

## The complete chloroplast genome of *Callicarpa rubella* Lindl. (Lamiaceae) from Guangxi, China: genome structure and phylogenetic analysis

Huimin Cai<sup>a</sup>, Xing Liu<sup>a</sup>, Zhiwei Su<sup>b</sup> and Zhonghui Ma<sup>a</sup>

<sup>a</sup>National Demonstration Center for Experimental Plant Science Education, Traditional Chinese Herbal Medicine Resources and Agriculturalization Research Institute, College of Agriculture, Guangxi University, Nanning, China; <sup>b</sup>Institute of Marine Drugs, Guangxi University of Chinese Medicine, Nanning, China

### ABSTRACT

The relationship between *Callicarpa rubella* Lindl. and its infraspecific taxa has troubled researchers for a long time. Here, we reported for the first time the complete chloroplast (cp) genome of *C. rubella* to investigate its phylogenetic position and provide more sequencing information for further studies about the relationship between *C. rubella* and its related species. The cp genome of *C. rubella* was 154,202 bp in length and displayed a typical quadripartite angiosperm structure, containing two inverted repeat (IR) regions of 25,701 bp, a large single-copy (LSC) region of 84,968 bp and a small single-copy (SSC) region of 17,832 bp. It contained 87 protein-coding genes, 37 tRNA genes and 8 rRNA genes. The analysis fully resolved *C. rubella* was in a clade with *C. bodinieri* and *C. nudiflora*. The results indicated *Callicarpa* formed a sister relationship with *Dicrastylis parvifolia* in Lamiaceae.

### ARTICLE HISTORY

Received 28 July 2021  
Accepted 7 October 2021

### KEYWORDS

*Callicarpa rubella*;  
chloroplast genome;  
Lamiaceae; phylogenetic  
relationship

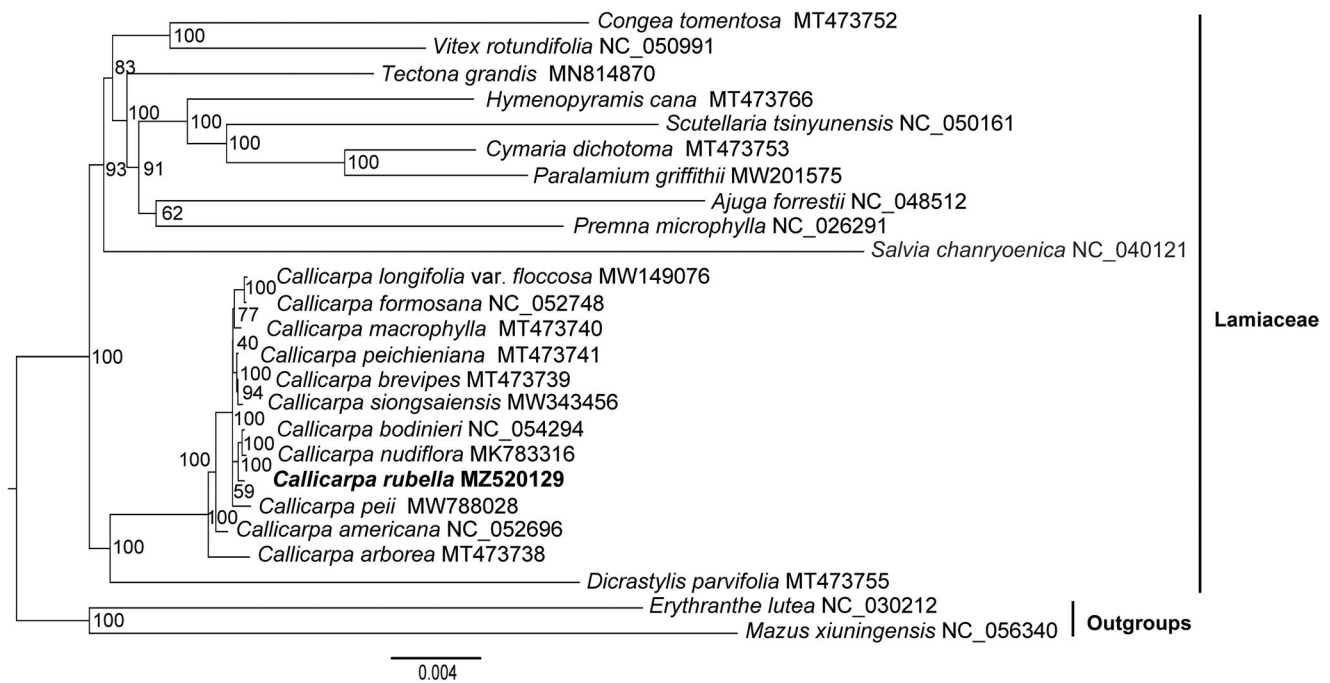
*Callicarpa* (Lamiaceae) with the nickname ‘beauty berry’ was first described by Linnaeus (1753) and due to its attractive purple fruits usually displaying in the autumn, the genus takes the name. The genus as traditionally medicinal plant group harbors significantly economic and pharmacological importances (Tu et al. 2013). *Callicarpa* was traditionally assigned to Verbenaceae, however subsequently Harley et al. (2004) adopted transferring it to Lamiaceae. *Callicarpa rubella* Lindl. 1825 is a shrub, distributed widely in eastern and southeastern Asia (China, Indonesia, Malaysia, Myanmar, Thailand and Vietnam) and always grows in some easily disturbed areas (roadsides or the margin of secondary forest). *C. rubella* is variable morphologically, especially in its indumentum, and the size and shape of the leaves. It is always a headache to identify *C. rubella* from its infraspecific taxa (*C. rubella* var. *subglabra*, *C. rubella* f. *angustata* and *C. rubella* f. *crenata*) and some related species (*C. longipes* and *C. mollis*) based on morphologies. In this study, we assembled and annotated the complete chloroplast (cp) genome of *C. rubella* for the first time to serve as a genetic resource for future studies on the taxonomy of *Callicarpa* and to get a better understanding of phylogenetic relationships in this pantropical genus.

The fresh leaves of *C. rubella* were collected from Guiping, Guangxi, China (23°41′10″N, 110°11′12″E). A specimen was deposited at the herbaria in College of Agriculture, Guangxi University (<https://nxy.gxu.edu.cn/>, Zhonghui Ma, [mazhonghui@gxu.edu.cn](mailto:mazhonghui@gxu.edu.cn)) under the voucher number H004. The total genomic DNA was extracted by modified CTAB method

(Doyle and Doyle 1987) and used for sequencing on Illumina NovaSeq 6000 platform at the Beijing Novogene Technology Co., Ltd. (Tianjin, China). The DNA sample was deposited in the Plant Systematic Evolution Laboratory, College of Agriculture, Guangxi University. The cp genome was assembled by GetOrganelle toolkit (Bankevich et al. 2012; Langmead and Salzberg 2012; Wick et al. 2015; Jin et al. 2020) and annotated using Plastid Genome Annotator (PGA) (Qu et al. 2019). Aligning the complete cp genome and manually adjusting annotation were performed in Geneious (Kearse et al. 2012). The annotated cp genome has been deposited in GenBank (accession number: MZ520129).

The cp genome of *C. rubella* was 154,202 bp in length and had a typical quadripartite angiosperm structure, containing two inverted repeats (IRA and IRB), each of 25,701 bp, a large single-copy (LSC) region of 84,968 bp and a small single-copy (SSC) region of 17,832 bp. The GC content of the whole cp genome was 38.1%. We recovered a total of 128 distinct genes, including 87 protein-coding genes (PCGs), 37 tRNA genes and 8 rRNA genes. Of these genes, 18 were duplicated, including 7 protein-coding, 7 tRNA and 4 rRNA genes.

To confirm the phylogenetic position of *C. rubella* with related species, 24 chloroplast (cp) genomes sequences were downloaded from GenBank, including 22 species from Lamiaceae representing 12 subfamilies recognized currently and 2 outgroup species from Mazaceae and Phrymaceae (Figure. 1) (Vallejo-Marin et al. 2016; Ha et al. 2018; Tao et al. 2019; Wang et al. 2019; Du et al. 2020; Zhao et al. 2020; Wang et al. 2021;



**Figure 1.** The maximum-likelihood (ML) phylogenetic tree of 25 taxa based on 78 CDS of the chloroplast genomes. Numbers near the branches are bootstrap support values. The *C. rubella* genome was marked in bold font.

Xie et al. 2021; Zhao, Chen, et al. 2021; Zhao, Wu, et al. 2021). Due to the rapid development of molecular phylogenetic studies during the past two decades, the traditionally systematic positions of some genera have dramatically changed. *Congea*, *Vitex* and *Tectona* (originally belong to Verbenaceae) have been moved to Lamiaceae (Harley et al. 2004) and *Mazus* and *Erythranthe* (formerly Scrophulariaceae) have been assigned to Mazaceae (Reveal 2011) and Phrymaceae (Beardsley and Olmstead 2002; Tank et al. 2006) separately, recognized by APG IV (Angiosperm Phylogeny Group 2016). The sequences were aligned with MAFFT (Katoh and Standley 2013) performed in Geneious and the phylogenetic relationship was conducted using RAXML-HPC2 on XSEDE at CIPRES Science Gateway (Miller et al. 2010) with the GTR Gamma and 1000 bootstrap replicates based on a data matrix of concatenation of 78 coding protein sequences (CDS). The analysis fully resolved *C. rubella* was in a clade with *C. bodinieri* and *C. nudiflora* (Figure 1). The results indicated *C. rubella* and other species of *Callicarpa* located at the bottom of the phylogenetic tree in Lamiaceae and formed sister relationship with *Dicrastylis parvifolia* which verified the previous conclusions about the original position of *Callicarpa* in the family Lamiaceae (Bramley et al. 2009; Drew and Sytsma 2012; Li et al. 2016) (Figure 1). This study also provided important sequence information for the coming solution about the relationship between *C. rubella* and its related species.

## Disclosure statement

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

## Funding

This work was supported by National Natural Science Foundation of China under Grant [31760045 and 31970220] and Natural Science Foundation of Guangxi Province under Grant [2018GXNSFAA281132].

## Data availability statement

The complete chloroplast genome data that supports the findings of this study are openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov/> under the accession no. MZ520129. The associated BioProject, BioSample and SRA numbers are PRJNA744154, SAMN20080096 and SRR15049219, respectively.

## References

- Angiosperm Phylogeny Group. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Bot J Linn Soc.* 181:1–20.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19(5):455–477.
- Beardsley PM, Olmstead RG. 2002. Redefining Phrymaceae: the placement of *Mimulus*, tribe Mimuleae, and *Phryma*. *Am J Bot.* 89(7): 1093–1102.
- Bramley GLC, Forest F, de Kok RPJ. 2009. Troublesome tropical mints: re-examining generic limits of *Vitex* and relations (Lamiaceae) in South East Asia. *Taxon.* 58(2):500–510.
- Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19(1):11–15.
- Drew BT, Sytsma KJ. 2012. Phylogenetics, biogeography, and staminal evolution in the tribe Mentheae (Lamiaceae). *Am J Bot.* 99(5):933–953.
- Du YX, Liu YF, Liu B, Wang TL. 2020. Complete chloroplast genome of *Callicarpa formosana* Rolfe, a famous ornamental plant and traditional medicinal herb. *Mitochondrial DNA Part B-Resources.* 5(3):3383–3384.
- Ha YH, Choi KS, Choi K. 2018. Characterization of complete chloroplast genome of endemic species of Korea Peninsular, *Salvia chanryoenica* (Lamiaceae). *Mitochondrial DNA B Resour.* 3(2):992–993.
- Harley RM, Atkins S, Budantsey AL, Cantino PD, Conn BJ, Grayer R, Harley MM, de Kok RPK, Krestovskaja T, Morales R, et al. 2004. Labiatae. In: Kubitzki K, Kadereit JW, editors. *Families and genera of vascular plants, vol. 7, flowering plants: dicotyledons; Lamiales (except Acanthaceae including Avicenniaceae)*. Berlin: Springer; p. 167–275.
- Jin JJ, Yu WB, Yang JB, Song Y, dePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21(1):31.

- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz 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.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 9(4):357–U354.
- Li B, Cantino PD, Olmstead RG, Bramley GLC, Xiang CL, Ma ZH, Tan YH, Zhang DX. 2016. A large-scale chloroplast phylogeny of the Lamiaceae sheds new light on its subfamilial classification. *Sci Rep.* 6: 34343.
- Linnaeus C. 1753. *Species plantarum*. 1st ed. Stockholm: Laurentii Salvii.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *Proceedings of the Gateway Computing Environments Workshop (GCE)*. 14 Nov; New Orleans, LA. p. 1–8.
- Qu XJ, Moore MJ, Li DZ, Yi TS. 2019. PGA: a software package for rapid, accurate, and flexible batch annotation of plastomes. *Plant Methods.* 15(1):12.
- Reveal JL. 2011. Summary of recent systems of angiosperm classification. *Kew Bull.* 66(1):5–48.
- Tank DC, Beardsley PM, Kelchner SA, Olmstead RG. 2006. Review of the systematics of Scrophulariaceae s.l. and their current disposition. *Aust Systematic Bot.* 19(4):289–307.
- Tao AE, Zhao FY, Xia CL. 2019. Characterization of the complete chloroplast genome of *Ajuga forrestii* (Lamiaceae), a medicinal plant in southwest of China. *Mitochondrial DNA B Resour.* 4(2):3969–3970.
- Tu YH, Sun LN, Guo ML, Chen WS. 2013. The medicinal uses of *Callicarpa* L. in traditional Chinese medicine: an ethnopharmacological, phytochemical and pharmacological review. *J Ethnopharmacol.* 146(2): 465–481.
- Vallejo-Marin M, Cooley AM, Lee MY, Folmer M, McKain MR, Puzeý JR. 2016. Strongly asymmetric hybridization barriers shape the origin of a new polyploid species and its hybrid ancestor. *Am J Bot.* 103(7): 1272–1288.
- Wang CY, Luo JH, He HL, Wang Q. 2021. The complete chloroplast genome of *Callicarpa bodinieri* (Lamiales, Lamiaceae), an ornamental and medicinal plant from Chongqing, China. *Mitochondrial DNA B Resour.* 6(3):1229–1230.
- Wang HX, Chen L, Cheng XL, Chen WS, Li LM. 2019. Complete plastome sequence of *Callicarpa nudiflora* Vahl (Verbenaceae): a medicinal plant. *Mitochondrial DNA B Resour.* 4(2):2090–2091.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics.* 31(20): 3350–3352.
- Xie YQ, Luo ST, Zhang LT, Huang H, Zhang Q, Lai MY, Deng CY. 2021. The complete chloroplast genome of *Callicarpa siongsaiensis* Metcalf (Lamiaceae) from Fujian Province, China: genome structure and phylogenetic analysis. *Mitochondrial DNA B Resour.* 6(5):1634–1635.
- Zhao F, Chen YP, Salmaki Y, Drew BT, Wilson TC, Scheen AC, Celep F, Brauchler C, Bendiksby M, Wang Q, et al. 2021. An updated tribal classification of Lamiaceae based on plastome phylogenomics. *BMC Biol.* 19(1):2.
- Zhao F, Li B, Drew BT, Chen YP, Wang Q, Yu WB, Liu ED, Salmaki Y, Peng H, Xiang CL. 2020. Leveraging plastomes for comparative analysis and phylogenomic inference within Scutellarioideae (Lamiaceae). *Plos One.* 15(5):e0232602.
- Zhao F, Wu YW, Drew BT, Yao G, Chen YP, Cai J, Liu ED, Li B, Xiang CL. 2021. Systematic placement of the enigmatic Southeast Asian genus *Paralamium* and an updated phylogeny of tribe Pogostemoneae (Lamiaceae Subfamily Lamioideae). *Front Plant Sci.* 12:646133.