Article

Intergeneric Relationships within the Family Salicaceae *s.l.* Based on Plastid Phylogenomics

Meng-Meng Li [1,†], **De-Yan Wang** [1,†], **Lei Zhang** [1], **Ming-Hui Kang** [1], **Zhi-Qiang Lu** [2], **Ren-Bin Zhu** [2], **Xing-Xing Mao** [1], **Zhen-Xiang Xi** [1] and **Tao Ma** [1,*]

1   Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China
2   CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China
*   Correspondence: matao.yz@gmail.com
†   These authors contributed equally to this study.

check for updates

Abstract: Many Salicaceae *s.l.* plants are recognized for their important role in the production of products such as wood, oils, and medicines, and as a model organism in life studies. However, the difference in plastid sequence, phylogenetic relationships, and lineage diversification of the family Salicaceae *s.l.* remain poorly understood. In this study, we compare 24 species representing 18 genera of the family. Simple sequence repeats (SSRs) are considered effective molecular markers for plant species identification and population genetics. Among them, a total of 1798 SSRs were identified, among which mononucleotide repeat was the most common with 1455 accounts representing 80.92% of the total. Most of the SSRs are located in the non-coding region. We also identified five other types of repeats, including 1750 tandems, 434 forward, 407 palindromic, 86 reverse, and 30 complementary repeats. The species in Salicaceae *s.l.* have a conserved plastid genome. Each plastome presented a typical quadripartite structure and varied in size due to the expansion and contraction of the inverted repeat (IR) boundary, lacking major structural variations, but we identified six divergence hotspot regions. We obtained phylogenetic relationships of 18 genera in Salicaceae *s.l.* and the 24 species formed a highly supported lineage. *Casearia* was identified as the basal clade. The divergence time between Salicaceae *s.l.* and the outgroup was estimated as ~93 Mya; *Salix*, and *Populus* diverged around 34 Mya, consistent with the previously reported time. Our research will contribute to a better understanding of the phylogenetic relationships among the members of the Salicaceae *s.l.*

Keywords: Salicaceae; phylogenetic relationship; plastid genome; comparative genomics; repeat sequences

1. Introduction

The previously defined Salicaceae *sensu stricto* (*s.s.*) only included *Salix* and *Populus* [1], but later more than 50 genera were classified into the Salicaceae *sensu lato* (*s.l.*) family, containing over 1000 species [2]. The family Salicaceae *s.l.* is a woody shrub plant, ranging in height from less than a few centimeters to tens of meters. The species in Salicaceae *s.l.* are primarily distributed in cold, tropical, and warm temperate regions and occupy extremely varied habitats [2,3]. This family's sexual systems are highly diverse. Most genera are dioecious, whereas, some are monoecious. Both XY and ZW sex determination systems have been reported in the dioecious species, indicating an amazingly diversified history of sex determination [4,5]. The species of this family have many uses. There are varied chemicals produced by the family. An abundant oil containing unsaturated fatty acids can be synthesized from *Idesia* fruits [6]. The early medicine, aspirin, was first isolated from willow and poplar

bark, and the willow species have been developed as bioenergy crops [7]. *Populus* species have become model organisms for basic research in molecular biology and genetics because of their small genome size and fast growth rates. Moreover, they play important roles in the ecosystem, plant domestication, and conservation, as well as being one of the most economically important groups of forest trees [8].

The plastid genome is widely used in plant genetic population and phylogeny analysis because of its slow rate of nucleotide substitution in gene coding genes and relatively conservative gene structure and content, which can also increase phylogenetic resolution at the lower taxonomic levels [9–14]. The plastid genome usually encodes about 80 unique proteins, 30 transfer RNAs (tRNAs), and four ribosomal RNAs (rRNAs) [15,16]. For terrestrial plants with photosynthesis, the plastid genome is 120–220 kb in size and has the typical feature of two inverted repeats (IRs), 20–28 kb in size, separated by small single copies (SSC) and large single copies (LSC) with sizes of 16–27 kb and 80–90 kb, respectively [17,18]. The size of IR in plastids varies widely for different groups, genus, family or species [13,19,20]. The IR copies recombine themselves in order to maintain or confer stability in the remaining plastome [17,18,21]. With the development of high-throughput sequencing technology in recent years, the number of complete sequenced plastid genomes has increased rapidly [22]. The whole plastid genome of 61 species in Salicaceae *s.l.* has been sequenced and stored in the GenBank.

In previous studies [23–26], their work mainly focused on the relationship of the main subclades in the genera of *Salix* and *Populus*, because the delimitation of species in these genera remains controversial. There are 29–70 species in the genus *Populus* and based on their morphological characteristics they have been grouped into the following six sections: *Abaso*, *Turanga*, *Populus*, *Leucoides*, *Aigeiros*, and *Tacamahaca* [27]. For the genus *Salix*, about 450 species have been published and two main subclades have been identified [28–30]. The most recent study used 42 species from six genera based on the complete plastomes in order to examine phylogenetic relationships of Salicaceae [31]. Although several other genera in Salicaceae were mentioned in their research, their main purpose was to determine the relationship of subclades in the genus of *Salix* and *Populus*. We know little about the phylogenetic relationships of the other 48 genera in Salicaceae *s.l.*

In this study, we sequenced 20 plastid genomes in the family, and added the following four previously published plastid genomes: *Populus euphratica*, *Salix interior*, *Idesia polycarpa maxim*, and *Poliothyrsis sinensis*. In total, there were 24 species from 18 genera of Salicaceae *s.l.*, as well as two outgroups, *Passiflora laurifolia* and *Passiflora ligularis*. We mainly aimed to: (1) determine the repeat sequence variations of plastid genomes, (2) examine structural changes in the plastomes of the Salicaceae *s.l.*, and (3) delimit intergeneric relationships within Salicaceae *s.l.*

2. Results

2.1. Characteristics of the Plastid Genomes

Twenty complete plastid genomes belonging to fourteen genera of the family Salicaceae *s.l.* were newly generated in this study (Table 1). All of the genome sequences have been submitted to the GenBank. In order to fully display the characteristics of the plastid genomes, we further collected four sequences of other species in the Salicaceae *s.l.* from the NCBI GenBank. In total, our subsequent comparative analysis included 24 species representing 18 genera of the family Salicaceae *s.l.* (Table 1). The total length of the chloroplast genome sequences ranged from 155,144 bp in *Flacourtia ramontchii* to 158,605 bp in *Bennettiodendron brevipes*. The structure of the genomes displayed a typical quadripartite structure, with a pair of inverted-repeat (IR) regions of 27,168–27,926 bp, separated by a large single copy (LSC) of 83,391–86,350 bp and a small single copy (SSC) of 16,305–17,889 bp. The LSC regions exhibited the greatest standard deviation in sequence length (s.d. = 872.58 bp), followed by the SSC regions (s.d. = 470.12 bp) and the the IR regions (s.d. = 217.50 bp). Nucleotide composition with an overall GC content of 36.8% was nearly identical in all plastid genomes.

Table 1. Characteristics of plastid genomes of 26 species.

Species	No. of Total Genes	Genome Size (bp)	LSC Length (bp)	SSC Length (bp)	IR Length (bp)	GC Content (%)	Protein-Coding Genes	No. of tRNAs Genes	No. of rRNAs Genes	Genes with Introns	GenBank Accesion Number
Populus euphratica *	131	157,839	85,858	16,649	27,666	36.5	86(8)	37(7)	8(4)	20(5)	MK267314
Salix interior *	127	156,620	85,979	16,305	27,168	37.0	80(5)	37(7)	8(4)	21(5)	NC_024681
Bennettiodendron leprosipes	131	157,872	85,643	16,439	27,895	36.7	86(8)	37(7)	8(4)	21(5)	MK281360
Bennettiodendron brevipes	131	158,605	86,350	16,501	27,877	36.7	86(8)	37(7)	8(4)	21(5)	MK046729
Idesia polycarpa Maxim *	130	157,017	84,787	16,512	27859	36.7	86(8)	36(7)	8(4)	19(5)	NC_032060
Olmediella betschleriana	131	158,171	85,783	16,588	27,900	36.7	85(8)	37(7)	8(4)	20(5)	MK044831
Carrierea calycina	131	158,085	84,663	17,570	27,926	36.8	86(8)	37(7)	8(4)	21(5)	MK263737
Itoa orientalis	131	157,875	84,495	17,742	27,819	36.7	86(8)	37(7)	8(4)	21(5)	MN078146
Poliothyrsis sinensis *	129	156,680	84,793	16,605	27641	36.8	85(8)	37(7)	8(4)	19(5)	NC_037412
Flacourtia inermis	131	155,535	83,706	16,577	27,626	36.9	86(8)	37(7)	8(4)	21(5)	MN078138
Flacourtia ramontchii	131	155,144	83,391	16,591	27,581	36.9	86(8)	37(7)	8(4)	21(5)	MN078145
Flacourtia rukam	131	155,582	83,854	16,598	27,565	36.9	86(8)	37(7)	8(4)	21(5)	MK281365
Xylosma congestum	131	155,986	84,180	16,650	27,578	36.8	86(8)	37(7)	8(4)	21(5)	MN078135
Scolopia chinensis	131	157,292	84,464	17,716	27,556	36.8	86(8)	37(7)	8(4)	21(5)	MN078144
Scolopia saeva	131	155,832	84,062	16,436	27,667	36.9	86(8)	37(7)	8(4)	21(5)	MN078143
Dovyalis caffra	131	155,182	83,593	16,683	27,453	37.0	86(8)	37(7)	8(4)	21(5)	MN078137
Homalium racemosum	131	156,961	85,818	16,547	27,298	36.6	86(8)	37(7)	8(4)	21(5)	MN078136
Homalium cochinchinense	131	155,339	83,701	16,566	27,536	36.8	86(8)	37(7)	8(4)	21(5)	MN078140
Banara guianensis	131	158,577	85,863	17,334	27,690	36.5	86(8)	37(7)	8(4)	21(5)	MH937752
Prockia crucis	131	157,406	85,622	17,024	27,380	36.7	86(8)	37(7)	8(4)	21(5)	MN078147
Azara serrata	131	158,306	85,059	17,889	27,697	36.5	86(8)	37(7)	8(4)	21(5)	MH719101
Abatia parviflora	131	156,426	84,800	16,724	27,451	36.5	86(8)	37(7)	8(4)	21(5)	MN078139
Casearia decandra Jacq	131	156,809	84,890	17,037	27,441	36.8	85(8)	37(7)	8(4)	21(5)	MN078142
Casearia velutina	131	156,008	84,446	17,220	27,171	36.8	85(8)	37(7)	8(4)	21(5)	MN078141
Passiflora laurifolia **	130	151,422	85,411	13,511	26,250	37.0	77(4)	37(7)	8(4)	20(5)	NC_038121
Passiflora ligularis **	128	150,827	85,471	13,502	25,927	37.0	77(4)	36(7)	8(4)	19(5)	NC_038122

Note: Numbers in brackets indicate genes duplicated in the IR regions. The species with * are previous published plastid genomes, and the species with ** are outgroups.

A total of 131 functional genes with the same order were annotated in each of the newly sequenced plastomes, of which 102 were unique genes, including 78 protein-coding genes, 30 tRNA genes, and four rRNA genes (Figure 1, Table 2). Most of these genes occurred as a single copy, while 19 genes were duplicated in the IR regions (Table 1). The gene *cemA* contained premature termination codons in *Casearia decandra Jacq* and *Casearia velutina*, while gene *atpF* contained premature termination codons in *Olmediella betschleriana*.
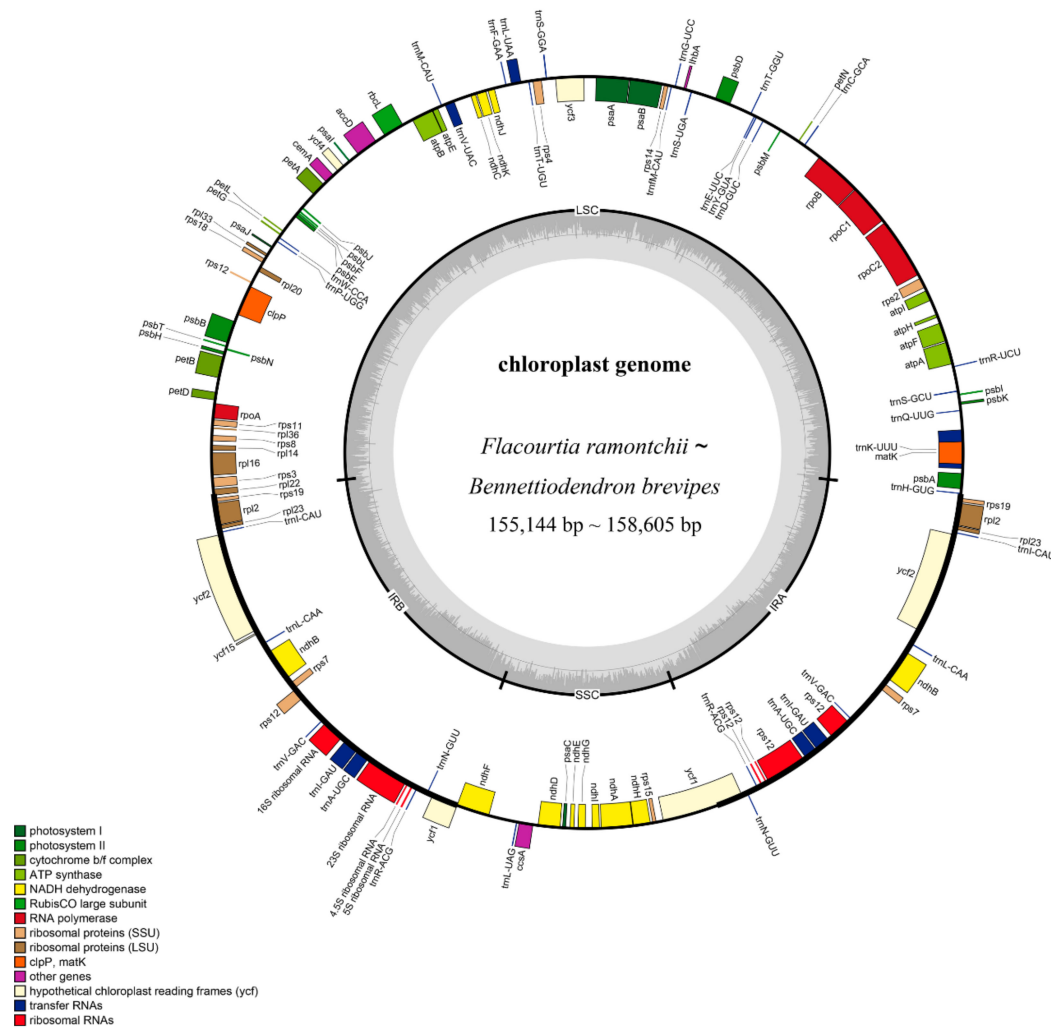
Figure 1. Gene map of 20 Salicaceae *s.l.* chloroplast genomes. Functional categories of genes are color-coded. Genes inside the circle are transcribed clockwise and genes outside the circle are transcribed counter clockwise. The dashed area in the inner circle indicates the GC content of the plastid genome.

2.2. Repeat Sequence Analysis

We identified a total of 1798 simple sequence repeat (SSR) loci across the 24 Salicaceae *s.l.* plastids (Figure 2a, Table S1). Of these, 1455 were mononucleotide repeats accounting for about 80.92% of the total SSRs, while 180 (10.01%), 98 (5.45%), 41 (2.28%), 18 (1.00%), and six (0.33%) were tetra-, di-, tri-, penta-, and hex-nucleotides repeats, respectively. The penta- and hex- nucleotides repeats were very rare across the plastid genomes. Most (74.36%) SSR loci were located in the intergenic regions, whereas, 9.07% were in intron and 16.57% were in the protein-coding regions.

Table 2. List of genes present in the plastid genomes of twenty newly sequenced Salicaceae *s.l.* species.

Category	Gene	Type Gene						
Self-replication	**Ribosomal RNA**	*rrn16*	*rrn23*	*rrn4.5*	*rrn5*			
	Transfer RNA	*trnA-UGC* * *trnC-GCA* *trnD-GUC* *trnE-UUC* *trnF-GAA*	*trnfM-CAU* *trnG-GCC* *trnG-UCC* * *trnH-GUG* *trnI-CAU*	*trnI-GAU* * *trnK-UUU* * *trnL-CAA* *trnL-UAA* * *trnL-UAG*	*trnM-CAU* *trnN-GUU* *trnY-GUA* *trnP-UGG* *trnQ-UUG*	*trnR-ACG* *trnW-CCA* *trnR-UCU* *trnS-GCU* *trnS-GGA*	*trnS-UGA* *trnT-GGU* *trnT-UGU* *trnV-GAC* *trnV-UAC* *	
	Small ribosomal units	*rps11* *rps19*	*rps12* * *rps2*	*rps14* *rps3*	*rps15* *rps4*	*rps18* *rps7*	*rps8*	
	Large ribosomal units	*rpl14* *rpl33*	*rpl16* * *rpl36*	*rpl2* *	*rpl20*	*rpl22*	*rpl23*	
	RNA polymerase subunits	*rpoA*	*rpoB*	*rpoC1* *	*rpoC2*			
Photosynthesis genes	**NADH dehydrogenase**	*ndhA* * *ndhG*	*ndhB* * *ndhH*	*ndhC* *ndhI*	*ndhD* *ndhJ*	*ndhE* *ndhK*	*ndhF*	
	photosystem I	*psaA*	*psaB*	*psaC*	*psaI*	*psaJ*	*ycf3* *	*ycf4*
	photosystem II	*psbA* *psbI* *psbZ*	*psbB* *psbJ*	*psbC* *psbK*	*psbD* *psbL*	*psbE* *psbM*	*psbF* *psbN*	*psbH* *psbT*
	cytochrome b/f complex	*petA*	*petB* *	*petD*	*petG*	*petL*	*petN*	
	ATP synthase	*atpA*	*atpB*	*atpE*	*atpF* *	*atpH*	*atpI*	
	Large subunit of rubisco	*rbcL*						
Other genes	**Maturase**	*matK*						
	Protease	*clpP* *						
	Acetyl-CoA-carboxylase subunit	*accD*						
	Envelope membrane protein	*cemA*						
	Component of TIC complex	*ycf1*						
	c-type cytochrome synthesis	*ccsA*						
Unknown	**hypothetical genes reading frames**	*ycf2*						

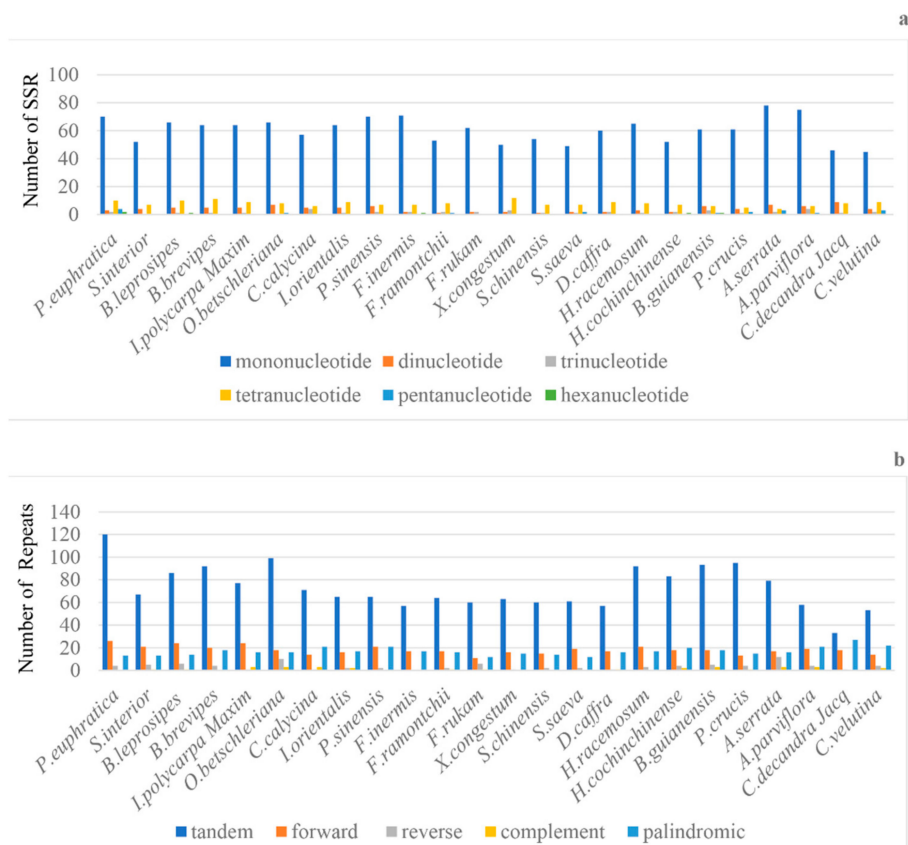Notes: the * symbols indicate genes with intron(s).

Figure 2. Comparison of repeat sequences among 24 plastomes. The *X*-axis represents the species and the *Y*-axis represents the number of repeats. (**a**) Frequency of selected motifs of simple sequence repeats (SSRs) and (**b**) number of each repeat type: tandem repeats, forward repeats, palindromic repeats, reverse repeats, and complement repeats.

We also identified 5 other types of repeats, including 1750 tandems, 434 forward, 407 palindromic, 86 reverse, and 30 complementary repeats (Figure 2b, Table S2, Table S3). Among these, tandem repeats are the most frequent type of repeats (64.65%), while complementary repeats are the least. We found that the tandem repeat sequences were mainly located in non-coding regions, whereas, only a few were located in the coding regions (*ycf2*, *rpl22*, *rpl14*, *rpoC2*, *ycf1*, *ndhD*, *ndhG*, *psbL*, *ndhF*, *rpoA*, *petD*, and *ccsA*). Most of the five types of repeats were concentrated in the intergenic regions (Table S3).

2.3. Inverted Repeats and Genome Comparison

Next, we conducted whole genome alignment using the program mVISTA (*Populus euphratica* as reference), and the results showed that both the content and order of the genes were highly conserved in the Salicaceae *s.l.* plastids (Figure S1). The IRa/SSC boundary positions for all species were located in the *ycf1* gene, while the border genes of IRa and LSC were *Rps19* and *trnH-GUG*, respectively. Only slight variations of the border structure were identified across these plastid genomes. For example, the IRb/LSC junction was located within the *rpl22* gene in all but six species (*Scolopia chinensis*, *Homalium racemosum*, *Homalium cochinchinense*, *Prockia crucis*, *Casearia velutina*, and *Casearia decandra Jacq*). The *ndhF* gene was located entirely in the SSC region of 16 species, while in the other eight species it extended into the IRb region (Figure 3).
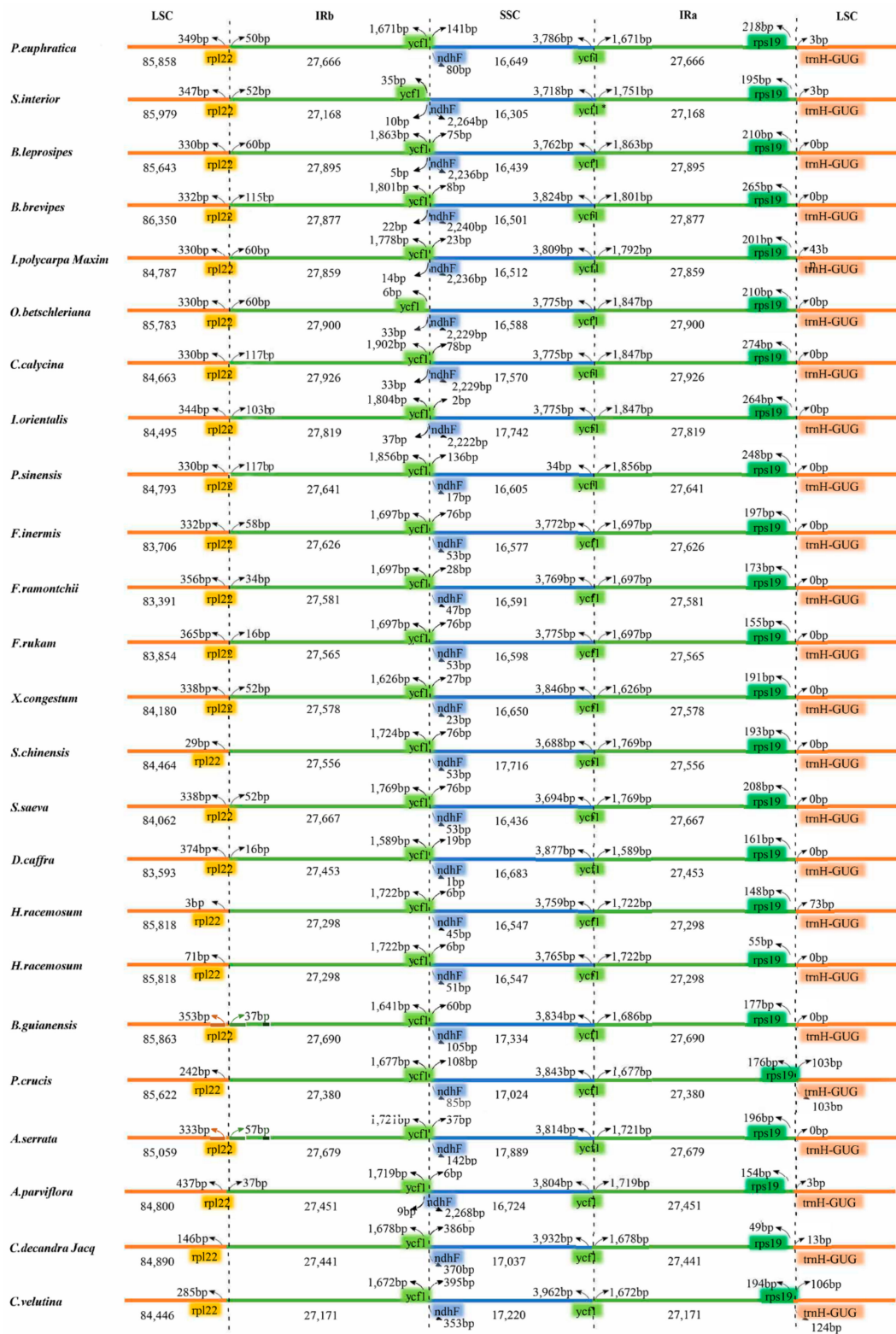
Figure 3. Comparison of the border positions of large single copy (LSC), small single copy (SSC), and inverted repeat (IR) region borders among plastid genomes of 24 Salicaceae *s.l.* species. Complete genes and portions of genes adjacent to the junctions are depicted by differently colored blocks.

2.4. Divergence Hotspots of Plastid Genomes

To evaluate the level of sequence divergence, we calculated the percentages of variation using a sliding window approach (Figure 4). Across the 24 species, the values of nucleotide variability ranged from 0 to 0.139, with an average of 0.023, suggesting a high conservation of plastid genomes within Salicaceae *s.l.* Six relatively high variable regions (divergence hotspots) were identified, which comprised one gene region (*ycf1 + ndhF*) and the following five intergenic regions: *matK-trnQ-UUG*, *trnS-GCU-trnG-UCC*, *psbZ-trnG-GCC*, and *ndhF-trnL-UAG*, *psbE-petL*.
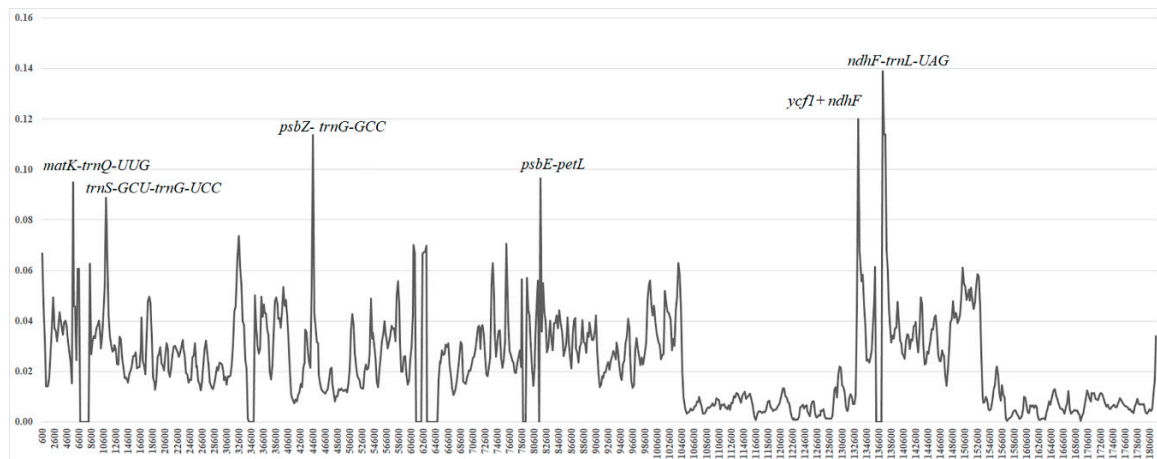
Figure 4. Comparison of the nucleotide variability (Pi) values of the plastomes. *X*-axis: position of the midpoint of a window, *Y*-axis: nucleotide diversity of each window (window length: 600 bp and step size: 200 bp).

2.5. Phylogenetic Analysis

To obtain an accurate phylogenetic relationship of Salicaceae *s.l.* species, we performed multiple sequence alignments of these 24 complete plastid genomes, including an additional two species from *Passiflora* as the outgroup. The final concatenated dataset, which included 63 protein-coding genes (Table S4) and 51,780 nucleotides, after trimming poorly aligned regions, produced a highly supported topology based on the maximum likelihood (ML) strategy (Figure 5). Two subfamilies in the phylogenetic tree, Samydoideae and Salicoideae, formed a highly supported monophyletic. *Casearia* (*Casearia decandra Jacq* and *Casearia velutina*) was the only genus from Samydoideae in this study and was identified as the basal clade. The remaining species belonged to Salicoideae and were divided into two main clades. Furthermore, we estimated divergence times from the plastome dataset using an approximate likelihood method. The Salicaceae *s.l.* was estimated to diverge from the outgroup around 93 Mya.
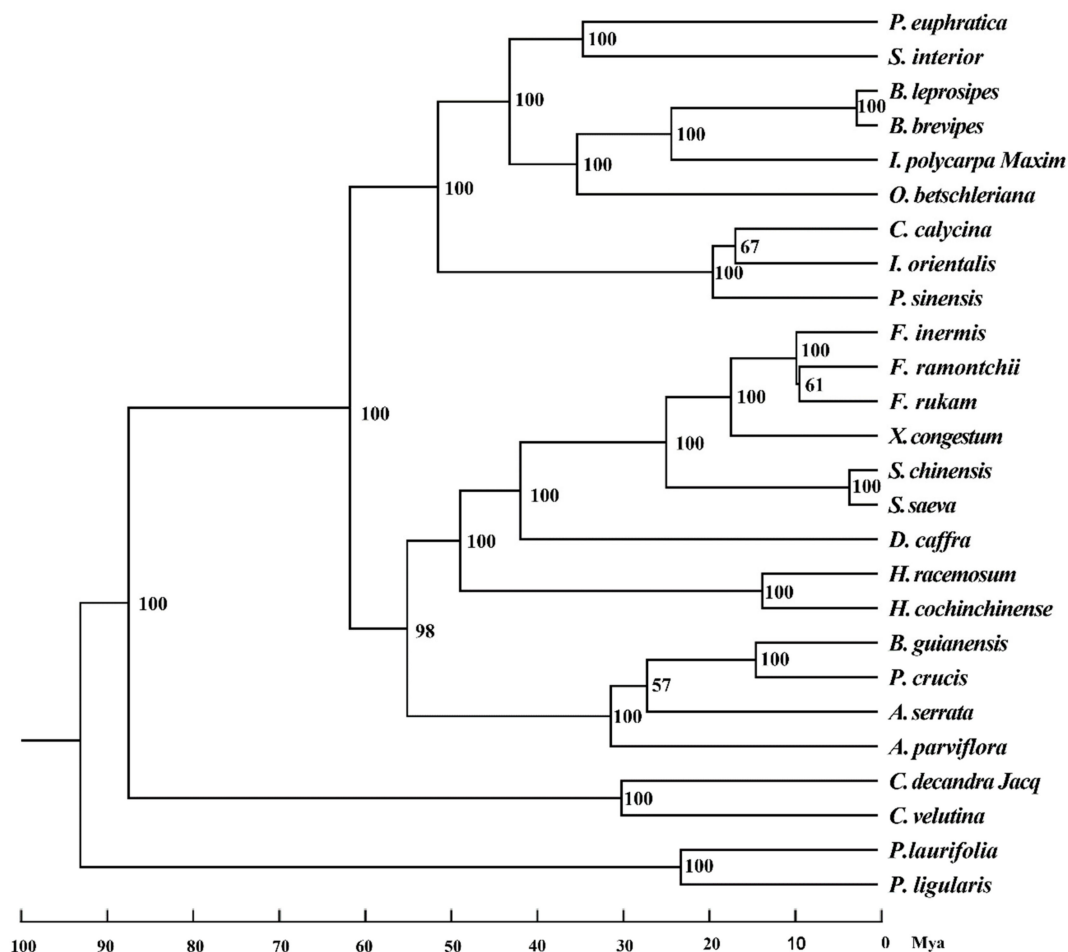
Figure 5. Phylogenetic tree of the 26 species based on 63 protein-coding genes using maximum likelihood (ML). Bootstrap values (1000 replications) between different genera are shown at the nodes. The branches are drawn in proportion to the posterior means of divergence times estimated under the GTR + I + G model with an independent relaxed clock and birth–death sampling, and the outgroup divergence constraint of 87 to 97 Mya.

3. Discussion

3.1. Features of Plastid Genomes

We used the information from complete plastid genome sequencing to research and analyze the plastomes of Salicaceae *s.l.* species. Generally, the plastids of angiosperms are considered to be highly conserved, have a typical tetragonal structure, usually contain about 110–130 unique genes, and the genome size, GC composition, as well as gene and intron content are identical in most angiosperms [32–35]. The shift of IR-SC boundaries was a common evolutionary event and played an important role in plastome size variation [36]. The IR regions of the Salicaceae *s.l.* species varied in size from 27,168 bp to 27,926 bp and the total length ranged from 155,144 bp in *Flacourtia ramontchii* to 158,605 bp in *Bennettiodendron brevipes*, as a result of the expansion and contraction of the IR borders.

Previous studies have revealed that many chloroplast genes, such as *infA*, *rpl22*, *rps19*, *rpl2* intron and *rpl23*, are transferred to the nucleus or lost [37–39]. In our study, only *infA* was not found in the plastid genome of Salicaceae *s.l. cemA* and contained premature termination codons in *Casearia decandra Jacq* and *Casearia velutina*, whereas, *atpF* was a pseudogene in *Olmediella betschleriana*. Gene duplication caused by IR is common in plastomes and is believed to be an important driving force in the evolution of genomes, leading to the creation of new genes and new gene functions [40]. Gene duplication has been previously reported in multiple angiosperm lineages and most of them are tRNAs [41–46].

Duplication of protein-coding genes outside of the IR is rare. In our results, seven tRNA genes and eight protein-coding genes were duplicated in the inverted-repeat region.

3.2. Repeat Sequence Variations

SSRs are a type one to seven nucleotide unit tandem repeat sequence frequently observed in plastid genome and the changes in copy number are usually polymorphic [47–50]. The SSRs are considered effective molecular markers for plant species identification and population genetics because they exhibit codominant inheritance, high repeatability, and high variation within the same species [47–50]. In this study, the mononucleotide repeats were the most common, of which 822 were T type accounting for about 57%, while 606 (41%), 16 (1%), 11 (0.7%), were A, C, and G type, respectively. Most of the SSR loci were located in the non-coding region, and only 16% of the SSR loci were found in the gene-coding region. These SSRs were located in 39 coding genes, and the genes with the highest SSR frequency are *ycf1*, *rpoc2*, and *ndhF*, which are 95, 48, and 26 times, respectively (Table S1). The plastome size variation was previously reported to contribute to tandem [36,51] and dispersed repeats [41,52–54]. We analyzed the direct repeats, inverted repeats, reverse repeats, complementary repeats, and tandem repeats in 24 Salicaceae *s.l.* species. The analysis showed that the number of tandem repeats is more than the other repeats, while complementary repeats are the least common in these species. Most of the repeats are distributed in the intergenic and intron regions, similar to those reported in other angiosperm lineages [55]. These repeat sequences and plastid SSRs provided molecular markers for studying the genetic diversity, population structure, and phylogenetics of Salicaceae *s.l.*

3.3. Comparative Analyses

Comparing chloroplast genomics not only provides insight into chloroplast evolution patterns [56], but also contributes to phylogenetic studies to understand the evolutionary relationships among taxa and apply them to species breeding and conservation [57]. We compared the 24 complete plastid genomes of 18 genera in Salicaceae *s.l.* The species in this family have a conserved plastid genome that lack major structural variations. The main cause of genomic length variation in higher plant plastids is the change in the position of the boundary between IR and SSC/LSC [58]. In this study, the location of the boundary and length of the IR regions showed that variation among the 24 species and the plastid genomes differed slightly in size, with *Bennettiodendron leprosipes*, *Bennettiodendron brevipes*, *Olmediella betschleriana*, *Carrierea calycina*, and *Ito orientalis* larger than other species, which may be related to their IR expansion. IR expansion/contraction also represents a highly variable region, which can be used to study the phylogenetic classification of plants and achieve molecular improvement of plant phenotypes. In addition, the position of the *trnH* gene, which was found in the LSC region of all 24 plastid genomes, was conserved. This result was consistent with the observation that the *trnH* gene is usually located in the IR region in monocots but is found within the LSC region in dicots.

3.4. Divergence Hotspot Analysis

The whole aligned sequences revealed relatively low divergence, however, six (*matK-trnQ-UUG*, *trnS-GCU-trnG-UCC*, *psbZ-trnG-GCC*, *ndhF-trnL-UAG*, *psbE-petL*, and *ycf1+ndhF*) displayed high variation. Further work is still necessary to determine whether these six variable loci could be used in phylogenetic analyses of related Salicaceae *s.l.* species or as potential molecular markers for population genetics and phylogenetics.

3.5. Phylogenetic Relationships

The deep relationships of mimosoids were poorly resolved by phylogenetic studies applying a few plastid markers [59,60]. Plastid phylogenomics has been proven to be efficient to resolve difficult relationships at the family level such as Orchidaceae [61] and the lower taxonomic level such as subfamilies Bambusoideae [62] and Chloridoideae of Poaceae. For the phylogenetic relationship between the species of Salicaceae, earlier studies focused on Salicaceae *s.s.* Preliminary phylogenetic

frameworks for *Salix* and *Populus* have previously been provided [23–26]. In addition, some studies of angiosperm and Malpighiales also involve the phylogenetic relationships of several genera of Salicaceae *s.l.* [63–66]. However, phylogenetic analyses focusing on Salicaceae *s.l.* have been limited to the use of one or a few genes obtained from plastid or nuclear genomes [2,67]. In this study, we obtained phylogenetic relationships of 18 genera in Salicaceae *s.l.* using 63 plastid genes from 26 species. According to the Angiosperm Phylogeny Group's most recent phylogeny, APG IV [2,68], the phylogenetic tree is divided into Salicoideae and Samydoideae. *Casearia* from Samydoideae was identified as the basal clade. The genera in one of the two main clades of Salicoideae included *Populus*, *Salix*, *Bennettiodendron*, *Idesia*, *Olmediella*, *Carrierea*, *Itoa*, and *Poliothyrsis*, all belonging to Saliceae. The other clade was divided into two subclades, one subclade included five genera, *Flacourtia*, *Xylosma*, and *Dovyalis* belonging to the Saliceae, *Scolopia* belonging to Scolopieae, and *Homolium* belonging to Homalieae. The other subclade consisted of four species of three tribes, of which *Banara* and *Prockia* belong to Prockieae, *Abatia* and *Azara* belong to Abatieae and Saliceae, respectively. The Saliceae with the most species is not monophyletic.

We further estimated divergence timescales of the major clades within the Salicaceae *s.l.* according to the calibrations of the species tree constructed on the basis of 63 protein-coding genes. The split between Salicaceae *s.l.* and the outgroup was estimated as ~93 Mya, and the basal *Casearia* was estimated to diverge from other clades around 87 Mya. The clade of Salicoideae is estimated to have originated around 61 Mya. *Salix* and *Populus* diverged around 34 Mya, consistent with the previously reported [26].

4. Materials and Methods

4.1. Sampling and Genome Sequencing

Fresh leaves and silica-gel dried materials were sampled from 20 species representing 14 genera of the family Salicaceae *s.l.* The voucher specimens for the ten fresh sampled plants collected from XiShuangBanNa Tropical Botanical Garden (Mengla, China) were deposited at the Key Laboratory of Bio-Resource and Eco-Environment of the Ministry of Education (Chengdu, Sichuan, China). The ten silica-gel-dried materials were obtained from Harvard University Herbaria and the Arnold Arboretum of Harvard University. For each species, we used the modified CTAB method [69] to extract the total genomic DNA from dry leaves. About 5 ug purified DNA was used to construct Illumina paired-end libraries with an insert size of 500 bp and sequenced using the HiSeq X Ten System by Novogene (Beijing, China). At least two gigabases (Gb) of 2 × 150 bp short reads data were generated for each sample. Reads with a Phred quality score <7 and more than 10% ambiguous nucleotides were filtered.

4.2. Plastid Genome Assembly and Annotation

The filtered reads were assembled by NOVOPlasty v2.7.1 [70], and we used the complete plastid genome *Itoa orientalis* and *Flacourtia indica* (NC_037411.1 and NC_037410.1) as the reference, and then used Geneious v11.1.5 [71] to correct the assemble. Then, we annotated the assembled plastid genome in Plann v1.1 [72]. The positions of starts, stops, introns, and exons were manually adjusted using Sequin v15.50. In addition, the physical map of the plastid genome was generated using OGDRAW v1.2 [73]. The complete plastid genome together with gene annotations was submitted to the GenBank. The accession numbers are shown in Table 1.

4.3. Repeat Sequence Analysis

The SSRs were evaluated using the online software MISA (https://webblast.ipk-gatersleben.de/misa/) [74] with thresholds: 10, 5, 4, 3, 3, 3 ten repeat units for mononucleotide SSRs, five repeat units for dinucleotide SSRs, four repeat units for trinucleotide SSRs and three repeat units for tetra-, penta-, and hexanucleotide SSRs. We used the web-based analysis tool REPuter (https://bibiserv.cebitec.uni-bielefeld.de/reputer) [75] to detect the repeat sequences, including reverse, forward

(direct), complement, and palindromic (inverted), with a minimal repeat size of 30 bp and Hamming distance less than or equal to 3 (90% or greater sequence identity). Tandem Repeats Finder v4.09 (http://tandem.bu.edu/trf/trf.submit.options.html) [76] was used to determine the trandem repeats, alignment parameters match, mismatch. Indel were set as 2, 7, and 7.

4.4. Genome Comparative Analysis

Furter, mVISTA (http://genome.lbl.gov/vista/mvista/submit.shtml) [77], a web-based program, was used to compare similarities and detect any rearrangements or inversions among different plastid genomes. We used this software (*Populus euphratica* as reference) to discover any significant interspecific and intergeneric variations among plastid genome sequences of Salicaceae *s.l.* Additionally, the IR expansion/contraction regions were compared among the 24 Salicaceae *s.l.* species (20 newly generated plastid genomes in this study and 4 previous published plastid genomes: *Populus euphratica*, *Salix interior*, *Idesia polycarpa Maxim*, and *Poliothyrsis sinensis*).

4.5. Sequence Divergence Analysis

After aligning all sequences using Mafft v7.313 [78], we located and counted the SNP sites in the plastid genomes using DnaSP v6.0 [79]. The nucleotide variability (Pi) and polymorphic sites (S) were evaluated using a sliding window analysis with a step size of 200 bp and window length of 600 bp.

4.6. Phylogenetic Analyses

There were 26 plastid genomes used for phylogenetic analyses, including 2 outgroups (*Passiflora laurifolia* and *Passiflora ligularis*) and 24 Salicaceae *s.l.* species (Table 1). We extracted 63 protein-coding genes in all 26 species using a python script. All the genes aligned by MAFFT and used RAxML v8.1.24 [80] to conduct maximum likelihood (ML) analyses with the GTR + I + G model. The best scoring ML tree was obtained with 1000 bootstrap replicates. Then, we estimated divergence times of the plastome dataset, using an approximate likelihood method as implemented in MCMCtree (in PAML version 4) [81], with an independent relaxed clock and birth–death sampling [82]. The divergence between Salicaceae *s.l.* species and the outgroup, *Passiflora*, was assigned an age constraint of 87 to 97 Mya, as has been previously used [83–86].

5. Conclusions

We compared 24 plastomes, 20 newly sequenced and four other species sequences were collected from the NCBI GenBank, all representing 18 genera of the family Salicaceae *s.l.* The SSRs are considered effective molecular markers for plant species identification and population genetics. A total of 1798 SSRs were identified. Mono-nucleotide repeats were the most common with 1455 repeats accounting for about 80.92% of the total. Most of repeats were located in the non-coding region. We also identified five other types of repeats, including 1750 tandems, 434 forward, 407 palindromic, 86 reverse, and 30 complementary repeats. The species in Salicaceae *s.l.* have a conserved plastid genome. Each plastome presented a typical quadripartite structure and varied in size due to the expansion and contraction of the IR boundary, lacking major structural variations, but we identified six divergence hotspot regions (*matK-trnQ-UUG*, *trnS-GCU-trnG-UCC*, *psbZ-trnG-GCC*, *ndhF-trnL-UAG*, *psbE-petL*, and *ycf1+ndhF*). We obtained phylogenetic relationships in 18 genera of Salicaceae *s.l.* and the 26 species formed a highly supported lineage. *Casearia* was identified as the basal clade. The divergence time between Salicaceae *s.l.* and the outgroup was estimated as ~93 Mya, and *Salix* and *Populus* diverged around 34 Mya, consistent with previously reported times. This study demonstrates the potential of plastid genome to address the genus relationship of Salicaceae *s.l.* This data will contribute to further understanding of the phylogenetic relationships among Salicaceae *s.l.*

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/20/15/3788/s1.

References

1. Ohashi, H. Salicaceae of Japan. *Sci. Rep.* **2001**, *40*, 269–396.
2. Chase, M.W.; Zmarzty, S.; Dolores Lledó, M.; Wurdack, K.J.; Fay, S.M.F. When in Doubt, Put It in Flacourtiaceae: A Molecular Phylogenetic Analysis Based on Plastid rbcL DNA Sequences. *Kew Bull.* **2002**, *57*, 141–181. [CrossRef]
3. Leskinen, E.; Alström-Rapaport, C. Molecular phylogeny of Salicaceae and closely related Flacourtiaceae: Evidence from 5.8 S, ITS 1 and ITS 2 of the rDNA. *Plant Syst. Evol.* **1999**, *215*, 209–227. [CrossRef]
4. Hou, J.; Ye, N.; Zhang, D.; Chen, Y.; Fang, L.; Dai, X.; Yin, T. Different autosomes evolved into sex chromosomes in the sister genera of Salix and Populus. *Sci. Rep.* **2015**, *5*, 9076. [CrossRef] [PubMed]
5. Kersten, B.; Pakull, B.; Groppe, K.; Lueneburg, J.; Fladung, M. The sex-linked region in Populus tremuloides Turesson 141 corresponds to a pericentromeric region of about two million base pairs on P. trichocarpa chromosome 19. *Plant Biol.* **2014**, *16*, 411–418. [CrossRef]
6. Li, R.J.; Gao, X.; Li, L.M.; Liu, X.L.; Wang, Z.Y.; Lü, S.Y. De novo Assembly and Characterization of the Fruit Transcriptome of Idesia polycarpa Reveals Candidate Genes for Lipid Biosynthesis. *Front. Plant Sci.* **2016**, *7*, 801. [CrossRef] [PubMed]
7. Abrahamson, L.P.; Robison, D.J.; Volk, T.A.; White, E.H.; Neuhauser, E.F.; Benjamin, W.H.; Peterson, J.M. Sustainability and environmental issues associated with willow bioenergy development in New York (U.S.A.). *Biomass Bioenergy* **1998**, *15*, 17–22.
8. Braatne, J.H.; Hinckly, T.M.; Stettler, R.F. Influence of soil water on the physiological and morphological components of plant water balance in Populus trichocarpa, Populus deltoides and their F1 hybrids. *Tree Physiol.* **1992**, *11*, 325–340. [CrossRef]
9. Besnard, G.; Hernández, P.; Khadari, B.; Dorado, G.; Savolainen, V. Genomic profiling of plastid DNA variation in the Mediterranean olive tree. *BMC Plant Biol.* **2011**, *11*, 80. [CrossRef]
10. Dexter, K.G.; Terborgh, J.W.; Cunningham, C.W. Historical effects on beta diversity and community assembly in Amazonian trees. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7787–7792. [CrossRef]
11. Howe, C.J.; Barbrook, A.C.; Lila Koumandou, V.; Nisbet, R.E.; Symington, H.A.; Wightman, T.F. Evolution of the chloroplast genome. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **2003**, *358*, 99–107. [CrossRef]
12. Liu, L.X.; Li, R.; Worth, J.R.P.; Li, X.; Li, P.; Cameron, K.M.; Fu, C.X. The Complete Chloroplast Genome of Chinese Bayberry (Morella rubra, Myricaceae): Implications for Understanding the Evolution of Fagales. *Front. Plant Sci.* **2017**, *8*, 968. [CrossRef]
13. Wicke, S.; Schneeweiss, G.M.; Depamphilis, C.W.; Kai, F.M.; 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] [PubMed]
14. Khan, A.; Khan, I.A.; Heinze, B.; Azim, M.K. The Chloroplast Genome Sequence of Date Palm (Phoenix dactylifera L. cv. 'Aseel'). *Plant Mol. Biol. Rep.* **2012**, *30*, 666–678. [CrossRef]
15. Bendich, A.J. Circular chloroplast chromosomes: The grand illusion. *Plant Cell* **2004**, *16*, 1661–1666. [CrossRef]
16. Chumley, T.W.; Palmer, J.D.; Mower, J.P.; Fourcade, H.M.; Calie, P.J.; Boore, J.L.; Jansen, R.K. The complete chloroplast genome sequence of Pelargonium x hortorum: Organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol. Biol. Evol.* **2006**, *23*, 2175–2190. [CrossRef] [PubMed]
17. Knox, E.B. The dynamic history of plastid genomes in the Campanulaceae *sensu lato* is unique among angiosperms. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11097–11102. [CrossRef]
18. Palmer, J.D. Chloroplast DNA exists in two orientations. *Nature* **1983**, *301*, 92–93. [CrossRef]

19. Gurdon, C.; Maliga, P. Two Distinct Plastid Genome Configurations and Unprecedented Intraspecies Length Variation in the accD Coding Region in Medicago truncatula. *DNA Res.* **2014**, *21*, 417–427. [CrossRef]

20. Vieira, L.D.N.; Helisson, F.; Marcelo, R.; Fraga, H.P.D.F.; Cardoso, R.L.A.; Emanuel Maltempi, D.S.; Fábio, D.O.P.; Rubens Onofre, N.; Miguel Pedro, G. The complete chloroplast genome sequence of Podocarpus lambertii: Genome structure, evolutionary aspects, gene content and SSR detection. *PLoS ONE* **2014**, *9*, e90618. [CrossRef]

21. Stein, D.B.; Palmer, J.D.; Thompson, W.F. Structural evolution and flip-flop recombination of chloroplast DNA in the fern genus Osmunda. *Curr. Genet.* **1986**, *10*, 835–841. [CrossRef]

22. Tangphatsornruang, S.; Sangsrakru, D.; Chanprasert, J.; Uthaipaisanwong, P.; Yoocha, T.; Jomchai, N.; Tragoonrung, S. The Chloroplast Genome Sequence of Mungbean (Vigna radiata) Determined by High-throughput Pyrosequencing: Structural Organization and Phylogenetic Relationships. *DNA Res.* **2009**, *17*, 11–22. [CrossRef] [PubMed]

23. Huang, Y.; Jun, W.; Yongping, Y.; Chuanzhu, F.; Jiahui, C. Phylogenomic Analysis and Dynamic Evolution of Chloroplast Genomes in Salicaceae. *Front. Plant Sci.* **2017**, *8*, 1050. [CrossRef] [PubMed]

24. Zong, D.; Gan, P.; Zhou, A.; Zhang, Y.; Zou, X.; Duan, A.; Song, Y.; He, C. Plastome Sequences Help to Resolve Deep-Level Relationships of Populus in the Family Salicaceae. *Front. Plant Sci.* **2019**, *10*, 5. [CrossRef] [PubMed]

25. Zong, D.; Zhou, A.; Zhang, Y.; Zou, X.; Li, D.; Duan, A.; He, C. Characterization of the complete chloroplast genomes of five Populus species from the western Sichuan plateau, southwest China: Comparative and phylogenetic analyses. *PeerJ* **2019**, *7*, e6386. [CrossRef] [PubMed]

26. Huang, D.I.; Hefer, C.A.; Kolosova, N.; Douglas, C.J.; Cronk, Q.C. 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] [PubMed]

27. Eckenwalder, E.J. Taxonomic Signal and Noise in Multivariate Interpopulational Relationships in Populus mexicana (Salicaceae). *Syst. Bot.* **1996**, *21*, 261–271. [CrossRef]

28. Chen, J.H.; Sun, H.; Wen, J.; Yang, Y.P. Molecular Phylogeny of Salix L. (Salicaceae) Inferred from Three Chloroplast Datasets and Its Systematic Implications. *Taxon* **2010**, *59*, 29–37. [CrossRef]

29. Lauron-Moreau, A.; Pitre, F.E.; Argus, G.W.; Labrecque, M.; Brouillet, L. Phylogenetic relationships of American willows (Salix L., Salicaceae). *PLoS ONE* **2015**, *10*, e0121965. [CrossRef]

30. Wu, J.; Nyman, T.; Wang, D.C.; Argus, G.W.; Yang, Y.P.; Chen, J.H. Phylogeny of Salix subgenus Salix s.l. (Salicaceae): Delimitation, biogeography, and reticulate evolution. *BMC Evol. Biol.* **2015**, *15*, 31. [CrossRef]

31. Lei, Z.; Xi, Z.X.; Wang, M.C.; Guo, X.Y.; Ma, T. Plastome phylogeny and lineage diversification of Salicaceae with focus on poplars and willows. *Ecol. Evol.* **2018**, *8*, 7817–7823.

32. Palmer, J.D. Plastid Chromosomes; Structure and Evolution. In *The Molecular Biology of Plastids*; Bogorad, L., Vasil, I.K., Eds.; Academic Press: San Diego, CA, USA, 1911; pp. 5–53.

33. Jansen, R.K.; Raubeson, L.A.; Boore, J.L.; Depamphilis, C.W.; Chumley, T.W.; Haberle, R.C.; Wyman, S.K.; Alverson, A.J.; Peery, R.; Herman, S.J. Methods for Obtaining and Analyzing Whole Chloroplast Genome Sequences. *Methods Enzym.* **2005**, *395*, 348–383.

34. Kim, K.J.; Jansen, R.K. ndhF sequence evolution and the major clades in the sunflower family. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10379–10383. [CrossRef]

35. Henry, R.J.; Henry, R.J. Plant diversity and evolution: Genotypic and phenotypic variation in higher plants. *Biologist* **2005**, 121.

36. Dugas, D.V.; Hernandez, D.; Koenen, E.J.M.; Schwarz, E.; Straub, S.; Hughes, C.E.; Jansen, R.K.; Nageswara-Rao, M.; Staats, M.; Trujillo, J.T. Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions, and accelerated rate of evolution inclpP. *Sci. Rep.* **2015**, *5*, 16958. [CrossRef]

37. Millen, R.S.; Olmstead, R.G.; Adams, K.L.; Palmer, J.D.; Lao, N.T.; Heggie, L.; Kavanagh, T.A.; Hibberd, J.M.; Gray, J.C.; Morden, C.W. Many parallel losses of infA from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. *Plant Cell* **2001**, *13*, 645–658. [CrossRef]

38. Thomas, F.; Massenet, O.; Dorne, A.M.; Briat, J.F.; Mache, R. Expression of the rpl23, rpl2 and rps19 genes in spinach chloroplasts. *Nucleic Acids Res.* **1988**, *16*, 2461–2472. [CrossRef]

39. Downie, S.R.; Olmstead, R.G.; Zurawski, G.; Soltis, D.E.; Palmer, J.D. Six independent loss of the chloroplast DNA rpl2 intron in Dicotyledons: Molecular and phylogenetic implications. *Evolution* **1991**, *45*, 1245–1259.

40. Xiong, A.S.; Peng, R.H.; Zhuang, J.; Gao, F.; Zhu, B.; Fu, X.Y.; Xue, Y.; Jin, X.F.; Tian, Y.S.; Zhao, W. Gene duplication, transfer, and evolution in the chloroplast genome. *Biotechnol. Adv.* **2009**, *27*, 340–347. [CrossRef]

41. Haberle, R.C.; Fourcade HMBoore, J.L.; Jansen, R.K. Extensive rearrangements in the chloroplast genome of Trachelium caeruleum are associated with repeats and tRNA genes. *J. Mol. Evol.* **2008**, *66*, 350–361. [CrossRef]

42. Hae-Lim, L.; Jansen, R.K.; Chumley, T.W.; Ki-Joong, K. Gene relocations within chloroplast genomes of Jasminum and Menodora (Oleaceae) are due to multiple, overlapping inversions. *Mol. Biol. Evol.* **2007**, *24*, 1161–1180.

43. Hipkins, V.D.; Marshall KANeale, D.B.; Rottmann, W.H.; Strauss, S.H. A mutation hotspot in the chloroplast genome of a conifer (Douglas-fir: Pseudotsuga) is caused by variability in the number of direct repeats derived from a partially duplicated tRNA gene. *Curr. Genet.* **1995**, *27*, 572–579. [CrossRef]

44. Lenka, D.; Jan, K.; Cestmír, V.; Václav, P. TrnL- trnF intergenic spacer and trnL intron define major clades within Luzula and Juncus (Juncaceae): Importance of structural mutations. *J. Mol. Evol.* **2004**, *59*, 1–10.

45. Lin, C.P.; Wu, C.S.; Huang, Y.Y.; Chaw, S.M. The complete chloroplast genome of Ginkgo biloba reveals the mechanism of inverted repeat contraction. *Genome Biol. Evol.* **2012**, *4*, 374–381. [CrossRef]

46. Vijverberg, K.; Bachmann, K. Molecular evolution of a tandemly repeated trnF (GAA) gene in the chloroplast genomes of Microseris (Asteraceae) and the use of structural mutations in phylogenetic analyses. *Mol. Biol. Evol.* **1999**, *16*, 1329–1340. [CrossRef]

47. Jiao, Y.; Jia, H.M.; Li, X.W.; Chai, M.L.; Jia, H.J.; Chen, Z.; Wang, G.Y.; Chai, C.Y.; Weg, E.V.D.; Gao, Z.S. Development of simple sequence repeat (SSR) markers from a genome survey of Chinese bayberry (Myrica rubra). *BMC Genom.* **2012**, *13*, 201. [CrossRef]

48. Kaundun, S.S.; Matsumoto, S. Heterologous nuclear and chloroplast microsatellite amplification and variation in tea, Camellia sinensis. *Genome* **2002**, *45*, 1041–1048. [CrossRef]

49. Powell, W.; Morgante, M.; Mcdevitt, R.; Vendramin, G.G.; Rafalski, J.A. Polymorphic simple sequence repeat regions in chloroplast genomes: Applications to the population genetics of pines. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 7759–7763. [CrossRef]

50. Pauwels, M.V.X.; Godé, C.; Frérot, H.; Castric, V.; Saumitou-Laprade, P. Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, Arabidopsis halleri (Brassicaceae). *New Phytol.* **2012**, *193*, 916–928. [CrossRef]

51. Yeong Deuk, J.; Jongsun, P.; Jungeun, K.; Wonho, S.; Cheol-Goo, H.; Yong-Hwan, L.; Byoung-Cheorl, K. Complete sequencing and comparative analyses of the pepper (Capsicum annuum L.) plastome revealed high frequency of tandem repeats and large insertion/deletions on pepper plastome. *Plant Cell Rep.* **2011**, *30*, 217–229.

52. Cosner, M.E.; Jansen, R.K.; Palmer, J.D.; Downie, S.R. The highly rearranged chloroplast genome of Trachelium caeruleum (Campanulaceae): Multiple inversions, inverted repeat expansion and contraction, transposition, insertions/deletions, and several repeat families. *Curr. Genet.* **1997**, *31*, 419–429. [CrossRef]

53. Mao-Lun, W.; Blazier, J.C.; Madhumita, G.; Jansen, R.K. Reconstruction of the ancestral plastid genome in Geraniaceae reveals a correlation between genome rearrangements, repeats, and nucleotide substitution rates. *Mol. Biol. Evol.* **2014**, *31*, 645–659.

54. Ogihara, Y.; Terachi, T.; Sasakuma, T. Intramolecular recombination of chloroplast genome mediated by short direct-repeat sequences in wheat species. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 8573–8577. [CrossRef]

55. Yang, Y.; Zhou, T.; Duan, D.; Yang, J.; Feng, L.; Zhao, G. Comparative Analysis of the Complete Chloroplast Genomes of Five Quercus Species. *Front. Plant Sci.* **2016**, *7*, 959. [CrossRef]

56. Daniell, H.; Lin, C.S.; Ming, Y.; Chang, W.J. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. *Genome Biol.* **2016**, *17*, 134. [CrossRef]

57. Wu, F.H.; Chan, M.T.; Liao, D.C.; Hsu, C.T.; Lee, Y.W.; Daniell, H.; Duvall, M.R.; Lin, C.S. Complete chloroplast genome ofOncidiumGower Ramsey and evaluation of molecular markers for identification and breeding in Oncidiinae. *BMC Plant Biol.* **2010**, *10*, 68. [CrossRef]

58. Ki-Joong, K.; Hae-Lim, L. Widespread occurrence of small inversions in the chloroplast genomes of land plants. *Mol. Cells* **2005**, *19*, 104–113.

59. Bouchenak-Khelladi, Y.; Maurin, O.; Hurter, J.; van der Bank, M. The evolutionary history and biogeography of Mimosoideae (Leguminosae): An emphasis on African acacias. *Mol. Phylogen. Evol.* **2010**, *57*, 495–508. [CrossRef]

60. Miller, J.T.; Seigler, D. Evolutionary and taxonomic relationships of *Acacia s.l.* (Leguminosae: Mimosoideae). *Aust. Syst. Bot.* **2012**, *25*, 217–224. [CrossRef]

61. Givnish, T.J.; Spalink, D.; Ames, M.; Lyon, S.P.; Hunter, S.J.; Zuluaga, A.; Iles, W.J.D.; Clements, M.A.; Arroyo, M.T.K.; Leebens-Mack, J. Orchid phylogenomics and multiple drivers of their extraordinary diversification. *Proc. Biol. Sci.* **2015**, *282*, 2108–2111. [CrossRef]

62. Zhang, Y.J.; Ma, P.F.; Li, D.Z. High-throughput sequencing of six bamboo chloroplast genomes: Phylogenetic implications for temperate woody bamboos (Poaceae: Bambusoideae). *PLoS ONE* **2011**, *6*, e20596. [CrossRef]

63. Barba-Montoya, J.; Dos Reis, M.; Schneider, H.; Donoghue, P.C.J.; Yang, Z. Constraining uncertainty in the timescale of angiosperm evolution and the veracity of a Cretaceous Terrestrial Revolution. *New Phytol.* **2018**, *218*, 819–834. [CrossRef]

64. Xi, Z.; Ruhfel, B.R.; Schaefer, H.; Amorim, A.M.; Sugumaran, M.; Wurdack, K.J.; Endress, P.K.; Matthews, M.L.; Stevens, P.F.; Mathews, S. Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation Malpighiales. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17519–17524. [CrossRef]

65. Wurdack, K.J.; Davis, C.C. Malpighiales phylogenetics: Gaining ground on one of the most recalcitrant clades in the angiosperm tree of life. *Am. J. Bot.* **2009**, *96*, 1551–1570. [CrossRef]

66. Cai, L.; Xi, Z.; Amorim, A.M.; Sugumaran, M.; Rest, J.S.; Liang, L.; Davis, C.C. Widespread ancient whole genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol.* **2017**, *221*, 565–576. [CrossRef]

67. Diffey, J.M. Phylogenetic Relationships of Salicaceae Based on Analyses of Nuclear DNA Data. *Honor. Theses* **2017**, *5*, 21.

68. Group, T.A.P.; Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Byng, J.W.; Judd, W.S.; Soltis, D.E.; Mabberley, D.J.; Sennikov, A.N.; Soltis, P.S.; et al. An update of the angiosperm phylogeny group classification for the orders and families of flowering plants: APG IV. *Bot. J. Linn. Soc.* **2016**, *181*, 1–21.

69. Allen, G.C.; Flores-Vergara, M.A.; Krasynanski, S.; Kumar, S.; Thompson, W.F. A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* **2006**, *1*, 2320–2325. [CrossRef]

70. Dierckxsens, N.; Mardulyn, P.; Smits, G. NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **2017**, *45*, e18.

71. Matthew, K.; Richard, M.; Amy, W.; Steven, S.H.; Matthew, C.; Shane, S.; Simon, B.; Alex, C.; Sidney, M.; Chris, D. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649.

72. Huang, D.I.; Cronk, Q.C.B. Plann: A Command-Line Application for Annotating Plastome Sequences. *Appl. Plant Sci.* **2015**, *3*, apps.1500026. [CrossRef]

73. Marc, L.; Oliver, D.; Sabine, K.; Ralph, B. OrganellarGenomeDRAW–a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Res.* **2013**, *41*, W575–W581.

74. Thiel, T.; Michalek, W.; Varshney, R.K.; Graner, A. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (Hordeum vulgare L.). *Appl. Genet.* **2003**, *106*, 411–422. [CrossRef]

75. 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]

76. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [CrossRef]

77. 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]

78. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2012**, *28*, 1647. [CrossRef]

79. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]

80. Alexandros, S. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313.

81. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **2007**, *24*, 1586–1591. [CrossRef]

82. Rannala, B.; Yang, Z. Inferring speciation times under an episodic molecular clock. *Syst. Biol.* **2007**, *56*, 453–466. [CrossRef]

83. Huang, C.H.; Sun, R.; Hu, Y.; Zeng, L.; Zhang, N.; Cai, L.; Zhang, Q.; Koch, M.A.; Al-Shehbaz, I.; Edger, P.P. Resolution of Brassicaceae Phylogeny Using Nuclear Genes Uncovers Nested Radiations and Supports Convergent Morphological Evolution. *Mol. Biol. Evol.* **2016**, *33*, 394–412. [CrossRef]

84. Hermant, M.; Hennion, F.; Bartish, I.V.; Yguel, B.; Prinzing, A. Disparate relatives: Life histories vary more in genera occupying intermediate environments. *Perspect. Plant Ecol. Evol. Syst.* **2012**, *14*, 283–301. [CrossRef]

85. Baraloto, C.; Chave, J. Using functional traits and phylogenetic trees to examine the assembly of tropical tree communities. *J. Ecol.* **2012**, *100*, 690–701. [CrossRef]

86. Vanneste, K.; Baele, G.; Maere, S.; Yves, V.D.P. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous-Paleogene boundary. *Genome Res.* **2014**, *24*, 1334–1347. [CrossRef]