

Article

# The Complete Chloroplast Genome of *Idesia polycarpa* and Comparative Analysis of Related Species

Xueqian Fu <sup>1,†</sup>, Jie Luo <sup>1,†</sup>, Yuan Guo <sup>1</sup>, Dalan Feng <sup>2</sup>, Yifei Deng <sup>1</sup>, Mi Kuang <sup>3</sup>, Houlin Zhou <sup>4</sup>, Xia Liu <sup>1,\*</sup> and Chong Sun <sup>1,5</sup>

- <sup>1</sup> Chongqing Key Laboratory of Economic Plant Biotechnology, Collaborative Innovation Center of Special Plant Industry in Chongqing, College of Smart Agriculture, Chongqing University of Arts and Sciences, Chongqing 402160, China; fxquick120@163.com (X.F.); luojie@cqwu.edu.cn (J.L.); 15923370686@163.com (Y.G.); 13594098664@163.com (Y.D.); zbgqsc1987@163.com (C.S.)
  - <sup>2</sup> Chongqing Key Laboratory of Forest Ecological Restoration and Utilization in the Three Gorges Reservoir Area, Chongqing Academy of Forestry Sciences, Chongqing 400036, China; fengdalang@163.com
  - <sup>3</sup> Chongqing Agricultural Technology Promotion General Station, Chongqing 401120, China; kim437@126.com
  - <sup>4</sup> Chongqing Wulipo National Nature Reserve Management Office, Chongqing 401147, China; z18996669717@163.com
  - <sup>5</sup> Hubei Key Laboratory of Spices & Horticultural Plant Germplasm Innovation and Utilization, Yangtze University, Jingzhou 434023, China
- \* Correspondence: 20170097@cqwu.edu.cn
- † These authors contributed equally to this work.

**Abstract:** Background/Objectives: The oil grape (*Idesia polycarpa*), often called the “golden tree”, is an essential woody plant valued for its edible oil. Although its economic significance is recognized, the specifics of its chloroplast genome and evolutionary connections remain unclear. This study sequenced the chloroplast genome of *I. polycarpa* and performed a comparative analysis of its genome structure, genetic diversity, and phylogenetics using chloroplast data from related species. Methods: In this study, we sequenced and annotated the whole chloroplast genome of *I. polycarpa* via GISEQ-500 sequencing and de novo assembly. Results: The chloroplast genome of *I. polycarpa* exhibits a typical tetrad structure, with a length of 155,899 bp and a GC content of 36.78%. It comprises 130 unique genes, including 85 coding genes, 37 tRNAs, and eight rRNAs, showing notable conservation in gene composition and arrangement compared to closely related species. However, the inverted repeat region boundaries are narrower. Phylogenetic analysis showed strong relationships among *I. polycarpa*, *Bennettiodendron brevipes*, *Poliothyrsis sinensis*, *Itoa orientalis*, and *Carrierea calycina* within the *Salicaceae* family. Additionally, positive selection analysis revealed that *rpl16*, *ycf1*, *rps18*, and *rpl22* are under significant selective pressure in related species, likely linked to adaptations for photosynthesis and environmental responses. Conclusions: This research provides vital molecular foundations for the conservation, classification, and enhancement of *I. polycarpa* germplasm resources, advancing the study of adaptive evolutionary mechanisms and broadening the genomic database for *I. polycarpa*.

**Keywords:** *Idesia polycarpa*; genome structure; genetic diversity; phylogenetic analysis



Academic Editor: Zhiqiang Wu

Received: 3 May 2025

Revised: 17 May 2025

Accepted: 17 May 2025

Published: 21 May 2025

**Citation:** Fu, X.; Luo, J.; Guo, Y.; Feng, D.; Deng, Y.; Kuang, M.; Zhou, H.; Liu, X.; Sun, C. The Complete Chloroplast Genome of *Idesia polycarpa* and Comparative Analysis of Related Species. *Genes* **2025**, *16*, 611. <https://doi.org/10.3390/genes16050611>

**Copyright:** © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Oil grape, *Idesia polycarpa* Maxim., the only species in the genus *Idesia*, is a high-quality, high-yield woody edible oil tree and an industrial oil raw material tree in China. The species is distributed in northeastern Asia (China, Japan, Korea, Russia). The natural range in China is extensive: the provinces south of the Qinling Mountains and Huai River, including Anhui, Zhejiang, and Jiangxi Provinces. This tree species is highly regarded for

its aesthetically pleasing crown structure and robust resistance to various environmental stressors. Furthermore, it is distinguished by its high fruit oil content, earning it the colloquial designation of the ‘tree oil depot’. The content of linoleic acid in the seed of *I. polycarpa* has been reported to be as high as 75%, making it highly valuable for health and nutritional applications [1]. Additionally, this oil serves as a viable raw material for the extraction of linoleic acid preparations and is used in the manufacturing of paints, coatings, and soaps. Given its multifunctional nature, *I. polycarpa* has significant market demand. It is an unusual economic tree species in that it offers value as an ornamental species as well as a source of timber and energy and is thus the focus of much research.

In recent years, with the growing global demand for sustainable resources, the exploration of germplasm resources and the breeding of superior varieties of *I. polycarpa* have become research hotspots. Liu et al. [2] used seeds and leaves of *I. polycarpa* as explants to study the effects of explant types, phytohormones, and other factors on various stages of tissue culture. Their results showed that seeds were the optimal explants for inducing callus formation, and they established a high-frequency tissue culture regeneration system for *I. polycarpa*. Yang et al. [3] investigated the role of explant types and plant growth regulators on organogenesis and plant regeneration under in vitro conditions, identifying the optimal media for subculture and proliferation. Jiang et al. [4] analyzed genetic variation among and within natural populations of *I. polycarpa*, laying a basis for the growth and application of its seed resources and superior variety breeding in China. Jia et al. [5] studied the germination of *I. polycarpa* var. *vestita* Diels seeds at six maturity stages, along with oil content and fatty acid composition in the fruits. Their results revealed significant differences in seed germination rates across maturity stages, providing a scientific basis for determining the optimal harvesting period in production. Fan et al. [6] analyzed the *FAD2* gene to elucidate the molecular mechanisms of fruit oil synthesis in *I. polycarpa*, offering theoretical support for genetic engineering in breeding. Su et al. [7] conducted a diversity analysis of fruit cluster and fruit traits across 16 *I. polycarpa* samples, finding significant differences among samples, which aids in superior variety selection. Cai et al. [8] analyzed the genetic diversity of 373 *I. polycarpa* germplasm resources, exploring correlations among traits and classifying morphological types. Ji et al. [9] investigated dormancy types and optimal methods for breaking dormancy in *I. polycarpa* var. *vestita* seeds, providing efficient germination strategies for seedling production. Wang et al. [10] studied floral characteristics to understand the effect of different flower types and sizes on the growth of trees. Zhang et al. [11] explored the microecological features and internal mechanisms of *I. polycarpa* growth through leaf–soil–microbe interactions, offering scientific insights for improving cultivation practices. Ping et al. [12] analyzed the impact of several drying processes on volatile flavor compounds present in *I. polycarpa* fruits and their oil. Luo et al. [13] focused on assessing the quality features of fruit oil samples from *I. polycarpa* with different phenotypic characteristics, revealing significant differences in oil content among three phenotypic types.

With the widespread adoption of sequencing technologies and advancements in bioinformatics, chloroplast genome data have become a core tool in botanical research due to its unique genetic characteristics (such as maternal inheritance, conserved structure, stable gene arrangement) and functional significance [14]. Pan-organelle genomes have provided new evolutionary insights [15]. Zhou et al. [16] newly assembled 513 complete plastomes of *Fagopyrum* and aligned with the published plastome of eight *Fagopyrum* species to construct a pan-plastome for *F. tartaricum* and resolve genomic variation. Jia et al. [17] newly assembled the complete chloroplast genomes of 65 *Hemerocallis citrina* Baroni, and conducted a comparative analysis with the published chloroplast genomes of *H. citrina*, the detailed study of the pan-plastome and the examination of the population

structure of *H. citrina* plastomes offer crucial information for upcoming breeding initiatives and the conservation of germplasm. Wang et al. [18] assembled 322 complete *Prunus mume* Siebold and Zucc. plastomes and executed a series of comparative analyses to elucidate the pan-plastomic patterns and population structure of *P. mume*; the research results provided a better understanding of the diversity at the genus level and its history, thereby clarifying the complex introgressive relationship between mei and other apricots/plums trees.

Although significant progress has been made in *I. polycarpa* research, its taxonomic classification has long been controversial. Traditional taxonomy places *I. polycarpa* within the Flacourtiaceae family, while phylogenetic trees based on nuclear genome data support its closer relationship to the Salicaceae family (e.g., poplars) [19]. The chloroplast (cp) genome, due to its maternally inherited characteristics, could provide independent evidence for resolving its taxonomic status. For instance, comparative analyses of inversions and repeat events in cp genomes may reveal the divergence pathways between *I. polycarpa* and other Salicaceae species, thereby resolving the long-standing taxonomic disputes. Yang et al. [20] first assembled and characterized the complete cp genome sequence of *I. polycarpa* (Qingchuan, Sichuan Province, China) based on the Illumina pair-end sequencing, thereby enriching the genetic information resource library of this species.

However, research on *I. polycarpa* has mainly focused on its biological [21] and ecological characteristics [22], plant tissue culture [23,24], and fruit and seed chemical composition [5]. There is a paucity of information on its molecular biology, and comprehensive studies of its cp genome are limited. This has seriously hindered the molecular breeding of *I. polycarpa* and the development and utilization of *I. polycarpa* resources. In this study, we sequenced and annotated the complete cp genome of *I. polycarpa* and analyzed the structural characteristics. Based on cp genome data, we constructed a phylogenetic tree to clarify its taxonomic position. By comparing the cp genome of *I. polycarpa* with those of closely related species, we identified key functional genes potentially under positive selection and elucidated the genetic basis of its environmental adaptation. Furthermore, we can screen highly variable regions or specific genomic fragments to provide a theoretical foundation for subsequent analyses of population genetic diversity, species identification, and resource conservation.

## 2. Materials and Methods

### 2.1. Materials

The fresh leaves of a single individual of *I. polycarpa* were obtained from Hani-Yi Autonomous Prefecture of Honghe (Yunnan, China; 103°34'42.02" E, 23°6'19.99" N). This tree grows on the sunny slope of a mountain at an altitude of 1342.3 m. Each sample was rapidly frozen using liquid nitrogen and kept in a lab freezer at −80 °C.

### 2.2. Chloroplast Genome Sequencing, Assembly, and Annotation

The extraction of high-quality total genomic DNA was performed using a Plant Genomic DNA Extraction Kit (Beijing Solarbio Science and Technology Co., Ltd., Beijing, China). After the extracted DNA was detected by agarose gel electrophoresis, a library suitable for the sequencing platform was constructed. The constructed library was sequenced using a GISEQ-500 sequencer (Huada Manufacturing Technology Co., Ltd., Shenzhen, China), and the PE150 reads were long. After sequencing, clean reads were used for de novo assembly of cp genomes using SPAdes Assembler (v3.9.0) [25].

GeSeq [26] and ARAGORN [27] were used for genome-wide annotation. To increase the accuracy of the annotation results, Geneious (v9.1.8) [28] was used for manual proof-reading to check the differential genes between the two annotation results, remove errors

and redundancies, and obtain the final annotation. The newly annotated cp genome data of *I. polycarpa* was sent to GenBank (PV495210).

### 2.3. Chloroplast Genome Characteristics

The GC content of the cp genome sequence was determined using EditSeq (v7.1.0) software [29]. The OGDRAW (v1.3.1) software [30] was used to draw the cp genome circle of *I. polycarpa*.

### 2.4. Repeat Sequence and Simple Repeat Sequence Analysis

Using REPuter [31] online tool (<https://bibiserv.cebitec.uni-bielefeld.de/reputer>, accessed on 3 May 2025), the cp genomes of *I. polycarpa* and six closely related species (*Populus davidiana*, *Salix linearistipularis*, *Bennettiodendron brevipes*, *Poliothyrsis sinensis*, *Itoa orientalis*, *Carrierea calycina*) were documented. Five repeating sequences were found, including forward, reverse, palindromic complement, and tandem sequences. The maximum number of repeats was set to 100 bp, the minimum repeat size was 30 bp, and the Hamming distance was 3. SSR locus information of the cp genome sequences of the seven species was analyzed using the MISA online tool [32] (<https://webblast.ipk-gatersleben.de/misa/index.php>, accessed on 3 May 2025). For mononucleotide, dinucleotide, trinucleotide, tetranucleotide, pentanucleotide, and hexanucleotide, the values were set at 10, 5, 4, 3, 3, and 3, respectively, and the minimum gap between two SSRs was established at 100 bp.

### 2.5. Analysis of Codon Preference

The software codonW (v1.4.4) [33] was used to detect the number of codons and the relative synonymous codon usage (RSCU) of synonymous codons in *I. polycarpa* and six closely related species, and the RSCU values were visualized.

### 2.6. Analysis of IR Boundary

IRscope (v3.1) [34] software was used to detect the shrinkage and expansion of the LSC, SSC, and IRs region boundaries in the genomes of *I. polycarpa* and the six closely related species.

### 2.7. Analysis of Selection Pressure

Ka and Ks calculations were performed using the *I. polycarpa* sequence as the reference sequence and six closely related species sequences, respectively. First, the coding and protein sequences of the cp genome of *I. polycarpa* were extracted using the PhyloSuite (v1.2.3) software. BLASTN v2.10.1 software was used to compare the protein sequences of six related species with the protein sequences of *I. polycarpa* to find the best match to obtain homologous protein sequences. Automatic alignment of homologous protein sequences using MAFFT (v7.526) software [35]. The perl script was used to map the aligned protein sequence to the coding sequence to obtain the aligned coding sequence, and the KaKs\_Calculator (v3.0) [36] is used to calculate the Ka and Ks values.

### 2.8. Comparative Analysis of CP Genomes

The mVISTA online tool [37] (<http://genome.lbl.gov/vista/mvista/submit.shtml>, accessed on 3 May 2025) was used to select seven genomes for visual analysis in Shuffle-LAGAN mode. DnaSP (v.6) [38] was employed to determine the nucleotide diversity ( $\pi$ ) of the genome sequence, and highly variable regions in the genome were screened out. The parameters were set with a sliding window of 600 and a step size of 200.

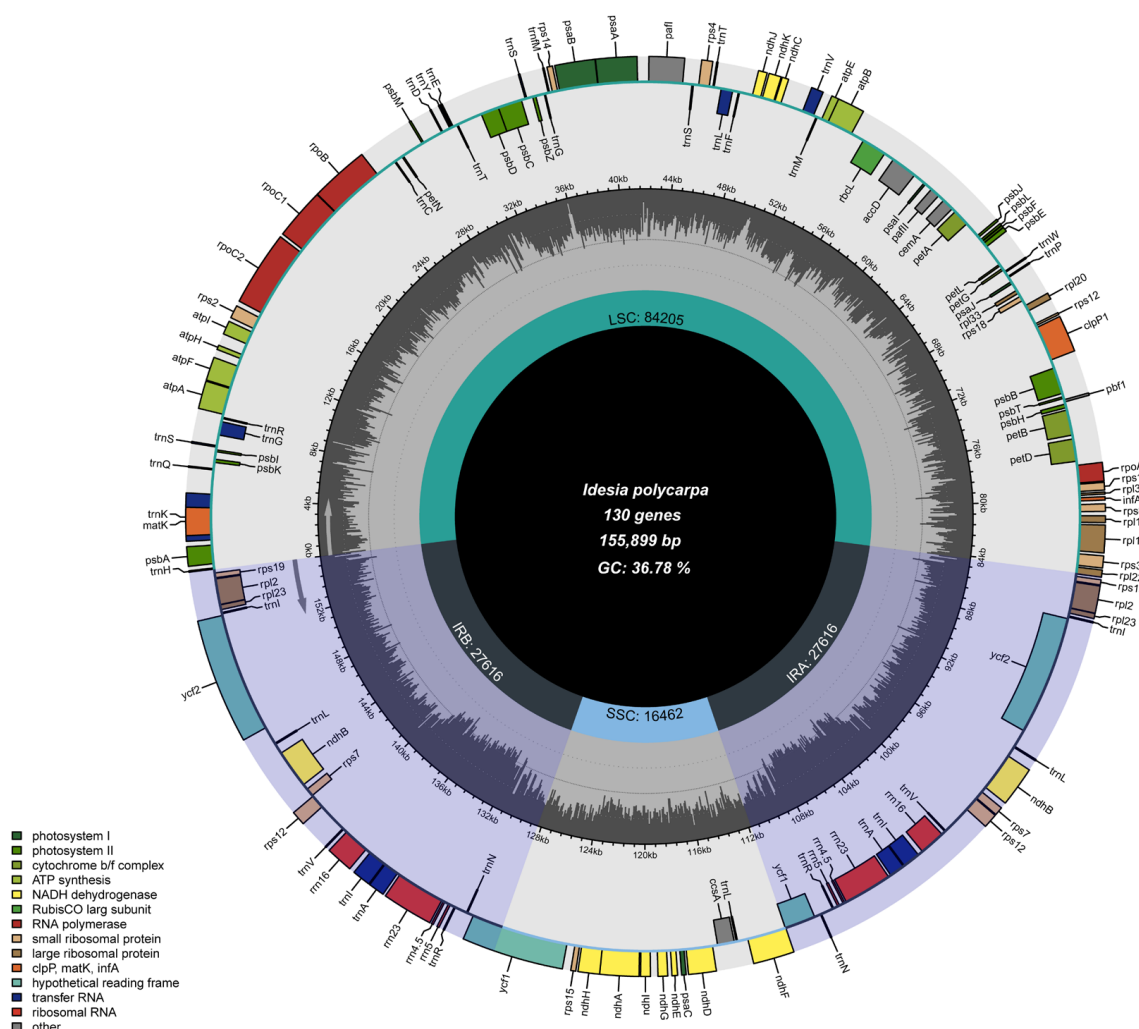
## 2.9. Phylogenetic Relationship Analysis

The genome sequences of 27 species were used, including one newly sequenced *I. polycarpa*, six Salicaceae species downloaded from NCBI, 18 Flacourtiaceae species, and two Linaceae species. Phylogenetic trees were built using these cp genome sequences, and the Linaceae family species were used as the outgroup. MAFFT was used to compare the cp genome sequences of 27 species, and MEGA11 was employed to build a phylogenetic tree. The bootstrap value was set to 1000 with LG + G as the best model.

## 3. Results

### 3.1. Chloroplast Genome Assembly and Genome Features

The cp genome of *I. polycarpa* has a typical quadruple structure (Figure 1). The total length of the genome sequence is 155,899 bp, including a large single copy region (LSC) of 84,205 bp, a small single copy region (SSC) of 16,462 bp, and two inverted repeat regions (IRa and IRb) with a combined length of 27,616 bp. The GC content of the cp genome was 36.78%, of which the CG content of the two IR regions was the highest, reaching 41.88%, the LSC region was 34.62%, and the SSC region was the lowest, 30.68%.



**Figure 1.** Gene map of the complete cp genome of *Idesia polycarpa*. The cp genome gene in the circle is transcribed clockwise, and the cp genome gene outside the circle is transcribed counter-clockwise. Different colors are used to encode genes from distinct functional groups. The inner ring's deeper gray signifies the GC content. Shown are small single copy (SSC) regions, large single copy (LSC) sequences, and inverted repeats (IRa and IRb).



The annotation included 130 genes in total, with 85 being protein-coding, 37 tRNA, and eight rRNA genes (Table 1). A total of 21 genes contained introns in the whole sequence: *trnK-UUU*, *trnG-UCC*, *atpF*, *rpoC1*, *trnL-UAA*, *trnV-UAC*, *rps12*, *rps12-2*, *petB*, *rpl16*, *rpl2*, *ndhB*, *trnI-GAU*, *trnA-UGC*, *ndhA*, *trnA-UGC-2*, *trnI-GAU-2*, *ndhB-2*, *rpl2-2*; a total of 19 genes contained one intron; *clpP1* and *pafl* contained two introns; and *petD* contained only two exons. Twelve cis-splicing genes were identified, including *atpF*, *rpoC1*, *pafl*, *clpP1*, *petB*, *petD*, *rpl16*, *rpl2* (two-copy regions), *ndhA* and *ndhB* (two-copy regions) (Table 2).

**Table 1.** Gene composition in cp genome of *Idesia polycarpa*.

Category	Gene Group	Gene Name	Number
Photosynthesis	Subunits of photosystem I	<i>psaA</i> , <i>psaB</i> , <i>psaC</i> , <i>psaI</i> , <i>psaJ</i>	5
	Subunits of photosystem II	<i>psbA</i> , <i>psbB</i> , <i>psbC</i> , <i>psbD</i> , <i>psbE</i> , <i>psbF</i> , <i>psbH</i> , <i>psbI</i> , <i>psbJ</i> , <i>psbK</i> , <i>psbL</i> , <i>psbM</i> , <i>psbT</i> , <i>psbZ</i>	14
	Subunits of NADH dehydrogenase	<i>ndhA</i> *, <i>ndhB</i> *(2), <i>ndhC</i> , <i>ndhD</i> , <i>ndhE</i> , <i>ndhF</i> , <i>ndhG</i> , <i>ndhH</i> , <i>ndhI</i> , <i>ndhJ</i> , <i>ndhK</i> ,	12
	Subunits of cytochrome b/f complex	<i>petA</i> , <i>petB</i> *, <i>petD</i> *, <i>petG</i> , <i>petL</i> , <i>petN</i>	6
	Subunits of ATP synthase	<i>atpA</i> , <i>atpB</i> , <i>atpE</i> , <i>atpF</i> *, <i>atpH</i> , <i>atpI</i>	6
	Large subunit of rubisco	<i>rbcL</i>	1
	Subunits photochlorophyllide reductase	—	
Self-replication	Proteins of large ribosomal subunit	<i>rpl14</i> , <i>rpl16</i> *, <i>rpl2</i> *(2), <i>rpl20</i> , <i>rpl22</i> , <i>rpl23</i> (2), <i>rpl33</i> , <i>rpl36</i>	10
	Proteins of small ribosomal subunit	<i>rps11</i> , <i>rps12</i> *(2), <i>rps14</i> , <i>rps15</i> , <i>rps18</i> , <i>rps19</i> (2), <i>rps2</i> , <i>rps3</i> , <i>rps4</i> , <i>rps7</i> (2), <i>rps8</i>	14
	Subunits of RNA polymerase	<i>rpoA</i> , <i>rpoB</i> , <i>rpoC1</i> *, <i>rpoC2</i>	4
	Ribosomal RNAs	<i>rrn16</i> (2), <i>rrn23</i> (2), <i>rrn4.5</i> (2), <i>rrn5</i> (2), <i>trnA-UGC</i> *(2), <i>trnC-GCA</i> , <i>trnD-GUC</i> , <i>trnE-UUC</i> , <i>trnF-GAA</i> , <i>trnG-GCC</i> , <i>trnG-UCC</i> *, <i>trnH-GUG</i> , <i>trnI-CAU</i> (2), <i>trnI-GAU</i> *(2), <i>trnK-UUU</i> *, <i>trnL-CAA</i> (2), <i>trnL-UAA</i> *, <i>trnL-UAG</i> , <i>trnM-CAU</i> , <i>trnN-GUU</i> (2), <i>trnP-UGG</i> , <i>trnQ-UUG</i> , <i>trnR-ACG</i> (2), <i>trnR-UCU</i> , <i>trnS-GCU</i> , <i>trnS-GGA</i> , <i>trnS-UGA</i> , <i>trnT-GGU</i> , <i>trnT-UGU</i> , <i>trnV-GAC</i> (2), <i>trnV-UAC</i> *, <i>trnW-CCA</i> , <i>trnY-GUA</i> , <i>trnM-CAU</i>	8
	Transfer RNAs		37
Other genes	Maturase	<i>matK</i>	1
	Protease	<i>clpP1</i> **	1
	Envelope membrane protein	<i>cemA</i>	1
	Acetyl-CoA carboxylase	<i>accD</i>	1
	c-type cytochrome synthesis gene	<i>ccsA</i>	1
	Translation initiation factor	<i>infA</i>	1
	other	<i>pafl</i> **, <i>paflI</i> , <i>pbfl</i>	3
Genes of unknown function	Conserved hypothetical chloroplast ORF	<i>ycf1</i> (2), <i>ycf2</i> (2)	4

\* Gene with one intron; \*\* gene with two introns; Gene (2): number of copies of multi-copy genes.

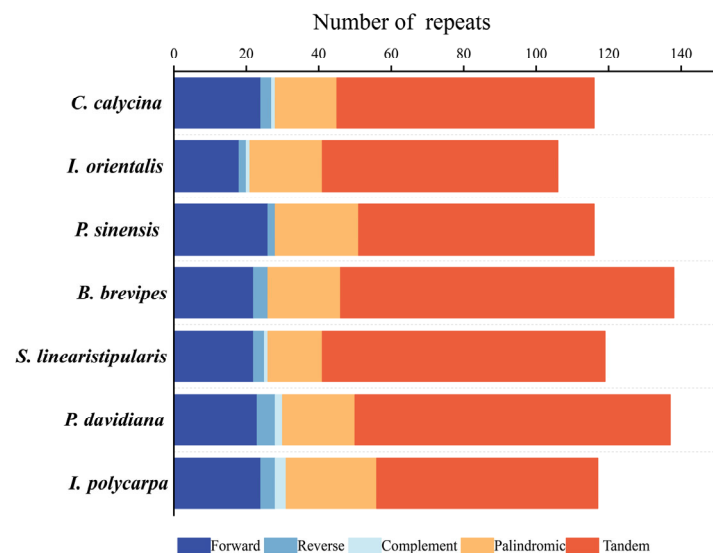
**Table 2.** Intron-containing genes and their exons and intron lengths in the cp genome of *Idesia polycarpa*.

Gene	Location	Exon1	Intron1	Exon2	Intron2	Exon3
<i>trnK</i> (UUU)	LSC	37	2544	35		
<i>trnG</i> (UCC)	LSC	23	708	48		
<i>atpF</i>	LSC	145	734	410		
<i>rpoC1</i>	LSC	432	776	1617		
<i>pafl</i>	LSC	124	722	230	712	153
<i>trnL</i> (UAA)	LSC	35	581	50		
<i>trnV</i> (UAC)	LSC	39	594	35		
<i>rps12</i>	LSC + IRa	114	536	232		26
<i>rps12-2</i>	LSC + IRb	114	536	232		26
<i>clpP1</i>	LSC	71	762	292	640	228
<i>petB</i>	LSC	6	814	642		
<i>petD</i>	LSC	8		490		
<i>rpl16</i>	LSC	9	1102	399		
<i>rpl2</i>	IRa	391	668	434		
<i>ndhB</i>	IRa	777	682	756		
<i>trnI</i> (GAU)	IRa	37	948	35		
<i>trnA</i> (UGC)	IRa	38	809	35		
<i>ndhA</i>	SSC	553	1033	545		
<i>trnA</i> (UGC)-2	IRb	38	809	35		
<i>trnI</i> (GAU)-2	IRb	37	948	35		
<i>ndhB-2</i>	IRb	777	682	756		
<i>rpl2-2</i>	IRb	391	668	434		

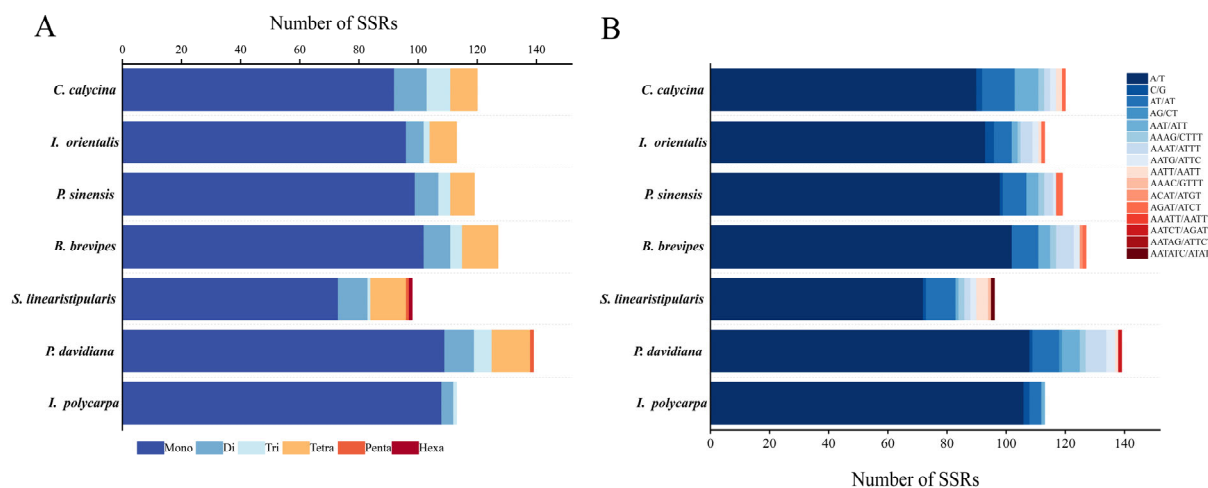
### 3.2. Repeat Sequences and SSR Analysis

A total of 330 long repetitive sequences were identified in the cp genomes of *I. polycarpa*, *P. davidiana*, *S. linearistipularis*, *B. brevipes*, *P. sinensis*, *I. orientalis*, *C. calycina*. The sequences included 159 forward repeats (F), 23 reverse repeats (R), eight complement repeats (C), and 140 palindromic repeats (P). Comparing the long repeat numbers and types in the cp genomes of seven different species. The results showed that the long-repeat types of *I. polycarpa*, *P. davidiana*, and *S. linearistipularis* were the same, and the number of long repeats was similar. The number of long repeats in *I. polycarpa* was 24F, 4R, 3C, 25P; *P. davidiana* was 23F, 5R, 2C, 20P; *S. linearistipularis* was 22F, 3R, 1C, 15P. In addition, the four closely related species, *I. polycarpa*, *B. brevipes*, *P. sinensis*, *I. orientalis*, *C. calycina* are also extremely close in the number of long repeats. *B. brevipes* and *P. sinensis* did not contain any complementary repeats. In addition, 519 tandem repeats (T) were detected in the seven species: *I. polycarpa* 61T, *P. davidiana* 87T, *S. linearistipularis* 78T, *B. brevipes* 92T, *P. sinensis* 65T, *I. orientalis* 65T, *C. calycina* 71T (Figure 2).

MISA detected SSRs in the cp genomes of the seven species, and 851 SSRs were detected. The number of *P. davidiana* SSRs was the highest (139), followed by *B. brevipes* (127), *C. calycina* (120), *P. sinensis* (119), *I. orientalis* (113), and *I. polycarpa* (113). The number of SSRs detected by *S. linearistipularis* was the lowest (98). Six types of nucleotide repeats were detected (Figure 3), including 679 mononucleotides, 58 dinucleotides, 26 trinucleotides, 63 tetranucleotides, two pentanucleotide repeats, and one hexanucleotide repeat. The SSRs of each species were mainly single-nucleotide repeats (72–109), accounting for 81.91% of the total SSRs, followed by trinucleotide SSRs, which were less abundant than dinucleotide and tetranucleotide SSRs, whereas pentanucleotide and hexanucleotide sequences were the least abundant; only one hexanucleotide was detected in *S. linearistipularis*.



**Figure 2.** Number and types of long repeats.



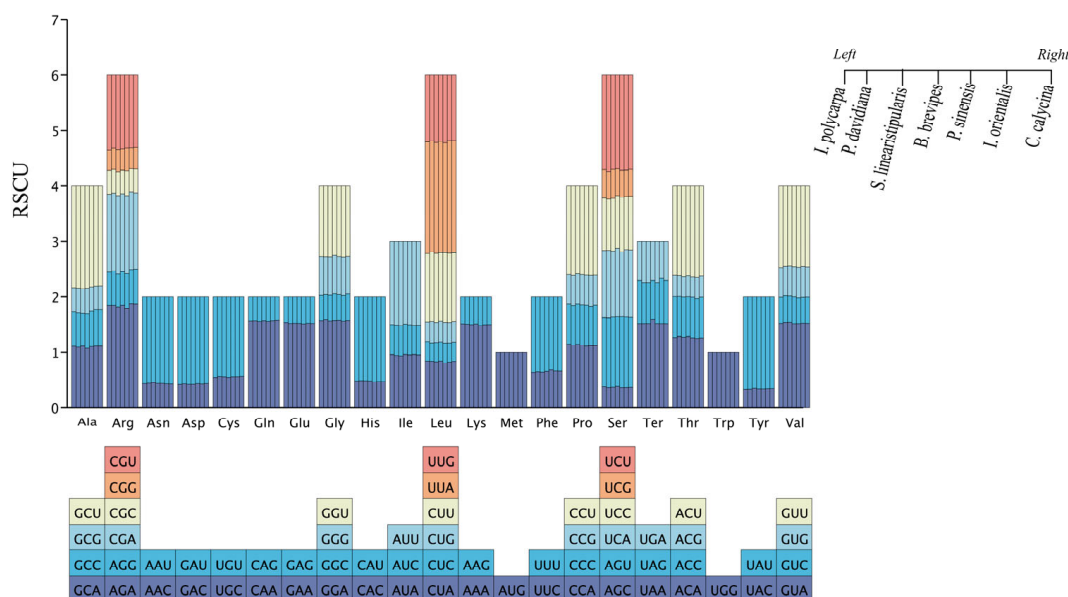
**Figure 3.** Number and types of nucleotide repeats. (A) Simple repeat sequence identification statistics. (B) Type and number of simple repetitions.

### 3.3. Codon Usage of Cp Genome

The relative synonymous codon usage values (RSCUs) of seven cp genome-coding regions of *I. polycarpa* and its related species were analyzed, and codon usage frequency and preference in the genome were detected (Figure 4). The sequenced genome encodes 21 amino acids with 22,703, 22,665, and 22,613 codons in *I. polycarpa*, *P. davidiana*, and *S. linearistipularis*, respectively. There were 22,761, 21,599, 22,764, and 22,757 codons in *B. brevipes*, *P. sinensis*, *I. orientalis*, *C. calycina*, respectively (including stop codons). Among the encoded amino acids of all species, leucine (Leu), serine (Ser), and arginine (Arg) accounted for 10.62–10.78%, 7.47–7.55%, and 5.79–5.85% of the total codons, respectively, and methionine (Met) 2.33–2.52% and tryptophan (Trp) 1.74–1.78% were the least.

In addition, the relative synonymous codon usage (RSCU) values ranged from 0.3325 to 2.0297, and most of the codons ended with A/U (A/T) (70.59–71.2%), which had a preference for A/U (A/T) bases. Arginine (Arg), leucine (Leu), and serine (Ser) have high codon preferences in individual species. Only two codons (UGG and AUG) had an RSCU of 1, indicating that methionine (Met) and tryptophan (Trp) had no codon preference.

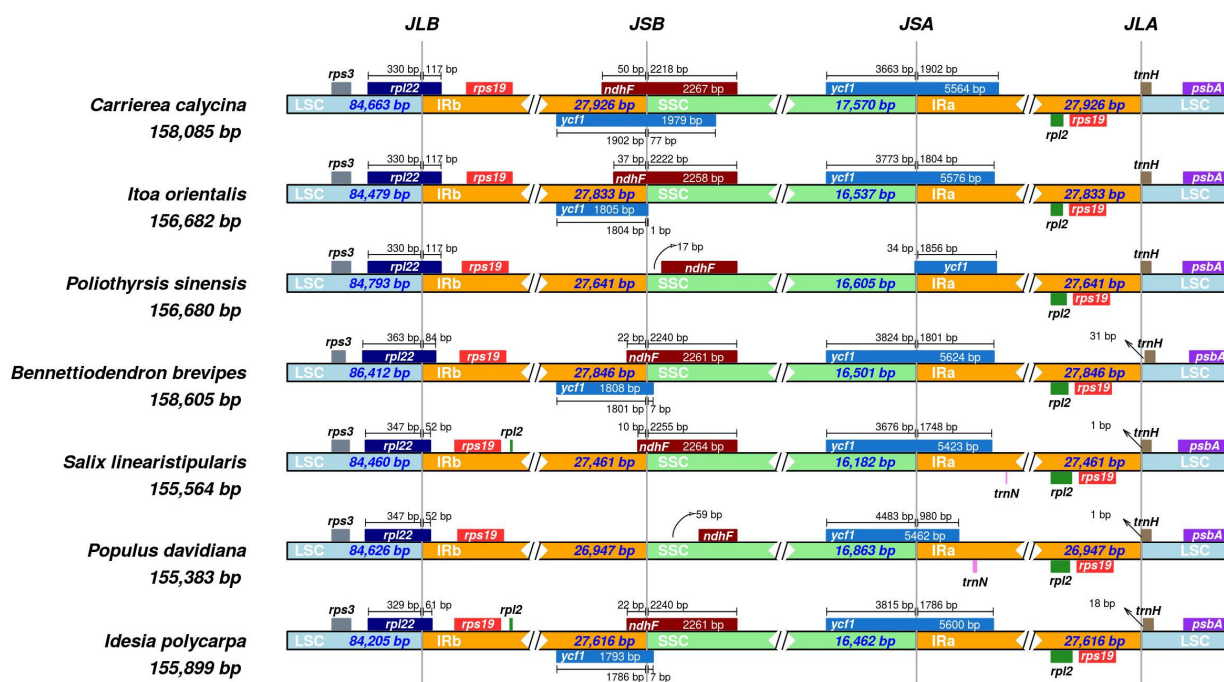




**Figure 4.** The protein-coding genes of the seven cp genomes were identified based on codon preferences and amino acid ratios of relatively synonymous codon usage (RSCU). Ter denotes the termination codon and the right side represents the ordinate species.

### 3.4. IR Boundary Analysis

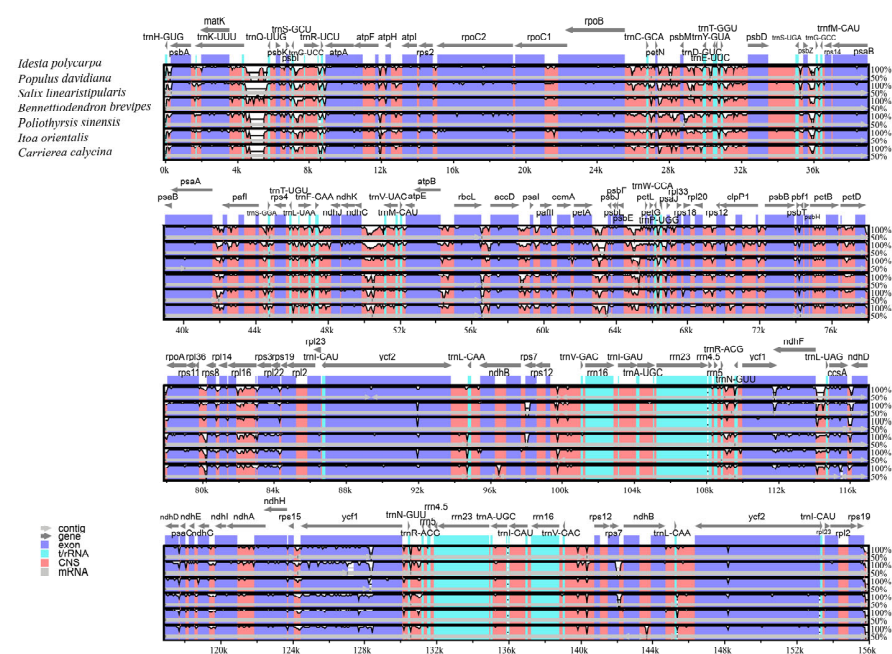
Many species commonly exhibit contraction/expansion in the boundaries of the IR region. By comparing the IR region boundaries of different species, similarities and differences between species can be identified, which can enrich genetic data for phylogenetic analysis to provide a basis for plant classification and evolutionary research. Therefore, the IR region boundaries of the cp genomes of the seven species were compared and analyzed (Figure 5).



**Figure 5.** Comparative analysis of the LSC, SSC, and IR boundaries of the cp genomes of seven species.

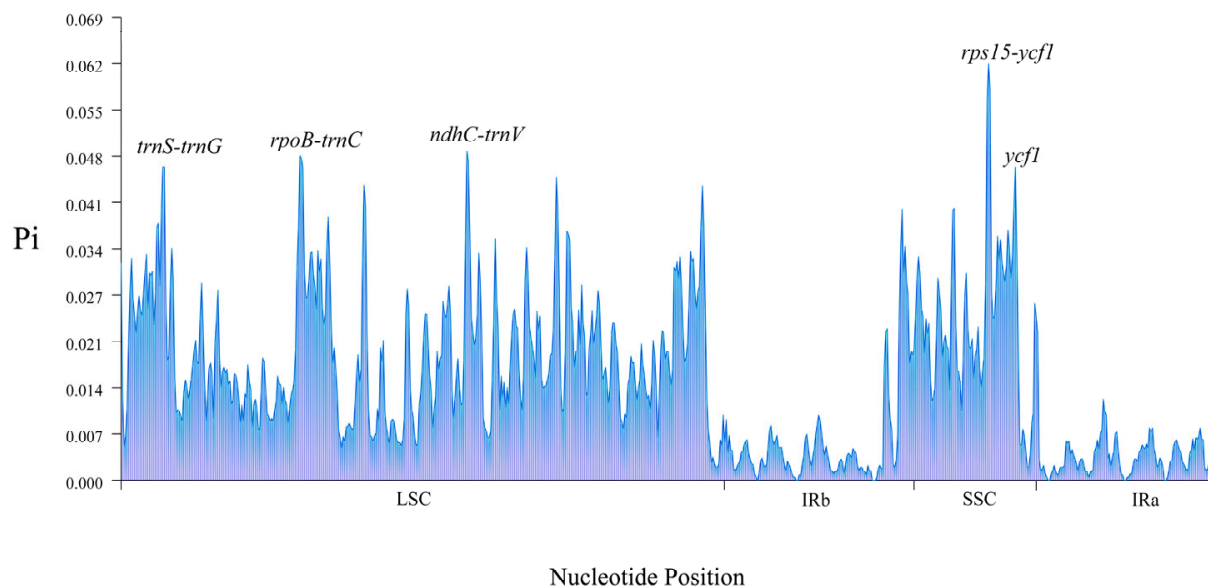
### 3.5. Comparative Analysis of Cp Genomes

The mVISTA tool was used to determine the degree of difference between the cp genome sequences of the seven species. The other six sequences were compared with the reference sequence of *I. polycarpa* (Figure 6). The results showed a certain degree of similarity among the seven sequences. Among them, the degree of variation of the IR region was lower than that of the other regions, and the degree of variation of the non-coding region was significantly higher than that of the coding region.



**Figure 6.** The mVISTA software was used to compare the cp genomes of the seven species. The gray arrow and thick black line above the alignment represent the direction and IR position of the gene, respectively. The Y axis represents the similarity percentage (50–100%). The color coding of the genomic region is protein coding (exon, purple), ribosomal RNA (rRNA, cyan), and conserved noncoding sequences (CNS, pink).

The cp genome nucleotide diversity (Pi) of the seven species was analyzed using DNAsp software. A total of 724 polymorphic sites were detected, with Pi values of 0.00048–0.06175 and an average Pi value of 0.01463. Additionally, five regions with high nucleotide variation rates ( $P_i > 0.046$ ) were screened out, which were *rps15-ycf1*, *ycf1*, *ndhC-trnV*, *rpoB-trnC*, and *trnS-trnG*, *trnG*. Five hypervariable regions were located in the LSC and SSC, and the Pi value of *rps15-ycf1* in the intergenic region was 0.06175 (Figure 7).



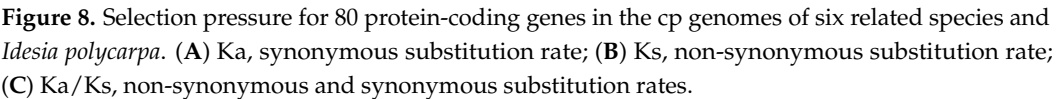
**Figure 7.** Comparison of nucleotide variability ( $P_i$ ) values of the seven cp genomes. The Y-axis represents the  $P_i$  value and the X-axis represents the region of the gene.

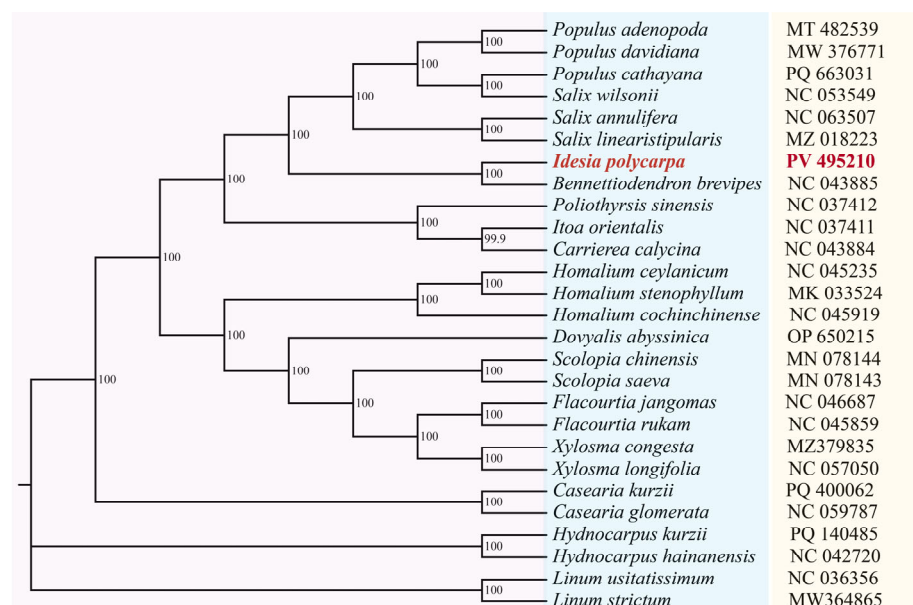
### 3.6. Selection Pressure Analysis

Synonymous ( $K_s$ ) and nonsynonymous ( $K_a$ ) substitution rates were analyzed using 80 protein-coding genes in the cp genomes of the seven species. The  $K_a/K_s$  ratios of all protein-coding genes ranged from 0 to 3.48857, with an average of 0.27. The  $K_a/K_s$  ratio of *rpl16* was the highest at 3.48857. In addition, the genes with  $K_a/K_s > 1$  were *ycf1*, *rps18*, and *cemA*, with  $K_a/K_s$  values of 1.32814, 1.09266, and 1.00319, respectively (Figure 8C).

### 3.7. Phylogenetic Analysis

The results showed that the maximum likelihood (ML) phylogenetic tree constructed with common protein-coding gene sequences was highly supported (Figure 9). The two species of *Linum usitatissimum* and *Linum strictum* of the Linaceae form independent branches as outgroups. Two species of *Hydnocarpus* genus in the Flacourtiaceae diverged first, followed by two species of *Casearia* in the Flacourtiaceae. Every species within the Salicaceae family formed a single monophyletic group, and sister groups with *I. polycarpa* and *B. brevipes*, followed by *P. sinensis*, *I. orientalis*, and *C. calycina*. The remaining Flacourtiaceae species were clustered into one branch with a relatively high support value (PP = 100).





**Figure 9.** A phylogenetic tree comprising 27 species was constructed. The species names marked in red color is the newly sequenced species identified in this study, the latter is the GeneBank accession number.

## 4. Discussion

### 4.1. Structural Features of the Cp Genome in *Idesia polycarpa*

This study involved the successful sequencing, assembly, and annotation of the entire cp genome sequence of a species of *I. polycarpa* and further analyzed its genomic characteristics, GC content, and gene structure. According to the results, the cp genome of *I. polycarpa* featured a typical tetrad configuration, with a total length of 155,899 bp, including a large single copy region (LSC) of 84,205 bp, a small single copy region (SSC) of 16,462 bp, and two inverted repeat regions (IRa and IRb) with a combined length of 27,616 bp. This result differs from that reported by Yang et al. [20], who documented a total chloroplast genome length of 157,017 bp for *I. polycarpa* samples from Qingchuan, Sichuan, China, with the LSC region measuring 84,787 bp, the SSC region 16,512 bp, and the IRs 27,859 bp each. This discrepancy could be due to the substantial genetic variation among *I. polycarpa* samples from different geographic provenances [4]. The GC content in the IR region was higher than that in the other regions (LSC and SSC), which is a common phenomenon in plants. The higher GC content in the IR region is mainly attributed to the abundance of rRNA and tRNA genes [39]. The annotation included 130 genes in total, with 85 being protein-coding, 37 tRNA, and eight rRNA genes (Table 1). This finding aligns with the total gene number of *I. polycarpa* reported by Yang et al. [20].

It is worth noting that the functions of the three genes, *pafI*, *pafII*, *pbfI* in the sequence have not been fully defined and are rarely mentioned in other studies. However, most of the sequences surrounding these three genes were tRNAs. The arrangement of this gene position may imply that it has specific functions in transcription, RNA processing, or translation regulation. Although the specific functions of these genes are not yet clear, the plant cp genome is highly conserved, indicating that they may play important roles in cp functions. It may also be a unique gene carried by the chloroplasts of *I. polycarpa*. Additional research is required to confirm the roles of these genes and explore their significance in plant physiology and evolution.



#### 4.2. Phylogenetic Relationship Between *Idesia polycarpa* and Related Species

Due to the conserved structure and slow evolution of plant cp genomes, species identification and phylogenetic evolution of cp genomes have become the developmental trend in plant taxonomic biology and have attracted increasing attention and recognition from researchers [40]. Simple sequence repeats (SSRs) play important roles in many applications, including species identification, population genetics, and phylogenetic research [14,41]. Through the study of cp genes, it was found that the number of single-nucleotide repeats was the highest in the SSRs sequence of the cp genome. This phenomenon is also true for the cp genomes of *I. polycarpa* and six related species, which contribute to A/T richness. These results are consistent with those reported for most angiosperms [42,43]. The number of tetranucleotides was slightly higher than the number of dinucleotides. There were 63 tetranucleotides, 58 dinucleotides, and 26 trinucleotides, indicating that dinucleotide and tetranucleotide repeats may provide stronger adaptive advantages or play important roles in genome stability. Long repeat sequence analyses identified 159 forward repeats and 140 palindromic repeats in the seven species. These repeat sequences may have been retained during evolution, indicating that they are important for species adaptability and survival. Codons play important roles in protein synthesis, gene regulation, and evolutionary biology [44,45]. Among the seven species, leucine (Leu), serine (Ser), and arginine (Arg) accounted for 10.62–10.78%, 7.47–7.55%, and 5.79–5.85% of the total codons, respectively. These are common amino acids that play important roles in the structure and function of photosynthesis-related proteins. Methionine (Met) and tryptophan (Trp) did not show codon preferences.

Although IR are more conserved in the cp genome than in the SSC and LSC regions, the evolutionary events and size changes in the cp genomes across various plants can also be linked to the expansion and contraction of IR boundaries [46–48]. In this study, the genes *rpl22*, *ndhF*, *ycf1*, and *trnH* at the LSC/IRb-IRa and SSC/IRb-IRa boundaries were clearly concluded to have shrunk in the IR region, whereas at the IRB/SSC boundary, *P. davidiana*, *S. linearistipularis*, *P. sinensis* species did not have *ycf1*, which may have undergone genomic recombination events during evolution, resulting in these genes being removed from the IRB/SSC boundary or relocated to other regions.

The Ka/Ks ratio is related to gene adaptive evolution, which is a meaningful marker in species evolution [49,50]. The Ka/Ks ratio can explain gene selection. In practice, some genes can be screened based on their Ka/Ks ratios for further functional studies. The Ka/Ks ratio has been widely used to study molecular evolution [51]. In the present study, the results of synonymous (Ks) and nonsynonymous (Ka) substitution rate analyses showed that Ks was much higher than Ka, indicating that the evolution of the seven species was relatively slow. The Ka and Ks values of *infA* were relatively high at 0.328445 and 1.52127, respectively. The average Ka/Ks ratio of *rpl16* was the highest, indicating strong positive selection. This gene may be rapidly evolving. However, their functions require further investigation. Similar to findings from previous research, genes involved in photosynthesis evolve more slowly than other protein-coding genes [52–54]. We identified five hyper-variable regions, *rps15-ycf1*, *ycf1*, *ndhC-trnV*, *rpoB-trnC*, and *trnS-trnG* by comparing the structures of seven cp genomes and analyzing nucleotide polymorphisms (Pi). Specific molecular markers for identifying *I. polycarpa* could be developed using these five highly variable regions.

In the traditional classification, plants of Flacourtiaceae belong to a morphologically diverse group containing approximately 90 genera and 800–1000 species, widely distributed in tropical and subtropical regions. Its morphological characteristics are highly heterogeneous, resulting in disputes over the taxonomic affiliations of many genera within this family. Some Flacourtiaceae species (such as *Idesia* and *Xylosma*) have morphological

features (such as unisexual flowers and capsules) similar to those of Salicaceae, leading researchers to speculate on the potential phylogenetic relationships between these two groups [19]. However, there is a lack of clear evidence supporting this speculation. With the development of molecular systematics, the application of DNA sequence analysis techniques (such as *rbcL*, *matK*, *ITS* transcribed spacer) has shown that the Flacourtiaceae family is not a monophyletic group. Its members are scattered across different branches of Malpighiales and some genera have close phylogenetic relationships with Salicaceae. However, previous studies were mainly based on a few short gene fragment sequences, which contained limited phylogenetic information and could not resolve interspecific phylogenetic relationships [55]. In the present study, we constructed a phylogenetic tree based on complete cp genome sequences and clarified the disputed interspecific relationships among several genera of Flacourtiaceae. This further verified that the core genera of Flacourtiaceae, such as *I. polycarpa*, *B. brevipes*, *P. sinensis*, *I. orientalis*, and *C. calycina*, had relatively close phylogenetic relationships with the Salicaceae, and the branch had a relatively high support rate (Figure 9). The phylogenetic analysis showed that *I. polycarpa* is closely related to the genera of *Salix* and *Populus*, consistent with the findings of Yang et al. [20]. These results are consistent with the official split of the Flacourtiaceae family by the APG IV classification system, where its core genera were incorporated into Salicaceae, and other genera were dispersed to different families within Malpighiales (such as Achariaceae and Salicaceae sensu lato).

## 5. Conclusions

The study is the first to reveal the structural characteristics of the cp genome of *I. polycarpa* and related species, along with the phenomenon of IR contraction. Based on the cp genome data, the classification status of *I. polycarpa* has been clarified, providing molecular evidence for the taxonomic disputes regarding the species. In the future, we will explore the molecular mechanisms causing the differences among various varieties of *I. polycarpa*, offering molecular-level support for variety selection and breeding. The five highly variable regions identified in this study lay the foundation for functional genomics and the development of molecular markers for *I. polycarpa*. We will analyze the functions of genes such as *ycf1*, *rps18*, *cemA*, and *rpl16*, and their roles in environmental adaptation, combined with population genetics, to investigate the association between the polymorphism of SSR loci and geographical distribution, thereby contributing to the efficient utilization and industrial upgrading of *I. polycarpa* resources.

**Author Contributions:** Conceptualization, X.L.; data curation, X.F. and Y.G.; formal analysis, X.F., J.L., Y.G., Y.D., H.Z. and C.S.; methodology, X.L.; project administration, X.L.; resources, J.L., D.F., Y.D., M.K. and H.Z.; software, X.F., J.L., Y.G., D.F., Y.D., M.K. and C.S.; supervision, C.S.; writing—original draft, X.F., J.L. and Y.G.; writing—review and editing, D.F., X.L. and C.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The funding for this study was provided by the Open Fund of State Key Laboratory of Tree Genetics and Breeding (Chinese Academy of Forestry) (Grant No. TGB2021009).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are openly available in GeneBank.

**Acknowledgments:** We are particularly grateful to Honghai Liang (Chongqing Wulipo National Nature Reserve Management Office) for his assistance in the field collection of the experimental materials.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Wen, L.-Y.; Mao, J.-M.; An, X.-F.; Yang, B.; Wang, Z.-R.; Kan, J.-Q. Dynamic analysis and evaluation of chemical constituents and fatty acids of *Idesia polycarpa* var. *vestita* Diels fruit. *Food Sci.* **2022**, *43*, 198–207. [\[CrossRef\]](#)
- Liu, X.-M. Study on Propagation and Tissue Culture Techniques of *Idesia polycarpa* Maxim. Master's Thesis, Northwest University of Agriculture and Forestry Science and Technology, Yangling, China, 2009.
- Yang, E.-R.; Liu, X.-M.; Xie, Q. Establishment of plantlet regeneration system of *Idesia polycarpa* in vitro. *J. Northwest For. Univ.* **2013**, *28*, 95–98. [\[CrossRef\]](#)
- Jiang, X.-B.; Gong, B.-C.; Li, D.-W.; Wu, K.-Y.; Zhao, X.-M. Variation analysis of phenotypic traits in natural population of *Idesia polycarpa*. *For. Res.* **2013**, *26*, 113–117. [\[CrossRef\]](#)
- Jia, R.-R.; Tang, X.-S.; Dong, N.; Tang, L.; Chen, F. Influence of different mature degrees on seeds germination and quality of fruit oil of *Idesia polycarpa* Maxim. var. *vestita* diels. *Seed* **2014**, *33*, 38–41. [\[CrossRef\]](#)
- Fan, R.-S.; Li, L.; Cai, G.; Ye, J.; Liu, M.-H.; Wang, S.-H.; Li, Z.-Q. Molecular cloning and function analysis of *FAD2* gene in *Idesia polycarpa*. *Phytochemistry* **2019**, *168*, 112114. [\[CrossRef\]](#) [\[PubMed\]](#)
- Su, S.; Li, Z.-J.; Ni, J.-W.; Geng, Y.-H.; Xu, X.-Q. Diversity analysis on characters of spikes and fruits of *Idesia polycarpa*. *J. Plant Resour. Environ.* **2021**, *33*, 35–44. [\[CrossRef\]](#)
- Cai, Q.-F.; Zhang, M.-M.; Xu, Z.-H.; Wang, S.; Xu, D.-H.; Liu, Z. Genetic diversity analysis of germplasm resources of *Idesia polycarpa* Maxim. *J. Henan Agric. Univ.* **2023**, *57*, 47–55. [\[CrossRef\]](#)
- Ji, Y.-F.; Song, S.-Q.; Liu, J.; Tian, X.-R. Study on characteristics of dormancy and germination in *Idesia polycarpa* seeds. *Chin. Wild Plant Resour.* **2023**, *42*, 23–28. [\[CrossRef\]](#)
- Wang, Y.-P.; Liu, C.-Y.; Hu, J.-S.; Wu, K.-Y.; Gong, B.-C.; Xu, Y. A flowering morphological investigation, fruit fatty acids, and mineral elements dynamic changes of *Idesia polycarpa* Maxim. *Plants* **2024**, *13*, 2663. [\[CrossRef\]](#)
- Zhang, T.; Wang, S.-S.; Rana, S.; Wang, Y.-M.; Liu, Z.; Cai, Q.-F.; Geng, X.-D.; Yuan, Q.-P.; Yang, Y.; Miao, C.; et al. Analysis of leaf and soil nutrients, microorganisms and metabolome in the growth period of *Idesia polycarpa* Maxim. *Microorganisms* **2024**, *12*, 746. [\[CrossRef\]](#)
- Ping, H.-R.; Ge, Y.-H.; Liu, W.-X.; Yang, J.-X.; Zhong, Z.-X.; Wang, J.-H. Effects of Different drying methods on volatile flavor compounds in *Idesia Polycarpa* Maxim fruit and oil. *Molecules* **2025**, *30*, 811. [\[CrossRef\]](#) [\[PubMed\]](#)
- Luo, T.-T.; Xu, J.-Q. Quality analysis of *Idesia polycarpa* fruit oil samples from cultivars with different phenotypes. *Food Chem. X* **2025**, *25*, 102127. [\[CrossRef\]](#)
- Daniell, H.; Lin, C.S.; Yu, M.; Chang, W.-J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [\[CrossRef\]](#)
- Wang, J.; Kan, S.-L.; Liao, X.-Z.; Zhou, J.-W.; Tembrock, L.R.; Daniell, H.; Jin, S.-X.; Wu, Z.-Q. Plant organellar genomes: Much done, much more to do. *Trends Plant Sci.* **2024**, *29*, 754–769. [\[CrossRef\]](#)
- Zhou, J.W.; He, W.-C.; Wang, J.; Liao, X.-Z.; Xiang, K.-L.; Ma, M.-C.; Liu, Z.; Li, Y.-Y.; Tembrock, L.R.; Wu, Z.-Q.; et al. The pan-plastome of tartary buckwheat (*Fagopyrum tataricum*): Key insights into genetic diversity and the history of lineage divergence. *BMC Plant Biol.* **2023**, *23*, 212. [\[CrossRef\]](#)
- Jia, M.-L.; Wang, J.; Cao, D.-M.; Jiang, C.-R.; Li, W.; Tembrock, L.-R.; Xing, G.-M.; Li, S.; Wu, Z.-Q. The pan-plastome of *Hemerocallis citrina* reveals new insights into the genetic diversity and cultivation history of an economically important food plant. *BMC Plant Biol.* **2024**, *24*, 44. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wang, J.; Kan, J.-H.; Wang, J.; Yan, X.-L.; Li, Y.; Soe, T.; Tembrock, L.R.; Xing, G.-M.; Li, S.; Wu, Z.-Q.; et al. The pan-plastome of *Prunus mume*: Insights into *Prunus* diversity, phylogeny, and domestication history. *Front. Plant Sci.* **2024**, *15*, 1404071. [\[CrossRef\]](#) [\[PubMed\]](#)
- Zuo, Y.; Liu, H.-B.; Li, B.; Zhao, H.; Li, X.-L.; Chen, J.-T.; Wang, L.; Zheng, Q.-B.; He, Y.-Q.; Zhang, J.-S.; et al. The *Idesia polycarpa* genome provides insights into its evolution and oil biosynthesis. *Cell Rep.* **2024**, *43*, 113909. [\[CrossRef\]](#)
- Yang, W.-L.; Wang, W.-W.; Zhang, L.; Chen, Z.-Y.; Guo, X.-Y.; Ma, T. Characterization of the complete chloroplast genome of *Idesia polycarpa*. *Conser. Gen. Res.* **2016**, *8*, 271–273. [\[CrossRef\]](#)
- Kikuchi, T.; Miura, O. Differentiation in vegetation related to micro-scale landforms with special reference to the lower side slope. *Ecol. Rev.* **1991**, *22*, 61–70.
- Shimada, K. Topographical distribution of five pioneer tree species and significance of their tree forms in natural forests on Mt. Takao, central Japan. *Jpn J. Ecol.* **1994**, *44*, 293–304. [\[CrossRef\]](#)
- Huang, X.-M.; Qiu, Z.-Q.; Wang, J.-J.; Wu, Y. Research progress on oil from *Idesia polycarpa*. *Cereals Oils.* **2017**, *30*, 11–13. [\[CrossRef\]](#)
- Wang, Z.-X. *Handbook of Chinese Resource Plants*, 1st ed.; Science Press: Beijing, China, 1989; pp. 233–234.
- Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Son, P.P.; Andrey, D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [\[CrossRef\]](#) [\[PubMed\]](#)

26. Tillich, M.; Lehwark, P.; Pellizzer, T.; Ulbricht-Jones, E.S.; Fischer, A.; Bock, R.; Greiner, S. GeSeq-versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **2017**, *45*, W6–W11. [[CrossRef](#)]
27. Laslett, D.; Canback, B. ARAGORN, a program to detect trna genes and tmrna genes in nucleotide sequences. *Nucleic Acids Res.* **2004**, *32*, 11–16. [[CrossRef](#)]
28. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)] [[PubMed](#)]
29. Burland, T.G. DNASTAR's lasergene sequence analysis software. *Methods Mol. Biol.* **2000**, *132*, 71–91. [[CrossRef](#)]
30. Lohse, M.; Drechsel, O.; Bock, R. Organellargenomedraw (ogdraw): A tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Curr. Genet.* **2007**, *52*, 267–274. [[CrossRef](#)]
31. Kurtz, S.; Choudhuri, J.V.; Ohlebusch, E.; Schleiermacher, C.; Stoye, J.; Giegerich, R. Reputer: The manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* **2001**, *29*, 4633–4642. [[CrossRef](#)]
32. Beier, S.; Thiel, T.; Munch, T.; Scholz, U.; Mascher, M. MISA-web: A web server for microsatellite prediction. *Bioinformatics* **2017**, *33*, 2583–2585. [[CrossRef](#)]
33. Sharp, P.M.; Li, W.-H. The codon adaptation index—a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* **1987**, *15*, 1281–1295. [[CrossRef](#)] [[PubMed](#)]
34. Amiryousefi, A.; Hyvonen, J.; Poczai, P. IRscope: An online program to visualize the junction sites of chloroplast genomes. *Bioinformatics* **2018**, *34*, 3030–3031. [[CrossRef](#)] [[PubMed](#)]
35. Katoh, K.; Toh, H. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* **2010**, *26*, 1899–1900. [[CrossRef](#)] [[PubMed](#)]
36. Zhang, Z. Kaks\_calculator 3.0: Calculating selective pressure on coding and noncoding sequences. *Genom. Proteom. Bioinform.* **2022**, *20*, 536–540. [[CrossRef](#)]
37. Frazer, K.A.; Pachter, L.; Poliakov, A.; Rubin, E.M.; Dubchak, I. VISTA: Computational tools for comparative genomics. *Nucleic Acids Res.* **2004**, *32*, W273–W279. [[CrossRef](#)]
38. Rozas, J.; Ferrer-Mata, A.; Carlos Sanchez-DelBarrio, J.; Guirao-Rico, S.; Librado, P.; Ramos-Onsins, S.E.; Sanchez-Gracia, A. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **2017**, *34*, 3299–3302. [[CrossRef](#)]
39. Liu, H.; Liu, X.; Sun, C.; Li, H.-L.; Li, Z.-X.; Guo, Y.; Fu, X.-Q.; Liao, Q.-H.; Zhang, W.-L.; Liu, Y.-Q. Chloroplast genome comparison and phylogenetic analysis of the commercial variety *Actinidia chinensis* 'Hongyang'. *Genes* **2023**, *14*, 2136. [[CrossRef](#)]
40. Nock, C.J.; Waters, D.L.E.; Edwards, M.A.; Bowen, S.G.; Cordeiro, G.M.; Henry, R.J. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.* **2011**, *9*, 328–333. [[CrossRef](#)]
41. Guo, Q.; Cao, S.; Li, D.; Li, X.-Y.; Zhang, J.-T.; Zhang, Y.-S.; Zhang, Z.-J.; Sun, Y.-H.; Cui, L.; Fan, Y.-M.; et al. Genetic diversity and population structure of *Robinia pseudoacacia* from six improved variety bases in China as revealed by simple sequence repeat markers. *J. For. Res.* **2022**, *33*, 611–621. [[CrossRef](#)]
42. Li, W.; Zhang, C.-P.; Guo, X.; Liu, Q.-H.; Wang, K.-L. Complete chloroplast genome of *Camellia japonica* genome structures, comparative and phylogenetic analysis. *PLoS ONE* **2019**, *14*, e0216645. [[CrossRef](#)]
43. Cui, Y.-X.; Nie, L.-P.; Sun, W.; Xu, Z.-C.; Wang, Y.; Yu, J.; Song, J.-Y.; Yao, H. Comparative and phylogenetic analyses of ginger (*Zingiber officinale*) in the family Zingiberaceae based on the complete chloroplast genome. *Plants* **2019**, *8*, 283. [[CrossRef](#)] [[PubMed](#)]
44. Hershberg, R.; Petrov, D.A. Selection on codon bias. *Annu. Rev. Genet.* **2008**, *42*, 287–299. [[CrossRef](#)] [[PubMed](#)]
45. Parvathy, S.T.; Udayasuriyan, V.; Bhadana, V. Codon usage bias. *Mol. Biol. Rep.* **2022**, *49*, 539–565. [[CrossRef](#)]
46. Raubeson, L.A.; Peery, R.; Chumley, T.W.; Dziubek, C.; Fourcade, H.M.; Boore, J.L.; Jansen, R.K. Comparative chloroplast genomics: Analyses including new sequences from the angiosperms *Nuphar advena* and *Ranunculus macranthus*. *BMC Genom.* **2007**, *8*, 174. [[CrossRef](#)]
47. Androsiuk, P.; Jastrzebski, J.P.; Paukszto, L.; Makowczenko, K.; Okorski, A.; Pszczolkowska, A.; Chwedorzewska, K.J.; Gorecki, R.; Gielwanowska, I. Evolutionary dynamics of the chloroplast genome sequences of six *Colobanthus* species. *Sci. Rep.* **2020**, *10*, 11522. [[CrossRef](#)]
48. Zhao, J.-T.; Xu, Y.; Xi, L.-J.; Yang, J.-W.; Chen, H.-W.; Zhang, J. Characterization of the chloroplast genome sequence of *Acer miaotaiense*: Comparative and phylogenetic analyses. *Molecules* **2018**, *23*, 1740. [[CrossRef](#)]
49. Li, W.; Wu, C.; Luo, C.-C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **1985**, *2*, 150–174. [[CrossRef](#)] [[PubMed](#)]
50. Raman, G.; Nam, G.H.; Park, S. Extensive reorganization of the chloroplast genome of *Corydalis platycarpa*: A comparative analysis of their organization and evolution with other *Corydalis* plastomes. *Front. Plant Sci.* **2022**, *13*, 1043740. [[CrossRef](#)] [[PubMed](#)]
51. Navarro, A.; Barton, N.H. Chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes. *Science* **2003**, *300*, 321–324. [[CrossRef](#)]

52. Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Müller, K.F.; Quandt, D. The evolution of the plastid chromosome in land plants: Gene content, gene order, gene function. *Plant Mol. Biol.* **2011**, *76*, 273–297. [[CrossRef](#)]
53. Saina, J.K.; Gichira, A.W.; Li, Z.-Z.; Hu, G.-W.; Wang, Q.-F.; Liao, K. The complete chloroplast genome sequence of *Dodonaea viscosa*: Comparative and phylogenetic analyses. *Genetica* **2017**, *146*, 101–113. [[CrossRef](#)] [[PubMed](#)]
54. Tian, N.; Han, L.; Chen, C.; Wang, Z.-Z. The complete chloroplast genome sequence of *Epipremnum aureum* and its comparative analysis among eight Araceae species. *PLoS ONE* **2018**, *13*, e0192956. [[CrossRef](#)] [[PubMed](#)]
55. Liu, X.; Wang, Z.-S.; Wang, D.-S.; Zhang, J.-G. Phylogeny of *Populus-Salix* (Salicaceae) and their relative genera using molecular datasets. *Biochem. Syst. Ecol.* **2016**, *68*, 210–215. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.