MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 10, 1384–1388
https://doi.org/10.1080/23802359.2024.2412239

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

The complete sequence of chloroplast genome of *Baeckea frutescens* Linaeus 1753 (Myrtoideae), a traditional folk medicinal plant

Chengcheng Ling, Menglei Zhang, Wenwen Ma and Menyu Liu

College of Food and Bioengineering, Bengbu University, Bengbu, China

ABSTRACT

Baeckea frutescens Linaeus 1753, as a traditional folk medicine in South East Asia, possesses sesquiterpenes, phloroglucinols, chromones, and essential oil, and is utilized for traditional Chinese medicinal purposes. The genetic diversity of the plant must be better understood, considering its significance. The complete chloroplast (cp) genome of *B. frutescens* was sequenced and assembled by using Illumina paired-end data, marking a significant advancement toward comprehending its genetic composition. The complete cp genome is 158,939 bp in length and contains 128 genes, consisting of 83 protein-coding genes, 8 ribosomal RNA genes, and 37 transfer RNA genes. Phylogenetic analyses indicated that *B. frutescens* and other the 13 were clustered to the family of Myrtaceae. These findings are crucial for the conservation and utilization of this important plant species. Additionally, they underscore the potential for future research on the evolution and preservation of *B. frutescens*, which could be advantageous in pharmaceutical applications.

ARTICLE HISTORY
Received 11 July 2024
Accepted 28 September 2024

KEYWORDS
Baeckea frutescens; complete chloroplast genome; Myrtoideae; phylogenetic analysis

Introduction

Baeckea frutescens Linaeus 1753, is a small aromatic low-growing tree found in mountainous regions, belonging to the Myrtaceae family. It is widely distributed in Peninsular Malaysia, Sumatra, Southern China, and Australia (Ahmad et al. 2012). As a traditional folk medicine in South East Asia, its leaves (Figure 1) contain sesquiterpenes, phloroglucinols, chromones, and essential oil, and are often used to treat rheumatism, snake bites, dermatitis, and the common cold (Hou et al. 2017). Additionally, cyclopentenones and phloroglucinols in the leaves were reported to have cytotoxic effects against human pancreatic, lung, and breast cancer cell lines (Fujimoto et al. 1996; Nisa et al. 2016). Despite its high medicinal value, there have been no studies on the complete chloroplast genome and phylogenetic analysis of *B. frutescens* up to now. Here, the chloroplast (cp) genome of *B. frutescens* was assembled, and it could facilitate the study of the phylogenetic relationships between *B. frutescens* and other plant species.

Materials and methods

Young *B. frutescens* leaf tissue was collected from Fujian Institute of Tropical Crops (Fujian, China, 117°65′52″E, 24°52′30″N), and was deposited in Bengbu University (voucher number No.BBU10071300, Chengcheng Ling, lingchengcheng402@163.com). To build an Illumina pair-end library, total genomic DNA was extracted using the modified CTAB method (Jinlu et al. 2013). We Weigh 1 g of dry plant tissue and ground it into a powder with sand using a mortar or a pestle. Remove the powder into a 2.0 mL microcentrifuge tube. Then, add 0.7 mL CI (chloroform: isoamyl alcohol = 24:1, v/v), and mix well for 10 min by inverting the tube gently. Centrifuge at 10 000 ×g for 10 min, carefully remove the supernatant to a new 1.5 mL microcentrifuge tube and add 0.5 mL pre-cooled isopropanol, carefully mix well. Incubate at −20 °C for 20 min and centrifuge at 10 000 ×g for 10 min, discard the supernatant. add 0.5 mL 75% ethanol, re-suspend the pellet, centrifuge at 10 000 ×g for 2 min, and discard the supernatant. Finally, add 0.1 mL TE to dissolve DNA after ethanol has evaporated. The resultant library was sequenced using HiSeq (Illumina, San Diego, CA, USA) at Beijing Genomics Institute (BGI, Shenzhen, China) and yielded approximately 16.05 GB of raw data. The raw paired-end data was filtered using the FastQC program (Andrews 2014). Subsequently, about 15.17 G of high-quality clean reads with Q20 = 98.01% were applied to assemble the cp genome using Getorganelle (Jin et al. 2020) and the cp genome annotation was conducted with the online program GeSeq (Tillich et al. 2017). The cp genome map was drawn using CPGview (http://www.1kmpg.cn/ cpgview). The BWA software was used to align Illumina short sequences to the chloroplast genome sequence. The coverage depth was then calculated using SAMtools software (Lin et al. 2009). The assembled

CONTACT Chengcheng Ling lingchengcheng402@163.com College of Food and Bioengineering, Bengbu University, Bengbu, China .
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2412239.

Figure 1. Distinctive characteristics of leaves and flowers from *B. frutescens*. The most characteristic feature of *B. frutescens* is its small, sessile or leafless leaves, which are narrow-linear or linear in shape. The flowers are small, white, with round petals, and the basal part narrows into a short stalk. Photographs by Meng-lei Zhang at Fujian Institute of Tropical Crops (Fujian, China, 117°65′52″E, 24°52′30″N).

genome sequence of *B. frutescens* has been deposited in GenBank with accession number ON470204.1. To conduct the overall sequence alignment of the *B. fretescens* and five Eucalyptus species cp genomes, we employed mVISTA software (Frazer et al. 2004), using the annotation of *B. fretescens* as a reference. The sequences were aligned using MAFFT 7 (Katoh and Standley 2013). The maximum likelihood (ML) analysis was performed using RaxML software v 8.2.9, of which the bootstrap values were calculated using 1000 replicates (Stamatakis 2006) using *Osbeckia stellate* and *Heterocentron elegans* as outgroups.

Results

The resulting graph shows that the chloroplast genome sequence has a high coverage depth of over 100X, indicating reliable data quality (Figure S1). Consequently, the whole complete circular cp genome size of *B. frutescens* is 158,939 bp with 37.5% GC content (Figure 2). The cp genome displays the common quadripartite structure (Figure 2). It comprises a large single-copy region (LSC) of 87,716 bp, a small single-copy region (SSC) of 18,417 bp, and a pair of inverted repeat regions (IRA and IRB) of 26,403 bp (Figure 2). A total of 128 genes were predicted in the chloroplast genome, consisting of 83 protein-coding genes, 37 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes (Figure 2). Among the 12 protein-coding genes containing introns, nine, two, or one are located in the LSC, IRA, or SSC region, respectively (Table S1). Ten genes have only one intron. Another two genes (*clp*P and *ycf*3) have two introns each. The *ndhA* gene possesses the largest intron (1070 bp) (Table S1). The *ndh*F gene displayed the highest level of variation (Figure S2A). The highly diverged regions were identified in the intergenic spacers within the noncoding genes, while the coding regions are highly conserved (Figure S2B). The cis-/trans-splicing and trans-splicing *rps12* genes maps in the Supplementary Materials Figure S3 and Figure S4. For phylogenetic analysis, 25 species published complete chloroplast genome of Myrtales were downloaded from the GenBank database (https://www.ncbi.nlm.nih.gov/genome/browse#!/organelles). The result indicated that *B. frutescens* and the other 13 were clustered to the family of Myrtaceae (Figure 3).

Discussion and conclusion

Here, the complete cp genome sequence of *B. frutescens* has been assembled, which has a gene content and size that falls within the range of other members of the Myrtaceae family (Asif et al. 2013; Bayly et al. 2013). The results related to the intron structure of the gene indicated that the most of genes are located in the LSC, IRA or SSC region. Similar results were observed in *Eugenia uniflora* cp genome (Eguiluz et al. 2017). The examination of variation discovered that the intergenic spacers within the noncoding genes exhibited high divergence, whereas the coding regions remained highly conserved, which is consistent with those reported by Eguiluz et al. (2010). In the ML phylogenetic analysis, we found that *B. frutescens* and other 13 were clustered to the family of Myrtaceae, providing the monophyly of the genus *Baeckea* (Berger et al. 2016). Of all the 13 Myrtoideae species

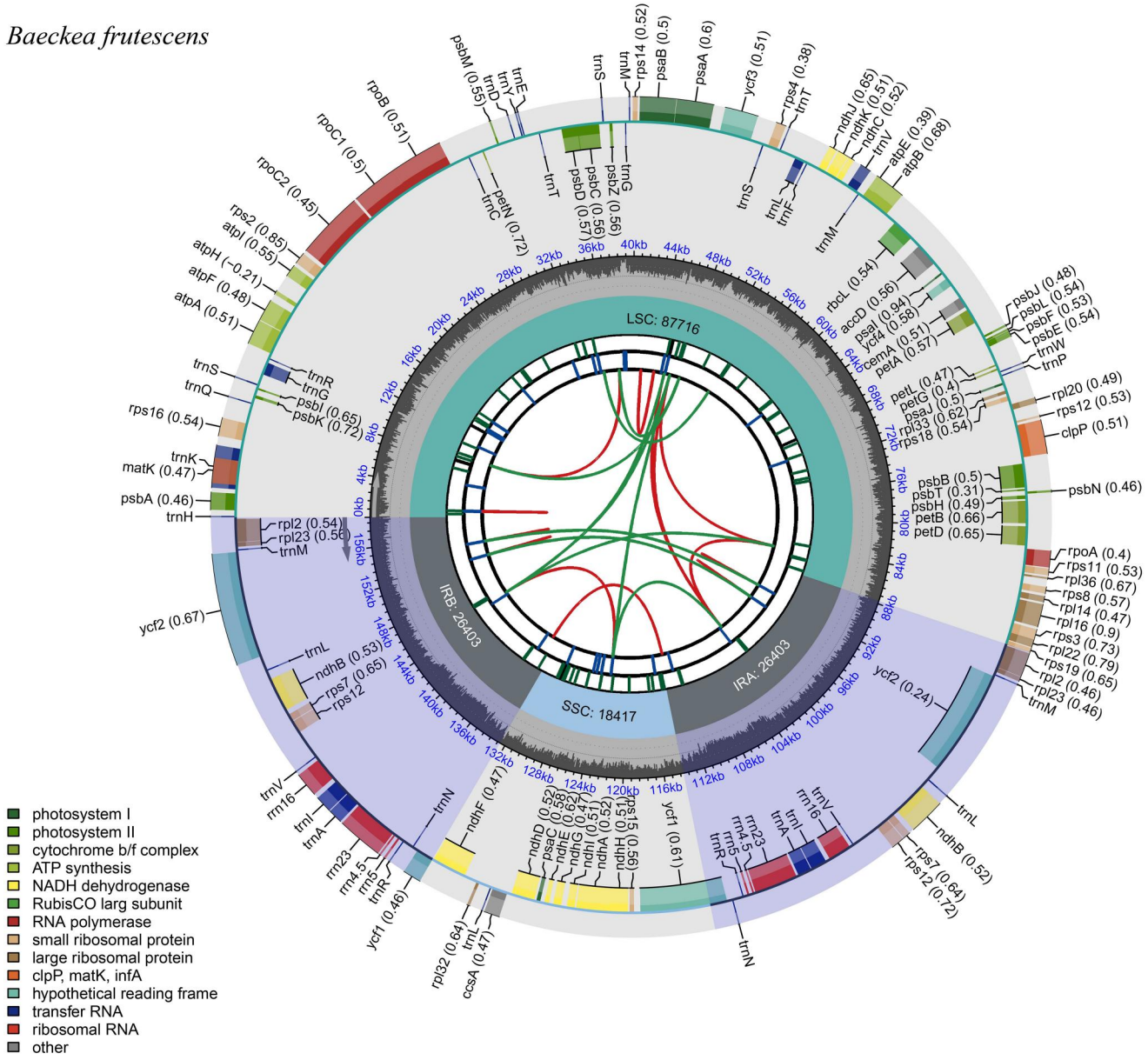

Figure 2. Gene map of the complete chloroplast genome of *B. frutescens*. The species name is shown in the top left corner. The map contains six tracks by default. From the center outward, the first track shows dispersed repeats, including direct and palindromic repeats connected by red and green arcs. The second track displays long tandem repeats as blue bars, while the third displays short tandem repeats or microsatellite sequences as differently colored short bars. These colors correspond to the type and description of each repeat, with black representing complex repeats, green for repeat unit size 1, yellow for size 2, purple for size 3, blue for size 4, orange for size 5, and red for size 6. The fourth track displays the SSC, IRa, IRb, and LSC regions. The fifth track shows the GC content along the genome, while the sixth track sounds the genes. The gene names are followed by optional information about codon usage bias and color-coded based on their functional classification. The inner genes are transcribed clockwise, and the outer genes are transcribed anticlockwise. The functional type of the genes is shown in the bottom left corner.

analyzed, *Baeckea frutescens* is evolutionarily one of the older ones. Our phylogenetic investigations using entire cp genomes have broadened the taxon sampling and provide novel resources of chloroplast genome sequence in the genus *Baeckea* for further comparative and evolutionary analysis.

Author contributions

Chengcheng Ling wrote the paper, conceptualized and designed the research and Menglei Zhang collected the sample. Menglei Zhang, Wenwen Ma, and Mengyu Liu participated in the analysis and interpretation of the data. All authors agree to be accountable for all aspects of the work.

Ethical approval

B. frutescens is not listed as a threatened or endangered species and is widespread in China. The permissions of plant material collection were not required. The field and laboratory studies comply with the guidelines provided by Bengbu University.

Disclosure statement

No potential conflict of interest was reported by the authors.

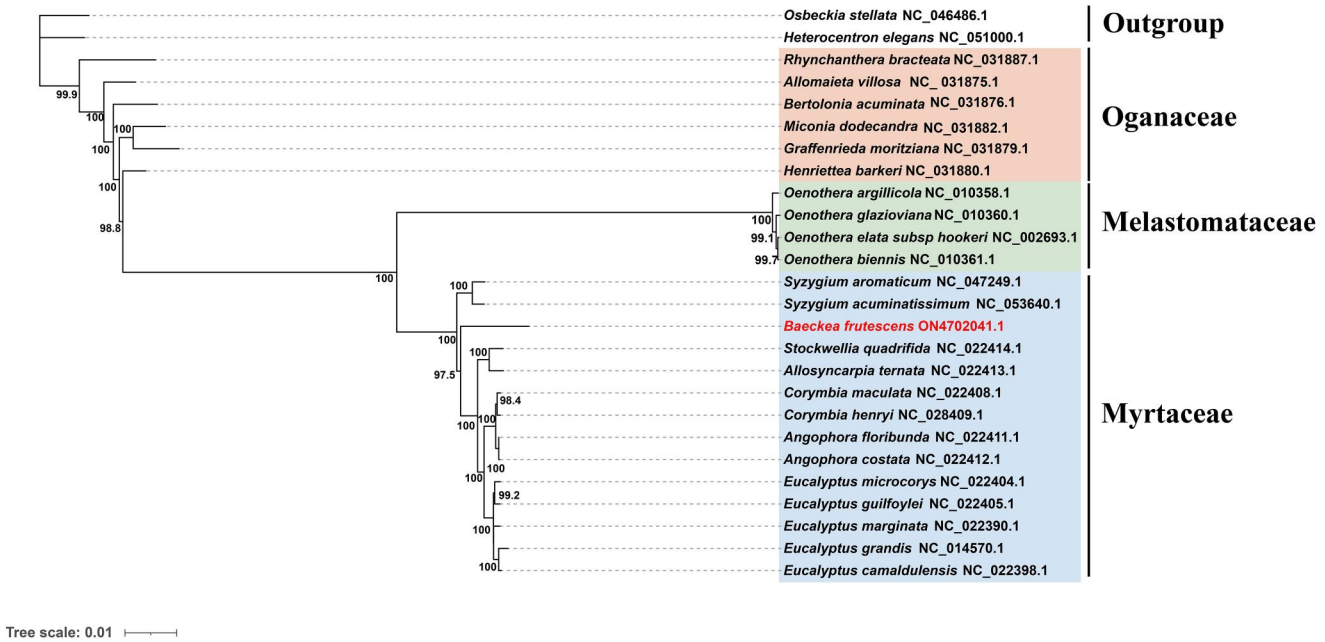

Figure 3. The maximum likelihood (ML) phylogenetic tree was constructed based on of the complete plant's chloroplast genomes data of 25 species in the Myrtales family. The corresponding phylogram tree is shown in panel B. We downloaded 24 campsis species chloroplast genomes from GenBank, *Osbeckia stellata* (NC_046486.1) (Liang et al. 2022) and *Heterocentron elegans* (NC_051000.1) (Zhang et al. 2020), from the nymphaeaceae, served as the outgroup. *Rhynchanthera bracteata* (NC_031887.1) (Reginato et al. 2016), *Allomaieta villosa* (NC_031875.1) (Reginato et al. 2016), *Bertolonia acuminata* (NC_031876.1) (Reginato et al. 2016), *Miconia dodecandra* (NC_031882.1) (Reginato et al. 2016), *Graffenrieda moritziana* (NC_031879.1) (Reginato et al. 2016), *Henriettea barkeri* (NC_031880.1) (Reginato et al. 2016), *Oenothera argillicola* (NC_010358.1) (Greiner et al. 2008), *Oenothera glazioviana* (NC_010360.1) (Greiner et al. 2008), *Oenothera elata subsp hookeri* (NC_002693.1) (Hupfer et al. 2000), *Oenothera biennis* (NC_010361.1) (Greiner et al. 2008), *Syzygium aromaticum* (NC_047249.1) (Nguyen 2023), *Syzygium acuminatissimum* (NC_053640.1) (Zeng et al. 2020), *Stockwellia quadrifida* (NC_022414.1) (Bayly et al. 2013), *Allosyncarpia ternata* (NC_022413.1) (Bayly et al. 2013), *Corymbia maculata* (NC_022408.1) (Bayly et al. 2013), *Corymbia henryi* (NC_028409.1) (Healey et al. 2018), *Angophora floribunda* (NC_022411.1) (Bayly et al. 2013), *Angophora costata* (NC_022412.1) (Bayly et al. 2013), *Eucalyptus microcorys* (NC_022404.1) (Bayly et al. 2013), *Eucalyptus guilfoylei* (NC_022405.1) (Bayly et al. 2013), *Eucalyptus marginata* (NC_022390.1) (Bayly et al. 2013), *Eucalyptus grandis* (NC_014570.1) (Bayly et al. 2013), *Eucalyptus camaldulensis* (NC_022398.1) (Bayly et al. 2013). The new chloroplast genomes of *Baeckea frutescens* in this study were labeled in red color.

Funding

This work was supported by the Bengbu University Foundation of High level Talent under Grant 2024YYX18QD and Bengbu University Foundation of 2024 Master's Construction "Biology and Medicine" Master's.

Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI at https://www.ncbi.nlm.nih.gov, reference number ON470204.1. The associated BioProject, SRA, and Bio-Sample numbers and are PRJNA905860 SRR22424258 and SAMN31886574, respectively.

References

Ahmad NS, A. Ghani MN, Manaf Ali A, Tajudin Tuan Johari SA, Harun MH. 2012. High performance liquid chromatography (HPLC) profiling analysis and bioactivity of *Baeckea frutescens* L.(Myrtaceae). JPS. 1(2):101. doi:10.5539/jps.v1n2p101.

Andrews S. 2014. FastQC a quality control tool for high throughput sequence data. http://wwwbioinformaticsbabrahamacuk/projects/fastqc/.

Asif H, Khan A, Iqbal A, Khan IA, Heinze B, Azim MK. 2013. The chloroplast genome sequence of *Syzygium cumini* (L.) and its relationship with other angiosperms. Tree Genet Genomes. 9(3):867–877. doi:10.1007/s11295-013-0604-1.

Bayly MJ, Rigault P, Spokevicius A, Ladiges PY, Ades PK, Anderson C, Bossinger G, Merchant A, Udovicic F, Woodrow IE, et al. 2013. Chloroplast genome analysis of Australian eucalypts—Eucalyptus, Corymbia, Angophora, Allosyncarpia and Stockwellia (Myrtaceae). Mol Phylogenet Evol. 69(3):704–716. doi:10.1016/j.ympev.2013.07.006.

Berger BA, Kriebel R, Spalink D, Sytsma KJ. 2016. Divergence times, historical biogeography, and shifts in speciation rates of Myrtales. Mol Phylogenet Evol. 95:116–136. doi:10.1016/j.ympev.2015.10.001.

Eguiluz M, Rodrigues NF, Guzman F, Yuyama P, Margis R. 2017. The chloroplast genome sequence from *Eugenia uniflora*, a Myrtaceae from neotropics. Plant Syst Evol. 303(9):1199–1212. doi:10.1007/s00606-017-1431-x.

Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 32(Web Server issue):W273–W279. doi:10.1093/nar/gkh458.

Fujimoto Y, Usui S, Makino M, Sumatra M. 1996. Phloroglucinols from *Baeckea frutescens*. Phytochemistry. 41(3):923–925. doi:10.1016/0031-9422(95)00659-1.

Hou J-Q, Guo C, Zhao J-J, Dong Y-Y, Hu X-L, He Q-W, Zhang B-B, Yan M, Wang H. 2017. Antiinflammatory meroterpenoids from *Baeckea frutescens*. J Nat Prod. 80(8):2204–2214. doi:10.1021/acs.jnatprod.7b00042.

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

Jinlu L, Shuo W, Jing Y, Ling W, Zhou S. 2013. A modified CTAB protocol for plant DNA extraction. Chin Bull Bot. 48(1):72–78. doi:10.3724/SP.J.1259.2013.00072.

Jo S, Kim H-W, Kim Y-K, Cheon S-H, Kim K-J. 2016. Complete plastome sequence of *Psidium guajava* L. (Myrtaceae). Mitochondrial DNA B Resour. 1(1):612–614. doi:10.1080/23802359.2016.1209096.

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

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics. 25(16):2078–2079. doi:10.1093/bioinformatics/btp352.

Liang R, Ruan F, Zhi X, Zhou Q. 2019. Characterization of the chloroplast genome of *Osbeckia stellata* (Melastomataceae). Mitochondrial DNA Part B. 4(1):1136–1137. doi:10.1080/23802359.2019.1574679.

Lima DF, Goldenberg R, Forest F, Cowan RS, Lucas EJ. 2021. Phylogeny and biogeography of *Myrcia sect. Aguava* (Myrtaceae, Myrteae) based on phylogenomic and Sanger data provide evidence for a Cerrado origin and geographically structured clades. Mol Phylogenet Evol. 157(2021):107043–107048. doi:10.1016/j.ympev.2020.107043.

Liu H, Geng M, Qin Y, Xiao Y, Li M. 2019. Characterization of the complete chloroplast genome of medicinal tea tree (*Melaleuca alternifolia*). Mitochondrial DNA B Resour. 4(2):3307–3308. doi:10.1080/23802359.2019.1673246.

Machado LdO, Vieira LdN, Stefenon VM, Faoro H, Pedrosa FdO, Guerra MP, Nodari RO. 2020. Molecular relationships of *Campomanesia xanthocarpa* within Myrtaceae based on the complete plastome sequence and on the plastid ycf2 gene. Genet Mol Biol. 43(2):e20180377. doi:10.1590/1678-4685-GMB-2018-0377.

Machado LdO, Vieira LdN, Stefenon VM, Oliveira Pedrosa Fd, Souza EMd, Guerra MP, Nodari RO. 2017. Phylogenomic relationship of feijoa (*Acca sellowiana* (O. Berg) Burret) with other Myrtaceae based on complete chloroplast genome sequences. Genetica. 145(2):163–174. doi:10.1007/s10709-017-9954-1.

Nisa K, Ito T, Kodama T, Tanaka M, Okamoto Y, Asakawa Y, Imagawa H, Morita H.,. 2016. New cytotoxic phloroglucinols, baeckenones D-F, from the leaves of Indonesian *Baeckea frutescens*. Fitoterapia. 109(2016):236–240. doi:10.1016/j.fitote.2016.01.013.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 22(21):2688–2690. doi:10.1093/bioinformatics/btl446.

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.

Zeng F, Deng Y, Liu X, Zhu X, Tan G. 2021. The complete chloroplast genome of *Syzygium acuminatissimum*. Mitochondrial DNA B Resour. 6(1):127–128. doi:10.1080/23802359.2020.1847615.

Zhang X, Liu X, Lee SY, Chen Y, Zhong Y, Zhou R, Zeng F. 2020. The complete chloroplast genomes of *Heterotis rotundifolia* and *Heterocentron elegans* (Melastomataceae). Mitochondrial DNA B Resour. 5(3):3094–3095. doi:10.1080/23802359.2020.1797582.