


Article

Comparative Analysis of the Characteristics, Phylogenetic Relationships of the Complete Chloroplast Genome, and Maternal Origin Track of White Poplar Interspecific Hybrid GM107

Bin Guo ^{1,2,†}, Tingting Chen ^{1,†}, Ying Li ¹, Shanwen Li ³, Wasif Ullah Khan ¹, Ren-Gang Zhang ⁴, Kai-Hua Jia ⁵ and Xinmin An ^{1,*} 

¹ State Key Laboratory of Tree Genetics and Breeding, National Engineering Research Center of Tree Breeding and Ecological Remediation, Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China

² Shanxi Academy of Forestry and Grassland Sciences, Taiyuan 030012, China

³ Shandong Academy of Forestry, Jinan 250014, China

⁴ Yunnan Key Laboratory for Integrative Conservation of Plant Species with Extremely Small Populations/Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China

⁵ Key Laboratory of Crop Genetic Improvement & Ecology and Physiology, Institute of Crop Germplasm Resources, Shandong Academy of Agricultural Sciences, Jinan 250100, China

* Correspondence: anxinmin@bjfu.edu.cn

† These authors contributed equally to this work.



Citation: Guo, B.; Chen, T.; Li, Y.; Li, S.; Khan, W.U.; Zhang, R.-G.; Jia, K.-H.; An, X. Comparative Analysis of the Characteristics, Phylogenetic Relationships of the Complete Chloroplast Genome, and Maternal Origin Track of White Poplar Interspecific Hybrid GM107. *Forests* **2023**, *14*, 587. <https://doi.org/10.3390/f14030587>

Academic Editor: Carol A. Loopstra

Received: 30 January 2023

Revised: 11 March 2023

Accepted: 12 March 2023

Published: 16 March 2023



Copyright: © 2023 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/>).

Abstract: White poplars are a dominant tree species in natural ecosystems throughout China, where they are also widely planted for agroforestry and industrial uses. However, the phylogenetic relationships among poplars in section *Populus* have a complex genetic background influenced by frequent hybridization events and, to date, only a few studies have attempted to clarify this background. In this study, we performed de novo assembly of the whole chloroplast (cp) genome of an elite individual GM107 with the size of 156,493 bp, which consists of a large single-copy (LSR) (84,610 bp), a small single-copy (SSC) (16,567 bp), and a pair of inverted repeats (27,658 bp). It comprises 127 genes, including 85 protein-coding genes, 36 tRNAs, and 6 rRNAs. Comparative analysis of the cp genomes was conducted among 7 poplars in section *Populus* and 4 cp DNA markers with >1% variable sites were detected. We found that *Populus alba* was the most closely related species to GM107 by phylogenetic analyses. RNA sequencing detected 66 genes that participated in translation, transcription, and photosynthesis. The expression levels of almost all 66 genes were higher in leaves than in other tissues, except for *PtatpF* and *PtatpH*. In all tissues, we detected higher transcript abundances of *PtndhF*, *PtpsbA*, *PtpsbB*, *Pttrps14*, *PtatpF*, and *PtatpH* than of other genes. Both cp genome and transcriptome data help understand evolution events in section *Populus* and unravel the origin of Chinese white poplars, and may contribute to the molecular genetic improvement of wood properties and carbon sink capacities in the breeding of poplars in this section.

Keywords: chloroplast genome; phylogenetic analysis; white poplars; transcriptome

1. Introduction

Genus *Populus*, which consists of nearly 30 species with abundant natural distribution in the Northern Hemisphere, is taxonomically classified into six sections: *Populus*, *Tacamahaca*, *Aigeiros*, *Turanga*, *Leucoides*, and *Abaso*. The *Populus* section, comprised of white poplar, includes *Populus tremuloides*, *Populus davidiana*, *Populus tremula*, *Populus adenopoda*, *Populus sieboldii*, *Populus alba*, and *Populus grandidentata*. These species are often segregated into geographical and morphological subspecies and varieties due to their wide natural

distribution [1]. Their capacity for outcrossing and wind pollination, as well as the absence of barriers to crossbreeding, has led to numerous hybrids [2].

Chinese white poplar is indigenous to China, with a wide distribution range [3]. It exhibits excellent features, such as rapid growth, a thick and straight trunk, strong resistance to environmental stresses, and a long lifespan (typically 100–200 years, but reaching 500 years). It is an industrial tree species, providing timber, pulp, and biofuel [4]. Thus, it is more economically, ecologically, and evolutionarily valuable than the other white poplars found in China [5].

Interspecific and intraspecific hybridization occurs frequently among white poplars due to long-distance pollination by wind and overlapping distribution ranges over large geographical extents [6]. Therefore, Chinese white poplar has a complex genetic background. It was originally thought to be a distinct species in the *Populus* section [2], but is now considered to be a hybrid of *P. alba* and *P. adenopoda* or a tri-hybrid of both species with *P. davidiana* [7,8]. Recent evidence indicated that another white poplar elite individual GM15 originated from a hybridization event of *P. adenopoda* and *P. alba* var. *pyramidalis* approximately 3.93 Mya [5], which suggests that the origin of Chinese white poplar is more complicated than previously described. The development of third-generation sequencing technology has significantly improved the study of poplar phylogeny. Analyzing the cp genome through phylogenetic analysis is a convincing method to unravel the mystery of hybrid origins.

Regarding the photosynthetic organelle, the chloroplast (cp) genome consists of circular DNA molecules that are highly similar and range in size from 115 to 165 kb [6,9]. The cp has a conserved genomic structure, gene sequence, and gene type, and consists of a large single-copy (LSC) region and a small single-copy (SSC) region separated by a pair of inverted repeats (IRa and IRb) [10,11]. The cp genome is a metabolic center with highly conserved physiological cell functions, especially for genes related to photosynthesis [9,12]. The cp genome is characterized by a slow molecular evolution rate and follows the single parental inheritance principle. Therefore, cp genomes are widely used in population taxonomy and phylogeny, genetics, and evolution studies [13,14]. The parental origin of hybrids with complex genetic backgrounds can be more accurately traced using cp genomes, which are considered more valuable and reliable markers than any others. With the advancement of third-generation sequencing technology, cp genomes are becoming increasingly popular.

In our study, we conducted a comprehensive analysis by performing de novo assembly of the cp genome of an elite individual GM107 and comparative analysis with seven other white poplar species to reveal its unique structures, variations, and repeat sequences, and to elucidate the phylogenetic relationships among white poplar species. In addition, we performed transcriptome sequencing and examined the gene expression profiles related to photosynthesis, transcription, and translation of cps in the roots, stems, leaves, and petioles of four poplar species. This study will offer substantial proof of the ancestry of the Chinese white poplar and aid in the preservation and optimization of its genetic resources.

2. Materials and Methods

2.1. Plant DNA Extraction and Sequencing

We collected young leaf tissue samples from 1-month-old tissue-cultured plantlets of an elite individual GM107 (male clone) and extracted genomic DNA immediately after sampling using the Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The quality of the genomic DNA was visualized by running agarose gel and quantified using a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) to determine its integrity, purity, and concentration. The short-insert polymerase chain reaction (PCR) free genomic DNA library (300–500 bp) was constructed according to the manufacturer's instructions (Illumina, San Diego, CA, USA) for paired-end sequencing. A 20 kb genomic DNA library was constructed for single-molecule, real-time (SMRT) sequencing according to the manufacturer's instructions (Oxford Nanopore Technologies (ONT), Oxford,

UK). Finally, sequencing was conducted on the Illumina HiSeq X Ten sequencer on the ONT platform.

2.2. Plant RNA Extraction and Transcriptome Sequencing

Tissue samples from the roots, stems, leaves, and petioles were collected from 4 1-month-old tissue-cultured plantlets of *P. tomentosa* GM15, an elite individual GM107, *P. alba* var. *pramidalis*, and *P. trichocarpa*, and treated instantly with liquid nitrogen for RNA extraction. Four replicates were obtained for each sample. Total RNA was extracted with TRIzol reagent and mRNAs were refined through NEBNext Ultra RNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA). Two micrograms of RNA from each sample were used to prepare the RNA sequencing libraries and then sequenced on the DNBSAQ platform. Salmon v1.6.0 software [15] was used to acquire transcript abundances (transcripts per million; TPM) for each sample and the number of reads originating from the transcripts was estimated. Additionally, Tximport software package was utilized to prepare inputs for expression profile analysis [16]. We used the mean of the four replicates to scan for expressed genes (TPM > 1 in at least 1 tissue), and then normalized the data ($\log_{10}(\text{TPM} + 1)$) to draw heatmaps using the *heatmap* package in R software (R Core Team, Vienna, Austria).

2.3. Chloroplast Genome Assembly and Annotation

After the filtering process, clean reads were produced, including those from nuclear and cp genomes. Reads from the cp genome were obtained by alignment with 14 related species cp genomes (including *Populus trichocarpa* NC_009143.1, *Populus balsamifera* NC_024735.1, *P. alba* NC_008235.1, *P. tremula* NC_027425.1, *P. tremula* × *P. alba* NC_028504.1, *P. davidiana* NC_032717.1, *Populus lasiocarpa* NC_036040.1, *P. adenopoda* NC_032368.1, *Populus rotundifolia* NC_033876.1, *Populus fremontii* NC_024734.1, *Populus qionghaoensis* NC_031398.1, *Populus ilicifolia* NC_031371.1, and *Populus euphratica* NC_024747.1) using Graphmap software [17]. The CANU de novo assembler was utilized to assemble the obtained reads [18]. Subsequently, the primary assembly was polished by ONT sequencing using next-generation sequencing data [19]. The cp genome was thoroughly annotated using the GeSeq software, utilizing its default parameters as specified on the Chlorobox website (<https://chlorobox.mpimp-golm.mpg.de/geseq.html>, accessed on 18 December 2022). The visualization of circular chloroplast genome maps was accomplished through the utilization of the Organelle Genome DRAW software [20]. The cp sequence generated in this study was submitted to the National Center for Biotechnology Information (NCBI) database with Genebank accession no. MK251149.

2.4. Comparative Genomic Analyses

We downloaded cp genome sequences of *Salix suchowensis* and 7 *Populus* species (*P. davidiana*, *P. tremula*, *P. alba*, *P. adenopoda*, *P. rotundifolia*, *P. tremula* × *P. alba*, and *P. alba* × *Populus glandulosa*) published in genebank. Using the mVISTA program (<http://genome.lbl.gov/vista/mvista/submit.shtml>, accessed on 18 December 2022), we compared the complete plastid genomes of these *Populus* species to the annotated *P. tomentosa* cp genome as a reference [6], and investigated the regions of high divergence as well as variable and parsimony informative sites across the complete cp genomes. Additionally, the LSC, SSC, and IR regions of 7 taxa were identified using DnaSP v5.0 software [21]. The codon usage of the elite individual GM107 cp genome was analyzed using the CodonW program in MEGA 7.0 [22].

2.5. Characterization of Repeat Sequences and SSRs

The numbers of large repeats, including forward, reverse, palindromic, and complementary repeats, were identified using the REPuter program [23] based on the following criteria: sequence identities of 90%, cutoff point of ≥ 30 bp, Hamming distance of 3, and minimum repeat size of 30 bp. Furthermore, the Tandem Repeats Finder online program was

utilized to identify tandem repeat sequences with a size of more than 10 bp, using default parameters [24] Lastly, SSRs were identified through MISA Perl script [25], with a minimum threshold of 10 repeat units for mono-, di-, tri-, tetra-, penta-, and hexanucleotides.

2.6. Phylogenetic Analyses

The phylogenetic analysis was performed using the complete cp genomes of 7 *Populus* species (*P. davidiana*, *P. tremula*, *P. alba*, *P. adenopoda*, *P. rotundifolia*, *P. tremula* × *P. alba*, and *P. alba* × *P. glandulosa*), where *Salix suchowensis* (NC_026462.1) was used as the outgroup. Complete cp genome alignment was performed using MAFFT v.7 software [26] and manually edited as required. Subsequently, phylogenetic analysis was conducted using maximum likelihood (ML) methods based on the 8 cp whole-genome sequence alignments. In order to generate a robust phylogenetic tree, the ML analyses were conducted using IQ-tree Inference of Quantitative Phylogenies Tree software [27] with 10,000 ultrafast bootstrap replicates [28].

3. Results

3.1. Cp Genome Structure

We conducted de novo assembly of the complete cp genome of the elite individual GM107. The primary assembly generated from ONT data was polished using Illumina data, resulting in a final cp genome assembly of 156,493 bp (Figure 1, Table 1).

Table 1. Summary of complete chloroplast genome of the elite individual GM107.

Characteristics	Elite Individual GM107
Length (bp)	156,493
GC content (%)	36.74
AT content (%)	63.26
LSC GC content (%)	39.31
SSC GC content (%)	30.49
IR GC content (%)	41.95
LSC length (bp)	84,610
SSC length (bp)	16,567
IR length (bp)	27,658
CDS length (bp)	80,562
tRNA length (bp)	2648
rRNA length (bp)	8599
Gene number	127
Gene number in IR regions	34
Protein-coding gene number	85
Protein-coding gene (%)	66.93
rRNA gene number	6
rRNA (%)	4.72
tRNA gene number	36
tRNA (%)	28.35

The chloroplast genome features a circular molecular structure composed of three regions. The largest region, referred to as the large single-copy (LSC) region, spans 84,610 base pairs and makes up 54.1% of the genome. The second region, the small single-copy (SSC) region, is much smaller, making up 10.6% of the genome, consisting of 16,567 base pairs. The remaining 17.7% of the genome is divided into two intergenic regions (IRa/b), each with a size of 27,658 base pairs (Table 1).

The cp genome contains 127 genes, including 85 protein-coding genes (66.93% of the total), 36 tRNA genes (28.35%), and 6 rRNA genes (4.72%). The noncoding structures, such as intergenic spacers and introns, occupy the remaining portion. The entire genome has GC content of 36.5% with increased GC concentration (41.95%) in the IR regions, followed by LSC (39.31%) and SSC (30.49%), respectively (Table 1).

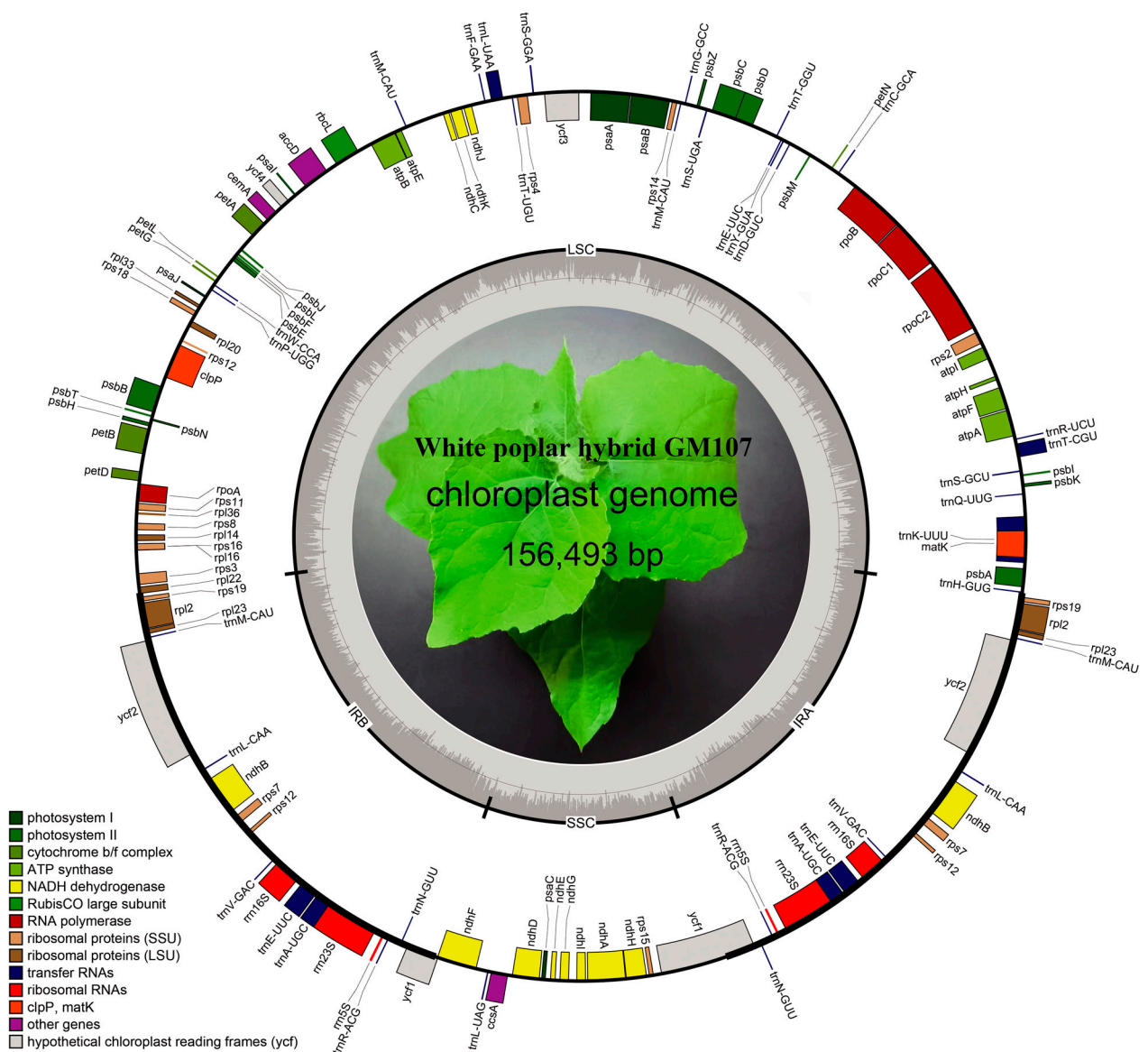


Figure 1. Gene circle map of the elite individual GM107 chloroplast (cp) genome. The colored bars in the diagram represent various functional groups. The bold lines of the main circle outline the extent of the inverted repeat regions (IRA and IRB), dividing the genome into small single-copy (SSC) and large single-copy (LSC) regions. Genes located on the inner and outer parts of the main circle are transcribed in a clockwise and counterclockwise direction, respectively. The inner circle features dark gray and light gray columns that denote the GC and AT content of the DNA, respectively.

All genes were classified into four categories: photosynthesis-performing genes, transcription-associated and translation-associated genes, genes related to RNA and its processing, and other genes. A total of 17 genes possessed introns, with 15 of these genes (comprising 9 protein-coding genes and 6 transfer RNA genes) possessing a single intron each, while *clpP* and *ycf3* possessed 2 introns each. Additionally, 18 genes were duplicated in the IR regions (Table 2). These findings align with those reported in other *Populus* chloroplast genomes [5,29,30].

Table 2. List of annotated genes in the chloroplast genome of the elite individual GM107.

Category of Genes	Family Name List of Genes	List of Genes
Photosynthesis-related genes	Subunit of rubisco	<i>rbcL</i>
	Subunits of photosystem I	<i>psaA, psaB, psaC, psaI, psaJ</i>
	Assembly of photosystem I	<i>ycf4, ycf3</i> **
	Subunits of photosystem II	<i>psbA, psbB, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ</i>
	Subunits of ATP synthase	<i>atpA, atpB, atpE, atpF</i> *, <i>atpH, atpI</i>
	Subunits of cytochrome b/f complex	<i>petA, petB</i> *, <i>petD, petG, petL, petN</i>
	c-type cytochrom synthesis gene	<i>ccsA</i>
Transcription-related and translation-related genes	Subunits of NADH-dehydrogenase	<i>ndhA</i> *, <i>ndhB</i> * (×2), <i>ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK</i>
	Subunits of protochlorophyllide reductase	-
	DNA-dependent RNA polymerase	<i>rpoA, rpoB, rpoC1, rpoC2</i>
	Small subunit of ribosome	<i>rps11, rps12</i> * (×2), <i>rps14, rps15, rps16, rps18, rps19</i> (×2), <i>rps2, rps3, rps4, rps7</i> (×2), <i>rps8</i>
	Large subunit of ribosome	<i>rpl14, rpl16, rpl2</i> * (×2), <i>rpl20, rpl22, rpl23</i> (×2), <i>rpl33, rpl36</i>
	Translational initiation factor	-
	Elongation factor	-
RNA and its processing	rRNA genes	<i>rrn16S</i> (×2), <i>rrn23S</i> (×2), <i>rrn5S</i> (×2)
	tRNA genes	<i>trnH-GUG, trnK-UUU</i> *, <i>trnQ-UUG, trnS-GCU, trnT-CGU</i> *, <i>trnR-UCU, trnC-GCA, trnD-GUC, trnY-GUA, trnE-UUC</i> *, <i>trnT-GGU, trnS-UGA, trnG-GCC, trnM-CAU, trnS-GGA, trnT-UGU, trnL-UAA</i> *, <i>trnF-GAA, trnM-CAU, trnW-CCA, trnP-UGG, trnM-CAU</i> (×2), <i>trnL-CAA</i> (×2), <i>trnV-GAC</i> (×2), <i>trnE-UUC</i> (×2), <i>trnA-UGC</i> * (×2), <i>trnR-ACG</i> (×2), <i>trnN-GUU</i> (×2), <i>trnL-UAG</i>
Other genes	RNA processing	<i>matK</i>
	Subunit of Acetyl-CoA-carboxylase	<i>accD</i>
	Envelop membrane protein	<i>cemA</i>
	Protease	<i>clpP</i> **
	Component of TIC complex	<i>ycf1</i> (×2)
	Hypothetical proteins	<i>ycf2</i> (×2)

* Genes with one intron; ** genes with two introns; (×2) genes with duplication in the IR regions.

3.2. Expansion and Contraction of IR Regions

Comparative sequence analysis of eight white poplar species indicated a high degree of conservation in the chloroplast genome structure, gene number, and sequence. Nevertheless, variations in structure and size at the IR boundaries were noted (as shown in Figure 2). The boundary sites between regions are designated as JLB (IRb/LSC), JSB (IRb/SSC), JSA (SSC/IRa), and JLA (IRa/LSC). The genes located in the boundary regions consist of *rpl22, rps19, trnL, ycf1, trnN, ndhF, trnH, and ycf15*. All eight analyzed species had functional *ycf1* and *trnN* genes in the JSA region, and six species had *ycf1* and *ndhF* genes in the JSB region, while the other two had *trnN* and *ycf1* (GM107) and *trnN* and *ndhF* genes (*P. rotundifolia*).

Seven species had functional *rpl22* and *rps19* genes in the JLB region, and *rps19* and *trnH* in the JLA region. Only *P. rotundifolia* had *trnL* and *ycf15* in both JLB and JLA regions. The genes' placement in the JLB and JLA boundary regions was comparatively stable, while the JSB and JSA boundary regions exhibited significant variation across the eight species. Such differences in the chloroplast (cp) genome often result in expansion and contraction, leading to divergence events between species [9].

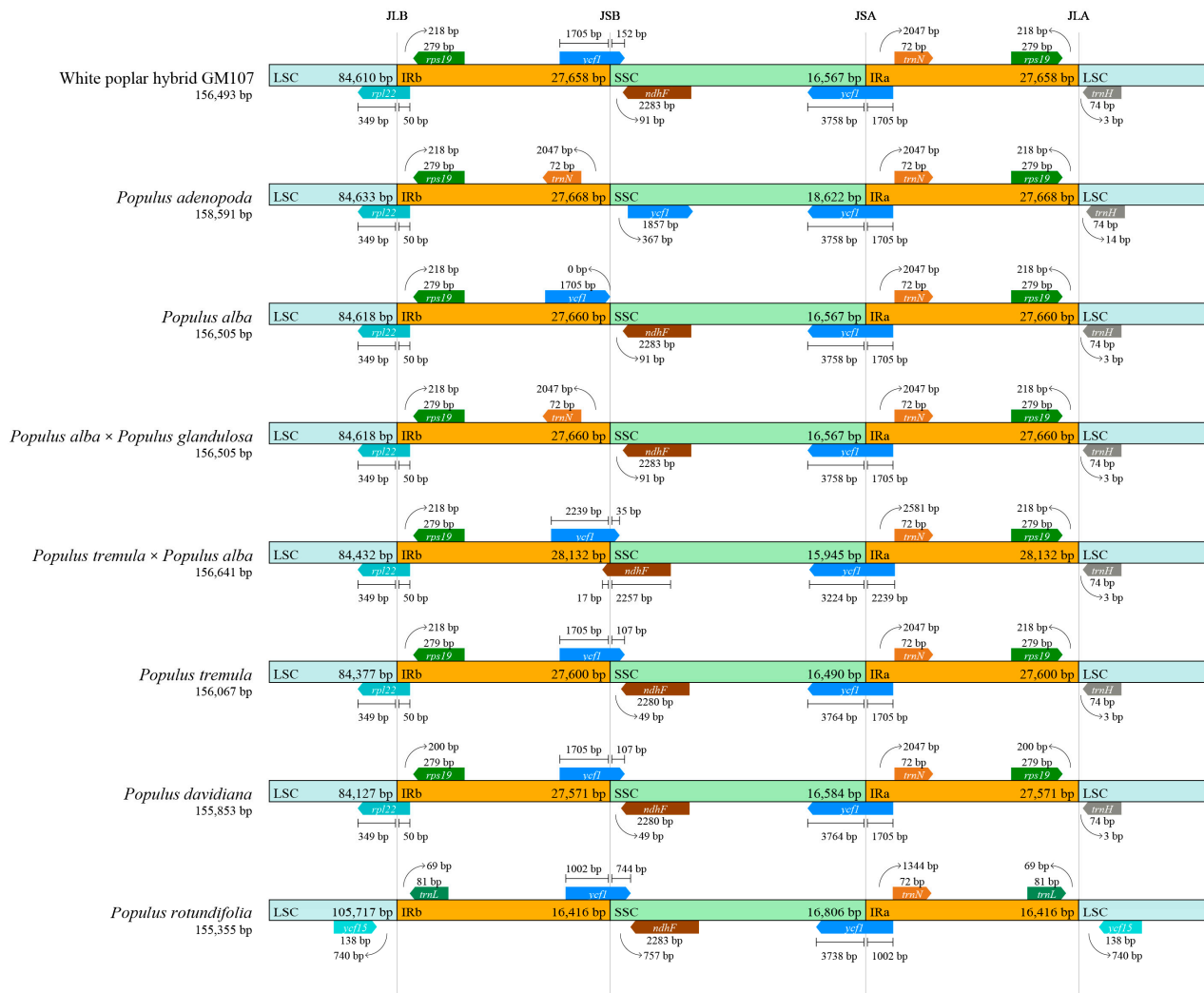


Figure 2. Comparison of LSC, IRs, and SSC region junction among the seven *Populus* cp genomes.

A common 50 bp extension to the IRb was detected in LSC gene *rpl22* of all species except *P. rotundifolia*. The SSC gene *ycf1* had a 1705 bp extension to the IRa regions of the elite individual GM107, *P. adenopoda*, *P. alba*, *P. alba* × *P. glandulosa*, *P. tremula*, and *P. davidiana*. In *P. rotundifolia* and *P. tremula* × *P. alba*, the SSC gene *ycf1* extended 2239 and 1002 bp to the IRa and IRb regions, respectively. The *ycf1* genes in the JSB border region presented some variation, with 35–744 bp extensions to the SSC region in *P. tremula* × *P. alba*, elite individual GM107, *P. tremula*, *P. davidiana*, and *P. rotundifolia*.

3.3. Repeat Structure and SSR Analyses

A comprehensive analysis of 8 chloroplast genomes revealed a total of 374 repeat sequences, with lengths varying from 30 to 76 base pairs. The majority of the repeats were found to be in the range of 30 to 42 base pairs (Table 3), and all repeat types (forward, reverse, palindromic, and complement) were found in each genome (Table 4). Repeats of

each type presented similar features across all eight cp genomes, except for complement repeats. *Populus alba*, *P. alba* × *P. glandulosa*, *P. tremula* × *P. alba*, and the elite individual GM107 had four complement repeats, while *P. rotundifolia*, *P. davidiana*, *P. adenopoda*, and *P. tremula* had two, one, zero, and zero complement repeats, respectively (Table 4). The forward repeats presented the highest proportion, accounting for 47.06% of all repeat types, followed by palindromic (36.10%), reverse (11.76%), and complement repeats (5.08%).

Table 3. Repeat length frequency among eight *Populus* chloroplast genomes.

Species	Number of Repeats																				
	30	31	32	33	34	35	36	37	38	39	40	41	42	44	45	46	47	48	55	68	76
<i>P. adenopoda</i>	13	4	4	1	3	1	1	0	0	6	0	1	6	0	1	0	0	1	1	1	1
<i>P. alba</i>	16	4	6	2	3	1	2	1	0	6	0	1	2	0	0	1	0	0	1	1	1
<i>P. alba</i> × <i>P. glandulosa</i>	16	4	6	2	3	1	2	1	0	6	0	1	2	0	0	1	0	0	1	1	1
<i>P. davidiana</i>	13	5	4	2	0	2	1	0	0	6	0	1	2	1	0	0	0	0	1	1	1
<i>P. rotundifolia</i>	12	8	11	3	0	2	1	1	1	6	2	1	4	1	0	0	1	0	1	1	1
Elite individual GM107	16	4	6	2	3	1	2	1	0	6	0	1	2	0	0	1	0	0	1	1	1
<i>P. tremula</i>	13	3	3	1	1	2	1	1	0	7	0	2	2	0	0	1	0	0	1	1	1
<i>P. tremula</i> × <i>P. alba</i>	16	4	6	2	3	1	2	1	0	6	0	1	2	0	0	1	0	0	1	1	1

Table 4. Repeat type frequency among eight *Populus* chloroplast genomes.

Species	Number of Repeats			
	Forward	Palindromic	Reverse	Complement
<i>P. adenopoda</i>	23	19	3	0
<i>P. alba</i>	22	16	6	4
<i>P. alba</i> × <i>P. glandulosa</i>	22	16	6	4
<i>P. davidiana</i>	21	15	3	1
<i>P. rotundifolia</i>	25	20	10	2
Elite individual GM107	22	16	6	4
<i>P. tremula</i>	19	17	4	0
<i>P. tremula</i> × <i>P. alba</i>	22	16	6	4
Total	176	135	44	19

A total of 1034 simple repeat sites were detected among the 8 *Populus* cp genomes, ranging from 126 to 135 SSRs per species. These repeats are comprised of a diverse set of one to six base pair units. Among these poplar species, the most abundant types of repeats were mononucleotides (101–106), followed by dinucleotides (8–10), trinucleotides (5–8), and tetranucleotides (8–1). Interestingly, only two species, *P. alba* and *P. rotundifolia*, had penta-base repeats, each with a single instance. There were one to two six-base repeats, except in *P. rotundifolia* (Figure 3a). In total, there were 1034 SSRs, among which 76.02% of loci were in intergenic spacer (IGS) regions, and 11.99% of which were distributed in introns and coding sequences (CDSs) (Figure 3b). In 8 species of *Populus*, the most abundant type of repeat unit found was mononucleotide A/T, with concentrations ranging from a high of 81.25% in *P. davidiana* to a low of 78.91% in *P. alba* × *P. glandulosa*. The presence of mononucleotide G repeats was limited to only *P. adenopoda*. Of the two-base pair repeats, AT was more commonly observed than TA. Additionally, the five-base repeat sequence ATAAT was found in *P. alba*, *P. alba* × *P. glandulosa*, and the elite individual GM107. Another five-base repeat identified was TAGAT, which was exclusive to *P. rotundifolia*. There were four six-base repeats: ATAAAA, ATATAG, ATCTAT, and TTTCTA. Among these, ATATAG and ATCTAT appeared in *P. adenopoda*, *P. alba*, *P. alba* × *P. glandulosa*, and the elite individual GM107; ATAAAA appeared in *P. davidiana*, *P. tremula*, and *P. tremula* × *P. alba*; and TTTCTA appeared only in *P. tremula* × *P. alba* (Figure 3a).

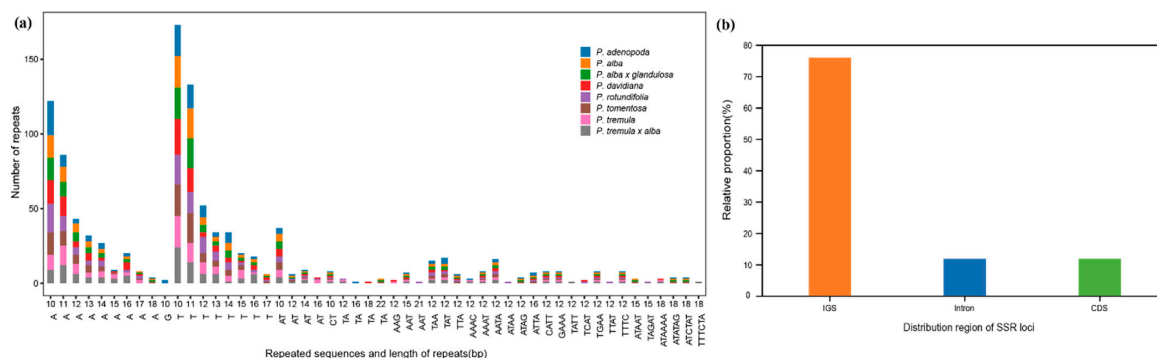


Figure 3. The microsatellites and oligonucleotide repeat in the eight *Populus* cp genomes. (a) Numbers of repeats by length; (b) distribution of microsatellites.

3.4. Sequence Divergence Analyses

To assess the diversity among the chloroplast genome sequences of eight white poplar individuals, we used the mVISTA program to visualize the degree of identity between the sequences and highlight highly conserved regions with color (Figure 4). As depicted in Figure 4, the sequences exhibited high sequence similarity (>90%) and gene order conservancy among poplar species. Moreover, the coding regions were more conserved than the noncoding regions, which is consistent with previous findings [6,31,32]. However, the coding regions of the *ycf1* gene were the most divergent among *Populus* section derivatives, indicating its potential use as a barcode for taxonomic and phylogenetic analysis of *Populus* species. Some highly divergent noncoding regions included: *trnS-trnT*, *petN-psbM*, *trnT-psbD*, *ndhC-trnM*, and *psbE-trnP*.

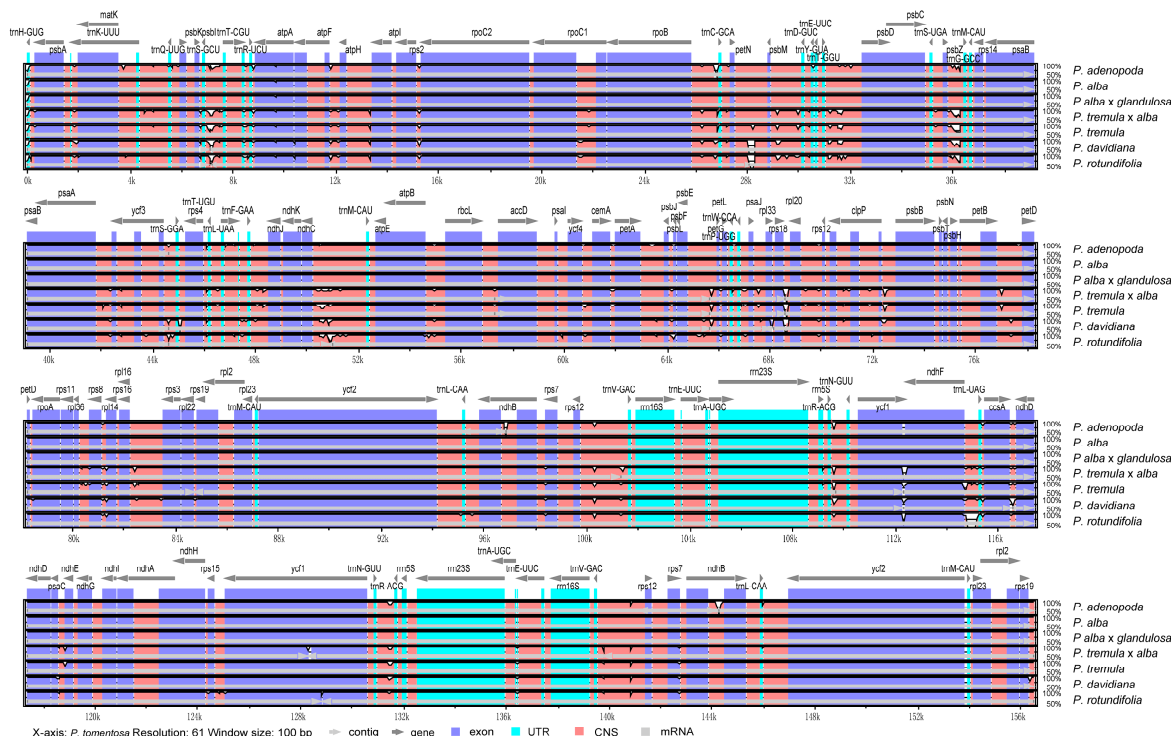


Figure 4. Seven *Populus* cp genomes' comparison through mVISTA program. White peaks indicate differences in genomics. The y-axis indicates the percentage identity (shown: 50–100%). Gray arrows and thick black lines above the alignment indicate gene orientation. Purple bars represent exons, blue bars represent untranslated regions (UTRs), pink bars represent noncoding sequences (CNSs), gray bars represent mRNA.

To gain a comprehensive understanding of sequence deviation, we conducted a sliding window analysis to examine the variation within nucleotides of eight chloroplast genomes. Figure 5 indicates that none of the CDS genes had Pi values above 0.02, and the intergenic sequences exhibited greater deviation than the gene regions. The figure clearly demonstrates that the SSC and LSC regions were more diverse compared to the IR regions. Intergenic sections with a Pi value greater than 1% included *psbI-trnT-CGU* (1.11%), *ndhC* (1.69%), *petA-psbL* (1.11%), and *ycf1-ndhF* (1.58%) (Figure 5).

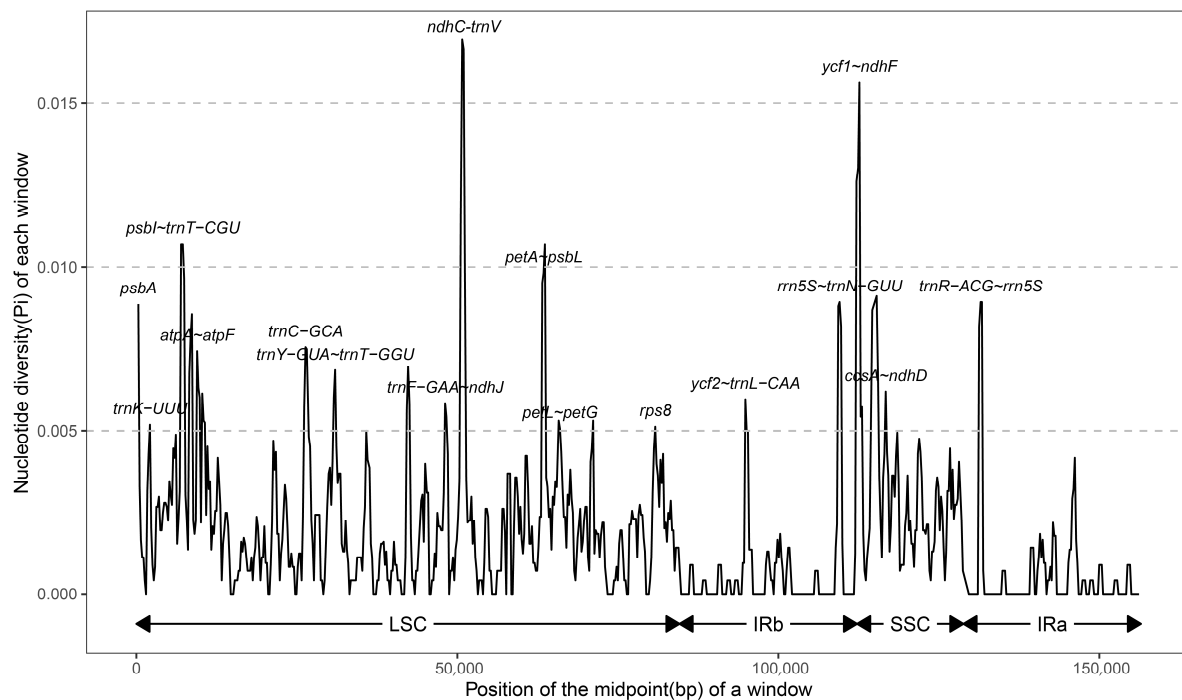


Figure 5. The nucleotide variability (Pi) values of eight *Populus* species.

3.5. Codon Biased Usage Analysis

The entire chloroplast genome of the elite individual GM107 contained 85 protein-coding genes, whose codons fit into 64 codon types. The analysis of amino acid rates exhibited that the top three most abundant amino acids were Leucine (Leu), Arginine (Arg), and Serine (Ser), with an occurrence of 9.4%. These were followed by Alanine (Ala), Glycine (Gly), Proline (Pro), Threonine (Thr), and Valine (Val), which had an occurrence of 6.3%. On the other hand, Tyrosine (Trp) and Methionine (Met) were the least abundant, with an occurrence of only 1.6%.

The calculation of the relative synonymous codon usage frequency was performed on protein-coding sequences derived from the cp genome of the elite plant individual GM107. The relative synonymous codon usage analysis, presented in Figure 6, revealed that 21 codons displayed RSCU values greater than 1, indicating a preference for certain codons. On the other hand, codons encoding Met and Trp had an RSCU value of one, indicating a lack of bias in usage. The Arg codon (AGA) had the highest RSCU value (2.02) among all codons, demonstrating a stronger preference compared to the other 5 Arg codons (AGG, CGA, CGC, CGG, and CGT).

Codon preference is prevalent in both prokaryotic and eukaryotic systems and understanding intraspecific or interspecific gene codon preference during species evolution is of significant importance [33]. The use of codons in gene translation is influenced by long-term natural selection, artificial domestication, and mutations. Highly expressed genes are mainly impacted by natural selection in determining their codon preference. Conversely, the preference of codons in genes with low expression is primarily determined by mutations [34].

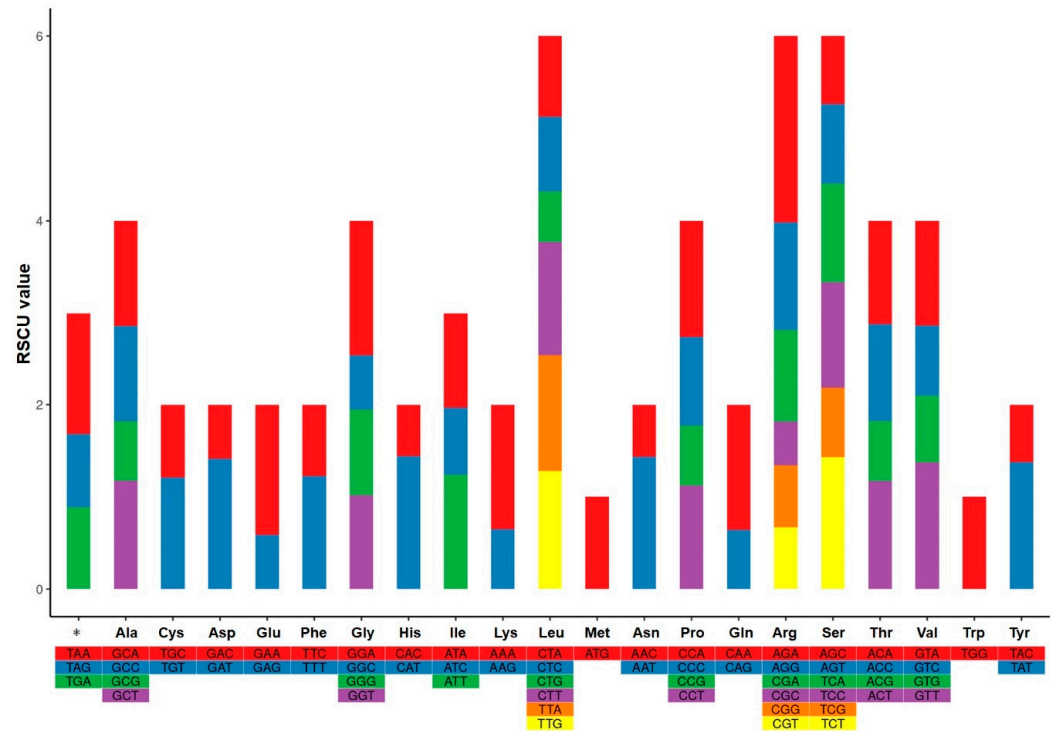


Figure 6. Relative synonymous codon usage and biased codon usage. * Stop code.

3.6. Phylogenetic Analyses

To determine the evolutionary relationship of the exceptional individual GM107 within the white poplar species, we chose seven additional cp genomes and utilized *Salix suchowensis* as a reference point. The maximum likelihood (ML) method was used to construct a phylogenetic tree, which is presented in Figure 7.

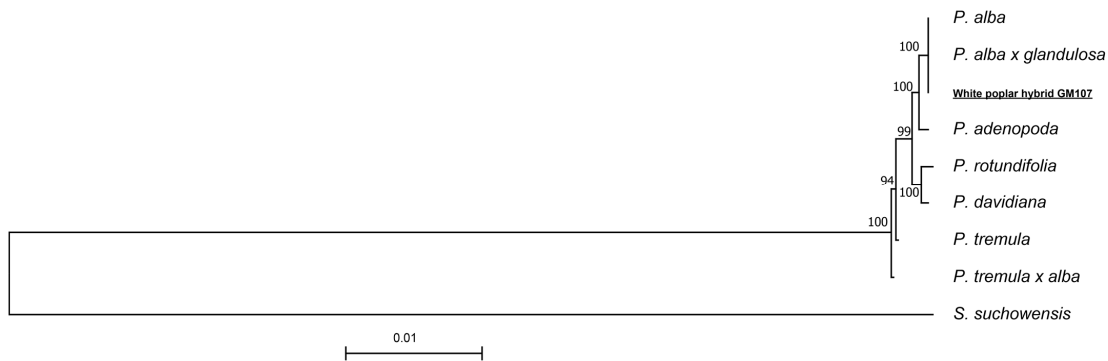


Figure 7. The phylogenetic tree of cp genomes from nine species. The tree was constructed using maximum likelihood. *S. suchowensis* was set as outgroup.

The study indicated that the divergence between *P. tremula* and other *Populus* species happened first, leading to the classification of other white poplars into two evolutionary branches. The first branch includes *P. rotundifolia* and *P. davidiana*, while the second branch contains *P. alba*, *P. adenopoda*, *P. tomentosa*, and *P. alba* × *P. glandulosa*. The results show that the elite individual GM107 is more closely related to *P. alba* and *P. alba* × *P. glandulosa* than to other species. Taking into account their separate geographical distributions, *P. alba* can be considered as the likely maternal ancestor of GM107, excluding the possibility of *P. alba* × *P. glandulosa*.

3.7. Transcription Profiles of cp Genes

To investigate the transcriptome profiles of cp genes, we collected roots, stems, leaves, and petioles from tissue-cultured plantlets of *P. alba* var. *pyramidalis*, *P. tomentosa* GM15, the elite individual GM107, and *P. trichocarpa*, and performed RNA sequencing. The transcripts of 66 genes were detected among the 4 tissue types, among which 44 genes were categorized into a low-expression group and 22 into a high-expression group. The expression levels of almost all 66 genes were higher in leaves than in other tissues, except for *PtatpF* and *PtatpH*. Among these genes, *PtndhF*, *Ptpsba*, *Ptpsbb*, *Ptrps14*, *PtatpF*, and *PtatpH* showed higher transcript abundance in all tissues than all other genes (Figure 8).

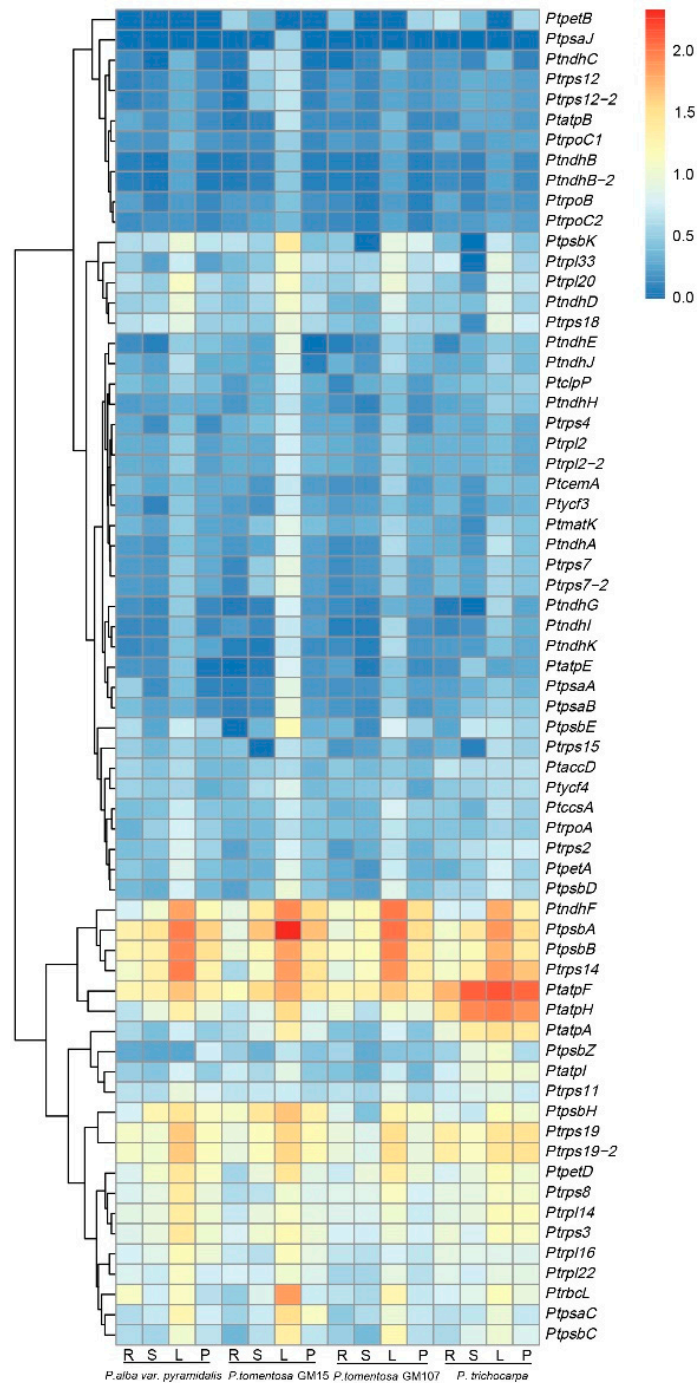


Figure 8. The transcription profiles of cp genes in roots, stems, leaves, and petioles from four poplars. R, S, L, and P represent roots, stems, leaves, and petioles, respectively.

4. Discussion

In this study, we presented the full chloroplast genome sequence of the elite individual GM107 and compared it to those of seven other white poplars. The results indicated that the quadruple structure, gene content, and organization of the cp genome were highly consistent, not only within poplar species [6,29,35], but also among other angiosperms [36,37]. The cp genome of GM107 was 156,493 bp long and encoded 127 genes, slightly different from other white poplars, which typically have a cp genome ranging from 155,355 to 158,591 bp and encode 131 genes, with the exceptions of *P. alba* (129) and *P. davidiana* (130).

Our research revealed minor variations in the positions of the IR and SC boundaries in different *Populus* species. These variations are a result of the expansion or contraction of the chloroplast genome, which is a common phenomenon among plants [37,38]. Compared with conserved boundary regions (except in *P. rotundifolia*) between the LSC and both IRs, the boundary region between the small single copy (SSC) and IRb was more variable among the eight poplar species. The *ycf1* gene, which is located on the SSC/IRa region, was relatively consistent among species, with a 1705 bp extension to IRa in 6 poplar species, and 2239 and 1002 bp extensions to IRa in *P. tremula* × *P. alba* and *P. rotundifolia*, respectively. By contrast, the *ycf1* gene, which is located on the SSC/IRb region, was more variable than that located on the SSC/IRa region (Figure 3). These findings suggest that changes in the IR region, mainly observed in white poplars, are due to variations in the SSC/IRb boundary.

We found 1034 SSRs in the chloroplast genomes of eight species of white poplar. The majority of these SSRs were A/T-rich and had a higher proportion of poly A/T than poly G/C. This aligns with the findings of previous studies [6,39,40]. Thus, SSRs located in the chloroplast genomes could be used as molecular markers for inferring phylogenetic relationships [41,42]. Large-scale overlapping of geographical distributions and easy crossing among white poplar species have led to abundant polymorphisms among section *Populus*. Although traditional phenotypic classification methods are intuitive and simple, they have some difficulty distinguishing hybrids or varieties with similar phenotypes and complex genetic backgrounds. SSRs in the cp genome are a powerful tool for the identification of white poplar species and their polymorphisms. DNA barcodes are effective molecular markers for species identification and genetic relationship analysis among species [43]. Generally, *nrITS*, *trnL-F*, *rbcL*, *matK*, and *trnH/psbA* are recommended for DNA barcoding in plants [44,45]. In this study, we identified 4 noncoding variable regions (*psbI-trnS*, *ndhC-trnV*, *petA-psbJ*, and *ycf1-ndhF*) showing >1% variation, making them potential candidates for use as molecular markers in phylogenetic analyses or DNA barcoding. Almost all species in section *Populus* are capable of interbreeding with at least one other species within the same section. This leads to a complex network of evolutionary change, making it challenging to distinguish between interspecific relationships [46] and identify hybrids and trace parental origins. Section *Populus* cp genomes show conserved quadruple structures and genes, have constant evolutionary rates [47,48], and, importantly, they follow maternal inheritance, which makes them suitable for inferring the phylogenetic status of the entire *Populus* section and tracing the parental origin of white poplar hybrids. In this study, we constructed a phylogenetic tree of the cp genomes of eight species in section *Populus* using ML methods. The results clearly indicated that interspecific relationships among section *Populus* species are highly stable. Notably, *P. alba* was found to be the maternal parent of the elite individual GM107. This finding differs from previously proposed origins for *P. tomentosa* [5,6,49]. There also appears to be variation within *P. tomentosa* with respect to its hybrid origin. In a previous study, *P. alba* was suggested to have played the role of the male parent species, with either *P. adenopoda* or *P. davidiana* as the female (maternal) parent [8]. However, in our recent study, *P. tomentosa* was suggested to have originated from hybridization between *P. adenopoda* (♀) and *P. alba* var. *pyramidalis* (♂) approximately 3.93 Mya [5]. Therefore, *P. tomentosa* may have multiple independent origins, suggesting a more complex evolutionary history.

5. Conclusions

In conclusion, the findings of this study provide valuable insights into the composition and organization of the cp genome of the elite individual GM107, a species that holds significant economic and ecological importance, and is widely found throughout China. We created a well-supported phylogenetic tree utilizing the existing chloroplast genomic information for section *Populus* and determined that the elite individual GM107 and *P. alba* were the most closely related species, followed by *P. alba* × *P. glandulosa*, which suggests that *P. alba* is the female parent of the elite individual GM107. These results will be helpful in elucidating the complex origin of Chinese white poplar and the phylogenetic relationships among poplar species in section *Populus*, and may contribute to future genetic improvements in the wood properties and carbon sink capacities of white poplar genetic resources.

Author Contributions: X.A. conceived the experiment. B.G., Y.L., T.C. and S.L. collected plant materials and conducted DNA and RNA extractions. R.-G.Z. assembled sequences and analyzed the data. K.-H.J. performed transcriptome analysis. X.A., B.G. and T.C. wrote and revised the manuscript. W.U.K. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Major Project of Agricultural Biological Breeding (2022ZD0401503), the National Natural Science Foundation of China (31870652, 31570661).

Data Availability Statement: The complete cp genome sequence of the elite individual GM107 cp sequence was submitted to the NCBI database with the GenBank accession number Pt OP894090.

Conflicts of Interest: The authors declare that they have no conflict of interest.

References

1. Lexer, C.; Fay, M.F.; Joseph, J.A.; Nica, M.S.; Heinze, B. Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): The role of ecology and life history in gene introgression. *Mol. Ecol.* **2005**, *14*, 1045–1057. [[CrossRef](#)] [[PubMed](#)]
2. Dickmann, D. *Poplar Culture in North America*; National Research Council Canada: Ottawa, ON, Canada, 2001.
3. An, X.M.; Wang, D.M.; Wang, Z.L.; Li, B.; Bo, W.H.; Cao, G.L.; Zhang, Z.Y. Isolation of a LEAFY homolog from *Populus tomentosa*: Expression of PtLFY in *P. tomentosa* floral buds and PtLFY-IR-mediated gene silencing in tobacco (*Nicotiana tabacum*). *Plant Cell Rep.* **2011**, *30*, 89–100. [[CrossRef](#)] [[PubMed](#)]
4. Yang, X.; Li, J.; Guo, T.; Guo, B.; Chen, Z.; An, X. Comprehensive analysis of the R2R3-MYB transcription factor gene family in *Populus trichocarpa*. *Ind. Crops Prod.* **2021**, *168*, 113614. [[CrossRef](#)]
5. An, X.; Gao, K.; Chen, Z.; Li, J.; Yang, X.; Yang, X.; Zhou, J.; Guo, T.; Zhao, T.; Huang, S. High quality haplotype-resolved genome assemblies of *Populus tomentosa* Carr., a stabilized interspecific hybrid species widespread in Asia. *Mol. Ecol. Resour.* **2022**, *22*, 786–802. [[CrossRef](#)] [[PubMed](#)]
6. Gao, K.; Li, J.; Khan, W.U.; Zhao, T.; Yang, X.; Yang, X.; Guo, B.; An, X. Comparative genomic and phylogenetic analyses of *Populus* section *Leuce* using complete chloroplast genome sequences. *Tree Genet. Genomes.* **2019**, *15*, 32. [[CrossRef](#)]
7. He, C. *Study on Genetic Diversity and Origin of Populus tomentosa Carr.*; Beijing Forestry University: Beijing, China, 2005.
8. Wang, D.; Wang, Z.; Kang, X.; Zhang, J. Genetic analysis of admixture and hybrid patterns of *Populus hopeiensis* and *P. tomentosa*. *Sci. Rep.* **2019**, *9*, 4821. [[CrossRef](#)]
9. Song, W.; Chen, Z.; He, L.; Feng, Q.; Zhang, H.; Du, G.; Shi, C.; Wang, S. Comparative Chloroplast Genome Analysis of Wax Gourd (*Benincasa hispida*) with Three *Benincaseae* Species, Revealing Evolutionary Dynamic Patterns and Phylogenetic Implications. *Genes* **2022**, *13*, 461. [[CrossRef](#)]
10. Wicke, S.; Schneeweiss, G.M.; dePamphilis, C.W.; Muller, 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](#)]
11. Shetty, S.M.; Md Shah, M.U.; Makale, K.; Mohd-Yusuf, Y.; Khalid, N.; Othman, R.Y. Complete Chloroplast Genome Sequence of Corroborates Structural Heterogeneity of Inverted Repeats in Wild Progenitors of Cultivated Bananas and Plantains. *Plant Genome* **2016**, *9*, 1–14. [[CrossRef](#)]
12. Howe, C.J.; Barbrook, A.C.; Koumandou, V.L.; Nisbet, R.E.; Symington, H.A.; Wightman, T.F. Evolution of the chloroplast genome. *Philos. Trans. R. Soc. B Biol. Sci.* **2003**, *358*, 99–106. [[CrossRef](#)]
13. Lossel, A.G.; Waheed, M.T. Chloroplast-derived vaccines against human diseases: Achievements, challenges and scopes. *Plant Biotechnol. J.* **2011**, *9*, 527–539. [[CrossRef](#)] [[PubMed](#)]
14. Ahmed, I. Chloroplast genome sequencing: Some reflections. *J. Next Gener. Seq. Appl.* **2015**, *2*, 119. [[CrossRef](#)]

15. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. [[CrossRef](#)] [[PubMed](#)]
16. Craciunas, L.; Pickering, O.; Chu, J.; Choudhary, M.; Zurauskiene, J.; Coomarasamy, A. The transcriptomic profile of endometrial receptivity in recurrent miscarriage. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2021**, *261*, 211–216. [[CrossRef](#)] [[PubMed](#)]
17. Sović, I.; Šikić, M.; Wilm, A.; Fenlon, S.N.; Chen, S.; Nagarajan, N. Fast and sensitive mapping of nanopore sequencing reads with GraphMap. *Nat. Commun.* **2016**, *7*, 11307. [[CrossRef](#)]
18. Koren, S.; Walenz, B.P.; Berlin, K.; Miller, J.R.; Bergman, N.H.; Phillippy, A.M. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **2017**, *27*, 722–736. [[CrossRef](#)]
19. Walker, B.J.; Abeel, T.; Shea, T.; Priest, M.; Abouelliel, A.; Sakthikumar, S.; Cuomo, C.A.; Zeng, Q.; Wortman, J.; Young, S.K.; et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **2014**, *9*, e112963. [[CrossRef](#)]
20. 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](#)]
21. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [[CrossRef](#)]
22. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)]
23. 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](#)] [[PubMed](#)]
24. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)] [[PubMed](#)]
25. Thiel, T.; Michalek, W.; Varshney, R.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **2003**, *106*, 411–422. [[CrossRef](#)]
26. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)] [[PubMed](#)]
27. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [[CrossRef](#)] [[PubMed](#)]
28. Hoang, D.T.; Chernomor, O.; von Haeseler, A.; Minh, B.Q.; Vinh, L.S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **2018**, *35*, 518–522. [[CrossRef](#)] [[PubMed](#)]
29. Han, X.M.; Wang, Y.M.; Liu, Y.J. The complete chloroplast genome sequence of *Populus wilsonii* and its phylogenetic analysis. *Mitochondrial DNA B Resour.* **2017**, *2*, 932–933. [[CrossRef](#)] [[PubMed](#)]
30. Kersten, B.; Faivre Rampant, P.; Mader, M.; Le Paslier, M.-C.; Bounon, R.; Berard, A.; Vettori, C.; Schroeder, H.; Leplé, J.-C.; Fladung, M. Genome sequences of *Populus tremula* chloroplast and mitochondrion: Implications for holistic poplar breeding. *PLoS ONE* **2016**, *11*, e0147209. [[CrossRef](#)]
31. Wang, X.; Zhou, T.; Bai, G.; Zhao, Y. Complete chloroplast genome sequence of *Fagopyrum dibotrys*: Genome features, comparative analysis and phylogenetic relationships. *Sci. Rep.* **2018**, *8*, 12379. [[CrossRef](#)]
32. Asaf, S.; Khan, A.L.; Lubna; Khan, A.; Khan, A.; Khan, G.; Lee, I.J.; Al-Harrasi, A. Expanded inverted repeat region with large scale inversion in the first complete plastid genome sequence of *Plantago ovata*. *Sci. Rep.* **2020**, *10*, 3881. [[CrossRef](#)]
33. Ma, Q.P.; Li, C.; Wang, J.; Wang, Y.; Ding, Z.T. Analysis of synonymous codon usage in *FAD7* genes from different plant species. *Genet. Mol. Res.* **2015**, *14*, 1414–1422. [[CrossRef](#)] [[PubMed](#)]
34. Wang, Q.; Mei, C.; Zhen, H.; Zhu, J. Codon preference optimization increases prokaryotic cystatin C expression. *J. Biomed. Biotechnol.* **2012**, 732017. [[CrossRef](#)] [[PubMed](#)]
35. Huang, D.I.; Hefer, C.A.; Kolosova, N.; Douglas, C.J.; Cronk, Q.C.B. Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytol.* **2014**, *204*, 693–703. [[CrossRef](#)]
36. Bausher, M.G.; Singh, N.D.; Lee, S.B.; Jansen, R.K.; Daniell, H. The complete chloroplast genome sequence of *Citrus sinensis* (L.) Osbeck var ‘Ridge Pineapple’: Organization and phylogenetic relationships to other angiosperms. *BMC Plant Biol.* **2006**, *6*, 21. [[CrossRef](#)] [[PubMed](#)]
37. Hansen, D.R.; Dastidar, S.G.; Cai, Z.; Penaflor, C.; Kuehl, J.V.; Boore, J.L.; Jansen, R.K. Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae). *Mol. Phylogenet. Evol.* **2007**, *45*, 547–563. [[CrossRef](#)] [[PubMed](#)]
38. Zeng, S.; Zhou, T.; Han, K.; Yang, Y.; Zhao, J.; Liu, Z.L. The Complete Chloroplast Genome Sequences of Six *Rehmannia* Species. *Genes* **2017**, *8*, 103. [[CrossRef](#)]
39. Martin, G.; Baurens, F.C.; Cardi, C.; Aury, J.M.; D’Hont, A. The complete chloroplast genome of banana (*Musa acuminata*, Zingiberales): Insight into plastid monocotyledon evolution. *PLoS ONE* **2013**, *8*, e67350. [[CrossRef](#)]
40. Melotto-Passarini, D.M.; Tambarussi, E.V.; Dressano, K.; De Martin, V.F.; Carrer, H. Characterization of chloroplast DNA microsatellites from *Saccharum* spp and related species. *Genet. Mol. Res.* **2011**, *10*, 2024–2033. [[CrossRef](#)]

41. Ahmed, I.; Matthews, P.J.; Biggs, P.J.; Naeem, M.; McLenachan, P.A.; Lockhart, P.J. Identification of chloroplast genome loci suitable for high-resolution phylogeographic studies of *Colocasia esculenta* (L.) Schott (Araceae) and closely related taxa. *Mol. Ecol. Resour.* **2013**, *13*, 929–937. [[CrossRef](#)]
42. Cavender-Bares, J.; Gonzalez-Rodriguez, A.; Eaton, D.A.; Hipp, A.A.; Beulke, A.; Manos, P.S. Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Mol. Ecol.* **2015**, *24*, 3668–3687. [[CrossRef](#)]
43. Hebert, P.D.; Ratnasingham, S.; deWaard, J.R. Barcoding animal life: Cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc. Biol. Sci.* **2003**, *270*, 96–99. [[CrossRef](#)] [[PubMed](#)]
44. Hollingsworth, P.M.; Forrest, L.L.; Spouge, J.L.; Little, D.M. A DNA barcode for land plants. *Biol. Sci.* **2009**, *106*, 12794–12797. [[CrossRef](#)]
45. Hollingsworth, P.M.; Graham, S.W.; Little, D.P. Choosing and using a plant DNA barcode. *PLoS ONE* **2011**, *6*, e19254. [[CrossRef](#)] [[PubMed](#)]
46. Eckenwalder, J.E. *Systematics and Evolution of Populus*; National Research Council Canada: Ottawa, ON, Canada, 1996.
47. Mejnartowicz, M. Inheritance of chloroplast DNA in *Populus*. *Theor. Appl. Genet.* **1991**, *82*, 477–480. [[CrossRef](#)]
48. Palmer, J.D. Chloroplast DNA evolution and biosystematic uses of chloroplast DNA variation. *Am. Nat.* **1987**, *130*, 6–29. [[CrossRef](#)]
49. Wang, Z.; Du, S.; Dayanandan, S.; Wang, D.; Zeng, Y.; Zhang, J. Phylogeny reconstruction and hybrid analysis of *populus* (*Salicaceae*) based on nucleotide sequences of multiple single-copy nuclear genes and plastid fragments. *PLoS ONE* **2014**, *9*, e103645. [[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.