Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Complete Chloroplast Genome of Critically Endangered *Chimonobambusa hirtinoda* (Poaceae: *Chimonobambusa*) and Phylogenetic Analysis

yanjiang liu
Guizhou University
Xiao Zhu
Guizhou University
Mingli Wu
Guizhou University
Xue Xu
Guizhou University
Zhaoxia Dai (daizhaoxia1999@163.com)
Guizhou University
Guangqian Gou
Guizhou University

Research Article

Keywords: Chloroplast Genome, Critically Endangered Chimonobambusa hirtinoda , Phylogenetic Analysis, simple sequence repeat (SSR)

Posted Date: November 12th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1019626/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Chimonobambusa hirtinoda is a threatened species and only naturally distributed in Doupeng Mountain, Duyun, Guizhou, China. Next-generation sequencing (NGS) is used obtained the complete chloroplast (cp) genome sequence of C. hirtinoda, and then the sequence was assembled and analyze for phylogenetic and evolutionary. We also analyzed comparing the cp genome among Chimonobambusa species with previously published. The complete cp genome of C. hirtinoda has the total length of 139, 561 bp, 38.90% GC content was detected. A total of 130 genes were founded in the cp genome, including 85 protein coding genes, 37 tRNA genes, 8 rRNA. Some genes are missing and the introns occur lost in the cp genome of C. hirtinoda. A total of 48 simple sequence repeat (SSR) were detected and by measuring the codon usage frequency of amino acids, the A/U preference of the third nucleotide in the cp genome of C. hirtinoda was obtained. Furthermore, phylogenetic analysis using complete cp sequences, matk gene exhibited genetic relationship within the Chimonobambusa genus.

Background

Chimonobambusa genus, most of the bamboo shoots in autumn, which not only delicious, but also contain various trace elements such as iron, sun, and zinc, and rich in nutrients[1, 2], it become a favorite food for people. Additionally, the material of *the bamboo* is rich in cellulose pulp fibers so that it is a high-quality raw material for papermaking, which can be produced for bamboo handicrafts, bamboo plywood and craft furniture, which has high economic, edible and cultural value[3, 4, 5].

Chimonobambusa hirtinoda, was listed in the red catalogue by IUCN in 2007 and rated as a national endangered plant. At present, it is only naturally distributed in Doupeng Mountain, Duyun, Guizhou, China[6, 7]. However, in recent years, with the local government's development of tourism in Doupeng mountain and the impact of natural environment, the living environment of *C. hirtinoda* has been damaged and on the verge of extinction[8]. Thus, it is very necessary to protect the natural resources of *C. hirtinoda* from all aspects. Most of the woody bamboos are mainly based on reproduction of the rhizomes, with flowering period is not fixed such as some species only bloom once in a lifetime[9]. There is a great deal of controversy in the division of the Bambusoideae because of their classification and identification generally depend on the morphological characteristics of vegetative organs. Therefore, the research from morphological identification to molecular perspective is of great significance to the classification and evolutionary relationship of bamboo species.

The origin of chloroplast (cp) is generally believed to be obtained from cyanobacteria through endosymbiosis[10]. As the photosynthetic organelle of plant cells, chloroplast not only plays a key role in photosynthesis, but also has important implications in plant physiology and development[11, 12]. The structure of chloroplast genome is very conservative, which is usually composed of a large single-copy (LSC) region, a small single-copy (SSC) region and two inverted repeat (IRs) region in opposite directions[13]. Therefore, chloroplast genomes have been widely used to as DNA barcodes to quickly

identify species, provided useful phylogenetic studies and chloroplast haplotypes are used to analyze the genetic diversity of species[14, 15, 16].

There are many reports on the chloroplast genome of the Arundinariatae in the Bambusoideae [17], but less data available on *Chimonobambusa* genus. Thus, this study is the first time reported the chloroplast genome of *C. hirtinoda*, including gene content, codon usage and compared with allied species. Besides that, phylogenetic relationship was constructed based on previously published cp genomes of Bambusoideae to clarify the taxonomic position of *C. hirtinoda*. These finding will provide valuable genetic resources for further research on the phylogenetic location of *C. hirtinoda* and investigating evolutionary relationships of the order Bambusoideae.

Results And Discussion

Assembly and annotation of the chloroplast genomes of *Chimonobambusa hirtinoda.*

Assembly resulted in a whole cp genome sequence of *C. hirtinoda* with a length of 139, 561 bp (Fig. 1), and consisted of an 83, 166-bp large single-copy region, a 12, 811-bp small single-copy region, and two 21,792-bp IR regions, respectively, comprising the typical quadripartite structure of terrestrial plants. The cp genome of *C. hirtinoda* was annotated with 130 genes, including 85 protein coding genes, 37 tRNA genes, and 8 rRNA genes (Table 1). Most of the 15 genes in the *C. hirtinoda* cp genome contain introns; of these, 13 genes contain one intron (*atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl2*, *rpl16*, *rps16*, *trnA-UGC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, *trnV-UAC*) and only gene *cyf3* includes two introns, while the intron of the gene *clpP* was found to be deleted(Supplementary TableS1). Unusually, it was determined that the *rps12* gene contained two copies, and the three exons were spliced into a trans-splicing gene [18].

Table 1

Summary of the chloroplast genome of *C. hirtinoda*.

Genome features	*C. hirtinoda*
Genome size (bp)	139,561
LSC size (bp)	83,166
SSC size (bp)	12,811
IR size (bp)	21,792
GC content (%)	38.9%
No. of genes	130
No. of PCGs	85
No. of tRNA	37
No. of rRNA	8

Note that the *accD*, *ycf1,* and *ycf2* genes are missing in the cp genome of *C. hirtinoda*, and that the introns in the genes *clpP* and *rpoC1* have been lost. This phenomenon is consistent with previous systematic evolutionary studies on the genome structure of plants in Poaceae [19]. Such a phenomenon of missing genes has also been reported in other plants [20-23].

The total GC content found for the *C. hirtinoda* cp genome was 38.90%, The content for each the four bases A, T, G, and C was 30.63%, 30.46%, 19.57%, and 19.33%, respectively (Table 2). The LSC region (36.98%) and SSC region (33.21%) have much lower values than that in the IR region (44.23%), indicating that distribution of the content in the cp genome is not uniform. This is probably because there are four rRNAs in the IR region, which in turn makes the GC content higher in the IR region. These values were similar to cp genome results previously reported for some Poaceae plants [24, 25].

Table 2

Base composition in the *C. hirtinoda* choloroplast genome.

Region	Length(bp)	A (%)	T (%)	G (%)	C (%)	GC (%)
LSC	83,166	31.24	31.78	18.76	18.22	36.98
SSC	12,811	36.02	30.78	16.17	17.04	33.21
IRA	21,792	27.96	27.81	21.19	23.04	44.23
IRB	21,792	27.96	27.81	21.19	23.04	44.23
Total genome	139,561	30.63	30.46	19.57	19.33	38.90
CDS	60,531	29.63	30.85	21.20	18.3	39.53

Repeat sequences and codon analysis.

SSR consists of approximately 10-bp-long base repeats and is widely used for exploring phylogenetic evolution and for genetic diversity analysis [26-29].

In total, 48 SSRs were detected in *C. hirtinoda*, including 27 mononucleotide versions, which accounted for 56.25% of the total, mainly comprised of A or T. There were 4 dinucleotide repeats comprised of AT/TA and TC/CT repeats, and 3 tri, 13 tetra, and 1penta-repeats shown (Fig. 2A). From the perspective of SSR distribution, the vast majority of SSRs are found in the LSC area, with 38 (79%); in the IR region there are 6 (13%) and in the SSC region there are 4 (8%), respectively (Fig. 2B). Previous research reports suggest that the distribution of SSR numbers in each region and the differences among locations in terms of GC content are related to the expansion or contraction of the IR boundary[30].

The REPuter program revealed that the cp genome of *C. hirtinoda* was identified with 61 repeats consisting of 15 palindromic repeats, 19 forward and no reverse and complement repeats (Fig.3). We notice that repeat analyses of three *Chimonobambusa* genus species showed a total of 61- 65 repeats, and there is only one reverse in C*. hejiangensis.* Most of the repeat lengths between 30 to 100 bp and almost all the repeat sequences were located in either IR or LSC region[31] (Supplementary TableS2).

There were 20,180 codons identified in the coding region of *C. hirtinoda* (Fig. 4, Supplementary TableS3). Among these, the codon AUU of Ile was most widely used, and the codon TER of UAG was least often used, not counting the termination codons (817 and 19). Of those amino acids encoded by codons, Leu had the highest presence with 2,170 and TER was lowest at 85. A relative synonymous codon usage (RSCU) value greater than 1.0 means a codon is used more frequently [32]. The RSCU values for 31 codons exceeded 1 in the *C. hirtinoda* cp genome, and of these the third most frequent codon was A/U with 29 (93.55%), and the codons ending in C and G had values of 1 (3.23%) and 1, respectively (3.23%).

Comparative analysis of genome structure.

The nucleotide variability (Pi) values for of the three cp genomes found in the *Chimonobambusa* genus species ranged from 0 to 0.021 with an average value of 0.000544, as found from analysis with the software package DnaSP 5.10. In Figure 5 that there are clearly five high peaks in the two single-copy regions, and the highest peak is in the *trnT-trnE-trnY* region of the LSC region. The Pi value for LSC and SSC is significantly higher than that of the IR region. In the IR region, no highly different sequences are found, and this is a highly conserved region. The sequences of these highly variable regions have also been reported in other plants during examinations for species identification, phylogenetic analysis, and population genetics research [33-35].

The structural information for the complete cp genomes among three *Chimonobambusa* genus species examined showed that those sequences in most regions were mostly conserved (Fig. 6). It can be seen from Figure 6 that the LSC and SSC regions show a large degree of variation, far higher than for the IR

region, and the noncoding region demonstrates higher variability than is found in the coding region. In noncoding regions, 7-9k, 28-30k, 36k and other gene loci differ greatly. In the protein coding region, genes *rpoC2*, *rps19*, *ndhJ* and other regions show high differences. However, the agreement between the tRNA and rRNA regions is almost 100%. A similar phenomenon has also been reported by others[36].

IR contraction and expansion in the chloroplast genome.

There are four regions and four boundaries in the cp genome of plants. During the process of species evolution, the stability of the two IR region sequences is ensured by the IR region of the chloroplast genome expanding and contracting to some degree, and this adjustment becomes the main reason for changes in chloroplast genome length[37, 38].

It can be seen from Figure 6 that the three *Chimonobambusa* genus chloroplast genomes were found to highly similar in organization, gene content and gene order. The size of IR ranges fom 21, 797 bp (*C. tumidissinoda*) to 21,835 bp (*C. hejiangensis*). The *ndhH* gene spans the IRa/SSC boundary and has a duplication of 181–224 bp in the IRa region. The gene *rps15* is located in the IR region (Fig. 7).

No inversion or translocation is found in the six genome sequences by mauve alignment, and the sequence is the same blocks, indicating that the cp genomes of the six species have not gene rearrangements (Fig. 8)

Phylogenetic analysis.

We performed phylogenetic analysis with both the complete chloroplast genomes and *matK* gene and observed complete chloroplast genome performed better to identify related species, consistent with previous study[39]. The maximum likehood (ML) analysis indicated 7 nodes with fully branch support (100% bootstrap values), however the three *Chimonobambusa* genus with moderately supported relationship as a result of less samples use, which supported *C. hirtinoda* to be closely related with *C. tumidissinoda* with 62 % bootstrap value more than *C. hejiangensis*. The result of phylogenetic tree based on *matK* gene showed that *Chimonobambusa* species clustered in one branch was consistent with the phylogenetic tree constructed by the complete cp genome tree (Fig. 10).

Conclusions

This study mainly explored the chloroplast genome of *C. hirtinoda* and compared it with those of related species within *Chimonobambusa*. These data provide useful genetic information that advances the genetic research on *Chimonobambusa*. Through successfully assembling, annotating, and analyzing the whole chloroplast genome sequence of *C. hirtinoda*, a phenomenon of genes loss was discovered. This loss is probably associated with the rapid evolution of the Poaceae species and the extensive rearrangements of chloroplast structures that took place during that process. The acquisition of these

data, particularly in terms of SSR, will enhance the study of the phylogenetic relationships of *Chimonobambusa* plants, cp genome variation among them, and the function of genes.

Comparative analysis of the *Chimonobambusa* species studied revealed that coding regions are more conservative than noncoding regions in the cp chloroplast genome. Such a change in genetic structure can reflect a relationship with the changes in species, but the mechanism that generates such variations and the subsequent results need further study.

The Poaceae family is generally divided into two large evolutionary branches (BEP and PACCMAD), among which the Bambusoideae, Pooideae, and Oryzoideae belong to the BEP branch. Panicoideae, Arundinoideae, Chloridoideae, Aristidoideae, Arundinoideae, and Micrairoideae belong to the PACCMAD branch. Here, the complete chloroplast genomes based phylogenetic tree (Fig. 10A) shows high values of bootstrap support (only three values less than 80 %), and these species can be polymerized into two clades and an outgroup with strong support. The genus *Bambusa* constitutes an isolated evolutionary branch, becoming a monophyletic group, a conclusion that is consistent with previous reports [40]. Also, the two phylogenetic trees revealed that *C. hirtinoda*, *C. hejiangensis* and *C. tumidissinoda* formed a group that was closely related to the group. On the evolutionary subclade of the second branch, the genus *Ampelocalamus* and *C. longiusculus* have a very close relationship.

Materials & Methods

Plant materials.

Four to six fresh leaf of native habitats are collected in Doupeng Mountain, Guizhou Province, China (N 26°22'32.55", E 107°22'9", altitude 1074.93 m). Those fresh leaf were dried with silica gel and stored at the Natural Museum of Guizhou University (accession number: GACP) for further DNA extraction (Fig. 9). The identification of the species by Professor Guangqian Gou, the director of Bamboo Research Institute in college of life sciences, Guizhou University.

The collection and experiments of plant materials has complied with relevant guidelines and regulations of Doupeng mountain virgin forest nature reserve.

DNA extraction, Chloroplast genome sequencing.

Total genomic DNAs were extracted from sample using the TIANGEN DNA extraction kit (TIANGEN BIOTECH CO., Beijing, China) and the DNA concentration was detected using spectrophotometry, and total DNA quality was detected via 1% agarose gel electrophoresis. All of the DNA obtained from *C. hirtinoda* was sent to BGI (Wuhan, China. https://www.genomics.cn), and the total DNA was sequenced using an Illumina sequencer with an HiSeq2500 system, with library type selected to be the De Novo Sequencing ≤800 bp conventional library.

Genome aassembly and annotation.

De novo assembly of the *C. hirtinoda* cp genome was performed in a GetOrganelle pipeline[41] (https://github.com/Kinggerm/GetOrganelle). Using the complete chloroplast genome of *C. hejiangensis* (#MT884004) as a reference sequence, PGA (https://github.com/quxiaojian/PGA) annotated the chloroplast genes of *C. hirtinoda,* followed by manual correction using the software Generous 10.0.5[42], with uploading to NCBI after the sequence was confirmed to be correct. The GenBank accession number of *C. hirtinoda* was OK046142.

Using the online software package Organellar Genome DRAW (OGDraw)[43] (http://ogdraw.mpimp-golm.mpg.de/index.shtml), a physical map of the chloroplast group was created.

Structural of the *C. hirtinoda* cp genome.

A simple repeat sequence (SSR) is also called a microsatellite. It can be detected using the identification tool MISA[44] and REPuter[45], respectively. The number for the repeat parameter was set to at least 10, 5, 4, 3, 3 and 3 repeat units, from mononucleotide to hexanucleotide. The codon bias for the chloroplast genome was analyzed using the software package CodonW1.4.2 (http://downloads.fyxm.net/CodonW-76666.html).

Sequence divergence.

Determining the nucleotide diversity of the whole cp genome can make the identification of related species more accurate and help to solve similar problems arising in the phylogenetic research [46, 47]. In order to compare the differences, three species of *Chimonobambusa* were selected, using *C. hejiangensis* as the control. The software package MAFFT [48] was used to compare the whole cp genomes of the three species, with results of the comparison manually truncated at both ends. Then the software package DnaSP5.10 was used to calculate the Pi values between species sequences [49], The sliding window was set to 600 and the step size was 200. The online program mVISTA[50] was used to compere three species, using *C. hejiangensis* annotation as the reference. The software MAUVE[51] provided rearrangements of those gene sequences.

Phylogenetic analyses.

For 14 sequences of complete chloroplast genome sequences and *matK* gene of Bambusoideae species and *Hypolytrum nemorum* (Cyperaceae: *Hypolytrum*) *w*as selected as an outgroup for construction the phylogenetic tree to identify the taxonomic position of the *C. hirtinoda*. All sequences were aligned using the tool MAFFT and the maximum likelihood (ML) phylogenetic tree constructed using the software package MEGA-X [52] and the bootstrap replicates parameter was set to 1,000.

Specimen collection statement.

The collection of fresh leaves obtained the permission of the nature reserve.

Declarations

Acknowledgements

This study was supported by The Breeding of Improved Bamboo Varieties for Shoot (Te-Lin-Yan 2020-17).

Author contributions

Yanjiang Liu conceived and designed the experiments, performed the experiments, analyzed the data, provided reagents/materials/analysis tools, prepared figures and/or tables, wrote and reviewed drafts of the manuscript, and approved the final draft.

Xiao Zhu analyzed the data, contributed reagents/materials/analysis tools, authored and reviewed drafts of the manuscript, and approved the final draft.

Mingli Wu and Xue Xu prepared figures and/or tables, authored and reviewed drafts of the manuscript, and approved the final draft.

Zhaoxia Dai and Guangqian Gou conceived and designed the experiments, authored or reviewed drafts of the manuscript, and approved the final draft.

Competing interests

The authors declare that there are no conflicts of interest.

Additional information

Correspondence and requests for materials should be addressed to Zhaoxia Dai and Yanjiang Liu.

References

1. Liu YJ, Wang LP, Fu B, *et al*. Nutrition components in shoots of different provenance of *Chimonobambusa hejiangensis* and *C. pachystachys*. *J Zhejiang for Sci Technol***32**(04):37-42 (2012).

2. Gou GQ, Ding YL, Yang L, *et al*. Nutrient analysis on bamboo shoots of three species in *Chimonobambusa.China Veg***16**:79-81 (2010).
3. Qi Gao, Xiang HZ, Ni LM, *et al*. Nitrogen self-doped activated carbons with narrow pore size distribution from bamboo shoot shells. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, **629**, 0927-7757 (2021).
4. Zhou QL. Effects of Three Kinds of Economic Forests on Soil Physical and Chemical Properties in Alpine Region. *Tibet Journal of Agricultural Sciences*, **42**:51-53 (2021).
5. Chen GJ, Li CF, Wang SS, *et al*. Characterization of physicochemical properties and antioxidant activity of polysaccharides from shoot residues of bamboo (*Chimonobambusa quadrangularis*): Effect of drying procedures. *Food Chemistry.* **292**: 281-293. https://doi.org/10.1016/j.foodchem.2019.04.060 (2019).
6. Editorial Committee of Chinese Flora, Chinese Academy of Sciences. Flora of China. *Beijing: Science Press*. 9[M]:333 (1996).
7. Wang S, Xie Y. Red List of Chinese Species. *Beijing: Higher Education Press*. 1[M] (2004).
8. Su CH, Zhu SX, Zhang H, Luo G. Research on Rhythm of Shooting and Growth for Critically Endangered *Chimonobambusa hirtinoda. Jour of Fujian Forestry Sci and Tech*. **43**:3 (2016).
9. de Carvalho AL, Nelson BW, Bianchini MC, Plagnol D, Kuplich TM, Daly DC. Bamboo-dominated forests of the southwest Amazon: detection, spatial extent, life cycle length and flowering waves. *PLoS One*.**8(1)**: e54852. https://doi.org/10.1371/journal.pone.0054852 (2013).
10. Zhu RM, Chai S, Zhang ZZ, Ma CL, Zhang Y, Li S. Arabidopsis Chloroplast protein for Growth and Fertility1 (CGF1) and CGF2 are essential for chloroplast development and female gametogenesis. *BMC Plant Biol*. **20(1)**:172. https://doi.org/10.1186/s12870-020-02393-5 (2020).
11. Bhattacharya O, Ortiz I, Walling LL. Methodology: an optimized, high-yield tomato leaf chloroplast isolation and stroma extraction protocol for proteomics analyses and identification of chloroplast co-localizing proteins. *Plant Methods*. **16**:131. https://doi.org/10.1186/s13007-020-00667-5 (2020).
12. Boudreau E, Nickelsen J, Lemaire SD, *et al*. The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in psbD mRNA stability. *EMBO J*. **19(13)**:3366-3376. https://doi.org/10.1093/emboj/19.13.3366 (2000).
13. Paudel YP, Hu ZX, Khatiwada JR, *et al*. Chloroplast genome analysis of Chrysotila dentata. *Gene.* **804**:145871. https://doi.org/10.1016/j.gene.2021.145871 (2021).
14. Wu L, Nie L, Wang Q. *et al*. Comparative and phylogenetic analyses of the chloroplast genomes of species of Paeoniaceae. *Sci Rep***11**, 14643. https://doi.org/10.1038/s41598-021-94137-0 (2021).
15. Kim S, Suh Y. Phylogeny of Magnoliaceae based on ten chloroplast DNA regions. *Journal of Plant Biology.* **56**:290–305. https://doi.org/10.1007/s12374-013-0111-9 (2013)
16. Zhang TT, Hou YK, Yang T, *et al*. Evolutionary analysis of chloroplast tRNA of Gymnosperm revealed the novel structural variation and evolutionary aspect. *PeerJ*. **8**: e10312. https://doi.org/10.7717/peerj.10312 (2020).

17. Choi KS, Chung MG, Park S. The complete chloroplast genome sequences of three veroniceae species (plantaginaceae): comparative analysis and highly divergent regions. *Front Plant Sci*. **7**:355. https://doi.org/10.3389/fpls.2016.00355 (2016).
18. Androsiuk P, Jastrzębski JP, Paukszto L, *et al*. The complete chloroplast genome of *Colobanthus apetalus* (Labill.) Druce: genome organization and comparison with related species. *PeerJ***6**: e4723. https://doi.org/10.7717/peerj.4723 (2018).
19. Su Q, Liu LX, Zhao MY, *et al*. The complete chloroplast genomes of seventeen Aegilops tauschii: genome comparative analysis and phylogenetic inference. *PeerJ* **8**: e8678. https://doi.org/10.7717/peerj.8678 (2020).
20. Lubna, Asaf S, Jan R, *et al*. Complete Chloroplast Genome Characterization of Oxalis Corniculata and Its Comparison with Related Species from Family Oxalidaceae. *Plants*. **9(8)**:928. https://doi.org/10.3390/plants9080928 (2020).
21. Herdenberger F, Pillay D T, Steinmetz A. Sequence of the trnH gene and the inverted repeat structure deletion site of the broad bean chloroplast genome. *Nucleic Acids Res***18(5)**:1297 (1990).
22. Li Y, Zhou JG, Chen XL, *et al*. Gene losses and partial deletion of small single-copy regions of the chloroplast genomes of two hemiparasitic Taxillus species. *Sci Rep***7(1)**:12834. https://doi.org/10.1038/s41598-017-13401-4 (2017).
23. Nie L, Cui Y, Wu L, *et al*. Gene losses and variations in chloroplast genome of parasitic plant *macrosolen* and phylogenetic relationships within santalales. *Int J Mol Sci***20(22)**:5812. https://doi.org/10.3390/ijms20225812 (2019).
24. Ogihara Y, Isono K, Kojima T, *et al*. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. *Mol Genet Genomics***266(5)**:740-746. https://doi.org/10.1007/s00438-001-0606-9 (2002).
25. Maier RM, Neckermann K, Igloi GL, Kössel H. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *J Mol Biol***251(5)**:614-628. https://doi.org/10.1006/jmbi.1995.0460 (1995).
26. Huang CJ, Chu FH, Liu SC, *et al*. Isolation and characterization of SSR and EST-SSR loci in *Chamaecyparis formosensis* (Cupressaceae). *Appl Plant Sci***6(8)**: e01175. https://doi.org/10.1002/aps3.1175 (2018).
27. Wen M, Wang H, Xia Z, *et al.* Developmenrt of EST-SSR and genomic-SSR markers to assess genetic diversity in Jatropha Curcas L. *BMC Res Notes***3**:42. https://doi.org/10.1186/1756-0500-3-42 (2010).
28. Zhang M, Mao W, Zhang G, Wu F. 2014. Development and characterization of polymorphic EST-SSR and genomic SSR markers for Tibetan annual wild barley. *PLoS One***9(4)**: e94881. https://doi.org/10.1371/journal.pone.0094881 (2014).
29. Parthiban S, Govindaraj P, Senthilkumar S. Comparison of relative efficiency of genomic SSR and EST-SSR markers in estimating genetic diversity in sugarcane. *3 Biotech***8(3)**:144. https://doi.org/10.1007/s13205-018-1172-8 (2018).

30. Feng KW. 2014. Comparative and phylogenetic analyses of chloroplast genomes in poaceae family. *Northwest A & F University.*
31. Saina, JK, Li, ZZ, Gichira, AW, *et al*. The Complete Chloroplast Genome Sequence of Tree of Heaven (*Ailanthus altissima* (Mill.) (Sapindales: Simaroubaceae), an Important Pantropical Tree. *Int. J. Mol. Sci.* ***19*****(4)**: 929. https://doi.org/10.3390/ijms19040929 (2019).
32. Sarkar I, Dey P, Sharma SK, *et al*. Turdoides affinis mitogenome reveals the translational efficiency and importance of NADH dehydrogenase complex-I in the Leiothrichidae family. *Sci Rep*. **10(1)**:16202. https://doi.org/10.1038/s41598-020-72674-4 (2020)
33. Bi Y, Zhang MF, Xue J, *et al*. Chloroplast genomic resources for phylogeny and DNA barcoding: a case study on Fritillaria. *Sci Rep***8(1)**:1184. https://doi.org/10.1038/s41598-018-19591-9 **(**2018).
34. Li Z, Long H, Zhang L, *et al*. The complete chloroplast genome sequence of tung tree (Vernicia fordii): organization and phylogenetic relationships with other angiosperms. *Sci Rep***7(1)**:1869. https://doi.org/10.1038/s41598-017-02076-6 (2017).
35. Souza UJB, Nunes R, Targueta CP, *et al*. The complete chloroplast genome of Stryphnodendron adstringens (Leguminosae - Caesalpinioideae): comparative analysis with related Mimosoid species. *Sci Rep***9(1)**:14206. https://doi.org/10.1038/s41598-019-50620-3 (2019).
36. Hand ML, Spangenberg GC, Forster JW, Cogan NO. Plastome sequence determination and comparative analysis for members of the Lolium-Festuca grass species complex. *G3 (Bethesda)***3(4)**:607-616. https://doi.org/10.1534/g3.112.005264 (2013).
37. Pang X, Liu H, Wu S, *et al*. Species Identification of Oaks (*Quercus* L., Fagaceae) from Gene to Genome. *Int J Mol Sci*. **20(23)**:5940. https://www.mdpi.com/1422-0067/20/23/5940 (2019).
38. Zhang X, Zhou T, Yang J, *et al*. Comparative Analyses of Chloroplast Genomes of Cucurbitaceae Species: Lights into Selective Pressures and Phylogenetic Relationships. *Molecules*. **23(9)**:2165. https://www.mdpi.com/1420-3049/23/9/2165 (2018).
39. Nguyen HQ, Nguyen TNL, Doan TN, *et al*, Complete chloroplast genome of novel *Adrinandra megaphylla* Hu species: molecular structure, comparative and phylogenetic analysis. *Scientific Reports*. **11**: 11731. https://doi.org/10.1038/s41598-021-91071-z (2020).
40. Zhang X, Chen SY, Chen P, Liang H. The complete chloroplast genome of Chusquea culeou (Poaceae: Bambusoideae: Bambuseae). *Mitochondrial DNA B* **4(1)**:91-92. https://doi.org/10.1080/23802359.2018.1536479 (2019).
41. Jin JJ, Yu WB, Yang JB, *et al*. GetOrganelle: a simple and fast pipeline for de novo assembly of a complete circular chloroplast genome using genome skimming data. *BioRxiv***256479**. https://doi.org/10.1101/256479 (2018).
42. Kearse M, Moir R, Wilson A, *et al*. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. *Bioinformatics***28(12)**:1647-1649 (2012).
43. Lohse M, Drechsel O, Bock R. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr Genet***52(5-6)**:267-

274. https://doi.org/10.1007/s00294-007-0161-y (2007).

44. Thiel T, Michalek W, Varshney RK, Graner A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor App Genet***106(3)**:411-422 (2003).
45. Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res***29(22)**: 4633-4642.
46. Collevatti RG, Terribile LC, Rabelo SG, Lima-Ribeiro MS. Relaxed random walk model coupled with ecological niche modeling unravel the dispersal dynamics of a Neotropical savanna tree species in the deeper Quaternary. *Front Plant Sci*. **6**:653. https://doi.org/10.3389/fpls.2015.00653 (2015).
47. Du YP, Bi Y, Yang FP, *et al*. Complete chloroplast genome sequences of Lilium: insights into evolutionary dynamics and phylogenetic analyses. *Sci Rep***7(1)**:5751. https://doi.org/10.1038/s41598-017-06210-2 (2017).
48. Katoh k, Rozewicki J, Yamada KD, MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization, *Briefings in Bioinformatics*, **20**: 4. 1160–1166. https://doi.org/10.1093/bib/bbx108 (2019).
49. Librado P, Rozas J. Dnasp v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. **25**:1451–1452. https://doi: 10.1093/bioinformatics/btp187 (2009).
50. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. mVISTA: computational tools for comparative genomics. *Nucleic Acids Res***32**: W273-W279 (2004).
51. Kurtz S, Phillippy A, Delcher AL. *et al*. Versatile and open software for comparing large genomes[J]. *Genome biology* 5: R12. https://doi.org/10.1186/gb-2004-5-2-r12 (2004).
52. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol***35(6)**:1547-1549. https://doi.org/10.1093/molbev/msy096 (2018).

Figures

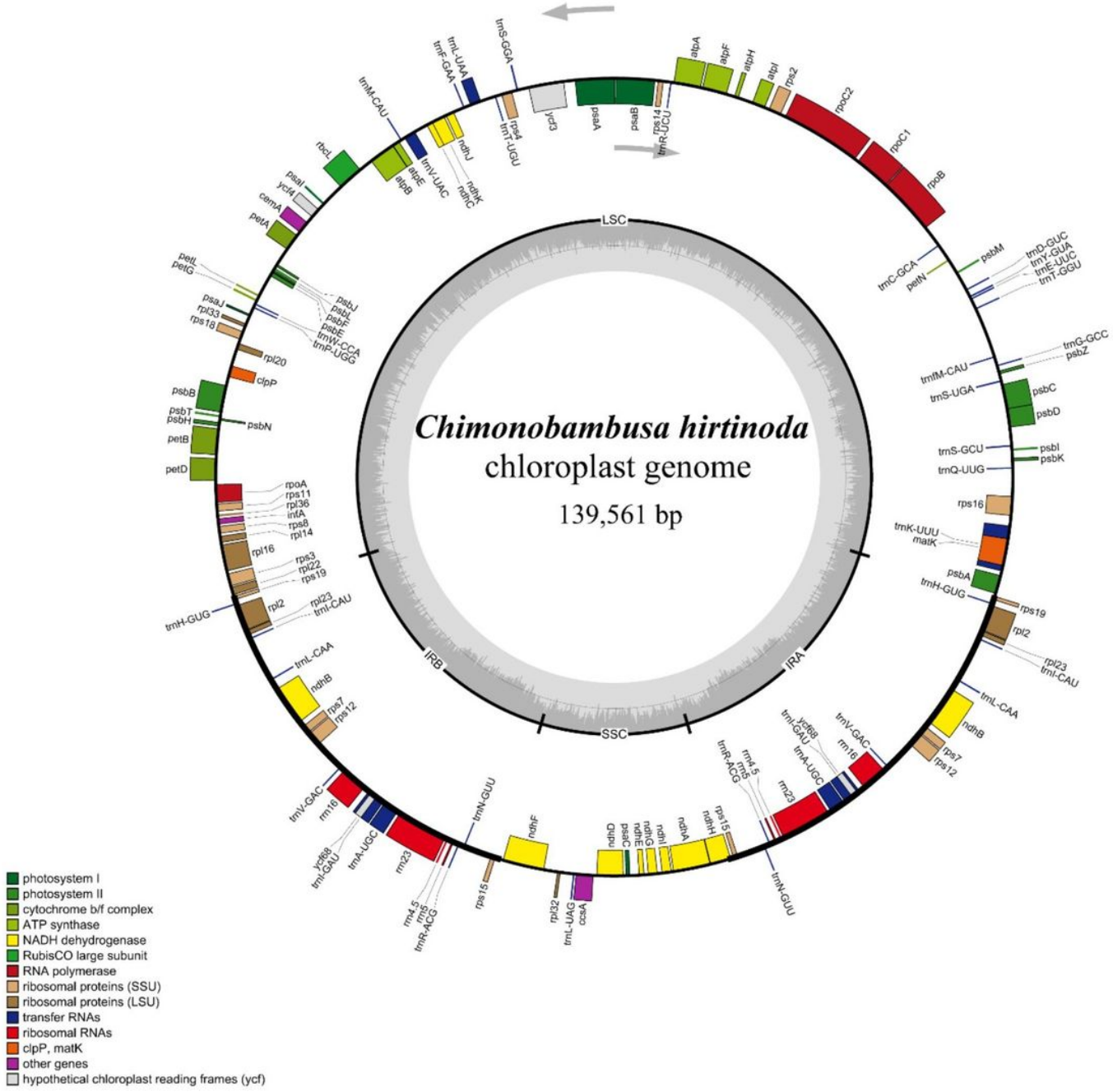

Figure 1

Chloroplast genome map of C. hirtinoda. Different colors represent different functional genes groups. Genes outside the circle indicate counterclockwise transcription, and genes inside the clockwise transcription. The thick black line on the outer circle represents the two IR regions. The GC content is the dark gray area within the ring.

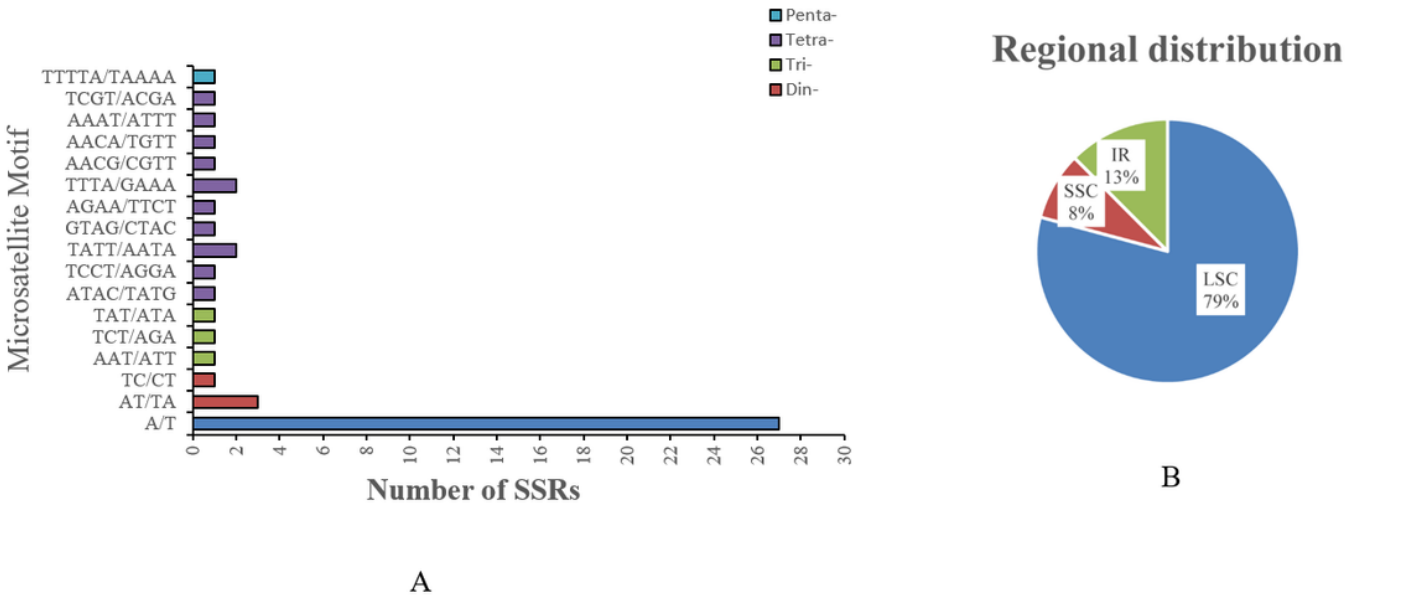

Figure 2

Analysis of simple sequence repeats in C. hirtinoda cp genome. (A) The percentage distribution of 45 SSRs in LSC, SSC, and IR regions. (B)

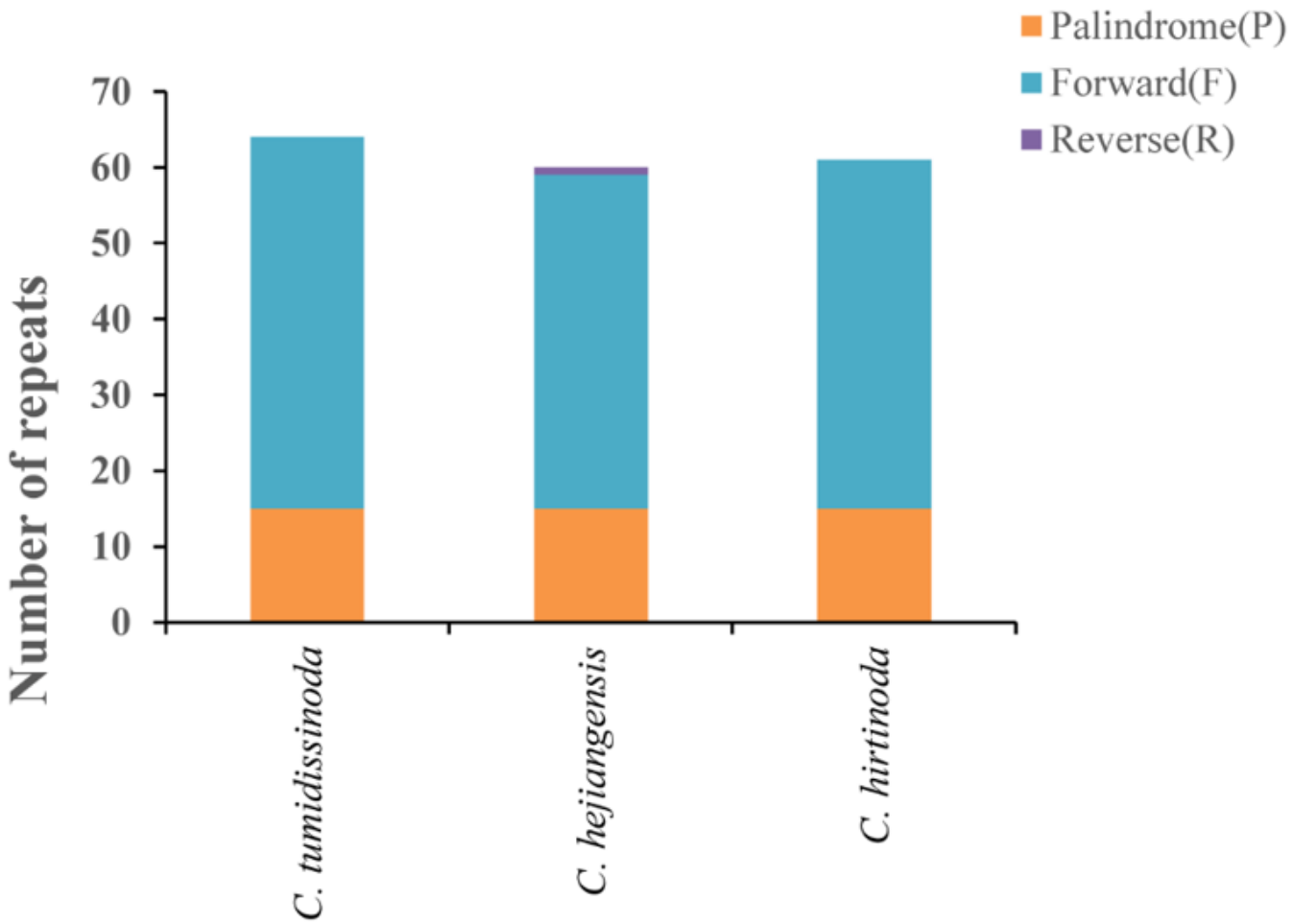

Figure 3

Information of chloroplast genome repeats of Chimonobambusa

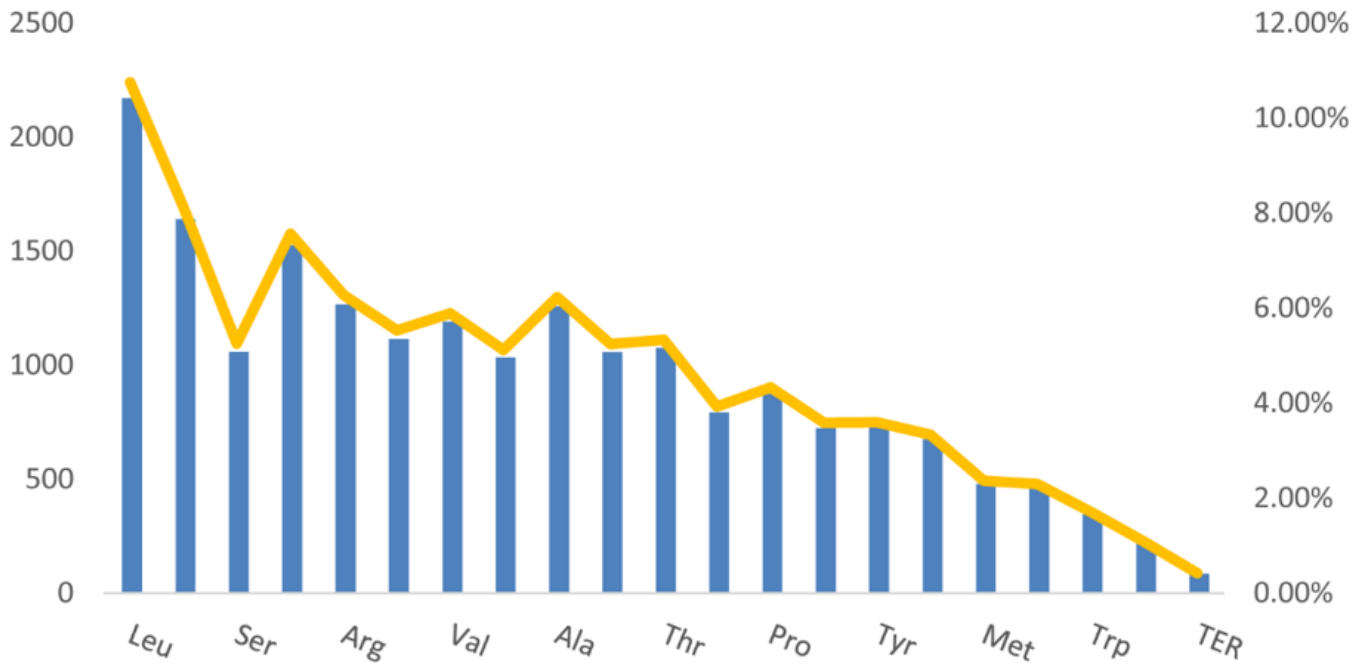

Figure 4

Amino acid frequencies in C. hirtinoda cp genome protein coding sequences. The column diagrams indicate the number of amino acid codes, and the broken line indicates the proportion of amino acid codes.

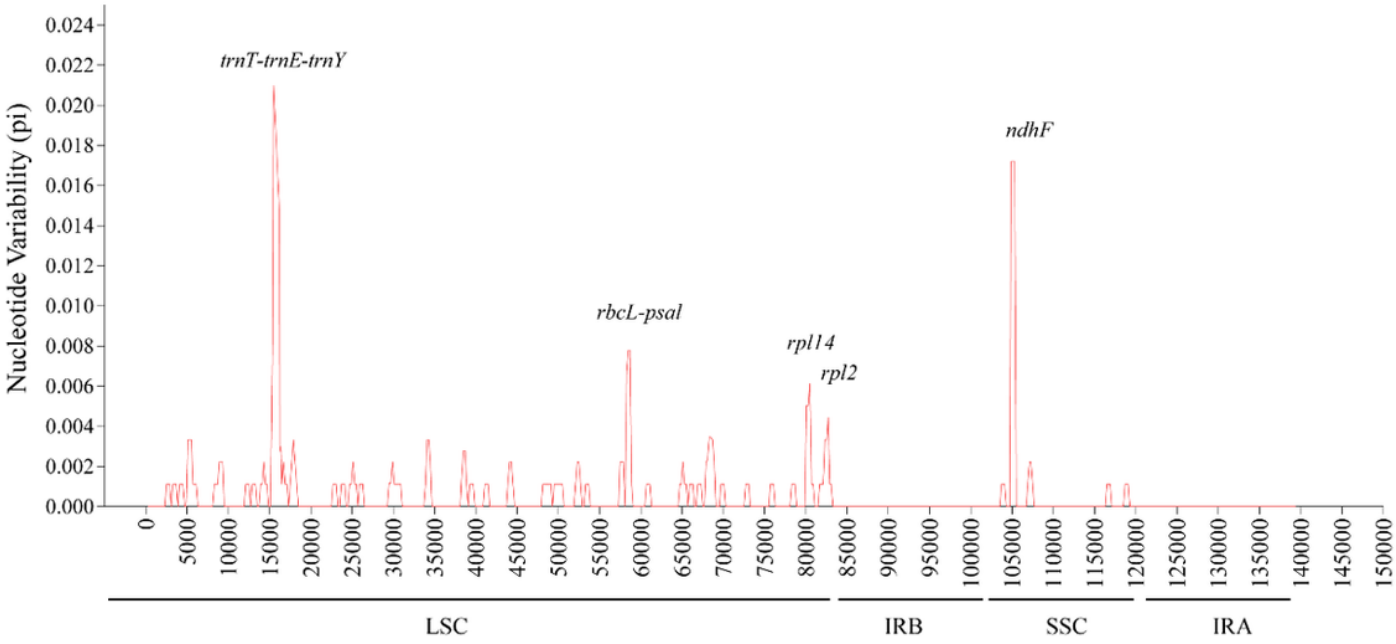

Figure 5

Sliding window analysis of Chimonobambusa genus complete chloroplast genome sequences. X-axis: position of the midpoint of a window, Y-axis: nucleotide diversity of each window.

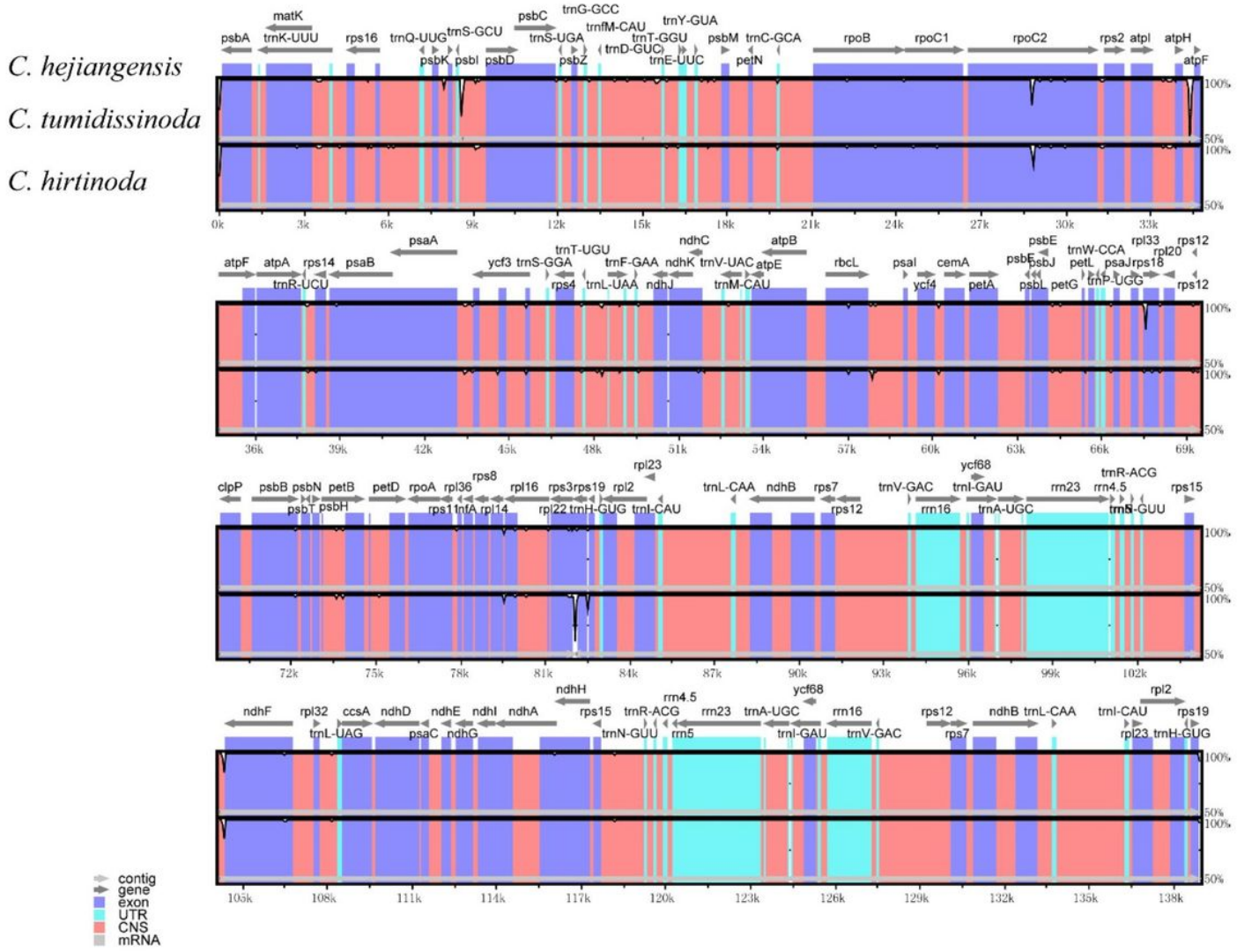

Figure 6

Visualization of genome alignment of three species chloroplast genome sequences using Chimonobambusa hejiangensis as reference. The vertical scale shows the percent of identity, ranging from 50% to 100%. The horizontal axis shows the coordinates within the cp genome. Those are some colors represents protein coding, intron, mRNA and conserved non-coding sequence, respectively.

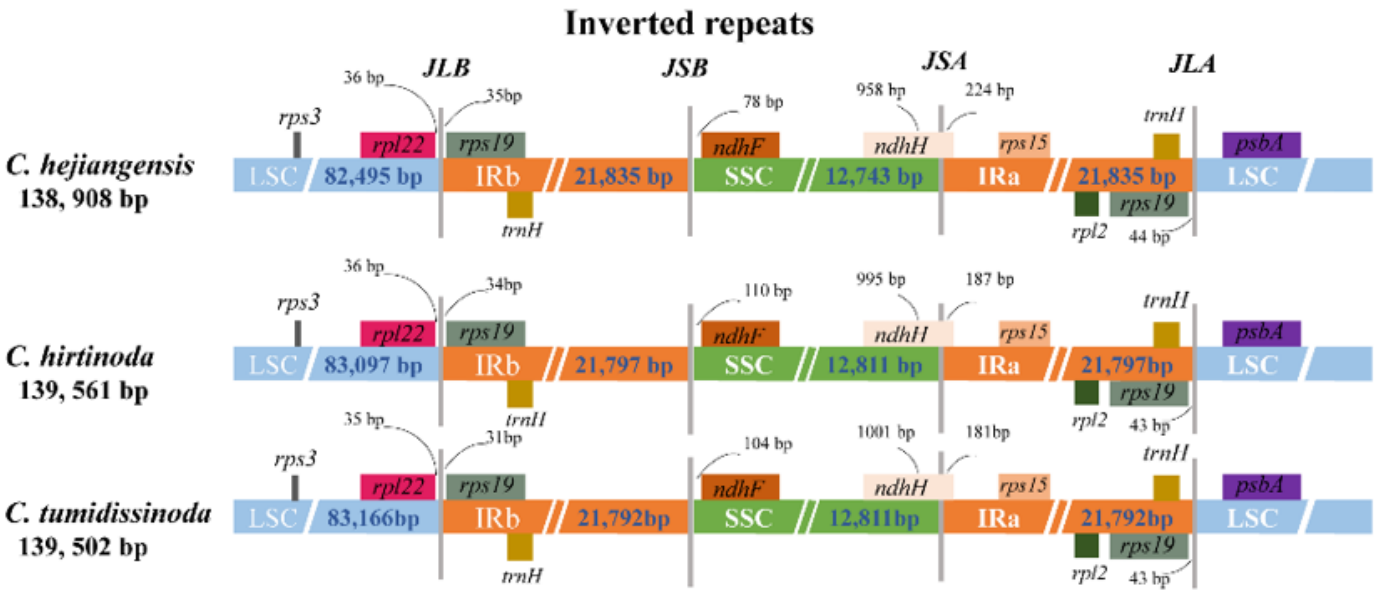

Figure 7

Comparison of LSC, SSC and IR boundaries of chloroplast genomes among the three Chimonobambusa species. The LSC, SSC and IRs regions are represented with different colors. JLB, JSB, JSA and JLA represent the connecting sites between the corresponding regions of the genome, respectively. Genes are showed by boxes.

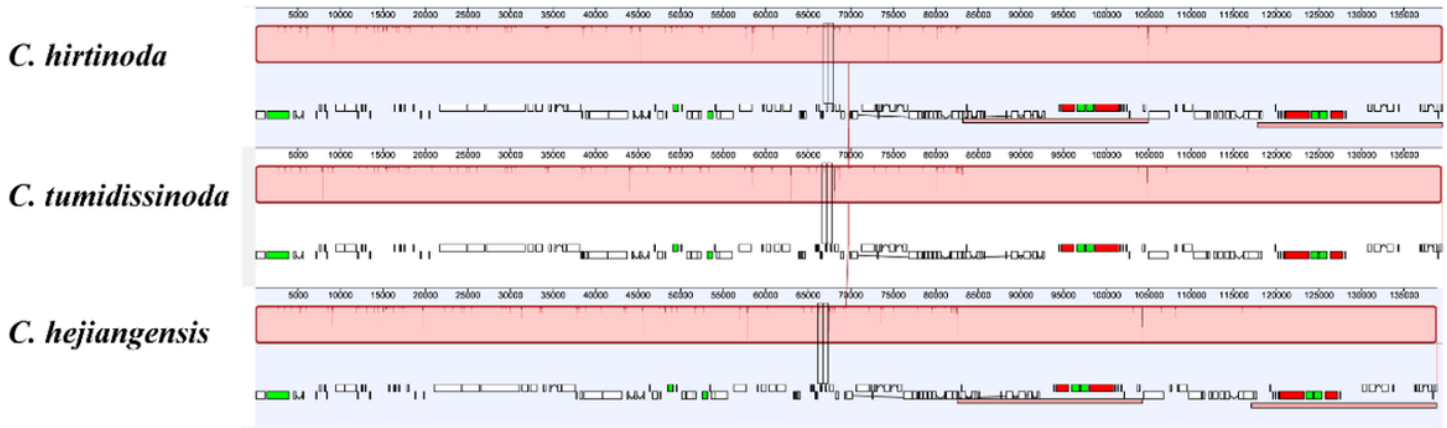

Figure 8

The chloroplast genomes of three Chimonobambusa species rearranged by the software MAUVE. Locally collinear blocks (LCBs) are represented by the same color blocks connected by lines. The vertical line indicates the degree of conservatism among position. The small red bar represents rRNA.

Figure 9

Morphological characteristic of C. hirtinoda. (A) Habit; (B) nodal ridges; (C) rings of root thorns.

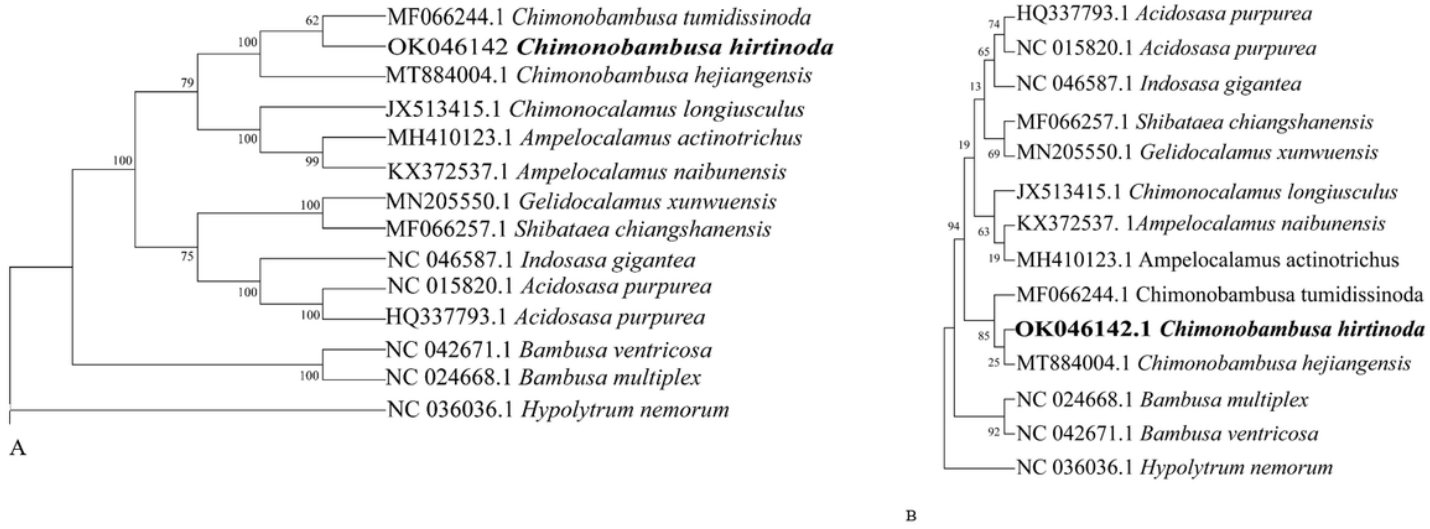

Figure 10

Maximum likelihood phylogenetic tree based on the complete chloroplast genomes (A) and matK gene (B).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.docx
- SupplementaryTableS2.docx
- SupplementaryTableS3.docx