

The chloroplast genome of the *Iris japonica* Thunberg (Butterfly flower) reveals the genomic and evolutionary characteristics of *Iris* species

Xinyi Zhang*, Heyu Yang*, Bin Wu and Haimei Chen

Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, P.R. China

ABSTRACT

Iris japonica Thunberg is one of the horticultural species belonging to the *Iris* genus and Iridaceae family. Previous studies have revealed its hepatoprotective activity and ornamental values. However, little genetic and genomic information about this species is available. Here, to decipher the chloroplast genome and reveal its evolutionary characteristics, we sequenced, *de novo* assembled, and comprehensively analyzed the chloroplast genome of *I. japonica*. The genome was 152,453 bp in length and displayed a circular structure with a large single-copy region, a small single-copy region, and two inverted repeat regions. It contained 131 genes, including 85 protein-coding genes, eight ribosomal RNA genes, and 38 transfer RNA genes. We also identified 23 microsatellite repeat sequences, 34 tandem repeat sequences, and 60 dispersed repeat sequences in the chloroplast genome of *I. japonica*. Sequence divergence analyses of the chloroplast genomes of 20 *Iris* species revealed that the top four most highly variable regions were *ndhC-trnV-UAC*, *rpl22-rps19*, *rps16-trnQ-UUG*, and *trnG-UCC-trnR-UCU*. Phylogenetic analysis showed that *I. japonica* was most closely related to *I. tectorum*. This study reported a new chloroplast genome of *I. japonica* and performed comparative analyses of 20 *Iris* chloroplast genomes. The results would facilitate the evolutionary research and development of molecular markers for *Iris* species.

ARTICLE HISTORY

Received 27 June 2022
Accepted 24 August 2022

KEYWORDS

Iris; chloroplast genome; hypervariable region; selective pressure analysis; phylogenetic analysis

1. Introduction

The butterfly flower (*Iris japonica* Thunberg) belongs to the Iridaceae family and is native to Japan and China. The butterfly flower has ornamental value for its beautiful butterfly-shaped flowers called 'Hudie Hua' in China. Moreover, *I. japonica* is a medicinal plant for treating bronchitis, internal injuries, rheumatism, and swelling. The butterfly flower has many biological activities, such as antioxidant, anti-mutagenic, anti-angiogenic, anti-inflammatory, and hypoglycemic functions (Xu et al. 2021). Although many studies about the medicinal value of *I. japonica* have been conducted, no work has studied its genetic and genomic information, limiting the species identification and evolutionary analysis of *I. japonica*.

Complete chloroplast genomes are circular, linear, or polycyclic double-stranded DNA molecules, with a length mostly around 120–160 kb (Bock 2015). The complete chloroplast genome sequence can be used for the development of DNA barcodes, the determination of phylogenetic relationships, and the identification of patterns of gene loss and adaptive changes that optimize photosynthesis (Olejniczak et al. 2016).

Recently, many chloroplast genomes have been deciphered due to the advanced genome sequencing technologies and bioinformatics tools. One study sequenced and

assembled the chloroplast genomes of 14 Korean-native *Iris* species (Kang et al. 2020). The 14 *Iris* chloroplast genomes were compared, and Bayesian phylogenetic trees were constructed. Five other studies reported the sequencing and assembly of the chloroplast genomes of *Iris sanguinea* (Lee et al. 2017), *Iris domestica* (Ai et al. 2019), *Iris loczyi* (Choi et al. 2020), *Iris tectorum* (Liu et al. 2020), and *Iris lactea* var. *chinensis* (Cai et al. 2021). Additionally, one assembly of the chloroplast genome of *I. japonica* was available in Genbank (NC_060499.1). However, a detailed analysis of this assembly was not reported.

Here, we sequenced and assembled the chloroplast genome of *I. japonica* and conducted a comparative analysis with 19 other *Iris* species. Our study presented the very first detailed analysis of the *I. japonica* chloroplast genome. The results will lay a solid foundation for the future development of *I. japonica*-based medicine.


2. Materials and methods

2.1. Plant materials and DNA extraction

We collected the fresh leaves of the *I. japonica* plants from Guangxi Province, China (E104°26', N20°54'). The voucher

CONTACT Haimei Chen  hmchen@implad.ac.cn; Bin Wu  bwu@implad.ac.cn  Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 151, Malianwa North Road, Haidian District, Beijing 100093, P.R. China

*These authors contributed equally to this work.

 Supplemental data for this article is available online at <https://doi.org/10.1080/23802359.2022.2118000>

© 2022 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 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

specimen (IMPLADS2021009040) and its DNA sample (IMPLADS2021009040_dna01) were deposited at the Herbarium of Institute of Medicinal Plant Development (IMD), China (<http://www.implad.ac.cn/>, Haimei Chen, e-mail: hmchen@implad.ac.cn). We extracted the total DNA of one *I. japonica* sample using the plant genomic DNA kit (Tiangen Biotech, Beijing, Co., Ltd.). Total DNA was separated with electrophoresis in 1.2% agarose gels and analyzed with the Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) to assess quality and purity. Our study, including sample collection, was conducted in compliance with relevant institutional, national, and international guidelines and laws.

2.2. Plastome sequencing, assembly, and annotation

The 1 µg total DNA of *I. japonica* was used to construct the sequencing libraries using the TruSeq DNA Sample Prep Kit (Illumina, Inc., San Diego, CA, USA) following the manufacturer's instructions and the libraries were sequenced on the Illumina NovaSeq PE150 instrument (Illumina Inc., San Diego, CA, USA) (Caporaso et al. 2012).

The genome was *de novo* assembled using NOVOPlasty (version 3.8.3) with the default parameters (Dierckxsens et al. 2017). The chloroplast genome of *I. japonica* was annotated with CPGAVAS2 (version 2.0) web service (Shi et al. 2019). The circular gene map of the chloroplast genome was drawn by the cpgavas2 web server (<http://47.96.249.172:16019/analyzer/view>). The raw data, final genome assembly sequence, and annotated information were deposited in CNCB-NGDC, with the accession numbers SRR18908135 and GWHBISG01000000.

2.3. Comparative genome analysis

The complete chloroplast genome sequences of 20 *Iris* species were downloaded from the GenBank database (Supplementary Material, Table S1), and the *I. japonica* chloroplast genome was obtained in this study (GWHBISG01000000). We conducted the comparative analysis of these 20 genomes using mVISTA with the ShuffleLAGAN mode (Brudno et al. 2003). The genes around the border of the IR, LSC, and SSC regions of the 21 *Iris* species were analyzed using IRscope online software (Amiryousefi et al. 2018) (<https://irscope.shinyapps.io/irapp/>). The genetic distance of the intergenic spacer (IGS) was calculated using the distmat program from EMBOSS (v6.3.1) with the Kimura 2-parameter (K2P) evolutionary model (Rice et al. 2000). We also compared our new chloroplast genome with the one deposited in the Genebank (NC_060499) using SeqMan software (version 7.1.0).

2.4. Selective pressure analysis of protein-coding genes

The selective pressure of the protein-coding genes (PCGs) in the chloroplast genome was analyzed based on nucleotide substitution rate variation (Song et al. 2020). To detect the genes of the *I. japonica* chloroplast genome that were under selection, we extracted 71 common PCGs among the 20 *Iris* chloroplast genomes, performed multiple sequence alignment

using MAFFT (v7.313), and constructed a maximum likelihood (ML) tree using IQTREE (v1.6.10) (Nguyen et al. 2015). The selective pressure of PCGs was analyzed using the aBSREL (adaptive branch-site random effects likelihood) model implemented in Hyphy software (v2.2.4) (Smith et al. 2015).

2.5. Phylogenetic analysis

Phylogenetic analysis was performed using the chloroplast genomes of 33 *Iris* species, and *Crocus cartwrightianus* (NC_041459) was selected as the outgroup. We used the PhyloSuite (v1.1.16) (Zhang et al. 2020) to extract the coding sequences of 66 common PCGs (*atpA*, *atpB*, *atpE*, *atpF*, *atpH*, *ccsA*, *cemA*, *infA*, *matK*, *nadA*, *ndhB*, *ndhC*, *ndhE*, *ndhF*, *ndhG*, *ndhH*, *ndhJ*, *ndhK*, *petA*, *petB*, *petD*, *petG*, *petL*, *petN*, *psaA*, *psaB*, *psaC*, *psaJ*, *psbA*, *psbC*, *psbD*, *psbE*, *psbF*, *psbH*, *psbI*, *psbJ*, *psbK*, *psbL*, *psbM*, *psbT*, *rbcL*, *rpl14*, *rpl16*, *rpl20*, *rpl22*, *rpl23*, *rpl32*, *rpl33*, *rpl36*, *rpoA*, *rpoB*, *rpoC1*, *rpoC2*, *rps2*, *rps3*, *rps4*, *rps7*, *rps8*, *rps11*, *rps12*, *rps14*, *rps15*, *rps16*, *rps18*, *ycf1* and *ycf2*) of the chloroplast genome of these 34 species. Then the coding sequences of the 66 shared PCGs were aligned using MAFFT (v7.313) (Kato et al. 2002). The phylogenetic tree was constructed using the ML method implemented in IQ-TREE (v1.6.8) with the TVM+I+G4 nucleotide substitution model (Smith et al. 2015). Finally, we assessed the reliability of the phylogenetic tree by the bootstrap test with 1000 replications and visualized the final tree using iTOL (<https://itol.embl.de/>) (Letunic and Bork 2019).

3. Results

3.1. Genome organization and compositions

The chloroplast genome of *I. japonica* presented a typical circular DNA molecule with a total length of 152,453 bp (Figure 1). It had a conserved tetrapartite structure, including an LSC region, an SSC region, and a pair of IR regions, with lengths of 83,252, 18,489, and 25,356 bp, respectively. The total GC content in the chloroplast genome of *I. japonica* was 37.84%. The GC content of the IR region (43.03%) was higher than those of the SSC (31.38%) and LSC (36.12%) regions.

131 genes, including 85 PCGs, 38 tRNA genes, and eight rRNA genes, which consisted of two copies of *rnn16S*, *rnn23S*, *rnn4.5S*, and *rnn5S* genes were identified in the chloroplast genome of *I. japonica*. The gene structures of 21 *cis*-splicing genes are shown in Figure S1 (Supplementary Material). The 21 *cis*-splicing genes contained 13 PCGs and eight tRNA genes. All the *cis*-splicing genes had only one intron each except *ycf3* and *clpP*, which contained two introns (Supplementary Material, Table S2).

3.2. Comparative analysis of 20 *Iris* chloroplast genomes

The chloroplast genome sequences of 21 species including the 19 species listed in Table S1 (Supplementary Material), *I. japonica* (GWHBISG01000000) obtained in this study, and *I. japonica* (NC_060499) were aligned using mVISTA to

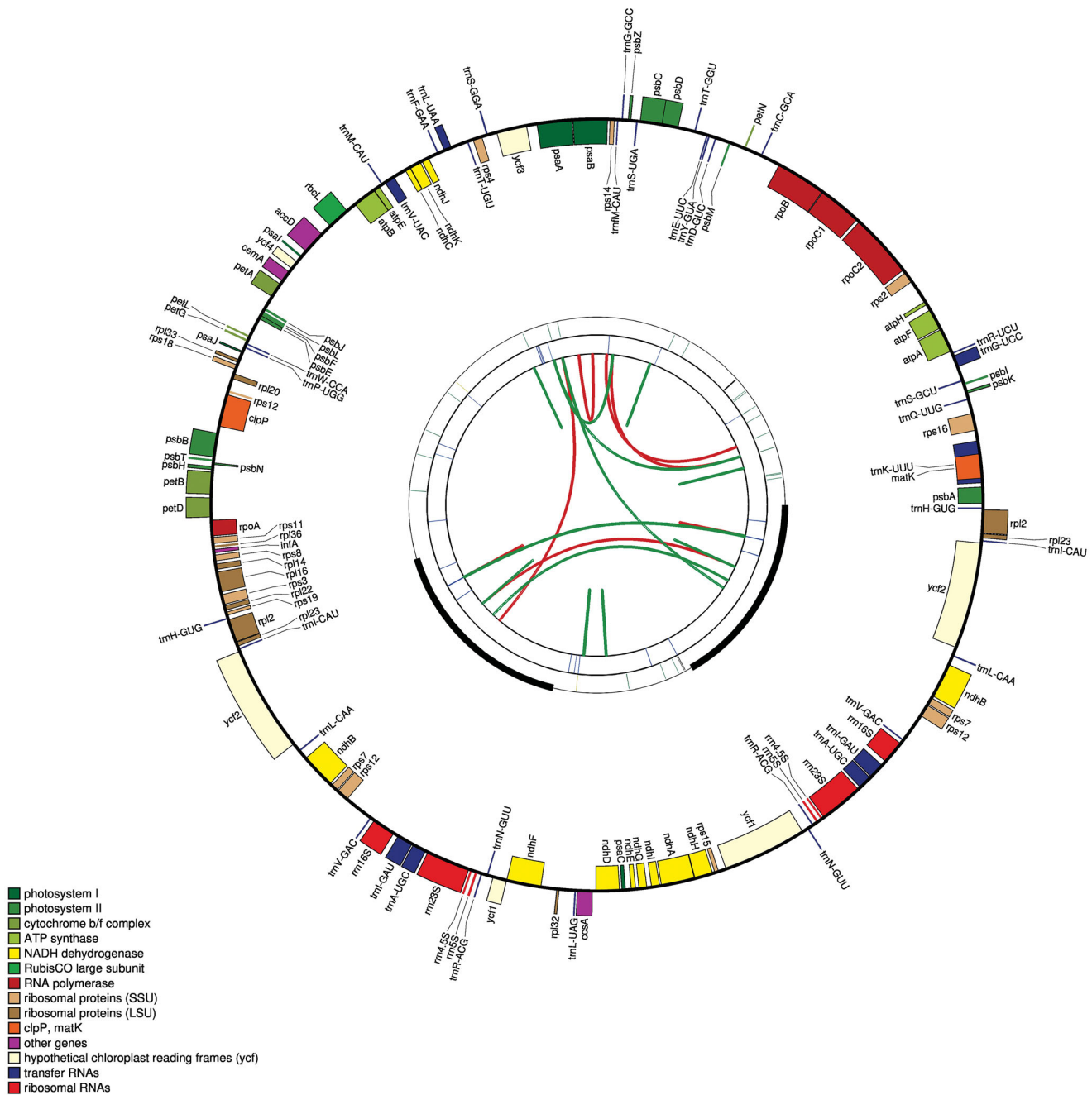


Figure 1. Graphic representation of features identified in *I. japonica* chloroplast genome using CPGAVAS2 (<http://47.96.249.172:16019/analyzer/view>). The map contains four circles. From the center going outward, the first circle shows the distributed repeats connected with red (the forward direction) and green (the reverse direction) arcs. The next circle shows the tandem repeats marked with short bars. The third circle shows the LSC, SSC, IRa, and IRb regions. The microsatellite sequences are shown as short bars on the circle. The fourth circle shows the genes having different colors based on the functional groups. The genes of the outside of the circle are transcribed clockwise. And the genes on the inside of the circle are transcribed anticlockwise. The functional classification is shown at the bottom left.

identify the variations in nucleotide sequence. The results showed that some highly variable regions were detected, including the protein-coding region of *ycf1*, *petB*, and *petD* and the IGS regions of *rps16-trnQ-UUG*, *rpoB-trnC-GCA*, *ndhC-trnV-UAC*, *rpl33-rps18*, and *rpl22-rps19* (Figure 2). The alignment of the chloroplast genomes of the 14 other species is shown in Figures S2 and S3 (Supplementary Material).

3.3. *Ir* structure analysis of 20 *Iris* species

IR boundary analysis of 21 complete chloroplast genomes of *Iris* species showed that most had the same boundary structure (Figure 3). For the 20 species except for *I. japonica*, both *rpl22*

and *rps19* genes were located around the border area of the LSC and IRb regions. By contrast, the *rps19* gene of *I. japonica* (GWHBISG01000000) was present in the LSC region. Besides, the *ndhF* and *ycf1* genes of *I. tectorum*, *I. japonica*, and *I. domestica* were located at the border area of IRb and SSC. The *ycf1* gene of most species spanned the SSC and IRb regions. The *psbA* genes were found at the border area of LSC and IRa.

3.4. Hypervariable region identification and selective pressure analysis

To find highly variable regions among the 20 *Iris* chloroplast genomes, we calculated the genetic distance of IGS regions

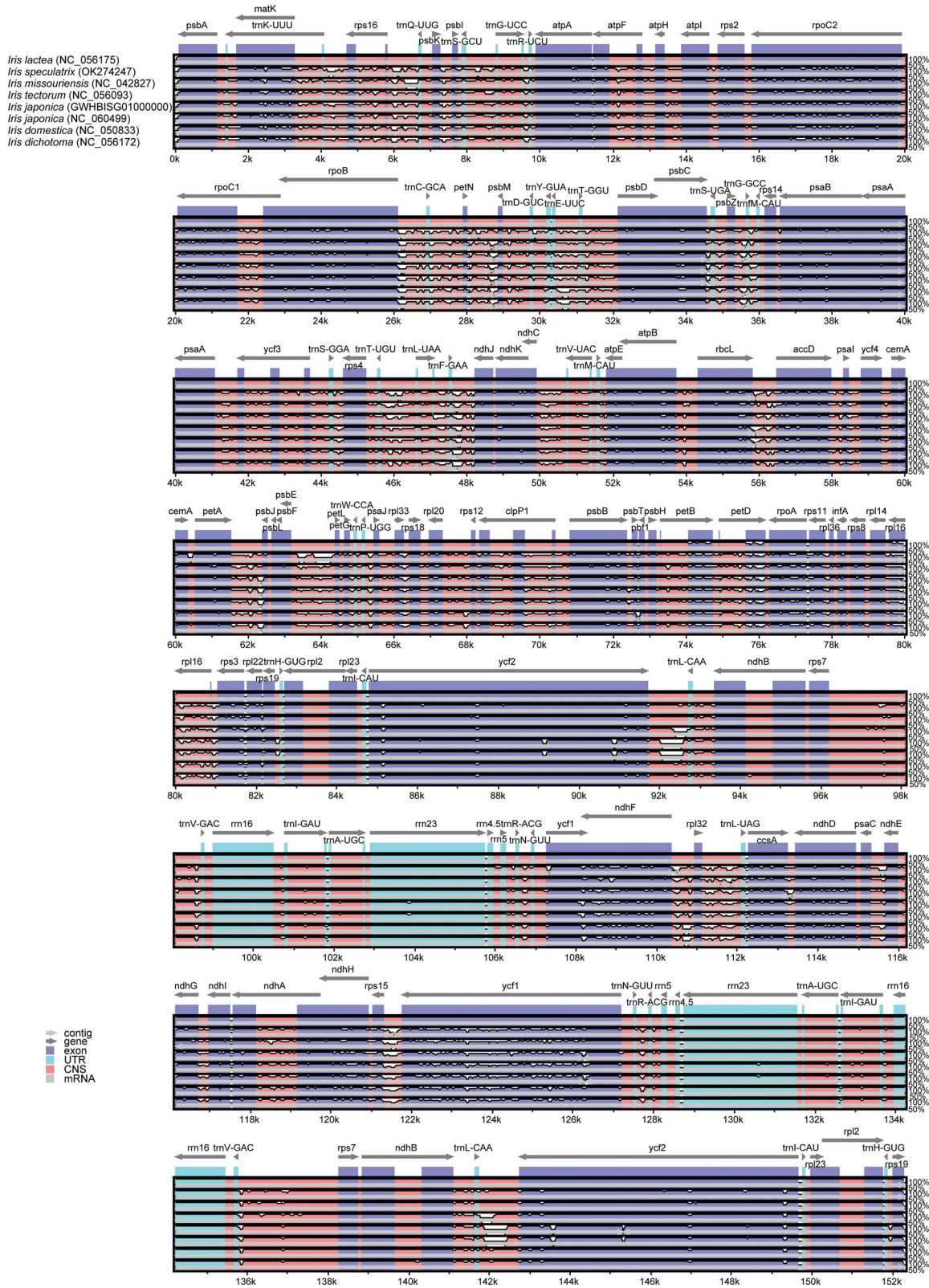


Figure 2. Sequence alignment of seven chloroplast genome of *Iris* species using mVISTA and chloroplast genome of *I. lactea* (NC_056175) as reference. The top arrow shows transcription direction, blue color indicates protein-coding regions, pink color shows non-coding sequences and light green indicates tRNAs and rRNAs. The x-axis represents the coordinates in the cp genome while y-axis represents percentage identity within 50–100%.

using the K2P model. K2P values of *ndhC-trnV-UAC*, *rpl22-rps19*, *rps16-trnQ-UUG*, *trnG-UCC-trnR-UCU*, *rpl36-infA*, *trnC-GCA-petN*, *rpoB-trnC-GCA*, *rpl33-rps18*, *rps19-trnH-GUG*, *trnH-GUG-rpl2*, *rpl2-trnH-GUG*, and *psbK-psbI* were high at 27.05, 19.79, 15.97, 11.69, 10.45, 8.25, 7.95, 6.42, 6.35, 6.12,

6.12, and 6.05, respectively (Supplementary Material, Figure S4). The variations in these IGS regions were large, which suggested that these regions could be used to develop potential molecular markers for species discrimination of *Iris* species.

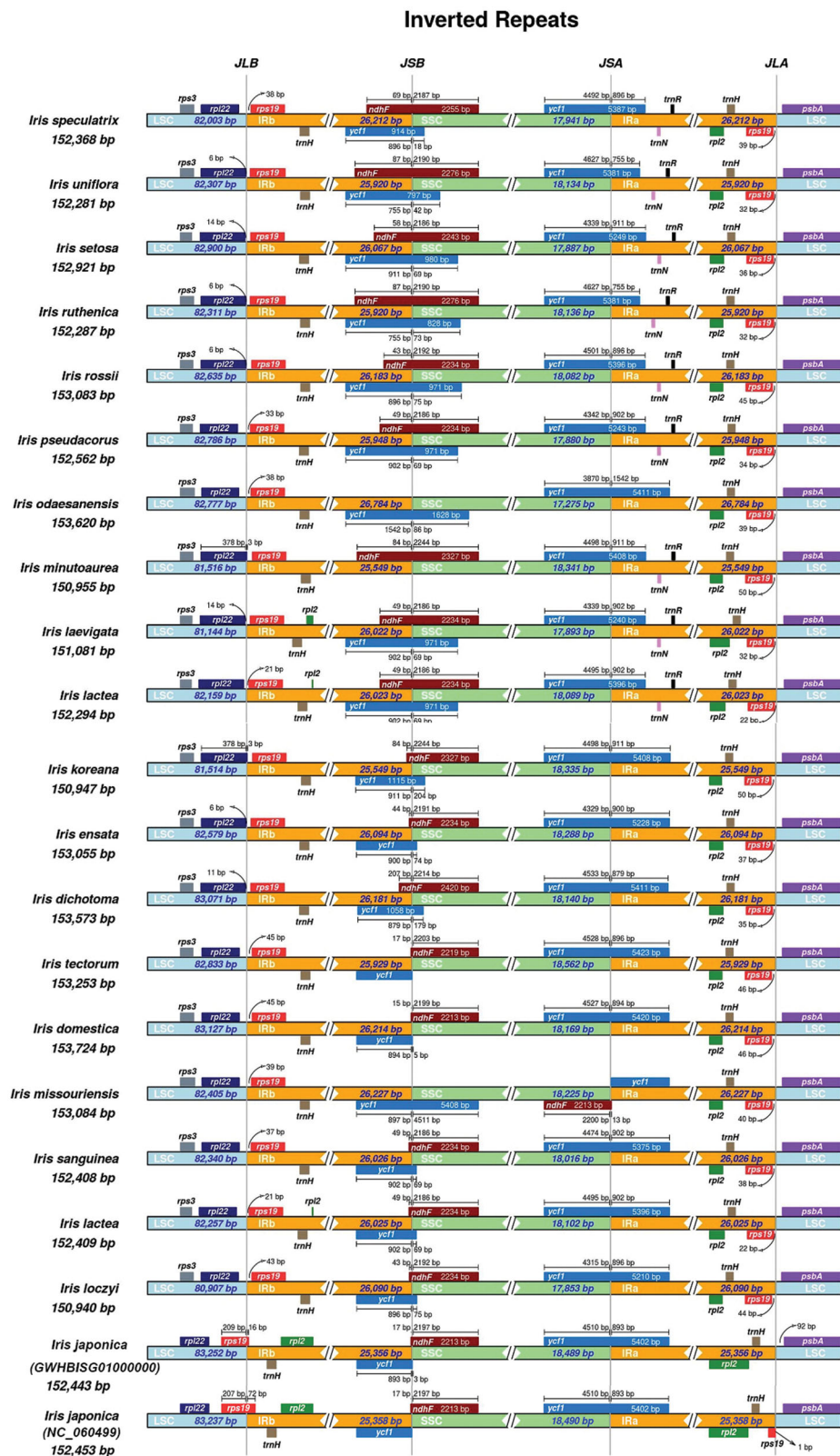


Figure 3. Comparison of LSC, SSC, and IR boundary areas of 21 *Iris* species.

Using the aBSREL model, we found that the genes *rpoC2* and *ycf1* were positively selected in *I. japonica*. Moreover, the genes *ycf2*, *ndhG*, and *rpl20* were positively selected in other *Iris* species. These genes may be related to the adaptation of these *Iris* species to diverse environmental conditions (Supplementary Material, Table S3).

3.5. Phylogenetic analysis

In this study, the chloroplast genome sequence of *I. japonica* (GWHBISG01000000) was aligned with 32 previously published genome sequences. The 32 published genome sequences are the chloroplast genome sequences of *I. sanguinea*

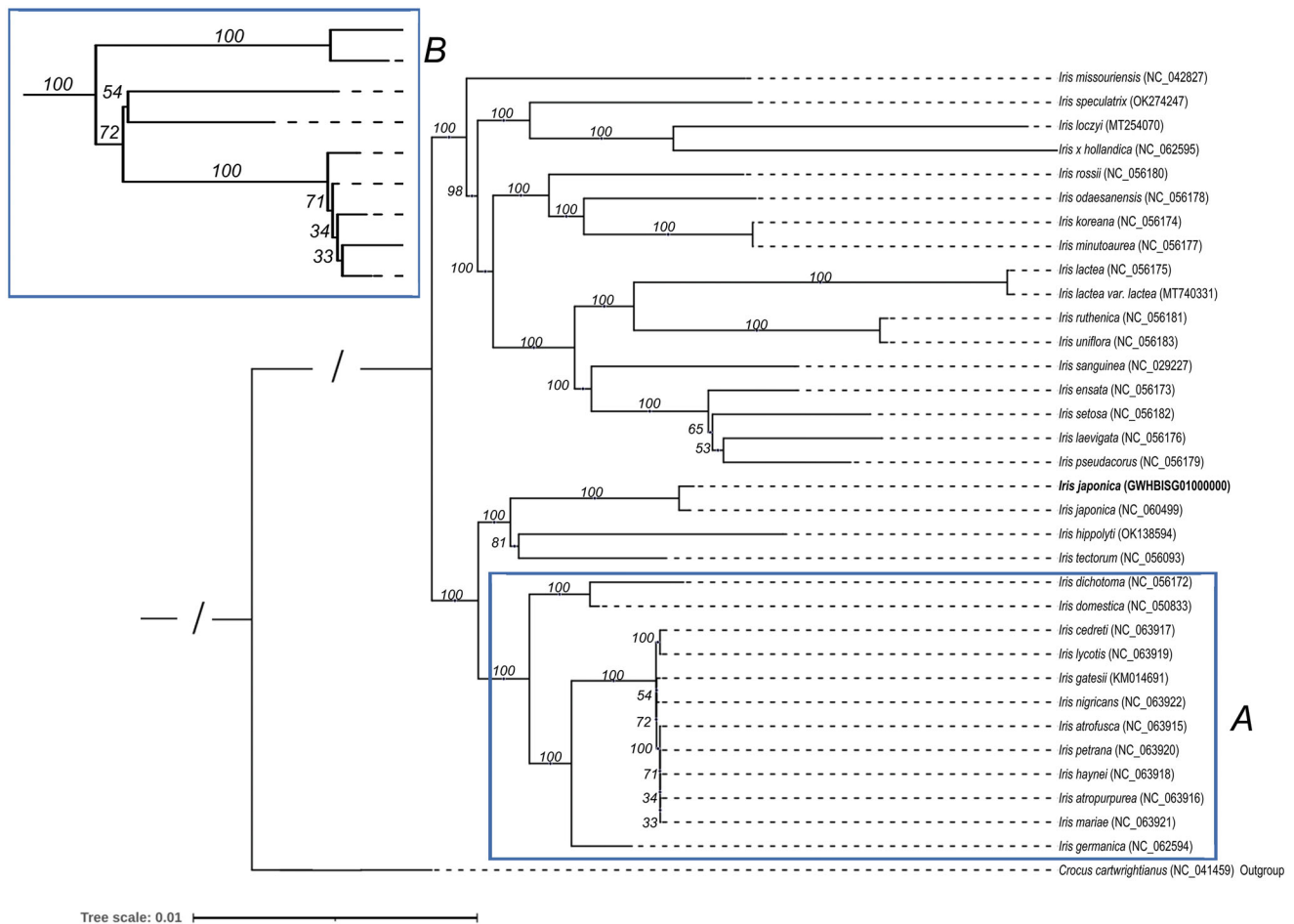


Figure 4. The phylogenetic tree of 21 *Iris* species. *Crocus cartwrightianus* was selected as the outgroup. The tree was constructed using the maximum-likelihood method based on the concatenated sequences from 66 shared proteins. Bootstrap support values were calculated from 1000 replicates. B box is a magnification of A box.

(Lee et al. 2017), *I. missouriensis* (Joyce et al. 2018), *I. speculatrix* (Kang et al. 2020), *I. loczyi* (Choi et al. 2020), *I. x hollandica*, *I. rossii* (Kang et al. 2020), *I. odaesanensis* (Kang et al. 2020), *I. koreana* (Kang et al. 2020), *I. minutoaurea* (Kang et al. 2020), *I. lactea* (Kang et al. 2020), *I. ruthenica* (Kang et al. 2020), *I. uniflora* (Kang et al. 2020), *I. ensata* (Kang et al. 2020), *I. setosa* (Kang et al. 2020), *I. hippolyti*, *I. tectorum*, *I. dichotoma* (Kang et al. 2020), *I. laevigata* (Kang et al. 2020) and *I. pseudacorus* (Kang et al. 2020), *I. lactea var lactea* (Cai et al. 2021), *I. japonica* (NC_060499), *I. domestica*, *I. cedreti* (Volis et al. 2022), *I. lycotis* (Volis et al. 2022), *I. gatesii*, *I. nigricans* (Volis et al. 2022), *I. atrofusca* (Volis et al. 2022), *I. petrana* (Volis et al. 2022), *I. haynei* (Volis et al. 2022), *I. atropurpurea* (Volis et al. 2022), *I. mariae* (Volis et al. 2022) and *I. germanica*. Moreover, *I. japonicum* was sister to moderately-supported clade containing *I. tectorum* and *I. hippolyti*. (Figure 4). During this analysis, we found that another sequence from the same species was released in GenBank (NC_060499.1). Phylogenetic analysis suggested that the two sequences were highly similar. Sequence comparison showed 254 variations between the two sequences (Supplementary Material, Table S4).

4. Discussion

In this study, we sequenced the chloroplast genome of *I. japonica* and conducted comparative and evolutionary

analyses of the chloroplast genomes with 19 other *Iris* chloroplast genomes. Overall, the *I. japonica* chloroplast genome was similar to those of the 19 other *Iris* species.

Notably, comparative analysis found some variations in genome organization. Different gene contents were noted at the border areas of the IR and LSC/SSC regions among the 21 *Iris* species. The *rps19* of only one species was present in the LSC region. However, the duplicated *rps19* of other 20 species were located in the IRa and IRb regions. This type of junction based on the different positions of *rps19* was also found in the chloroplast genomes of *Aphelandra knappiae* and *Blepharis ciliaris*, which caused by the contraction and expansion in the IR regions (Alzahrani et al. 2020).

Highly variable regions in the chloroplast genome can serve as the DNA barcodes for evaluating phylogeny (Dong et al. 2012) and taxa classification (Dong et al. 2014). On the basis of genetic distances, some highly variable regions were observed in the 20 *Iris* chloroplast genomes, including *ndhC-trnV-UAC*, *rpl22-rps19*, *rps16-trnQ-UUG*, *trnG-UCC-trnR-UCU*, *rpl36-infA*, and *trnC-GCA-petN*. In the future, we will test the possibility of distinguishing these *Iris* species using molecular markers developed from these regions.

As described previously, only one comprehensive analysis of the phylogenetic relationships was reported. In that study, 14 Korean *Iris* species were separated into four clades (Kang et al. 2020). Here, we conducted phylogenetic analysis using 20 *Iris*

species, including the 14 reported species. Our study included a systematic analysis with six additional *Iris* chloroplast genomes. In our analysis, the 14 species reported previously were also clustered into four clades. For the newly added five species, the sister relationships between *I. loczyi* and *I. speculatrix*, and between *I. tectorum* and *I. japonica* were observed. In summary, the results obtained from this study will help us unveil the chloroplast genome evolution of *Iris* species.

Ethics statements

I. japonica was not listed as a national and provincial key protected wild plant in China nor a threatened species on the IUCN Red List. Therefore, no specific permissions or licenses were needed for the sampling of *I. japonica* for research purpose according to the regulations of the People's Republic of China on the protection of wild plants. During the sampling process, we followed the local sampling guideline to ensure no substantial harm to the collecting individual.

Author contributions

B.W. and H.M.C. conceived the study. X.Y.Z. and H.Y.Y. performed the data analysis and drafted the manuscript. B.W. critically reviewed the article. All authors approved the final version of the manuscript.

Disclosure statement

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

Funding

This work was supported by the Chinese Academy of Medical Sciences Innovation Funds for Medical Sciences (CIFMS) [2021-I2M-1-022], National Science & Technology Fundamental Resources Investigation Program of China [2018FY100705], National Science Foundation Funds [81872966]. The funders were not involved in the study design, data collection, analysis, publishing decision, or manuscript preparation.

Data availability statement

The sample was stored at the Herbarium of the Institute of Medicinal Plant Development, Beijing, China, with the voucher numbers Implad201808209. The chloroplast genome sequence and the related annotations were available at <https://ngdc.cncb.ac.cn/gwh> in the Genome Warehouse in China National Genomics Data Center [38] with accession number GWHBISG01000000. The raw sequencing data for the Illumina platforms were deposited in CNGB-NGDC (<https://ngdc.cncb.ac.cn/?lang=en>). The accession numbers of Bioproject, Biosample, and GSA are PRJCA009243, SAMC781189, and CRA007000, respectively.

References

Ai HL, Ye K, Zhang X, Lv X, Li ZH, Zhang SD. 2019. The complete plastid genome of *Iris domestica*: a traditional Chinese medicine. *Mitochondrial DNA B Resour.* 4(2):4214–4215.

Alzahrani DA, Yaradua SS, Albokhari EJ, Abba A. 2020. Complete chloroplast genome sequence of *Barleria prionitis*, comparative chloroplast genomics and phylogenetic relationships among Acanthoideae. *BMC Genomics.* 21(1):1–19.

Amiryousefi A, Hyvönen J, Poczei P. 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes. *Bioinformatics.* 34(17):3030–3031.

Bock R. 2015. Engineering plastid genomes: methods, tools, and applications in basic research and biotechnology. *Annu Rev Plant Biol.* 66:211–241.

Brudno M, Malde S, Poliakov A, Do CB, Couronne O, Dubchak I, Batzoglu S. 2003. Glocal alignment: finding rearrangements during alignment. *Bioinformatics.* 19(Suppl 1):i54–62.

Cai X, Zhang B, Wang S, Cheng Y, Wang H. 2021. Characterization and phylogenetic analysis of the chloroplast genome of *Iris lactea* var. chinensis. *Mitochondrial DNA B Resour.* 6(4):1490–1491.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, et al. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme J.* 6(8):1621–1624.

Choi TY, Oh SH, Jang CG, Kim HW, Kim A, Lee SR. 2020. The complete chloroplast genome sequences of the *Iris loczyi* kanitz (Iridaceae). *Mitochondrial DNA B Resour.* 5(3):2876–2877.

Dierckxnsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* 45(4):e18–e18.

Dong W, Liu H, Xu C, Zuo Y, Chen Z, Zhou S. 2014. A chloroplast genomic strategy for designing taxon specific DNA mini-barcodes: a case study on ginsengs. *BMC Genet.* 15(1):138–138.

Dong W, Liu J, Yu J, Wang L, Zhou S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS One.* 7(4):e35071.

Joyce EM, Crayn DM, Lam VK, Gerelle WK, Graham SW, Nauheimer L. 2018. Evolution of Geosiris (Iridaceae): historical biogeography and plastid-genome evolution in a genus of non-photosynthetic tropical rainforest herbs disjunct across the Indian Ocean. *Aust Systematic Bot.* 31(6):504–522.

Kang YJ, Kim S, Lee J, Won H, Nam GH, Kwak M. 2020. Identification of plastid genomic regions inferring species identity from de novo plastid genome assembly of 14 Korean-native *Iris* species (Iridaceae). *PLoS One.* 15(10):e0241178.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–3066.

Lee HJ, Nam GH, Kim K, Lim CE, Yeo JH, Kim S. 2017. The complete chloroplast genome sequences of *Iris sanguinea* donn ex Hornem. *Mitochondrial DNA A DNA Mapp Seq Anal.* 28(1):15–16.

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47(W1):W256–W259.

Liu Z, Yu X, Cui P, Tian X. 2020. The complete chloroplast genome of *Iris tectorum* (Iridaceae). *Mitochondrial DNA Part B.* 5(2):1561–1562.

Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 32(1):268–274.

Olejniczak SA, Łojewska E, Kowalczyk T, Sakowicz T. 2016. Chloroplasts: state of research and practical applications of plastome sequencing. *Planta.* 244(3):517–527.

Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European molecular biology open software suite. *Trends Genet.* 16(6):276–277.

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.

Smith MD, Wertheim JO, Weaver S, Murrell B, Scheffler K, Kosakovsky Pond SL. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol.* 32(5):1342–1353.

Song F, Li T, Burgess KS, Feng Y, Ge X-J. 2020. Complete plastome sequencing resolves taxonomic relationships among species of *Calligonum* L. (Polygonaceae) in China. *BMC Plant Biol.* 20(1):1–15.

Volis S, Zhang Y, Deng T, Yusupov Z. 2022. Dark-colored *Oncocycclus* irises in Israel analyzed by AFLP, whole chloroplast genome sequencing and species distribution modeling. *Isr J Ecol Evol.*

Xu Z, Wang Y, Chen J, Huang N, Cheng D. 2021. Medicinal plant *Iris japonica* as histidine kinase inhibitor candidate can enhance antibiotic susceptibility of *Escherichia coli* resistant to β -lactam antibiotics. *Research Square.* DOI: 10.21203/rs.3.rs-149416/v1.

Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT. 2020. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. *Mol Ecol Resour.* 20(1):348–355.