Lilium souliei* (Liliaceae, Liliales; GenBank accession number: ON409199; Li et al. 2022) included as outgroups. For the maximum likelihood, the GTR + G substitution model was employed, and branch supports were calculated under 1,000 bootstrap replicates; for the Bayesian inference, a 4-by-4 nucleotide model was applied, and MCMC was performed under 2,000,000 generations with sampling taken at every 100 cycles. The bootstrap support (left) and posterior probability (right) values are indicated at each branch node. The following sequences were used: *Acanthochlamys bracteata* (GenBank accession number: MN905940; Wanga et al. 2021), *Benstonea copelandii* (GenBank accession number: MN905941; Soto-Gomez et al. 2020), *Carludovica palmata* (GenBank accession number: KP462882; Mennes et al. 2015), *Croomia japonica* (GenBank accession number: MN905942; Soto-Gomez et al. 2020), *Dioscorea glabra* (GenBank accession number: OL638497; Wonok et al. 2023), *Lilium souliei* (GenBank accession number: ON409199; Li et al. 2022), *Pandanus amaryllifolius* (GenBank accession number: ON920703; this study), *Pandanus odorifer* (GenBank accession number: OK648473; Darshetkar et al. 2022), *Pandanus tectorius* (GenBank accession number: MH748568, ON040650; Tan et al. 2019), *Stemona parviflora* (GenBank accession number: MZ151339; Wei and Li 2022), and *Xerophyta schlechteri* (GenBank accession number: MK279914; Wang and Lanfear 2019).

the potential of the complete chloroplast genome sequences in reconstructing a well-resolved phylogenetic tree. Nevertheless, the genome data obtained from this study could significantly enhance our understanding of the phylogenetics and evolutionary patterns within Pandanaceae.

Author contributions

ATH, TFC, TT, SYL conception and design; JHC, XJW, PM, SYL analysis and interpretation of the data; JHC, XJW, PM drafting of the paper; ATH, TFC, TT, SYL critical revision; all the authors approved the final version; and all authors agree to be accountable for all aspects of the work.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the INTI IU Research Seeding Grant under Grant [number INTI-FHLS-01-14-2022] and Mahasarakham University.

Permission for sample collection

Pandanus amaryllifolius is not a protected plant. Permission for sample collection at the TCM Garden of INTI International University is acknowledged by the Director of the School.

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at http://www.ncbi.nlm.nih.gov under the accession number ON920703. The associated BioProject, SRA, and BioSample numbers are PRJNA853926, SRR19895280, and SAMN29416416, respectively.

References

Bhuyan B, Sonowal R. 2021. An overview of *Pandanus amaryllifolius* Roxb. ex Lindl. and its potential impact on health. Curr Trends Pharm Res. 8(1):138–157.

Buerki S, Callmander MW, Devey DS, Chappell L, Gallaher T, Munzinger J, Haevermans T, Forest F. 2012. Straightening out the screwpines: a first step in understanding phylogenetic relationships within Pandanaceae. Taxon. 61(5):1010–1020. doi:10.1002/tax.615008.

Buerki S, Gallaher T, Booth T, Brewer G, Forest F, Pereira JT, Callmander MW. 2016. Biogeography and evolution of the screw-pine genus *Benstonea* Callm. & Buerki (Pandanaceae). Candollea. 71(2):217–229. doi:10.15553/c2016v712a8.

Callmander MW, Booth TJ, Beentje H, Buerki S. 2013. Update on the systematics of *Benstonea* (Pandanaceae): when a visionary taxonomist foresees phylogenetic relationships. Phytotaxa. 112(2):57–60. doi:10.11646/phytotaxa.112.2.4.

Callmander MW, Keim AP, Heatubun CD, Homot P, Buerki S. 2016. Lifting the curtain on our knowledge on New Guinean *Benstonea* (Pandanaceae). Phytotaxa. 275(2):168–174. doi:10.11646/phytotaxa.275.2.8.

Callmander MW, Lowry PP, Forest F, Devey DS, Beentje H, Buerki S. 2012. *Benstonea* Callm. & Buerki (Pandanaceae): characterization, circumscription, and distribution of a new genus of screw-pines, with a synopsis of accepted species. Candollea. 67(2):323–345. doi:10.15553/c2012v672a12.

Darshetkar AM, Patil SS, Pable AA, Nadaf AB, Barvkar VT. 2022. Chloroplast genome sequence of *Pandanus odorifer* (Forssk.) Kuntze: genome features, mutational hotspots and phylogenetic analyses. Biologia. 77(12):3397–3412. doi:10.1007/s11756-022-01155-y.

Dewanti NI, Sofian FF. 2017. Review artikel: Aktivitas farmakologi ekstrak daun pandan wangi (*Pandanus amaryllifolius* Roxb.). Farmaka. 15(2):186–193.

Han P, Kumar P, Ong BL. 2014. Remediation of nutrient-rich waters using the terrestrial plant, *Pandanus amaryllifolius* Roxb. J Environ Sci (China). 26(2):404–414. doi:10.1016/s1001-0742(13)60426-x.

Hillis DM, Pollock DD, McGuire JA, Zwickl DJ. 2003. Is sparse taxon sampling a problem for phylogenetic inference? Syst Biol. 52(1):124–126. doi:10.1080/10635150390132911.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowits S, Duran C, et al. 2012. Geneious Basic:

an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12): 1647–1649. doi:10.1093/bioinformatics/bts199.

Lee SY, Turjaman M, Chaveerach A, Subasinghe S, Fan Q, Liao W. 2022. Phylogenetic relationships of *Aquilaria* and *Gyrinops* (Thymelaeaceae) revisited: evidence from complete plastid genomes. Bot J Linn Soc. 200(3):344–359. doi:10.1093/botlinnean/boac014.

Li Y, Zhang L, Wang T, Zhang C, Wang R, Zhang D, Xie Y, Zhou N, Wang W, Zhang H, et al. 2022. The complete chloroplast genome sequences of three lilies: genome structure, comparative genomic and phylogenetic analyses. J Plant Res. 135(6):723–737. doi:10.1007/s10265-022-01417-5.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

Mennes CB, Lam VK, Rudall PJ, Lyon SP, Graham SW, Smets EF, Merckx VS. 2015. Ancient Gondwana break-up explains the distribution of the mycoheterotrophic family Corsiaceae (Liliales). J Biogeogr. 42(6):1123–1136. doi:10.1111/jbi.12486.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE). New Orleans: IEEE. p. 1–8.

Nakbanpote W, Kutrasaeng N, Jitto P, Prasad MNV. 2024. Potential of ornamental plants for phytoremediation and income generation. In Prasad MNV, editors. Bioremediation and bioeconomy. 2nd ed. Amsterdam: Elsevier. p. 211–256.

Ngadi N, Yahya NY, Abd Rahman R. 2014. Sustainable removal of heavy metal using extracted *Pandanus amaryllifolius* Roxb. AMM. 695:28–31. doi:10.4028/www.scientific.net/AMM.695.28.

POWO. 2024. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. http://www.plantsoftheworldonline.org/ [accessed 2024 March 20]

Pradopo S, Sinaredi BR, Januarisca BV. 2017. Pandan leaves (*Pandanus amaryllifolius*) aromatherapy and relaxation music to reduce dental anxiety of pediatric patients. J Int Dental Med Res. 10(3):933–937.

Rambaut A. 2018. FigTree, a graphical viewer of phylogenetic trees (Version 1.4.4). Edinburgh, UK: Institute of evolutionary biology, University of Edinburgh.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 61(3):539–542. doi:10.1093/sysbio/sys029.

Soto-Gomez M, Lin Q, da Silva Leal E, Gallaher TJ, Scherberich D, Mennes CB, Smith SY, Graham SW. 2020. A bi-organellar phylogenomic study of Pandanales: inference of higher-order relationships and unusual rate-variation patterns. Cladistics. 36(5):481–504. doi:10.1111/cla.12417.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

Tan XH, Wang JH, Zhao KK, Zhu ZX, Wang HF. 2019. Complete plastome sequence of *Pandanus tectorius* Parkinson ex Du Roi (Pandanaceae). Mitochondrial DNA Part B. 4(1):68–69. doi:10.1080/23802359.2018.1536460.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11. doi:10.1093/nar/gkx391.

Wakte KV, Zanan RL, Saini A, Jawali N, Thengane RJ, Nadaf AB. 2012. Genetic diversity assessment in *Pandanus amaryllifolius* Roxb. populations of India. Genet Resour Crop Evol. 59(7):1583–1595. doi:10.1007/s10722-012-9882-y.

Wang W, Lanfear R. 2019. Long-reads reveal that the chloroplast genome exists in two distinct versions in most plants. Genome Biol Evol. 11(12):3372–3381. doi:10.1093/gbe/evz256.

Wanga VO, Dong X, Oulo MA, Mkala EM, Yang JX, Onjalalaina GE, Gichua MK, Kirika PM, Gituru RW, Hu GW, et al. 2021. Complete chloroplast genomes of *Acanthochlamys bracteata* (China) and *Xerophyta* (Africa) (Velloziaceae): comparative genomics and phylogenomic placement. Front Plant Sci. 12:691833. doi:10.3389/fpls.2021.691833.

Wei R, Li Q. 2022. The complete chloroplast genome of endangered species *Stemona parviflora*: insight into the phylogenetic relationship and conservation implications. Genes (Basel). 13(8):1361. doi:10.3390/genes13081361.

Wonok W, Sudmoon R, Tanee T, Lee SY, Chaveerach A. 2023. Complete chloroplast genome of four Thai native *Dioscorea* species: structural, comparative and phylogenetic analyses. Genes (Basel). 14(3):703. doi:10.3390/genes14030703.

Wu P, Xu C, Chen H, Yang J, Zhang X, Zhou S. 2021. NOVOWrap: an automated solution for plastid genome assembly and structure standardization. Mol Ecol Resour. 21(6):2177–2186. doi:10.1111/1755-0998.13410.

Zhong Y, Zhang A, Qin X, Yu H, Ji X, He S, Zong Y, Wang J, Tang J. 2022. Effects of Intercropping *Pandanus amaryllifolius* on soil properties and microbial community composition in *Areca catechu* plantations. Forests. 13(11):1814. doi:10.3390/f13111814.