

Characteristics and phylogenetic analysis of the complete chloroplast genome of *Cornus hongkongensis* Hemsl. 1888 (Cornaceae)

Ze Liu^a, Liping He^b, Xiaorui Cui^c, Yong Zhou^b, Wenxia He^a, Yong Li^a, Haimei Chen^d and Junjie Wang^e

^aSchool of Nursing, Xiangnan University, Chenzhou, PR China; ^bSchool of Public Health, Xiangnan University, Chenzhou, PR China;

^cDepartment of Neurology, Affiliated Hospital of Xiangnan University, Chenzhou, PR China; ^dInstitute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing, PR China; ^eSchool of Pharmacology, Xiangnan University, Chenzhou, PR China

ABSTRACT

Cornus hongkongensis Hemsl. 1888, native to Hong Kong, belongs to the subgenus *Syncarpea* within the *Cornus* genus of the *Cornaceae* family. The complete chloroplast genome of *C. hongkongensis* spans 156,954 bp, comprising four subregions: a large single-copy region (86,290 bp), a small single-copy region (18,394 bp), and a pair of inverted repeats (26,135 bp). Within the chloroplast genome of *C. hongkongensis*, we identified 113 unique genes, including 80 protein-encoding genes, four ribosomal RNA (rRNA) genes, and 30 transfer RNA (tRNA) genes. Phylogenetic analysis based on the complete chloroplast genome of 30 related taxa of the *Cornus* genus indicates that *C. hongkongensis* has not formed a monophyletic lineage. Analyses of sequence divergence found three intergenic regions including *rps19-rpl22*, *ccsA-ndhD*, and *atpH-atpI*, exhibiting a high degree of variations. The first chloroplast genome of *C. hongkongensis* was reported in this work contributes to the enrichment of genomic data for the genus *Cornus*.

ARTICLE HISTORY

Received 13 November 2023
Accepted 29 February 2024

KEYWORDS

Chloroplast genome;
Cornaceae; *Cornus hongkongensis*; phylogenetic relationships

Introduction



Cornus hongkongensis Hemsl. 1888, a member of the subgenus *Syncarpea* within the *Cornus* genus of the *Cornaceae* family, is a small evergreen tree or shrub with distinctive four-flowered flowers native to regions including Hong Kong, Jiangxi, Fujian, and Zhejiang, typically thriving at altitudes ranging from 250 to 1200 m (Jijun et al. 2017). Renowned for its ecological and economic significance, this species is a visually captivating tree, harmoniously combining colorful flowers, leaves, and fruits, enhancing its ornamental appeal (Jijun et al. 2017). The fruit of *C. hongkongensis* is not only nutritionally rich and sweet but also suitable for winemaking. Moreover, its medicinal properties include promoting of menstrual regularity, enhancing of blood circulation, and anti-inflammatory effects. These attributes establish it as a rare and valuable wild fruit tree germplasm resource (Yu-Hong et al. 2003). Furthermore, its hardwood and robust root system make it an exceptional choice for specialized timber applications, soil stabilization, and water retention. Consequently, it plays a pivotal role in afforestation, landscape beautification, and soil fertility restoration (Kaveriappa et al. 1997).


As a result, the demand for *C. hongkongensis* in ornamental gardens has steadily increased (Dan-Qi and Cheng-Xiang

2008). Notably, *C. hongkongensis*, a species within this category found in Hong Kong, lacks a publicly available chloroplast genome, hindering a precise phylogenetic understanding. To address this knowledge gap, our study aims to construct a high-quality assembled chloroplast genome, offering deeper molecular insights into germplasm relationships, genetic diversity, and phylogeny within this botanical context.

Materials and methods

Fresh leaves of *C. hongkongensis* (Figure 1) were collected with permission from the Guidong Botanical Garden, located in Chenzhou City, Hunan Province, China (latitude 24.98N, longitude 112.89E). The specimen was deposited at the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing (contact person: Haimei Chen, hmchen@implad.ac.cn) under the voucher number Implad20230711004. Subsequently, the sequencing library with 200–400 bp insert sizes was constructed and sequenced on the DNBSEQ-T7RS platform (MGI) at Grandomics (Beijing, China). This process yielded a total of 10.8 Gb of raw data. The raw reads were then filtered using Trimmomatic (Bolger et al. 2014) to remove adaptors and low-quality sequences. As a result, a total of 10.6 GB of paired-end data were retrieved. The clean data were

CONTACT Junjie Wang  422967089@qq.com  No. 889, Chenzhou Avenue, Suxian District, Chenzhou 423000, PR China

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/23802359.2024.2326825>.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

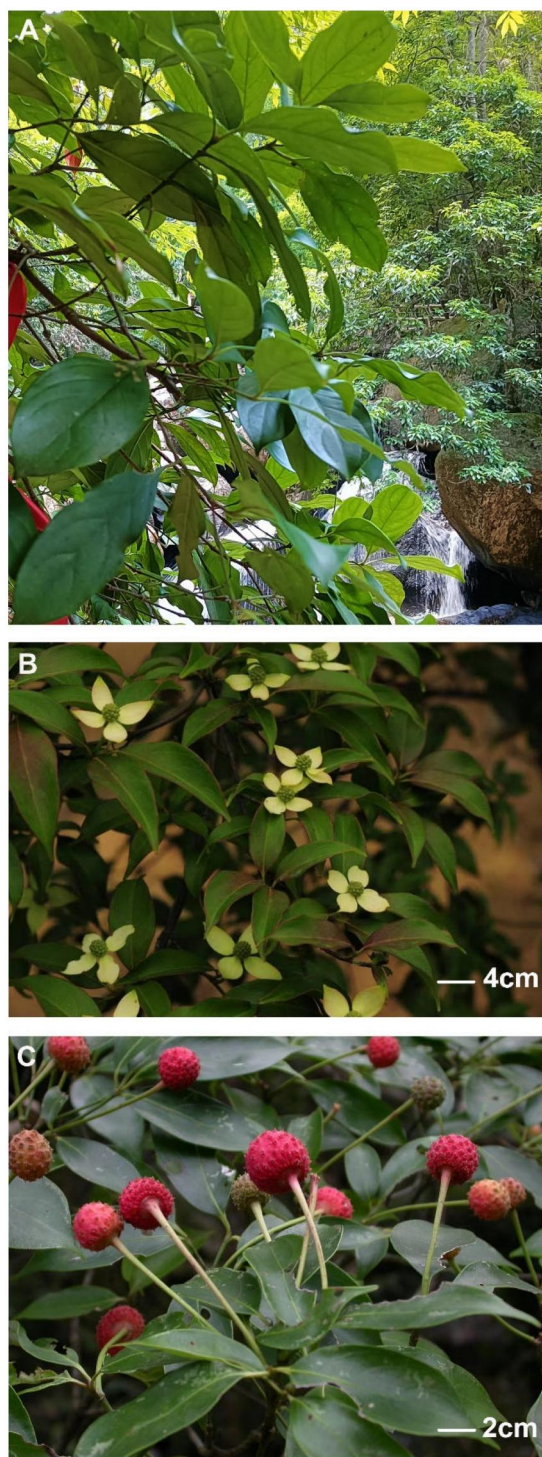


Figure 1. Photograph of *C. hongkongensis* leaf, flower, and fruit. (A) Leaf. This image was taken by Lichai Yuan at Guidong Botanical Garden, Chenzhou city, Hunan province, China. The leaf blade displays an elliptic, oblong-elliptic, or obovate-oblong shape with a thinly to thickly leathery. The leaf base is cuneate or broadly cuneate to rounded, and the apex is shortly acuminate to caudate. Typically, there are 3 or 4 (or occasionally 5) veins that ascend in a curved pattern. (B) Flower. This image was taken by Jun Liu at Hangzhou City, Zhejiang Province, China. Capitulate cymes are globose, and the bracts are yellowish or white, broadly elliptic, broadly ovate, or orbicular to obovate. (C) Fruit. This image was taken by Yan Liu. The compound fruit, which is globose in shape, turns red or yellowish-red at maturity.

employed to *de novo* assemble the plastid genome using GetOrganelle (Jin et al. 2020) with default parameters. To assess the quality of the assembly, the clean data were

mapped onto the assembly result using minimap2 (Li 2018). The aligned reads were extracted, and the depth and coverage were calculated using Samtools (v1.12; Danecek et al. 2021). The annotations for the plastid genome of *C. hongkongensis* were generated by CPGAVAS2 (Shi et al. 2019) with the reference annotation of *Cornus capitata* (NC_060994.1) and adjusted manually using Apollo (Misra and Harris 2005). The map of the plastid genome, cis-spliced genes, and trans-spliced genes of *C. hongkongensis* was drawn by CPGview (Liu et al. 2023). The phylogenetic relationships among *C. hongkongensis* and 29 other *Cornus* species were analyzed based on the complete plastome sequences. Two *Alangium* species, namely, *Alangium alpinum* (MG525003.1) and *Alangium chinense subsp. strigosum* (OR197702.1) were chosen as outgroups. The complete plastomes of 32 species were aligned by MAFFT (Katoh and Standley 2013) with default parameters. The poorly aligned positions and divergent regions of the alignment were eliminated using Gblocks (v0.91b; Talavera and Castresana 2007). The maximum-likelihood (ML) tree was constructed in IQtree (v2.1.2 (Minh et al. 2020)) with ultrafast bootstrap (1000 replicates) and best-fit model TVM + F + R3. The intergenic spacer (IGS) regions of 25 *Cornus* cp genomes were extracted with our custom script. Then, Distmat program that was implemented in EMBOSS (v6.3.1) (Rice et al. 2000) was used to determine the pairwise distance of the intergenic regions with the Kimura 2-parameters (K2p) evolution model.

Results

A total of 751,194 paired reads were assembled to the complete plastid genome of *C. hongkongensis*, revealing a median depth of 1413, with a minimum depth of 540 and a maximum depth of 2213 (Figure S1). The assembled plastid genome has been deposited to the NCBI (National Center for Biotechnology Information, <https://www.ncbi.nlm.nih.gov/>) under the accession number OR545260. The plastome exhibits a quadripartite structure with 156,954 bp in length, consisting of a pair of inverted repeats (IRs, 26,135 bp), a large single-copy (LSC, 86,290 bp) region, and a small single-copy (SSC, 18,394 bp) (Figure 2). The overall GC content of *C. hongkongensis* plastome was 38%, with the LSC, SSC, and IRs regions exhibiting GC contents of 36%, 32%, and 43%, respectively. A total of 132 genes (114 unique genes) were annotated, including 85 protein-coding genes (PCGs, 80 are unique), 37 transfer RNA genes (tRNAs, 30 are unique), eight ribosomal RNA genes (rRNAs, four are unique). Among them, a total of 15 unique genes (including *rps16*, *atpF*, *rpoC1*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhB*, *ndhA*, *trnK-UUU*, *trnG-UCC*, *trnL-UAA*, *trnV-UAC*, *trnI-GAU*, and *trnA-UGC*) containing one intron and two unique genes containing two introns (including *ycf3* and *clpP*) (Figure S2). The *rps12* is a trans-spliced gene as shown in Figure S3. Sixteen genes were duplicated in IR regions, including seven PCGs (*rpl2*, *rpl23*, *ycf15*, *ycf2*, *ndhB*, *rps7*, and *rps12*), six tRNA genes (*trnN-GUU*, *trnR-ACG*, *trnA-UGC*, *trnI-GAU*, *trnV-GAC*, and *trnL-CAA*), and four rRNA genes (*rnr16*, *rnr23*, *rnr4.5*, and *rnr5*).

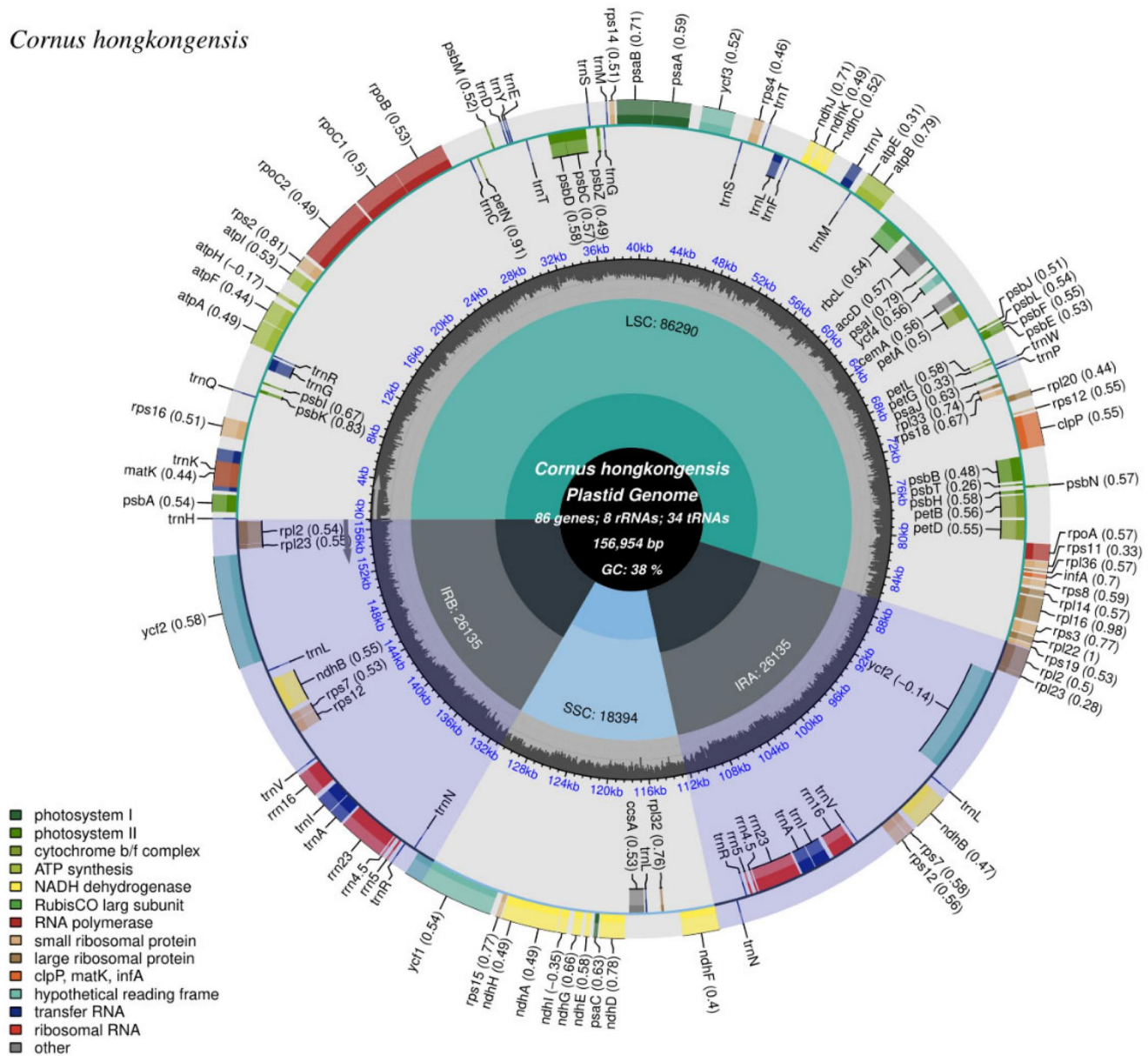
Cornus hongkongensis

Figure 2. The complete plastome map of *C. hongkongensis*, which was generated by CPGview. LSC, SSC, and IRs (IRa and IRb) with their length are represented on the first circle. The second circle showed the GC ratio in dark gray. The outermost circle indicates gene name color-coded by their functional classification. The transcription directions for the inner and outer genes are clockwise and anticlockwise, respectively. The functional classification of the genes is shown in the left bottom corner. The optional codon usage bias is displayed in the parenthesis after the gene name.

A ML phylogenetic analysis based on the complete chloroplast genome of 25 related taxa of the *Cornus* genus showed that the genus *Cornus* forms four major clades: the big-bracted (BB) group, the dwarf dogwoods (DW), the cornelian cherries (CC), and the blue- or white-fruited group (BW) (Figure 3). Subgenus *Syncarpea*, encompassing *C. hongkongensis*, constitutes a monophyletic group. *C. hongkongensis* (OR545260) forms a distinct cluster alongside *subsp. elegans*, *subsp. gigantea*, and *subsp. tonkinensis*, exhibiting a robust statistical support value of 88%. While *C. hongkongensis subsp. ferruginea* clusters with *Cornus kousa* and *C. hongkongensis subsp. melanotricha* clusters with *Cornus sunhangii*.

The pairwise comparison of IGS regions was conducted among the 25 *Cornus* species to identify hypervariable regions using the K2p model. The K2p distance, spanning from 0.00 to 44.46, was observed among the 124 IGSs of the

25 *Cornus* species (Table S1). Among them, the IGS regions *rps19-rpl22*, *ccsA-ndhD*, and *atpH-atpI* showed average distances of 11.88, 11.25, and 9.85, respectively, among the 25 *Cornus* species (Figure 4). These specific IGS regions present potential utility as molecular markers for discerning plant phylogeny at lower taxonomic levels and for DNA barcoding within *Cornus* species.

Discussion and conclusions

Cornus encompasses approximately 58 species, primarily comprising hermaphroditic shrubs and small trees. These *Cornus* species widely distribute in the temperate and subtropical regions of the Northern Hemisphere, with rare occurrences in tropical areas. The taxonomic composition, ranking, and relationships of subgroups within the genus have been

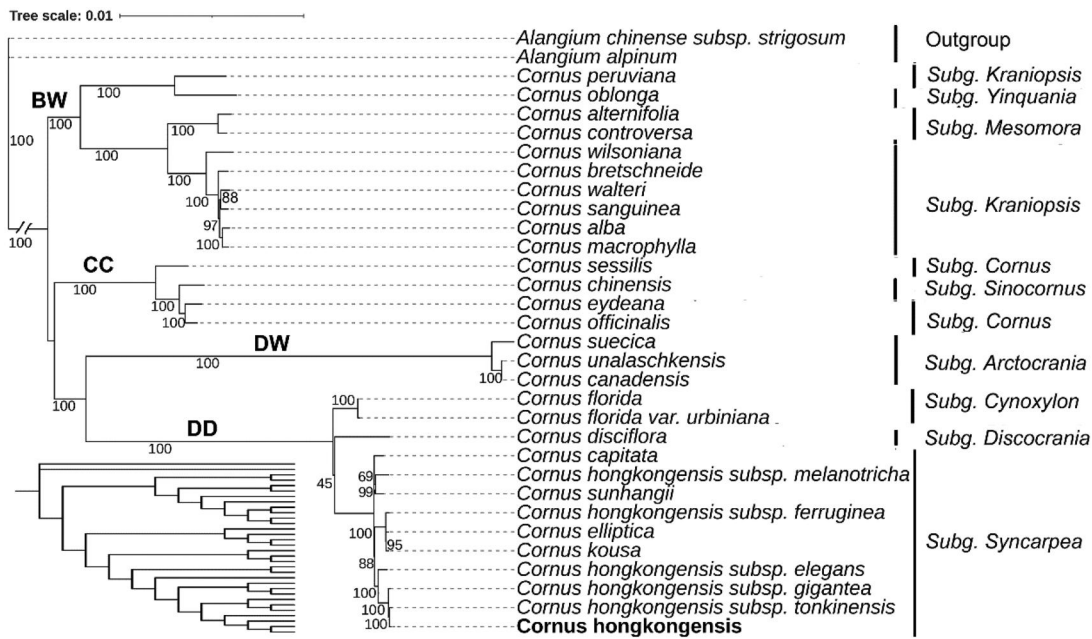


Figure 3. A maximum-likelihood (ML) based phylogenetic tree of *C. hongkongensis* and related *Cornus* species. The cladogram was shown in the lower left corner. The scale bar represents the number of nucleotide substitutions per site. The numbers on each node indicated the ML bootstrap support with 1000 replicates. The subgenus is labeled next to the species name. The bold font indicates that genus *Cornus* was classified into four major clades: the big-bracted (BB) group, the dwarf dogwoods (DW), the cornelian cherries (CC), and the blue- or white-fruited group (BW). The following sequences were used: *Alangium alpinum* (MG525003.1) (Fu et al. 2017), *Alangium chinense* subsp. *strigosum* (OR197702.1) (Yang et al. 2022), *Cornus peruviana* (NC_044825.1) (Fu et al. 2019), *Cornus oblonga* (NC_044811.1) (Fu et al. 2019), *Cornus alternifolia* (NC_044812.1) (Fu et al. 2019), *Cornus controversa* (NC_030260.1), *Cornus wilsoniana* (NC_063837.1), *Cornus bretschneideri* (NC_060800.1) (Li et al. 2020), *Cornus walteri* (NC_058318.1), *Cornus sanguinea* (NC_044817.1) (Fu et al. 2019), *Cornus alba* (NC_059720.1) (Yuan et al. 2021), *Cornus macrophylla* (NC_044810.1) (Fu et al. 2019), *Cornus sessilis* (NC_044814.1) (Fu et al. 2019), *Cornus chinensis* (NC_044815.1) (Fu et al. 2019), *Cornus eydeana* (NC_044816.1) (Fu et al. 2019), *Cornus officinalis* (NC_042746.1), *Cornus suecica* (NC_044823.1) (Fu et al. 2019), *Cornus unalaschkensis* (NC_044824.1) (Fu et al. 2019), *Cornus canadensis* (NC_044822.1) (Fu et al. 2019), *Cornus disciflora* (NC_044819.1) (Fu et al. 2019), *Cornus florida* (NC_044820.1) (Fu et al. 2019), *Cornus florida* var. *urbiniana* (MN380671.1) (Fu et al. 2019), *Cornus sunhangii* (NC_060994.1) (Lv et al. 2019), *Cornus capitata* (MG524998.1) (Fu et al. 2017), *Cornus hongkongensis* (OR545260.1), *Cornus elliptica* (NC_056267.1) (Lu et al. 2021), *Cornus kousa* (NC_044818.1) (Fu et al. 2019), *Cornus hongkongensis* subsp. *Gigantea* (OR597583.1), *Cornus hongkongensis* subsp. *Tonkinensis* (OR597581.1), *Cornus hongkongensis* subsp. *Elegans* (OR597582.1), *Cornus hongkongensis* subsp. *Ferruginea* (OR597579.1), and *Cornus hongkongensis* subsp. *Melanotricha* (OR597580.1).

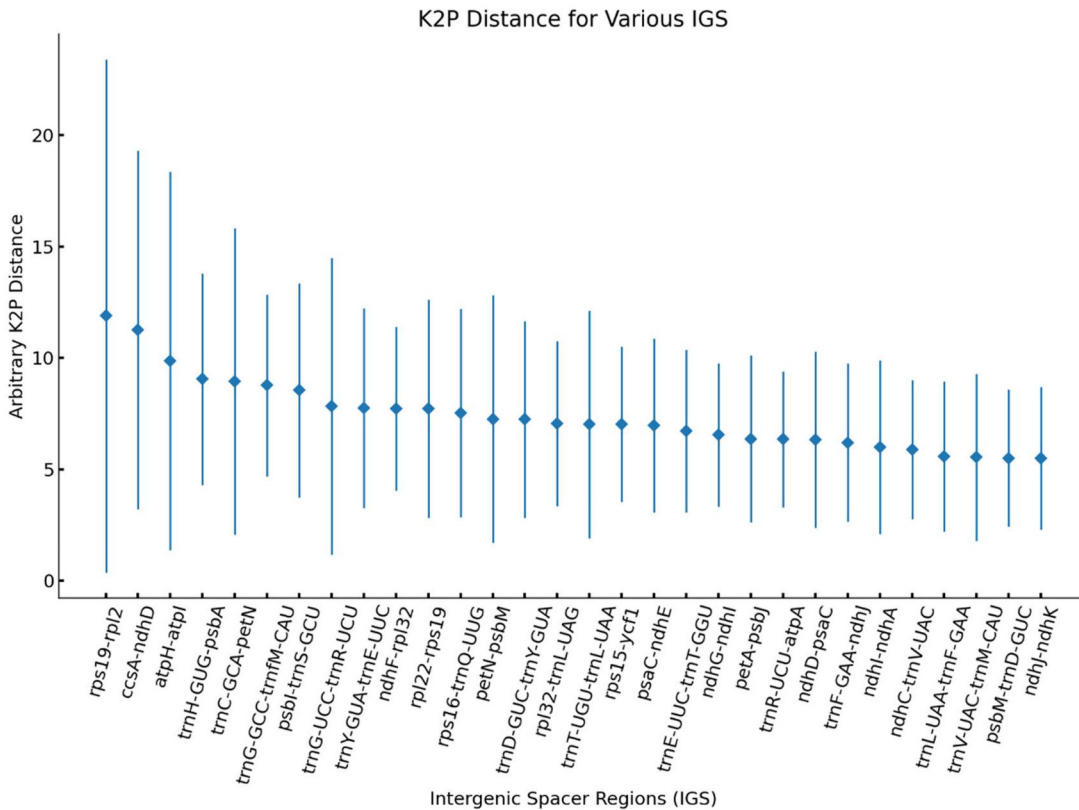


Figure 4. Results of genetic distance analysis of intergenic spacer regions in *Cornus* species. The top 30 IGS of K2p distances are shown. The X-axis indicates the name of IGS regions. And the Y-axis shows the range of K2p distances between different pairs of species. The diamond shows the average K2p distance.

subjects of contention for nearly a century. The phylogenetic analysis utilizing complete plastome sequences in this study supported the monophyly of most subgenera, except *subg. Kraniopsis* and *subg. Cornus*. *C. peruviana* belonging to *Subg. Kraniopsis* was clustered with *C. oblonga* belonging to *Subg. Yinquania*. Additionally, *subg. Sinocornus* were nested within *subg. Cornus*. This finding is consistent with previous studies, which integrated morphology, *matK*, *ITS*, and previously published *rbcl* and 26S rDNA sequences (Xiang et al. 2006).

Xiang (1987) recognized six subspecies within *C. hongkongensis*, including *subsp. gigantea*, *subsp. tonkinensis*, *subsp. elegans*, *subsp. ferruginea*, *subsp. melanotricha*, and *subsp. hongkongensis* based on their habits and buds. *Subsp. C. hongkongensis* is much controversial owing to the diversity of evaluation of characters. Our phylogenetic analysis based on the whole cp genomes showed that six subspecies within *C. hongkongensis* were not clustered together. As a result, a reasonable system reflecting the infra-subspecies relationships is still lacking and further data will be needed to reevaluate the relationship of subspecies within *C. hongkongensis*. The result enriches the genomic data for the genus *Cornus*, which will contribute to phylogenetic and evolutionary studies in the future. In this study, we found that the K2p values were particularly high for the three IGS regions: *rps19-rpl22*, *ccsA-ndhD*, and *atpH-atpI*. These IGS regions can be used to distinguish *Cornus* species.

Author contributions

ZL and JJW made substantial contributions to the design of this work and final approval of the version to be published; LPH, XRC, and YZ analyzed the data; WXH and YL drafted this work; HMC and JJW reviewed this work critically. All authors agreed to be accountable for all aspects of this work.

Ethical approval

The sample of *Cornus hongkongensis* was collected with the permission from the Guidong Botanical Garden, Chenzhou city, Hunan Province, China and strictly complied with local and Chinese regulations. No ethical approval is required in this study.

Disclosure statement

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

Funding

This work was financial supported by The Research Foundation for Advanced Talents of Xiangnan University [063, 6011901] and The Scientific Research Plan Project of Hunan Provincial Health Commission [202203053563].

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at <https://www.ncbi.nlm.nih.gov/> under the accession number OR545260.1. The associated BioProject, BioSample, and SRA numbers are PRJNA1033361, SAMN38031310, and SRR26559992, respectively.

References

- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30(15):2114–2120. doi:10.1093/bioinformatics/btu170.
- Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience*. 10(2):giab008. doi:10.1093/gigascience/giab008.
- Dan-Qi X, Cheng-Xiang LU. 2008. Germplasm resource distribution and recent research advance in *Dendrobenthamia*. *J Gansu Forest Sci Technol*. 33:24–26, 68.
- Fu CN, Li HT, Milne R, Zhang T, Ma PF, Yang J, Li DZ, Gao LM. 2017. Comparative analyses of plastid genomes from fourteen Cornales species: inferences for phylogenetic relationships and genome evolution. *BMC Genomics*. 18(1):956. doi:10.1186/s12864-017-4319-9.
- Fu CN, Mo ZQ, Yang JB, Ge XJ, Li DZ, Xiang QJ, Gao LM. 2019. Plastid phylogenomics and biogeographic analysis support a trans-Tethyan origin and rapid early radiation of Cornales in the Mid-Cretaceous. *Mol Phylogenet Evol*. 140:106601. doi:10.1016/j.ympev.2019.106601.
- Jijun X, Biting HE, Zhaoyou XU, Huanwei C, Jun LI. 2017. A study on container seedling quality grading of *Cornus hongkongensis* Hemsl. ssp. *elegans*. *J Fujian Forest Sci Technol*. 44(4):102–106. doi:10.13428/j.cnki.fjlk.2017.04.022.
- 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.
- 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.
- Kaveriappa KM, Phillips LM, Trigiano RN. 1997. Micropropagation of flowering dogwood (*Cornus florida*) from seedlings. *Plant Cell Rep*. 16(7):485–489. doi:10.1007/bf01092771.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 34(18):3094–3100. doi:10.1093/bioinformatics/bty191.
- Li X, Ma Q, Zhou H, Yang Y, Li H, Wang J. 2020. Characterization of the complete chloroplast genome of *Cornus bretschneideri* (Cornaceae). *Mitochondrial DNA B Resour*. 5(1):543–544. doi:10.1080/23802359.2019.1710281.
- 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.
- Lu H, Tang J, Li D. 2021. The complete chloroplast genome sequences of *Cornus elliptica* (Cornaceae). *Mitochondrial DNA B Resour*. 6(3):931–932. doi:10.1080/23802359.2021.1888333.
- Lv ZY, Huang XH, Luo J, Zhang X, Deng T, Li ZM. 2019. The complete chloroplast genome sequence of *Cornus sunhangii* (Cornaceae). *Mitochondrial DNA B Resour*. 4(2):3242–3243. doi:10.1080/23802359.2019.1669090.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 37(5):1530–1534. doi:10.1093/molbev/msaa015.
- Misra S, Harris N. 2005. Using Apollo to browse and edit genome annotations. *Curr Protoc Bioinform*. 9(1):9.5.1–9.5.28. doi:10.1002/0471250953.bi0905s12.
- Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet*. 16(6):276–277. doi:10.1016/s0168-9525(00)02024-2.
- Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. *Nucleic Acids Res*. 47(W1):W65–W73. doi:10.1093/nar/gkz345.
- Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol*. 56(4):564–577. doi:10.1080/10635150701472164.
- Xiang Q-Y, Thomas DT, Zhang W, Manchester SR, Murrell Z. 2006. Species level phylogeny of the genus *Cornus* (Cornaceae) based on molecular and morphological evidence – implications for taxonomy and tertiary intercontinental migration. *TAXON*. 55(1):9–30. doi:10.2307/25065525.

- Xiang Q-Y. 1987. System and synopsis of *Cornus* subgen. *Syncarpea* (Nakai) Q. Y. Xiang (Cornaceae). Bull Bot Res. 7(2):33–52.
- Yang X, Liu C, Zeng X, Liu X, Zhao H, Feng T, Zhou Y. 2022. Sequence structure and phylogenetic analysis of the chloroplast genomes of *Alangium chinense* (Lour.) Harms and its different subspecies. Yaoxue Xuebao. 57:3229–3239.
- Yuan W, He S, Zhang S, Chang D, He Y. 2021. The complete chloroplast genome sequence of *Cornus alba* L. (Cornaceae). Mitochondrial DNA B Resour. 6(7):1997–1998. doi:10.1080/23802359.2021.1938727.
- Yu-Hong Z, Guo-Zheng T, Ming-Shu XU, Yong-Mei Yi. 2003. Determination of main nutritious ingredients in the fruit of *Dendrobenthamia angustata*. J Hubei Inst Natl. (4):21–22.