Complete chloroplast genome assembly of *Alternanthera denticulata* R.Br. 1810 (Amaranthaceae): insight to potential marker and phylogenetic analysis

Krishnamoorthy Devanathan[1], Nithaniyal Stalin[2], Govindan Ganesan[3] and Balaji Raju[4*]
1. Plant Biology and Systematics Lab, CSIR-Central Institute of Medicinal and Aromatic Plants (CIMAP), Research Centre, Bengaluru, INDIA
2. Botanical Survey of India, Western regional Centre, Pune - 411001, Maharashtra, INDIA
3. Department of Genetic Engineering, College of Engineering and Technology, Faculty of Engineering and Technology, SRM Institute of Science and Technology, SRM Nagar, Kattankulathur, Kanchipuram, Chennai, Tamil Nadu, INDIA
4. Center for Global Health Research, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, Tamil Nadu, INDIA
*balajiramasrinivasan@gmail.com; balajir24.smc@saveetha.com

Abstract
Alternanthera denticulata R.Br. is a unique traditional medicinal plant from Amaranthaceae requiring potential genetic exploration. This study aimed to assemble the complete chloroplast (cp) genome of A. denticulata and compare it with the existing published cp genome of the congenerics A. philoxeroides and A. sessilis. The assembled cp genome, with a size of 151,978 bp, exhibits a typical quadripartite structure. The large single-copy (LSC) region spans 84,529 bp and the small single-copy (SSC) region spans 17,281 bp, together comprising of 67% of the genome. The pair of inverted repeats (IRA and IRB) is each 25,084 bp, covering the remaining 33% of the genome. A total of 129 valuable genes were identified, including 88 protein-coding genes, 33 tRNA genes and 8 rRNA genes.

Phylogenetic analysis using complete cp genomes from 11 species in the Amaranthaceae indicated that A. denticulata, A sessilis and A. philoxeroides form a monophyletic clade. Notably, the ycf1 gene in the LSC region was more divergent in A. denticulata, A. sessilis and A. philoxeroides. Given the low levels of taxonomic expertise and clarity in species delineation within the Alternanthera genus, the genetic dataset generated from this study provides nucleotide insights for identifying new DNA barcode marker for herbal drug authentication of Alternanthera. The DNA barcode based identification enhances the discriminatory power of species identification within the congeners beyond morphological methods and thus will be useful in the wide aspects of plants breeding, authentication of functional foods and evolutionary studies.

Keywords: *Alternanthera denticulata*, Amaranthaceae, Chloroplast Genome, Conservation, Species delineation.

Introduction
Amaranthaceae consists of 183 genera and the genus *Alternanthera* consists of 106 accepted species. The genus *Alternanthera* has been widely studied not only for ornamental or vegetable purposes, but also as functional foods and medicines[25]. The species *Alternanthera denticulata* R.Br. (Synonym of *Alternanthera sessilis* var. *denticulata*), commonly known as Lesser Joyweed, is a low scrambling herb, grows up to 50 cm in height and generally possesses papery, white or pinkish flower[3,8,28]. Most of the species are distributed from the tropical to dry climate regions and are native to Australia and New Zealand[8,20].

Traditionally, *A. denticulata* is commonly used as leafy vegetables and is also used as potential medicine in folk and ayurvedic pharmacopoeia. This cultivar is believed to be able to reduce the risk of cardiovascular disease and is used in the treatment of diabetes[24]. Therefore, identification of novel genes from the genome of this species will be helpful for understand the genetic mechanisms.

Chloroplast is an organelle that performs primarily photosynthesis and its genome sequence is well conserved, so it is a major material utilized in the studies of species classification, differentiation and evolutionary process[2,12,14]. Chloroplasts are a type of plastid distinguished by their double-layered membrane, independent DNA and thylakoid structures[17]. They originated through endosymbiosis between a photosynthetic bacterium and a non-photosynthetic host[6], preserving their unique genomic information[7]. These intracellular organelles are crucial for photosynthesis, supplying energy to plants and algae and facilitating the biosynthesis of primary metabolites.

Plastids exhibit non-recombinative behavior and are inherited uniparentally[4]. Typically, an angiosperm chloroplast genome is quadripartite, consisting of a large single-copy (LSC) region and a small single-copy (SSC) region, which are separated by a pair of inverted repeats (IRs)[16,19]. The number of chloroplast genomes reported and stored in the National Center for Biotechnology Information (NCBI) database is steadily increasing. Compared to nuclear and mitochondrial genomes, chloroplast genomes are the most conserved in terms of DNA sequences, organization and structure, making them valuable for phylogenetic analysis, species identification, authentication of herbal products and molecular taxonomy[22,27].

Jiang et al[19], reported the complete chloroplast genomes of *Alternanthera philoxeroides* and their comparison with

congenerics (Amaranthaceae), marking *Alternanthera* chloroplast sequence. Presently, chloroplast genome sequences are available for *A. philoxeroides* and *A. sessilis* from the *Alternanthera* genus. However, there are no reports on the chloroplast genome sequence of *A. denticulata* (Syn. *A. sessilis* var. *denticulata*) in the organellar genome resource at NCBI[18]. In this study, the chloroplast complete genome sequence assembly of *A. denticulata* was performed to be used as an important material for further studies on the evolution, biodiversity of the *Alternanthera* species and the *Amaranthaceae* family.

Material and Methods

The whole genome sequence raw data of *A. denticulata* was downloaded from NCBI SRA database (GenBank accession: SRR8666550) using fastq-dump tool of SRA-toolkit in two forward and reverse reads. We accessed the European Nucleotide Archive (ENA) (https://www.ncbi.nlm.nih.gov/sra) database and obtained WGS (whole-genome sequencing) data for other *Alternanthera* species. Using *de novo* and reference-based assembly methods, the chloroplast genome assembly was done using NovoPlasty v.4.3.2[5] and GetOrganelle v1.7.7.0[10], with ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbc*L) gene from *A. philoxeroides* (GenBank accession no. NC_042798.1) as a seed sequence.

The assembled chloroplast genome of *A. denticulata* was annotated with GeSeq[26]. The predicted transfer RNAs (tRNAs) were confirmed by tRNAscan-SE 2.0[15]. Chloroplot tool was used to visualize the chloroplast genome map (https://irscope.shinyapps.io/chloroplot/)[30]. In addition, the CPGView (www.1kmpg.cn/cpgview/)[13,14], was applied to structures to visualize the intron-containing genes. The *de novo* assembled chloroplast genome of *A. denticulata* was compared with two reported chloroplast genome of *Alternanthera* species. DNaSP version 6.0[21], was used to calculate nucleotide diversity (*Pi*) among the three *Alternanthera* chloroplast genomes. Only the protein coding genes, more than 1000 bp in size were considered.

The complete cp genome sequence of *A. denticulata* (GenBank ID: PP869626.1), two other species of the *Alternanthera*, such as *A. philoxeroides* (GenBank ID: NC_042798.1), *A. sessilis* (GenBank ID: PP239384.1), 14 other species from the *Amaranthaceae* family and one species of *Arabidopsis* thaliana as outgroup were downloaded from NCBI database. The 18 complete cp sequences were aligned using MAFFT v7.4.0.9 (https://mafft.cbrc.jp/alignment/software/index.html) with default parameters[11]. The aligned sequences were further trimmed to equal length and the maximum likelihood method was followed to infer the phylogenetic relationship with 1000 bootstrap replicated in MEGA 11 and a phylogenetic tree was generated[23].

Results and Discussion

The *A. denticulata* cp genome is 151,978 bp in length and follows the typical quadripartite structure, comprising of two inverted repeat regions (IRs), a large single copy region (LSC) and a small single copy region (SSC) (Figure 1).

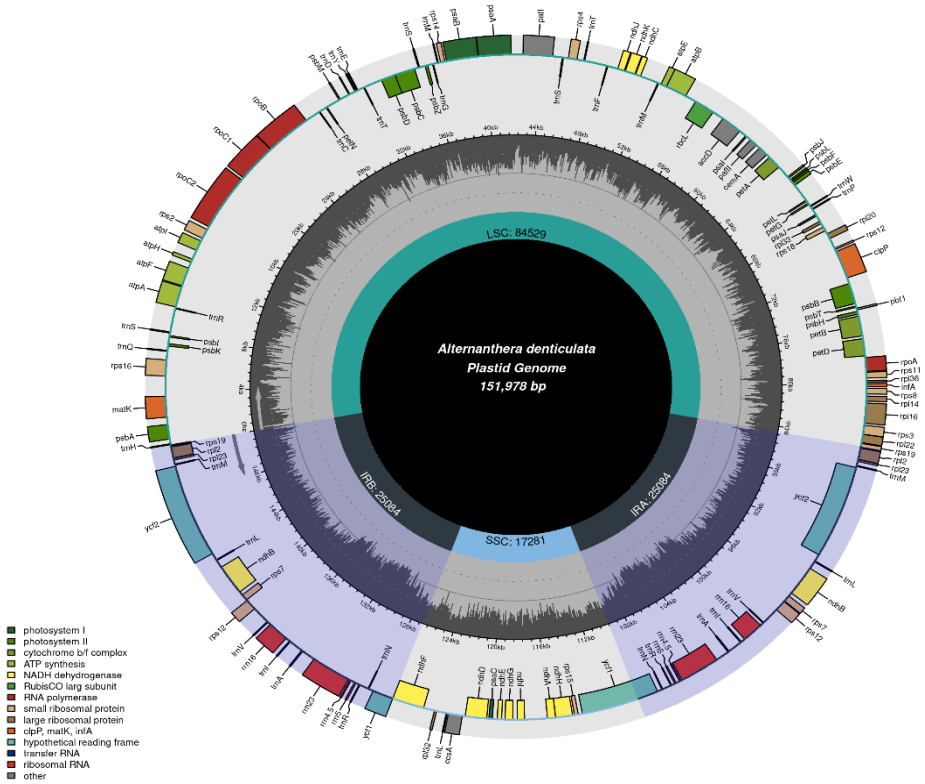

Figure 1: Chloroplast genome maps of *A. denticulata* Genes coding forward are on the outer circle, while genes coding backward are on the inner circle. The grey circle inside represents the GC content

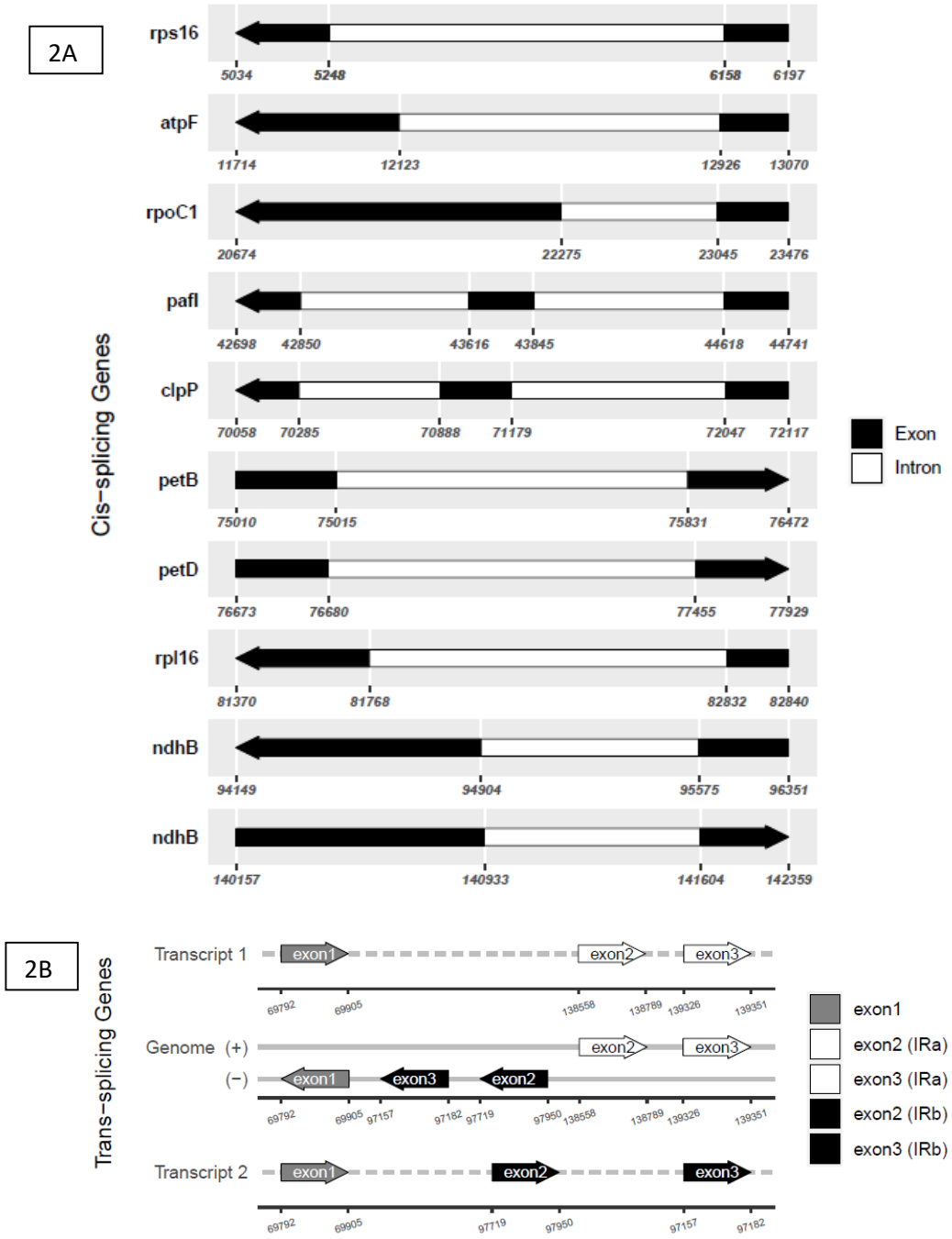

Figure 2: Details of intron containing genes and trans-splicing genes of cp genome of *A. denticulata*. Cis-splicing genes map (2A) and Trans-splicing gene of *rps*12 (2B)

The LSC region spans 84,529 bp, while the SSC region is 17,281 bp long, with a pair of IRs, each covering 25,084 bp. A total of 129 unique genes were identified, including 88 protein-coding genes, 33 tRNA genes and 8 rRNA genes. Seven protein-coding genes (*rps*16, *atp*F, *rpo*C1, *pet*B, *pet*D, *rpl*16 and *ndh*B) were single-intron genes and two genes (clpP1 and pafI) had two introns (Figure 2). The complete chloroplast genome of *A. denticulata* with supportive gene annotations was submitted to GenBank under the accession number PP869626.1.

Nucleotide diversity (*Pi*) is an indicator of the degree of variation of DNA sequence and also represents the genetic diversity of species[1]. In the study of *A. sessilis*, the ycf2-trnL intergenic region in the IRa region and the ycf2 of coding gene in the IRb region showed highest nucleotide diversity. A total of 59 chloroplast gene sequences from 28 closely related species were used to reconstruct the phylogeny to understand the species relationship[29].

However, the low level of sequence variation provided only the limited information and could not resolve the intra-genus relationship. By further understanding of the nucleotide variability (*Pi*), we also calculated the DNA polymorphism among these three *Alternanthera* species. There were six variable regions that showed high *Pi* value in *ycf1* gene

(0.0142), followed by *mat*K (0.0127) and *ndh*F (0.0125) in the *Alternanthera* chloroplast genomes (Figure 3).

Therefore, these five high-resolution regions, especially *ycf1* (*Pi* = 0.0142) are screened according to nucleotide polymorphisms which are statistically significant (p=<0.0001). The *ycf*1 gene can be used as effective molecular markers for species identification and phylogeny within the *Alternanthera* genus. These hotspot regions could be developed as species-specific marker not only for *Alternanthera* species identification, but also can be useful in plant breeding, authentication of functional foods, phylogenetic analysis and developing novel DNA barcodes.

Phylogenetic Analysis: In order to confirm the evolutionary relationship of *A. denticulata*, a maximum likelihood (ML) phylogenetic tree was inferred based on *ycf1* gene and the complete protein-coding genes of chloroplast genome, of which 17 species were from the Amaranthaceae family including 3 species of genus *Alternanthera*, 14 species of Amaranthaceae and *Arabidopsis thaliana* that served as the outgroups. The 18 sequences were aligned using the default settings using MAFFT tool. The maximum likelihood phylogenetic analyses were performed based on GTR+G+I model in the MEGA software, with 1,000 bootstrap replicates. The phylogenetic tree was constructed based on the *ycf1* gene formed two separate clades. The clade I consists of species from four sub-families such as Gomphrenoidae, Amaranthoideae, Corispermoideae and Chenopodioideae. Our results were similar to previously reports and support the same clustering pattern in the phylogeny[29].

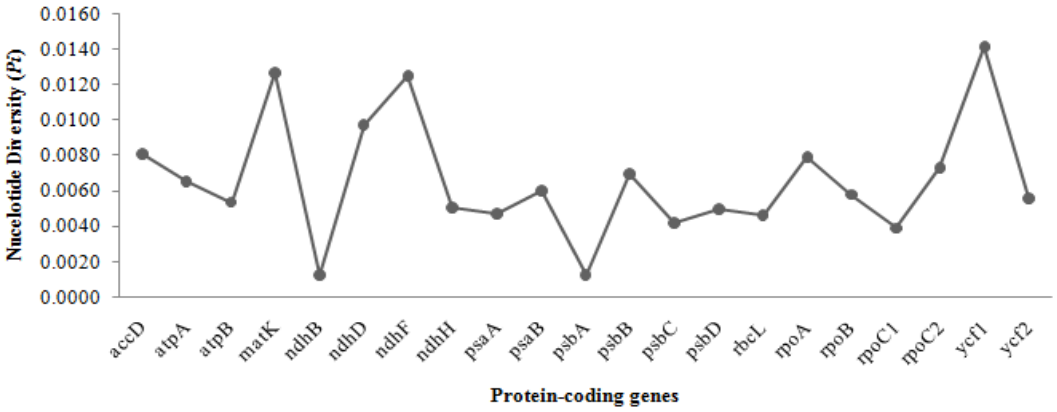

Figure 3: The nucleotide polymorphism for chloroplast genomes of *Alternanthera* species was calculated using DnaSP 6.0. The protein-coding genes were selected which are more than 1000 bp in size. Three most divergent regions (*ycf1*, *mat*K and *ndh*F) are suggested as mutational hotspots

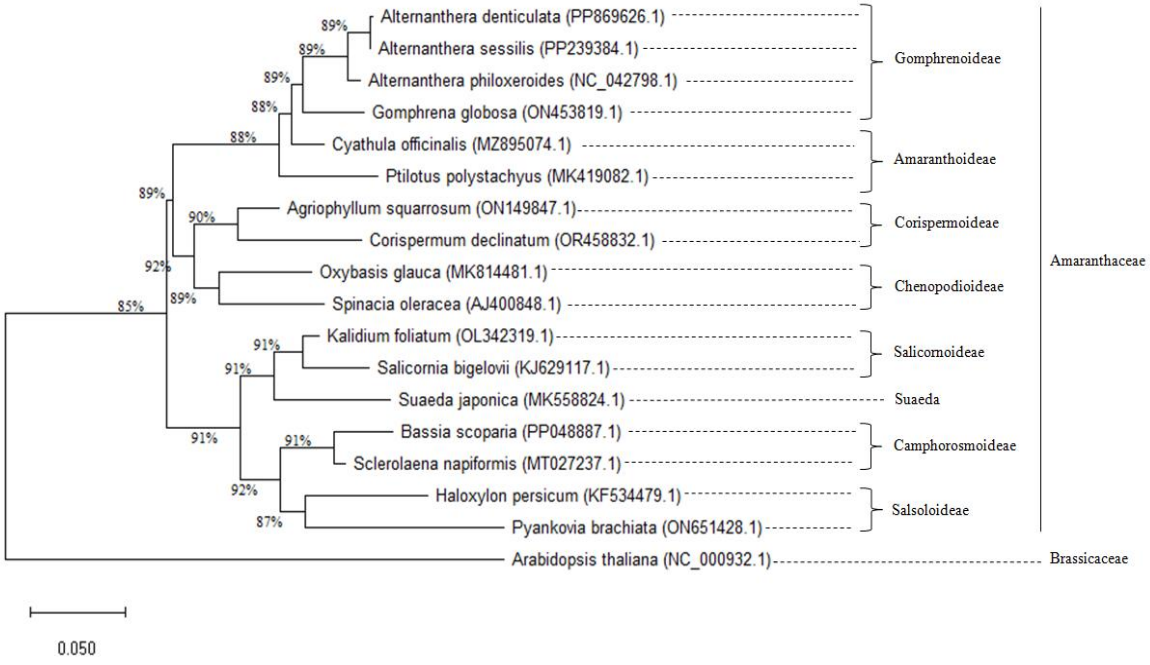

Figure 4: Phylogenetic tree of 10 *Amaranthaceae* species and their related species based on the *ycf1* gene of chloroplast genomes. The numbers on the branches indicate the bootstrap values

Conclusion

Three *Alternanthera* species *A. denticulate*, *A. sessilis* and *A. philoxeroides* were clustered in Gomphrenoideae sub-family clade with *Gomphrena globosa*, a basal clade. The clade II was formed by the Salicornoideae, Camphorosmoideae and Salsoloideae. Results of phylogenetic analysis showed that the genomes of *Alternanthera* formed a well-supported clade (Figure 4). Among the three *Alternanthera* species in the phylogenetic tree, *A. denticulata* is close to *A. sessilis*. The complete chloroplast genomes of *Alternanthera* are too few to correctly infer the phylogenetic relationships of the members of the genus, considering the number of species of the genus.

However, our phylogenetic analysis suggests that *A. denticulata* is distinct from the other species of Amaranthaceae. The complete chloroplast genome of *A. denticulata* determined in this study will provide useful information for further studies on evolution and biodiversity with *Alternanthera* species and Amaranthaceae which contain many economically important plants.

Acknowledgement

The author R.B. acknowledges the Center for Global Health Research, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University for providing the computational bioinformatics facility and support for this study. The author S.N. acknowledges the Director, Botanical survey of India, Kolkata, India.

References

1. Akhunov E.D., Akhunova A.R., Anderson O.D., Anderson J.A., Blake N., Clegg M.T., Coleman-Derr D., Conley E.J., Crossman C.C., Deal K.R., Dubcovsky J., Gill B.S., Gu Y.Q., Hadam J., Heo H., Huo N., Lazo G.R., Luo M.C., Ma Y.Q., Matthews D.E., McGuire P.E., Morrell P.L., Qualset C.O., Renfro J., Tabanao D., Talbert L.E., Tian C., Toleno D.M., Warburton M.L., You F.M., Zhang W. and Dvorak J., Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes, *BMC Genomics*, **11**, 702 **(2010)**

2. Cheng Y., Zhang L., Qi J. and Zhang L., Complete chloroplast genome sequence of *Hibiscus cannabinus* and comparative analysis of the Malvaceae family, *Front. Genet*., **11**, 227 **(2020)**

3. Curtis W.M., The Student's Flora of Tasmania, Part 3, Government Printer, Hobart, 250 **(1967)**

4. Daniell H., Lin C.S., Yu M. and Chang W.J., Chloroplast genomes: diversity, evolution and applications in genetic engineering, *Genome Biol*., **17(1)**, https://doi.org/10.1186/s13059-016-1004-2 **(2016)**

5. Dierckxsens N., Mardulyn P. and Smits G., NOVOPlasty: de novo assembly of organelle genomes from whole genome data, *Nucleic Acids Res*., **45(4)**, e18, https://doi.org/10.1093/nar/gkw955 **(2017)**

6. Guo X., Wang Z., Cai D., Song L. and Bai J., The chloroplast genome sequence and phylogenetic analysis of *Apocynum venetum* L., *PLoS ONE*, **17(3)**, e0261710 **(2022)**

7. Howe C.J., Barbrook A.C., Koumandou V.L., Nisbet R.E.R., Symington H.A. and Wightman T.F., Evolution of the chloroplast genome, *Philos. Trans. R. Soc. Lond. Ser. B. Biol. Sci*., **358(1429)**, 99–107, https://doi.org/10.1098/rstb.2002.1176 **(2003)**

8. Jacobs S.W.L. and Lapinpuro L., PlantNET - FloraOnline: *Alternanthera* denticulate, Retrieved from http://plantnet.rbgsyd.nsw.gov.au **(2024)**

9. Jiang P., Xu G., He Y., Sun T., Liu C., Chen C., Zuo R. and Sun C., Complete chloroplast genome of *Alternanthera philoxeroides* by de novo sequencing, *Mitochondrial DNA*, Part B, **6(7)**, 1826–1828, https://doi.org/10.1080/23802359.2021.1931512 **(2021)**

10. Jin J.J., Yu W.B., Yang J.B., Song Y., dePamphilis C.W., Yi T.S. and Li D.Z., GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes, *Genome Biol*., **21(1)**, 241, https://doi.org/10.1186/s13059-020-02154-5 **(2020)**

11. Katoh K. and Standley D.M., MAFFT multiple sequence alignment software version 7: improvements in performance and usability, *Mol. Biol. Evol*., **30(4)**, 772–780, https://doi.org/10.1093/molbev/mst010 **(2013)**

12. Kim S.T., Oh S.H. and Park J., The complete chloroplast genome of *Diarthron linifolium* (*Thymelaeaceae*), a species found on a limestone outcrop in eastern Asia, *Korean J. Pl. Taxon*., **51**, 345–352 **(2021)**

13. Liu S., Ni Y., Li J., Zhang X., Yang H., Chen H. and Liu C., CPGView: a package for visualizing detailed chloroplast genome structures, *Mol. Ecol. Resour*., https://doi.org/10.1111/1755-0998.13729 **(2023)**

14. Liu X.F., Zhu G.F., Li D.M. and Wang X.J., Complete chloroplast genome sequence and phylogenetic analysis of Spathiphyllum 'Parrish', *PLoS ONE*, **14**, e0224038 **(2019)**

15. Lowe T.M. and Chan P.P., tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes, *Nucleic Acids Res*., **44(W1)**, W54–W57, https://doi.org/10.1093/nar/gkw413 **(2016)**

16. Maheswari P., Kunhikannan C. and Yasodha R., Chloroplast genome analysis of angiosperms and phylogenetic relationships among *Lamiaceae* members with particular reference to teak (*Tectona grandis* L.f), *J. Biosci*., **46**, 43 **(2021)**

17. Mustardy L., Buttle K., Steinbach G. and Garab G., The three-dimensional network of the thylakoid membranes in plants: quasihelical model of the granum-stroma assembly, *Plant Cell*, **20(10)**, 2552–2557, https://doi.org/10.1105/tpc.108.059147 **(2008)**

18. NCBI, Organelle resources at NCBI, Retrieved from http://www.ncbi.nlm.nih.gov/genome/organelle **(2024)**

19. Oldenburg D.J. and Bendich A.J., The linear plastid chromosomes of maize: terminal sequences, structures and implications for DNA replication, *Current Genetics*, **62**, 431–442 **(2016)**

20. POWO, Plants of the World Online, Royal Botanic Gardens, Kew, Retrieved from https://powo.science.kew.org **(2024)**

21. Rozas J., Ferrer-Mata A., Sánchez-DelBarrio J.C., Guirao-Rico S., Librado P., Ramos-Onsins S.E. and Sánchez-Gracia A., DnaSP 6: DNA sequence polymorphism analysis of large data sets, *Mol. Biol. Evol.*, **34(12)**, 3299–3302, https://doi.org/10.1093/molbev/msx248 **(2017)**

22. Singh N.V., Patil P.G., Sowjanya R.P., Parashuram S., Natarajan P., Babu K.D., Pal R.K., Sharma J. and Reddy U.K., Chloroplast genome sequencing, comparative analysis and discovery of unique cytoplasmic variants in pomegranate (*Punica granatum* L.), *Front. Genet*., **12**, 704075, https://doi.org/10.3389/fgene.2021.704075 **(2021)**

23. Tamura K., Stecher G. and Kumar S., MEGA11: Molecular evolutionary genetics analysis version 11, *Mol. Biol. Evol*., **38**, 3022–3027, https://doi.org/10.1093/molbev/msab120 **(2021)**

24. Tan K.K. and Kim K.H., A. sessilis red ethyl acetate fraction exhibits antidiabetic potential on obese type 2 diabetic rats, *Evid*.-*Based Complement. Altern. Med*., doi: 10.1155/2013/845172 **(2013)**

25. Tanveer A., Khali A. and Siddiqui M.H., A review on genus *Alternanthera* weeds implications, *Pak. J. Weed. Sci. Res*., **19(1)**, 53–58 **(2013)**

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

27. Vassou S.L., Nithaniyal S., Raju B. and Parani M., Creation of reference DNA barcode library and authentication of medicinal plant raw drugs used in Ayurvedic medicine, BMC Complement, *Altern. Med*., **18(Suppl 1)**, 186, https://doi.org/10.1186/s12906-016-1086-0 **(2016)**

28. Walsh N.G. and Entwisle T.J., Flora of Victoria, Volume 3, Dicotyledons: *Winteraceae* to *Myrtaceae*, Inkata Press, Melbourne, 150 **(1996)**

29. Wang Y., Zhao X., Chen Q., Yang J., Hu J., Jia D. and Ma R., Complete chloroplast genome of *Alternanthera sessilis* and comparative analysis with its congeneric invasive weed *Alternanthera philoxeroides*, *Genes (Basel)*, **15(5)**, 544, https://doi.org/10.3390/genes15050544 **(2024)**

30. Zheng S., Poczai P., Hyvönen J., Tang J. and Amiryousefi A., Chloroplot: an online program for the versatile plotting of organelle genomes, *Front. Genet*., **11**, 576124, https://doi.org/10.3389/fgene.2020.576124 **(2020)**.

(Received 14th October 2024, accepted 24th December 2024)