diversity

MDPI

Article

A High-Quality Genome Assembly of the Mitochondrial Genome of the Oil-Tea Tree *Camellia gigantocarpa* (Theaceae)

Cui Lu, Li-Zhi Gao and Qun-Jie Zhang *

Institution of Genomics and Bioinformatics, South China Agricultural University, Guangzhou 510642, China
* Correspondence: zhangqunjie@scau.edu.cn

Abstract: *Camellia gigantocarpa* is one of the oil-tea trees whose seeds can be used to extract high-quality vegetable oil. To date, there are no data on the mitochondrial genome of the oil-tea tree, in contrast to the tea-tree *C. sinensis*, which belongs to the same genus. In this paper, we present the first complete mitochondrial genomes of *C. gigantocarpa* obtained using PacBio Hi-Fi (high-fidelity) and Hi-C sequencing technologies to anchor the 970,410 bp genome assembly into a single sequence. A set of 44 protein-coding genes, 22 non-coding genes, 746 simple sequence repeats (SSRs), and more than 201 kb of repetitive sequences were annotated in the genome assembly. The high percentage of repetitive sequences in the mitochondrial genome of *C. gigantocarpa* (20.81%) and *C. sinensis* (22.15%, tea tree) compared to *Arabidopsis thaliana* (4.96%) significantly increased the mitogenome size in the genus *Camellia*. The comparison of the mitochondrial genomes between *C. gigantocarpa* and *C. sinensis* revealed genes exhibit high variance in gene order and low substitution rate within the genus *Camellia*. Information on the mitochondrial genome provides a better understanding of the structure and evolution of the genome in *Camellia* and may contribute to further study of the after-ripening process of oil-tea trees.

Keywords: mitochondrial genome; *Camellia*; *Camellia gigantocarpa*; PacBio Hi-Fi

Citation: Lu, C.; Gao, L.-Z.; Zhang, Q.-J. A High-Quality Genome Assembly of the Mitochondrial Genome of the Oil-Tea Tree *Camellia gigantocarpa* (Theaceae). *Diversity* **2022**, *14*, 850. https://doi.org/10.3390/d14100850

Academic Editors: Hong-Xiang Zhang, Xiao-Long Jiang and Bei Gao

Received: 29 August 2022
Accepted: 29 September 2022
Published: 8 October 2022

1. Introduction

Camellia is the largest genus in Theaceae having more than 200 species, which include many economically important and worldwide cultivated species, such as tea tree, oil-tea tree, and camellia flower. Oil-tea trees are a group of traditional woody edible-oil crop species in China from whose seeds high-quality camellia oil can be extracted. Camellia oil is famous for its nutritional and health benefits because it is rich in unsaturated fatty acids (more than 70% oleic acid and 5–10% linoleic acid) and because of its antioxidant activity [1,2]. The main cultivated oil-tea tree species include *C. oleifera*, *C. gigantocarpa*, *C. chekiangoleosa*, *C. yuhsienensis*, and so on [3–6]. The seeds of *C. gigantocarpa* have an oil content of more than 40% and an oleic acid content of more than 60%, making them an excellent oil-yielding woody plant [4]. As an oil crop, *C. gigantocarpa* has enormous economic and development potential.

Plant mitochondria provide energy and metabolites to the cell. As a source of ATP energy and the intracellular calcium pool, mitochondria carry out a number of cellular functions in plant growth and development [7,8]. Mitochondria also play important functions in seed development [9]. Mutations in some mitochondrial ribosomal proteins have caused gametophytic or seed development defects, such as RPL21M [10] and RPS9M [9]. The seed development of *C. gigantocarpa* affects not only seed fate, but also final seed yield and quality.

Land plant chloroplast genomes have revealed conserved genome structure, gene order, essential gene content, and corresponding gene functions. However, due to extensive rearrangement and repeat sequences, the mitochondrial genome of land plants has very low collinearity, whereas its protein-coding genes are relatively conserved [11]; thus, research

on the evolution and function of the mitochondrial genome can be challenging [12]. The mitochondrial genomes in *Camellia* spp. have a high rearrangement and a large size. To date, the chloroplast genomes of more than 30 species of *Camellia* spp. [13,14] and only one mitochondrial genome of *C. sinensis* have been identified and published [15].

In this work, we sequenced and annotated the complete mitochondrial genome of *C. gigantocarpa*, which is the first mitochondrial genome sequence published for oil-tea tree. The mitogenome characteristics, repetitive sequences, SSR identification, and RNA editing prediction were investigated. Further analyses regarding species synteny and phylogeny were carried out for the determination of phylogenic positions and molecular diversity of the genus *Camellia*. This comparative analysis provided a more comprehensive perspective on the complexity of the mitochondrial genome of the genus *Camellia*.

2. Materials and Methods

2.1. Plant Materials and Genome Sequencing

Fresh and healthy young leaves of *C. gigantocarpa* were collected on 4 April 2022, at the Jinhua International *Camellia* Species Garden of Zhejiang province (geographic coordinates: 29°9′8″ N, 61 1119°35′51.86″ E). The samples were immediately frozen in liquid nitrogen and stored at −80 °C before DNA extraction.

High-molecular-weight genomic DNA was prepared by the CTAB method and followed by purification with QIAGEN® genomic kit (cat#13343, QIAGEN) for regular sequencing, according to the standard operating procedure provided by the manufacturer.

SMRTbell target-size libraries were constructed for Hi-Fi sequencing according to PacBio's standard protocol (Pacific Biosciences, Menlo Park, CA, USA) using 15 kb preparation solutions.

For Hi-C sequencing, the chromosomal structure was fixed by formaldehyde crosslinking, and then the *MboI* enzyme was used to shear DNA. The Hi-C library with insert size 200–600 bp was constructed and then sequenced on the Illumina HiSeq platform. The Hi-C sequence data were qualified with HIC-pro [16].

2.2. Genome Assembly

The mitochondrial genome assembly was started according to the pipeline of Kovar et al. [17]. The Hi-Fi reads were aligned using BLASR ver. 5.3.5 [18] to the mitochondrial genome of 14 plant (*Vaccinium macrocarpon* [19], *Ricinus communis* [20], *Carica papaya* [21], *Citrullus lanatus* [22], *Vitis vinifera* [23], *Glycine max* [24], *Zostera marina* [25], *Sorghum bicolor*, *Zea mays* [26], *Triticum aestivum* [27], *Nicotiana tabacum* [28], *C. sinensis*, *A. thaliana*, and *Salvia miltiorrhiza*) (Table S1). Seqtk (https://github.com/lh3/seqtk (accessed on 5 June 2022)) was used to extract filtered hits into a new fastq file from hits with lengths longer than 500 bp. Canu ver. 2.2 [29] was used to assemble selected Hi-Fi reads to contigs, then the assembled contigs used as the reference genome in the next round (Figure 1). The seventh round assigned the longest contig, N50 with 323,549 bp.

To anchor contigs to mitochondrial genome, Hi-C reads were mapped to the Hi-Fi contigs by Bowtie2 ver. 2.4.2 [30]. Imported reads were sorted and indexed by SAMtools ver. 1.11 [31] and BEDTools ver. 2.30.0 [32]. The mapped reads were analyzed using the Juicer ver. 1.6 [33]. With the Juicer output files, Hi-C scaffolding was performed using 3D-DNA ver. 180,922 (https://github.com/theaidenlab/3d-dna (accessed on5 June 2022)). Inversions and misjoins in the assemblies that occurred during the Hi-C scaffolding process were corrected by using Juicebox ver. 1.11.08 [34] based on the frequency of Hi-C contacts. Finally, the complete mitochondrial genome of *C. gigantocarpa* used 71,338 Hi-Fi reads with approximately 1070X coverage.

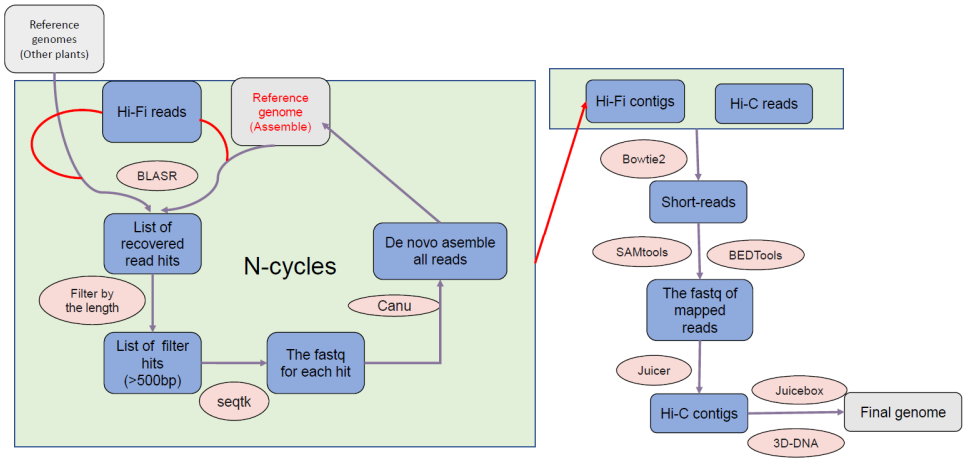

Figure 1. Flowchart of *C. gigantocarpa* mitochondrial genome assembly.

2.3. Genome Annotation and Visualization

MITOFY [22] was used to characterize the complement of protein-coding and rRNA genes in the mitochondrial genome and a tRNA gene search was carried out using the tRNAscan-SE ver. 1.3.1 [35]. The complete mitochondrial genome circular map was created on the web server CGView (http://wishart.biology.ualberta.ca/cgview/ (accessed on 1 July 2022)) [36].

Repetitive elements were identified based on homologous detection and de novo searches. RepeatModeler ver. 2.0.1 [37] was used to identify and model repeat families. Then, RepeatMasker ver. 4.1.0 (http://www.repeatmasker.org/ (accessed on 1 July 2022)) was used to annotate and mask repetitive elements using the library generated by RepeatModeler. Repeat sequences, including forward and palindromic repeats, were also searched by REPuter (https://bibiserv.cebitec.uni-bielefeld.de/reputer (accessed on 1 July 2022)) [38] with the following parameters: minimal length 50 nt and Hamming distance 3 nt. Simple sequence repeats (SSRs) were identified and located using MISA (http://pgrc.ipk-gatersleben.de/misa/ (accessed on 1 July 2022)). All the annotated SSRs were classified by the size and copy number of their tandemly repeated monomer (one nucleotide, $n \geq 8$), dimer (two nucleotides, $n \geq 4$), trimer (three nucleotides, $n \geq 4$), tetramer (four nucleotides, $n \geq 3$), pentamer (five nucleotides, $n \geq 3$), and hexamer (six nucleotides, $n \geq 3$).

2.4. Prediction of RNA-Editing Sites

The predictive RNA Editor for Plants (PREP) (http://prep.unl.edu/ (accessed on 1 July 2022)) was used to predict potential RNA editing sites in protein-coding genes with a cutoff value of 0.2 [39].

2.5. Synteny Analysis

Homologous genes from different plant species were combined using all vs. all BLASTP (BLAST + ver. 2.90) [40], and then synteny blocks were identified and drawn as a graph with MCscanX (Python version) [41]. We adopted the mitochondrial genomes of *C. gigantocarpa*, *C. sinensis*, and *A. thaliana* (Supplementary Material Table S1) for synteny analysis. Synteny analysis of genomes was carried out at the nucleic acid level using Mauve ver. 2.4.0 [42].

2.6. Phylogenetic Analysis

A total of 12 conserved mitochondrial protein-coding genes [15] among *C. gigantocarpa* and 15 other plant species (*V. macrocarpon* [19], *R. communis* [20], *C. papaya* [21], *C. lanatus* [22], *V. vinifera* [23], *G. max* [24], *Z. marina* [25], *S. bicolor*, *Z. mays* [26], *T. aestivum* [27], *N. tabacum* [28], *C. sinensis*, *A. thaliana*, *S. miltiorrhiza*, and *Ginkgo biloba*) (Table S1) were indi-

vidually aligned with MAFFT ver. 7.475 (L-INS-I algorithm) [43], and then concatenated to construct a contiguous sequence in the order of *cob*, *cox1*, *cox2*, *cox3*, *nad2*, *nad3 nad5*, *nad6*, *nad7*, *nad9*, *atp1*, and *atp9*. The HIVw + I + G + F model of amino acid substitution was found to be the best fit by Protest ver. 3.4.2 (coverage threshold=0.5), [44]. A maximum likelihood (ML) phylogenetic tree was produced using RAxML ver. 8.2.12 [45] with *G. biloba* as the outgroup.

3. Results

3.1. Genome Assembly and Genome Annotation

The circular mitochondrial genome of *C. gigantocarpa* was 970,410 bp in length (GenBank: OP270590) and a GC content of 45%, and it contained 44 protein-coding genes and 22 non-coding genes (Figure 2). The protein-coding genes of mitochondrial genome of *C. gigantocarpa* included 15 NADH dehydrogenase genes (*nad1–nad7*, *nad4L*, *nad9*; there are three copies of *nad1*, *nad2*, and *nad5*), two succinate dehydrogenase genes (*sdh3* and *sdh4*), one cytochrome c reductase gene (*cob*), three cytochrome c oxidase genes (*cox1–cox3*), five ATP synthase synthesis genes (*atp1*, *atp4*, *atp6*, *atp8*, and *atp9*), four cytochrome c biogenesis genes (*ccmB*, *ccmC*, *ccmFC*, and *ccmFn*), one maturase gene (*matR*), one transporter gene (*mttB*), and 12 ribosomal protein genes (*rpl10*, *rpl16*, *rpl2*, *rpl5*, *rps1*, *rps3*,*rps4*, *rps7*, *rps12–rps14*, and *rps19*). The non-coding genes of mitochondrial genome of *C. gigantocarpa* include three rRNA genes (*rrn5*, *rrn18*, and *rrn26*) and 19 tRNA genes that transferred 16 amino acids. In comparison to the mitochondrial genome of *C. sinensis* (GenBank: OM809792.1), the mitochondrial genome of *C. sinensis* contained 47 protein-coding genes and 33 non-coding genes, which was more than *C. gigantocarpa*, but two protein-coding genes (*rpl2*, and *rps3*) were exclusively found in the mitochondrial genome of *C. gigantocarpa* (Table S2).

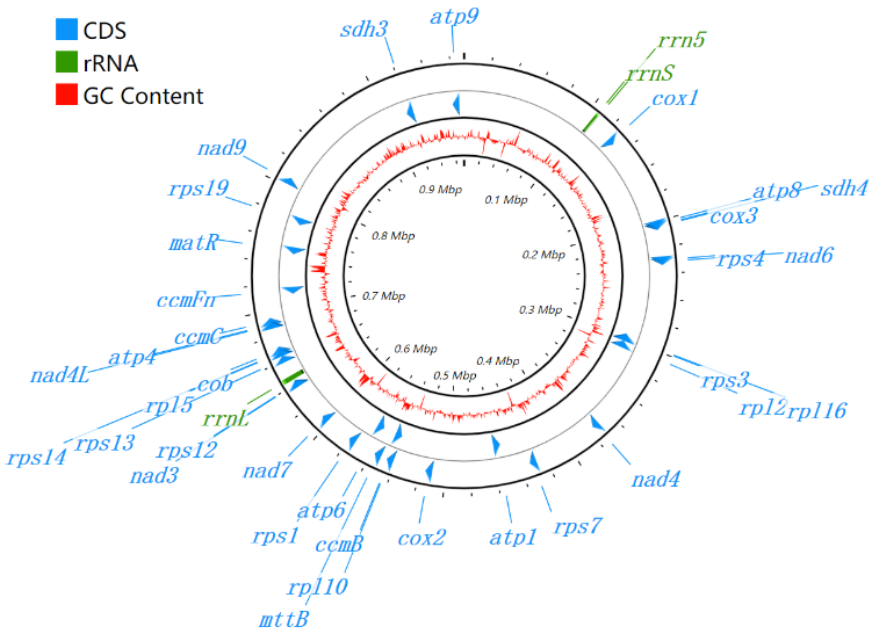

Figure 2. Mitochondrial gene map of *C. gigantocarpa*.

3.2. Identified Repetitive Sequences

The total length of *C. gigantocarpa* mitochondrial genomes was 201 kb (20.81%), of which long terminal repeats retrotransposons (LTR-RTs) accounted for 20.39% (197 kb) (Table 1). While for *C. sinensis*, the repetitive sequence was 22.15% (202 kb), of which LTR retrotransposons accounted for 21.68% (198 kb) in its mitochondrial genome. Compared to *A. thaliana* 3.51% (12 kb), the mitochondrial genome in *Camellia* spp. showed a large expansion which may have been caused by LTR retrotransposons insertion.

Table 1. Comparisons of repetitive sequence categories and contents of *C. gigantocarpa*, *C. sinensis*, and *A. thaliana* mitochondrial genome.

Type	*C. gigantocarpa*		*C. sinensis*		*A. thaliana*	
	Length (bp)	Percentage (%)	Length (bp)	Percentage (%)	Length (bp)	Percentage (%)
RNA/non-LTR-RTs	432	0.04	992	0.11	1567	0.43
RNA/LTR-RTs	197,860	20.39	198,371	21.68	12,887	3.51
DNA transposons	0	0	0	0	104	0.03
Other repeats	3622	0.38	3315	0.36	3670	0.99
Total	201,914	20.81	202,678	22.15	18,228	4.96

Long repeat sequences (repeat unit > 50 bp) of forward and palindromic repeats were further annotated: 50 paired repeats were distributed throughout the genome, including 27 paired forward repeats and 23 paired palindromic repeats (Table S2). These repeats ranged from 434 to 2631 bp in length.

Short repeats are abundant in plant mitochondrial genomes, particularly in higher plants [46]. In the mitochondrial genome of *C. gigantocarpa*, 746 SSRs were found, with 32.44% being monomers, 44.5% dimers, 4.96% trimers, 14.21% tetramers, 3.08% pentamers, and 0.8% hexamers (Table 2). In addition, the two most abundant SSR motif was A/T (28.28%) and AG/CT (31.64%) (Table S4).

Table 2. Statistics of SSR motifs in the *C. gigantocarpa* mitochondrial genomes.

SSR Motif	SSR Number	SSR (%)
Monomer	242	32.44
Dimer	332	44.5
Trimer	37	4.96
Tetramer	106	14.21
Pentamer	23	3.08
Hexamer	6	0.8
Total	746	100

3.3. The Prediction of RNA Editing

In mitochondria, RNA editing is common. A single base can modify a codon, which, in turn, alters an amino acid, and changes the content, structure, or function of the protein. RNA editing frequently results in the unintentional addition of a stop codon, which prevents the protein from being fully translated, making the protein non-functional [47]. Our results show that all 44 protein-coding genes had RNA edits, all of which were C-U transitions. All 483 C-U RNA editing sites were unevenly distributed among different genes, ranging from 1 (*nad5*) to 37 (*ccmFn*) (Figure 3). There were four cases of RNA editing, of *atp6*, *atp9*, *cox2*, and *rpl16*, in which the results were stop codons. The amino acid characteristics were modified by 55% RNA editing, such as switching from hydrophilic to hydrophobic amino acids (Table S5).

3.4. Comparison of the Genome Structure

The sizes of the mitochondrial genomes of *Camellia* spp. (*C. gigantocarpa* and *C. sinensis*) were significantly larger than those of *A. thaliana* (367,808 bp), but the differences in the types and numbers of protein-coding genes were not significant (Table S2).

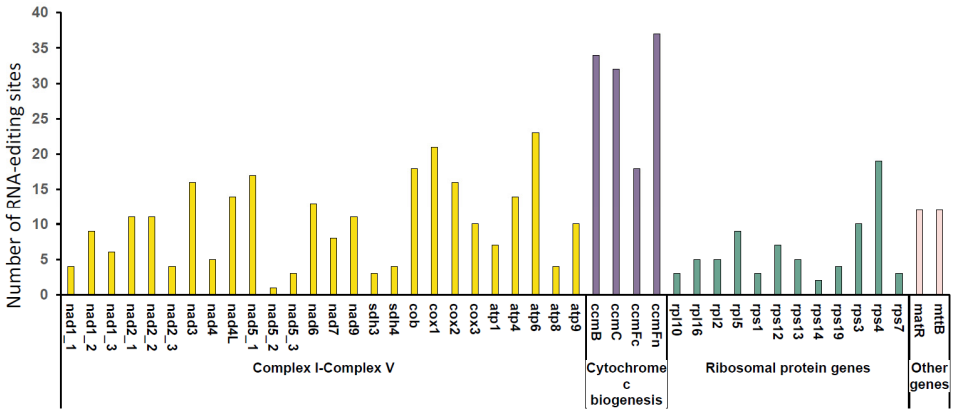

Figure 3. The number of RNA-editing sites; *nad1*, *nad2*, and *nad5* have three copies.

The collinearity analysis revealed that the protein-coding genes of the mitochondrial genomes of *C. gigantocarpa*, *C. sinensis*, and *A. thaliana* were highly conserved, but with high variance in the order of mitochondrial genes among these three species. The number of co-linear gene pairs between *C. gigantocarpa* and *C. sinensis* was 36, and the similarity of their corresponding coding regions was 99.8%. The number of co-linear gene pairs between *C. gigantocarpa* and *A. thaliana* was 34, and the similarity of their coding regions was 95.2%. In the mitochondrial genome of *C. gigantocarpa*, 34% protein-coding genes had the same gene order (three, and more genes arranged in that order) as *C. sinensis* (*atp8-cox3-sdh4-rps4-nad6*, *rpl10-ccmB-mttB-atp6*, *rpl5-rps14-cob*, and *nad4L-atp4-ccmC*), however, no protein-coding genes shared the same gene order as *A. thaliana* (Figure 4). A MAUVE graphic of the structural alignments of complete mitochondrial genomes of three species also revealed divergences. We found complex genome rearrangements in two *Camellia* spp., despite high sequence similarity (Figure 5a). In contrast, in the mitochondrial genome of *A. thaliana* and *C. gigantocarpa*, only protein-coding gene regions could be aligned (Figure 5b). In contrast to the dramatic expansion of intergenic regions and rapid evolution of gene order in the *Camellia* spp., most functional genes were highly conserved in plant mitochondrial genomes.

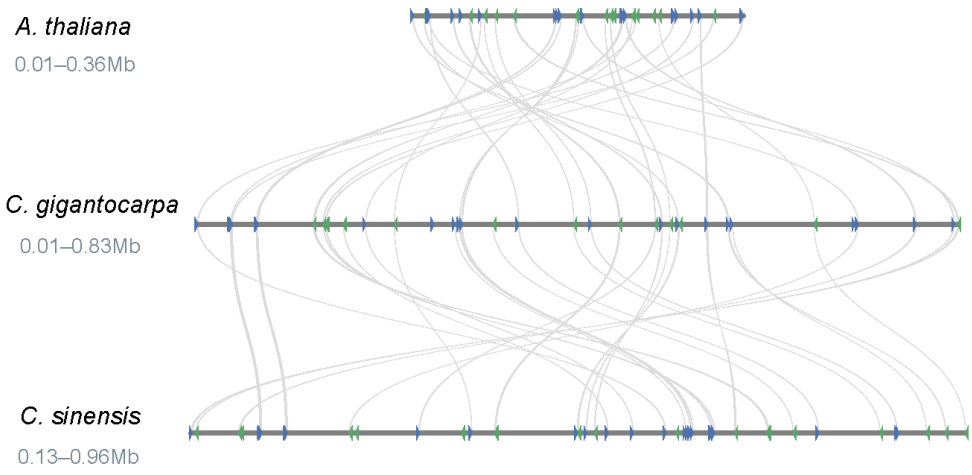

Figure 4. Synteny analysis of *C. gigantocarpa*, *C. sinensis*, and *A. thaliana* mitochondrial genomes.

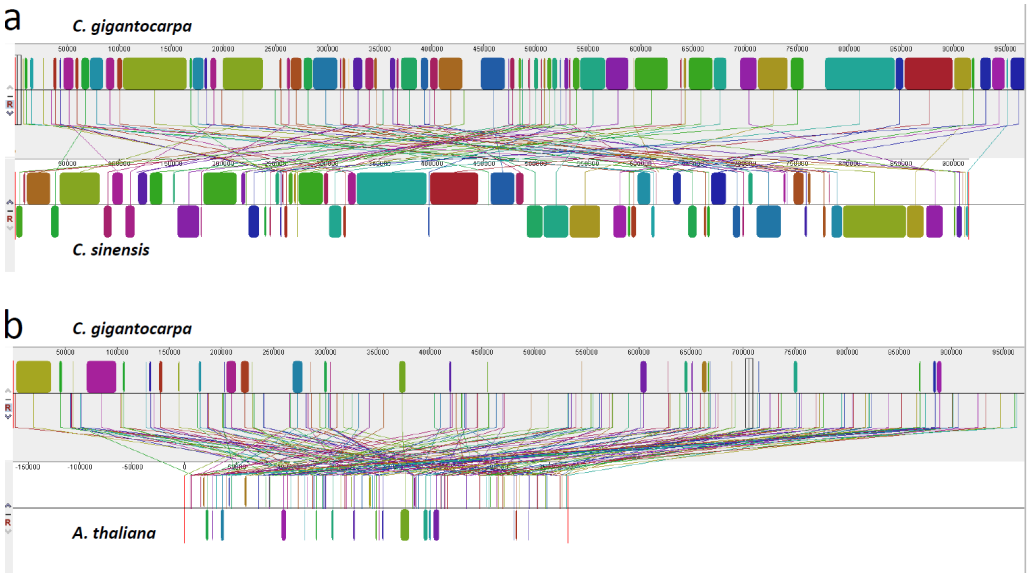

Figure 5. Whole mitochondrial alignments of three species. (**a**) Whole mitochondrial alignments of *C. gigantocarpa* and *C. sinensis*; and (**b**) whole mitochondrial alignments of *C. gigantocarpa* and *A. thaliana*.

3.5. Phylogenetic Analysis

ML trees were built for 12 protein sequences shared by 16 plant mitochondrial genomes (Table S1) [15]. The ML phylogeny tree with *G. biloba* as the outgroup formed two clades: monocotyledons and dicotyledons. We discovered that *C. gigantocarpa* and *C. sinensis* were clustered together with *V. macrocarpon* [20], and that these three species were members of the order Ericales (Figure 6). The pairwise distance (Poisson model) of these 12 protein sequences of *C. gigantocarpa* and *C. sinensis* is 0.00611, showing that, despite the mitochondrial genome structures being relatively different, mitochondrial protein-coding genes among *C. gigantocarpa* and *C. sinensis* are conserved.

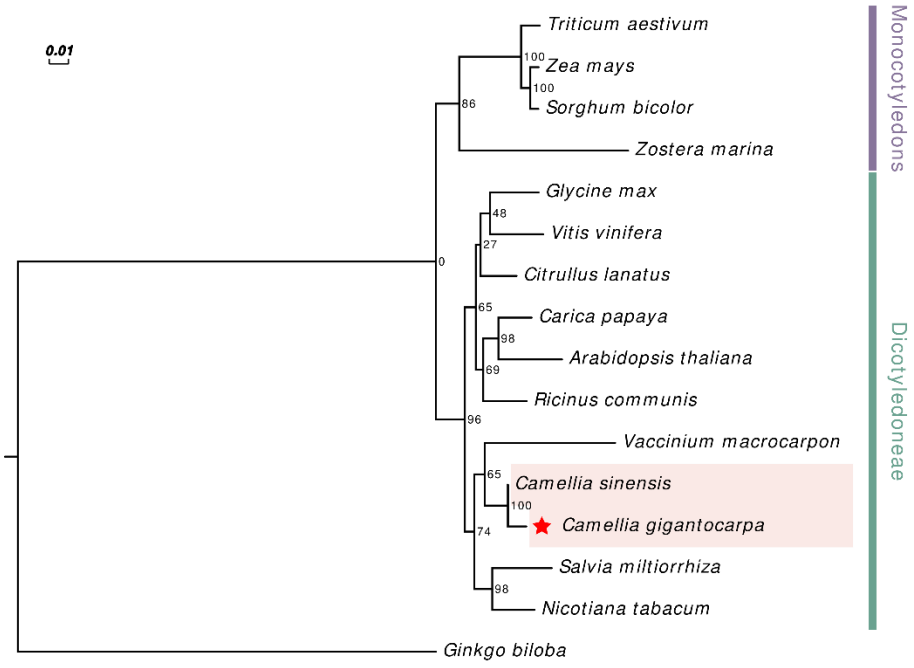

Figure 6. Maximum likelihood tree based on 12 genes common in the 16 plant mitochondrial genomes.

4. Discussion

Camellia spp. mitochondrial genomes are harder to assemble than chloroplast genomes due to their large size and high repetitive rate [48]. Compared to dozens of published chloroplast genomes [49], only the mitochondrial genome of *C. sinensis* was assembled [15]. Utilizing long-read genome sequencing allows mitochondrial genome assembly to achieve high-sequence contiguity as well as high-scaffold contiguity [50]. Here, we used a combination of sequencing technologies, including PacBio Hi-Fi and Hi-C, to assemble the mitochondrial genome for *C. gigantocarpa*, and present workflows for the accurate and complete assembly of the large and complex plant mitochondrial genome with highly repetitive sequences (Figure 1).

Mitochondria play key roles in energy supply during seed development for encoded components of the TCA cycle and ETC complexes [51,52]. Especially for *Camellia* spp., the physiological maturity of *C. gigantocarpa* seeds is impactful to germination efficiency. Our work assembled the mitochondrial genome of *C. gigantocarpa* into a complete mitochondrial genome of 970,410 bp and annotated a total of 44 protein-coding genes, 22 non-coding genes, compared with *C. sinensis*, whose two protein-coding genes (*rpl2*, and *rps3*) are exclusively found in the mitochondrial genome of *C. gigantocarpa* (Figure 2). Mitochondrial genome organization, core protein-coding genes, and RNA editing provided rich genetic information for understanding the genetics and evolution of *C. gigantocarpa*. In all, 483 RNA-editing sites were identified in this mitochondrial genome; the editing sites were distributed among all 44 protein-coding genes (Figure 3).

In plant mitochondria, repetitive sequences are typical and frequently quite long [53]. The occurrence of such repeats can partially account for variations in the size of the mitochondrial genome [25]. We annotated high-confidence prediction of 746 simple sequence repeats (SSRs) and more than 201 kb (31.16%) of repetitive sequences in genome assembly. Compared with *A. thalian* (367 kb), *C. gigantocarpa* and *C. sinensis* (>900 kb) have bigger mitochondrial genomes and contain more repetitive sequences. The mitochondrial genome of *Camellia* spp. is abundant with repetitive elements, accounting for more than 20% of the genome, whereas *A. thalian* has only 4.96% (Table 1). The variation in mitochondrial genome size of this tree species can be partially explained by its repetitive content. Plant mitochondria are made up of a heterogeneous population of highly branching, circularly permuted linear molecules [54]. Large-size repeats (>1 kb) conduct a multipartite structure of plant mitochondrial genomes [55]. There are six and eight paired large-size repeats in the mitochondrial genomes of *C. gigantocarpa* and *C. sinensis* [15], which may result in the diversification of mitochondrial structure in *Camellia* spp. and make assembling the mitochondrial genome difficult.

In higher plants, the gene order of chloroplast genomes is often highly conserved, especially among closely related species. Although the genes contained in the mitochondrial genomes of higher plants are largely conserved, the size, structure, and gene order of mitochondrial genomes are highly variable [56]. Gene-order comparisons frequently reflect the high rate of mitochondrial genome rearrangement between plant species. According to this study, *C. gigantocarpa* has fewer protein-coding genes with the same gene order as *A. thalian* than *C. sinensis* (Figure 4) indicating that the discovered gene order is less conserved when compared to more distantly related species. Despite usually slow rates of sequence evolution, plant mitochondrial genomes develop rapidly in terms of genome rearrangement [57]. Our work showed high-sequence homology and abundant genomic rearrangement between the *C. gigantocarpa* and *C. sinensis* mitochondrial genomes (Figure 5).

Mitochondrial genomes contain valuable information that can be used for understanding the evolution of these mitochondria. We built ML trees for 12 protein sequences shared by 16 plant mitochondrial genomes (Table S1) and found that *C. gigantocarpa* was grouped with *C. sinensis* with 100% bootstrap support, which means the protein-coding genes are conserved despite the high rate of recombination in the mitochondrial gene order in the genus *Camellia* (Figure 6).

Camellia spp. are highly self-incompatible plants, and many species are polyploid. The identification of species classification and evolutionary relationships of the genus *Camellia* is still challenging because of the widespread hybridization. Numerous researches have recently used whole-genome resequencing [58] and RNA-seq [59] to investigate the evolutionary relationships of the *Camellia* genus. Thus far, the high cost of sequencing has limited genome research of the genus *Camellia*, and few *Camellia* spp. have had their genomes sequenced, with the majority of the work focused on *C. sinensis* [58,60–63] and *C. oleifera* [64]. Comparing the organelle genomes (chloroplast DNA and mitochondrial DNA) to the nuclear genome, the organelle genomes have several advantages over the nuclear genome, including a smaller size, a lower sequencing cost, a simpler assembly method, and a matrilineal inheritance [65]. Our complete mitochondrial genome of *C. gigantocarpa* in this study is the second mitochondrial genome sequence to be published in the genus *Camellia*, and it offers new insight into the evolution of the mitochondrial genome of the genus *Camellia*.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/d14100850/s1, Table S1: The species used for synteny analysis and Phylogenetic analysis; Table S2: Gene content of the *C. gigantocarpa*, *C. sinensis* and *A. thaliana* mitochondrial genome; Table S3: Long repeats (repeat unit > 50 bp) in the *C. gigantocarpa* mitochondrial genome; Table S4: Frequency of classified repeat types; Table S5: Prediction of RNA editing sites.

Author Contributions: Q.-J.Z. designed and managed the project; C.L. performed the genome assembly, genome annotation and subsequent data analyses; C.L. and Q.-J.Z. wrote the manuscript; L.-Z.G. and Q.-J.Z. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Natural Science Foundation of China (32170625) (to Q.-J.Z.) and the Guangdong Special Support 318 Program (to Q.-J.Z.).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The genome assembly has been deposited in the National Center for Biotechnology Information (NCBI) GenBank database (https://www.ncbi.nlm.nih.gov/genbank (accessed on 10 August 2022); accession number: OP270590).

Acknowledgments: We appreciate the anonymous reviewers for their comments on this manuscript and all editors for assistance in editing the manuscript. The authors thank Hai-Hua He for providing us with the Plant Materials used in this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ye, Z.; Wu, Y.; Ul, H.M.Z.; Yan, W.; Yu, J.; Zhang, J.; Yao, G.; Hu, X. Complementary transcriptome and proteome profiling in the mature seeds of *Camellia oleifera* from Hainan island. *PLoS ONE* **2020**, *15*, e226888. [CrossRef] [PubMed]
2. Sieber, J.; Lindenmeyer, M.T.; Kampe, K.; Campbell, K.N.; Cohen, C.D.; Hopfer, H.; Mundel, P.; Jehle, A.W. Regulation of podocyte survival and endoplasmic reticulum stress by fatty acids. *Am. J. Physiol. Ren. Physiol.* **2010**, *299*, F821–F829. [CrossRef] [PubMed]
3. Lin, P.; Wang, K.; Zhou, C.; Xie, Y.; Yao, X.; Yin, H. Seed transcriptomics analysis in *Camellia oleifera* uncovers genes associated with oil content and fatty acid composition. *Int. J. Mol. Sci.* **2018**, *19*, 118. [CrossRef] [PubMed]
4. Shen, T.F.; Huang, B.; Xu, M.; Zhou, P.Y.; Ni, Z.X.; Gong, C.; Wen, Q.; Cao, F.L.; Xu, L.A. The reference genome of *Camellia chekiangoleosa* provides insights into Camellia evolution and tea oil biosynthesis. *Hortic. Res.* **2022**, *9*, uhab083. [CrossRef]
5. Li, J.; Luo, Z.; Zhang, C.; Qu, X.; Chen, M.; Song, T.; Yuan, J. Seasonal variation in the rhizosphere and non-rhizosphere microbial community structures and functions of *Camellia yuhsienensis* hu. *Microorganisms* **2020**, *8*, 1385. [CrossRef]
6. Xie, Y. Fruit economic characters and seed oil components of seven kinds of oil-used *Camellia*. *Chin. J. Trop. Crops* **2016**, *2*, 427–431.
7. Epstein, C.B.; Waddle, J.A.; Hale, W.; Davé, V.; Thornton, J.; Macatee, T.L.; Garner, H.R.; Butow, R.A. Genome-wide responses to mitochondrial dysfunction. *Mol. Biol. Cell* **2001**, *12*, 297–308. [CrossRef]
8. Carafoli, E. The fateful encounter of mitochondria with calcium: How did it happen? *Biochim. Biophys. Acta* **2010**, *1797*, 595–606. [CrossRef]
9. Lu, C.; Yu, F.; Tian, L.; Huang, X.; Tan, H.; Xie, Z.; Hao, X.; Li, D.; Luan, S.; Chen, L. RPS9M, a mitochondrial ribosomal protein, is essential for central cell maturation and endosperm development in Arabidopsis. *Front. Plant Sci.* **2017**, *8*, 2171. [CrossRef]

10. Portereiko, M.F.; Sandaklie-Nikolova, L.; Lloyd, A.; Dever, C.A.; Otsuga, D.; Drews, G.N. NUCLEAR FUSION DEFECTIVE1 encodes the Arabidopsis RPL21M protein and is required for karyogamy during female gametophyte development and fertilization. *Plant Physiol.* **2006**, *141*, 957–965. [CrossRef]
11. Knoop, V. The mitochondrial DNA of land plants: Peculiarities in phylogenetic perspective. *Curr. Genet.* **2004**, *46*, 123–139. [CrossRef]
12. Marechal, A.; Brisson, N. Recombination and the maintenance of plant organelle genome stability. *New Phytol.* **2010**, *186*, 299–317. [CrossRef]
13. Yang, J.B.; Yang, S.X.; Li, H.T.; Yang, J.; Li, D.Z. Comparative chloroplast genomes of *Camellia* species. *PLoS ONE* **2013**, *8*, e73053. [CrossRef]
14. Li, W.; Zhang, C.; Guo, X.; Liu, Q.; Wang, K. Complete chloroplast genome of *Camellia japonica* genome structures, comparative and phylogenetic analysis. *PLoS ONE* **2019**, *14*, e216645. [CrossRef]
15. Zhang, F.; Li, W.; Gao, C.; Zhang, D.; Gao, L. Deciphering tea tree chloroplast and mitochondrial genomes of *Camellia sinensis* var. *assamica*. *Sci. Data* **2019**, *6*, 209. [CrossRef]
16. Servant, N.; Varoquaux, N.; Lajoie, B.R.; Viara, E.; Chen, C.; Vert, J.; Heard, E.; Dekker, J.; Barillot, E. Hic-pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **2015**, *16*, 259. [CrossRef]
17. Kovar, L.; Nageswara-Rao, M.; Ortega-Rodriguez, S.; Dugas, D.V.; Straub, S.; Cronn, R.; Strickler, S.R.; Hughes, C.E.; Hanley, K.A.; Rodriguez, D.N.; et al. PacBio-based mitochondrial genome assembly of *Leucaena trichandra* (leguminosae) and an intrageneric assessment of mitochondrial RNA editing. *Genome Biol. Evol.* **2018**, *10*, 2501–2517. [CrossRef]
18. Chaisson, M.J.; Tesler, G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): Application and theory. *BMC Bioinform.* **2012**, *13*, 238. [CrossRef]
19. Fajardo, D.; Schlautman, B.; Steffan, S.; Polashock, J.; Vorsa, N.; Zalapa, J. The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. *Gene* **2014**, *536*, 336–343. [CrossRef]
20. Rivarola, M.; Foster, J.T.; Chan, A.P.; Williams, A.L.; Rice, D.W.; Liu, X.; Melake-Berhan, A.; Huot Creasy, H.; Puiu, D.; Rosovitz, M.J.; et al. Castor bean organelle genome sequencing and worldwide genetic diversity analysis. *PLoS ONE* **2011**, *6*, e21743. [CrossRef]
21. Magee, A.M.; Aspinall, S.; Rice, D.W.; Cusack, B.P.; Sémon, M.; Perry, A.S.; Stefanović, S.; Milbourne, D.; Barth, S.; Palmer, J.D.; et al. Localized hypermutation and associated gene losses in legume chloroplast genomes. *Genome Res.* **2010**, *20*, 1700–1710. [CrossRef] [PubMed]
22. Alverson, A.J.; Wei, X.; Rice, D.W.; Stern, D.B.; Barry, K.; Palmer, J.D. Insights into the evolution of mitochondrial genome size from complete sequences of *Citrullus lanatus* and *Cucurbita pepo* (cucurbitaceae). *Mol. Biol. Evol.* **2010**, *27*, 1436–1448. [CrossRef] [PubMed]
23. Goremykin, V.V.; Salamini, F.; Velasco, R.; Viola, R. Mitochondrial DNA of *Vitis vinifera* and the Issue of Rampant Horizontal Gene Transfer. *Mol. Biol. Evol.* **2008**, *26*, 99–110. [CrossRef] [PubMed]
24. Chang, S.; Wang, Y.; Lu, J.; Gai, J.; Li, J.; Chu, P.; Guan, R.; Zhao, T. The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels. *PLoS ONE* **2013**, *8*, e56502.
25. Petersen, G.; Cuenca, A.; Zervas, A.; Ross, G.T.; Graham, S.W.; Barrett, C.F.; Davis, J.I.; Seberg, O. Mitochondrial genome evolution in Alismatales: Size reduction and extensive loss of ribosomal protein genes. *PLoS ONE* **2017**, *12*, e177606. [CrossRef]
26. Clifton, S.W.; Minx, P.; Fauron, C.M.R.; Gibson, M.; Allen, J.O.; Sun, H.; Thompson, M.; Barbazuk, W.B.; Kanuganti, S.; Tayloe, C.; et al. Sequence and Comparative Analysis of the Maize NB Mitochondrial Genome. *Plant Physiol.* **2004**, *136*, 3486–3503. [CrossRef]
27. Cui, P.; Liu, H.; Lin, Q.; Ding, F.; Zhuo, G.; Hu, S.; Liu, D.; Yang, W.; Zhan, K.; Zhang, A.; et al. A complete mitochondrial genome of wheat (*Triticum aestivum* cv. Chinese Yumai), and fast evolving mitochondrial genes in higher plants. *J. Genet.* **2009**, *88*, 299–307. [CrossRef]
28. Sugiyama, Y.; Watase, Y.; Nagase, M.; Makita, N.; Yagura, S.; Hirai, A.; Sugiura, M. The complete nucleotide sequence and multipartite organization of the tobacco mitochondrial genome: Comparative analysis of mitochondrial genomes in higher plants. *Mol. Genet. Genomics* **2005**, *272*, 603–615. [CrossRef]
29. 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]
30. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [CrossRef]
31. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef]
32. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [CrossRef]
33. Durand, N.C.; Shamim, M.S.; Machol, I.; Rao, S.S.P.; Huntley, M.H.; Lander, E.S.; Aiden, E.L. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **2016**, *3*, 95–98. [CrossRef]
34. Durand, N.C.; Robinson, J.T.; Shamim, M.S.; Machol, I.; Mesirov, J.P.; Lander, E.S.; Aiden, E.L. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **2016**, *3*, 99–101. [CrossRef]
35. Lowe, T.M.; Eddy, S.R. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **1997**, *25*, 955–964. [CrossRef]

36. Stothard, P.; Wishart, D.S. Circular genome visualization and exploration using CGView. *Bioinformatics* **2005**, *21*, 537–539. [CrossRef]
37. Flynn, J.M.; Hubley, R.; Goubert, C.; Rosen, J.; Clark, A.G.; Feschotte, C.; Smit, A.F. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 9451–9457. [CrossRef]
38. 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]
39. Mower, J.P. The PREP suite: Predictive RNA editors for plant mitochondrial genes, chloroplast genes and user-defined alignments. *Nucleic Acids Res.* **2009**, *37*, W253–W259. [CrossRef]
40. Johnson, M.; Zaretskaya, I.; Raytselis, Y.; Merezhuk, Y.; Mcginnis, S.; Madden, T.L. NCBI BLAST: A better web interface. *Nucleic Acids Res.* **2008**, *36*, W5–W9. [CrossRef]
41. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H.; et al. *MCScanX*: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef]
42. Darling, A.C.; Mau, B.; Blattner, F.R.; Perna, N.T. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* **2004**, *14*, 1394–1403. [CrossRef]
43. 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]
44. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. Prottest 3: Fast selection of best-fit models of protein evolution. *Bioinformatics* **2011**, *27*, 1164–1165. [CrossRef]
45. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef]
46. Odahara, M.; Kuroiwa, H.; Kuroiwa, T.; Sekine, Y. Suppression of repeat-mediated gross mitochondrial genome rearrangements by RecA in the moss physcomitrella patens. *Plant Cell.* **2009**, *21*, 1182–1194. [CrossRef]
47. Wu, Z.; Stone, J.D.; Atorchová, H.; Sloan, D.B. High transcript abundance, RNA editing, and small RNAs in intergenic regions within the massive mitochondrial genome of the angiosperm *Silene noctiflora*. *BMC Genomics.* **2015**, *16*, 938. [CrossRef]
48. Zhao, N.; Wang, Y.; Hua, J. The roles of mitochondrion in intergenomic gene transfer in plants: A source and a pool. *Int. J. Mol. Sci.* **2018**, *19*, 547. [CrossRef]
49. Li, L.; Hu, Y.; He, M.; Zhang, B.; Wu, W.; Cai, P.; Huo, D.; Hong, Y. Comparative chloroplast genomes: Insights into the evolution of the chloroplast genome of *Camellia sinensis* and the phylogeny of *Camellia*. *BMC Genom.* **2021**, *22*, 138. [CrossRef]
50. Weisenfeld, N.I.; Kumar, V.; Shah, P.; Church, D.M.; Jaffe, D.B. Direct determination of diploid genome sequences. *Genome Res.* **2017**, *27*, 757–767. [CrossRef]
51. Law, S.R.; Narsai, R.; Taylor, N.L.; Delannoy, E.; Carrie, C.; Giraud, E.; Millar, A.H.; Small, I.; Whelan, J. Nucleotide and RNA metabolism prime translational initiation in the earliest events of mitochondrial biogenesis during Arabidopsis germination. *Plant Physiol.* **2012**, *158*, 1610–1627. [CrossRef] [PubMed]
52. Logan, D.C.; Millar, A.H.; Sweetlove, L.J.; Hill, S.A.; Leaver, C.J. Mitochondrial biogenesis during germination in maize embryos. *Plant Physiol.* **2001**, *125*, 662–672. [CrossRef] [PubMed]
53. Satoh, M.; Kubo, T.; Nishizawa, S.; Estiati, A.; Itchoda, N.; Mikami, T. The cytoplasmic male-sterile type and normal type mitochondrial genomes of sugar beet share the same complement of genes of known function but differ in the content of expressed ORFs. *Mol. Genet. Genomics.* **2004**, *272*, 247–256. [CrossRef] [PubMed]
54. Oldenburg, D.J.; Bendich, A.J. Mitochondrial DNA from the liverwort marchantia polymorpha: Circularly permuted linear molecules, head-to-tail concatemers, and a 5′ protein 1 1edited by n.-M. Chua. *J. Mol. Biol.* **2001**, *310*, 549–562. [CrossRef] [PubMed]
55. Gualberto, J.M.; Mileshina, D.; Wallet, C.; Niazi, A.K.; Weber-Lotfi, F.; Dietrich, A. The plant mitochondrial genome: Dynamics and maintenance. *Biochimie* **2014**, *100*, 107–120. [CrossRef] [PubMed]
56. Gualberto, J.M.; Newton, K.J. Plant mitochondrial genomes: Dynamics and mechanisms of mutation. *Annu. Rev. Plant. Biol.* **2017**, *68*, 225–252. [CrossRef] [PubMed]
57. Cole, L.W.; Guo, W.; Mower, J.P.; Palmer, J.D.; Purugganan, M. High and variable rates of repeat-mediated mitochondrial genome rearrangement in a genus of plants. *Mol. Biol. Evol.* **2018**, *35*, 2773–2785. [CrossRef]
58. Zhang, X.; Chen, S.; Shi, L.; Gong, D.; Zhang, S.; Zhao, Q.; Zhan, D.; Vasseur, L.; Wang, Y.; Yu, J.; et al. Haplotype-resolved genome assembly provides insights into evolutionary history of the tea plant *Camellia sinensis*. *Nat. Genet.* **2021**, *53*, 1250–1259. [CrossRef]
59. Wu, Q.; Tong, W.; Zhao, H.; Ge, R.; Li, R.; Huang, J.; Li, F.; Wang, Y.; Mallano, A.I.; Deng, W.; et al. Comparative transcriptomic analysis unveils the deep phylogeny and secondary metabolite evolution of 116 *Camellia* plants. *Plant J.* **2022**, *111*, 406–421. [CrossRef]
60. Wang, X.; Feng, H.; Chang, Y.; Ma, C.; Wang, L.; Hao, X.; Li, A.L.; Cheng, H.; Wang, L.; Cui, P.; et al. Population sequencing enhances understanding of tea plant evolution. *Nat. Commun.* **2020**, *11*, 4447. [CrossRef]
61. Xia, E.; Tong, W.; Hou, Y.; An, Y.; Chen, L.; Wu, Q.; Liu, Y.; Yu, J.; Li, F.; Li, R.; et al. The reference genome of tea plant and resequencing of 81 diverse accessions provide insights into its genome evolution and adaptation. *Mol. Plant.* **2020**, *13*, 1013–1026. [CrossRef]

62. Zhang, Q.J.; Li, W.; Li, K.; Nan, H.; Shi, C.; Zhang, Y.; Dai, Z.Y.; Lin, Y.L.; Yang, X.L.; Tong, Y.; et al. The chromosome-level reference genome of tea tree unveils recent bursts of non-autonomous LTR retrotransposons in driving genome size evolution. *Mol. Plant* **2020**, *13*, 935–938. [CrossRef]
63. Zhang, W.; Zhang, Y.; Qiu, H.; Guo, Y.; Wan, H.; Zhang, X.; Scossa, F.; Alseekh, S.; Zhang, Q.; Wang, P.; et al. Genome assembly of wild tea tree DASZ reveals pedigree and selection history of tea varieties. *Nat. Commun.* **2020**, *11*, 3719. [CrossRef]
64. Lin, P.; Wang, K.; Wang, Y.; Hu, Z.; Yan, C.; Huang, H.; Ma, X.; Cao, Y.; Long, W.; Liu, W.; et al. The genome of oil-camellia and population genomics analysis provide insights into seed oil domestication. *Genome Biol.* **2022**, *23*, 14. [CrossRef]
65. Sandhya, S.; Srivastava, H.; Kaila, T.; Tyagi, A.; Gaikwad, K. Methods and tools for plant organelle genome sequencing, assembly, and downstream analysis. In *Legume Genomics*; Humana: New York, NY, USA, 2020; Volume 2107, pp. 49–98.